MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF PHYSICS
8.033 FALL 2021

LECTURE 7
4-MOMENTUM AND 4-VELOCITY

7.1 Transforming energy and momentum between reference frames

The requirement that all observers measure the speed of light to be c has led us to rather different formulations of energy and momentum: a body of *rest mass* m (i.e., the mass that we measure it to have when it is at rest with respect to us) moving with velocity $\mathbf{u}$ has an energy E and a momentum $\mathbf{p}$ given by

$$E = \gamma(u)mc^2\ , \qquad \mathbf{p} = \gamma(u)m\mathbf{u}\ . \tag{7.1}$$

These quantities respect conservation laws: a system's total E and $\mathbf{p}$ are conserved as its constituents interact with one another. In the limit $u/c \ll 1$, these formulas reduce to

$$E = mc^2 + \frac{1}{2}mu^2\ , \qquad \mathbf{p} = m\mathbf{u} + O(u^3)\ . \tag{7.2}$$

This makes it clear that Newtonian momentum agrees with relativistic momentum for speeds much smaller than c. The energies likewise agree in this limit, provided we account for the body's *rest energy* mc^2. In the vast majority of circumstances a body's rest energy is bound up in the body, and cannot be "used" for anything in their interaction, so it can be ignored. The relativistic quantities and the Newtonian quantities thus agree perfectly when $u \ll c$.

Suppose we measure a body to have energy E_L and momentum $\mathbf{p}_L$ in our laboratory. What energy E_T and momentum $\mathbf{p}_T$ will an observer moving past our lab in a train with velocity $\mathbf{v} = v\mathbf{e}_x$ measure the body to have? To figure this out, follow this recipe:

1. Deduce the 3-velocity $\mathbf{u}_L$ of the body in the lab from the values of E_L and $\mathbf{p}_L$.
2. Use the velocity addition formulas to compute the 3-velocity of $\mathbf{u}_T$ of the body as measured by observers on the train.
3. From $\mathbf{u}_T$, compute E_T and $\mathbf{p}_T$.

You will work through these steps on a problem set. The result you find is

$$\begin{aligned} E_T = \gamma(E_L - vp^x_L)\ , \qquad & p^x_T = \gamma(p^x_L - vE_L/c^2)\ , \\ p^y_T = p^y_L\ , \qquad & p^z_T = p^z_L\ . \end{aligned} \tag{7.3}$$

Tweaking notation slightly, we rewrite this

$$\begin{pmatrix} E_T/c \\ p^x_T \\ p^y_T \\ p^z_T \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E_L/c \\ p^x_L \\ p^y_L \\ p^z_L \end{pmatrix}\ . \tag{7.4}$$

In other words, the *relativistic formulations of energy and momentum form a set of quantities that transform under a Lorentz transformation.*

7.2 An invariant for energy and momentum

Recall that we found $\Delta s^2 = -c^2\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$ is a Lorentz invariant: all Lorentz frames agree on the value of Δs^2 between two events. Can we do something similar with energy and momentum?

Looking at how E and $\mathbf{p}$ behave under a Lorentz transformation, let's think of energy as the "timelike" component of momentum (E/c actually — which hopefully makes sense since we need our quantities to have the right dimensions). Let's see what happens when we examine "negative time bit squared" plus "space bit squared":

$$-\frac{E^2}{c^2} + (p^x)^2 + (p^y)^2 + (p^z)^2 = -\frac{E^2}{c^2} + |\mathbf{p}|^2 \ . \tag{7.5}$$

Plug into this

$$E^2 = \gamma^2 m^2 c^4 = \frac{m^2c^4}{1-u^2/c^2} \ , \tag{7.6}$$

$$|\mathbf{p}|^2 = \gamma^2 m^2 u^2 = \frac{m^2u^2}{1-u^2/c^2} \ . \tag{7.7}$$

Putting these together, we have

$$\begin{aligned} -\frac{E^2}{c^2} + |\mathbf{p}|^2 &= \frac{m^2u^2 - m^2c^2}{1-u^2/c^2} \\ &= -m^2c^2 \ . \end{aligned} \tag{7.8}$$

Let us reorganize this expression slightly:

$$E^2 - |\mathbf{p}|^2c^2 = m^2c^4 \qquad \text{or} \qquad E^2 = |\mathbf{p}|^2c^2 + m^2c^4 \ . \tag{7.9}$$

In other words, although different Lorentz frames will measure E and $\mathbf{p}$ differently, all frames agree that E^2 and $|\mathbf{p}|^2$ are related by the expressions given in Eq. (7.9).

Notice that if $m = 0$, then $|\mathbf{p}| = E/c$: massless bodies carry non-zero momentum. This relationship corresponds perfectly to the energy and momentum carried by electromagnetic radiation; compare with the Poynting vector if you need a refresher in this concept.

7.3 The 4-momentum

By virtue of the way in which E/c and $p^{x,y,z}$ transform, we can see that they behave exactly like the components of the displacement 4-vector. This tells us that we really should define a 4-vector whose components all have the dimensions of momentum:

$$\vec{p} = \sum_{\mu=0}^{3} p^\mu \vec{e}_\mu \ , \tag{7.10}$$

with

$$p^0 = E/c \ , \qquad p^1 = p^x \ , \qquad p^1 = p^y \ , \qquad p^3 = p^z \ . \tag{7.11}$$

This $\vec{p}$ is then a geometric object: observers in all Lorentz frames use this 4-vector to describe the system's energy and momentum, but break it up into components and unit vectors differently. If the components and unit vectors according to $\mathcal{O}$ are p^μ and $\vec{e}_\mu$, then an observer $\mathcal{O}'$ constructs $\vec{p}$ using

$$p^{\mu'} = \Lambda^{\mu'}{}_{\alpha} p^\alpha \ , \qquad \vec{e}_{\mu'} = \Lambda^{\alpha}{}_{\mu'} \vec{e}_\alpha \tag{7.12}$$

(switching to the Einstein summation convention). The matrix elements $\Lambda^{\mu'}{}_{\alpha}$ perform the Lorentz transformation of events from the frame of $\mathcal{O}$ to the frame of $\mathcal{O}'$; the matrix elements $\Lambda^{\alpha}{}_{\mu'}$ perform the inverse transformation.

The reason why this is useful for us is that conservation of energy and conservation of momentum are now combined into a single law: the *conservation of 4-momentum.* Suppose N_i bodies interact, resulting in N_f bodies afterwards. Then,

$$\sum_{j=1}^{N_i} \vec{p}_j^{\,\rm init} = \sum_{j=1}^{N_f} \vec{p}_j^{\,\rm final} \ , \tag{7.13}$$

where $\vec{p}_j^{\,\rm init}$ is the initial 4-momentum of particle j, and $\vec{p}_j^{\,\rm final}$ is the final 4-momentum of particle j.

7.4 4-vectors in general; scalar products of 4-vectors

Let's pause a momentum to reflect on the logic by which we assembled the 4-momentum. We essentially followed the following recipe:

1. We found that a grouping of 4 quantities plays a meaningful role in physics: $p^0 = E/c$, $p^{1,2,3} = p^{x,y,z}$.
2. We found that when we change reference frames, these 4 quantities are transformed to the new frame by the Lorentz transformation exactly as the components of the 4-displacement are: $p^{\mu'} = \Lambda^{\mu'}{}_{\alpha} p^\alpha$.
3. Since it behaves under the transformation law exactly like the 4-vector we discussed previously, we define p^μ as the components of a new 4-vector, $\vec{p}$, and use this 4-vector as a tool in our physics moving forward.

We can do this for *any* set of 4 quantities that turns out to be meaningful for our analysis. In other words,

> If any set b^μ with $\mu \in [0, 1, 2, 3]$ has the property that when we change reference frames their values are related by a Lorentz transformation, $b^{\mu'} = \Lambda^{\mu'}{}_{\alpha} b^\alpha$, then b^μ represent the components of a 4-vector: $\vec{b} = b^\mu \vec{e}_\mu$.

Once we have identified these quantities as the components of a 4-vector, we can start identifying invariants. Whatever the vector $\vec{b}$ represents, we are guaranteed that all Lorentz frames agree on the value of $-(b^0)^2 + (b^1)^2 + (b^2)^2 + (b^3)^2$. In fact, it is not hard to show that we can define a more general notion of an invariant. Suppose $\vec{a} = a^\mu \vec{e}_\mu$ and $\vec{b} = b^\mu \vec{e}_\mu$. Then,

$$\vec{a} \cdot \vec{b} \equiv -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3 \tag{7.14}$$

is a Lorentz invariant: all Lorentz frames agree on the value of $\vec{a} \cdot \vec{b}$. This is simply proven by transforming the components of $\vec{a}$ and $\vec{b}$ to another frame and then showing that the right-hand side of (7.14) in the new frame is unchanged from its value in the original frame.

Equation (7.14) defines what we call the "scalar product" between two 4-vectors. A (rather obvious) corollary is that the scalar product of any 4-vector with itself is a Lorentz invariant. Two quantities we've recently examined can be rephrased using this definition:

$$\Delta\vec{x}\cdot\Delta\vec{x} = \Delta s^2\ , \tag{7.15}$$
$$\vec{p}\cdot\vec{p} = -m^2c^2\ . \tag{7.16}$$

The resemblance to the invariant interval Δs^2 gives us a convention for describing 4-vectors. For any 4-vector $\vec{a}$, if

$$\vec{a}\cdot\vec{a} < 0 \tag{7.17}$$

then we say that $\vec{a}$ is *timelike*. This means that we can find a Lorentz frame in which only the time component of $\vec{a}$ is non-zero: $\vec{a}$ has no spatial components in that frame. If

$$\vec{a}\cdot\vec{a} > 0 \tag{7.18}$$

then we say that $\vec{a}$ is *spacelike*. There exists a[1] Lorentz frame in which $\vec{a}$ has no component in the time direction; it points purely in a spatial direction. Finally, if

$$\vec{a}\cdot\vec{a} = 0 \tag{7.19}$$

then $\vec{a}$ is *lightlike* or *null*. In all Lorentz frames, $\vec{a}$ points along light cones.

Notice that $\vec{p}$ is either timelike or lightlike, and is only lightlike for $m = 0$.

7.5 4-velocity

In Newtonian mechanics, velocity and momentum were related by a factor of the body's mass. Let's do the same thing using the 4-momentum; we will define the quantity that results as the 4-velocity:

$$\vec{u} = \frac{1}{m}\vec{p}\ . \tag{7.20}$$

What does this quantity mean? Let's look at its components:

$$u^0 = \frac{p^0}{m} = \frac{E}{mc} = \gamma\, c\ , \tag{7.21}$$
$$u^1 = \frac{p^1}{m} = \gamma\,(\mathbf{u})^x\ , \tag{7.22}$$
$$u^2 = \frac{p^2}{m} = \gamma\,(\mathbf{u})^y\ , \tag{7.23}$$
$$u^3 = \frac{p^3}{m} = \gamma\,(\mathbf{u})^z\ . \tag{7.24}$$

(Note the somewhat unusual notation on the spatial components: $(\mathbf{u})^x$ means the x component of the 3-vector $\mathbf{u}$, and likewise for the y and z components.) The spatial components of $\vec{u}$ look just like "normal" 3-velocity, but multiplying by a factor of γ. How do we interpret this factor?

[1] Actually, *many* such Lorentz frames: once we find one, any Lorentz frame that is related to the first by a rotation will do the trick.

Consider someone passing by with 3-velocity $\mathbf{u}$. That person's clocks run slow according to you: as an interval $d\tau$ passes on their clock, an interval dt passes on your clock, with

$$dt = \gamma\, d\tau \ . \tag{7.25}$$

Notice that their clock is running slow — if $\gamma = 2$, then we measure 2 seconds passing for every 1 second interval that they record.

We define the interval $d\tau$ as the *proper time*: it is an interval of time according to the clock of the observer (or object) who we say is moving. Note that the word "proper" in this case comes from an older meaning denoting "belonging to oneself." Hence an observer's proper time is the time which that observer measures.

Proper time is a useful quantity because it is a Lorentz invariant: *all* Lorentz frames agree that the observer in motion measures a time interval $d\tau$. That won't be the time interval we measure as observer $\mathcal{O}$ whizzes by us at 90% of the speed of light; it won't be what our friend measures as they whizz by at 90% of the speed of light in another direction; but we all agree that it *is* what $\mathcal{O}$ measures. Hence it is a useful benchmark which we can agree on.

With this in mind, let's re-examine the spatial components of the 4-velocity:

$$u^x = \gamma\,(\mathbf{u})^x = \gamma\frac{dx}{dt} = \frac{dx}{d\tau} \ , \tag{7.26}$$

$$u^y = \gamma\,(\mathbf{u})^y = \gamma\frac{dy}{dt} = \frac{dy}{d\tau} \ , \tag{7.27}$$

$$u^z = \gamma\,(\mathbf{u})^z = \gamma\frac{dz}{dt} = \frac{dz}{d\tau} \ , \tag{7.28}$$

Let's also look at the timelike component:

$$u^t = \gamma\, c = \gamma\, c\frac{dt}{dt} = c\frac{dt}{d\tau} \ . \tag{7.29}$$

Comparing with how we defined the displacement 4-vector, we see that

$$\vec{u} = \frac{d\vec{x}}{d\tau} \ . \tag{7.30}$$

The 4-velocity is the rate at which something moves through spacetime *per unit proper time.*

It's worth computing the invariant associated with the 4-velocity:

$$\vec{u}\cdot\vec{u} = \frac{1}{m^2}\vec{p}\cdot\vec{p} = -\frac{m^2c^2}{m^2} = -c^2 \ . \tag{7.31}$$

Notice that the 4-velocity of a body which is at rest in some Lorentz frame has the same $\vec{u}\cdot\vec{u}$ as a body which is moving $0.99999999999c$ in that frame.

7.6 4-velocity contrasted with 3-velocity

We now have two important ways to characterize a moving body's motion:

- 3-velocity $\mathbf{u} = d\mathbf{x}/dt$ describes motion through **space** per unit **time**. Both "space" and "time" are frame-dependent concepts, and so $\mathbf{u}$ depends on the frame in which it is measured.

- 4-velocity $\vec{u} = d\vec{x}/d\tau$ describes motion through **spacetime** per unit **proper time**. It is a frame-independent, geometric object; the same $\vec{u}$ is used by all observers.

A major conceptual difference between these two quantities is how we regard them when observed in different Lorentz frames:

- As a frame-independent geometric object, all observers agree on an object's 4-velocity $\vec{u}$. They assign it different components, however, and use different unit vectors in expanding $\vec{u}$ into components:

$$\vec{u} = u^\mu \vec{e}_\mu = u^{\alpha'} \vec{e}_{\alpha'} \ , \tag{7.32}$$

  where

$$u^{\alpha'} = \Lambda^{\alpha'}{}_\mu u^\mu \ , \qquad \vec{e}_{\alpha'} = \Lambda^\mu{}_{\alpha'} \vec{e}_\mu \ . \tag{7.33}$$

- The 3-vector is actually different in the two frames. Given $\mathbf{u}$, we find the components of $\mathbf{u}'$ which describe the body's motion in a new frame by applying the velocity addition formulas: if the relative motion of the two frames is given by $\mathbf{v} = v\mathbf{e}_x$, then

$$(\mathbf{u})^{x'} = \frac{(\mathbf{u})^x + v}{1 + (\mathbf{u})^x\, v/c^2} \ , \tag{7.34}$$

$$(\mathbf{u})^{y'} = \frac{(\mathbf{u})^y}{\gamma(v)(1 + (\mathbf{u})^x\, v/c^2)} \ , \tag{7.35}$$

$$(\mathbf{u})^{z'} = \frac{(\mathbf{u})^z}{\gamma(v)(1 + (\mathbf{u})^x\, v/c^2)} \ . \tag{7.36}$$