Scott Hughes *15 March 2005*

Massachusetts Institute of Technology
Department of Physics
8.022 Spring 2005

Lecture 11: Spacetime: Introduction to Special Relativity

Preamble: The vast majority of material presented in this course consists of subjects that are challenging, but that I am convinced every 8.022 student can — and should! — learn well. *Special relativity is the only exception to this rule.* The material that we will begin covering today (and that we will conclude on Thursday) is, in my experience, by *far* the most confusing subject that we cover all term. Further, it is not *strictly* necessary to learn this material inside and out in order to understand and excel in 8.022.

You may wonder, given this, why we are bothering to study it at all. The reason is that special relativity "explains" the magnetic field: what we find at the end of all of this is that magnetic forces are nothing but electric forces viewed in another frame of reference. Magnetic fields are just electric fields "in drag", so to speak. Conceptually, this is an *extremely* important point: even if you do not fully grasp the details of where many of the formulas we are about to derive come from, I firmly believe every 8.022 student should understand what magnetic forces "look like" in different frames of reference.

The punchline of this preamble is that you should not be too worried if 8.022's discussion of special relativity makes your brain hurt. Those of you who take more physics courses *will* encounter these concepts several more times; if you don't quite get it now, you will have plenty of chances to get it later[1]. Try to get the flavor of what is happening now, and practice manipulating these formulas, even if it is not 100% clear what is going on.

11.1 Principles of special relativity

Special relativity depends upon two postulates:

1. The laws of physics are the same in all frames of reference.
2. The speed of light is the same in all frames of reference.

These postulates beg the question — what is a frame of reference? A frame of reference (more properly an *inertial* frame of reference) is any set of coordinates that can be used to describe a non-accelerating observer. In the following discussion we will compare measurements in two frames of reference: a station, and a train that is moving through this station in the x direction with constant speed v.

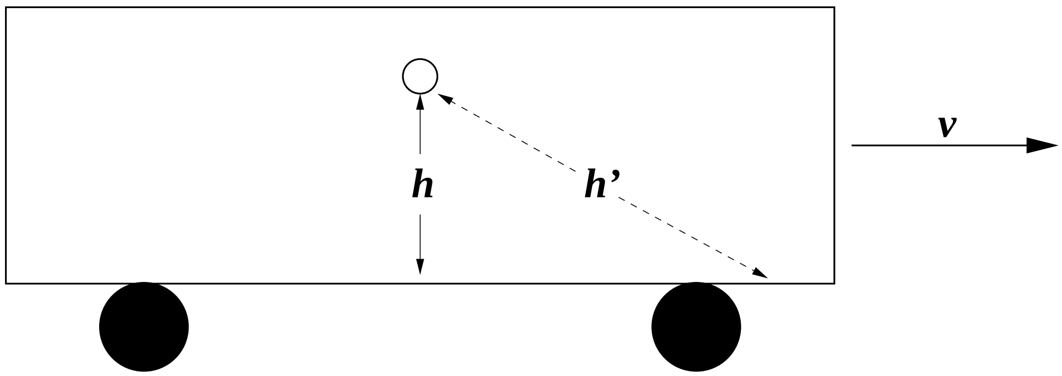
11.2 Time dilation

These postulates have consequences that are rather amazing. In particular, it means that inertial observers in different frames of reference measure *different intervals of time* between

[1] I studied special relativity in 4 separate courses as an undergrad; I honestly can't say I really knew what I was doing until about halfway through the third time. Even then, the fourth study was useful to me.

events, and *different spatial separations* between events. As we'll see in a moment, clocks that appear to be moving run slow; rulers that appear to be moving are shrunk.

Consider a pulse of light emitted from a bulb in the center of a train. The train is moving at speed v in the x direction with respect to a station. The pulse of light is directed toward a photosensor on the floor of the train, directly under the bulb. Two *events* are of interest to observers on the train and in the station: the emission of light from the bulb, and the reception of this light by the photodetector.



Our first question is: How much time passes between the emission of this pulse and its reception on the floor?

The amount of time that is measured to pass is different for observers on the train (at rest with respect to the bulb and the photosensor) than it is for observers in the station (who see the bulb and the photosensor in motion). Let's look at things in the train frame first: the amount of time between emission and reception is just the vertical distance h the light pulse travels divided by the speed of light c:

$$\Delta t_{\text{train}} = \frac{h}{c}$$

How do things look to observers in the station, watching the train go by? They agree that the pulse of light moves vertically through a distance h. However, they claim it also moves through *horizontal* distance because the train is moving with speed v. If the elapsed time between emission and reception as seen in the station frame is $\Delta t_{\text{station}}$, then the pulse moves through a total distance $h' = \sqrt{h^2 + (v\Delta t_{\text{station}})^2}$. Since the speed of light is the same in all frames of reference, the time interval $\Delta t_{\text{station}} = h'/c$:

$$\Delta t_{\text{station}} = \frac{1}{c}\sqrt{h^2 + (v\Delta t_{\text{station}})^2}$$

Substitute $h = c\Delta t_{\text{train}}$:

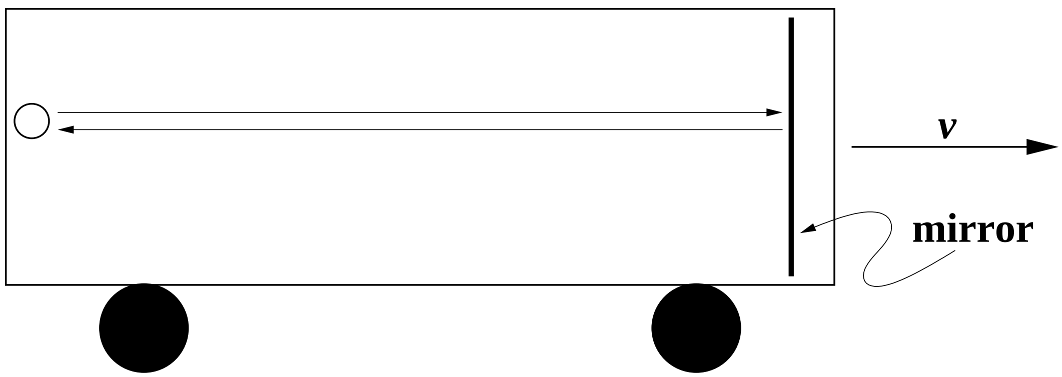
$$(\Delta t_{\text{station}})^2 = (\Delta t_{\text{train}})^2 + \frac{v^2}{c^2}(\Delta t_{\text{station}})^2$$

$$\begin{aligned} \Delta t_{\text{station}} &= \frac{\Delta t_{\text{train}}}{\sqrt{1 - v^2/c^2}} \\ &\equiv \gamma \Delta t_{\text{train}} \end{aligned}$$

where we have defined $\gamma = 1/\sqrt{1 - v^2/c^2}$; note that $\gamma \geq 1$. This relation tells us that *the moving clock runs slow*: the time interval measured by observers who see the clock moving (those in the station) is larger than the time interval measured by observers who are at rest with respect to it.

11.3 Length contraction

Let's move our light source to the back of the train. We flash it in the forward direction, where it reflects off a mirror at the front of the train. The light then returns to the light source, where we measure it. This is a way to measure the length of the railway car.



The two events of interest in this case are the emission of light by the bulb, and its reception after it bounces back to us.

Let's examine what's going on in the train's reference frame first. If the elapsed time between emission of the light and reception is Δt_{train}, we would infer that the length of the train is

$$\Delta x_{\text{train}} = c\Delta t_{\text{train}}/2$$

How do things look in the station frame? We break down the light travel into two pieces: emission to the mirror, then mirror back to reception. Let's say it takes a time $\Delta t_{\text{station, 1}}$ for the first piece — emission to the mirror. An interval $\Delta t_{\text{station, 2}}$ passes following reflection to travel back to reception. If the length of the train as measured in the station is $\Delta x_{\text{station}}$, then the travel times in each bounce is given by

$$\begin{aligned} \Delta t_{\text{station},1} &= (\Delta x_{\text{station}} + v\Delta t_{\text{station},1})/c \\ \Delta t_{\text{station},2} &= (\Delta x_{\text{station}} - v\Delta t_{\text{station},2})/c \end{aligned}$$

The time from emission to mirror is longer since the mirror is moving in the same direction as the light — the light has to chase the mirror on the first piece of the round trip, which takes extra time. On the second piece, the back of the train is rushing towards the light, so it takes less time.

We rearrange these expressions to isolate the travel times:

$$\begin{aligned} \Delta t_{\text{station},1} &= \frac{\Delta x_{\text{station}}}{c - v} \\ \Delta t_{\text{station},2} &= \frac{\Delta x_{\text{station}}}{c + v} \end{aligned}$$

Then, we add them together to get the total travel time:

$$\begin{aligned}\Delta t_{\text{station}} &\equiv \Delta t_{\text{station},1} + \Delta t_{\text{station},2} \\ &= \frac{2\Delta x_{\text{station}}/c}{1 - v^2/c^2} \\ &= 2\gamma^2 \Delta x_{\text{station}}/c\end{aligned}$$

From our discussion of time dilation, we also know that $\Delta t_{\text{station}} = \gamma \Delta t_{\text{train}} = 2\gamma \Delta x_{\text{train}}/c$. Combining this with our result for $\Delta t_{\text{station}}$, we get

$$\gamma \Delta x_{\text{train}} = \gamma^2 \Delta x_{\text{station}}$$

or

$$\Delta x_{\text{station}} = \Delta x_{\text{train}}/\gamma$$

Since $\gamma \geq 1$, this tells us that *moving objects are contracted*: the train's length according to observers that see it moving is less than the length measured in the "rest frame".

11.3.1 No contraction for perpendicular directions

What about the dimensions of the train that are perpendicular to the train's motion? Without getting too technical, a simple thought experiment can convince us that the perpendicular directions cannot be affected by the motion.

Imagine a train moving towards the mouth of a tunnel. Suppose that the tunnel is 4 meters tall in its rest frame, and that the train is 3.5 meters tall in its rest fram. Suppose the train is moving at 90% of the speed of light. What would happen if the Lorentz contraction shrunk the vertical size of the train?

In the rest frame of the tunnel, the train would appear very short: its height would be Lorentz contracted to $h_{\text{train, contracted}} = 3.5\,\text{meters}/\gamma = 1.53\,\text{meters}$. It fits into the tunnel with no problem.

However, it is equally valid to say that the train is at rest and the tunnel is moving toward it at $0.9c$. Observers in the rest frame of the tunnel claim that the tunnel is shrunk to a height $h_{\text{tunnel, contracted}} = 4.0\,\text{meters}/\gamma = 1.74\,\text{meters}$ — far smaller than the train itself!!!

In one frame, the train fits; in the other, it gets smashed. This is *not* consistent physics! Lorentz contraction of the perpendicular directions does not happen since it does not follow the rules that physics be the same in all reference frames.

11.3.2 Example: Cosmic ray muons

Muons are elementary particles that act, for all practical purposes, like heavy electrons. They are commonly created on earth when high energy protons (cosmic rays) slam into the upper reaches of the earth's atmosphere, several tens of kilometers about the surface. Muons decay into an electron and a neutrino in about 2×10^{-6} seconds. This suggests a mystery: how can a particle that is produced 20 kilometers about the earth possibly reach the ground if it decays in 2 microseconds? Even if the particle were moving at nearly the speed of light, it could only move about 600 meters before decaying.

Special relativity saves the day. Suppose the muon is moving at 99.99% of the speed of light. As seen on the surface of the earth, time dilation stretches out its lifetime to

$$\Delta t_{\text{muon}} = \frac{2 \times 10^{-6}\,\text{sec}}{\sqrt{1 - 0.9999^2}} = 1.4 \times 10^{-4}\,\text{sec}$$

With this lifetime, the muon can travel a distance

$$\Delta x_{\text{muon}} = 0.9999 \times c \times 1.4 \times 10^{-4}\,\text{sec} = 42\,\text{kilometers}$$

before decaying. The muon can easily make it from high up in the atmosphere to the ground.

It's also interesting to think about this from the muon's perspective. In this frame, the muon is at rest, and the earth's atmosphere is rushing past it at 99.99% of the speed of light. This means that the atmosphere is length contracted by a factor of $1/\gamma = 1/71$. From the muon's perspective, the atmosphere is so thin that it can easily reach the ground.

The two perspectives shown here encapsulate what the slogan "*The laws of physics are the same in all frames of reference*" really means. On the earth's surface, we say that the lifetime of the muon is time dilated until it lives long enough to reach the earth's surface. In the muon's rest frame, its lifetime is only 2 microseconds, so there's no way it could make it through tens of kilometers of atmosphere. However, in this frame, the atmosphere is length contracted until the muon can easily penetrate it in 2 microseconds. The two observers interpret the *details* of what is going on very differently. But both agree on the most important aspect of the phenomena: the muon reaches the ground.

11.4 Lorentz transformation

Length and time dilation are specific manifestations of a *general* consequence of special relativity: what we consider to be "time" and "space" in one reference frame is mixture of the notions of "time" and "space" as measured in another reference frame. The general relation between time and space coordinates in difference frames of reference is described by the *Lorentz transformation.*

The Lorentz transformation is a linear relation between the coordinates of the "rest frame" (for example, the train station) and the "moving frame" (for example, on the train). Let us denote the coordinates of the rest frame as (t, x, y, z); the coordinates of the moving frame are (t', x', y', z'). The moving frame moves with speed v in the x direction, as seen in the rest frame.

The general form of the transformation that relates the two coordinates is

$$\begin{aligned} x' &= Ax + Bt \\ t' &= Cx + Dt \\ y' &= y \\ z' &= z \end{aligned}$$

The y and z coordinates are the same in both frames since the relative motion is in the x direction. Note that the transformation is linear — there are no terms that go proportional to t^2, for example. This is necessary to ensure that the reference frames are inertial — if there were nonlinear terms, the frames would appear to be accelerating with respect to one another rather than moving at a relative velocity.

Our goal now is to calculate A, B, C, and D. We first require that the two coordinate systems must match at the origin: the coordinate $x' = 0$ must match $x = vt$:

$$\begin{aligned} x' &= Ax + Bt \\ &= Avt + Bt \\ &= 0 \end{aligned}$$

The first line is the general Lorentz tranformation relating x' to x and t. On the second line, we plug in $x = vt$. Setting $x' = 0$ and solving, we find.

$$\rightarrow B = -Av \ .$$

Matching at the origin also requires that $x = 0$ must match $x' = -vt'$:

$$\begin{aligned} x' &= A(x - vt) \qquad \text{(General transform for } x'\text{)} \\ &= -Avt \qquad (x = 0) \\ &= -vt' \qquad (x' = -vt') \\ &= -v(Cx + Dt) \qquad \text{(General transform for } t'\text{)} \\ &= -vDt \qquad (x = 0) \end{aligned}$$

Comparing the second and fifth lines, we have

$$A = D$$

OK, so far we've determined two of the four constants, and so we can rewrite the Lorentz transformation as

$$\begin{aligned} x' &= A(x - vt) \\ t' &= Cx + At \end{aligned}$$

To determine the remaining two constants, we think about the properties of light pulses as described by the two coordinate systems.

First, consider a light pulse that is sent out in the x direction. We shoot this pulse out at $t = 0$. The location of the pulse in the unprimed coordinate system is $x = ct$. This pulse will also move out parallel to the x' coordinate system since the relative motion of the two systems is in this direction.

Since the speed of light is the same in all reference frames, its position in the primed coordinate system is given by $x' = ct'$:

$$\begin{aligned} x' &= ct' \\ A(x - vt) &= c(Cx + At) \\ Act - Avt &= c^2Ct + cAt \\ -Av &= c^2C \\ \rightarrow C &= -Av/c^2 \end{aligned}$$

The Lorentz transformation now becomes

$$\begin{aligned} x' &= A(x - vt) \\ t' &= A(t - vx/c^2) \end{aligned}$$

To finally determine A, we shoot a light pulse out along the y axis. In the unprimed coordinate system, the location of the pulse is given by $y = ct$, $x = 0$. What is its location in the primed system? The motion of the frame itself will cause it to move in the x' direction. The total distance that the pulse moves is ct'. We therefore must have $(x')^2 + (y')^2 = c^2(t')^2$:

$$
\begin{aligned}
(x')^2 + (y')^2 &= c^2(t')^2 \\
A^2(x - vt)^2 + y^2 &= c^2A^2(t - vx/c^2)^2 \\
A^2(-vt)^2 + c^2t^2 &= c^2A^2t^2 \qquad (x = 0, y = ct) \\
A^2 &= \frac{1}{1 - v^2/c^2} = \gamma^2
\end{aligned}
$$

We thus have finally fixed the Lorentz transformation's form:

$$
\begin{aligned}
x' &= \gamma(x - vt) \\
t' &= \gamma(t - vx/c^2)
\end{aligned}
$$

This can be written more symmetrically by definining $\beta = v/c$ and multiplying the t transformation by c:

$$
\begin{aligned}
x' &= \gamma[x - \beta(ct)] \\
ct' &= \gamma[ct - \beta x]
\end{aligned}
$$

The inverse of this transformation is quite simple. Observers at rest on the train see the station moving in the $-x$ direction with speed v. This means that the Lorentz transformation is inverted by just reversing the direction of v — i.e., we replace v with $-v$:

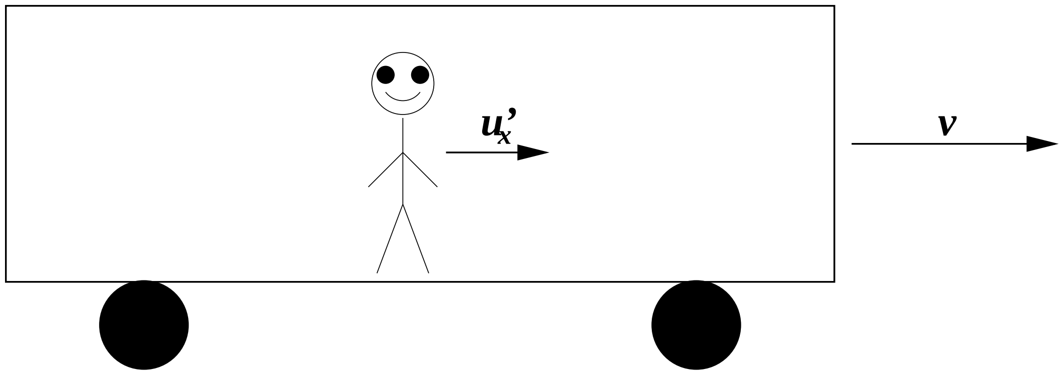
$$
\begin{aligned}
x &= \gamma(x' + vt') \\
t &= \gamma(t' + vx'/c^2)
\end{aligned}
$$

or

$$
\begin{aligned}
x &= \gamma[x' + \beta(ct')] \\
ct &= \gamma[ct' + \beta x']
\end{aligned}
$$

11.5 Transformation of velocity

An important result that follows from this is a formula for addition of velocity. Suppose I am in the moving train, walking from one end to the other with a speed $dx'/dt' = u'_x$:



With what speed u_x does someone in the station frame see me move? In old fashioned mechanics, we would just add my speed to the speed of the moving train, and so we'd have $u_x = u'_x + v$. To do it correctly in special relativity, we use the Lorentz transformation:

$$\begin{aligned} u_x &= \frac{dx}{dt} \\ &= \frac{\gamma(dx' + v dt')}{\gamma(dt' + v dx'/c^2)} \\ &= \frac{(dx'/dt') + v}{1 + (v/c^2)(dx'/dt')} \\ &= \frac{u'_x + v}{1 + (v u'_x/c^2)} \end{aligned}$$

This formula shows how it is impossible for anything to move faster than the speed of light. Suppose the train is moving at $v = 0.95c$. Suppose further that my walking speed within the train is $u'_x = 0.95c$. In old fashioned mechanics, we would predict that my walking speed as seen from the ground would be $1.9c$ — almost twice the speed of light. In special relativity, we find

$$\begin{aligned} u_x &= \frac{0.95c + 0.95c}{1 + 0.95 \times 0.95} \\ &= \frac{1.9c}{1.9025} \\ &= 0.9987c \end{aligned}$$

Close to, but still less than, the speed of light.

Suppose that I walk perpendicular to the train's motion: I walk with

$$u'_y = \frac{dy'}{dt'}$$

What is this as seen from the ground? Again, we use the Lorentz transformation:

$$u_y = \frac{dy}{dt}$$

$$
\begin{aligned}
&= \frac{dy'}{\gamma(dt' + v dx'/c^2)} \\
&= \frac{dy'/dt'}{\gamma[1 + (v/c^2)(dx'/dt')]} \\
&= \frac{u'_y}{\gamma[1 + (v u'_x/c^2)]}
\end{aligned}
$$

Likewise,

$$
u_z = \frac{u'_z}{\gamma[1 + (v u'_x/c^2)]}
$$

We will use these formulas in the next lecture to explore how things transform from the “viewpoint” of moving charges and currents.