- office hours
- HW due Thursday

Relativity Lecture 3

At the end of last lecture, we introduced the idea of how conservation laws led to constraints on the description of a physical system via Noether's theorem. For example, we discussed how conservation of angular momentum required that the physical description of a system is in ~~terms~~ terms of vector dot products. Under a rotation implemented by a matrix $\mathbb{M}$, a vector transforms as:

$$\vec{a} \to \mathbb{M}\vec{a} \quad \text{and the dot product is}$$

$$\vec{b}\cdot\vec{a} = \vec{b}^T\vec{a} \to (\mathbb{M}\vec{b})^T \mathbb{M}\vec{a} = \vec{b}^T \mathbb{M}^T \mathbb{M}\vec{a} = \vec{b}\cdot\vec{a}$$

because $\mathbb{M}^T\mathbb{M} = \mathbb{1}$ for rotations. Any rotation matrix satisfies this requirement. Another way to say it is that rotation matrices leave the identity matrix invariant:

$$\mathbb{M}^T \mathbb{1} \mathbb{M} = \mathbb{1}.$$

Now, on to relativity. We want to describe our particle physics system so that relativistic energy, momentum, and angular momentum are conserved. Naturally, we should expect then to be able to construct vectors and the notion of a dot product which will naturally express these conservation laws; that is, the symmetries associated with the conservation laws. We have the relativistic energy-momentum relationship

$$E^2 = m^2 + \vec{p}^2 \quad \text{or} \quad m^2 = E^2 - \vec{p}^2,$$

expressed in the Lorentz-invariant way. $E^2 - \vec{p}^2$ looks like a kind of dot product with a kind of vector.

If we define a four-vector

$$p = (E, \vec{p}) = (E, p_x, p_y, p_z)$$

Then $p \cdot p \equiv p^2 = E^2 - \vec{p}^2$ is Lorentz invariant. This defines the four-vector dot product.

Like we can with familiar space-vectors, we can represent the four-vector dot product as matrix multiplication. We will denote individual elements of a four-vector with Greek indices, as

$$p_\mu \, , \, p^\mu$$

which is the μ^{th} element of p. Then, the four-vector dot product can be expressed as:

$$p \cdot p \equiv p_\mu \eta^{\mu\nu} p_\nu \equiv p_\mu p^\mu = \sum_{\substack{\mu=0 \\ \nu=0}}^{3} p_\mu \eta^{\mu\nu} p_\nu = p^\mu \eta_{\mu\nu} p^\nu$$

Here, we use the standard notation that $p_0 = E$, and $p_1 = p_x$, $p_2 = p_y$, $p_3 = p_z$. We also employ the Einstein summation notation, where repeated indices are summed over. The matrix $\eta^{\mu\nu}$ implements the dot product for four-vectors:

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}^{\mu\nu}$$

In this class, we won't much distinguish upper from lower indices. We can use η to change upper to lower indices as

$$p_\mu = \eta_{\mu\nu} p^\nu$$ via matrix multiplication.

So, what did we gain from this? If we want to preserve relativistic energy, momentum, and angular momentum, then we should express the physical system in terms of four-vector dot products, as they are Lorentz invariant. Lorentz transformations can be implemented by a matrix Λ on a four-vector as:

$$p_\mu \to \Lambda_\mu{}^\nu p_\nu \text{ , using Einstein summation notation.}$$

Note that this is a linear transformation on the vector p_μ. If we demand that the four-vector dot product of two four vectors p and q is Lorentz invariant, we have:

$$p\cdot q = p_\mu \eta^{\mu\nu} q_\nu \to p_\rho \Lambda^\rho{}_\mu \eta^{\mu\nu} \Lambda_\nu{}^\sigma q_\sigma = p_\mu \eta^{\mu\nu} q_\nu .$$

That is, just by relabling indices:

$$p_\mu \Lambda^\mu{}_\rho \eta^{\rho\sigma} \Lambda_\sigma{}^\nu q_\nu = p_\mu \eta^{\mu\nu} q_\nu \quad \text{for all } p, q.$$

Therefore, a Lorentz transformation implemented by a matrix Λ leaves η invariant:

$$\Lambda^\mu{}_\rho \eta^{\rho\sigma} \Lambda_\sigma{}^\nu = \eta^{\mu\nu}.$$

The set of matrices Λ that satify this constraint are called $O(3,1)$. Also, I will often call η the "flat-space" or Minkowski metric, or sometimes just the metric. Let's see how these Lorentz transformations work in a couple of examples.

Ex One type of Lorentz transformation is just a rotation along a fixed axis. Let's consider a rotation about the x-axis by an angle θ.

Such a rotation is implemented by a matrix $\mathbb{M}$ acting on the momentum vector $\vec{p}$:

$$\vec{p} \to \mathbb{M}\vec{p} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} p_x \\ p_y\cos\theta - p_z\sin\theta \\ p_y\sin\theta + p_z\cos\theta \end{pmatrix}$$

Energy is a scalar, and so remains unchanged under a rotation. Therefore, the matrix Λ that implements a rotation on the four-vector p is

$$p_\mu \to \Lambda_\mu{}^\nu p_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} E \\ p_x \\ p_y\cos\theta - p_z\sin\theta \\ p_y\sin\theta + p_z\cos\theta \end{pmatrix}$$

The rotation can also be implemented by acting from the right as:

$$p_\mu \to p_\nu \Lambda^\nu{}_\mu = \begin{pmatrix} E & p_x & p_y & p_z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

Note that the matrix Λ does indeed satisfy:

$$\Lambda^\mu{}_\rho \eta^{\rho\sigma} \Lambda_\sigma{}^\nu = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}$$

and so is indeed a Lorentz Transformation.

Ex Let's now consider a different type of Lorentz transformation: a boost. A Lorentz Boost is a change of ~~an~~ inertial reference frame implemented by moving with a relative velocity. Let's consider boosting along the z-axis by changing the relative velocity of the ~~an~~ inertial frame by ~~an~~ a velocity $\vec{\beta} = \beta\hat{z}$. In natural units, the speed β is a fraction of the speed of light, c and so $-1 \leq \beta \leq 1$ (negative because we could boost in the opposite direction). $|\beta| = 1$ means the boost is by the speed of light c (which of course isn't possible).

Boosting along the z axis means that p_x and p_y are unchanged by the boost:

$$p_x \to p_x$$
$$p_y \to p_y$$

while the energy E and p_z change. Under this boost, the energy E becomes

$$E \to \gamma(E + \beta p_z)$$, where γ is called the boost factor

and is $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. (In SI units this would be: $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$, where v is the velocity of the boost.)

The z-component of momentum transforms as:

$$p_z \to \gamma(p_z + \beta E)$$

Note that both of these transformations are linear.

they are represented by a linear combination of energy E and momentum p_z. This can therefore be implemented by a matrix Λ acting on the four-vector p. Acting from the left, we have:

$$p_\mu \to \Lambda_\mu{}^\nu p_\nu = \begin{pmatrix} \gamma & 0 & 0 & \gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\beta & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \gamma E + \gamma\beta p_z \\ p_x \\ p_y \\ \gamma\beta E + \gamma p_z \end{pmatrix}$$

Acting instead from the right, we have,

$$p_\mu \to p_\nu \Lambda^\nu{}_\mu = \begin{pmatrix} E & p_x & p_y & p_z \end{pmatrix} \begin{pmatrix} \gamma & 0 & 0 & \gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma\beta & 0 & 0 & \gamma \end{pmatrix}$$

In homework, you will show that the matrix Λ does satisfy the constraint that all Lorentz transformations must satisfy:

$$\Lambda^\mu{}_\rho \eta^{\rho\sigma} \Lambda_\sigma{}^\nu = \eta^{\mu\nu}.$$

Break //

Okay, so we know how to use and transform four-vectors. What can we do with them? Four-vectors combine the energy and momentum of a particle, and so naturally encode application of energy and momentum conservation. Well, for a system that conserves relativistic energy, momentum, and angular momentum, we describe the system with Lorentz transformation ~~transformation~~ invariant four-vector dot products, $p \cdot q$, for particles with momenta p and q.

A very common problem that we want to analyze in particle physics is the decay of an unstable particle. The vast majority of particles decay to two or more particles. ("Only" electrons and protons do not decay.) For example, the pion is a particle composed of up and down quarks that decays to two photons. The pion has a mass of about 135 MeV and in its decay to photons energy and momentum are (of course) conserved. (We'll talk about angular momentum conservation a bit later.) How can we understand this decay using four-vectors?

Of course, four-vectors depend on the frame in which they are evaluated, so we need to pick a frame to analyze this decay. The pion is massive so we can boost to the frame where the pion is at rest. In this frame, the four-momentum of the pion is:

$$p_\pi = (m_\pi, 0, 0, 0)$$

That is, when the pion is at rest, its energy is just set by the pion mass, m_π. Now, we want to determine the four-momentum of the two ~~pions~~ photons in the pion decay. We can again perform a rotation (= Lorentz transformation) to have the photons travel along the z-axis. By momentum conservation, the photons must be traveling back-to-back from the pion decay:

$$\gamma \longleftarrow \pi \longrightarrow \gamma$$

We can express the four-vectors of the two photons as:

$$p_{\gamma_1} = (E_{\gamma_1}, 0, 0, p_{\gamma_1})$$

$$p_{\gamma_2} = (E_{\gamma_2}, 0, 0, p_{\gamma_2})$$

By energy and momentum conservation the sum of the four-vectors of the photons must add up to the pion's four momentum:

$$p_\pi = p_{\gamma_1} + p_{\gamma_2} \Rightarrow \begin{aligned} m_\pi &= E_{\gamma_1} + E_{\gamma_2} \\ 0 &= p_{\gamma_1} + p_{\gamma_2} \end{aligned}$$

We can denote $p_{\gamma_1} = p_z$ and $E_{\gamma_1} = E$, then the four vectors of the photons are:

$$p_{\gamma_1} = (E, 0, 0, p_z)$$

$$p_{\gamma_2} = (m_\pi - E, 0, 0, -p_z)$$

Now, we enforce the masslessness of photons. We have

$$p_{\gamma_1} \cdot p_{\gamma_1} = m_{\gamma_1}^2 = 0 = E^2 - p_z^2$$

$$p_{\gamma_2}^2 = 0 = (m_\pi - E)^2 - p_z^2$$

Then, $E = p_z$ and $E = \frac{m_\pi}{2}$. Then, the four-vectors of the photons are

$$p_{\gamma_1} = \left(\frac{m_\pi}{2}, 0, 0, \frac{m_\pi}{2}\right)$$

$$p_{\gamma_2} = \left(\frac{m_\pi}{2}, 0, 0, -\frac{m_\pi}{2}\right)$$

Note that $p_\pi^2 = (p_{\gamma_1} + p_{\gamma_2})^2 = \cancelto{0}{p_{\gamma_1}^2} + 2 p_{\gamma_1} \cdot p_{\gamma_2} + \cancelto{0}{p_{\gamma_2}^2} = 2 p_{\gamma_1} \cdot p_{\gamma_2}.$

One can check this with the photons' four-vectors. We will refer to a four-vector for which

$$p^2 = m^2 \quad (\text{like for } p_\pi)$$

as "on the mass shell" or just "on-shell" for short. To find the four-vector in any other frame, we just Lorentz transform appropriately. While we just considered the case when the decay products are massless, one can also consider the case when the decay products are massive.

For the remainder of this lecture, we will introduce relativistic wave equations. In homework, you will work with four-vectors more, and in the next lecture we will talk more about relativistic equations of motion. As particle physics is the realm of both quantum mechanics and special relativity, we should have (or would expect to have) a wave-function equation like the Schrödinger equation that describes time evolution of the system. Recall that the Schrödinger equation is

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

There are a couple of problems in attempting to use the Schrödinger equation to describe particle physics. The first, and biggest problem is that the Schrödinger equation represents conservation of non-relativistic energy. To write the Schrödinger equation, we have identified

$$E \leftrightarrow +i\hbar\frac{\partial}{\partial t}, \quad \vec{p} \leftrightarrow -i\hbar\vec{\nabla}.$$

and so the Schrödinger equation represents

$$E = \frac{p^2}{2m} + V,$$ which is <u>not</u> invariant to Lorentz transformations.

Additionally, the Schrödinger Equation treats space and time differently. The Schrödinger equation describes the <u>time</u> evolution of the wave-function. Why time only? Time shouldn't be special in relativity, though in the Schrödinger equation, time just marches invariably on. As a final point, what is a potential in relativity?

So, it appears that, in order to develop quantum mechanics relativistically, we need to abandon the Schrödinger equation. How do we proceed, then? Just like the Schrödinger equation encoded non-relativistic energy conservation, we should find a wave equation that encodes relativistic energy conservation. So, starting from the relativistic energy-momentum relation (putting back the c's for now):

$$E^2 - \vec{p}^2c^2 - m^2c^4 = 0,$$ we replase with the

canonical quantum mechanical operators:

$$E \leftrightarrow +i\hbar\frac{\partial}{\partial t}, \quad \vec{p} \leftrightarrow -i\hbar\vec{\nabla}$$

and act on a wave-function $\phi(x,t)$. We then find the Klein-Gordon Equation:

$$\left(-\hbar^2\frac{\partial^2}{\partial t^2} + \hbar^2c^2\nabla^2 - m^2c^4\right)\phi(x,t) = 0.$$

Because this equation has both c and $\hbar$, it is indeed a relativistic quantum mechanical wave equation. Going back to natural units, this is

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi(x,t) = 0.$$

This equation is solved by the solution

$$\phi(x,t) = e^{ip\cdot x}$$

where $p\cdot x = Et - \vec{p}\cdot\vec{x}$, is the four-vector dot product. To see that this is indeed a solution, note that

$$\frac{\partial^2}{\partial t^2}\phi(x,t) = -E^2\phi(x,t)$$

and $$-\nabla^2\phi(x,t) = \vec{p}^{\,2}\phi(x,t).$$

Then, the Klein-Gordon equation implies:

$$\left(E^2 - \vec{p}^{\,2} - m^2\right)\phi(x,t) = 0$$

This is indeed true if the four vector p is on-shell: $p^2 = m^2$. Thus, we often say that solutions to the Klein-Gordon equation are on-shell solutions. Note also that the Klein-Gordon equation is Lorentz invariant: the solution $e^{ip\cdot x}$ depends on a four-vector dot product.

A couple other things to note: $p\cdot x$ has dimensions of angular momentum, $\hbar$, and a classical action. Interesting. Also, as a second-order equation, the Klein-Gordon equation has two solutions: $E>0$ and $E<0$. Huh? What does $E<0$ mean? More next time...