GR lecture 2-2

More on rotations; Special Relativity

I. MORE ON ROTATIONS

There is a bit more to say about rotations of orthonormal bases; in particular, there are some insights to be gained from using matrix notation. A rotation matrix acts on vectors as:

$$x_i \;\to\; R_{ij}x_j \quad \Longleftrightarrow \quad \mathbf{x} \;\to\; R\mathbf{x} \; . \tag{1}$$

Acting on a matrix, a rotation does the same to each of the matrix's indices, i.e.:

$$M_{ij} \;\to\; R_{ik}R_{jl}M_{kl} \quad \Longleftrightarrow \quad M \;\to RMR^T \; . \tag{2}$$

The defining property of rotation (and reflection) matrices is that they preserve the identity matrix δ_{ij}, which defines the scalar product in an orthonormal basis:

$$R_{ik}R_{jk} = R_{ik}R_{jl}\delta_{kl} = \delta_{ij} \quad \Longleftrightarrow \quad RR^T = R\,1R^T = 1 \; . \tag{3}$$

Now let's consider a uniform rotation R_{ij} in some plane, with a constant angular velocity. The infinitesimal rotation over a time interval dt is given by a matrix of the form $\delta_{ij} + \omega_{ij}dt$, where ω_{ij} should clearly have the interpretation of angular velocity. In matrix notation, this can be written as $1 + \omega dt = e^{\omega dt}$. If we continue the same rotation over a finite time interval t, then we must multiply all the infinitesimal $e^{\omega dt}$ rotations, to obtain a finite rotation of the form:

$$R = e^{\omega t} \; ; \quad \dot{R} = \omega R = R\omega \; . \tag{4}$$

In terms of the matrix ω, the orthogonality condition on R becomes:

$$R^T = R^{-1} \iff e^{\omega^T t} = e^{-\omega t} \iff \omega^T = -\omega \; , \tag{5}$$

i.e. we've rediscovered the fact that the angular velocity $\omega_{ij} = \epsilon_{ijk}\omega_k$ is an antisymmetric matrix! In Lie group language, the rotation matrices R form the Lie group $SO(3)$, while the angular velocities ω form the Lie algebra $so(3)$, and are called the generators of the group

$SO(3)$. What we've discovered is that orthogonal group elements arise from exponentiating antisymmetric generators.

We might also consider a rotation process with non-constant angular velocity. In particular, the angular velocity at different times may be along different planes, which means that the rotations at different times do not commute with each other. Then we have two candidate definitions of the angular velocity matrix ω at time t, both antisymmetric: $\omega = R^T\dot{R}$ vs. $\omega = \dot{R}R^T$. However, these two definitions are related via $\omega \to R\omega R^T$, which means that they are actually the same angular velocity, but measured in two different orthonormal bases: before and after the rotation R.

Now we are ready to properly address Exercise 5 from Lecture 1-2. Consider an arbitrary change of reference frame in Newtonian physics:

$$\mathbf{x} \;\to\; \mathbf{x}' = R(t)\mathbf{x} + \mathbf{u}(t) \;, \tag{6}$$

where $R(t)$ is an orthogonal rotation matrix, and $\mathbf{u}(t)$ is a displacement of the origin. Consider a particle with a particular trajectory $\mathbf{x}(t)$. Our goal was to find its acceleration $\ddot{\mathbf{x}}'$ in the new frame, and identify the various inertial forces. To begin, we differentiate twice with respect to time and find:

$$\ddot{\mathbf{x}}' = \ddot{R}\mathbf{x} + 2\dot{R}\dot{\mathbf{x}} + R(t)\ddot{\mathbf{x}} + \ddot{\mathbf{u}} \;. \tag{7}$$

The 4th term is the uniform acceleration of the frame's origin. The 3rd term is the acceleration of the particle in the original frame, simply rotated into the new, rotated, basis. The other terms can be brought into the same form, using:

$$\dot{R} = R\omega \;; \quad \ddot{R} = \dot{R}\omega + R\dot{\omega} = R(\omega^2 + \dot{\omega}) \;. \tag{8}$$

We then get:

$$\ddot{\mathbf{x}}' = R(\omega^2\mathbf{x} + \dot{\omega}\mathbf{x} + 2\omega\dot{\mathbf{x}} + \ddot{\mathbf{x}}) + \ddot{\mathbf{u}} \;. \tag{9}$$

The first term corresponds to the centrifugal force, the second term to the angular acceleration of the frame's rotation, and the third term – to the Coriolis force.

II. THE LORENTZ TRANSFORMATIONS OF SPECIAL RELATIVITY

In the last lecture, we discussed the Galilean boost symmetry:

$$t \;\to\; t \;; \quad \mathbf{x} \;\to\; \mathbf{x} - \mathbf{v}t \;, \tag{10}$$

which is not respected by electromagnetism. One fine day, Lorentz noticed that there is a deformed version of the boost transformations (10) under which Maxwell's equations and the electromagnetic force law are invariant:

$$t \;\to\; \frac{t - \mathbf{v}\cdot\mathbf{x}/c^2}{\sqrt{1-v^2/c^2}} \;; \quad \mathbf{x} \;\to\; \mathbf{x} - \frac{\mathbf{v}}{\sqrt{1-v^2/c^2}}\left(t + \frac{\mathbf{x}\cdot\mathbf{v}}{v^2}\left(\sqrt{1-v^2/c^2}-1\right)\right) \; . \tag{11}$$

Slightly less horrifyingly, if we set $\mathbf{v}$ along the x direction, this becomes:

$$t \;\to\; \frac{t - vx/c^2}{\sqrt{1-v^2/c^2}} \;; \quad x \;\to\; \frac{x - vt}{\sqrt{1-v^2/c^2}} \;; \quad y \;\to\; y \;; \quad z \;\to\; z \; . \tag{12}$$

For Lorentz, this discovery was just a mathematical curiosity. Einstein's greatness was to say:

- Electromagnetism is more like gravity than like hydrodynamics: it's a fundamental law of Nature.
- The laws of Nature do have a boost symmetry, and it's important. However, the correct boost symmetry is given by (12), and (10) is just an approximation.

Before electromagnetism, we simply never had to notice corrections of the order v^2/c^2. Magnetism brought them to our attention by historical accident, because the tiny v^2/c^2 was being multiplied by the enormous magnitude of the electric force (which is usually almost perfectly canceled out, since matter is neutral overall).

So, what are these ugly formulas (12)? They're a rotation in the (t, x) plane in spacetime. Indeed, we can bring an ordinary rotation to the same kind of ugly form. Consider the rotation:

$$x \;\to\; x\cos\theta - y\sin\theta \;; \quad y \;\to\; y\cos\theta + x\sin\theta \; . \tag{13}$$

Now, instead of the angle θ, let's use the slope $v = \tan\theta$ (remember: velocity is slope in spacetime!). Then our rotation becomes:

$$x \;\to\; \frac{x - vy}{\sqrt{1+v^2}} \;; \quad y \;\to\; \frac{y + vx}{\sqrt{1+v^2}} \; . \tag{14}$$

Finally, suppose that x measures horizontal distance along the surface of the Earth, and y measures vertical distance. Normally, humans and their transport devices travel horizontally much further than vertically. So it's often useful to use larger units for the horizontal axis.

E.g. on a plane, the height is often given in feet, and the horizontal distance – in miles. Then the slope v is measured in feet/miles, and we need to insert a unit conversion constant $c = 5,280$feet/miles. Our rotation formula finally becomes:

$$x \;\rightarrow\; \frac{x - vy/c^2}{\sqrt{1 + v^2/c^2}} \;; \quad y \;\rightarrow\; \frac{y + vx}{\sqrt{1 + v^2/c^2}} \;. \tag{15}$$

So now we understand how to clean up the Lorentz boost formula (12). First, we need to stop using different units for space and time, and realize that the "speed of light" $c \approx 3 \cdot 10^8$m/s is just a conversion constant: there are about $3 \cdot 10^8$ meters in a second, and that's that. So we redefine our units such that $c = 1$, and forget about it. Then the boost formula simplifies to:

$$t \;\rightarrow\; \frac{t - vx}{\sqrt{1 - v^2}} \;; \quad x \;\rightarrow\; \frac{x - vt}{\sqrt{1 - v^2}} \;, \tag{16}$$

where we forgot about the (y, z) axes for simplicity. As the next step, we may want to identify v as $\tan\theta$ of some spacetime angle. However, the signs in (16) are not the same as in (14), so that doesn't work. To take care of the signs, we pass from ordinary trigonometry into hyperbolic trigonometry, and define $v = \tanh\theta$. The boost formula then simplifies further:

$$t \;\rightarrow\; t\cosh\theta - x\sinh\theta \;; \quad x \;\rightarrow\; x\cosh\theta - t\sinh\theta \;, \tag{17}$$

and we're beginning to glimpse the basic idea of spacetime geometry. Note that $|v|$ is restricted in the range $(0, c = 1)$, which corresponds to the full range $(0, \infty)$ of the so-called "rapidity" angle θ.

Setting $c = 1$ means that various quantities now have the same units. In particular, in SR units, we have:

$$\begin{gathered} [\text{length}] = [\text{time}] \;; \quad [\text{velocity}] = 1 \;; \quad [\text{acceleration}] = 1/[\text{length}] \;; \\ [\text{mass}] = [\text{momentum}] = [\text{energy}] \;; \quad \frac{[\text{charge}]}{[\text{volume}]} = \frac{[\text{current}]}{[\text{area}]} \;. \end{gathered} \tag{18}$$

III. TIME DILATION, LENGTH CONTRACTION, RELATIVITY OF SIMULTANEITY, INVARIANCE OF THE SPEED OF LIGHT

Let's quickly work out the two famous effects of a Lorentz boost. First, let's focus on two of my birthdays in my rest frame: $(t_1, x_1) = (0, 0)$ and $(t_2, x_2) = (\tau, 0)$. In other words, we consider a process that happens to me while I'm at rest in the origin, and takes time τ in my reference frame. Now let's pass to a frame moving at velocity $-v$ (so that I'm now moving at velocity v).

Exercise 1. *Find the time interval $t'_2 - t'_1$ in the new reference frame. Is it longer or shorter than τ?*

The frame shift also affects the x coordinates – in the new frame, the two events no longer happen in the same place:

$$x'_1 = 0 \; ; \quad x'_2 = \frac{v\tau}{\sqrt{1-v^2}} \neq x'_1 \; . \tag{19}$$

Apart from the $1/\sqrt{1-v^2}$, this phenomenon is familiar from Galilean physics: for two events at different times, whether or not they happened at the same place depends on the reference frame.

Now, consider the opposite situation: in some reference frame, two events happen at the same time, but in different places: $(t_1, x_1) = (0, 0)$ and $(t_2, x_2) = (0, L)$. Now we pass into a new frame, moving at velocity v. In Galilean physics, time never transforms, so the events will be simultaneous in the new frame as well. However, in SR, space and time are placed on an equal footing, so something like the effect (19) should apply again: for two events at different places, whether or not they happened at the same time should depend on the reference frame.

Exercise 2. *Work this out explicitly. What is the time shift $t'_2 - t'_1$ between the events $(t_1, x_1) = (0, 0)$ and $(t_2, x_2) = (0, L)$ in the new frame?*

Exercise 3. *What is the spatial distance between the events in the new frame? Is it longer or shorter than L?*

From the last exercise, it may be unclear why people talk about relativistic length contraction. That actually refers to a slightly more involved scenario:

Exercise 4. *Consider a stick of length L in its rest frame, such that its endpoints stay at $x_1 = 0$ and $x_2 = L$ at all times. Now pass into a reference frame moving at velocity $-v$, so that the stick now moves at velocity v. What's the length of the stick in the new frame, defined as the distance $x'_2 - x'_1$ between its endpoints at a fixed time t'?*

The above effects can all be understood intuitively by imagining Lorentz boosts as rotations in spacetime, which tilt the various spacetime vectors. However, the sign of the effect may come out opposite from our naive Euclidean intuition. For example, if we draw x and t axes on Euclidean paper, then the Lorentz transformation (17) doesn't look like an actual rotation; instead, it tilts both axes either towards or away from each other, while preserving the midline at 45^o.

Exercise 5. *Let's understand the last statement physically. Consider a photon with trajectory $x = ct$, i.e. simply $x = t$. Find the trajectory $x'(t')$ in a boosted frame with velocity v along the x direction. What is the photon's velocity in the new frame?*

Exercise 6. *Now let's find the transformation rule for general velocities, not necessarily along the x axis. Consider a particle with trajectory $\mathbf{x} = \mathbf{u}t$. Find the paritcle's new velocity $\mathbf{u}'$ in a boosted frame with velocity $\mathbf{v}$. Check that the speed of light is still invariant, regardless of the boost's direction.*

IV. CAUSALITY

The relativity of simultaneity, which we've seen in Exercise 2, raises a pertinent question about causality. Two events that were simultaneous in one frame may occur at different times in another, and their time ordering will depend on the chosen frame. Thus, if we insist on consistent ordering of causes and effects, then we must never allow instantaneous influence across a spatial distance: that would allow us to flip the order of cause and effect by changing the reference frame! This why SR tends to guide us towards field theory, where interactions are always local in both space and time.

Of course, even if the microscopic laws are local, their effects will end up propagating far away from the original cause. We should therefore understand when is it possible to speak of a consistent causal relationship between events in spacetime. Consider two events $(0, 0)$ and $(t, \mathbf{x})$. A key property of Lorentz boosts is that they can change the events' separation in space only within certain ranges, never crossing from one range to the other. These different types of separation are:

- $|t| > |\mathbf{x}|$, $t > 0$: timelike-separated, second event is in the future.
- $|t| > |\mathbf{x}|$, $t < 0$: timelike-separated, second event is in the past.

- $|t| = |\mathbf{x}|$, $t > 0$: lightlike-separated, second event is in the future.
- $|t| = |\mathbf{x}|$, $t < 0$: lightlike-separated, second event is in the past.
- $|t| < |\mathbf{x}|$: spacelike-separated (the time ordering depends on the reference frame).

Proving this assertion is not too hard, but it will become obvious once we introduce 4-vector language below. The point for now is that causal relationship is non-ambiguous between timelike- or lightlike-separated events. In other words, causal influence can only travel at, or below, the speed of light.

V. 4-VECTOR NOTATION, THE MINKOWSKI METRIC AND THE LORENTZ GROUP

Enough of this non-relativistic notation! Time to grow up, as well as to start putting index notation to good use. From now on, we unify t and $\mathbf{x}$ into a 4-vector $x^\mu = (t, x, y, z)$, where the index μ runs over (t, x, y, z), or $(0, 1, 2, 3)$. The crucial property of Lorentz boosts is that they preserve the spacetime interval:

$$x_\mu x^\mu = \eta_{\mu\nu} x^\mu x^\nu = -t^2 + x^2 + y^2 + z^2 \ , \tag{20}$$

where $\eta_{\mu\nu}$ is the metric of flat spacetime, known as the Minkowski metric:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ . \tag{21}$$

Exercise 7. *Write down the inverse metric $\eta^{\mu\nu}$.*

Exercise 8. *Prove that Lorentz boosts indeed preserve $x_\mu x^\mu$.*

More precisely, the full group of "spacetime rotations" that preserve the "squared distance from the origin" $x_\mu x^\mu$ consists of:

- Ordinary spatial rotations in the xy, yz and zx planes.
- Lorentz boosts in the x,y and z directions. We think of these as rotations in the tx, ty and tz planes.

Together, these symmetries are called the Lorentz group. Both the rotations and boosts take the general linear form $x^\mu \to \Lambda^\mu{}_\nu x^\nu$. If we combine the Lorentz group with translations of the origin, we obtain the Poincare group $x^\mu \to \Lambda^\mu{}_\nu x^\nu + w^\mu$, which preserves the spacetime distance $\Delta x_\mu \Delta x^\mu$ between pairs of points.

The minus sign in the metric (21) is the one and only distinction between space and time in SR. It cannot be removed by any change of basis, unless we allow complex numbers. Thus, the (t, x, y, z) basis is "as orthonormal as it gets", and we're forced to pay attention to upper vs. lower indices: $x^\mu = (t, x, y, z)$ vs. $x_\mu = (-t, x, y, z)$.

The distance-squared $\Delta x_\mu \Delta x^\mu$ can be positive, negative or zero. This corresponds precisely to the notions of spacelike, timelike and lightlike (also known as null) separation from the previous section. The causality statements from the previous section follow from the invariance of $\Delta x_\mu \Delta x^\mu$, together with the fact that one cannot rotate a future-pointing timelike or null vector into a past-pointing one without going through the spacelike region in between. Due to the different possible signs, we sometimes define the actual distance between events as $s = \sqrt{\Delta x_\mu \Delta x^\mu}$ (if we expect it to be spacelike), and sometimes as $\tau = \sqrt{-\Delta x_\mu \Delta x^\mu}$ (if we expect it to be timelike). Since τ is the same as time in the rest frame $\mathbf{x} = \text{const}$, it is often called "proper time". The notation s is often used for timelike separations as well. Also, some people prefer to define the entire metric with opposite signs, i.e. $t^2 - x^2 - y^2 - z^2$, so be careful.

In Euclidean space, we define a sphere as the surface of constant distance $\sqrt{\mathbf{x} \cdot \mathbf{x}}$ from the origin. In Minkowski spacetime, we can similarly consider "spheres" at constant $x_\mu x^\mu$, but their topology is now different. When drawn on Euclidean paper, they look like hyperboloids. A surface of constant positive $x_\mu x^\mu$ is a single-sheeted hyperboloid, while the surface of negative constant $x_\mu x^\mu$ is a two-sheeted hyperboloid. The limiting case, $x_\mu x^\mu = 0$, is a cone – the lightcone. It demarcates the boundary between spacelike-separated and timelike-separated regions.

VI. 4-VELOCITY, 4-ACCELERATION AND 4-MOMENTUM

Many familiar quantities arrange themselves into spacetime vectors and tensors. We refer to a quantity v^μ as a spacetime vector, or a 4-vector, if its components transform under Lorentz boosts like the components of $x^\mu = (t, \mathbf{v})$.

Let us start with velocity. The standard non-relativistic notion of velocity $d\mathbf{x}/dt$ is not a good starting point: the denominator is just one component of the 4-vector dx^μ. Instead, it's better to treat $\mathbf{v}$ as an approximation to the spatial components of the 4-vector:

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{\sqrt{-dx_\mu dx^\mu}} = \frac{(1, \mathbf{v})}{\sqrt{1-v^2}} \ , \tag{22}$$

which is called the 4-velocity. This is just the unit tangent vector to the (timelike!) trajectory of a massive particle through spacetime, which is also known as its worldline. Similarly, the appropriate generalization of acceleration given by the 4-vector:

$$\alpha^\mu = \frac{du^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} \ , \tag{23}$$

called the 4-acceleration.

Exercise 9. *Express the space and time components of α^μ in terms of the ordinary velocity $\mathbf{v}$ and the ordinary acceleration $\mathbf{a}$.*

The appropriate generalization of momentum $\mathbf{p} = m\mathbf{v}$ is the 4-momentum:

$$p^\mu = mu^\mu = \left(\frac{m}{\sqrt{1-v^2}}, \frac{m\mathbf{v}}{\sqrt{1-v^2}}\right) \ . \tag{24}$$

What's the meaning of the t component of this 4-vector? Let us Taylor-expand in small velocities $v \ll 1$:

$$p^t = m + \frac{1}{2}mv^2 + \dots \tag{25}$$

We recognize the second term as the non-relativistic kinetic energy. The first term is new: it is the rest energy $m = mc^2$ of a particle with mass m. Together, p^t is just the total energy of the particle! To see why this makes sense, consider the rule from analytical mechanics for the variation of the action under changing the location $(t, \mathbf{x})$ of the trajectory's final point:

$$dS = -E dt + \mathbf{p} \cdot \mathbf{dx} \ , \tag{26}$$

which relativistically gets rewritten simply as $dS = p_\mu dx^\mu$. Note that the mass m is just the (timelike) length of the 4-momentum:

$$m = \sqrt{-p_\mu p^\mu} = E^2 - \mathbf{p} \cdot \mathbf{p} \ . \tag{27}$$

When several particles interact, the total 4-momentum $P^\mu = \sum p^\mu$ is conserved. However, in SR, there is no conservation of the total mass $\sum m$. Mass can be created and destroyed, such as an electron-positron pair turning into two photons. In fact, the "total mass" of a given system is an ambiguous concept. Is it the sum $\sum m$ of individual particle masses, or the length $\sqrt{-P_\mu P^\mu}$ of the total 4-momentum? What if we zoom out of the composite system so that it appears as a point particle? For the 4-momentum itself, such problems do not arise: it is additive and conserved.

For massless particles $m = 0$, such as photons, a slightly different approach is needed, which we won't spell out in full here. While massive particles travel along timelike trajectories, massless particles travel along lightlike ones. A free massless particle travels along a lightlike geodesic, which we call simply a lightray. The proper time along a lightlike line vanishes, and thus the 4-velocity is ill-defined. Though the particle's worldline has a direction, there is no special unit vector $u^\mu = dx^\mu/d\tau$ along this direction. The particle's 4-momentum is then just an arbitrary lightlike vector p^μ tangent to the worldline, satisfying:

$$p_\mu p^\mu = E^2 - p^2 = 0 \ . \tag{28}$$

A pair of parallel lightlike 4-momenta can describe e.g. a pair of photons that are both flying at the speed of light in the same direction, but with different energies (i.e. different frequencies of light).