Physics 106b: Lecture 2 9 January, 2014

Relativity: Geometric Approach

I discuss a frame independent (geometric) formulation of special relativity. This is discussed briefly in Hand and Finch §12.11, in Taylor $15.8-11, and is the approach taken in *Spacetime Physics* by Taylor and Wheeler. For a more advanced account consult Chapter 1 of the notes to Ph136a.

Frame Independent Quantities

Event: A precise location in space (a point) at a precise moment in time, denoted by $\mathcal{P}$,$\mathcal{Q}$ etc..

We often like to think of an action to pinpoint the event, such as "an atom emits a flash of light". Observers in all inertial frames of reference agree on the event, but will label it by different space and time coordinates. Note that measurement of the length of an object is not an event, since it is not located at one point in space. Two events are involved: the coincidence of one end of object and ruler, and the coincidence of the other end of object and ruler.

Wordline: The motion of a particle is represented by its worldline connecting events at which the particle is present. It is convenient to parameterize the worldline by the *proper time*: the time measured by an ideal clock moving with the particle (and therefore agreed on by all observers) The worldline is then $\mathbf{x}(\tau)$ with $\mathbf{x}$ the displacement 4-vector measured from some reference origin event. Since light travels with a constant speed c, the worldline of a light pulse is a straight line of slope 1 in relativistic units.

Displacement 4-vector: The straight arrow from one event $\mathcal{P}$ to another event $\mathcal{Q}$, denoted by $\boldsymbol{\Delta x}$. In these notes I will use **bold** for 4-vectors; in class I use a double arrow. We can add 4-vectors to give a new 4-vector and multiply by a scalar in the obvious way. In a frame of reference S the 4-vector $\boldsymbol{\Delta x}$ has components $(\Delta t, \Delta x, \Delta y, \Delta z)$. Think of this ordering as the "zeroth, first, second, third" components, and I will sometimes write the components as $(\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3)$, with $\Delta x^0 \equiv \Delta t$, $\Delta x^1 \equiv \Delta x$, etc. (The reason for the superscript notation will be discussed in a later lecture.) In another frame of reference S' the 4-vector $\boldsymbol{\Delta x}$ has components $(\Delta t', \Delta x', \Delta y', \Delta z')$. This is a *different* component representation of the *same* 4-vector. The components are related by the Lorentz transformation introduced in the previous lecture.

The displacement 4-vector serves as the template for more general 4-vectors.

Principle of Relativity

The principle of relativity can be restated as:

> Every law of physics must be expressible as a geometric frame independent relationship between geometric frame independent objects.

We expect that the laws of physics will be most elegant if formulated in this way. Sometimes however, it is easier to do calculations in some chosen inertial frame using frame dependent quantities.

Light cone

All the possible light worldlines leaving some event $\mathcal{O}$ together with all the possible light worldlines arriving at the event give a frame independent partition of spacetime known as the *lightcone*, see Fig. 1. The speed

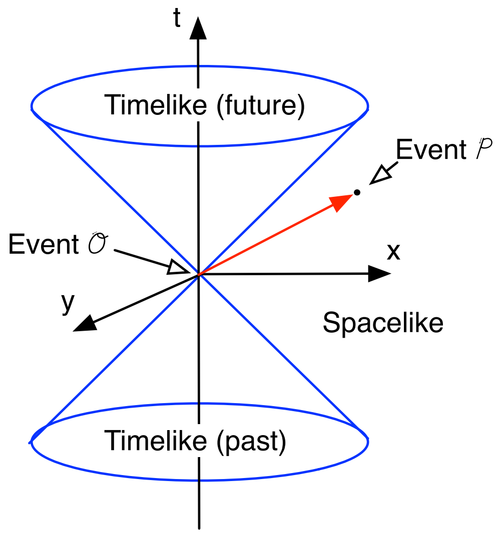

Figure 1: Light cone partitioning of space time for the event $\mathcal{O}$. Only two of the three space dimensions are drawn.

of light provides a maximum speed at which information or conventional particles can travel[1]. To see this consider the following.

> Two events $\mathcal{O}$, $\mathcal{P}$ are separated by Δx, $\Delta t > 0$ in some frame of reference S, with $\Delta x/\Delta t > 1$. If information could travel faster than light (speed > 1) the event $\mathcal{O}$ could causally influence the event $\mathcal{P}$. Now consider some other frame S' for which the time interval between the events is
>
> $$\Delta t' = \gamma(\Delta t - v\Delta x) = \gamma\,\Delta t\left(1 - v\frac{\Delta x}{\Delta t}\right). \tag{1}$$
>
> Since $\Delta x/\Delta t > 1$, we can find some frame S' given by a physical $v < 1$ for which $\Delta t' < 0$, i.e. the order of events is reversed. This is inconsistent with our basic idea of causality.

With the speed of light the maximum speed at which information can propagate, only events inside the light cone on the positive time side can be influenced by the event $\mathcal{O}$, and only events within the light cone on the negative time side can influence the event $\mathcal{O}$: these events are said to have a *timelike* relationship to $\mathcal{O}$. Events outside the light cone, such as the event $\mathcal{P}$ in the figure, are causally disconnected from the event $\mathcal{O}$ – they can neither influence or be influenced by $\mathcal{O}$. These events are said to have a *spacelike* relationship with $\mathcal{O}$.

Interval

Just as the length of a regular 3-vector is independent of rotations, there is a quantitative measure of the size of a 4-vector $\mathbf{\Delta x}$ that is the same in all inertial frames. This is the *interval* with the symbol Δs^2 and we think of it as the "length squared" or "magnitude squared" of the 4-vector $\Delta s^2 \equiv \mathbf{\Delta x}^2$, even though it is not always positive[2].

---

[1] See Assignment 2 and search the internet on *tachyons* for hypothetical exotic particles that always travel at speeds faster than light. In 2011 there were reports from CERN of neutrinos traveling faster than the speed of light, but these have now been disproved.

[2] Think of Δs^2 as a single symbol; we do not consider Δs.

The interval can be defined in a frame independent way as follows. Consider the 4-vector $\mathbf{\Delta x}$ joining the events $\mathcal{O}$, $\mathcal{P}$.

- For a timelike 4-vector (i.e. event $\mathcal{P}$ within the lightcone from $\mathcal{O}$) it is possible to find an inertial clock that is present at both events. The interval is defined from the time interval $\Delta\tau$ measured by this clock
$$\Delta s^2 = \Delta\tau^2. \tag{2}$$
- For a spacelike 4-vector (i.e. event $\mathcal{P}$ outside the lightcone from $\mathcal{O}$) it is possible to find a stationary ruler joining the two events (i.e. there is an inertial frame in which the two events are simultaneous, and their separation can be measured). The interval is defined from the spatial separation Δl measured by this ruler[3]
$$\Delta s^2 = -\Delta l^2. \tag{3}$$
- For a lightlike 4-vector (event $\mathcal{P}$ on the lightcone of $\mathcal{O}$) the interval is zero.

Using a Lorentz transformation from the special frames of reference of the inertial clock or ruler used to define the interval, it is easy to show that in an inertial frame in which the 4-vector $\mathbf{\Delta x}$ has components $(\Delta t, \Delta\vec{x})$ the interval is
$$\Delta s^2 = \Delta t^2 - \Delta\vec{x}^2 = \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2\ . \tag{4}$$
You can check that the Lorentz transformation $(\Delta t, \Delta\vec{x}) \to (\Delta t', \Delta\vec{x}')$ gives the same expression for the interval in terms of the primed coordinates, so that the interval is indeed a geometric, frame independent quantity.

Having defined the length of a 4-vector we can define the scalar product of two 4-vectors
$$\mathbf{\Delta x}_1 \cdot \mathbf{\Delta x}_2 = \tfrac{1}{2}\left[(\mathbf{\Delta x}_1 + \mathbf{\Delta x}_2)^2 - \mathbf{\Delta x}_1^2 - \mathbf{\Delta x}_2^2\right] \tag{5}$$
where $(\mathbf{\Delta x}_1 + \mathbf{\Delta x}_2)^2$ means the interval corresponding to the sum 4-vector, etc. From the construction the scalar product is clearly a scalar — the same in all inertial frames. Evaluating this expression in terms of the components in some inertial frame S gives
$$\mathbf{\Delta x}_1 \cdot \mathbf{\Delta x}_2 = \Delta t_1 \Delta t_2 - \Delta\vec{x}_1 \cdot \Delta\vec{x}_2 = \Delta t_1 \Delta t_2 - \Delta x_1 \Delta x_2 - \Delta y_1 \Delta y_2 - \Delta z_1 \Delta z_2. \tag{6}$$
Performing a Lorentz transformation on all the components to an S' frame would give the same expression in terms of the primed components
$$\mathbf{\Delta x}_1 \cdot \mathbf{\Delta x}_2 = \Delta t_1' \Delta t_2' - \Delta\vec{x}_1' \cdot \Delta\vec{x}_2' = \Delta t_1' \Delta t_2' - \Delta x_1' \Delta x_2' - \Delta y_1' \Delta y_2' - \Delta z_1' \Delta z_2', \tag{7}$$
which is what we mean by the scalar product being an invariant.

Constructing 4-Vectors

Other 4-vectors can be constructed in a number of ways:

1. Multiply a known 4-vector by a scalar (i.e. an invariant quantity, the same in all inertial frames);
2. If we know that $\mathbf{B}$ is a 4-vector and an object $\mathbf{A}$ with four components gives a scalar product $\mathbf{A} \cdot \mathbf{B}$ that is a scalar (Lorentz invariant) for every value of $\mathbf{B}$, then $\mathbf{A}$ is a 4-vector (see Taylor, §15.11 for the proof);

[3] Sign conventions differ between different texts. Some texts define the interval to be always positive, and do not introduce the minus sign in the definition of spacelike intervals.

3. A quantity with four components that are related in different inertial frames by a Lorentz transformation.

We can formulate relativistic laws of physics by forming new 4-vectors and hypothesizing appropriate laws that must then be tested against known limits (e.g. Newtonian physics for low velocities) and experiment. In this class we are particular interested in momentum, energy, and relativistic mechanics.

I will denote the components of a general 4-vector **A** in some inertial frame as (A^0, A^1, A^2, A^3) with A^0 the timelike component and A^1, A^2, A^3 the spacelike components forming a 3-vector $\vec{A}$ (i.e. they transform under rotations as a 3-vector does).

4-velocity: Defined by the derivative along the worldline

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau} \tag{8}$$

This is the 4-vector $d\mathbf{x}$ multiplied by the scalar $d\tau^{-1}$, and so is a 4-vector.

The magnitude-squared $\mathbf{u}^2$ (corresponding to the interval for the displacement 4-vector) is invariant. It can be calculated in any inertial frame. Choosing the rest frame of the particle $d\mathbf{x} \Rightarrow (d\tau, 0, 0, 0)$ so that $\mathbf{u}^2 = 1$.

In a frame of reference in which the 3-velocity is $\vec{u}$ we have[4]

$$u^0 = \frac{dt}{d\tau} = \gamma_u \qquad \text{with } \gamma_u = \frac{1}{\sqrt{1-u^2}}, \tag{9}$$

$$u^1 = \frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = \gamma_u u_x \tag{10}$$

$$\vdots \tag{11}$$

(remember $c = 1$), so that $\mathbf{u} \Rightarrow (\gamma_u, \gamma_u \vec{u})$.

4-momentum: This is simply mass $\times$ velocity

$$\mathbf{p} = m\mathbf{u} \tag{12}$$

with m a scalar — the value of the mass you would look up in tables.[5] In some inertial frame $\mathbf{p} \Rightarrow (m\gamma_u, m\gamma_u\vec{u})$ with $\vec{u}$ the 3-velocity of the particle.

We identify the *relativistic 3-momentum* as

$$\vec{p} = m\gamma_u\vec{u} = \frac{m\vec{u}}{\sqrt{1-u^2}}. \tag{13}$$

This of course is γ_u times the Newtonian expression.

The time component is $p^0 = m\gamma_u$. Expanding for small $\vec{u}$

$$p^0 = m\gamma_u \to m + \tfrac{1}{2}mu^2 + \cdots . \tag{14}$$

The Newtonian kinetic energy appears in the small u expansion, and we tentatively identify p^0 as the total energy E of the particle, with m (mc^2 in conventional units) the energy equivalent of the mass. This identification is amply confirmed by experiment! Thus the momentum 4-vector has components

[4] I will use superscripts 0, 1, 2, 3 for components of a 4-vector and subscripts x, y, z for components of a 3-vector.

[5] You might have seen this called the *rest mass* and then the momentum was written $\mathbf{p} = (m_u, m_u\vec{u})$ with $m_u = \gamma_u m$ a *relativistic mass*, but this usage has gone out of fashion.

$(E, \vec{p})$. We therefore now know how energy and momentum in different inertial frames are related: they transform in the same way as $(\Delta t, \Delta\vec{x})$. For our standard configuration of frames S, S'

$$p'_x = \gamma(p_x - vE) \tag{15}$$
$$E' = \gamma(E - vp_x) \tag{16}$$

where $\gamma = (1 - v^2)^{-1/2}$ with v the relative speed of the frames. (For conventional units put factors of c in to make dimensionally homogeneous, e.g. $E \to E/c$.)

4-acceleration: The 4-acceleration is defined as

$$\mathbf{a} = \frac{d\mathbf{u}}{d\tau}. \tag{17}$$

The components of the 4-acceleration in the frame S in which the particle velocity is $\vec{u}$ are complicated

$$a^\mu = (\gamma^4\vec{u}\cdot\vec{a}, \gamma^2\vec{a} + \gamma^4(\vec{u}\cdot\vec{a})\vec{u}). \tag{18}$$

We will study the relationship of this to 4-forces and 3-forces in Lecture 3.

Gradient 4-vector: The result here is a little unexpected, so let's work through a direct derivation in terms of Lorentz transformations. From the chain rule, we have

$$\frac{\partial}{\partial x'} = \left(\frac{\partial x}{\partial x'}\right)\frac{\partial}{\partial x} + \left(\frac{\partial y}{\partial x'}\right)\frac{\partial}{\partial y} + \left(\frac{\partial z}{\partial x'}\right)\frac{\partial}{\partial z} + \left(\frac{\partial t}{\partial x'}\right)\frac{\partial}{\partial t}. \tag{19}$$

The partial derivatives are evaluated from the $x', y, z', t' \to x, y, z, t$ Lorentz transformation $x = \gamma(x' + vt')$, etc., giving

$$\frac{\partial}{\partial x'} = \gamma\frac{\partial}{\partial x} + \gamma v\frac{\partial}{\partial t}, \tag{20}$$
$$\frac{\partial}{\partial t'} = \gamma\frac{\partial}{\partial t} + \gamma v\frac{\partial}{\partial x}, \tag{21}$$
$$\frac{\partial}{\partial y'} = \frac{\partial}{\partial y}, \tag{22}$$
$$\frac{\partial}{\partial z'} = \frac{\partial}{\partial z}. \tag{23}$$

Thus the quantity defined by components $(\frac{\partial}{\partial t}, -\vec{\nabla})$ is a 4-vector. I don't have a good symbol for this and will usually write it in component form ∂^μ. Note the minus sign appearing in front of the spatial gradient. The reason for writing the index μ in the "up" position and the minus sign will appear more natural once we consider co- and contra- variant components in Lecture 4.

D'Alembertian: The quantity

$$\Box^2 = (\frac{\partial}{\partial t}, -\vec{\nabla})\cdot(\frac{\partial}{\partial t}, -\vec{\nabla}) = \frac{\partial^2}{\partial t^2} - \nabla^2 \tag{24}$$

is a scalar product of two 4-vectors and is therefore a scalar (invariant). It is known as the d'Alembertian and appears in the wave equation for electromagnetic waves.

Wave 4-vector: The solutions to the wave equation $\Box^2\Phi = 0$ are of the form $\Phi \propto \cos(\omega t - \vec{k}\cdot\vec{x})$ with the phase speed $\omega/k = 1$ (i.e. the speed of light). The argument of the cosine determines the space-time points at which the wave is a maximum or minimum, and is a frame invariant quantity. Hence $\mathbf{k}$ with components in some frame $(\omega, \vec{k})$ is a 4-vector: then the argument of the cosine is $\mathbf{k}\cdot\mathbf{x}$, a scalar. This can be used to calculate the Doppler effect.

Charge-current 4-vector: Charge in some region can only change if there is a flow of charge (electric current) into that region. This can be expressed as the differential equation for *charge conservation*

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \tag{25}$$

with ρ the charge density and $\vec{j}$ the current density. This can be written

$$(\frac{\partial}{\partial t}, -\vec{\nabla}) \cdot (\rho, \vec{j}) = 0 \qquad \text{an invariant} \tag{26}$$

so that $(\rho, \vec{j})$ form a 4-vector.

Electromagnetic potential 4-vector: The electric potential ϕ and magnetic (vector) potential $\vec{A}$ form a 4-vector[6] **A** with components $(\phi, \vec{A})$.

Electromagnetic field tensor: You shouldn't think that you can take any 3-vector in physics and just find the right "time component" to give a 4-vector. This does not work, for example, for the electric field $\vec{E}$ and magnetic field $\vec{B}$. We know this since in a inertial frame in which there are only stationary charge sources there is only an electric field, whereas in any other inertial frame the charges will be moving and so there will also be a magnetic field. Thus $\vec{E}$, $\vec{B}$ are mixed by Lorentz transformation. In fact $\vec{E}$, $\vec{B}$ in some inertial frame are given by the 6 components in that frame of the antisymmetric second rank *electromagnetic field tensor* **F**. Second rank means that it has two component indices $F^{\alpha\beta}$ with $\alpha, \beta = 0, 1, 2, 3$. Then in some inertial frame

$$F^{\alpha\beta} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix}. \tag{27}$$

I'll discuss 4-tensors in more generality in Lecture 4. The electromagnetic field tensor can be derived from the potential 4-vector

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha. \tag{28}$$

If you know the expressions for $\vec{E}$, $\vec{B}$ in terms of ϕ, $\vec{A}$ (you will probably study/review this later in Ph106c) you can check this result.

Michael Cross, January 8, 2014

[6] Providing a relativistically covariant choice of "gauge" is used — see Ph106c.