Lecture 2: Special Relativity

Today and next lecture, we'll go through a lightning review of special relativity. You will have seen much of this in at least two other courses, so our coverage will be brief, but from a general point of view. This will provide the baseline for general relativity and we'll revisit these ideas throughout the course.

Special relativity follows from the simple observation:

<u>Axiom</u>: The speed of light c is the same in every inertial reference frame.

This is a very dense statement, so let's break it apart. First, the speed of light is just the distance that light travels in a set amount of time. That is in a time interval dt, light travels a distance dx so that

$$\frac{dx}{dt} = c$$

In writing this expression, we have implicitly assumed that time interval "dt" and distance "dx" are well-defined. This follows from the last phrase in the axiom: "inertial reference frame." The time interval dt and distance dx are measured in an inertial reference frame. An *inertial reference frame* is a frame in which there are no forces acting on you. By Newton's second law, another way to say this is that an inertial reference frame is one in which you are traveling at a constant velocity.

However, the definition via forces is more general. For instance, are you, sitting in your seat, in an inertial reference frame? Though you don't appear to be accelerating, you should feel a force; otherwise you wouldn't be stuck in your seat. We'll dive into consequences of this observation in this course, but for now, we will just use this definition.

So, for <u>any</u> inertial reference frame in which we measure distances and times, we always measure the ratio of the distance light travels dx to the time it takes dt to equal the constant c: $\frac{dx}{dt} = c$.

Let's massage this expression. This can equivalently be written as:

$$dx = c\,dt \quad \text{or} \quad dx^2 = c^2 dt^2,$$

where we have just squared the two sides of the equation. Let's further move everything to one side:

$$0 = -c^2 dt^2 + dx^2$$

This is an extremely useful expression. This quantity

$ds^2 \equiv -c^2 dt^2 + dx^2$ is called the "space-time interval".

If $ds^2 = 0$, we call the interval "light-like" because dx and dt represent the distance and time, respectively, that light travels. In general, ds^2 for any space interval dx and time interval dt is independent of inertial frame.

Working in 3 space and one time dimension, the space-time interval is:

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \equiv -c^2dt^2 + d\vec{x}^2$$

where $\vec{x}$ is the vector of spatial coordinates. This kind of looks like the square of a vector. Let's call the space-time interval vector ds:

$$ds = (c\,dt, dx, dy, dz)$$

This is called a four-vector because it has four components; one for each space-time dimension. The way that it is squared is interesting. The dot product of ds with itself is defined as:

$$ds \cdot ds = ds^2 = -c^2dt^2 + d\vec{x}^2.$$

This can be accomplished by multiplying by a matrix η which is called the "flat metric" as:

$$(c\,dt,\ dx\ \ dy\ \ dz)\begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}\begin{pmatrix} c\,dt \\ dx \\ dy \\ dz \end{pmatrix} \doteq ds^2.$$

$$\eta = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

and another way to define it is as the matrix of coefficients of the space-time interval.

This can be equivalently be expressed with Einstein summation notation as

$$ds^2 = ds^\mu \eta_{\mu\nu} ds^\nu = \sum_{\substack{\mu=0,\dots,3 \\ \nu=0,\dots,3}} ds^\mu \eta_{\mu\nu} ds^\nu$$

μ and ν are spacetime indices that range over 0 (time), 1,2,3 (space) coordinates and repeated indices mean implicit summation.

With this powerful notation established, we can then discuss <u>Lorentz transformations</u>; or relationships between different inertial frames. Let's call the space-time interval in one frame ds and in another frame ds'. Different inertial frames are related by a relative velocity or a relative angle between the coordinate bases. These are linear relationships: that is the space-time interval in the two frames are related to one another by matrix multiplication. That is, there exists a matrix Λ for which

$$ds' = \Lambda ds, \text{ or with indices } ds'^{\mu} = \Lambda^{\mu}{}_{\nu} ds^{\nu}$$

The space-time interval is independent of inertial frame so

$$ds^2 = (ds')^2 \text{ or that}$$

$$(ds')^2 = (ds')^T \eta \, ds' = (ds^T \Lambda^T) \eta (\Lambda ds) = ds^T \Lambda^T \eta \Lambda ds.$$

As this must hold for <u>any</u> interval ds, this produces a constraint on the matrix Λ:

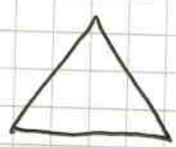
$$\Lambda^T \eta \Lambda = \eta, \text{ or that } \Lambda_{\mu}{}^{\rho} \eta_{\rho\sigma} \Lambda^{\sigma}{}_{\nu} = \eta_{\mu\nu}.$$

The set of all matrices $\{\Lambda\}$ that satisfy this requirement, or leave the metric invariant, is called the Lorentz group and denoted by $O(3,1)$. This is analogous to rotations implemented by a matrix $\mathbb{M}$ defined by:

$\mathbb{M}^T \mathbb{1} \mathbb{M} = \mathbb{1}$. Such a matrix $\mathbb{M}$ is called "orthogonal".

Any Lorentz transformation (boost or rotation) can be represented as an appropriate matrix Λ. Your book discusses this, so I won't go through examples here.

Finally for today, I want to introduce vectors and tensors. More generally, these objects are representations of the Lorentz group. The Lorentz group, the set of all possible Lorentz transformations, corresponds to the transformations of a system which leaves the physics unchanged. In our discussion thus far, we started with the constraint (speed of light c is the same in all inertial reference frames) to determine the possible transformations we could perform that left this observation (i.e. physics) unchanged. Such a transformation is called a symmetry. Symmetries are very familiar in colloquial conversation; for example, an equilateral triangle has a high degree of symmetry:



We can rotate by 120°, for example, and the triangle appears unchanged.

Lorentz transformations represent the symmetries of flat space-time. We can rotate and boost all we want and we will always observe the same phenomena in our experiments.

Vectors and tensors defined in space-time take different values depending on the particular coordinate system used.

For example, the interval four-vector ds is changed if we define a rotated coordinate system. Nevertheless, the interval squared is invariant to coordinate changes.

We can denote a general four-vector as $V = (V_0, V_1, V_2, V_3)$ or with indices as V_μ. The length or square of the four-vector is Lorentz invariant:

$$V^2 = V \cdot V = V_\mu V^\mu = V_\mu \eta^{\mu\nu} V_\nu = -V_0^2 + V_1^2 + V_2^2 + V_3^2.$$

However, under a change of coordinates to a new inertial frame, the four-vector transforms with one $O(3,1)$ matrix Λ:

$$V \to \Lambda V \quad \text{or} \quad V_\mu \to \Lambda_\mu{}^\nu V_\nu, \text{ in components.}$$

That one matrix Λ transforms V defines it to be a four-vector. That is, one matrix Λ represents the transformation of V to a new inertial frame. We equivalently say that a four-vector V transforms in the vector representation of the Lorentz group.

A tensor, at its simplest, is just an object with multiple indices, and so transforms with multiple $O(3,1)$ matrices. For example, a two-index tensor T transforms as:

$T \to \Lambda \Lambda T$ or, with indices
$$T_{\mu\nu} \to \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma T_{\rho\sigma}$$

An adage to keep in mind for determining what an object is is: "By their transformations thou shalt know them."

That is, the transformation properties of an object under Lorentz transforming *defines* what object it is. The two index tensor $T_{\mu\nu}$ transforms in the tensor representation of the Lorentz group.

We'll pick up from tensors next lecture and use this formalism to express electromagnetism and Maxwell's equations.