Introduction and the line element

These notes follow Jim Hartle's "Gravity: An Introduction to Einstein's General Relativity" somewhat closely. Some of the figures are borrowed directly from the textbook. For more advanced material, see Sean Carroll's "Spacetime and Geometry: An Introduction to General Relativity."

There are four fundamental forces in nature:

- Gravitational
- Electromagnetic
- Weak
- Strong

The first two are long-range and the last two are short-range.

Even though gravity was known as a force well before the other three (formalized by Newton), it is actually the least understood one (understanding a quantum theory of gravity occupies most of my and many other physicists' time). It is simlar to electromagnetism in that it is long range, with a similar-looking force law:

$$F = G\frac{m_1 m_2}{r^2} \quad \longleftrightarrow \quad F = k\frac{q_1 q_2}{r^2}\,. \tag{0.1}$$

A crucial difference, related to what makes a quantum theory of gravity so difficult, is that *everything* gravitates. Relatedly, positive masses attract (and mass can't be negative), whereas same-sign electric charges repel.

Newton's force law is good enough for everyday purposes, and even not-everyday purposes like launching rockets to the moon, but it's not precisely correct. If you used it to calculate distances for GPS, you would be off by a few miles.

Please review/use chapters 1-3 of Hartle, I will assume you know it except for line elements which we'll get to now.

Line elements and geometry

When we get to general relativity, we will see that geometry plays a starring role. So we should acquaint ourselves with it. Line elements will also be crucial in special relativity.

In ordinary geometry, the line element is familiar from the Pythagorean theorem:

$$\Delta s^2 = \Delta x^2 + \Delta y^2\,. \tag{0.2}$$

This is often useful to think of infinitesimally, i.e. for small Δs, Δx and Δy:

$$ds^2 = dx^2 + dy^2\,. \tag{0.3}$$

This may not look familiar, but you've used it to calculate the length of curves in the plane

$$\text{Length } = \int ds = \int \sqrt{dx^2 + dy^2} = \int dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \tag{0.4}$$

This just adds up a bunch of little ds lengths and accounts for curving. So this line element encodes geometry.

Example 1: A simple example is computing the radius of a circle defined as $x^2 + y^2 = R^2$:

$$\text{Radius } = \int_\gamma ds = \int \sqrt{dx^2 + dy^2} = \int_0^R dx = R \tag{0.5}$$

where we had the curve go straight along the x-axis from $x = 0$ to $x = R$, meaning $dy/dx = 0$. This may seem complicated but this kind of thing will become more important when we consider circles living on a sphere.

Example 2: We can also compute the circumference of the circle:

$$\text{Circumference } = \int_{x^2+y^2=R^2} ds = \int_{x^2+y^2=R^2} \sqrt{dx^2 + dy^2} = 2\int_{-R}^{R} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \tag{0.6}$$

In this step, we parameterized the curve as $y = \sqrt{R^2 - x^2}$, which as x goes from $-R$ to R, covers half of the circle. We put in the factor of 2 to account for the other half. Plugging in $y = \sqrt{R^2 - x^2}$ and doing the integral gives $2\pi R$ for the circumference.

As we'll discuss over the next 10 weeks, the line element encodes *all* aspects of the geometry. So we will be seeing it a LOT.

Let's consider the flat plane in polar coordinates:

$$ds^2 = dr^2 + r^2 d\theta^2 \tag{0.7}$$

This is the *same* line element as $ds^2 = dx^2 + dy^2$, it is just written in *different coordinates*. To see this, we can treat these differentials as derivatives and perform the coordinate transformation $x = r\cos\theta$, $y = r\sin\theta$ which relates the two coordinate systems. To do this, it will help to recall the chain rule for a coordinate transformation $x = f(r(t), \theta(t))$:

$$\frac{dx}{dt} = \frac{\partial x}{\partial r}\frac{dr}{dt} + \frac{\partial x}{\partial \theta}\frac{d\theta}{dt} \tag{0.8}$$

For a differential, we just multiply both sides by dt:

$$dx = \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta \tag{0.9}$$

You can begin doing what you were told never to do as you learned calculus ;). For a single variable coordinate change $u = f(x)$ you should be familiar with $du = f'(x)dx$ from u-substitutions in evaluating integrals.

Returning to the coordinate transformation $x = r\cos\theta$, $y = r\sin\theta$ and using the formula above gives

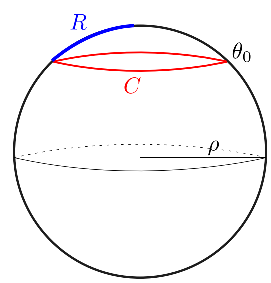
$$dx = \cos\theta dr - r\sin\theta d\theta\,, \qquad dy = \sin\theta dr + r\cos\theta d\theta\,. \tag{0.10}$$

Plugging into $ds^2 = dx^2 + dy^2$ gives

$$\begin{aligned} ds^2 &= \cos^2\theta dr^2 + r^2\sin^2\theta d\theta^2 - 2r\sin\theta\cos\theta dr d\theta + \sin^2\theta dr^2 + r^2\cos^2\theta d\theta^2 + 2r\sin\theta\cos\theta dr d\theta \\ &= dr^2 + r^2 d\theta^2\,. \end{aligned} \tag{0.11}$$

These coordinate changes simply use the chain rule of calculus, and we will be using them a lot.

Example 3: Let's consider a sphere of radius ρ. What is the ratio between the circumference C of a circle of constant latitude and its radius R?



To solve this, we will use three-dimensional spherical coordinates, which can be obtained from $ds^2 = dx^2 + dy^2 + dz^2$ through the coordinate transformations $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$:

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi \tag{0.12}$$

Since we are restricting to the surface of the sphere, $r = \rho$, we find

$$ds^2 = \rho^2 d\theta^2 + \rho^2\sin^2\theta\, d\phi \tag{0.13}$$

Now we compute the radius and circumference as before:

$$R = \text{Radius } = \int_{\phi=const.} \rho\sqrt{d\theta^2 + \sin^2\theta\, d\phi^2} = \int_0^{\theta_0} \rho\, d\theta = \rho\theta_0\,. \tag{0.14}$$

$$C = \text{Circumference } = \int_{\theta=\theta_0} \rho\sqrt{d\theta^2 + \sin^2\theta\, d\phi^2} = \int_0^{2\pi} \rho\sin\theta_0\, d\theta = 2\pi\rho\sin\theta_0\,. \tag{0.15}$$

So we find

$$\frac{\text{Circumference}}{\text{Radius}} = \frac{C}{R} = 2\pi\frac{\sin\theta_0}{\theta_0}\,. \tag{0.16}$$

For small θ_0, we do not notice the curvature of the sphere and should recover the flat-space answer, and indeed $\frac{C}{R} \to 2\pi$ as we expect. But in general the ratio $\frac{C}{R} < 2\pi$.

Special relativity

We will begin with the Michelson-Morley experiment and velocity addition. As you may know, the understanding of E&M through Maxwell's equations was one of the triumphs of physics in the 1800s. It made clear that light is an E&M wave, with speed $c = 299,792,458\ m/s$. It is similar to ordinary waves, like waves in the ocean, except the speed of those depends on the frame of reference: the velocity of the wave according to a surfer is zero! This is because velocities *add* in Newtonian physics.

Example 4: Say the velocity of the wave relative to the sea is $\vec{v}_{ws}$, and the velocity of a dolphin relative to the sea is $\vec{v}_{ds}$. Then the velocity of the wave relative to the dolphin is

$$\vec{v}_{wd} = \vec{v}_{ws} - \vec{v}_{ds} = \vec{v}_{ws} + \vec{v}_{sd} \tag{0.17}$$

where $\vec{v}_{sd}$ is the velocity of the sea as seen by the dolphin.

This is so intuitive that we don't usually stop to think about it. Let's derive where it comes from. Consider three points A, B, and C which move as a function of time. We have the (oriented) distances

$$\vec{x}_{CB}(t) = \vec{x}_C(t) - \vec{x}_B(t)\,, \qquad \vec{x}_{BA}(t) = \vec{x}_B(t) - \vec{x}_A(t)\,, \qquad \vec{x}_{CA}(t) = \vec{x}_C(t) - \vec{x}_A(t) \tag{0.18}$$

This means the velocity of C relative to A is

$$\vec{v}_{CA} = \frac{d\vec{x}_{CA}(t)}{dt} = \vec{v}_C - \vec{v}_A = (\vec{v}_C - \vec{v}_B) + (\vec{v}_B - \vec{v}_A) = \vec{v}_{CB} + \vec{v}_{BA}\,. \tag{0.19}$$

There seemed to be no assumptions made at all! There is sort of an implicit one, which is that the notions of space and time are the same for the three points.

This intuition was strong in the late 1800s; people were curious about the "sea" which light waves traveled in. This was called the "ether" and the **speed of light was interpreted to be constant relative to this ether**. But presumably the Earth moves with respect to this ether, so the speed of light as measured on Earth should be slightly different when light points in the direction of the Earth's motion vs. orthogonal to the Earth's motion. Michelson and Morley set out to find these differences and discover the ether: a beam splitter was used to send light in two orthogonal directions and measure their $x/t = v$ in each direction. They did this at what is now Case Western in Ohio, and they found **no difference** in the two directions.[1] Apparently the Earth was at rest with respect to the ether. Getting desperate, they tried again six months later when the Earth would be moving in the opposite direction. It was still at rest with respect to the ether!

Then people got really desperate, claiming the Earth dragged the ether with it, so went on to do the experiments on mountaintops where it would drag the ether less...in retrospect, this

[1] Notice that if light is going in the same direction as the Earth relative to the ether, then on one leg of the light's journey it gets an increased velocity $c + v$ and on the other leg it gets a decreased velocity $c - v$, but these don't cancel: the times along the two legs are $\Delta t_1 = \frac{2L}{c+v}$ and $\Delta t_2 = \frac{2L}{c-v}$, which don't sum to $2L/c$ which would be the time for both legs without any boost.

sounds ridiculous, but it is a persistent, incredibly scientific pursuit of all logical possibilities. These wonderful experiments, being continued today, give us confidence that there is no such ether.

Then how do we interpret the fact that we always measure the speed of light to be the same? The essence of Einstein's idea (apparently developed independently of knowledge of the Michelson and Morley experiments) is to abandon the idea of an ether and to claim that the speed of light is the same in any reference frame. This means that velocities do not simply add! But if velocities do not add then space and/or time must be different in different inertial frames of reference.

Simultaneity

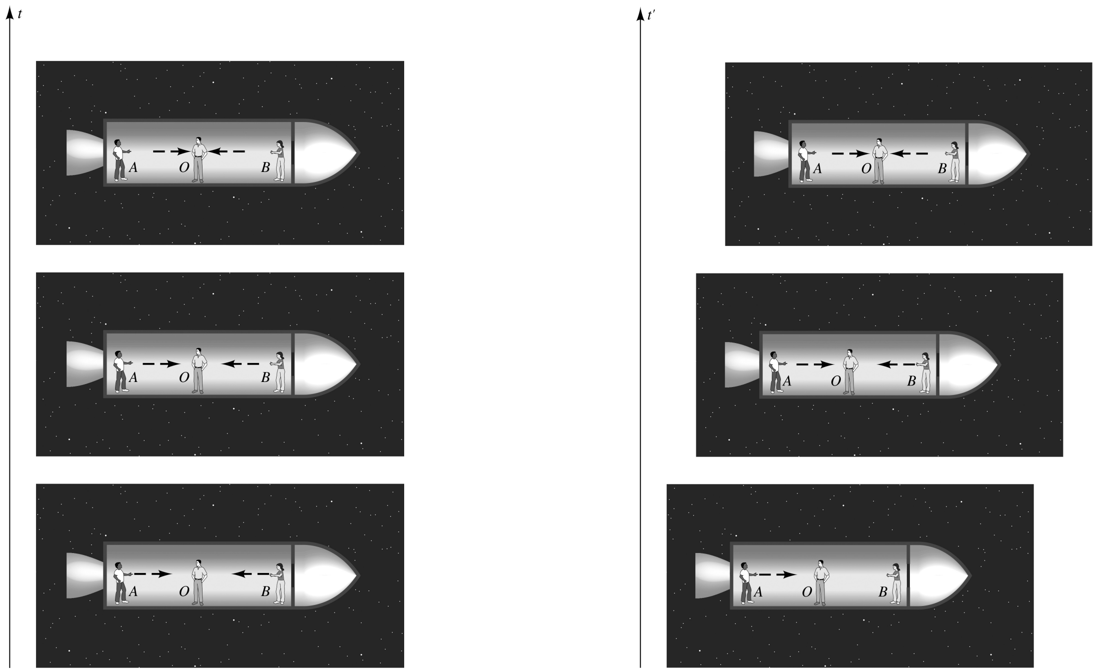
We will begin with the key assumption behind special relativity: the speed of light is the same in all inertial reference frames. Recall from Ch. 3 in Hartle that an inertial reference frame is defined by the existence of some time parameter t such that coordinates in that reference frame obey

$$\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0\,. \tag{0.20}$$

In particular, this implies that inertial reference frames move at constant velocities with respect to each other (and objects with no net forces acting on them move at constant velocities in any inertial reference frame). And this definitely means that velocities do not add, since the speed of light can't go faster or slower if you run next to it and look at it. [2]

The easiest way to see that space and time are different in different inertial frames is through the loss of simultaneity. Say we have observers A, O, and B. A is 10 feet to the left of O who is 10 feet to the left of B. Say they are on a rocket moving to the right at some very fast speed. In the frame of the rocket (i.e. we are inside of it), let A and B turn on flashlights at the same time toward observer O. Then O will receive the light at the same time from both sides. This is represented on the left in the figure below. But now analyze this in the frame of your lawn chair, watching the rocket whiz by. Since the speed of light is finite, the light that leaves A's flashlight has a *further* distance to travel than the light that leaves B's flashlight, since O is rushing toward B. If velocities added, then A's flashlight would be boosted by the rocket's speed, and B's would be slowed down, so they would still reach O at the same time. But velocities don't add, the speed of light is still c as you watch from your lawn chair! But it still must be objectively true that O receives the signals the same time, since you can just stop the rocket and ask O what happened afterward, and he'd better have just one answer.

[2]When he was a teenager (years before developing the special theory of relativity), Einstein imagined what it would be like to ride alongside a lightbeam at the speed of light. Using the Newtonian velocity addition formula he concluded that the beam of light would be stationary, and thought that this contradicted Maxwell's equations (which it does if you assume they always apply with the same numerical value c for the speed of light).



The only logical possibility left is that the light did not leave the flashlights at the same time. It had to have left A's flashlight at an earlier time (in the lawnchair's frame) than it left B's flashlight, so that the beams get to O at the same time. (Notice if you stop the rocket and chat with the observers there won't be any paradoxes the way there would be if O received the signals at different times in the lawnchair's frame.) This is represented on the right in the figure above. We see, as we predicted, that the notion of time is different for different frames.

Spacetime diagrams

Let's draw the previous example on a spacetime diagram, a very useful tool that shows the path of things through space and time. A point P on a spacetime diagram is called an "event" since it corresponds to a particular place in space *and* time. On the y-axis we plot ct and on the x-axis we plot x. So lightbeams traveling in the x direction are at 45 degree lines on the spacetime diagram, and objects at rest $\Delta x = 0$ are vertical lines.

In the rocket's frame, the lightbeams travel according to spacetime diagram on the left of figure 1, whereas in the lawnchair's frame the lightbeams travel according to right diagram of figure 1. In both cases O *receives* the signals at the same time, but only in the rocket's frame do they *leave* the flashlights at the same time.

The geometry of flat spacetime

Adapting the words of Archimedes, "Give me a lever long enough...and I shall move the world." The constancy of the speed of light is a long lever: from it comes out all sorts of

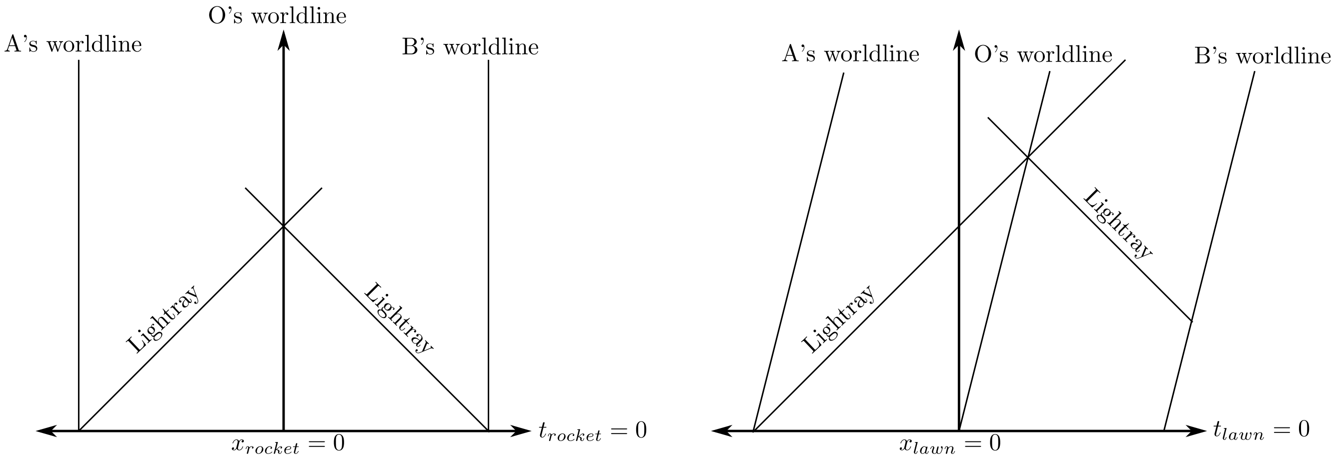


Figure 1: On the left we have the spacetime diagram in the reference frame of the rocket; the flashlights are turned on simultaneously $t_{rocket} = 0$, and they reach observer O at the same instant. On the right we have the same events now in the reference frame of us in our lawnchair watching the rocket whiz by. O still needs to observe the lights arriving at the same time, so it must be the case that they do not leave at the same lawn time t_{lawn}.

beautiful physics! We saw a clever thought experiment which exhibited the loss of the notion of simultaneity for spatially separated events. All the consequences of this new starting point can be worked out by stating the new line element which governs spacetime in special relativity. But let's motivate it by one more thought experiment.

Consider two mirrors separated by distance $\Delta z = L$ with light bouncing back and forth between them. This is a clock: the time is measured by the number of bounces, e.g. a return to the starting point gives $c\Delta t = 2L \implies \Delta t = 2L/c$. Let's say this clock is put onto a rocket which flies perpendicular to the z direction, say in the x direction, at velocity v. Then from your lawnchair's perspective, the light is not bouncing straight up and down – it is going both in the x and z directions. (See Figure 4.7 in Hartle.)

Let's analyze the two frames. In the rocket's frame we have

$$\Delta t = 2L/c\,, \qquad \Delta x = \Delta y = \Delta z = 0\,. \tag{0.21}$$

In the lawnchair's frame the light also travels in the x direction an amount $\Delta x' = v\Delta t'$. We use primes on space and time in the lawnchair's frame since the notions of space and time are different in different frames! The time it takes for the light to come back to its starting point is obtained by doubling the time it takes to go from one mirror to the other one:

$$c\Delta t' = 2\sqrt{L^2 + \left(\frac{\Delta x'}{2}\right)^2} \tag{0.22}$$

where the distance is just the hypotenuse of a triangle. So altogether, in the lawnchair's frame, we have

$$\Delta t' = \frac{2}{c}\sqrt{L^2 + \left(\frac{\Delta x'}{2}\right)^2}\,, \qquad \Delta x' = v\Delta t'\,, \qquad \Delta y' = \Delta z' = 0\,. \tag{0.23}$$

We immediately see the notion of time dilation:

$$\text{Time dilation: } \Delta t' > \Delta t\,. \tag{0.24}$$

More importantly we can construct the following quantity which is the **same** whether we compute them in the rocket frame or the lawnchair frame, i.e. whether we compute them with the unprimed distances/times or the primed distances/times

$$-(c\Delta t)^2 + (\Delta x)^2 = -4L^2 = -(c\Delta t')^2 + (\Delta x')^2 = -4L^2\,. \tag{0.25}$$

Generalizing this to include the y and z directions gives

$$-(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = (-c\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2\,, \tag{0.26}$$

where (t, x, y, z) and (t', x', y', z') represents the coordinates of two inertial reference frames. This is the fundamental **invariant** of special relativity, which we encode in the line element of flat spacetime as[3]

$$ds^2 = -(cdt)^2 + dx^2 + dy^2 + dz^2\,. \tag{0.27}$$

This is called Minkowski space or Minkowski geometry. This is to be contrasted with Euclidean geometry, which has the metric $ds^2 = dx^2 + dy^2 + dz^2$.

——— END LEC 1, ASKED TO DISPROVE PERP LENGTH CONTRACTION ———

Example 5: In the above argument we assumed the distance L is the same in the two reference frames. Here is a thought experiment which shows that has to be the case. Let's say L got shorter or taller. Now let's consider two trains A and B moving past one another, in opposite directions. From the perspective of train A, train B should be shorter. And from the perspective of train B, train A should be shorter. This by itself is no problem – as we know, things can look different in different frames of reference. But in the end when you stop all the motion and bring everyone together to compare the results of experiments, everything should be consistent!

To see that things won't be consistent in this case, I'll hang off the side of train A with a red paintbrush at some height above the ground, and I'll have one of you hang off the side of train B with a blue paintbrush at the same height above the ground. Without changing the height of the paintbrush, we'll each paint a horizontal stripe on our train and the other train as it whizzes by.

[3] In many references, especially ones related to particle physics, you will see the convention $ds^2 = (cdt)^2 - dx^2 - dy^2 - dz^2$. This has led to many arguments on Twitter and other places.

As you pass me by, you will look shorter, so my red paint will be above your blue paint on both trains. But from your perspective, I will look shorter, so your blue paint will be above my red paint on both trains. This is clearly a paradox because we can stop the trains and *observe where the paint is!* There can't be more than one answer to this question.

This device of putting markings on objects, or having objects blow up, is often a very useful pedagogical device in deciding what is paradoxical and can't be the case vs. what is just a very strange, new feature of special relativity. There is often a fine line between the two so we have to be very careful!

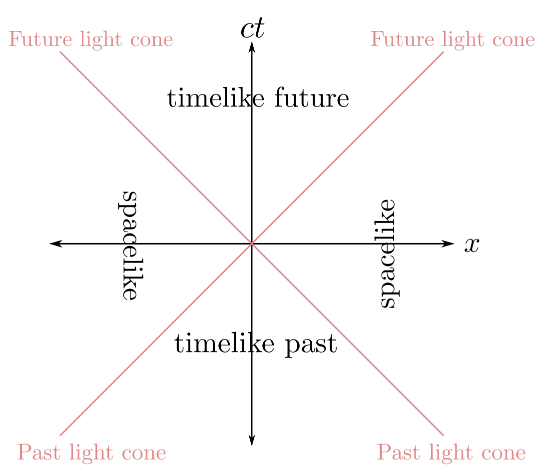
Now that we have the line element, let's put it to use! We will see how length contraction and time dilation, two of the most famous special relativity effects, are encoded in it. But first, light cones.

Light cones, proper time, proper distance

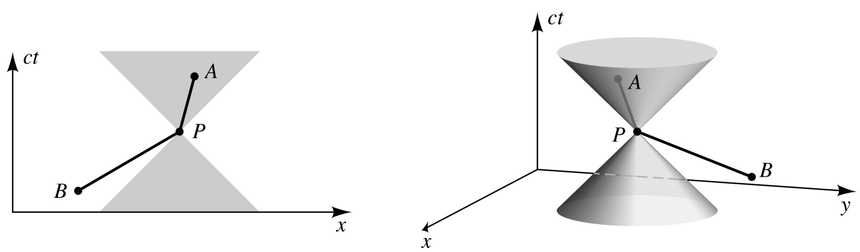
There are three possibilities for the line element:

$$ds^2 > 0 \quad \text{(spacelike)}\,, \qquad ds^2 = 0 \quad \text{(null)}\,, \qquad ds^2 < 0 \quad \text{(timelike)}\,. \tag{0.28}$$

The first case is if the separation is "mostly space" and the third case is if the separation is "mostly time." Massless particles travel with $v = c$, so $dx^2 + dy^2 + dz^2 = cdt^2$ and $ds^2 = 0$. They follow null trajectories. This marginal case defines a lightcone. The apex of the light-cone and any point within the lightcone defines a timelike trajectory, while the apex and any point outside the lightcone defines a spacelike trajectory. Physical particles cannot follow spacelike trajectories, i.e. they cannot travel outside the lightcone. (Hypothetical particles known as tachyons, which have not been observed and most likely don't make sense, travel faster than light and can therefore follow spacelike trajectories.) Instead, physical particles have $v < c$ and therefore follow timelike trajectories which are constrained to lie within the lightcone. Clearly c is special: not only is it the same in every frame of reference, but in special relativity nothing travels faster than c.



So, while time and space are flexible in special relativity, there is still a sharp **causal structure**, which encodes cause and effect, built into the line element. This is represented in the 2d image above, where the lightcone is represented in red. Any spacetime event defines lightcones, and therefore spacelike separated and timelike separated events. This is shown in 2d and 3d below.



For timelike trajectories we often speak of the *proper time*

$$d\tau^2 \equiv -ds^2/c^2 > 0\,. \tag{0.29}$$

This is called a time because it is literally the time a clock would measure if it followed this trajectory; in the clock's frame of reference $\Delta t = \Delta\tau$ and $\Delta x = \Delta y = \Delta z = 0$. Furthermore, timelike separated events can be made ot occur at the same spatial point by appropriately choosing a reference frame, hence they are separated by "pure time" in some frame.

For spacelike-separated events we often speak of the *proper distance* or *proper length*

$$d\sigma^2 \equiv ds^2 > 0\,. \tag{0.30}$$

We can always choose a reference frame so that spacelike separated events occur at the same point in time, i.e. simultaneously, hence they are separated by "pure space" in some frame. The proper distance is the distance of something (say a rod) measured in this frame (i.e. with $\Delta t = 0$)

Time dilation

We already saw time dilation, but let's see it more directly from the line element. Say Alice is whizzing along in a rocket at speed v. In Alice's frame with coordinates (t_A, x_A, y_A, z_A), her elapsed time between two events is

$$\Delta s^2 = -c^2\Delta t_A^2 + \Delta x_A^2 + \Delta y_A^2 + \Delta z_A^2 = -c^2\Delta t_A^2\,. \tag{0.31}$$

where the spatial distances vanish since Alice is at rest in her frame (and therefore not going anywhere, $\Delta x_A = \Delta y_A = \Delta z_A = 0$).

Bob is relaxing in his lawnchair, in a frame with coordinates (t_B, x_B, y_B, z_B). For him, Alice travels $v\Delta t_B$ between the two events, so

$$\Delta s^2 = -c^2\Delta t_B^2 + v^2\Delta t_B^2 = (v^2 - c^2)\Delta t_B^2\,. \tag{0.32}$$

But the invariance of the line element means Δs^2 is always the same (this is why I didn't write Δs_A^2 and Δs_B^2!). So we have

$$-c^2\Delta t_A^2 = (v^2 - c^2)\Delta t_B^2 \implies \Delta t_B = \frac{\Delta t_A}{\sqrt{1 - v^2/c^2}} > \Delta t_A\,. \tag{0.33}$$

So the same two events seem to take longer according to Bob!

We can use calculus (i.e. integration) to find the elapsed proper time even for non-inertial (i.e. accelerating) reference frames:

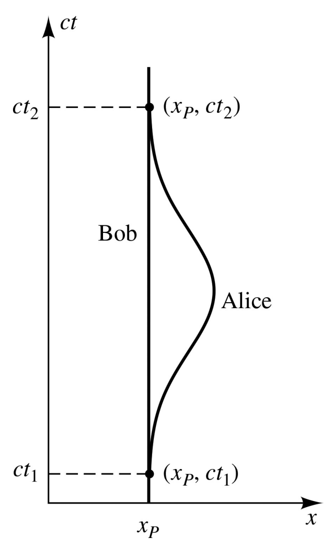
$$\Delta\tau = \int d\tau = \int \sqrt{-ds^2/c^2} = \int \sqrt{dt^2 - \frac{dx^2 + dy^2 + dz^2}{c^2}} = \int \sqrt{dt^2 - v(t)^2 dt^2/c^2} \tag{0.34}$$

$$= \int dt\sqrt{1 - v(t)^2/c^2} \; < \int dt = \Delta t\,. \tag{0.35}$$

The clock is accelerating with respect to the coordinates (t, x, y, z).

This has been verified experimentally! In the 1960's scientists took a heavy atomic clock onto a plane and flew it around. The speed $v \sim 600mph \sim 300m/s \sim 10^{-6}c$ which means $\sqrt{1 - v^2/c^2} \approx 1 - \frac{1}{2} \times 10^{-12}$. So this is a difference of 0.0000000001%, which is easily measurable by modern technology.

Example 6: The twin paradox is one of the most famous paradoxes of special relativity. It stems from the discussion above: Bob sees Alice whizzing by and we find $\Delta t_B > \Delta t_A$ for two events which have no spatial separation in Alice's frame, $\Delta x_A = \Delta y_A = \Delta z_A = 0$. But there is a symmetry: in Alice's frame, Bob is whizzing by, so we will find $\Delta t_A > \Delta t_B$ for two events which have no spatial separation in Bob's frame, $\Delta x_B = \Delta y_B = \Delta z_B = 0$. So what happens if we have two twins Alice and Bob, who are the same age, and Alice departs on a high-speed journey at time t_1 and returns at time t_2? Who is older?



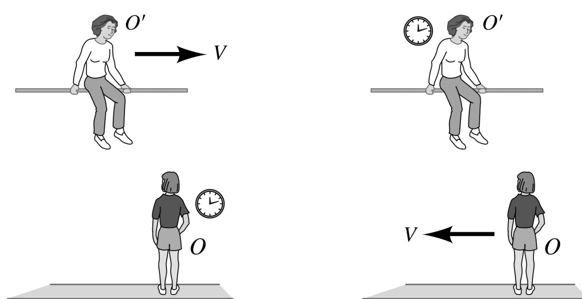
When drawn on a spacetime diagram as in the above, we see that the two situations are not symmetric. In particular, Alice has to accelerate to leave Bob and decelerate to return to Bob. Integrating the *proper time* for each observer, we will find that Bob is older.

Another way to see the results from the example and main text above is to look back at the proper time $d\tau^2 = -ds^2/c^2$. It is *maximized* for curves at constant x, y, z, since traversing in these coordinates *decreases* $d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$ due to the minus signs on the spatial coordinates. In other words, adding wiggles in the spatial direction decreases the proper time. In ordinary Euclidean geometry we are used to the fact that straight lines minimize the distance between points, and adding wiggles *increases* the distance.

Proper lengths in Minkowski geometry are increased by adding spatial wiggles and decreased by adding temporal wiggles. This is because $d\sigma^2 = -dt^2 + \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$, and space comes with positive signs while time comes with negative signs.

Length contraction

Given that time gets all mixed up in different frames of reference, we should also expect space to get all mixed up, since the two are linked together through the invariant line element $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$. This indeed happens, and the most famous consequence is the notion of length contraction. The simplest way to see this is to use our result for time dilation.



Say we have a rod of length L in its rest frame, and it is traveling at speed v (relative to a clock) parallel to its length. We can measure the length by measuring the proper time it takes a clock to travel the full length of the rod. In observer O's rest frame (left figure above), the time a clock measures is $v\Delta t = L'$, where L' is the length of the rod in O's rest frame. (Observer O is like Alice in the time dilation section.) In the rod's rest frame (right figure above), the time a clock in that frame measures is given by $v\Delta t' = L$. (Observer O' is like Bob in the time dilation section.) Using $\Delta t' = \Delta t/\sqrt{1 - v^2/c^2}$. and combining these two equations gives[4]

$$\frac{L'}{L} = \frac{\Delta t}{\Delta t'} = \sqrt{1 - v^2/c^2} \implies L' = L\sqrt{1 - v^2/c^2}\,. \tag{0.36}$$

This is also problem 17 in Chapter 4 of Hartle. The fact that length contracts and time dilates seems to conflict with the symmetry between space and time. The difference is *not* due to the minus sign in the metric; instead it is because the question being asked is *asymmetric*.

Lorentz transformations

While the thought experiments are a fun and physical way to deduce aspects of special relativity from the constancy of the speed of light, it is more economical to have some universal formulas, which we will see are the Lorentz transformations. These relate the coordinates in two different inertial frames. These coordinate changes are a **symmetry** of physics; physical laws are the same in all inertial frames. Let's warm up with a familiar example.

Example 7: What are the symmetries of the 3d Euclidean plane? Intuitively, we know that rotations, translations, and reflections don't change anything about the geometry of the plane. We can see this from the metric

$$ds^2 = dx^2 + dy^2 + dz^2\,. \tag{0.37}$$

The transformations stated above act on the metric as

$$\begin{aligned} &\text{Translation (in } x\text{):} && x \to x + x_0 && (0.38)\\ &\text{Reflection (in } x\text{):} && x \to -x && (0.39)\\ &\text{Rotation (in } x-y \text{ plane):} && x \to x\cos\theta + y\sin\theta\,, \quad y \to y\cos\theta - x\sin\theta && (0.40)\\ & && && (0.41) \end{aligned}$$

We also have translations and reflections in y, z and rotations in the $x - z$ plane and $y - z$ plane. These all leave the metric invariant. We can check it for the rotations in the $x - y$ plane:

$$dx^2 + dy^2 \to (dx\cos\theta + dy\sin\theta)^2 + (dy\cos\theta - dx\sin\theta)^2 = dx^2 + dy^2\,. \tag{0.42}$$

How do we understand the symmetries of Minkowski space? The metric

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \tag{0.43}$$

[4] Why did we use $\Delta t' = \Delta t/\sqrt{1 - v^2/c^2}$ instead of $\Delta t = \Delta t'/\sqrt{1 - v^2/c^2}$? After all the two frames seem pretty symmetric. What chose the direction of the time dilation is that in O's frame, the two spacetime events (front of rod passes O, back of rod passes O) have zero separation in space; recall the discussion of Alice and Bob in the time dilation section above. A spacetime diagram really helps with the analysis; see Example 10 for an alternative derivation of length contraction with the aid of a spacetime diagram.

clearly has all the symmetries of the 3d Euclidean plane. It also has time reflection and time translation. But can we "rotate" space into time? Yes! It is called a hyperbolic rotation[5], and is given by

$$ct \to (\cosh\theta)ct - (\sinh\theta)x \tag{0.44}$$
$$x \to -(\sinh\theta)ct + (\cosh\theta)x \tag{0.45}$$

Notice the similarity with ordinary rotations, up to trig functions becoming hyperbolic trig functions and the signs being slightly different. This is known as a **Lorentz boost**.

Example 8: Let's check that this Lorentz boost keeps the Minkowski line element invariant. We can ignore the y and z directions since we are not transforming them. We have

$$-c^2dt^2 + dx^2 \to -c^2\left[(\cosh\theta)cdt - (\sinh\theta)dx\right]^2 + \left[(-\sinh\theta)cdt + (\cosh\theta)dx\right]^2 \tag{0.46}$$
$$= -c^2(\cosh^2\theta - \sinh^2\theta)dt^2 + (\cosh^2\theta - \sinh^2\theta)dx^2 = -c^2dt^2 + dx^2\,. \tag{0.47}$$

There are also Lorentz boosts between time and the y, z directions. The full set of Lorentz transformations change our frame of reference, and the laws of physics are independent of the change of frame. Here are some examples

- The laws of physics are the same here as ten miles from here (spatial translations).
- The laws of physics are the same now as ten minutes from now (time translations).
- The Large Hadron Collider would measure the same laws of physics if it was rotated by some angle (spatial rotations).
- The laws of physics are the same whether I'm on my lawnchair or whizzing by in a rocket at constant velocity (Lorentz boosts).

We haven't quite seen that Lorentz boosts correspond to changing into an inertial frame at some velocity, as stated above. We can see this by observing what the line of constant spatial position looks like for the Lorentz transformed observer:

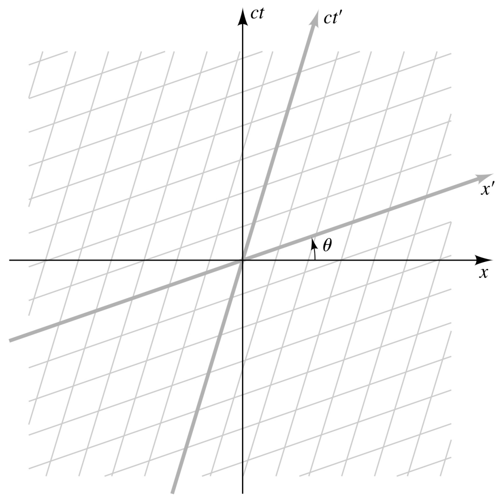
$$ct' = (\cosh\theta)ct - (\sinh\theta)x\,, \qquad x' = (\cosh\theta)x - (\sinh\theta)ct \tag{0.48}$$

Consider the line of constant x' for our primed observer, say $x' = 0$. This is a spatial point that doesn't move relative to our observer, even as time progresses. The above transformation gives

$$0 = (\cosh\theta)x - (\sinh\theta)ct \implies \frac{x}{t} = c\tanh\theta\,. \tag{0.49}$$

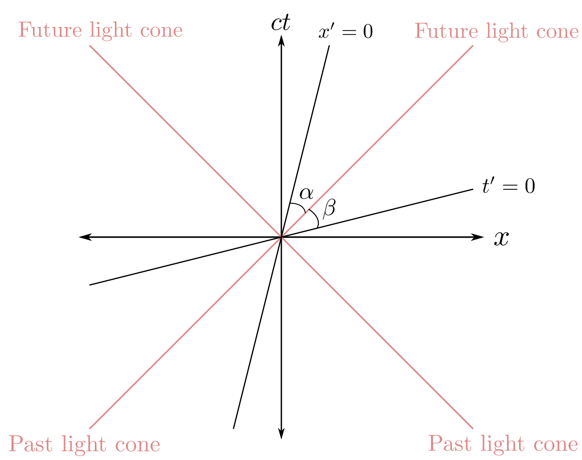
[5]An ordinary rotation keeps invariant spheres, which satisfy $\Delta x^2 + \Delta y^2 + \Delta z^2 = const.$ A hyperbolic rotation keeps invariant hyperbolas, which satisfy $\Delta x^2 + \Delta y^2 + \Delta z^2 - \Delta t^2 = const.$

This means the unprimed frame is moving at speed $v = x/t = c\tanh\theta$ relative to the primed frame of reference![6] On a spacetime diagram, the new coordinates are represented in the image below.



With these formulas, we can now return to our spacetime diagrams and derive a very powerful computational aid.

Example 9: Say we have a boosted observer moving at velocity $v = c\tanh\theta$ relative to another observer. We would like to draw the boosted observer's worldline and constant-time slice (sometimes called a "line of simultaneity"). Our spacetime diagram is given by:



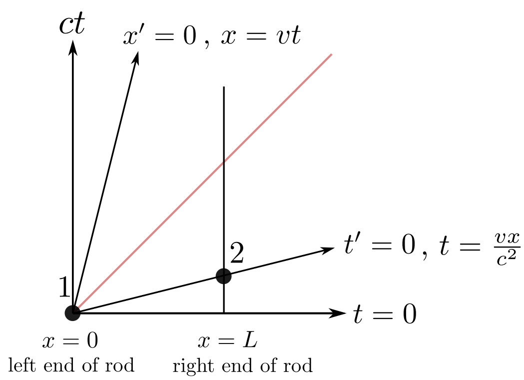
We can show $\beta = \alpha$ as follows. Using (0.48) and (0.49), the lines $x' = 0$ and $t' = 0$ can be written in the unprimed frame as $x = ct\tanh\theta = (ct)\frac{v}{c}$ and $ct = x\tanh\theta = x\,\frac{v}{c}$,

[6]We will often speak of a "boosted" observer, but notice it is symmetric: each observer looks boosted to the other one, since the only meaningful quantity is the relative velocity between the two.

respectively. In other words $\frac{x}{ct} = \tanh\theta$ for the $t' = 0$ line equals $\frac{ct}{x} = \tanh\theta$ for the $x' = 0$ line. So the $x' = 0$ line is bent by the same angle from the $x = 0$ line as the $t' = 0$ line is bent from the $t = 0$ line. Therefore $\alpha = \beta$, and the boosted observer's constant-space slice (i.e. worldline) is symmetrically reflected across the lightcone from their constant-time slice (i.e. line of simultaneity). This had to be true since it is trivially true in the unprimed coordinates.

———————————— END LEC 2 ————————————

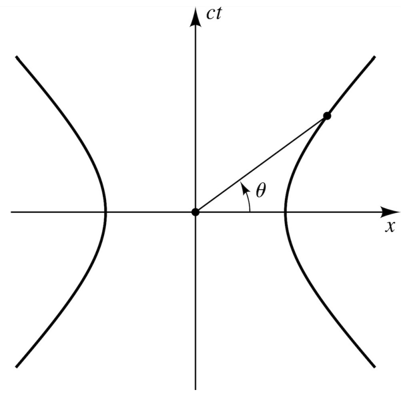
Example 10: We want to use spacetime diagrams and the formulas from Example 8 to derive the formula for length contraction. We have a rod of length L in its rest frame. These are the unprimed coordinates in the diagram below. We want to measure the length in the primed frame, which is moving at speed v parallel to the length of the rod. It is important to remind ourselves *what it means* to measure the length of something. It is the distance between two spatial points **at the same instant in time**. In the diagram below, this means we want the distance between spacetime events 1 and 2 on the line $t' = 0$:



This distance is given by

$$\Delta s^2 = -c^2\Delta t^2 + \Delta x^2 = -c^2\frac{v^2L^2}{c^4} + L^2 = L^2(1 - v^2/c^2) \implies L' = L\sqrt{1 - v^2/c^2}\,. \quad (0.50)$$

In what sense is this hyperbolic angle an angle? Let us recall the definition of an angle in Euclidean geometry. Drawing a circle around the origin, we define an angle θ as the ratio of the length of the portion of the circle it cuts out to the radius. In Minkowski geometry the invariant objects are hyperbolas, so θ is now the ratio of the length of the portion of the hyperbola it cuts out to the distance from the origin, see below.



Familiar trigonometric formulas become hyperbolic trigonometric formulas, e.g.

$$ct = R\sinh\theta\,, \qquad x = R\cosh\theta\,, \tag{0.51}$$

where R is the distance from the origin to the hyperbola.

We see that our formulas for worldlines and lines of simultaneity are useful in deriving some of the consequences of special relativity. Another easy consequence is the loss of simultaneity, which by now should hopefully start feeling natural. To see this, revisit Figure 1, and notice that on the right-hand-side the location of the spacetime event of the lightray leaving B's worldline can be calculated by reflecting A's worldline across the lightray and intersecting it with B's worldline (this is just using $\alpha = \beta$ from Example 8).

Another way of writing the Lorentz boosts is as follows. Using $v = c\tanh\theta$ and $\cosh^2\theta - \sinh^2\theta = 1$, we have

$$\cosh\theta = \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \gamma\,, \qquad \sinh\theta = \frac{v/c}{\sqrt{1 - v^2/c^2}}\,. \tag{0.52}$$

We can use these to write the boosts as

$$t' = \gamma(t - vx/c^2)\,, \qquad x' = \gamma(x - vt)\,. \tag{0.53}$$

Example 11: Here is a very fun thought experiment. Say we shoot a nail with length ℓ at a wall of thickness L with a hole in it. The hole is big enough such that the nail can go through, although the back of the nail will run into the wall (e.g. the nail's sharp part is cylindrical with radius r_1 the hole in the wall is circular with radius r_2, and the back of the nail is circular with radius r_3 such that $r_1 < r_2 < r_3$. On the other side of the wall is a balloon. Does the nail pop the balloon?

At first it seems the two reference frames can disagree. If the nail is going fast enough, then in the nail's frame, the wall is very length-contracted so the nail will definitely get through and pop the balloon. In the wall's frame the nail is length contracted, and if the nail is going fast enough it will become shorter than the thickness of the wall (length contraction can make things arbitrarily short), so it seems the wall will stop it from popping the balloon.

The resolution is very instructive. The basic point is that nothing travels faster than light, not even information. So even though the back of the nail hits the wall, *that information needs to be communicated to the front of the nail.* Until it is, the front of the nail keeps sailing along. This effect is what leads to the balloon popping in the wall's reference frame as well.

Let's analyze it more quantitatively. In the wall's frame, the length of the nail is

$$\ell' = \ell/\gamma \,. \tag{0.54}$$

We have a race: the front of the nail wants to traverse the distance $L - \ell'$ going at speed v, and light wants to traverse the distance L going at speed c. The balloon will pop if light beats the nail.

$$vt_{nail} = L - \ell' \,, \quad ct_{light} = L \,, \qquad \text{Balloon pops if} \quad t_{nail} < t_{light} \tag{0.55}$$

So the equation we get is

$$\frac{L - \ell/\gamma}{v} < \frac{L}{c} \,. \tag{0.56}$$

We can compute this same equation in the nail's reference frame, but notice that with the aid of the invariant line element, **we know that if we compute an invariant in one frame of reference that it will be the same in the other frame of reference!** In this case, the invariant we are computing is Δs^2 for two spacetime events: the back of the nail hitting the wall and the front of the nail hitting the balloon. In the wall's frame, we put the first event at $x = t = 0$, and the second event at $t = (L - \ell/\gamma)/v$ and $x = L$. The spacetime interval is therefore

$$\Delta s^2 = -c^2 \left(\frac{L - \ell/\gamma}{v} \right)^2 + L^2 \tag{0.57}$$

If this is a timelike distance, $\Delta s^2 < 0$, then light from event 1 can reach event 2 and stop the balloon from popping. However, if they are spacelike separated, $\Delta s^2 > 0$, then there is no way event 1 can influence event 2, so the balloon will pop. This condition gives

$$\Delta s^2 > 0 \implies \frac{L}{c} > \frac{L - \ell/\gamma}{v} \,, \tag{0.58}$$

exactly the same as before! While it is extremely fun to analyze the physics in various reference frames, the invariant line element Δs^2 often saves us the trouble of having to do so.

Relativistic addition of velocities

Hyperbolic angles are a very useful way to describe boosts, just like ordinary angles are very useful ways to describe rotations. They add in a simple way!

Example 12: Let's show using a rotation matrix that a rotation by θ_1 followed by a rotation by θ_2 (in the same direction) is equivalent to a rotation by $\theta_1 + \theta_2$. The first rotation is given by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{0.59}$$

A second rotation gives

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{0.60}$$

Altogether we have

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{0.61}$$

$$= \begin{pmatrix} \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & \cos\theta_2\sin\theta_1 + \sin\theta_2\cos\theta_1 \\ -\sin\theta_2\cos\theta_1 - \cos\theta_2\sin\theta_1 & -\sin\theta_2\sin\theta_1 + \cos\theta_2\cos\theta_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{0.62}$$

$$= \begin{pmatrix} \cos(\theta_1+\theta_2) & \sin(\theta_1+\theta_2) \\ -\sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{0.63}$$

We have enough intuition for spatial rotations that this result is fairly obvious.

While velocities don't simply add as in Newtonian physics, the hyperbolic angles θ in $v = c\tanh\theta$ do. To see this, we write our Lorentz boost in the $x - t$ directions as a matrix transformation:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\theta & -\sinh\theta \\ -\sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{0.64}$$

Writing two successive boosts by hyperbolic angles θ_1 and θ_2 gives

$$\begin{pmatrix} ct'' \\ x'' \end{pmatrix} = \begin{pmatrix} \cosh\theta_2 & -\sinh\theta_2 \\ -\sinh\theta_2 & \cosh\theta_2 \end{pmatrix} \begin{pmatrix} \cosh\theta_1 & -\sinh\theta_1 \\ -\sinh\theta_1 & \cosh\theta_1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{0.65}$$

$$= \begin{pmatrix} \cosh\theta_2\cosh\theta_1 + \sinh\theta_2\sinh\theta_1 & -\cosh\theta_2\sinh\theta_1 - \sinh\theta_2\cosh\theta_1 \\ -\sinh\theta_2\cosh\theta_1 - \cosh\theta_2\sinh\theta_1 & \sinh\theta_2\sinh\theta_1 + \cosh\theta_2\cosh\theta_1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{0.66}$$

$$= \begin{pmatrix} \cosh(\theta_1+\theta_2) & -\sinh(\theta_1+\theta_2) \\ -\sinh(\theta_1+\theta_2) & \cosh(\theta_1+\theta_2) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{0.67}$$

The velocities themselves add as follows. We pick $v_1 = c\tanh\theta_1$ and $v_2 = c\tanh\theta_2$, both in the same direction. They combine to

$$v = c\tanh(\theta_1+\theta_2) = c\,\frac{\tanh\theta_1 + \tanh\theta_2}{1 + \tanh\theta_1\tanh\theta_2} = c\,\frac{\frac{v_1}{c} + \frac{v_2}{c}}{1 + \frac{v_1}{c}\frac{v_2}{c}} = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}\,. \tag{0.68}$$

This is called the *Law of Relativistic Addition of Velocities.* Notice that if the velocities are small, $v_1 \ll c$, $v_2 \ll c$, this becomes the ordinary Newtonian (or Galilean) addition of velocities, $v \approx v_1 + v_2$. And if one of the velocities is already at the speed of light, say $v_1 = c$, we have

$$v = \frac{c + v_2}{1 + \frac{cv_2}{c^2}} = c\,\frac{1 + \frac{v_2}{c}}{1 + \frac{v_2}{c}} = c\,. \tag{0.69}$$

As promised, the speed of light is the same in every reference frame!

Example 13: Let's use the formula for Lorentz boosts

$$t' = \gamma(t - vx/c^2)\,, \qquad x' = \gamma(x - vt)\,, \qquad y' = y\,, \qquad z' = z \tag{0.70}$$

to derive the addition of velocities formula from before, and to obtain the formula for motion in the y or z directions. Let's call $v \equiv v_1^{(x)}$ to indicate it is in the x direction (and to distinguish it from $v_2^{(x)}$ which will appear below).

Velocity in the primed frame is given by

$$\frac{dx'}{dt'} = \frac{d(\gamma(x - v_1^{(x)}t))}{d(\gamma(t - v_1^{(x)}x/c^2))} = \frac{dx - v_1^{(x)}dt}{dt - \frac{v_1^{(x)}}{c^2}dx} = \frac{\frac{dx}{dt} - v_1^{(x)}}{1 - \frac{v_1^{(x)}}{c^2}\frac{dx}{dt}} = \frac{v_2^{(x)} - v_1^{(x)}}{1 - \frac{v_1^{(x)}v_2^{(x)}}{c^2}}\,, \tag{0.71}$$

where $v_1^{(x)}$ is the speed between the primed and unprimed frames (the superscript denoting that it is purely in the x direction), and $v_2^{(x)} = \frac{dx}{dt}$ is the x component of the velocity of something in the unprimed frame. This formula has some sign differences from the one from before since $v_1^{(x)}$ is the (t', x', y', z') frame so is traveling in the positive x direction, and $v_2^{(x)} = \frac{dx}{dt}$, so if it is positive then it is also traveling in the positive x direction. Therefore the velocities should subtract not add. (To recover the formula from before just take $v_1^{(x)} \to -v_1^{(x)}$ to flip its direction and therefore add velocities.)

We can repeat this in the other directions:

$$\frac{dy'}{dt'} = \frac{dy}{d(\gamma(t - v_1^{(x)}x/c^2))} = \frac{dy}{\gamma\left(dt - \frac{v_1^{(x)}}{c^2}dx\right)} = \frac{\frac{dy}{dt}}{\gamma\left(1 - \frac{v_1^{(x)}}{c^2}\frac{dx}{dt}\right)} = \frac{v_2^{(y)}}{\gamma\left(1 - \frac{v_1^{(x)}v_2^{(x)}}{c^2}\right)} \tag{0.72}$$

$$\frac{dz'}{dt'} = \frac{v_2^{(z)}}{\gamma\left(1 - \frac{v_1^{(x)}v_2^{(x)}}{c^2}\right)} \tag{0.73}$$

where $\gamma = \frac{1}{\sqrt{1 - v_1^2/c^2}}$ as before.

Before we go onto relativistic kinematics (momenta, forces, etc.), it will help to introduce some notation, applicable to both the non-relativistic and relativistic contexts.

4-vectors

3-vectors and index notation

You are used to vectors and dot products. We will introduce a notation $\mathbf{x} = \vec{x} \equiv x^i$. The index i can be $i = 1, 2, 3$ in three dimensions, and we have $x^1 = x$, $x^2 = y$, $x^3 = z$. Then we can write vector equations as e.g.

$$\mathbf{p} = m\mathbf{v} \implies p^i = mv^i \tag{0.74}$$

and dot products as

$$\mathbf{x} \cdot \mathbf{x} = \sum_i x^i x^i = x^2 + y^2 + z^2 \,. \tag{0.75}$$

This notation is useful because it lets us represent matrix multiplication compactly; for example the rotation matrix from Example 10 can be written as

$$R(\theta)\mathbf{x} = \sum_j R^{ij}(\theta) x^j \,. \tag{0.76}$$

We can simplify things by introducing the *Einstein summation convention*, which is:

$$a_i b^i \equiv \sum_i a^i b^i \,. \tag{0.77}$$

In other words, when you see the same index appearing somewhere up and somewhere else down, you are instructed to perform the sum on the RHS of the above. Thus our matrix multiplication for rotation becomes

$$R(\theta)\mathbf{x} = R^{ij} x_j \,. \tag{0.78}$$

It is also important to note that when writing things this way the **order doesn't matter**, so $R^{ij} x_j = x_j R^{ij}$. This feels a little funny at first since matrix multiplication is non-commutative, but once you have written it in terms of components it is just scalar multiplication! This notation will be **essential** as we move onto general relativity, so let's get some practice with it.

Example 14: Let's check matrix multiplication,

$$\begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} \begin{pmatrix} B^{11} & B^{12} \\ B^{21} & B^{22} \end{pmatrix} = \begin{pmatrix} A^{11}B^{11} + A^{12}B^{21} & A^{11}B^{12} + A^{12}B^{22} \\ A^{21}B^{11} + A^{22}B^{21} & A^{21}B^{12} + A^{22}B^{22} \end{pmatrix} \equiv C \tag{0.79}$$

This can be written very compactly as

$$AB = A^{ij} B_j^k \equiv C^{ik} \,. \tag{0.80}$$

This equation has the j index repeated, so we sum over it. It has two non-repeated indices, i and k, and that is because the resulting object is a matrix, which has two indices. Let's check:

$$AB = A^{ij}B_j^k = \sum_j A^{ij}B^{jk} = A^{i1}B^{1k} + A^{i2}B^{2k}\,. \tag{0.81}$$

Now we can check entry by entry if this matches the matrix multiplication. For example,

$$C^{12} = A^{11}B^{12} + A^{12}B^{22}\,. \tag{0.82}$$

This is precisely the top right entry of the matrix for C, as required! Whenever in doubt, expand out the sum and plug in particular values for the "free" (i.e. non-repeated) indices. After you do this many times you will get the hang of it.

———————————— END LEC 3 ————————————

Example 15: Two rotation matrices multiply as

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta_2 & \sin\theta_2 \\ -\sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{0.83}$$

We can write this as

$$\mathbf{X} = R(\theta_2)R(\theta_1)\mathbf{x} \implies X^i = R^{ij}(\theta_2)R_{jk}(\theta_1)x^k \tag{0.84}$$

To check this, we expand out using our summation convention rule. j and k are repeated so we sum over them. i is not repeated because this is a vector equation. We will check this vector equation component by component:

$$X^1 = R^{11}(\theta_2)R^{11}(\theta_1)x^1 + R^{12}(\theta_2)R^{21}(\theta_1)x^1 + R^{11}(\theta_2)R^{12}(\theta_1)x^2 + R^{12}(\theta_2)R^{22}(\theta_1)x^2\,, \tag{0.85}$$

$$X^2 = R^{21}(\theta_2)R^{11}(\theta_1)x^1 + R^{22}(\theta_2)R^{21}(\theta_1)x^1 + R^{21}(\theta_2)R^{12}(\theta_1)x^2 + R^{22}(\theta_2)R^{22}(\theta_1)x^2\,. \tag{0.86}$$

You can check that this is precisely the rule for multiplying the two rotation matrices together and multiplying them with the vector x^i.

4-vectors and index notation

In Newtonian physics three-vectors like velocity, momentum, and force are very important. In special relativity these will be generalized to four-vectors which also include the time direction, since time and space are inextricably linked. We will use the same notation for four-vectors, although often we will use a Greek index to denote a four-vector ($\mu = 0, 1, 2, 3$) and a Latin index for a three vector ($i = 1, 2, 3$), e.g.

$$a^\mu = (a^0, a^1, a^2, a^3) = (a^t, a^x, a^y, a^z)\,, \qquad a^i = (a^1, a^2, a^3) = (a^x, a^y, a^z)\,. \tag{0.87}$$

We can expand this in a basis of unit 4-vectors which we will call $\mathbf{e}_t$, $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$. The latter three you have seen before, $\mathbf{e}_x = \hat{\imath}$, $\mathbf{e}_y = \hat{\jmath}$, $\mathbf{e}_z = \hat{k}$. Our 4-vector is represented as

$$\mathbf{a} = a^\mu = a^t\mathbf{e}_t + a^x\mathbf{e}_x + a^y\mathbf{e}_y + a^z\mathbf{e}_z = a^\alpha\mathbf{e}_\alpha = a^\alpha e^\mu_\alpha \tag{0.88}$$

In the second-to-last expression we used the Einstein summation convention and introduced a somewhat confusing object, $\mathbf{e}_\alpha$. It is really a 4×4 matrix, since for each value of $\alpha = 0,1,2,3$ it is a 4-vector. The final expression is an equivalent way of writing it which makes this clear (lose the boldface, add an index). While we will use boldface both for 3-vectors and 4-vectors, hopefully it is clear from context which we mean.

4-vectors are **defined** to transform in the same way as coordinates do. So, for a boost in the x direction by speed v, we have

$$a^{t'} = \gamma(a^t - va^x/c^2)\,, \qquad a^{x'} = \gamma(a^x - va^t)\,, \qquad a^{y'} = a^y\,, \qquad a^{z'} = a^z\,. \tag{0.89}$$

We can write this compactly with the summation convention as we did with rotations,

$$a^{\mu'} = \Lambda^{\mu'}_\nu a^\nu\,. \tag{0.90}$$

Λ^μ_ν is the Lorentz transformation matrix – it is a 4×4 matrix depending on v whose form was given below Example 12. We will often drop the primes on indices when writing equations in this way.

For the rest of the course we will set $c = 1$. This means time and space have the same units, so velocities are unitless. In particular it means 1 sec = 299792458 meters. You can use dimensional analysis to restore the factors of c in your formulas if you wish. To go from length to time you divide by c, and to go from time to length you multiply by c. Similarly, if you see e.g. $1 - v^2$, since 1 is unitless and v has units of velocity, then you know to bring in factors of c to turn it into $1 - v^2/c^2$. Bringing back these factors is often more natural: you would look at me funny if I said this lecture is 1.6×10^{12} m long.

We can define a dot product or "scalar product" in special relativity in analogy to the dot product of ordinary three-vectors:

$$\mathbf{a}\cdot\mathbf{b} = a_ib^i = a^1b^1 + a^2b^2 + a^3b^3 \;\longrightarrow\; \mathbf{a}\cdot\mathbf{b} = a_\mu b^\mu = -a^0b^0 + a^1b^1 + a^2b^2 + a^3b^3\,. \tag{0.91}$$

Notice the minus sign mimics the Minkowski geometry $ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$, whereas the Euclidean dot product on the left mimics the Euclidean geometry $ds^2 = dx^2 + dy^2 + dz^2$. We can also write these two dot products as

$$a_ib^i = a^ib^j\delta_{ij}\,, \qquad a_\mu b^\mu = a^\mu b^\nu\eta_{\mu\nu} \tag{0.92}$$

where

$$\delta_{ij} = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\,, \qquad \eta_{\mu\nu} = \begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\,. \tag{0.93}$$

When using Euclidean dot products, whether indices are superscripts or subscripts is irrelevant. But in relativity we define

$$a_\beta = a^\alpha \eta_{\alpha\beta} \implies \mathbf{a} \cdot \mathbf{a} = a_\beta a^\beta = a^\alpha \eta_{\alpha\beta} a^\beta \tag{0.94}$$

which means $a_t = -a^t$. (Notice dot products defined with δ_{ij} in Euclidean geometry do not lead to this minus sign, and so $a_x = a^x$ and the index placement does not matter.)

Here is the key point about dot products: they are invariant, meaning they are the same in any frame of reference! This is just like we saw with the invariant line element Δs^2. This is very familiar from Euclidean geometry: consider the dot product of two vectors $\mathbf{a} \cdot \mathbf{b} = ab\cos\theta$. Performing a spatial rotation, translation, or reflection won't change this quantity since it leaves the lengths of the vectors and the angle between them unchanged. It's the same for our 4-vector dot product, now generalized to include boosts!

Example 16: Let's show the invariance of $\mathbf{a} \cdot \mathbf{b}$. Since we know Δs^2 is invariant, we have

$$\Delta s^2 = \Delta s^\alpha \eta_{\alpha\beta} \Delta s^\beta = \Delta s^{\gamma'} \eta_{\gamma'\delta'} \Delta s^{\delta'} = \Delta s^\alpha \Lambda^{\gamma'}_{\alpha} \eta_{\gamma'\delta'} \Lambda^{\delta'}_{\beta} \Delta s^\beta \tag{0.95}$$

$$\implies \eta_{\alpha\beta} = \Lambda^{\gamma'}_{\alpha} \eta_{\gamma'\delta'} \Lambda^{\delta'}_{\beta} \tag{0.96}$$

We can now write

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{4}\left[(\mathbf{a}+\mathbf{b})^2 - (\mathbf{a}-\mathbf{b})^2\right] \tag{0.97}$$

and use the fact that $(\mathbf{a}+\mathbf{b})^2$ and $(\mathbf{a}-\mathbf{b})^2$ are invariant. Another way to see the result is to write

$$\mathbf{a} \cdot \mathbf{b} = a^\alpha \eta_{\alpha\beta} b^\beta \to (a^{\gamma'} \Lambda^{\alpha}_{\gamma'}) \eta_{\alpha\beta} (\Lambda^{\beta}_{\delta'} b^{\delta'}) = a^{\gamma'} \eta_{\gamma'\delta'} b^{\delta'} \,, \tag{0.98}$$

where in the last equality we used $\Lambda^{\alpha}_{\gamma'} \eta_{\alpha\beta} \Lambda^{\beta}_{\delta'} = \eta_{\gamma'\delta'}$, which is just (0.96) with relabeled indices.

Relativistic kinematics

In Newtonian physics we describe the motion of a particle as $x^i(t)$. We can promote this to a 4-vector as

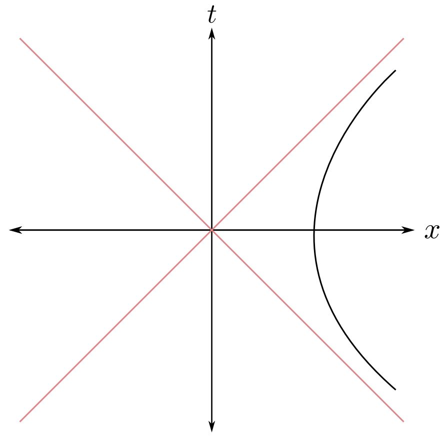
$$x^\mu(t) = (x^0(t), x^1(t), x^2(t), x^3(t)) = (t, x(t), y(t), z(t)) \,. \tag{0.99}$$

This makes time seem special, but it's just a choice for how to parameterize a curve. We can instead parameterize in terms of some arbitrary parameter σ:

$$x^\mu(\sigma) = (t(\sigma), x(\sigma), y(\sigma), z(\sigma)) \,. \tag{0.100}$$

For example, a hyperbolic worldline can be parameterized as

$$x = a^{-1}\cosh\sigma\,, \qquad t = a^{-1}\sinh\sigma\,, \qquad y = z = 0\,. \tag{0.101}$$



This is similar to the parameterization of a circle of radius a^{-1} in the Euclidean plane, $x = a^{-1}\cos\sigma$, $y = a^{-1}\sin\sigma$. For timelike worldlines, like the hyperbolic one above, it is useful to parameterize in terms of the *proper time* τ defined by

$$d\tau = \sqrt{-ds^2} = \sqrt{-dx^\alpha dx^\beta \eta_{\alpha\beta}}\,. \tag{0.102}$$

For the hyperbolic worldline we have

$$dt = a^{-1}\cosh\sigma d\sigma\,, \qquad dx = a^{-1}\sinh\sigma d\sigma\,, \qquad dy = dz = 0 \tag{0.103}$$

$$\implies d\tau = \sqrt{dt^2 - dx^2} = \sqrt{a^{-2}d\sigma^2} = a^{-1}d\sigma\,. \tag{0.104}$$

Integrating $d\tau = a^{-1}d\sigma$ with the boundary condition $\tau = 0$ when $\sigma = 0$ gives $\tau = \sigma/a$, so our worldline is parameterized in terms of proper time as

$$t = a^{-1}\sinh(a\tau)\,, \qquad x = a^{-1}\cosh(a\tau)\,. \tag{0.105}$$

4-velocity & 4-acceleration

In Newtonian physics we use $\mathbf{v} = d\mathbf{x}/dt$ and $\mathbf{a} = d\mathbf{v}/dt$ all the time. In analogy, we define the 4-velocity

$$u^\alpha = \frac{dx^\alpha}{d\tau} \tag{0.106}$$

This is a 4-vector since the numerator is a 4-vector and the denominator is an invariant (proper time):

$$dx^{\alpha'} = \Lambda^{\alpha'}_{\beta} dx^\beta \implies u^{\alpha'} = \Lambda^{\alpha'}_{\beta}\frac{dx^\beta}{d\tau} = \Lambda^{\alpha'}_{\beta} u^\beta\,. \tag{0.107}$$

However, since $d\tau^2 = -dx^\alpha dx^\beta \eta_{\alpha\beta}$, u^α only have 3 independent components due to an overall constraint relating the 4 components:

$$u^2 = u^\alpha u^\beta \eta_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} \eta_{\alpha\beta} = -\frac{d\tau^2}{d\tau^2} = -1\,. \tag{0.108}$$

If we use $d\tau = \sqrt{1-v^2}dt = \gamma^{-1}dt$ (see e.g. (0.35)) for an inertial frame with time t, we can write the 4-velocity in terms of the Newtonian 3-velocity:

$$u^\alpha = \gamma \frac{dx^\alpha}{dt} = (\gamma, \gamma \mathbf{v}) = \left(\frac{1}{\sqrt{1-v^2}}, \frac{\mathbf{v}}{\sqrt{1-v^2}} \right), \quad \text{inertial observer with velocity } \mathbf{v}. \tag{0.109}$$

In the case of the hyperbolic worldline, using (0.103) - (0.104) we have

$$u^t = \frac{dt}{d\tau} = \cosh(a\tau)\,, \qquad u^x = \frac{dx}{d\tau} = \sinh(a\tau)\,, \qquad u^y = u^z = 0\,. \tag{0.110}$$

Notice that

$$v^i \equiv \frac{dx^i}{dt} = \frac{dx^i/d\tau}{dt/d\tau} = \frac{u^i}{u^t} = (\tanh(a\tau), 0, 0) \implies v = \sqrt{v_i v^i} = \tanh(\alpha\tau) \implies \theta = a\tau\,. \tag{0.111}$$

where we recall the boost parameter $v = \tanh\theta$. So it seems the boost keeps increasing at a steady rate: a uniform boost per unit proper time. This is a uniform acceleration! The fact that there is a positive acceleration at all is clear from the plot of the hyperbolic worldline above: as we go up the hyperbolic worldline, the tangent to the hyperbola goes from vertical to 45°.

To make this rigorous, we observe that the worldline is a constant proper distance from the origin

$$\Delta s^2 = x^2 - t^2 = a^{-2}\cosh^2\sigma - a^{-2}\sinh^2\sigma = a^{-2}\,. \tag{0.112}$$

Proper distance is invariant under Lorentz boosts, so the hyperbolic curve is invariant! Points along it can be shuffled around, however. For example, in the above frame the rocket is at rest at $\tau = t = 0$, $u^\alpha = (1,0,0,0)$, but we can go to another frame where $\tau = t+1$ (since $\theta = a\tau$ and hyperbolic angles add under boosts, we boost $\theta \to \theta + a \implies \tau \to \tau + 1$). So the *physics experienced on the rocket* must be the same at $\tau = 0$ and $\tau = 1$. In other words, the acceleration must be the same at all times! Notice that unlike in Newtonian physics the velocity doesn't increase without bound: $a = \mathit{const.}$ but $v = c\tanh(a\tau) < c$.

——————————— END LEC 4 ———————————

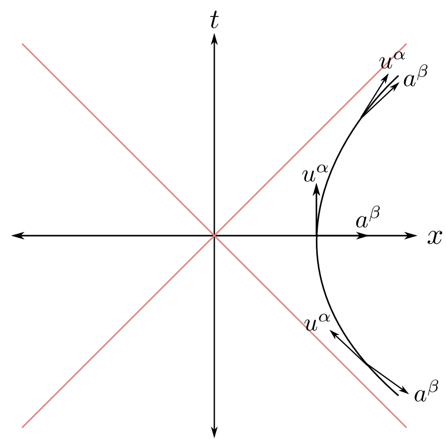
We define the 4-acceleration as

$$a^\beta = \frac{du^\beta}{d\tau}\,. \tag{0.113}$$

Like the 4-velocity, this is a 4-vector since the numerator is a 4-vector and the denominator is an invariant. For the hyperbolic worldline we have

$$a^\beta = (a\sinh(a\tau), a\cosh(a\tau), 0, 0)\,. \tag{0.114}$$

The 4-velocity and 4-acceleration are plotted on the worldline below:



Notice that u^α and a^β are perpendicular at $\tau = 0$. This must remain true at all times.[7] But this can be checked by calculating $u^\alpha a^\beta \eta_{\alpha\beta} = 0$. This orthogonality is in fact true for *any* worldline:

$$u^2 = -1 \implies \frac{d}{d\tau}(u^\alpha u^\beta \eta_{\alpha\beta}) = a^\alpha u^\beta \eta_{\alpha\beta} + u^\alpha a^\beta \eta_{\alpha\beta} = 0 \implies \mathbf{u}\cdot\mathbf{a} = u^\alpha a^\beta \eta_{\alpha\beta} = 0\,. \tag{0.115}$$

This orthogonality also says there are only three independent components of a^β. The **proper acceleration** is defined as the magnitude of a^α:

$$a = \sqrt{\mathbf{a}\cdot\mathbf{a}} = \sqrt{a^\alpha a^\beta \eta_{\alpha\beta}}\,. \tag{0.116}$$

Computing this for our hyperbolic worldline we find a uniform proper acceleration equal to a, which is why we used this variable!

4-momentum & 4-force

If a worldline is instantaneously at rest, i.e. $v = 0$ with respect to some fixed frame, then we have $\tau = t \implies d\tau = dt$ momentarily and therefore

$$u^\alpha = (1, 0, 0, 0)\,, \qquad a^\beta = \frac{du^\beta}{d\tau} = \frac{du^\beta}{dt} = \left(0, \frac{d\mathbf{v}}{dt}\right), \tag{0.117}$$

[7]The vectors u^α and a^β don't look perpendicular based on our Euclidean notions away from $t = 0$. But to see that they are perpendicular, recall that $\mathbf{u}\cdot\mathbf{a} = u^\alpha a^\beta \eta_{\alpha\beta} = -u^0 a^0 + u^1 a^1 + u^2 a^2 + u^3 a^3$. So it is OK for u^α and a^β to both have components in, say, the x direction as drawn above, as long as they have an "equal" amount in the t direction to cancel it off with the minus sign in the Minkowski dot product.

where $a^0 = 0$ since $\mathbf{u} \cdot \mathbf{a} = u^\beta a_\beta = 0$. So we see that in this frame the proper acceleration reduces to the ordinary acceleration! This type of frame is called a "comoving inertial frame" – it is the constant-velocity frame in which the observer at a certain time is instantaneously at rest in that frame. If the observer is accelerating then the comoving inertial frame is a function of time. The reduction to the ordinary acceleration makes sense: the observer's velocity is infinitesimally small, so relativistic corrections should vanish.

With a 4-acceleration in hand we can define a 4-force

$$f^\alpha = m a^\alpha \tag{0.118}$$

The definitions from before tell us $\mathbf{f} \cdot \mathbf{u} = f^\beta u_\beta = 0$, so the 4-force must depend on u. We write

$$f^\alpha = m \frac{du^\alpha}{d\tau} = \frac{dp^\alpha}{d\tau}\,, \tag{0.119}$$

where

$$p^\alpha = m u^\alpha = (E, \mathbf{p}) \tag{0.120}$$

is the 4-momentum or energy-momentum 4-vector. Since $p^\alpha = m u^\alpha = (\gamma m, \gamma m \mathbf{v})$ we can expand for non-relativistic speeds $v \ll 1$ to get

$$p^0 = \gamma m = \frac{m}{\sqrt{1-v^2}} \approx m + \frac{1}{2} m v^2 + \cdots\,, \qquad \mathbf{p} = \gamma m \mathbf{v} \approx m \mathbf{v} + \cdots\,. \tag{0.121}$$

This is why we call p^0 an energy and $\mathbf{p}$ a 3-momentum. Restoring physical units gives

$$E = mc^2 + \frac{1}{2} m v^2 + \cdots\,. \tag{0.122}$$

We see Einstein's famous equation $E = mc^2$, the energy contained in any massive object even if it is at rest! Notice also that $\gamma \to \infty$ as $v \to c$, which means $E = p^0 = m\gamma \to \infty$ and $\mathbf{p} = m\gamma\mathbf{v} \to \infty$, so any massive object requires infinite energy or momentum to approach the speed of light (so $v > c$ is not possible). Furthermore, $u^2 = -1$ gives

$$-m^2 = -E^2 + |\mathbf{p}|^2 \implies E^2 = m^2 + |\mathbf{p}|^2 \tag{0.123}$$

where $\mathbf{p}$ is the relativistic 3-momentum $\gamma m \mathbf{v}$.

Let's return to meaning of 4-force. In Newtonian mechanics we have the force as the time-derivative of momentum. We define similarly

$$f^\alpha = (f^0, \mathbf{f}) = \left(f^0, \frac{d\mathbf{p}}{d\tau}\right) = \left(f^0, \frac{d\mathbf{p}}{dt}\frac{dt}{d\tau}\right) = \left(f^0, \gamma \frac{d\mathbf{p}}{dt}\right) = (f^0, \gamma \mathbf{F})\,. \tag{0.124}$$

where $\mathbf{p}$ is the relativistic 3-momentum $\gamma m \mathbf{v}$ and our initial derivative was with respect to proper time. Don't forget that 3-vectors like $\mathbf{p}$ and $\mathbf{F}$ are the *relativistic* 3-vectors, so have corrections by factors of γ compared to the Newtonian momentum and force. Using

$a^\beta u_\beta = 0 \implies f^\beta u_\beta = 0$ gives $-f^0\gamma + \gamma^2 \mathbf{F}\cdot\mathbf{v} = 0 \implies f^0 = \gamma\mathbf{F}\cdot\mathbf{v}$. We can therefore see the work-energy relationship from the time component of $f^\alpha = dp^\alpha/d\tau$:

$$\gamma\mathbf{F}\cdot\mathbf{v} = \frac{dE}{d\tau} \implies \frac{dE}{dt} = \mathbf{F}\cdot\mathbf{v}\,. \tag{0.125}$$

We will find energy and momentum as much more useful concepts than force in this (and probably many of your subsequent) course(s)!

Skipping sections 5.4 & 5.5, please read them on your own.

Non-inertial observers & orthonormal frames

A central assumption is that the physics of non-inertial (i.e. accelerating) observers can be obtained by integrating up inertial observers over infinitesimal periods (we saw this in the twin paradox!). So all measurements on worldlines can be phrased in terms of u^α without having to bring in a^α. We define an orthonormal basis for the observer $\mathbf{e}_{\hat\alpha}$, with

$$\mathbf{e}_{\hat 0} = \mathbf{u}_{\rm obs}\,, \qquad \mathbf{e}_{\hat\alpha}\cdot\mathbf{e}_{\hat\beta} = \eta_{\hat\alpha\hat\beta}\,. \tag{0.126}$$

Think of the spatial basis vectors as arrows painted on the walls of the observer's laboratory. The hats on the subscripts are simply to denote that the frame is orthonormal. Physical quantities for observers in this frame are the components of 4-vectors *in this orthonormal basis*. For example, the energy and momentum an observer with orthonormal basis $\mathbf{e}_{\hat\alpha}$ measures is given by the coefficients in the expansion[8]

$$\mathbf{p} = p^{\hat\alpha}\mathbf{e}_{\hat\alpha} = E(\tau)\mathbf{e}_{\hat 0}(\tau) + p^{\hat 1}(\tau)\mathbf{e}_{\hat 1} + p^{\hat 2}(\tau)\mathbf{e}_{\hat 2} + p^{\hat 3}(\tau)\mathbf{e}_{\hat 3} \tag{0.127}$$

Whenever we expand a vector in some basis, we pick off components of the vector in that basis by taking a dot product with the basis elements:

$$p^{\hat 0} = -\mathbf{p}\cdot\mathbf{e}_{\hat 0}\,, \qquad p^{\hat 1} = \mathbf{p}\cdot\mathbf{e}_{\hat 1}\,, \qquad p^{\hat 2} = \mathbf{p}\cdot\mathbf{e}_{\hat 2}\,, \qquad p^{\hat 3} = \mathbf{p}\cdot\mathbf{e}_{\hat 3}\,. \tag{0.128}$$

In particular, $E = -\mathbf{p}\cdot\mathbf{u}_{\rm obs}$.

Example 17: Let's quickly show that the better way to write (0.128) is

$$p_{\hat\alpha} = \eta_{\hat\alpha\hat\beta}p^{\hat\beta} = \mathbf{p}\cdot\mathbf{e}_{\hat\alpha}\,. \tag{0.129}$$

This equation has the indices balanced on each side as we have come to expect. Notice that $p_{\hat\alpha} = \eta_{\hat\alpha\hat\beta}p^{\hat\beta} \implies p_{\hat\imath} = \delta_{\hat\imath\hat\jmath}p^{\hat\jmath} \implies p_{\hat\imath} = p^{\hat\imath}$ for $\hat\imath = \hat 1, \hat 2, \hat 3$. So the latter three equations in (0.128) are correct, and are now written covariantly. For the first equation, we simply see that $p_{\hat 0} = \eta_{\hat 0\hat 0}p^{\hat 0} = -p^{\hat 0}$, which explains the relative minus sign between (0.128) and (0.129) for $\hat\alpha = \hat 0$.

[8] I may sometimes use $\hat 0, \hat 1, \hat 2, \hat 3$ interchangeably with $\hat t, \hat x, \hat y, \hat z$ and $\hat t, \hat\imath, \hat\jmath, \hat k$.

Example 18: Let's return to our uniformly accelerated observer and write down their basis:

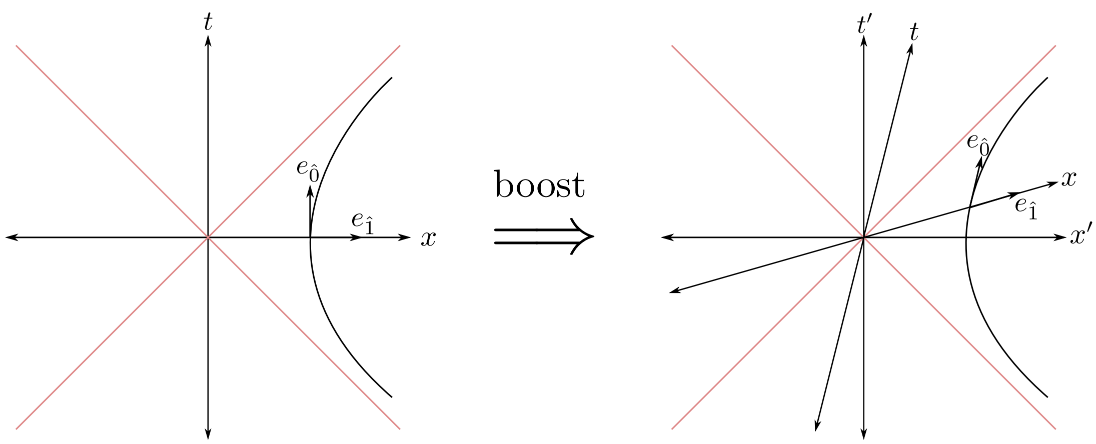
$$\mathbf{e}_{\hat{0}} = \mathbf{u}_{\text{obs}} = (\cosh(a\tau), \sinh(a\tau), 0, 0)\,, \qquad \mathbf{e}_{\hat{1}} = (\sinh(a\tau), \cosh(a\tau), 0, 0)\,, \tag{0.130}$$

$$\mathbf{e}_{\hat{2}} = (0, 0, 1, 0)\,, \qquad \mathbf{e}_{\hat{3}} = (0, 0, 0, 1)\,. \tag{0.131}$$

where $\mathbf{e}_{\hat{1}}$ was chosen by $\mathbf{e}_{\hat{1}} \cdot \mathbf{e}_{\hat{0}} = 0$, $\mathbf{e}_{\hat{1}} \cdot \mathbf{e}_{\hat{1}} = 1$, and 0's in the third and fourth entries (this fixes the vector up to an overall sign).

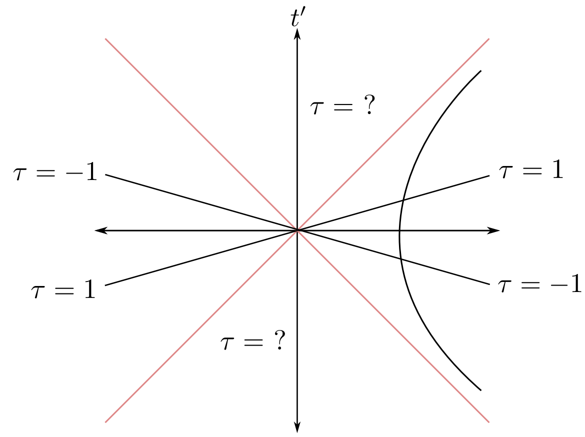
Accelerating observers and the Rindler wedge

We recall that since $\theta = a\tau$ for the hyperbolic angle θ, shifting τ along the worldline $\tau \to \tau + \tau_1$ is achieved by a boost $\theta \to \theta + a\tau_1$. In particular this means $\mathbf{e}_{\hat{1}}$ always points away from the origin.



- This makes sense since $\mathbf{e}_{\hat{1}} \cdot \mathbf{u} = 0$ means $\mathbf{e}_{\hat{1}}$ defines the line of simultaneity for the inertial frame instantaneously at rest with respect to our accelerated observer.
- This works for stuff close to the worldline but not far away, e.g. the origin does not have a well-defined value of τ!
- In fact, in the left wedge of the spacetime diagram time runs backwards, and the lines of simultaneity don't even enter the top and bottom wedges.

The conclusion is that extending locally inertial frames far from the worldline gives nonsense in general, see the diagram below:



We can add different accelerating observers with different values of a; they correspond to hyperbolas with the same asymptotics and the same lines of simultaneity. Since each line of simultaneity has a given value of θ, and $\theta = a\tau$ for τ the time along the hyperbola and a^{-1} the distance of the hyperbola from the origin, we have

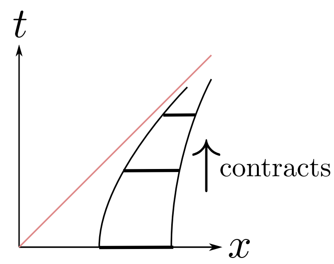
$$a_1\tau_1 = \theta = a_2\tau_2 \implies \tau_1 = \frac{a_2}{a_1}\tau_2 \tag{0.132}$$

for two observers at distance a_1^{-1} and a_2^{-1} from the origin. At $t = \tau_1 = \tau_2 = 0$, say the distance between them is L. By the boost symmetry (i.e. the physics should be the same when we boost $\theta \to \theta + \theta_1$), this length remains the same along any other line of simultaneity! So the distance between the two observers remains constant.

Pictorially we can imagine several observers spaced out along a rocket accelerating upward. To maintain a fixed proper distance between them, the observers have different accelerations. Whoever is at the top has the smallest acceleration. This is because the acceleration is given by a as we saw in (0.116), while the distance is given by a^{-1}. (This type of "rigid" body is sometimes referred to as Born rigidity.) In fact, if we have a distance from the origin to the bottom of the rocket as a_{bottom}^{-1}, and the length of the rocket is L, then the distance from the origin to the top of the rocket is

$$a_{\text{top}}^{-1} = a_{\text{bottom}}^{-1} + L \implies a_{\text{top}} = \frac{a_{\text{bottom}}}{1 + a_{\text{bottom}}L} < a_{\text{bottom}} \, . \tag{0.133}$$

So we see explicitly that the acceleration at the top is smaller than the acceleration at the bottom. This seems weird! Non-relativistic intuition says the distance between two observers should remain the same if their accelerations are the same. But let's analyze the situation from an inertial frame watching the rocket go by. The rocket becomes more and more length-contracted:



If the accelerations were the same at top and bottom the length would remain constant (as viewed from an inertial observer).

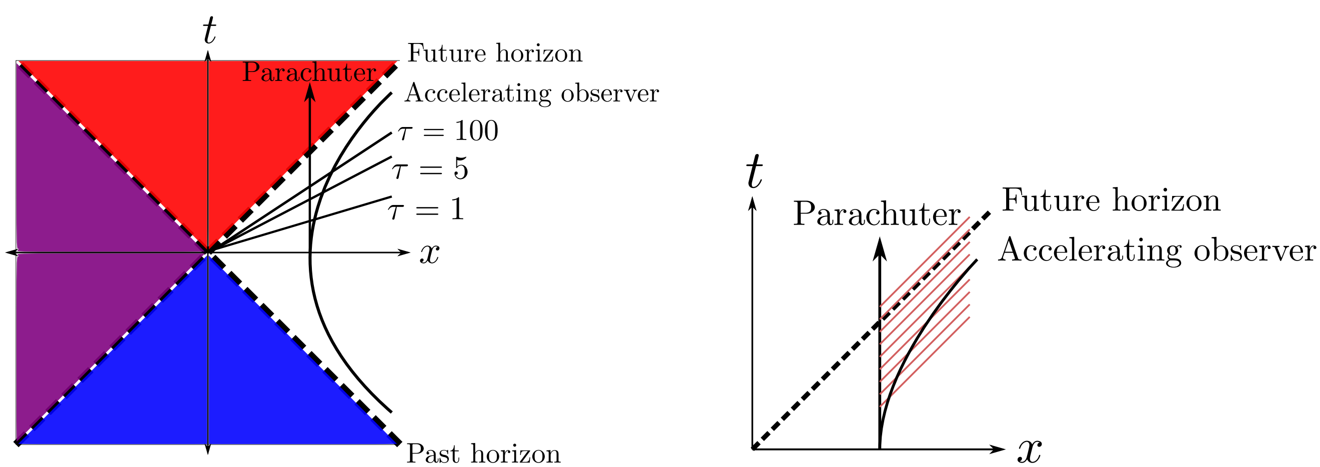
If all observers synchronize such that $\tau = 0$ when $t = 0$, then we also see that clocks run faster at the top of the rocket:

$$a_{\text{top}}\tau_{\text{top}} = a_{\text{bottom}}\tau_{\text{bottom}} \implies \tau_{\text{top}} = \tau_{\text{bottom}}(1 + a_{\text{bottom}}L) > \tau_{\text{bottom}}\,, \tag{0.134}$$

where we used (0.133). We can intuitively understand this result by having an observer at the top of the rocket emit signals to an observer at the bottom. The observer at the bottom will receive them at a faster rate than they are emitted at the top, since the observer at the bottom is accelerating and catching up with the signals more and more quickly. So a given interval of time at the bottom of the rocket corresponds to a longer interval of time at the top; the observer on the top ages more!

A physical way to understand why the coordinate system cannot be extended past the origin is that the acceleration diverges there.

Notice the "horizons" that exist for the accelerated observers:



They cannot receive signals from beyond the future horizon, and cannot send signals to beyond the past horizon. The purple region above is beyond both horizons, the red is beyond just the future horizon, and the blue is beyond just the past horizon (red and blue make purple). The horizons are denoted by the dashed lines. They are the edges of the rocket's spacetime. If somebody parachutes out of the rocket, however, they can go past the rocket's future horizon. But the people in the rocket will never see it happen, since their lines of simultaneity never coincide with the future horizon at any finite proper time (this is denoted in the left diagram by τ getting large but not passing the future horizon)! Moreover, they will see the parachuting astronaut become redder and dimmer as they approach the future horizon. As we see in the image on the right, if the parachuting astronaut releases a fixed number of photons per unit time in their frame (denoted by 45° pink lines), then fewer and fewer photons will reach the accelerating rocket per unit rocket time, so the image will be dimmer. Using the equations for Doppler shift from Chapter 5.5 of Hartle, we also see that the light becomes redder and redder (in the instantaneous rest frame of the rocket, the photons are emitted in the opposite direction of the motion of the parachuting astronaut).

———————————— END LEC 5 ————————————

Why general relativity?

Seeing that time and space are very flexible in special relativity, and that the maximum speed is that of light, we see that Newton's force law

$$F = G\frac{m_1 m_2}{r^2} \tag{0.135}$$

has to be modified. This is because it acts instantaneously, and refers to some privileged distance r. If the sun disappeared, the Earth would instantaneously move off its orbit, whereas SR says it should take 8 minutes before that happens.

Let's say we tried to build an E&M-like theory for it (since E&M is consistent with SR). The equations there are

$$\nabla \cdot \mathbf{E} = +\rho\,, \qquad \nabla \times \mathbf{B} = +\mathbf{J}\,, \qquad \nabla \cdot \mathbf{B} = 0\,, \qquad \nabla \times \mathbf{E} = 0 \tag{0.136}$$

with the Lorentz force law given by

$$\mathbf{F} = +q\mathbf{E} + q\mathbf{v} \times \mathbf{B}\,. \tag{0.137}$$

For gravity we want like charges to attract, so we can either change the sign of $\rho, \mathbf{J}$ in Maxwell's equations, or change the sign of q in the Lorentz force law. It doesn't matter which we do, so let's do the former. Now assume a positive energy wave arrives at a charge. In E&M, the charge is excited and oscillates up and down, emitting an E&M wave that is 180° out of phase with the original wave. This attenuates the incoming radiation. This is how polarized sunglasses work. But if we flip the sign of Maxwell's equations then it will be perfectly in phase. This means the wave **gains** more and more energy. So the theory is unstable.

The equivalence principle

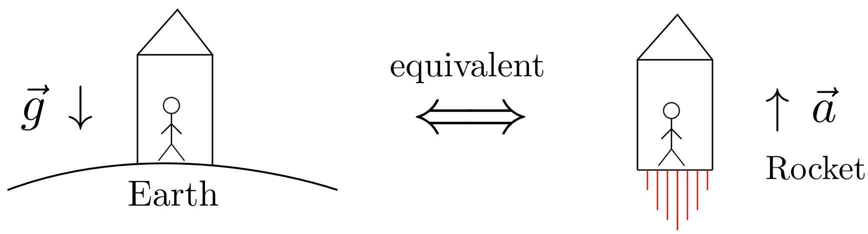
In pondering how to make gravity consistent with SR, Einstein was struck by the equality of gravitational and inertial mass, $m_I = m_G$, where these quantities appear as

$$\mathbf{F} = m_I \mathbf{a}\,, \qquad \mathbf{F} = -G\frac{M_G m_G}{r^2}\hat{\mathbf{r}}\,. \tag{0.138}$$

These two masses are logically distinct (e.g. even electromagnetic charges accelerate according to $\mathbf{F} = m_I\mathbf{a}$), but they seem to be equal (by now verified to 1 part in 10^{12}). This is sometimes called the weak equivalence principle.

This equivalence is why all objects fall the same way under gravity: defining $\mathbf{g} = -\frac{GM_G}{r^2}\hat{\mathbf{r}}$, we have $\mathbf{a} = \mathbf{g}$ by equating the two forces above and cancelling $m_I = m_G$. This means that we can **exactly cancel** the effects of a gravitational field by moving into an accelerating reference frame which also falls at $\mathbf{a}$. In other words:

$$\boxed{\text{Effects of uniform } \mathbf{g} \text{ equivalent to zero } \mathbf{g} \text{ but with acceleration } \mathbf{a} = -\mathbf{g}} \tag{0.139}$$



The "happiest thought" of Einstein's life was the realization that a freely falling observer does not experience gravity, i.e. if she lets go of an object then it will remain at rest in her vicinity. Einstein proposed the following as a fundamental principle of physics, sometimes called the strong equivalence principle:

No local experiment can distinguish between having a gravitational field

and being in an accelerated reference frame.

Local means the experiment is performed over short times/spatial separation so we can ignore inhomogeneities in time/space. For example the field around the Earth points in different directions depending on where you are, so no single choice of $\mathbf{a}$ can eliminate this; locally, however, you can go into a freely falling frame.

Clocks in a gravitational field

We can use the equivalence principle to translate our results from accelerating frames to results in gravitational fields. Previously we found (0.134), which said $\tau_{\text{top}}/\tau_{\text{bottom}} = 1 + a_{\text{bottom}}L$. Let's make L very small, so that $a_{\text{top}} = a_{\text{bottom}} + O(L)$ (you can get this by solving $a_{\text{top}}^{-1} = a_{\text{bottom}}^{-1} + L \implies a_{\text{top}} = a_{\text{bottom}}/(1 + a_{\text{bottom}}L)$ and expanding for small L). Then if we ignore terms of order L^2, we can write $\tau_{\text{top}}/\tau_{\text{bottom}} = 1 + a_{\text{bottom}}L = 1 + a_{\text{top}}L + O(L^2)$. The requirement of *local* experiments in the equivalence principle means we can ignore terms of order L^2. So in a gravitational field of strength g, higher clocks run faster according to the formula

$$\frac{\Delta\tau_{\text{top}}}{\Delta\tau_{\text{bottom}}} = 1 + gh + O(h^2)\,, \tag{0.140}$$

where h is the height above, say, ground level on the Earth's surface. For this formula to be accurate h needs to be small (compared to c^2/g).

Similarly, if the bottom of the rocket sends signals at frequency $\omega(0)$, then the frequency at the top will be

$$\omega(h) = \omega(0)[1 - gh + O(h^2)]\,, \tag{0.141}$$

which is obtained by $\omega(0)\Delta\tau(0) = \#$ of signals $= \omega(h)\Delta\tau(h)$.

We can integrate this effect over long distances to get the result for an inhomogeneous gravitational field:

$$\frac{\Delta\tau(h_2)}{\Delta\tau(h_1)} = 1 + \int_{h_1}^{h_2} gdh + O(h^2) = 1 + (\Phi(h_2) - \Phi(h_1)) + O(\Phi^2)\,, \tag{0.142}$$

where we used $\int gdh = -\int \mathbf{g} \cdot d\mathbf{h}$ (since $\mathbf{g}$ points down and $d\mathbf{h}$ points up given that we integrate from a lower point h_1 to a higher point h_2), $\mathbf{g}(\mathbf{x}) = -\nabla\Phi(\mathbf{x})$, and the fundamental theorem of calculus. So the rate of clocks is determined by the gravitational potential. The term $O(\Phi^2)$ is a term of order g^2h^2/c^4, and since $g^2/c^4 \sim 10^{-15}$ this term is usually very small and can be ignored.

This effect has been verified in several ways:

- Pound-Rebke experiment: measured frequency shift of photons going up a tower at Harvard
- GPS clocks
- Atomic clocks at the National Bureau of Standards in Denver run differently than those in DC
- Shift in frequency of emitted light from white dwarf stars (gravitational potential difference between surface of star and us)

Another effect we can use the equivalence principle to see is the bending of light due to gravity. To see this, imagine light entering a slit on the left of a rocket accelerating upward and exiting a slit on the right (see Figure 6.5 of Hartle). In the rest frame of the rocket, the right slit must be further down than the left slit. In the frame of the Earth, the light ray is going straight across. But in the frame of the rocket, the light ray is bending downward to go past the right slit. Using the equivalence principle, this must mean that a gravitational field bends light! This effect is responsible for "gravitational lensing," one of the best ways we have of mapping out the distribution of dark matter in our universe. Here is an image of lensing due to clusters of galaxies: it's quite dramatic!

Newtonian gravity in terms of curved spacetime

How can we incorporate the fact that clocks at different locations run at different rates into a spacetime line element? Let's define coordinate time t as the proper time elapsed for vanishing gravitational potential $\Phi = 0$. The time elapsed at some other point should be

$$\Delta\tau(x) = (1 + \Phi(x))\Delta t + O(\Phi^2) \tag{0.143}$$

as we saw before. This leads to $d\tau^2 = (1 + 2\Phi(x))dt^2 + O(\Phi^2)$. For the line element on the space terms there are many choices that will be consistent with Newtonian gravity, but we will choose the one predicted by general relativity (see Chapter 22 of Hartle). Altogether this gives

$$ds^2 = -\left(1 + \frac{2\Phi(x^i)}{c^2}\right)(cdt)^2 + \left(1 - \frac{2\Phi(x^i)}{c^2}\right)(dx^i dx_i) \tag{0.144}$$

This is a good approximation to the full machinery of general relativity when the sources are *static* and *weak*.

We can use this line element to extract the equation of motion for a particle. Recall from Section 5.4 of Hartle that a free particle in flat spacetime extremizes the proper time between any two points. We want to apply that principle to the line element (0.144). The proper time between points A and B is given as

$$\tau_{AB} = \int_A^B d\tau = \int_A^B \sqrt{-\frac{ds^2}{c^2}} = \int_A^B \sqrt{\left(1+\frac{2\Phi}{c^2}\right) - \frac{1}{c^2}\left(1-\frac{2\Phi}{c^2}\right)\frac{dx^i dx_i}{dt^2}} \tag{0.145}$$

Expanding to first order in $1/c^2$ gives

$$\tau_{AB} \approx \int_A^B dt \left[1 - \frac{1}{c^2}\left(\frac{v^2}{2} - \Phi\right)\right] \tag{0.146}$$

where v is the nonrelativistic velocity $\sqrt{dx^i dx_i/dt^2}$. Treating the integrand as a Lagrangian

$$\tau_{AB} = \int_A^B dt\, L\left(\frac{dx^i}{dt}, x^i\right) \tag{0.147}$$

means we can extremize it via the Euler-Lagrange equations:

$$\partial_t \frac{\partial L}{\partial\left(\frac{dx^i}{dt}\right)} = \frac{\partial L}{\partial x^i} \implies \frac{d^2\mathbf{x}}{dt^2} = -\nabla\Phi\,. \tag{0.148}$$

Multiplying both sides by the mass of the particle m gives $\mathbf{F} = m\mathbf{a}$. So Newton's gravitational law is reproduced from a curved line element! If we set $\Phi = 0$, we recover the geometry of Minkowski space, and the Euler-Lagrange equations simplify to $d^2\mathbf{x}/dt^2 = 0$. This is expected: in the absence of a gravitational field the acceleration vanishes.

Example 19: Let's consider the case $\Phi = 0$ further, which reduces to the ordinary Minkowski geometry of special relativity. For simplicity we will ignore the y, z directions. The action is

$$S = \int d\tau = \int L dt = \int \sqrt{1 - v^2/c^2}\, dt\,. \tag{0.149}$$

The Euler-Lagrange equations are unchanged if we multiply by an overall constant, so let's multiply by $-mc^2$, where m is the mass of the particle. Does this action have anything to do with the one you're familiar with from Newtonian physics? Let's see:

$$S = -mc^2 \int d\tau = -mc^2 \int \sqrt{1 - v^2/c^2}\, dt \approx \int \left(\frac{1}{2}mv^2 - mc^2\right) dt\,. \tag{0.150}$$

where we expanded for small v/c. Notice the integrand is simply $L = T - V$ with the potential energy given by mc^2! This is the correct relativistic action for a massive point particle, as can be checked by computing the canonical momentum and energy:

$$p_x = \frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial v} = \frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}\,, \quad E = p_x\dot{x} - L = \frac{mv^2}{\sqrt{1-\frac{v^2}{c^2}}} + mc^2\sqrt{1-\frac{v^2}{c^2}} = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\,. \tag{0.151}$$

These are precisely the components of the 4-momentum we saw in (0.121)!

Notice, however, that the gravitational potential is still determined from Poisson's equation $\nabla^2\Phi = 4\pi G\mu$ This equation takes place at one time, so influences propagate instantaneously ($v = \infty > c$). To see this more directly, consider the Green function solution of Poisson's equation $\Phi(\mathbf{x}) = -\int d^3x'\, G\mu(\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|$, which is an integral over all of space at an instant in time. We have therefore not fully incorporated relativity into Newton's laws.

Part of Einstein's great insight was to realize the force of gravity in purely geometric terms, encapsulating its effects in a general line element

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta\,. \tag{0.152}$$

We will eventually see that this line element is determined by some **relativistic** dynamical equation, called Einstein's equation. This dispenses with the non-relativistic Poisson equation.

——————— END LEC 6———————

Curved spacetime

The general line element for a gravitational field in general relativity describes a *curved* spacetime; it can't just be Minkowksi space since observers in this spacetime would see each other moving with constant velocity, but observers accelerate in gravitational fields. This "bending" of worldlines away from straight lines (i.e. $d^2x/dt^2 \neq 0$) is precisely analogous to the bending of lines on e.g. the surface of a sphere. These lines are called "geodesics" and they extremize the distance between two points in space or spacetime.

Given a line element

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta\,, \tag{0.153}$$

we can represent it in arbitrary coordinates $y^\alpha = y^\alpha(x)$, which gives

$$ds^2 = \tilde{g}_{\alpha\beta}dy^\alpha dy^\beta \tag{0.154}$$

for some $\tilde{g}_{\alpha\beta}$. So, for example, flat Minkowski space $ds^2 = \eta_{\alpha\beta}dx^\alpha dx^\beta = -dt^2 + dx^2 + dy^2 + dz^2$ can be disguised in many ways simply by performing coordinate transformations. One example is to use the coordinates along a hyperbolic worldline (0.105):

$$x = \rho\cosh\theta\,, \qquad t = \rho\sinh\theta \qquad \Longrightarrow \qquad ds^2 = -\rho^2 d\theta^2 + d\rho^2\,. \tag{0.155}$$

This is just Minkowski space again, but in different coordinates! (This is like the Lorentzian version of polar coordinates.) But there are many line elements $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$ that do **not** represent flat spacetime. In general, $g_{\alpha\beta}$ is called the **spacetime metric**. It is a symmetric matrix, $g_{\alpha\beta} = g_{\beta\alpha}$, with one negative eigenvalue (corresponding to time) and three positive eigenvalues (corresponding to space). It is symmetric because it measures lengths, and $dx^\alpha dx^\beta = dx^\beta dx^\alpha$.

Say we made another coordinate transformation in (0.155):

$$r = \rho^2 \qquad \Longrightarrow \qquad ds^2 = -r d\theta^2 + \frac{dr^2}{4r} \qquad \Longrightarrow \qquad g_{\alpha\beta} = \begin{pmatrix} -r & 0 \\ 0 & \frac{1}{4r} \end{pmatrix} . \tag{0.156}$$

We see that the metric **diverges** at $r = 0$! What is going on? Notice $r = 0 \implies \rho^2 = x^2 - t^2 = 0$. These are just the lightcones from the origin, which correspond to the horizons of our accelerating observer! We know nothing goes wrong here in the spacetime (recall our parachuting observer who can cross these horizons). Instead, our coordinates are just problematic: t and x are not smooth functions of ρ, since e.g. $x = \sqrt{r}\cosh\theta$. This is known as a **coordinate singularity**. Generally, if a metric diverges somewhere, we need to be careful to distinguish whether it is a real, physical singularity (we will get to this soon in the context of black holes), or just a coordinate singularity.

Curved space examples

Since we will spend much of the rest of the quarter extracting physics from metrics describing curved spacetimes, it will help to warm up with the case of curved space. This means a metric with all positive eigenvalues, i.e. no time component. In flat space, we have the familiar rules of geometry, circles have circumference $C = 2\pi R$ and area $A = \pi R^2$. If these relations are violated, space must be curved!

Let's introduce some mathematical notation. We will refer to lengths, areas, volumes, etc. as n-volumes. A 1-volume is a length, a 2-volume is an area, a 3-volume is what you usually call a volume, and an n-volume for $n > 3$ is sometimes called a hypervolume. So the circumference $2\pi R$ of a circle is the 1-volume of the circle or 1-sphere. The area of the circle πR^2 is really the 2-volume of the disk or 2-ball (called that since it is a 2-dimensional version of an ordinary 3d ball), which is the *interior* of the circle. The area of a sphere $4\pi R^2$ is called the 2-volume of the 2-sphere. The volume of the sphere $\frac{4}{3}\pi R^3$ is called the 3-volume of the 3-ball, which is the interior of the 2-sphere. Generally we have n-dimensional geometries known as n-spheres ($n = 1$ is a circle, $n = 2$ is the ordinary sphere, etc.), and we can compute their n-volumes. Their interiors are $(n+1)$-dimensional manifolds known as $(n+1)$-balls, and we can compute their $(n+1)$-volumes. Sometimes the n-volume of an n-sphere is called the surface area (see e.g. https://en.wikipedia.org/wiki/N-sphere), but calling things n-volumes lets you generalize more easily to other geometries.

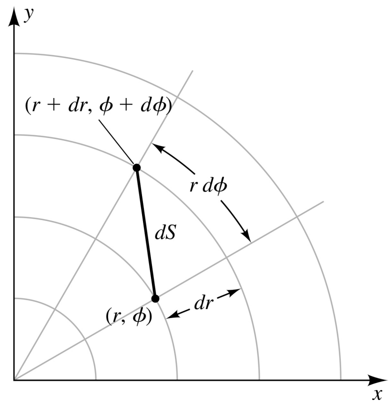
Example 20: Let's consider flat space in polar coordinates:

$$ds^2 = dr^2 + r^2 d\phi^2 \qquad \Longrightarrow \qquad g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \tag{0.157}$$

The curve $r = const.$ is a circle centered around the origin, while the curve $\phi = const.$ is a straight line measuring the distance of the circle from the origin. This means we can compute the radius R and circumference C (or 1-volume) as

$$R = \int_{\phi=const.} ds = \int_0^R dr = R\,, \qquad C = \int_{r=const.} ds = \int_0^{2\pi} R d\phi = 2\pi R\,. \tag{0.158}$$

To compute the area of the circle (or 2-volume of the disk in our new notation), notice we want to integrate up infinitesimal slivers $rdrd\phi$.



Since our metric is diagonal, this just corresponds to $\sqrt{g_{rr}g_{\phi\phi}}\,drd\phi$. More generally, we use the determinant of $g_{\alpha\beta}$, which is often just denoted g:

$$\boxed{dV = \sqrt{g}dx^1dx^2\cdots dx^d \qquad g \equiv \det(g_{\alpha\beta}) \qquad \text{(d-dimensional metric)}\,.} \tag{0.159}$$

We therefore obtain the area or 2-volume as

$$V_2 = \int dV = \int \sqrt{g}\,dx^1dx^2 = \int_0^R \int_0^{2\pi} rdrd\phi = \pi R^2 \tag{0.160}$$

Example 21: The metric on a sphere is given by

$$ds^2 = R_0^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{0.161}$$

We actually saw this in Example 3, by restricting the spherical coordinates for 3d Euclidean space to a constant-r slice. Constant-latitude circles are $\theta = \mathit{const.}$. Lines connecting the north pole to the south pole are $\phi = \mathit{const}$. We can compute the radius, circumference, and area of a constant-latitude circle as in Example 3:

$$R = \int ds = \int_0^{\theta_0} \underbrace{R_0}_{\sqrt{g_{\theta\theta}}}\,d\theta = R_0\theta_0 \tag{0.162}$$

$$C = \int ds = \int_0^{2\pi} \underbrace{R_0\sin\theta_0}_{\sqrt{g_{\phi\phi}}}\,d\phi = 2\pi R_0\sin\theta_0 < 2\pi R_0 \tag{0.163}$$

$$V_2 = \int \sqrt{g}d\theta d\phi = \int_0^{\theta_0} d\theta \int_0^{2\pi} d\phi\, R_0^2\sin\theta = 2\pi R_0^2(1-\cos\theta_0) \tag{0.164}$$

Notice that for small θ_0 we have $\cos\theta_0 \approx 1 - \frac{1}{2}\theta_0^2$ and $\sin\theta_0 \approx \theta_0$. Using $\theta = R/R_0$ gives $C \approx 2\pi R$ and $V_2 \approx \pi R^2$, recovering the flat-space formulas. As $\theta_0 \to \pi$ we get $A = 4\pi R_0^2$, the area of a 2-sphere. Again, this is phrased as the integral of an infinitesimal volume dV because it is the volume of a 3-ball, which is a fancy way to refer to the interior of a 2-sphere.

Example 22: Consider the metric

$$ds^2 = \frac{dr^2}{1 - r^2/R_0^2} + r^2 d\phi^2, \qquad r \in [0, R_0)\,, \quad \phi \sim \phi + 2\pi\,. \tag{0.165}$$

Circles correspond to $r = const.$, while straight lines to the origin $r = 0$ correspond to $\phi = const.$ We compute the radius, circumference, and area (using $r = \rho = const.$ as the location of the circle)

$$R = \int ds = \int_0^\rho \frac{dr}{\sqrt{1 - r^2/R_0^2}} = R_0 \sin^{-1}(\rho/R_0) \tag{0.166}$$

$$C = \int ds = \int_0^{2\pi} \rho d\phi = 2\pi\rho = 2\pi R_0 \sin(R/R_0) \tag{0.167}$$

$$V_2 = \int \sqrt{g} d\theta d\phi = \int_0^\rho dr \int_0^{2\pi} d\phi \frac{r}{\sqrt{1 - \frac{r^2}{R_0^2}}} = 2\pi R_0^2 \left[1 - \sqrt{1 - \frac{\rho^2}{R_0^2}}\right] = 2\pi R_0^2 \left(1 - \cos\frac{R}{R_0}\right). \tag{0.168}$$

We see that V_2 and C are the same as in the case of the 2-sphere! This is a clue that maybe this is the sphere metric in disguise. In fact it is, defining $r = R_0 \sin\theta$ gives

$$dr^2 = R_0^2 \cos^2\theta d\theta^2 = R_0^2(1 - r^2/R_0^2) d\theta^2 \implies ds^2 = R_0^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{0.169}$$

Example 23: We can use line elements to compute 3-volumes and 4-volumes in the same way. Let's calculate the 3-volume of a 3-sphere of radius R. The 3-sphere is defined by

$$x^2 + y^2 + z^2 + w^2 = R^2 \tag{0.170}$$

in flat Euclidean 4-space with metric $ds^2 = dx^2 + dy^2 + dz^2 + dw^2$. We can change to spherical coordinates

$$w = r\cos\theta\,, \qquad z = r\sin\theta\cos\chi\,, \qquad y = r\sin\theta\sin\chi\cos\phi\,, \qquad x = r\sin\theta\sin\chi\sin\phi \tag{0.171}$$

which gives metric

$$ds^2 = dr^2 + \underbrace{r^2\left[d\theta^2 + \sin^2\theta(d\chi^2 + \sin^2\chi d\phi^2)\right]}_{\text{3-sphere of radius } r} \tag{0.172}$$

The 3-volume is computed by

$$V_3 = \int \sqrt{g} d\theta d\chi d\phi = \int \sqrt{g_{\theta\theta} g_{\chi\chi} g_{\phi\phi}} d\theta d\chi d\phi = R^3 \int_0^\pi d\theta \int_0^\pi d\chi \int_0^{2\pi} d\phi \sin^2\theta \sin\chi \tag{0.173}$$

$$= 4\pi R^3 \int_0^\pi \sin^2\theta d\theta = 2\pi^2 R^2 \tag{0.174}$$

Local inertial frames

While there are no "global" inertial frames in generic curved spacetime (since observers accelerate), there are frames which act *locally* like inertial frames – these are known as freely falling frames. A restatement of the equivalence principle is the following: **Experiments in a sufficiently small freely falling laboratory, carried out over a sufficiently short time, give results that are indistinguishable from those of the same experiments in an inertial frame in empty space**. This is a restatement of the equivalence principle since it simply equates gravity with acceleration, and then ensures the spatial and temporal regions are small enough that the change in velocity due to the acceleration is negligible. In that case you may as well be in an inertial frame in flat space!

Mathematically, this means the following. Given any spacetime geometry and a point x_0, there exists a set of coordinates x^α (corresponding to a freely falling frame at x_0) in which

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = \eta_{\alpha\beta} dx^\alpha dx^\beta + O[(x - x_0)^2] \,. \qquad (0.175)$$

Recall that $\eta_{\alpha\beta}$ is the flat spacetime metric.

It is trivial to set $g_{\alpha\beta} = \eta_{\alpha\beta}$ at one point, i.e. up to $O(x - x_0)$. To see this, notice that $g_{\alpha\beta}$ can be diagonalized (since it is a symmetric matrix). In the coordinates that diagonalize the metric, say we have $ds^2 = \tilde{g}_{\alpha\beta} d\tilde{x}^\alpha d\tilde{x}^\beta$, with eigenvalues $\tilde{g}_{00} = -f_0(x_0)$, $\tilde{g}_{11} = f_1(x_0)$, $\tilde{g}_{22} = f_2(x_0)$, $\tilde{g}_{33} = f_3(x_0)$. We can then simply rescale $\tilde{x}^i \to \frac{1}{\sqrt{f_i(x_0)}} \tilde{x}^i$ so that the eigenvalues become $\{-1, 1, 1, 1\}$. This gives $g_{\alpha\beta} = \eta_{\alpha\beta} + O(x - x_0)$.

——————————— END LEC 7 ———————————

It is more nontrivial to see that we can achieve (0.175). We won't prove it, but let's do a counting argument that shows it's reasonable. Think of Taylor expanding $g_{\alpha\beta}$ around $x = x_0$. The linear term we want to vanish is $\partial_\gamma g_{\alpha\beta}|_{x=x_0}$. A 4×4 symmetric matrix has 10 independent components. The 3-index object $\partial_\gamma g_{\alpha\beta}$ has 4 times as many (10 for each value of γ which takes on 4 values), so $4 \times 10 = 40$ components. Consider a coordinate transformation $x^\alpha \to x^\alpha + f^\alpha(x^\mu)$. The metric at $x = x_0$ is modified by $\partial_\beta f^\alpha|_{x=x_0}$, which we fix to zero to keep the metric as Minkowski, $g_{\alpha\beta}|_{x_0} = \eta_{\alpha\beta}$:

$$dx^\alpha \to d(x^\alpha + f^\alpha) = dx^\alpha + \partial_\beta f^\alpha|_{x=x_0} dx^\beta = dx^\alpha \implies \partial_\beta f^\alpha|_{x=x_0} = 0 \qquad (0.176)$$

This means $f^\alpha(x^\mu) = f^\alpha(x_0) + O[(x - x_0)^2]$, i.e. the linear term vanishes. The quadratic term is given by the 3-index object $\partial_\gamma \partial_\beta f^\alpha$, which is symmetric in γ and β since partial derivatives commute. This means it has 40 components, just like $\partial_\gamma g_{\alpha\beta}$! By suitably picking those 40 components, we can make $\partial_\gamma g_{\alpha\beta}$ vanish and achieve (0.175). At next order we will have $\partial_\nu \partial_\gamma \partial_\beta f^\alpha$ ($\partial_\nu \partial_\gamma \partial_\beta$ is symmetric in all indices so has 20 components, while f^α has 4 components, so altogether there are $20 \times 4 = 80$ components), which does not have enough components to set $\partial_\nu \partial_\gamma g_{\alpha\beta}$ ($10 \times 10 = 100$ components) equal to zero.[9] So the second derivative of the metric has **coordinate-invariant information** in it, since it cannot be removed

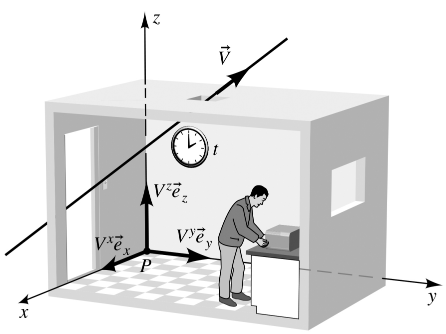
[9] A totally symmetric tensor with n indices which run from 1 to d has $\binom{d+n-1}{n}$ independent components.

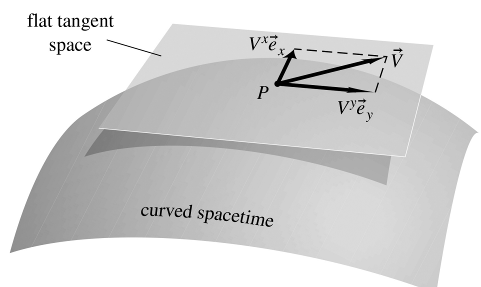
by a coordinate transformation. We will see later, when we introduce the Riemann tensor which calculates curvature, that it has precisely 20 ($= 100 - 80$) independent components in it!

Coordinates in which the metric looks like (0.175) will be referred to as a locally inertial frame at x_0.

Vectors in curved spacetime

How do we define vectors in curved spacetime? It is difficult to say they "point" along a certain direction. The trick is to work in a very small patch where the spacetime is approximately flat, and use our usual notions of vectors from flat space(time). These small vectors (which like in SR can be thought of as painted on an observer's laboratory wall) can serve as a basis, and then they can be scaled up to bigger vectors. The mathematical way to say this is that the basis vectors are defined in the *tangent space.* See Hartle Chapter 20 or Carroll's book for more mathematical details. Some ideas don't carry over: there isn't a natural way to add vectors defined at different spacetime points, since they live in *different* tangent spaces.





We can have coordinate bases, denoted $\mathbf{e}_\alpha$, or orthonormal bases, denoted $\mathbf{e}_{\hat{\alpha}}$. They are defined by

$$\mathbf{e}_\alpha(x) \cdot \mathbf{e}_\beta(x) = g_{\alpha\beta}(x)\,, \qquad \mathbf{e}_{\hat{\alpha}}(x) \cdot \mathbf{e}_{\hat{\beta}}(x) = \eta_{\hat{\alpha}\hat{\beta}}(x) \tag{0.177}$$

Notice that in flat spacetime these two are the same thing, $g_{\alpha\beta} = \eta_{\alpha\beta}$. Similarly, the coordinate basis of a locally inertial frame is an orthonormal basis, again since $g_{\alpha\beta} = \eta_{\alpha\beta}$ locally. As in special relativity, the observer's laboratory defines an orthonormal basis, with $\mathbf{e}_{\hat{0}} = \mathbf{u}_{\text{obs}}$.

Any vector $\mathbf{v}$ at a given point can be expanded in either basis defined in the tangent space at that point:

$$\mathbf{v}(x) = v^\alpha(x)\mathbf{e}_\alpha(x) = v^{\hat{\alpha}}(x)\mathbf{e}_{\hat{\alpha}}(x)\,. \tag{0.178}$$

Scalar products work the same way, although as usual they can only be taken between vectors defined *at the same point.* They can be represented in either the coordinate basis or an orthonormal basis:

$$\mathbf{v} \cdot \mathbf{w} = \eta_{\hat{\alpha}\hat{\beta}} v^{\hat{\alpha}} w^{\hat{\beta}} = g_{\alpha\beta} v^\alpha w^\beta\,. \tag{0.179}$$

When referring to physical observations, we will want to use orthonormal bases. We write the 4-momentum of a particle $\mathbf{p}$ as

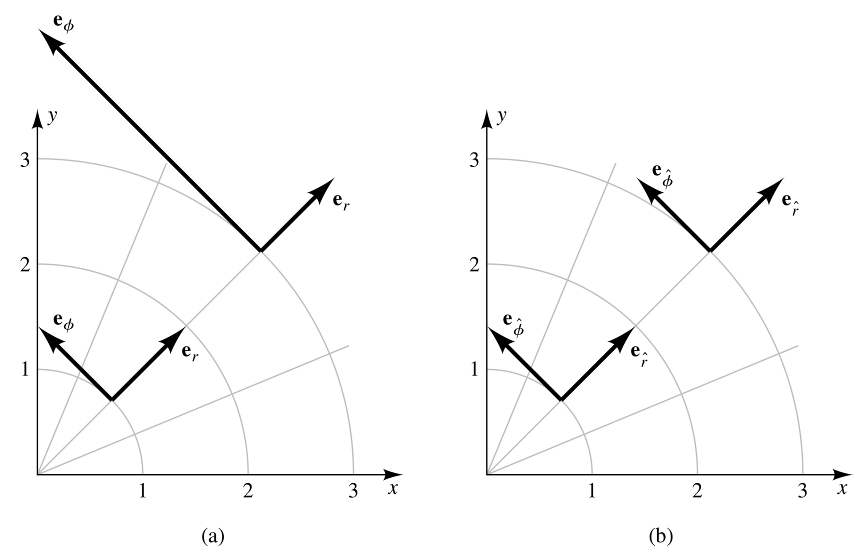
$$\mathbf{p} = p^{\hat{\alpha}}\mathbf{e}_{\hat{\alpha}}\,. \tag{0.180}$$

The *observed* energy and three-momentum are the components of $\mathbf{p}$ in this basis, i.e. $E = p^{\hat{t}}$ is the observed energy and the observed 3-momentum is given by $(p^{\hat{x}}, p^{\hat{y}}, p^{\hat{z}})$. We can extract these components, as usual, by dot products, e.g.

$$E = -\mathbf{p}\cdot\mathbf{u}_{\text{obs}} \tag{0.181}$$

While orthonormal bases refer to physical observations, we will find in calculations that coordinate bases tend to be more useful.

Example 24: This is an elaboration of part of Example 7.10 in Hartle. Let's consider the two-dimensional polar coordinates shown below, with metric $ds^2 = dr^2 + r^2 d\phi^2$. This means $g_{11} = g_{rr} = 1$, $g_{22} = g_{\phi\phi} = r^2$, and $g_{12} = g_{r\phi} = g_{21} = g_{\phi r} = 0$. The corresponding orthonormal basis vectors $\{\mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\phi}}\}$ and coordinate basis vectors $\{\mathbf{e}_r, \mathbf{e}_\phi\}$:



The components of the coordinate basis vectors *in the coordinate basis* are (i ranges over $1 = r$ and $2 = \phi$)

$$(\mathbf{e}_r)^i = (1,0)\,, \qquad (\mathbf{e}_\phi)^i = (0,1)\,. \tag{0.182}$$

Often expression like this are written more simply as

$$\mathbf{e}_r = (1,0)\,, \qquad \mathbf{e}_\phi = (0,1) \tag{0.183}$$

with the added specification that the components of the vectors are represented in the coordinate basis. The fact that they come out $(1,0)$ and $(0,1)$ is basically the definition of the vectors, but at the risk of being pedantic there are two ways we can see that this is correct.

One is to compute the scalar product in the coordinate basis (i, j, k, l all range over $1 = r$ and $2 = \phi$)

$$\mathbf{e}_i \cdot \mathbf{e}_j = g_{kl}(\mathbf{e}_i)^k(\mathbf{e}_j)^l \implies \begin{cases} \mathbf{e}_r \cdot \mathbf{e}_r = g_{11}(\mathbf{e}_r)^1(\mathbf{e}_r)^1 = 1\,, \\ \mathbf{e}_r \cdot \mathbf{e}_\phi = 0\,, \\ \mathbf{e}_\phi \cdot \mathbf{e}_\phi = g_{22}(\mathbf{e}_\phi)^2(\mathbf{e}_\phi)^2 = r^2 \end{cases} \implies \mathbf{e}_i \cdot \mathbf{e}_j = g_{ij} \tag{0.184}$$

Thus the defining relation for a coordinate basis is satisfied. The other is simply to expand the vector $\mathbf{e}_i$ in the coordinate basis:

$$\mathbf{e}_r = e^i_r \mathbf{e}_i, \quad \mathbf{e}_\phi = e^i_\phi \mathbf{e}_i \implies e^r_r = e^\phi_\phi = 1, \quad e^\phi_r = e^r_\phi = 0\,. \tag{0.185}$$

The notation is a little funny, but these are just the components of the vectors $(\mathbf{e}_r)^i$ and $(\mathbf{e}_\phi)^i$.

The components of the orthonormal basis vectors in the orthonormal basis can be similarly shown to be

$$(\mathbf{e}_{\hat{r}})^{\hat{i}} = (1, 0)\,, \qquad (\mathbf{e}_{\hat{\phi}})^{\hat{i}} = (0, 1)\,. \tag{0.186}$$

Again, this is basically definitional: the components of a basis vector **in that basis** are always going to look this way. However, the components of a basis vector **in another basis** will be more nontrivial; the coordinate components of the orthonormal basis vectors are

$$(\mathbf{e}_{\hat{r}})^i = (1, 0)\,, \qquad (\mathbf{e}_{\hat{\phi}})^i = (0, 1/r)\,. \tag{0.187}$$

We check that this is correct by calculating the scalar product in the coordinate basis

$$\mathbf{e}_{\hat{i}} \cdot \mathbf{e}_{\hat{j}} = g_{kl}(e_{\hat{i}})^k(e_{\hat{j}})^l \implies \begin{cases} \mathbf{e}_{\hat{r}} \cdot \mathbf{e}_{\hat{r}} = g_{11}(\mathbf{e}_{\hat{r}})^1(\mathbf{e}_{\hat{r}})^1 = 1\,, \\ \mathbf{e}_{\hat{r}} \cdot \mathbf{e}_{\hat{\phi}} = 0\,, \\ \mathbf{e}_{\hat{\phi}} \cdot \mathbf{e}_{\hat{\phi}} = g_{22}(\mathbf{e}_{\hat{\phi}})^2(\mathbf{e}_{\hat{\phi}})^2 = 1 \end{cases} \implies \mathbf{e}_{\hat{i}} \cdot \mathbf{e}_{\hat{j}} = \delta_{\hat{i}\hat{j}}\,. \tag{0.188}$$

Notice that we get $\delta_{\hat{i}\hat{j}}$ and not $\eta_{\hat{i}\hat{j}}$ since this computation was in Euclidean space, where the metric has purely positive eigenvalues.

Light cones in curved spacetime

We can separate vectors according to whether $\mathbf{v} \cdot \mathbf{v} < 0$ (timelike), $\mathbf{v} \cdot \mathbf{v} > 0$ (spacelike), or $\mathbf{v} \cdot \mathbf{v} = 0$ (null). Light rays follow worldlines whose tangent vectors are null, i.e. $ds^2 = 0$ everywhere along the curve. Notice that this means the coordinate speed of light can differ from $c = 1$! To see this, let's return to our Newtonian metric

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)(dx_i dx^i) = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)(dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)) \tag{0.189}$$

Suppose we have spherical symmetry, like a planet centered at the origin $r = 0$, so that $\Phi(r)$ only depends on r. We consider a *radial* null ray (i.e. one with constant θ, ϕ):

$$0 = ds^2 = -(1+2\Phi)dt^2 + (1-2\Phi)dr^2 \implies \left(\frac{dr}{dt}\right)^2 = \frac{1+2\Phi}{1-2\Phi} \approx 1 + 4\Phi + O(\Phi^2) \quad (0.190)$$

$$\implies v = \frac{dr}{dt} \approx 1 + 2\Phi + O(\Phi^2)\,. \quad (0.191)$$

So if we have $\Phi = -GM/r < 0$, which corresponds to an attractive force like that due to a planet, then $v < 1$. One way to interpret this is to say light cones are "smaller" closer to the planet (smaller r), and so light travels "more slowly." The physical implication is that light signals passing through small r take **longer** to propagate than in flat spacetime!

Shapiro time delay has been measured by various NASA missions, originally with the Viking landers on Mars, and more recently with the Cassini spacecraft near Saturn. The basic idea is that a signal was sent from Earth, past the sun, to the spacecraft, and returned to Earth. The roundtrip time was calculated and compared at various points in the Earth's orbit when the signal would pass more or less close to the Sun.

Notice that flipping the sign of M means light cones will "open up," allowing light to travel faster than c. This requires negative mass, and there are general theorems relating faster-than-light travel to negative energy, which is forbidden in classical general relativity.

Even though $c \neq 1$ in the coordinate sense above, freely falling observers making **local** measurements of c will always find $c = 1$. This is because their local metric is Minkowski space:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + O[(x-x_0)^2] \implies ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 + O[(x-x_0)^2] \quad (0.192)$$

so null rays on this geometry travel on curves

$$ds^2 = 0 \implies \left(\frac{d\mathbf{x}}{dt}\right)^2 = 1 + O[(x-x_0)^2] \implies c = 1 \text{ locally at } x_0 \quad (0.193)$$

Penrose diagrams

It is still useful, if possible, to choose coordinates such that the **coordinate** speed of light is precisely 1. This lets us keep drawing spacetime diagrams with light rays at 45°. We will also try to choose coordinates such that the entire spacetime (which can be infinite!) can be drawn on a diagram of finite size. (Think of the function $y = \tanh(x)$; x can have infinite range but $y \in (-1, 1)$.) Such diagrams are called Penrose diagrams or conformal diagrams, and a theorist's chalkboard is always full of them.

Let's draw the Penrose diagram for flat Minkowski space as an example:

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2\,, \qquad d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2\,. \quad (0.194)$$

Radial light rays already travel at $dr/dt = \pm 1$, so that part is solved. But we want to find better coordinates (T, R) whose range is finite and preserves $dR/dT = \pm 1$. We'll do this in three steps.

1. Let $u = t - r$, $v = t + r$, which are known as null coordinates. The metric becomes

$$ds^2 = -dudv + \underbrace{\frac{1}{4}(u - v)^2 d\Omega^2}_{\text{ignore from now on}} \tag{0.195}$$

   Note that $-\infty < u \leq v < +\infty$ since $v - u = 2r \geq 0$.

2. Let $U = \tan^{-1} u \in (-\pi/2, \pi/2)$, $V = \tan^{-1} v \in (-\pi/2, \pi/2)$ which means $V \geq U$. The metric becomes

$$ds^2 = -\frac{du}{dU}\frac{dv}{dV}\, dUdV + (\cdots)d\Omega^2 = -\frac{1}{\cos^2 U \cos^2 V}\, dUdV + (\cdots)d\Omega^2\,. \tag{0.196}$$

3. Let $U = T - R$, $V = T + R$, so the metric becomes

$$ds^2 = \frac{1}{\cos^2(T - R)\cos^2(T + R)}\,(-dT^2 + dR^2) + (\cdots)d\Omega^2\,. \tag{0.197}$$

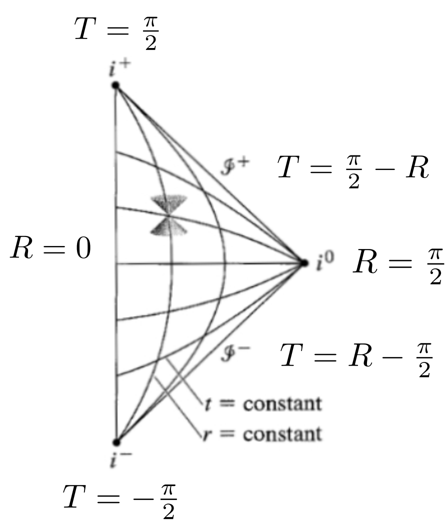
That's it! We see that radial null rays still satisfy

$$ds^2 = 0 \implies \frac{dR}{dT} = \pm 1\,, \tag{0.198}$$

and the coordinates satisfy

$$R = \frac{1}{2}(V - U) \implies 0 \leq R < \pi/2\,, \qquad T = \frac{1}{2}(V + U) \implies R - \frac{\pi}{2} < T < \frac{\pi}{2} - R \tag{0.199}$$

since $-\pi/2 < U \leq V < \pi/2$. Since the ranges of R and T are finite, we should be able to draw this geometry on a piece of paper. Here it is:



Notice that infinity comes in many flavors. We have

- $i^{\pm}$: past & future timelike infinity, where all timelike geodesics begin and end.
- i^0: spacelike infinity, where all spacelike geodesics begin and end.
- $\mathscr{I}^{\pm}$: past & future null infinity, where all null geodesics begin and end.

——————————— END LEC 8 ———————————

Geodesic Equation

Straight lines are to flat space(time) as geodesics are to curved space(time). In curved space, geodesics measure the shortest distance between two points. For example, the geodesic between two points on the equator of the Earth is the equatorial line connecting them. More generally, any two points on the sphere have a geodesic given by the portion of the great circle which the two points lie on (a great circle is the largest possible circle that can be drawn on the sphere, e.g. the equator is a great circle but other lines of constant latitude are not).

In spacetime, as we saw, we need to generalize "shortest" to "extremal." This is because of the minus sign in the metric. Inertial observers, for example, *maximize* their proper time. We can capture all these cases by saying that geodesics are defined as **stationary points of the proper time or length**. We will focus on timelike geodesics, since this corresponds to the motion of massive particles. We introduce a parameter along the curve λ and write the proper time as

$$\tau = \int \sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}\, d\lambda\,. \tag{0.200}$$

Think of this as an action with Lagrangian

$$L(x^\alpha, \dot{x}^\alpha) = \sqrt{-g_{\alpha\beta}(x)\dot{x}^\alpha\dot{x}^\beta}\,. \tag{0.201}$$

where we defined $\dot{x}^\alpha = \frac{dx^\alpha}{d\lambda}$. The Euler-Lagrange equations are therefore

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0\,. \tag{0.202}$$

We can use the chain rule to compute the various derivatives (introducing the notation $\partial_\alpha = \frac{\partial}{\partial x^\alpha}$):

$$\frac{\partial L}{\partial \dot{x}^\alpha} = \frac{1}{2L}\left(-2g_{\alpha\beta}\frac{dx^\beta}{d\lambda}\right), \qquad \frac{\partial L}{\partial x^\alpha} = \frac{1}{2L}\left(-\partial_\alpha g_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda}\right) \tag{0.203}$$

The factors of L are annoying, so let's use $L = \frac{d\tau}{d\lambda}$ (from $\tau = \int L d\lambda$) to write

$$\frac{1}{L}\frac{dx^\beta}{d\lambda} = \frac{dx^\beta}{d\tau}\,. \tag{0.204}$$

Multiplying the Euler-Lagrange equations by $1/L$ gives

$$0 = \underbrace{\frac{1}{L}\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\alpha}\right)}_{(1)} - \underbrace{\frac{1}{L}\frac{\partial L}{\partial x^\alpha}}_{(2)} \tag{0.205}$$

Using (0.204) we write the first term as

$$(1) = \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\alpha}\right) = \frac{d}{d\tau}\left(-g_{\alpha\beta}\frac{dx^\beta}{d\tau}\right) = -g_{\alpha\beta}\frac{d^2x^\beta}{d\tau^2} - \partial_\gamma g_{\alpha\beta}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau}\,. \tag{0.206}$$

We can write

$$\partial_\gamma g_{\alpha\beta}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau} = \partial_\gamma g_{\alpha\beta}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = \partial_\beta g_{\alpha\gamma}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau}\,, \tag{0.207}$$

where in the first equality we just commuted the derivatives past each other and in the second equality we renamed the indices $\gamma \leftrightarrow \beta$. Since these indices are repeated they are to be summed over, so you can call them whatever you want; they sometimes go by the name of “dummy” indices. If you’re nervous about it, write out the sums explicitly and verify that they are the same. We finally have

$$(1) = -g_{\alpha\beta}\frac{d^2x^\beta}{d\tau^2} - \frac{1}{2}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau}\left(\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\alpha\gamma}\right)\,. \tag{0.208}$$

(We’ll see soon why we rewrote the second term in this way; for now it looks a little crazy.) Next we have

$$(2) = \frac{1}{L}\frac{\partial L}{\partial x^a} = \frac{1}{2L^2}\left(-\partial_\alpha g_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda}\right) = -\frac{1}{2}\partial_\alpha g_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau}\,. \tag{0.209}$$

So altogether the Euler-Lagrange equations give

$$g_{\alpha\beta}\frac{d^2x^\beta}{d\tau^2} + \frac{1}{2}\frac{dx^\gamma}{d\tau}\frac{dx^\beta}{d\tau}\left(\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\alpha\gamma} - \partial_\alpha g_{\beta\gamma}\right) = 0\,. \tag{0.210}$$

We can go one step further and multiply the equation by the “inverse metric” $g^{\mu\alpha}$. This is defined by

$$g^{\mu\alpha}g_{\alpha\beta} = \delta^\mu_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{cases} 1\,, & \text{if } \mu = \beta\,, \\ 0\,, & \text{if } \mu \neq \beta\,, \end{cases} \tag{0.211}$$

where we gave two representations of the “Kronecker delta” δ^μ_β. Notice that $g_{\alpha\beta}$ is definitely invertible since its eigenvalues are $-+++$, i.e. none of them vanish. We multiply our Euler-Lagrange equation by this quantity and get for the first term

$$g^{\mu\alpha}g_{\alpha\beta}\frac{d^2x^\beta}{d\tau^2} = \delta^\mu_\beta\frac{d^2x^\beta}{d\tau^2} = \frac{d^2x^\mu}{d\tau^2}\,. \tag{0.212}$$

To understand the final equality, we note that

$$\delta^{\mu}_{\beta} X^{\beta} = X^{\mu} , \tag{0.213}$$

where X^{β} can be any vector. We can check this component-by-component, e.g.

$$\mu = 1 : \qquad \delta^{1}_{\beta} X^{\beta} = \delta^{1}_{1} X^{1} + \delta^{1}_{2} X^{2} + \delta^{1}_{3} X^{3} + \delta^{1}_{4} X^{4} = X^{1} . \tag{0.214}$$

The second term becomes

$$\frac{1}{2} g^{\mu\alpha} \frac{dx^{\gamma}}{d\tau} \frac{dx^{\beta}}{d\tau} \left(\partial_{\gamma} g_{\alpha\beta} + \partial_{\beta} g_{\alpha\gamma} - \partial_{\alpha} g_{\beta\gamma} \right) . \tag{0.215}$$

We define the Christoffel symbol

$$\Gamma^{\mu}_{\gamma\beta} = \frac{1}{2} g^{\mu\alpha} \left(\partial_{\gamma} g_{\alpha\beta} + \partial_{\beta} g_{\alpha\gamma} - \partial_{\alpha} g_{\beta\gamma} \right) \tag{0.216}$$

and write our Euler-Lagrange equation as

$$\boxed{\frac{d^2 x^{\mu}}{d\tau^2} = -\Gamma^{\mu}_{\gamma\beta} \frac{dx^{\gamma}}{d\tau} \frac{dx^{\beta}}{d\tau} .} \tag{0.217}$$

This is known as the **geodesic equation**: it gives the 4-acceleration in terms of geometry (encoded in the Christoffel symbol) and the 4-velocity.

An important property of the Christoffel symbol is its symmetry in the lower indices, $\Gamma^{\mu}_{\gamma\beta} = \Gamma^{\mu}_{\beta\gamma}$. For flat space in Cartesian coordinates, the derivatives of the metric vanish and $\Gamma^{\mu}_{\gamma\beta} = 0$, so we recover the fact that geodesics don't undergo any acceleration, $d^2 x^{\mu}/d\tau^2 = 0$.

This procedure is easy to repeat for spacelike geodesics, and we will get the same equation, except with τ replaced by σ, the proper distance.

Example 25: Let's calculate geodesics in flat space in *polar coordinates*. We already know they should be straight lines, but it will be instructive to see this come out of the geodesic equation. The metric is

$$ds^2 = dr^2 + r^2 d\phi^2 \implies g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} , \qquad g^{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix} . \tag{0.218}$$

We calculate the Christoffel symbols, component-by-component (we use $r = 1$ and $\phi = 2$):

$$\Gamma^{1}_{11} = \frac{1}{2} g^{1\alpha} (\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_{\alpha} g_{11}) = 0 \tag{0.219}$$

$$\Gamma^{1}_{12} = \frac{1}{2} g^{1\alpha} (\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_{\alpha} g_{21}) = 0 = \Gamma^{1}_{21} \tag{0.220}$$

$$\Gamma^{1}_{22} = \frac{1}{2} g^{1\alpha} (\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_{\alpha} g_{22}) = -\frac{1}{2} g^{11} \partial_1 g_{22} = -r \tag{0.221}$$

$$\Gamma^{2}_{22} = \frac{1}{2} g^{2\alpha} (\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_{\alpha} g_{22}) = 0 \tag{0.222}$$

$$\Gamma^2_{12} = \frac{1}{2}g^{2\alpha}(\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_\alpha g_{21}) = \frac{1}{2}g^{22}\partial_1 g_{22} = \frac{1}{r} = \Gamma^2_{21} \tag{0.223}$$

$$\Gamma^2_{11} = \frac{1}{2}g^{2\alpha}(\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_\alpha g_{11}) = 0 \tag{0.224}$$

Our geodesic equation therefore becomes

$$\frac{d^2x^1}{d\sigma^2} = -\Gamma^1_{22}\frac{dx^2}{d\sigma}\frac{dx^2}{d\sigma} \implies r''(\sigma) = r(\sigma)\phi'(\sigma)^2\,, \tag{0.225}$$

$$\frac{d^2x^2}{d\sigma^2} = -2\Gamma^2_{12}\frac{dx^1}{d\sigma}\frac{dx^2}{d\sigma} \implies \phi''(\sigma) = -\frac{2}{r(\sigma)}\,\phi'(\sigma)r'(\sigma)\,. \tag{0.226}$$

Notice these are the same as (8.6a), (8.6b) in Example 8.1 of Hartle (although what Hartle calls S we are calling σ!). In that example, Hartle extremizes the proper distance in polar coordinates. Let's redo it in a manner similar to our derivation of the geodesic equation above

$$S = \int d\lambda \sqrt{\left(\frac{dr}{d\lambda}\right)^2 + r^2\left(\frac{d\phi}{d\lambda}\right)^2} \implies L = \sqrt{\dot{r}^2 + r^2\dot{\phi}^2} \tag{0.227}$$

Notice that the Lagrangian depends on r, $\dot{r}$ and $\dot{\phi}$ – in particular it doesn't depend on ϕ. This is because there was no dependence on ϕ in the metric, and it means we have a symmetry of the geometry: ϕ is the same as $\phi + const$. This is similar to the translation symmetries we explored in Euclidean space and Minkowski space.

Anyway, our Euler-Lagrange equations become

$$0 = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} \implies 0 = \frac{d}{d\sigma}\left(\frac{\partial L}{\partial \dot{x}^\alpha}\right) - \frac{1}{L}\frac{\partial L}{\partial x^\alpha} \tag{0.228}$$

where we used $\frac{1}{L}\frac{d}{d\lambda} = \frac{d}{d\sigma}$. Let's work out the $\alpha = 2 = \phi$ case using

$$\frac{\partial L}{\partial \phi} = 0\,, \qquad \frac{\partial L}{\partial \dot{\phi}} = \frac{1}{2L}r^2 2\dot{\phi} = r^2\frac{d\phi}{d\sigma}\,. \tag{0.229}$$

Our Euler-Lagrange equation becomes

$$\frac{d}{d\sigma}\left(r^2\frac{d\phi}{d\sigma}\right) = 0\,. \tag{0.230}$$

which is the same as (0.226). The $\alpha = 1 = r$ case gives

$$\frac{d^2r}{d\sigma^2} - r\left(\frac{d\phi}{d\sigma}\right)^2 = 0\,, \tag{0.231}$$

which agrees with (0.225). We can actually read off the nonvanishing Christoffel symbols by comparing these equations to the general geodesic equation.

In the example above, we had a symmetry $\phi \to \phi + const.$ and it led to a conservation law: the "angular momentum" $J = r^2 d\phi/d\sigma$ is **constant along the geodesic**. There is a general theorem, due to Emmy Noether, that relates symmetries (in the Lagrangian formulation of a theory) to conservation laws. The conservation law above is a consequence of this theorem.

Due to the conservation law, we only need to solve (0.225), which can be written

$$\frac{d^2r}{d\sigma^2} = \frac{J^2}{r^3} \tag{0.232}$$

where J is the constant angular momentum. To solve this, we will use a "free" conserved quantity we always have access to:

$$d\sigma^2 = dr^2 + r^2 d\phi^2 \implies 1 = \frac{d\sigma^2}{d\sigma^2} = \left(\frac{dr}{d\sigma}\right)^2 + r^2\left(\frac{d\phi}{d\sigma}\right)^2 \implies \frac{dr}{d\sigma} = \sqrt{1 - \frac{J^2}{r^2}}\,. \tag{0.233}$$

We can check that another σ derivative of this recovers (0.232), but we have reduced the second-order equation to a first order equation. Now we can just integrate

$$\sigma - \sigma_0 = \int d\sigma = \int \frac{dr}{\sqrt{1 - J^2/r^2}} = \frac{1}{2}\int \frac{d(r^2)}{\sqrt{r^2 - J^2}} = \sqrt{r^2 - J^2} \tag{0.234}$$

This gives

$$r^2 = J^2 + (\sigma - \sigma_0)^2\,. \tag{0.235}$$

We can also integrate the conservation equation

$$\frac{d\phi}{d\sigma} = \frac{J}{J^2 + (\sigma - \sigma_0)^2} \implies \phi - \phi_0 = \tan^{-1}\left(\frac{\sigma - \sigma_0}{J}\right) \implies \sigma - \sigma_0 = J\tan(\phi - \phi_0). \tag{0.236}$$

Plugging this into (0.235) gives

$$r^2 = J^2[1 + \tan^2(\phi - \phi_0)] = \frac{J^2}{\cos^2(\phi - \phi_0)} \implies r\cos(\phi - \phi_0) = J\,. \tag{0.237}$$

This is actually a straight line! For example, for $\phi_0 = 0$ we get $x = r\cos\phi = const.$ These are vertical lines whose locations are fixed to be $x = J$. For general ϕ_0 we have lines that are rotated by ϕ_0.

This is also a lesson in choosing the correct coordinates: it is much easier to find straight lines as the geodesics in flat space if we work in Cartesian coordinates.

Symmetries and conservation laws

We already mentioned a connection, due to Emmy Noether, between symmetries (in the Lagrangian formulation) and conservation laws. For a d-dimensional Lagrangian system with d conservation laws, the Euler-Lagrange equations reduce to 1st order equations which are easily soluble. Geodesics *always* have the conservation law $\mathbf{u} \cdot \mathbf{u} = const.$ (where $const = +1$ for spacelike geodesics, -1 for timelike geodesics, and 0 for null geodesics).[10]

A simple way to find symmetries is the following: if you can find some coordinates x^α in which the metric $g_{\alpha\beta}$ is invariant under a constant shift $x^\alpha \to x^\alpha + \xi^\alpha$ (i.e. ξ^α is independent of x^α), then we say that spacetime has a symmetry. The symmetry is a physical thing that doesn't depend on the coordinates; this is just a convenient way to discover the symmetry! For example, the rotational symmetry $\phi \to \phi + const.$ was trivial to see in polar coordinates $ds^2 = dr^2 + r^2 d\phi^2$. But it is still there in Cartesian coordinates. If we insist on seeing it, we can refer to Example 7, where we have the transformation $x \to x\cos\phi_0 + y\sin\phi_0$, $y \to y\cos\phi_0 - x\sin\phi_0$, which keeps the metric in Cartesian coordinates $ds^2 = dx^2 + dy^2$ invariant. Conversely, translations $x \to x + const.$ are hard to see in polar coordinates, but they are still a symmetry.

Let's derive the conserved quantity associated to a symmetry. We assume there exist coordinates x^α with the property that the metric – and therefore L – is invariant under $x^\alpha \to x^\alpha + \xi^\alpha$. Then we can take the scalar product of ξ^α with $1/L$ times our Euler-Lagrange equation to get

$$0 = \xi^\alpha \frac{1}{L}\frac{d}{d\lambda}\left(\frac{\partial L}{\partial(dx^\alpha/d\lambda)}\right) - \frac{1}{L}\xi^\alpha \frac{\partial L}{\partial x^\alpha}\,. \tag{0.238}$$

The second term vanishes, because say the Lagrangian is independent of x^3. Then we can take $\xi^\alpha = (0,0,1,0)$ and the product $\xi^\alpha \frac{\partial L}{\partial x^\alpha}$ picks out $\frac{\partial L}{\partial x^3}$, which vanishes since the L is independent of x^3. We use $L = d\tau/d\lambda$ to rewrite the first term and get

$$0 = \frac{d}{d\tau}\left(\xi^\alpha \frac{\partial L}{\partial(dx^\alpha/d\lambda)}\right). \tag{0.239}$$

Since ξ^α is a constant we moved it inside the derivative. The quantity inside the brackets is our conserved quantity! The derivative of the Lagrangian is like a canonical momentum, but let's make that more precise. Use $L = \sqrt{-g_{\beta\gamma}\dot{x}^\beta \dot{x}^\gamma}$ (dots refer to λ derivatives) to write

$$-\xi^\alpha \frac{\partial L}{\partial \dot{x}^\alpha} = -\xi^\alpha \frac{1}{2L}\left(-g_{\beta\gamma}\delta^\beta_\alpha \dot{x}^\gamma - g_{\beta\gamma}\dot{x}^\beta \delta^\gamma_\alpha\right) = \xi^\alpha \frac{1}{2L}(2g_{\alpha\beta}\dot{x}^\beta) = \xi^\alpha g_{\alpha\beta}\frac{dx^\beta}{d\tau} = \boldsymbol{\xi}\cdot\mathbf{u} \tag{0.240}$$

where $\mathbf{u}$ is the 4-velocity. We could also multiply this by the mass to get $\boldsymbol{\xi}\cdot\mathbf{p}$ as a conserved quantity where $\mathbf{p}$ is the 4-momentum. The vector $\boldsymbol{\xi}$ is a Killing vector, named after Wilhelm Killing. We used a particular coordinate system in which the components of $\boldsymbol{\xi}$ were pretty simple, but the final conserved quantity $\boldsymbol{\xi}\cdot\mathbf{u}$ is independent of coordinates.

——————————— END LEC 9 ———————————

[10] Technically the symmetry that this conservation law corresponds to is a translation symmetry $\lambda \to \lambda + const.$ This is like "time translations" in the freely falling frame, and the corresponding "energy" is conserved, but we won't use this language since we want time translations to refer to $t \to t + const.$ for coordinate t.

Null geodesics

We discussed timelike geodesics as stationary points of the action

$$\tau = \int \sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}\,d\lambda\,, \qquad \delta\tau = 0 \implies \frac{d^2x^\mu}{d\tau^2} = -\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = -\Gamma^\mu_{\alpha\beta}u^\alpha u^\beta \quad (0.241)$$

with $\mathbf{u}\cdot\mathbf{u} = g_{\alpha\beta}u^\alpha u^\beta = -1$. Spacelike geodesics are the same, except the action is the proper distance

$$\sigma = \int \sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}\,d\lambda\,, \qquad \delta\sigma = 0 \implies \frac{d^2x^\mu}{d\sigma^2} = -\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} = -\Gamma^\mu_{\alpha\beta}u^\alpha u^\beta \quad (0.242)$$

with $\mathbf{u}\cdot\mathbf{u} = g_{\alpha\beta}u^\alpha u^\beta = +1$. What about null geodesics, which have $\mathbf{u}\cdot\mathbf{u} = g_{\alpha\beta}u^\alpha u^\beta = 0$? This will be important for understanding e.g. light rays around stars, black holes, dark matter, etc. There is a bit of a problem, since $ds^2 = 0 \implies \Delta\tau = \Delta\sigma = 0$ for a light ray. Here is a slick way to derive the null geodesic equation from the timelike one:

1. Write the timelike geoedesic equation in terms of an "affine" parameter $\lambda = a\tau$ for some constant a:

$$\frac{d^2x^\mu}{d\lambda^2} = -\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\,. \qquad (0.243)$$

   The equation takes the same form as before since the factors of a all cancel.

2. Write

$$\frac{dx^\alpha}{d\lambda} = \frac{1}{a}\frac{dx^\alpha}{d\tau} = \frac{1}{a}u^\alpha \implies g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = -\frac{1}{a^2}\,. \qquad (0.244)$$

3. We take the "double scaling limit" $a \to \infty$, $\tau \to 0$ with $\lambda_{\text{null}} = a\tau$ held finite, and define

$$u^\alpha_{\text{null}} \equiv \frac{dx^\alpha}{d\lambda_{\text{null}}} \implies \mathbf{u}_{\text{null}}\cdot\mathbf{u}_{\text{null}} = 0\,. \qquad (0.245)$$

We therefore see that (0.243) holds with $\lambda = \lambda_{\text{null}}$. We will henceforth drop the "null" subscript and reserve λ for an affine parameter for null geodesics.

Example 26: Let's write an affine parameterization for null geodesics in $(1+1)$-dimensional Minkowski space:

$$ds^2 = -dt^2 + dx^2\,. \qquad (0.246)$$

We already know the answer: light rays follow $x = \pm t$. Pick $x = t$ and write the parameterization

$$x = \lambda\,, \qquad t = \lambda \qquad \text{(more generally can write } x^\alpha = u^\alpha\lambda) \qquad (0.247)$$

The vector $u^\alpha = dx^\alpha/d\lambda$ is tangent to the worldline and $\mathbf{u}\cdot\mathbf{u} = 0$. We have

$$\frac{d^2x^\alpha}{d\lambda^2} = \frac{du^\alpha}{d\lambda} = 0\,, \tag{0.248}$$

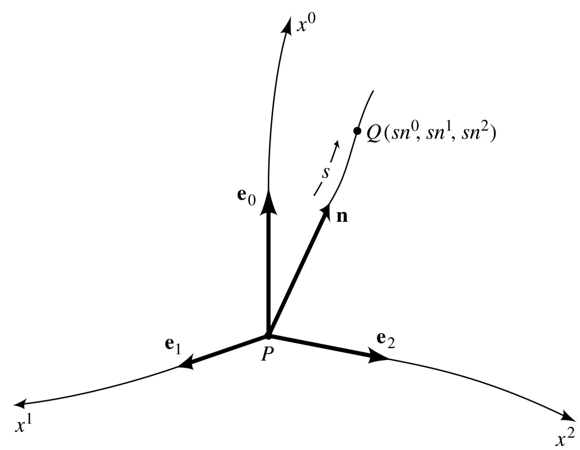
so the geodesic equation (0.243) is satisfied (recall that the Christoffel symbols vanish for flat space in Cartesian coordinates).

Riemann normal coordinates and freely falling frames

By the equivalence principle, the geodesic equation at x_0 should simplify when expressed in a locally inertial frame (LIF) at P. Mathematically this is because an LIF satisfies

$$g_{\alpha\beta}|_P = \eta_{\alpha\beta}\,, \qquad \partial_\gamma g_{\alpha\beta}|_P = 0 \implies \Gamma^\alpha_{\beta\gamma}\Big|_P = 0 \implies \frac{d^2x^\alpha}{d\tau^2}\bigg|_P = 0 \tag{0.249}$$

where the Christoffel symbols vanish since they are built out of $\partial_\gamma g_{\alpha\beta}$, and the final equation is just the geodesic equation with vanishing Christoffels. So the geodesic is **locally** a straight line in an LIF.



We can reverse this logic to *construct* an LIF as shown above. These coordinates will be called Riemann normal coordinates at P. We begin by picking a point x_0 in the spacetime and choosing a basis of four orthonormal vectors $\{\mathbf{e}_\alpha\}$ at P. We consider all geodesics through P, and for each spacelike or timelike one we consider a vector $\mathbf{n} = n^\alpha\mathbf{e}_\alpha$ which is tangent to the geodesic . We go a proper distance or time s along each geodesic and **define** the point to have coordinates $x^\alpha = sn^\alpha$. Points along null geodesics are coordinatized by continuity. The basis vectors $\{\mathbf{e}_\alpha\}$ point along the coordinates and so constitute the coordinate basis vectors of the frame at P, so $\mathbf{e}_\alpha\cdot\mathbf{e}_\beta = g_{\alpha\beta}$, and since they are orthonormal we have $g_{\alpha\beta}|_P = \eta_{\alpha\beta}$. Furthermore, the curves $x^\alpha = sn^\alpha$ (with $n^\alpha = const.$) are geodesics and therefore satisfy

$$\frac{d^2x^\alpha}{ds^2}\bigg|_P = 0 = -\Gamma^\alpha_{\beta\gamma}\Big|_P\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} = -\Gamma^\alpha_{\beta\gamma}\Big|_P n^\beta n^\gamma \implies \Gamma^\alpha_{\beta\gamma}\Big|_P = 0 \tag{0.250}$$

since this is true for arbitrary n^β. This means that $\partial_\gamma g_{\alpha\beta}|_P = 0$, recovering that the metric locally is Minkowski and its first derivative vanishes.

The equivalence principle says that even more than this is possible: one can get $g_{\alpha\beta} = \eta_{\alpha\beta}$ and $\partial_\gamma g_{\alpha\beta} = 0$ not just at a point but **along a geodesic**. These are the coordinates of a freely falling observer, who carries an orthonormal frame along their worldline. We won't discuss how these coordinates are constructed.

Schwarzschild geometry

We will now study the first nontrivial solution to Einstein's equations, which are the equations that determine the gravitational field (i.e. spacetime line element or metric). We won't confront the equations themselves until later in the course. Karl Schwarzschild discovered this solution while volunteering for the German army during WWI. He wrote to Einstein from the trenches:

The war treated me kindly enough, in spite of the heavy gunfire, to allow me to get away from it all and take this walk in the land of your ideas.

He died about a year later.

Here is Schwarzschild's solution, in "Schwarzschild coordinates":

$$\boxed{ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2\,.} \qquad (0.251)$$

Two symmetries are manifest: $t \to t + const.$ and $\phi \to \phi + const.$. The full spherical symmetry is hidden in $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. This metric describes the solution outside **any** spherically symmetric, static distribution of matter. It is the analog of Coulomb's Law in general relativity. Notice that if we expand for small GM/r, i.e. a weak gravitational potential, we get

$$ds^2 \approx -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 + \frac{2GM}{r}\right) dr^2 + r^2 d\Omega^2. \qquad (0.252)$$

This looks like the static weak-field metric

$$ds^2 = -\left(1 + 2\Phi(\tilde{r})\right) dt^2 + (1 - 2\Phi(\tilde{r}))(d\tilde{r}^2 + \tilde{r}^2 d\Omega^2) + O(\Phi^2) \qquad (0.253)$$

with $\Phi(\tilde{r}) = -GM/\tilde{r}$. We put the radial coordinate in the weak field metric as $\tilde{r}$ as a hint: define

$$r \equiv \tilde{r}(1 - \Phi(\tilde{r})) = \tilde{r} + GM \implies r^2 = \tilde{r}^2(1 - 2\Phi(\tilde{r})) + O(\Phi^2)\,, \qquad (0.254)$$

which gives

$$\Phi(r) = -\frac{GM}{r} = \frac{-GM}{\tilde{r} + GM} = \frac{\Phi(\tilde{r})}{1 - \Phi(\tilde{r})} = \Phi(\tilde{r}) + O(\Phi^2) = \frac{GM}{\tilde{r}} + O(\Phi^2)\,, \qquad dr = d\tilde{r}\,. \qquad (0.255)$$

This puts (0.252) in the form of (0.253). The mass parameter M therefore agrees with the Newtonian mass in the Newtonian limit.

There is a special length scale, $r = R_S = 2GM$, which we will call the Schwarzschild radius (for mass M). Something interesting is happening in the metric at $r = R_S$ and $r = 0$: we will come to this when we study black holes. For stars and planets these radii are deep within the star or planet, where the solution is modified from the metric above. For example, the mass of the Earth is $M_{\text{Earth}} \sim 6 \times 10^{24} kg$ which gives a Schwarzschild radius $(R_S)_{\text{Earth}} = \frac{2GM_{\text{Earth}}}{c^2} \sim 1cm$. This is much less than the radius of the Earth!

From now on we will also set $G = 1$, similar to how we set $c = 1$. We can restore factors of G by dimensional analysis (each factor of G converts a mass to a length or time). The Schwarzschild radius is therefore $R_S = 2M$.

Gravitational time dilation and redshift

An observer at fixed position (r, θ, ϕ) experiences a proper time

$$\Delta\tau = \sqrt{-g_{tt}}\,\Delta t = \sqrt{1 - \frac{2M}{r}}\,\Delta t\,, \tag{0.256}$$

where t is coordinate time. Clocks run more slowly deeper in gravitational fields (essentially stopping at $r = 2M$). This is just like what we saw with the weak-field metric (0.143). Similarly, since frequency is related to the inverse period, we have

$$\omega_1 = \frac{1}{\Delta\tau_1} = \frac{1}{\sqrt{-g_{tt}(r_1)}}\Delta t \implies \frac{\omega_1}{\omega_2} = \sqrt{\frac{1 - 2M/r_2}{1 - 2M/r_1}}\,. \tag{0.257}$$

Another way to see this same result is to use the conservation of $\mathbf{p} \cdot \boldsymbol{\xi}$ for any photon along its worldline, where $\xi = (1, 0, 0, 0)$ (in the coordinate basis) is the Killing vector for time translation symmetry $t \to t + const$. We also use

$$u^\alpha_{\text{obs}}(r) = ((1 - 2M/r)^{-1/2}, 0, 0, 0) = (1 - 2M/r)^{-1/2}\xi^\alpha \tag{0.258}$$

which comes from $g_{\alpha\beta}u^\alpha_{\text{obs}}u^\beta_{\text{obs}} = -1$ from normalization of the 4-velocity of a timelike observer and $u^i_{\text{obs}}(r) = \frac{dx^i_{\text{obs}}}{d\tau_{\text{obs}}} = 0$ for a stationary observer. Conservation gives

$$(\mathbf{p} \cdot \boldsymbol{\xi})_1 = (\mathbf{p} \cdot \boldsymbol{\xi})_2 \implies \sqrt{1 - 2M/r_1}\,\mathbf{p} \cdot \mathbf{u}_{\text{obs}}(r_1) = \sqrt{1 - 2M/r_2}\,\mathbf{p} \cdot \mathbf{u}_{\text{obs}}(r_2). \tag{0.259}$$

The observed energy of the photon is $E = \hbar\omega = -\mathbf{p} \cdot \mathbf{u}_{\text{obs}}$, which altogether gives

$$\omega_1 = \omega_2 \frac{\sqrt{1 - 2M/r_2}}{\sqrt{1 - 2M/r_1}}\,. \tag{0.260}$$

More generally, for stationary observers in a metric with time-translation symmetry (and therefore energy conservation), we have

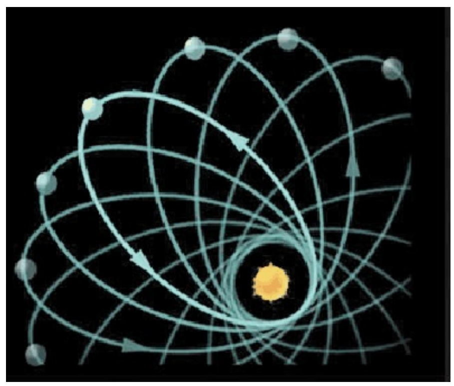
$$\omega_1 = \omega_2\sqrt{g_{tt}(r_2)/g_{tt}(r_1)}\,. \tag{0.261}$$

The photon loses energy when climbing out of a gravitational well, similar to what you would expect for a massive particle.

——————————————— END LEC 10———————————————

Timelike orbits & geodesics in Schwarzschild

The orbits of all the planets "precess" a little, which means that the ellipses move around. This is often called the "precession of the perihelion," i.e. the point of closest approach to the sun (perihelion) precesses:



Historically, there was a known anomaly in the precession of the perihelion of Mercury before Einstein developed his general theory of relativity. It was measured to be 5600 arcseconds per century, which is 1.6° per *century.* This was almost accounted for by Newton's theory: the tug of various planets, the rotation of the sun, etc., affect the Keplerian ellipses and cause some precession. Accounting for all of these led to 5557 arcseconds per century, leaving an unexplained 43 arcseconds = .01° per century (!!). That's tiny, but it was known even then, and people worried about it. The form of the dominant explanation was that there was some new matter (maybe a planet, maybe dust between the Sun and Mercury) which explained the remaining 43 arcseconds per century. To understand this, we will undertake a study of orbits and geodesics in the Schwarzschild geometry.

The Schwarzschild geometry has so much symmetry that we can solve for the geodesics analytically. We will work in the equatorial plane $\theta = \pi/2$. This is kept fixed by the (discrete) symmetry transformation $\theta \to \pi - \theta$. This means if we begin with $\dot{\theta} = 0$ and $\theta = \pi/2$, then we must remain that way, since there is nothing to determine whether θ should become greater or less than $\pi/2$.

The symmetries that lead to conserved quantities are time-translation and ϕ-translation. The time translation Killing vector $\boldsymbol{\xi}_t = (1, 0, 0, 0)$ leads to the conservation of

$$e = -\boldsymbol{\xi}_t \cdot \mathbf{u} = -g_{tt} u^t = (1 - 2M/r)\frac{dt}{d\tau}\,. \tag{0.262}$$

Notice at large r this becomes the energy per unit rest mass since $E = mu^t = m(dt/d\tau)$. So we will call this the energy per unit rest mass everywhere.

The ϕ translation Killing vector $\boldsymbol{\xi}_\phi = (0, 0, 0, 1)$ leads to the conservation of

$$\ell = \boldsymbol{\xi}_\phi \cdot \mathbf{u} = g_{\phi\phi} u^\phi = r^2 \sin^2\theta \frac{d\phi}{d\tau} \,. \tag{0.263}$$

This is the angular momentum per unit rest mass (since it reduces to this at low velocity). I will sometimes drop the modifier "per unit rest mass."

We also always have

$$\mathbf{u} \cdot \mathbf{u} = -1 \implies -\left(1 - \frac{2M}{r}\right)\left(\frac{dt}{d\tau}\right)^2 + \left(1 - \frac{2M}{r}\right)^{-1}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{d\phi}{d\tau}\right)^2 = -1 \tag{0.264}$$

since $d\theta/d\tau = 0$ and $\theta = \pi/2$. We can multiply this by $-(1 - 2M/r)$ and use our conserved quantities to write it as

$$e^2 - \left(\frac{dr}{d\tau}\right)^2 - \frac{\ell^2}{r^2}\left(1 - \frac{2M}{r}\right) = 1 - \frac{2M}{r} \,. \tag{0.265}$$

We therefore have an additional relation that can help determine $r(\tau)$, which we can write as

$$\mathcal{E} \equiv \frac{e^2 - 1}{2} = \underbrace{\frac{1}{2}\left(\frac{dr}{d\tau}\right)^2}_{\text{“Kinetic energy”}} + \underbrace{\frac{1}{2}\left[\left(1 - \frac{2M}{r}\right)\left(1 + \frac{\ell^2}{r^2}\right) - 1\right]}_{V_{\text{eff}}(r)} \,. \tag{0.266}$$

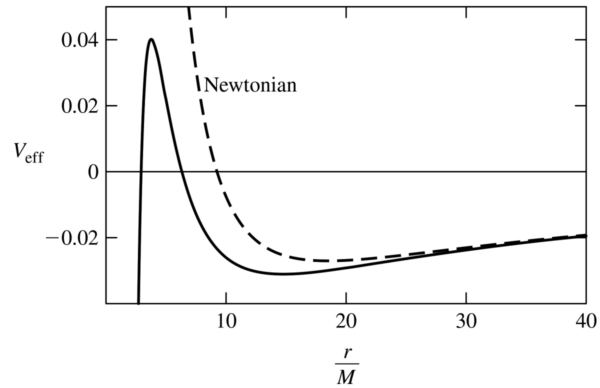
This is like ordinary energy conservation in an effective potential

$$\boxed{V_{\text{eff}}(r) = -\frac{M}{r} + \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3}} \,. \tag{0.267}$$

The first term is the Newtonian potential Φ, the second term is the Newtonian centrifugal barrier. The source of this term, if you don't remember, is actually the part of the kinetic energy (per unit mass) $\frac{1}{2}v^2$ coming from motion in the ϕ direction:

$$\frac{1}{2}v^2 = \frac{1}{2}\dot{r}^2 + \frac{r^2}{2}\dot{\phi}^2 = \frac{1}{2}\dot{r}^2 + \frac{\ell^2}{2r^2} \,. \tag{0.268}$$

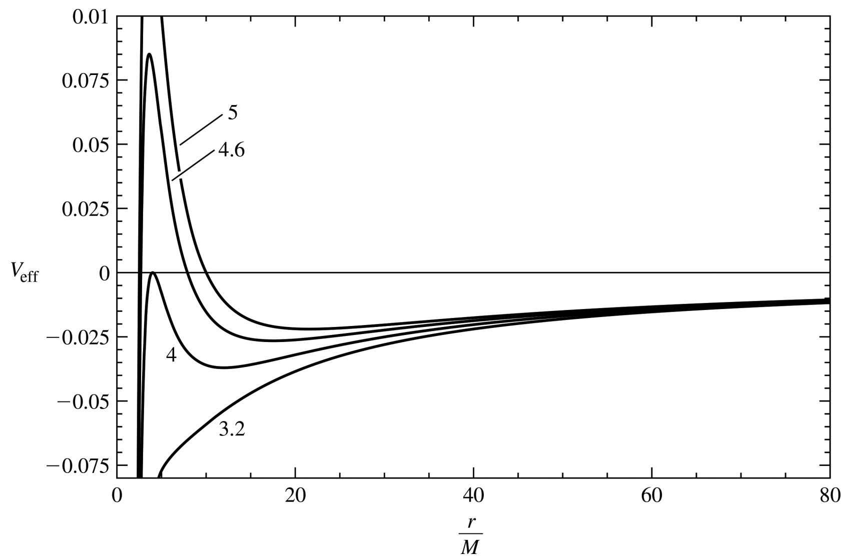
The term $-M\ell^2/r^3$, on the other hand, is a brand new effect coming from general relativity! Let's plot the prediction from general relativity against the one from Newtonian gravity, for $\ell = 4.3M$:



Notice that at large r the effect of the new term diminishes (we enter the weak-field limit) and we get agreement between the two curves. Thankfully we still have a minimum, so stable orbits still exist. The extrema are calculated as

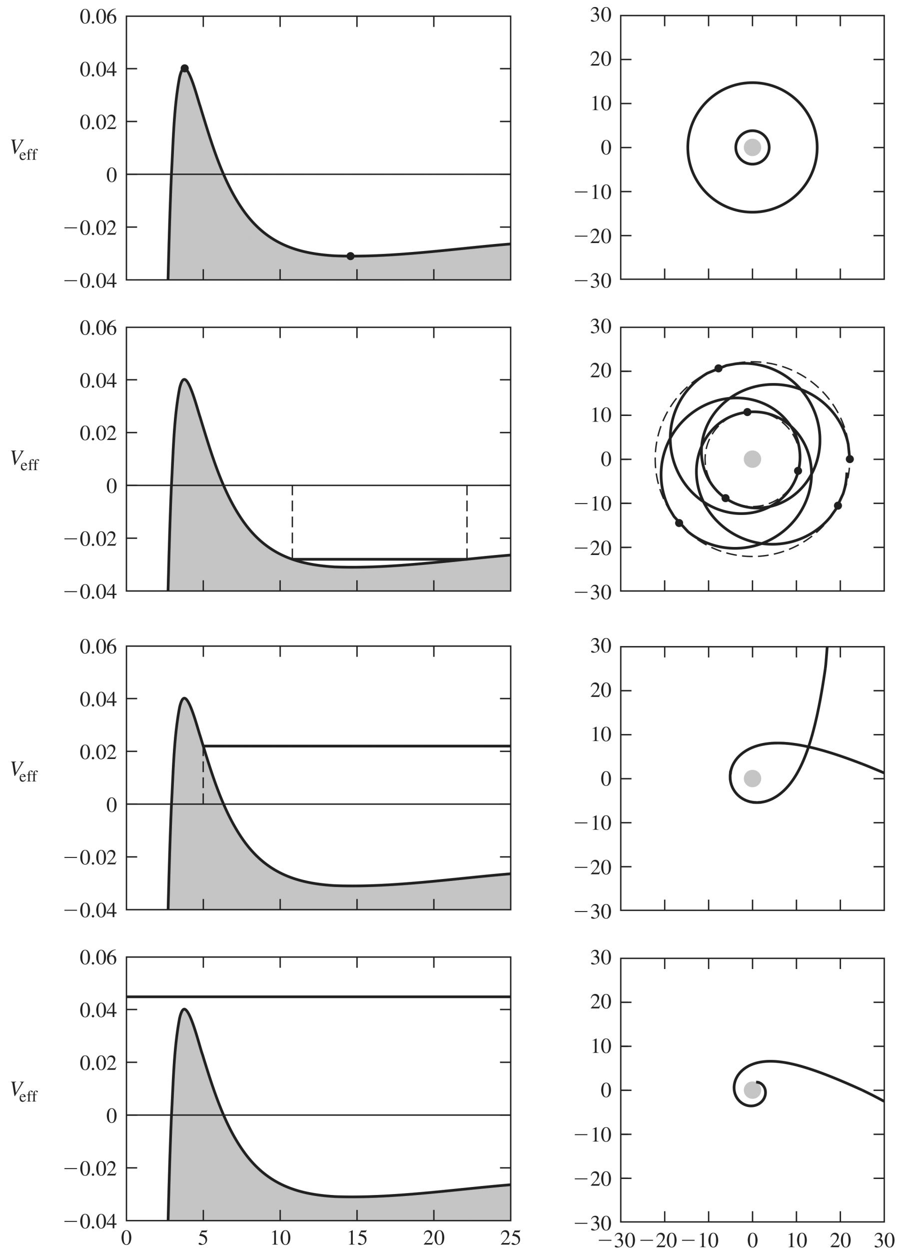
$$0 = V'_{\text{eff}}(r) = \frac{M}{r^2} - \frac{\ell^2}{r^3} + \frac{3M\ell^2}{r^4} \implies r = r_\pm = \frac{\ell^2}{2M}\left(1 \pm \sqrt{1 - \frac{12M^2}{\ell^2}}\right) . \qquad (0.269)$$

For $\ell^2 > 12M^2$, both roots are real, and r_+ corresponds to the minimum in the plot above while r_- corresponds to the maximum. The maximum is an unstable orbit – a small perturbation will push it to larger or smaller r. For $\ell^2 < 12M^2$ the roots are complex, and the effective potential is monotonically increasing and negative everywhere. We plot some effective potentials below with ℓ/M labeling the curves



The key thing to notice in these curves is that for sufficiently small ℓ the minima and maxima are lost, and particles will fall straight toward $r = 0$ (they just roll down the hill). The new term in general relativity overwhelms the centrifugal barrier, which in the Newtonian case always keeps particles from rolling to $r = 0$ (unless ℓ is *exactly* zero).

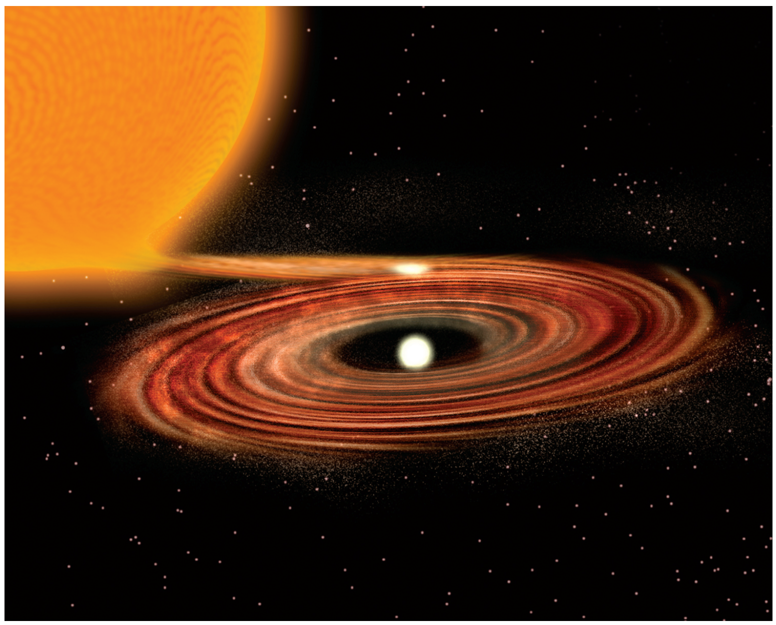
There are four qualitatively different kinds of orbits, plotted below.



The bold lines on the left indicate values of $\mathcal{E}$, and turning points are at $\mathcal{E} = V_{\text{eff}}(r_{\text{tp}})$ since the radial velocity vanishes there. Bound orbits require $\mathcal{E} < 0 \implies e^2 < 1$.

The first orbits are ordinary circular orbits (the larger one stable, the smaller one unstable). The second are elliptical orbits, although the perihelion precesses (we will come back to this). The third are "scattering states" – they have enough energy to slingshot around, not getting trapped.[11] The final case spirals into the center. The way to understand this is that compared to the previous case, we are keeping ℓ fixed but increasing $\mathcal{E}$, which means we are increasing $dr/d\tau$. The particle therefore has a higher radially inward velocity, making it more likely it will fall into the black hole.

The barrier $\ell^2 = 12M^2$ between the existence and non-existence of stable orbits defines the "innermost stable circular orbit" (ISCO), which is at $r_+ = r_- = 6M$. This is very important for accretion disks in very compact objects, displayed below. Beyond $r = 6M$ the matter just falls into the star and doesn't accrete in stable orbits around it. So the inner edge of the accretion disk is at $r_{ISCO} \equiv 6M$.



Example 27: Let's calculate the 4-velocity of a particle in a stable circular orbit (constant $r = r_+$ and constant $\theta = \pi/2$):

$$u^\alpha = (u^t, 0, 0, u^\phi) = \left(\frac{dt}{d\tau}, 0, 0, \frac{d\phi}{d\tau}\right) = \left(\frac{e}{1 - 2M/r_+}, 0, 0, \frac{\ell}{r_+^2}\right) \tag{0.270}$$

We want to write this just in terms of M and r_+, which should determine ℓ and e. We can solve for them through the pair of equations

$$r_+ = \frac{\ell^2}{2M}\left(1 + \sqrt{1 - \frac{12M^2}{\ell^2}}\right), \qquad \mathcal{E}(r_+) = \frac{e^2 - 1}{2} = -\frac{M}{r_+} + \frac{\ell^2}{2r_+^2} - \frac{M\ell^2}{r_+^3}, \tag{0.271}$$

[11] This is *not* the same as a "gravity assist" which NASA uses to cut costs in rocketry. In that case momentum or energy is *extracted* from the body the rocket slingshots around: for example the rocket can extract energy from the rotation of a planet, slowing down its rotation a tiny bit.

which together give

$$e = \frac{1 - 2M/r_+}{\sqrt{1 - 3M/r_+}}\,, \qquad \ell = \frac{\sqrt{Mr_+}}{\sqrt{1 - 3M/r_+}}\,. \tag{0.272}$$

Plugging these into the 4-velocity gives

$$u^\alpha = (1 - 3M/r_+)^{-1/2}\left(1, 0, 0, \sqrt{\frac{M}{r_+^3}}\right). \tag{0.273}$$

We can also use these formulas to calculate $\Omega = \frac{d\phi}{dt}$, which is the angular velocity of the particle **with respect to the coordinate time** t, i.e. as measured by an observer at infinity (since out there $g_{tt} = 1$ and $\Delta\tau = \Delta t$). We have

$$\Omega = \frac{d\phi}{dt} = \frac{d\phi/d\tau}{dt/d\tau} = \frac{\ell/r_+^2}{e/(1 - 2M/r_+)} = \frac{1}{r_+^2}\left(1 - \frac{2M}{r_+}\right)\frac{\ell}{e} = \sqrt{\frac{M}{r_+^3}}\,. \tag{0.274}$$

Example 28: Let's calculate the escape velocity for a radial projectile from some radius R to infinity. We can use the conservation of energy per unit mass $e = (1 - 2M/r)\frac{dt}{d\tau}$. At infinity the proper time τ is just the coordinate time t, so $e = 1$. At some radius R we have

$$e = 1 \implies \mathcal{E} = 0 = \frac{1}{2}v_{\text{escape}}^2 + V_{\text{eff}}(R) \implies v_{\text{escape}} = \sqrt{\frac{2M}{R}}\,, \tag{0.275}$$

where we set $\ell = 0$ since we consider a radial trajectory. Notice that we got the same result as in Newtonian mechanics, because the new term in the effective potential has no effect for $\ell = 0$.

END LEC 11

Precession of the perihelion of Mercury

The second pair of plots in the previous section showed the precession of elliptical elliptical orbits. How do we see it from the equations? We want to compute $\phi(r)$. We can obtain this from the differential equation for

$$\frac{d\phi}{dr} = \frac{d\phi/d\tau}{dr/d\tau} = \frac{\ell/r^2}{\sqrt{2(\mathcal{E} - V_{\text{eff}}(r))}} = \frac{\ell}{r^2}\left[e^2 - \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\ell^2}{r^2}\right)\right]^{-1/2} \tag{0.276}$$

The Newtonian calculation is instead

$$\frac{d\phi}{dr} = \frac{\ell}{r^2}\left[e^2 - \left(1 - \frac{2M}{r}\right) - \frac{\ell^2}{r^2}\right]^{-1/2} \tag{0.277}$$

which integrates to

$$\tan\phi = \frac{M - \ell^2/r}{\ell\sqrt{e^2 - 1 + 2M/r - \ell^2/r^2}} \tag{0.278}$$

Notice the denominator equals $\ell\sqrt{2(\mathcal{E} - V_{\text{eff}}(r))}$, so it vanishes whenever $dr/d\tau$ vanishes, i.e. at the turning points of the orbit. Picking the inner turning point to be r_1 and the outer turning point to be r_2, we have that $\tan\phi$ diverges at $\phi(r_1)$ and then again at $\phi(r_2)$ (and that is the next time it diverges). Successive divergences of the tangent function are spaced out by π, so we find that a full orbit covers twice this much angle (since we counted half the orbit), giving 2π. The orbits of the periods are closed!

Technically, all that remains to solve the problem in general relativity is to do the integral

$$\Delta\phi = 2\ell\int_{r_1}^{r_2}\frac{dr}{r^2}\left[e^2 - \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\ell^2}{r^2}\right)\right]^{-1/2}, \tag{0.279}$$

where we put in the factor of 2 to account for a full orbit (orbit here means roundtrip from an inner radius to an outer radius and back – the final point will not have the same coordinates as the initial point due to the precession). For the orbit to precess we should get $\Delta\phi = 2\pi + \delta\phi_{\text{prec}}$. If there is time, we will discuss the calculation of this integral, which is Problem 15, Chapter 9 of Hartle. The leading correction to the Newtonian answer is

$$\delta\phi_{\text{prec}} = \frac{6\pi G}{c^2}\frac{M}{a(1-\epsilon^2)}, \tag{0.280}$$

where a is the semimajor axis (1/2 the length of the long axis of an ellipse) and ε is the eccentricity ($\varepsilon = \sqrt{1 - b^2/a^2}$ where b is the semi-minor axis, i.e. 1/2 the length of the short axis of an ellipse).

Let's calculate the numerical result from general relativity:

$$\delta\phi_{\text{prec}} = \frac{6\pi(6.7\times 10^{-11}N\cdot m^2/kg^2)}{(3\times 10^8 m/s)^2}\frac{2\times 10^{30}kg}{(5.8\times 10^{10}m)(1-0.21^2)} \approx 5.1\times 10^{-7}. \tag{0.281}$$

This is in radians, per orbit of Mercury. Mercury orbits the Sun once every 88 Earth days = 0.24 Earth years, so we multiply this 100/.24 and by $360°/2\pi$ to get

$$\delta\phi_{\text{prec, century}} = (5.1\times 10^{-7})\,\frac{100}{.24}\,\frac{360°}{2\pi} \approx 0.01° \text{ (!!!)} \tag{0.282}$$

Imagine Einstein's amazement when he calculated the above quantity for Mercury and found this precise agreement.

Notice also the fact that a appears in the denominator means the effect is biggest for planets closest to the Sun, which is why the anomaly was measurable in Mercury's orbit. Intuitively this bigger effect is simply because the strength of the gravitational field is bigger closer to the Sun, so the effects of general relativity will be stronger. The further we get from massive bodies the more the Newtonian approximation becomes valid. (Mercury also has the most eccentric orbit, so ϵ is closest to 1, but this is a smaller enhancement.)

Bending & delay of light

We already predicted the bending of light in gravitational fields using the equivalence principle and an accelerating rocket. We also argued for the (Shapiro) time delay of light simply from an analysis of the lightcones in the weak-field limit (they get narrower near gravitational fields). Both effects are studied in detail for the Schwarzschild metric in Hartle's section 9.4. The key to the analysis is to study null geodesics, in a very similar fashion to the timelike geodesics we studied in the previous section. We still have the conserved quantities

$$e = -\boldsymbol{\xi} \cdot \boldsymbol{u} = \left(1 - \frac{2M}{r}\right)\frac{dt}{d\lambda}\,, \qquad \ell = \boldsymbol{\eta} \cdot \boldsymbol{u} = r^2 \sin^2\theta \frac{d\phi}{d\lambda} \tag{0.283}$$

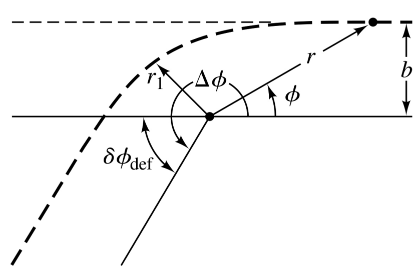
due to the metric's independence of t and ϕ, respectively. λ is some affine parameter. The differential equation for $dr/d\lambda$ comes from $\frac{ds^2}{d\lambda^2} = 0$ or $\boldsymbol{u} \cdot \boldsymbol{u} = 0$, which gives

$$\left(\frac{dr}{d\lambda}\right)^2 = e^2 - \frac{\ell^2}{r^2}\left(1 - \frac{2M}{r}\right)\,. \tag{0.284}$$

To find the deflection of light, we divide

$$\frac{d\phi}{dr} = \frac{d\phi/d\lambda}{dr/d\lambda} = \pm\frac{1}{r^2}\left[\frac{e^2}{\ell^2} - \frac{1}{r^2}\left(1 - \frac{2M}{r}\right)\right]^{-1/2}\,, \tag{0.285}$$

which we can integrate to find $\Delta\phi$. If we come in from infinity at $\phi = 0$ in the equatorial plane $\theta = \pi/2$, then no deflection should correspond to ending up at $\phi = \pi$. So we have $\Delta\phi = \pi + \delta\phi_{\text{def}}$, see below:



where $b \equiv |\ell/e|$. Hartle finds the formula

$$\delta\phi_{\text{def}} = \frac{4GM}{c^2 b} \qquad \text{for } \frac{GM}{c^2 b} \ll 1\,. \tag{0.286}$$

Interestingly, there is a calculation in Newtonian mechanics you can do which gets precisely 1/2 of this answer. Newton actually predicted it. You simply use $v = c$ and assume light undergoes gravitational acceleration according to $GM/r^2 = a$.

Einstein's initial prediction was the same as Newton's, and he asked for astronomers to go look for it. Two astronomers, William Wallace Campbell and Erwin Finlay-Freundlich set off to make the observations of a solar eclipse in early 1914. Campbell went to Kiev while

Freundlich went to Crimea. Soon after, Austrian Archduke Franz Ferdinand was assassinated, an event often marked as the beginning of WWI. As a German, Freundlich was taken as a prisoner of war by the Russians (freed soon after in a prisoner swap), and Campbell had his equipment confiscated (and in any case was thwarted by weather). In 1918 Campbell tried again, this time in Washington state, using equipment from Lick Observatory (just east of San Jose and managed by UCSC!) since his original equipment was confiscated, but the new equipment was not precise enough to measure the deflection. A year later, Sir Arthur Eddington photographed the eclipse and claimed the first confirmation of the deflection of light. Doubt remained due to the accuracy of the observation, but these doubts were laid to rest by Campbell in a 1922 expedition to Australia. These delays were fortuitous for Einstein, who revisited his calculations and revised the estimate for the deflection of light by a factor of 2. The observational confirmation of his theory launched him to instant worldwide fame.

A journalist once asked him what he would do if Eddington's observations failed to match his theory. He replied "Then I would feel sorry for the dear Lord. The theory is correct." Since we've seen how the 43 arcseconds per century were retrodicted by Einstein's theory, we can understand his steadfast confidence.

For the delay of light, we want to calculate

$$\frac{dt}{dr} = \frac{dt/d\lambda}{dr/d\lambda} = \pm b \left(1 - \frac{2M}{r}\right)^{-1} \left[\frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{2M}{r}\right)\right]^{-1/2} \tag{0.287}$$

and integrate to get Δt. This can then be compared to the case with $M = 0$ to figure out Δt_{excess}.

Schwarzschild black hole

Black holes were first predicted by John Michell in 1783. This guy is really underrated: in addition to black holes he argued that earthquakes are seismic waves, the magnetic force falls off as an inverse square law $1/r^2$, and predicted binary star systems.

Using Newton's "corpuscular" theory of light (i.e. light as particle), he calculated when the kinetic energy of light would not be enough to escape the gravitational potential energy well:

$$\frac{GMm}{R} > \frac{1}{2}mc^2 \implies R < \frac{2GM}{c^2}\,. \tag{0.288}$$

He wrote

We could have no information from sight; yet, if any other luminous bodies should happen to revolve about them we might still perhaps from the motions of these revolving bodies infer the existence of the central ones with some degree of probability, as this might afford a clue to some of the apparent irregularities of the revolving bodies...

People began calling these objects "dark stars." As a simple calculation, he pointed out that an object with the mass density of the sun $\rho = \rho_{\rm sun}$ would be dark if its radius was 500 times bigger than that of our sun, $R = 500 R_{\rm sun}$. Check this prediction!

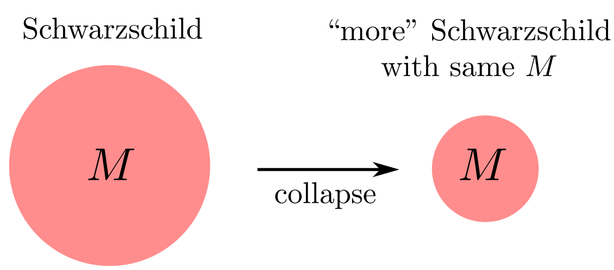
Later Laplace made the same prediction – he was much more famous so he was often credited with the idea.

The modern understanding of black holes, based on Einstein's general theory of relativity, begins with Karl Schwarzschild and his solution to Einstein's equations:

$$ds^2 = -\left(1 - \frac{2GM}{c^2 r}\right) dt^2 + \frac{dr^2}{1 - \frac{2GM}{c^2 r}} + r^2 d\Omega^2\,. \tag{0.289}$$

I put back in the factors of G and c so we can see the special point $r = 2GM/c^2 \equiv R_S$, which appeared in John Michell's calculation of the radius of a "dark star." We will call this the Schwarzschild radius R_S or just $2M$ (setting $G = c = 1$). Let's begin to understand why this geometry also describes a black hole.

Recall that the Schwarzschild geometry describes the solution outside *any* spherically symmetric distribution of matter in terms of the total mass M. This is similar to Coulomb's Law in E&M, which gives the field outside a point charge in terms of total charge Q. For example let's say you compress a star a bit:



If the radius of a spherical mass changes with time, all that happens outside the mass is that we uncover more or less of the Schwarzschild solution (recall that the solution changes once we reach the surface of the body, the pink region above). This happens in nature: a spherically symmetric star collapses when it runs out of thermonuclear fuel. If the star is not sufficiently massive, then it can reach an equilibrium supported by some non-thermal pressure (e.g. white dwarfs and neutron stars are supported by Fermi pressure due to the Pauli exclusion principle). But if the star is sufficiently massive, $M \gtrsim 2M_{\rm sun}$, no known degeneracy pressure can support the object and it keeps collapsing (for white dwarfs it is a bit smaller, 1.4 $M_{\rm sun}$, and is known as the Chandrasekhar limit). Once the radius of the object is less than R_S, it becomes a black hole! The geometry is then given by the Schwarzschild geometry, basically everywhere (as long as we ignore the collapsing matter creating the black hole and the singularity, which we will get to later.)

To see something funny at $r = R_S$, let's investigate our timelike geodesics in the Schwarzschild geometry again. We will focus on radial plunge orbits, $\ell = 0$, for which we have

$$\mathcal{E} \equiv \frac{e^2 - 1}{2} = \frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 + V_{\rm eff}(r) = \frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 - \frac{M}{r}\,. \tag{0.290}$$

Nothing weird happens at $r = R_S$; it is reached (and passed) in finite proper time $\tau = \tau_*$. But what happens in Schwarzschild time t? This is the time measured by an observer at infinity. We have

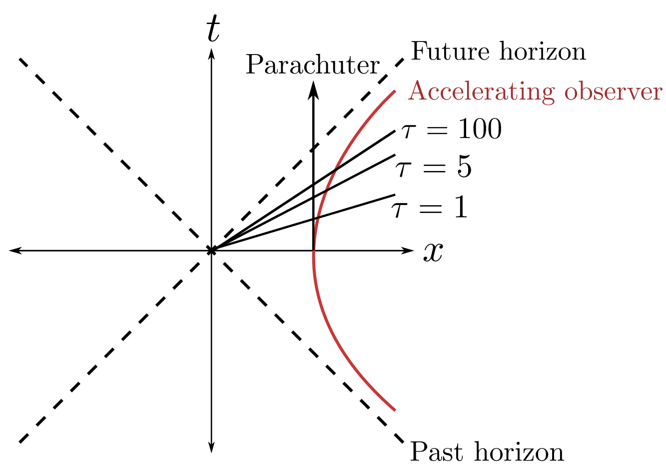
$$e = \left(1 - \frac{2M}{r}\right)\frac{dt}{d\tau} \implies dt = \frac{e d\tau}{1 - \frac{2M}{r}} \,. \tag{0.291}$$

Notice it is singular at $r = 2M$! We can say more about the singularity. We expand $r(\tau)$ around $\tau = \tau_*$: $r = 2M + r'(\tau_*)(\tau - \tau_*) + O(\tau - \tau_*)^2 = 2M + u^r(\tau - \tau_*) + O(\tau - \tau_*)^2$. This gives

$$dt = \frac{e r d\tau}{r - 2M} \approx \frac{2Me}{u^r(\tau - \tau_*)}\left(1 + O(\tau - \tau_*)\right) d\tau \implies t = \frac{2Me}{u^r}\ln(\tau - \tau_*) + \text{ finite at } \tau_* \,. \tag{0.292}$$

So the Schwarzschild time diverges logarithmically. If we're at infinity, we never see our plunging friend cross $r = R_S$, even though in their proper time they sail past the horizon!

Where have we seen this behavior before? Acceleration horizons in flat Minkowski space!



Proper acceleration in Schwarzschild

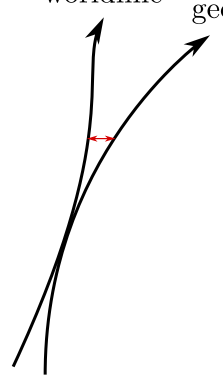
We want to make a sharp comparison between $r = 2M$ (the black hole event horizon) and acceleration horizons. Recall that uniformly accelerating observers with proper acceleration $a = 1/\ell$ stay a proper distance ℓ from the intersection of past and future acceleration horizons. We will see that the same is true for static observers near $r = 2M$ in the Schwarzschild geometry. Let's calculate the acceleration or force felt by an observer at some fixed r, i.e. a static observer.

The acceleration **felt** by an observer is **not** $d^2x^\alpha/d\tau^2$. There are a couple ways to see this. First, it would be zero for a stationary observer at fixed r, θ, ϕ, which is just as wrong as saying we don't feel any acceleration at our fixed position on the Earth's surface. To see this, notice that fixed position means the proper time becomes

$$d\tau = \left(1 - \frac{2M}{r}\right) dt \implies \tau = \left(1 - \frac{2M}{r}\right) t\,, \tag{0.293}$$

so τ is just some constant times t, and we get $d^2x^\alpha/d\tau^2 = 0$ (the non-time components are trivially zero since $dr/d\tau = d\theta/d\tau = d\phi/d\tau = 0$ from stationarity of the observer) . Another way to see this is not the right expression is that the acceleration felt by a *freely falling observer*, i.e. one following a geodesic, should be zero (picture weightless astronauts in a space shuttle), but $d^2x^\alpha_{\text{geodesic}}/d\tau^2 \neq 0$ in general.

Instead, the correct notion of acceleration felt is to compare the acceleration of the observer's worldline *away from its tangent geodesic*. This seems reasonable: we feel a force or acceleration on the Earth's surface because our freely falling frame wants to plunge through, but the surface of the Earth (and its electromagnetic forces) keep us from doing so. Our worldline deviates from its intended geodesic. A general situation is drawn below, with the deviation marked in red.



So the physical acceleration is

$$a^\alpha = \frac{d^2x^\alpha}{d\tau^2} - \frac{d^2x^\alpha_{\text{geodesic}}}{d\tau^2} \tag{0.294}$$

where x_{geodesic} is the tangent geodesic at the point of our worldline we are considering. The geodesic equation gives

$$\frac{d^2x_{\text{geodesic}}}{d\tau^2} = -\Gamma^\alpha_{\beta\gamma}u^\beta u^\gamma \tag{0.295}$$

where since we consider a geodesic tangent to our worldline, the 4-velocity u^α is simply the 4-velocity of our worldline! So we have

$$a^\alpha = \frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}u^\beta u^\gamma\,. \tag{0.296}$$

A few things to note

- a^α measures the amount by which we fail to satisfy the geodesic equation, and therefore deviate from our geodesic, since $a^\alpha = 0$ recovers the geodesic equation.

- In a locally inertial frame $\Gamma^\alpha_{\beta\gamma}\Big|_{x_0} = 0$, so $a^\alpha = d^2x^\alpha/d\tau^2$. In a freely falling frame the geodesic equation is satisfied, so we furthermore have $d^2x^\alpha/d\tau^2 = 0 \implies a^\alpha = 0$.

- Similarly, this definition reduces to our definition of acceleration in special relativity $a^\alpha = d^2x^\alpha/d\tau^2$, where the Minkowski metric in coordinates $ds^2 = -dt^2+dx^2+dy^2+dz^2$ has vanishing Christoffel symbols.

- The same notion (deviation of worldlines from geodesics) defines acceleration in Newtonian physics, but it gives the opposite answers. So your intuition for the above may be shaky! For example, in Newtonian physics we have $a = 0$ for a stationary observer outside a massive body, and $a > 0$ for a radially plunging observer. The reason for the opposite answers is that spacetime is not curved in Newtonian physics (notice we mean ordinary Newtonian physics, not the weak-field limit of general relativity, i.e. "Newtonian Gravity in Spacetime Terms" a la Hartle 6.6); the geodesics are just straight lines in spacetime. So a stationary observer is a geodesic, and a radially plunging one is not! Notice this Newtonian notion does not correspond with the acceleration we feel, which can be quantified by e.g. how much we weigh on a scale.

——————————— END LEC 12———————————

Now that we have our formula in hand, let's compute the acceleration for a static observer in the Schwarzschild geometry. By spherical symmetry, we have $a^\theta = a^\phi = 0$, so our worldline 4-velocity is

$$u^\alpha = \left(\frac{dt}{d\tau}, 0, 0, 0\right) = \left(\frac{1}{\sqrt{1-\frac{2M}{r}}}, 0, 0, 0\right) \tag{0.297}$$

where u^t was fixed by the normalization $u^\alpha u_\alpha = -1$ as usual. The components of our 4-acceleration become

$$a^t = \frac{du^t}{d\tau} + \Gamma^t_{\beta\gamma}u^\beta u^\gamma = \frac{du^t}{d\tau} + \Gamma^t_{tt}u^t u^t\,, \qquad a^r = \frac{du^r}{d\tau} + \Gamma^r_{\beta\gamma}u^\beta u^\gamma = \frac{du^r}{d\tau} + \Gamma^r_{tt}u^t u^t\,. \tag{0.298}$$

where the sum reduced to the tt component since only u^t is nonzero. We calculate

$$\Gamma^t_{tt} = \frac{1}{2}g^{tt}\left(\partial_t g_{tt} + \partial_t g_{tt} - \partial_t g_{tt}\right) = 0\,, \tag{0.299}$$

$$\Gamma^r_{tt} = \frac{1}{2}g^{rr}\left(\partial_t g_{rt} + \partial_t g_{rt} - \partial_r g_{tt}\right) = -\frac{1}{2g_{rr}}\partial_r g_{tt} = \frac{M}{r^2}\left(1-\frac{2M}{r}\right). \tag{0.300}$$

Altogether we get $a^\alpha = (0, M/r^2, 0, 0)$ and the invariant acceleration

$$a \equiv |a| = \sqrt{g_{\alpha\beta}a^\alpha a^\beta} = \sqrt{g_{rr}a^r a^r} = \frac{1}{\sqrt{1-\frac{2M}{r}}}\frac{M}{r^2}\,. \tag{0.301}$$

See Box 12.2 or Example 20.8 of Hartle for different perspectives on this same calculation.

We want to compare this relativistic expression to the "acceleration felt by an observer" in Newtonian physics. This is the acceleration that would be required (say via some rockets) to keep the observer stationary. (Recall that the actual acceleration of a stationary observer in Newtonian physics is 0.) The answer is given by $GMm/r^2 = ma \implies a = GM/r^2 = M/r^2$, where we set $G = 1$ in the end. So the relativistic expression has an enhancement by $1/\sqrt{1-2M/r}$, which diverges at the Schwarzschild radius $r = 2M$. It takes an infinite amount of force to hold up an object as we approach the event horizon. This is reminiscent of the accelerating rocket case approaching the origin! Physically, a star collapsing to $r = 2M$ will not be able to stop collapsing further.

So we can compare:

- Schwarzschild geometry has a special place $r = 2M$, where $a = \frac{1}{\sqrt{1-2M/r}}\frac{M}{r^2} \to \infty$. It takes a finite proper time for Alice to cross but infinite time according to Bob far away.

- Flat geometry according to accelerating observer has a special place $\sigma = 0$, where $a = 1/\sigma \to \infty$. It takes finite proper time for Alice to cross if she jumps out of the accelerating rocket, although it takes infinite time for Bob in the accelerating rocket to see Alice cross.

The expressions for acceleration look very different. This is actually because σ is the proper distance to the horizon, and we haven't measured things in the Schwarzschild case in terms of proper distance. We do that by considering the distance along a radial line at a fixed moment in time,

$$ds = \frac{dr}{\sqrt{1-\frac{2M}{r}}}\,. \tag{0.302}$$

Proper distance from $r = 2M$ is measured by

$$d\sigma = ds = \frac{dr}{\sqrt{1-\frac{2M}{r}}} \implies \sigma = \int_{2M}^{r} \frac{dr}{\sqrt{1-\frac{2M}{r}}}. \tag{0.303}$$

We can do this integral exactly or we can approximate it. Since in the end we will work to leading order in $\Delta = r - 2M$ let's expand around there. We have

$$\frac{1}{\sqrt{1-2M/r}} = \sqrt{\frac{\Delta+2M}{\Delta}} = \sqrt{\frac{2M}{\Delta}}(1+O(\Delta)) \tag{0.304}$$

$$\implies \sigma = \int_0^{\Delta} d\Delta \sqrt{\frac{2M}{\Delta}}(1+O(\Delta)) = 2\sqrt{2M\Delta}(1+O(\Delta)) \tag{0.305}$$

$$\implies \Delta = \frac{\sigma^2}{8M}(1+O(\sigma^2))\,. \tag{0.306}$$

The acceleration can therefore be written as

$$a = \frac{1}{\sqrt{1-\frac{2M}{r}}}\frac{M}{r} = \sqrt{\frac{2M}{\Delta}}\,(1+O(\Delta))\,\frac{M}{(2M)^2} = \frac{4M}{\sigma}\frac{M}{4M^2}(1+O(\sigma^2)) = \frac{1}{\sigma}+O(\sigma). \quad (0.307)$$

We see that we recover the formula $a = 1/\sigma$ with σ the proper distance to the horizon!

Smooth coordinates across the horizon

Is $r = 2M$ in some sense an acceleration horizon? Let's write

$$1-\frac{2M}{r} = \frac{\Delta}{\Delta+2M} = \frac{\Delta}{2M}(1+O(\Delta)) = \frac{\sigma^2}{16M^2}(1+O(\sigma^2)) \quad (0.308)$$

$$\implies r^2 = 4M^2(1+O(\sigma^2)), \quad (0.309)$$

for $\Delta = r-2M \approx \frac{\sigma^2}{8M}$. The metric therefore becomes

$$ds^2 \approx -\frac{\sigma^2}{16M^2}dt^2 + d\sigma^2 + 4M^2 d\Omega^2\,. \quad (0.310)$$

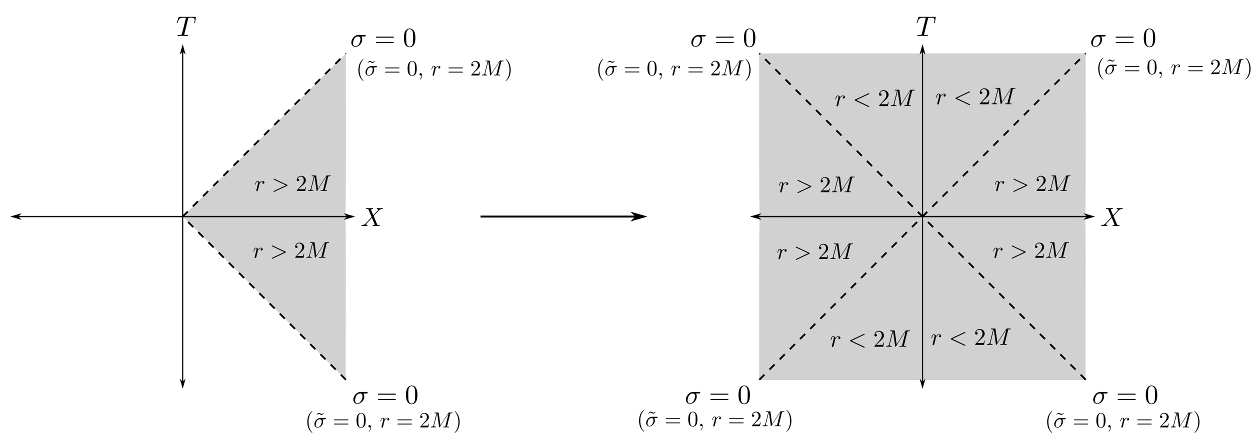
Defining a new time coordinate $\eta = \frac{t}{4M}$ gives

$$ds^2 = \underbrace{-\sigma^2 d\eta^2 + d\sigma^2}_{\text{1+1D Minkowski space}} + 4M^2 d\Omega^2\,, \quad (0.311)$$

where the Minkowski space is written in accelerating coordinates. The horizon $r = 2M$ is now mapped to the acceleration horizon $\sigma = 0$! We can extend these accelerating coordinates past the horizon by recalling ordinary Minkowski coordinates:

$$X = \sigma\cosh\eta\,, \qquad T = \sigma\sinh\eta \qquad \implies \qquad ds^2 = -dT^2 + dX^2 + 4M^2 d\Omega^2\,. \quad (0.312)$$

Why does this extend the coordinates past the horizon? Well, let's plot it.



We have $\sigma^2 = X^2 - T^2$, and we've been plotting the region $\sigma > 0 \implies X^2 > T^2$, which corresponds to $r > 2M$. This is, for example, the shaded quadrant in the figure on the left above. But in these new coordinates there's nothing wrong with $\sigma = 0$ ($r = 2M$) or $\sigma < 0$ ($r < 2M$). That just corresponds to $X^2 < T^2$, and our metric $ds^2 = -dT^2 + dX^2 + 4M^2 d\Omega^2$ is perfectly well-behaved there. More precisely, for $\Delta < 0$, we can write

$$8M\Delta = -\tilde{\sigma}^2 = X^2 - T^2 \tag{0.313}$$

in the Schwarzschild metric and recover (0.312) again. In fact, according to the metric $ds^2 = -dT^2 + dX^2$ there are 4 quadrants, as shown in the right image above.

In this discussion we were working close to $r = 2M$. How do we write down coordinates that cover all of $r > 2M$ and $r < 2M$? Thankfully, there exists a set of coordinates, known as Eddington-Finkelstein coordinates, which do this. We define a new time v through

$$t = v - r - 2M \log \left| \frac{r}{2M} - 1 \right| . \tag{0.314}$$

Plugging this into the Schwarzschild metric gives

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dv^2 + 2dv\, dr + r^2(d\theta^2 + \sin^2\theta d\phi^2) . \tag{0.315}$$

This is the metric you get *regardless of whether* $r < 2M$ *or* $r > 2M$. And it doesn't diverge when $r = 2M$ like with Schwarzschild coordinates, it simply has the g_{vv} component vanish.

Hartle discusses an exact coordinate system, Kruskal coordinates, which cover all 4 quadrants of the entire Schwarzschild spacetime (so, including higher order terms in $\sigma = r - 2M$ which we dropped).

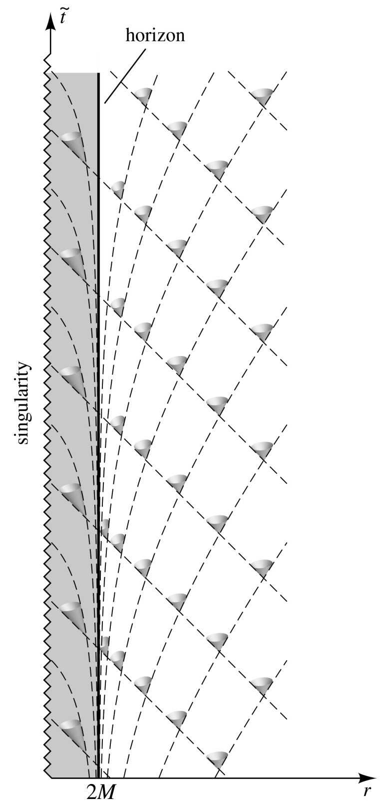
So $r = 2M$ is just a *coordinate singularity*. How could we have known that it is actually well-behaved without this long calculation? The determinant of the metric! It gives $r^4 \sin^2\theta$, indicating that $r = 2M$ is totally fine. We will soon come to a better diagnostic, the curvature. It will indicate for us that the region $r = 0$, which is problematic both in Schwarzschild and Eddington-Finkelstein coordinates (g_{tt} and g_{vv} diverge), is a true physical singularity, i.e. a place where the curvature diverges.

Tipping of light cones

Let's calculate what happens to the light cones as we approach $r = 2M$. This will illustrate in another way that nothing can escape a black hole. Restricting to $\theta = \phi = const.$ gives

$$ds^2 = 0 = -\left(1 - \frac{2M}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} \implies \frac{dr}{dt} = 1 - \frac{2M}{r} . \tag{0.316}$$

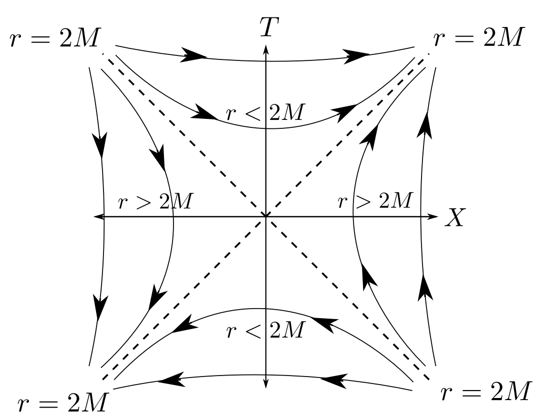
We see that as we approach $r = 2M$ the light cone is totally closing up, indicating that it is becoming impossible to escape the pull of the black hole. But since the Schwarzschild coordinates are singular at $r = 2M$ we should use better coordinates, like the Eddington-Finkelstein ones. Hartle does this and produces the plot below, where $\tilde{t} = v - r$.



The plot shows that for $r < 2M$ all light rays are doomed to fall toward $r = 0$.

Full Schwarzschild geometry

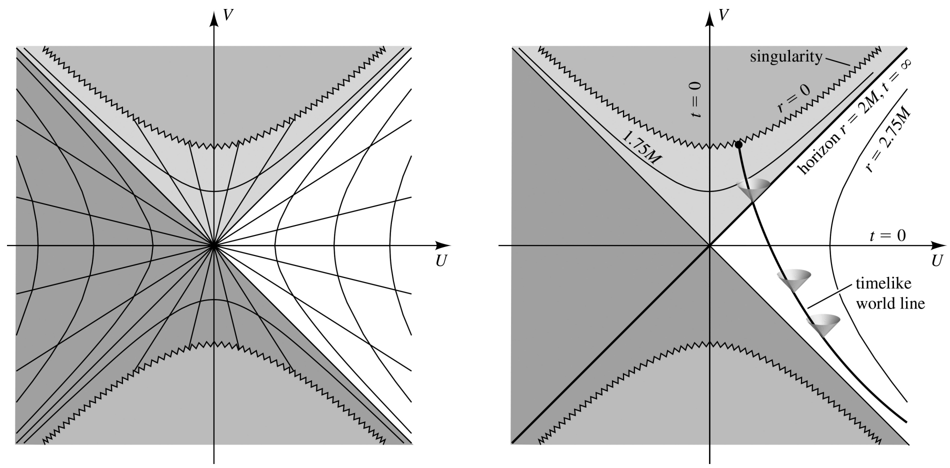
The geometry in the various patches look as below:



The hyperbolas that are drawn are at constant r values, and the arrows indicate the direction of increasing t. This is a symmetry since $t \to t + const.$ leaves the metric unchanged. Recall that we saw this same structure for Lorentz boosts in special relativity! They left invariant the hyperbolic worldline of an accelerating observer, and acted as $\theta = a\tau \to a(\tau + const.)$.

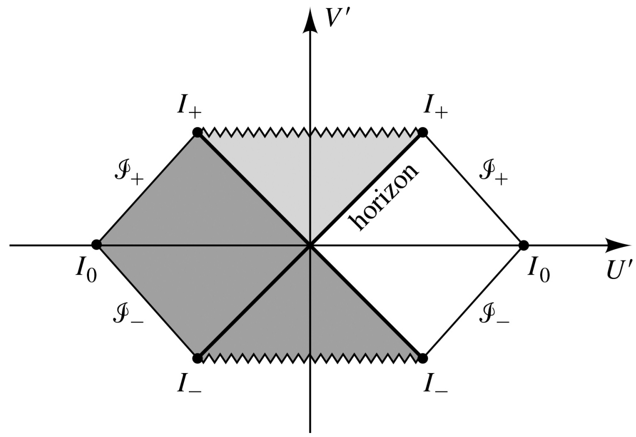
Notice also that time translation is *spacelike* for $r < 2M$! This is weird, but agrees with the Schwarzschild metric, for which g_{tt} becomes positive and g_{rr} becomes negative once we go into the region $r < 2M$. For this reason, the full solution is not static: the r-dependence of the metric is actually a time dependence in the region $r < 2M$.

Even though in this sense r is a crummy name for the coordinate, it is always the radius of the transverse sphere $r^2 d\Omega^2$ in the metric. To analyze this in more detail, we would like the full spacetime geometry, covering all four quadrants for arbitrary r (i.e. not just for $r - 2M$ small, which is most of what we focused on above). A good set of coordinates on that geometry are given by Kruskal-Szekeres coordinates, which lead to the following diagram



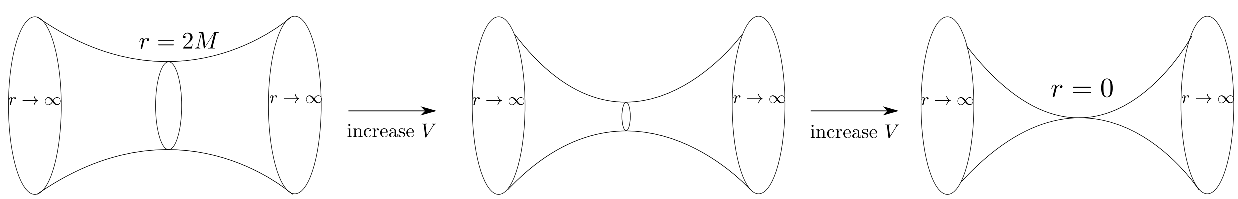
On the left curves of constant Schwarzschild r and t are drawn. There are singularities at the jagged curves $r = 0$ (we will get to this shortly). On the right an infalling observer is drawn, with local light cones which don't tip over. This is because in these coordinates, the horizon is at 45°, so even if the light cones don't tip over, as soon as the infalling observer crosses $r = 2M$ they can't get back out. A stationary observer at $r = 2.75M$ is also drawn.

These are useful coordinates to construct the Penrose diagram of the black hole, which is done in Box 12.5:



These diagrams help us understand the global structure of the solution. Notice the edges of the diagram, besides the singularity, are just the pieces of the flat-space Penrose diagram we've seen before. But now there are *two* copies of the various infinities! What is going on? In fact, an analysis of the Kruskal coordinates, or just an extrapolation of the local analysis we did above, implies that there are two exteriors of the black hole! These are the two regions denoted by $r > 2M$ in the diagrams above. Alice, in the (white) diamond on the right, sees a black hole with a future horizon (and a white hole with a past horizon). Bob, in the (gray) diamond on the left, also sees a black hole with a future horizon (and a white hole with a past horizon). There is both a final singularity ($r = 0$ to the future) and an initial singularity ($r = 0$ to the past).

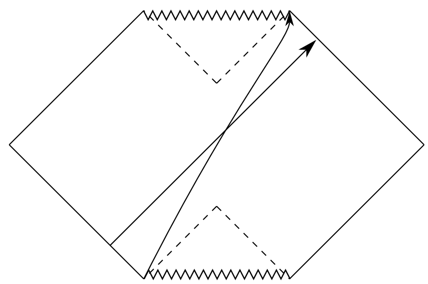
So what we actually have is **2 universes connected by a wormhole**. This wormhole can be seen on the constant-V slices of the Kruskal diagram above. At $V = 0$ it has a minimum size of $r = 2M$, and as V increases or decreases, this minimum size decreases and eventually hits $r = 0$ at the singularity: the wormhole closes off!



If we start from $V = -\infty$, then the wormhole begins with zero size, grows to a max size of $r = 2M$ at the waist, and then collapses back down to zero size. It turns out that it closes off so quickly that you cannot traverse it. This can also be seen from the Penrose diagram above, since light rays are at 45° in Penrose diagrams and no light ray can get from the right diamond (in white) to the left diamond.

——————————— END LEC 13———————————

Is it possible for a wormhole like this one to be traversable? We would need a Penrose diagram that looked like the one below:



A null and timelike trajectory are displayed that go from the universe on the left to the universe on the right. Such geometries are not possible in general relativity without introducing a source of negative energy. Such sources are allowed in quantum mechanics. In that case, the wormhole can become traversable!

Let's return to the Schwarzschild black hole, with its non-traversable wormhole. What is going on at $r = 0$ (say in the future)? First notice that observers reach $r = 0$ in finite proper time. For simplicity let's consider radial geodesics (i.e. $\ell = 0$), we have

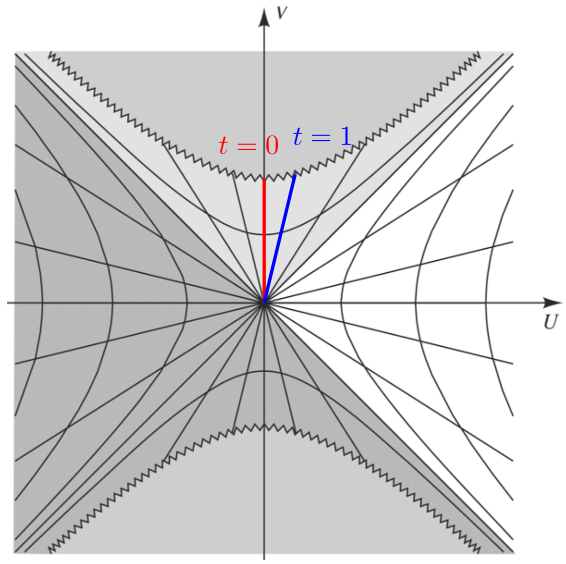
$$\mathcal{E} = \frac{1}{2}\left(\frac{dr}{d\tau}\right)^2 - \frac{M}{r}\,. \tag{0.317}$$

This can be integrated to obtain $\tau(r)$. The exact expression is a little messy, but if we measure the time from $r_0 \ll 1$ to $r = 0$ then we can ignore the constant $\mathcal{E}$ and it simplifies to

$$\Delta\tau \approx \frac{\sqrt{2}}{3}\frac{r_0^{3/2}}{\sqrt{M}}\,. \tag{0.318}$$

This is clearly finite. The expression with finite $\mathcal{E}$ can be checked and is also finite. Geodesics with $\ell \neq 0$ have a finite proper time to the singularity as well.

Now let's jump into the black hole with a friend. We will follow $t = 0$ and our friend will follow $t = 1$. Recall that t is a *spatial* coordinate for $r < 2M$, see below.



The proper distance between these two trajectories, *at a given moment in time* r^*, is

$$\int d\sigma = \int_0^1 \sqrt{g_{tt}}\, dt = \int_0^1 \sqrt{-(1-2M/r^*)}dt \implies \Delta\sigma = \sqrt{-(1-2M/r^*)}\,. \qquad (0.319)$$

As $r^* \to 0$ we see that the proper distance $\Delta\sigma \to \infty$. Similarly, the relative acceleration diverges as well. This means that if you tried to tie two rocks together with a piece of rope, one on the trajectory $t = 0$ and one on $t = 1$, then the rope would tear as you approach $r = 0$. In fact, any physical object will have a finite tensile or breaking strength, and will be ripped apart as it approaches $r = 0$. On the other hand, the transverse sphere $r^2 d\Omega^2$ is shrinking to zero size, so there are infinte crushing forces in those directions. Yikes! This is why this region is called a true singularity, or a physical singularity, or a curvature singularity (we'll get to curvature singularities later), and what happens to objects that fall in is called "spaghettification." Compare this to the event horizon of a black hole, which we already discussed is a coordinate singularity. For an incredibly massive black hole, the gravitational force or acceleration at $r = 2GM$ can be very weak: $F = GM/r^2 \sim GM/(2GM)^2 \sim 1/(GM) \to 0$ as $M \to \infty$.

Finally, notice that $r = 0$ is a "moment in time" not a "location in space." This is often conflated in discussions (especially popular ones) of the black hole. The singularity is **not** at the center of the black hole in any useful sense. If it were at the center, then once you cross the horizon you should be able to accelerate outward to maximize your proper time (i.e. your life) before hitting $r = 0$. But a calculation shows that any such acceleration only shortens your proper lifetime.

What did Einstein think?

Einstein hated black holes. More precisely, he hated the curvature singularity of black holes. That's reasonable, because he knew they implied a breakdown in his theory. In his usual style, he came up with a clever argument to try to disprove their existence. He assumed that a black hole can be formed quasi-statically, e.g. by slowly compressing a bunch of matter circling around some common center of mass. He assumed that at each point in the compression there should be some equilibrium configuration. The conclusion was that the matter would have to exceed the speed of light well before the Schwarzschild radius is reached, and therefore the formation of a black hole would be impossible. The error was that black holes are not formed through stages of equilibrium configurations: it is a chaotic implosion of matter.[12]

Here is the introduction to his paper:

ANNALS OF MATHEMATICS
Vol. 40, No. 4, October, 1939

ON A STATIONARY SYSTEM WITH SPHERICAL SYMMETRY CONSISTING OF MANY GRAVITATING MASSES

BY ALBERT EINSTEIN

(Received May 10, 1939)

If one considers Schwarzschild's solution of the static gravitational field of spherical symmetry

$$ds^2 = -\left(1 + \frac{\mu}{2r}\right)^4 (dx_1^2 + dx_2^2 + dx_3^2) + \left(\frac{1 - \frac{\mu}{2r}}{1 + \frac{\mu}{2r}}\right)^2 dt^2 \tag{1}$$

it is noted that

$$g_{44} = \left(\frac{1 - \frac{\mu}{2r}}{1 + \frac{\mu}{2r}}\right)^2$$

vanishes for $r = \mu/2$. This means that a clock kept at this place would go at the rate zero. Further it is easy to show that both light rays and material particles take an infinitely long time (measured in "coördinate time") in order to reach the point $r = \mu/2$ when originating from a point $r > \mu/2$. In this sense the sphere $r = \mu/2$ constitutes a place where the field is singular. (μ represents the gravitating mass.)

There arises the question whether it is possible to build up a field containing such singularities with the help of actual gravitating masses, or whether such regions with vanishing g_{44} do not exist in cases which have physical reality.

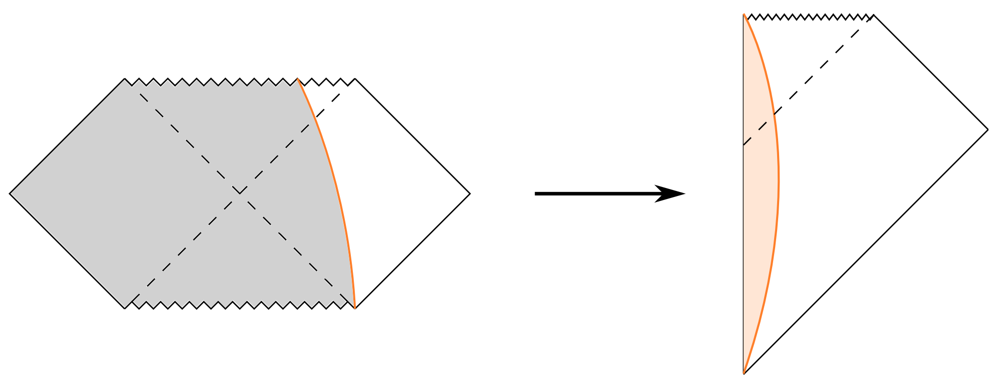
He concludes the paper with...

[12]In fact, almost simultaneously, Oppenheimer and his graduate student Hartland Snyder published a paper showing that when a star runs out of fuel it collapses indefinitely (via an out-of-equilibrium process), although they didn't explicitly identify the end state as Schwarzschild's solution.

> The essential result of this investigation is a clear understanding as to why the "Schwarzschild singularities" do not exist in physical reality. Although the theory given here treats only clusters whose particles move along circular paths it does not seem to be subject to reasonable doubt that more general cases will have analogous results. The "Schwarzschild singularity" does not appear for the reason that matter cannot be concentrated arbitrarily. And this is due to the fact that otherwise the constituting particles would reach the velocity of light.

Black holes from gravitational collapse

These black holes, like the ones in nature, are **not** described by the full Schwarzschild geometry, with its two universes. Instead, there is just one exterior. The geometry can be obtained by patching one of the two exteriors of the full Schwarzschild geometry to some other geometry that describes the collapsing star (this geometry is too difficult to solve for exactly!) In the picture below, this corresponds to excising the gray region on the left, and replacing it with something that roughly looks like what's on the right.



What happens when additional matter falls into a black hole? Since $r = 2M$, this means that the black hole must *grow*. This means that to identify the true event horizon of the black hole, we have to know the entire future of the universe. Otherwise, say you make a claim as to where the black hole event horizon will be. I can always show up and throw a bunch of mass into the black hole to increase the event horizon beyond where your prediction was. For this reason, the event horizon is often referred to as "teleological." telos = $\tau\acute{\varepsilon}\lambda o\varsigma$ is a Greek word meaning completion or end.

A view from quantum gravity

The fact that matter falling into a black hole leads to its area increasing is a special case of a general feature: the area of a black hole never decreases with time (classically). This is known as Hawking's Area Theorem. For example, when two black holes collide, they merge to form a larger black hole whose area must be at least the sum of the areas of the two original black holes.

The Hawking Area Theorem sounds like the 2nd Law of Thermodynamics, which says that entropy cannot decrease with time. Jacob Bekenstein was the first person to take the analogy seriously. Motivated by this and the fact that you can violate the 2nd Law of Thermodynamics by dropping a system with entropy into a black hole, he argued that black holes must have an entropy proportional to their area. Stephen Hawking disliked the idea and tried to disprove it. Instead, by bringing in quantum mechanics he discovered that black holes have an entropy given by

$$S = \frac{k_B c^3 A}{4G\hbar}\,. \tag{0.320}$$

This formula is beautiful for two reasons. The first is that the entropy scales with the area of the black hole instead of extensively with the volume. This is the conceptual root behind holography or AdS/CFT. The second is that it combines all of the fundamental constants in physics: Boltzmann's constant k_B from statistical mechanics, the speed of light c from special relativity/electromagnetism, G from gravity, and $\hbar$ from quantum mechanics.

Because of the fundamental constants this is a huge amount of entropy! The sun has roughly $S \sim 10^{58}$ coming from thermal entropy. A black hole of the same mass of the sun would be much smaller and have $S \sim 10^{78}$!

But if black holes have an entropy in analogy to thermodynamics, and they have a mass as we've discussed, then they must have a temperature (and therefore radiate)! We can see this from the 1st Law of Thermodynamics:

$$dE = TdS\,. \tag{0.321}$$

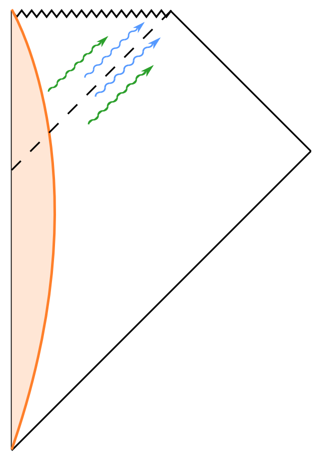
Using $E = M$ and $dS = dA/4$ with $A = 4\pi R_S^2 = 16\pi M^2$ gives

$$dM = 8\pi T M dM \implies T = \frac{1}{8\pi M} = \frac{\hbar c^3}{8\pi k_B G M}\,. \tag{0.322}$$

This is incredibly cold. A black hole with the mass of the sun would have temperature $T = 6\times 10^{-8}$ Kelvin. And they get colder as they get bigger/more massive. This is unconventional from the statistical mechanics perspective, since it means a negative specific heat.

Because of the positive temperature $T > 0$, black holes are radiating a gas of particles at that temperature. This means they are not entirely black, and as they radiate particles away they have to lose energy, so they shrink and eventually evaporate away (as long as we don't feed it any additional matter). This is a quantum-mechanical violation of Hawking's Area Theorem, and Hawking himself discovered it. It is therefore known as Hawking radiation.

Even though classically nothing can get out of a black hole, we are used to the fact that quantum mechanics allows classically forbidden processes, for example tunneling through a potential barrier. In quantum mechanics, the vacuum state of the theory is a highly complicated place. It has particles being pair-produced into existence, for example an electron and a positron, and then annihilating one another and therefore disappearing again. This happens near the horizon of a black hole as well, except it occurs with one particle on the outside of the horizon and one on the inside. A picture is below:



The blue particles were pair produced and the green particles were pair produced. The pairs are actually highly quantum-mechanically entangled with one another. An observer outside the black hole sees the black hole emitting particles, so must see the black hole losing mass and therefore area. So the black hole shrinks to zero size and disappears.

The black hole information paradox is mostly sharply stated in terms of the von Neumann entropy of the observable part of the universe. You can begin with a system in a quantum-mechanical pure state, say a star, and have it collapse to form a black hole. Then the Hawking process begins and spits out these Hawking quanta, which are entangled with stuff inside the black hole. But then the black hole evaporates away! This seems to imply that what is left – the exterior Hawking quanta – is in a mixed state. You have thus evolved from a pure quantum state into a mixed quantum state, which cannot be done by unitary evolution. This is a violation of quantum mechanics, and in particular it says that the information of whatever fell into the black hole is lost forever. Understanding how quantum mechanics is preserved, and how the information that falls into a black hole actually comes out, is the black hole information paradox. It has been a guiding light in theoretical physics for almost 50 years.

The physical or curvature singularity is a fancy name that hides the fact that **we don't know what's going on there.** The theory is breaking down. The belief is that a more complete theory should "resolve the singularity," i.e. make reasonable, finite predictions for what is going on. It is believed that a quantum theory of gravity should be able to resolve the singularity. Some progress has been made on this question from the perspective of string theory, but we still do not have a satisfactory picture of what is happening to spacetime at the singularity.

String theory has provided another interesting perspective on the full Schwarzschild geometry, with its two universes connected by a wormhole. The perspective is that this has an equivalent description as two universes which are highly entangled quantum-mechanically.

The wormhole is a connection between the two universes in a similar way to how entanglement is a connection between two particles. The fact that the wormhole is not traversable is just the statement that quantum entanglement does not permit superluminal propagation. What if we introduce negative energy as in quantum mechanics, making the wormhole traversable? This actually requires having the two universes directly interact with each other. In turn, sending signals from one universe to the other then corresponds to quantum teleportation, a protocol that transmits quantum information through entanglement!

—————————— END LEC 14——————————

Einstein's equations

So far we have just focused on *solutions* in general relativity without discussing where they came from. We will now discuss where they come from: Einstein's equations.

The qualitative idea is that mass/energy acts as a source for spacetime curvature. Without any mass or energy we will simply have flat Minkowski space, and adding masses will lead to the appearance of a more general curved metric or line element.

Riemann curvature tensor

Before we get to the equations we have to quantitatively discuss curvature, which we have only discussed qualitatively so far. A good measure is the relative acceleration between two nearby freely falling observers. Intuitively this is like measuring what parallel lines do on your geometry: in flat space they remain parallel, so there is no curvature. On a sphere they will converge (picture lines of longitude), so it is positively curved. On a hyperbola they will diverge, so it is negatively curved.

The notion of curvature we will be discussing is known as *intrinsic curvature*. It is intrinsic to the geometry itself, not in how it may be embedded in a higher-dimensional space. For example take a piece of paper and roll it up into a cylinder. This cylinder sure looks curved, but that notion is known as *extrinsic curvature*; the paper wasn't intrinsically curved before you rolled it up and it still isn't afterward. To convince yourself of this draw a pair of parallel lines on the cylinder – they remain parallel!

To make this notion a quantitative measure of curvature we pick a spacetime point x_0, a 4-velocity for the observers (they will have the same one since they are initially parallel geodesics), and the observer's separation Δ^α which will be infinitesimally small. We want to compute the relative acceleration, so we will use the geodesic equation

$$\frac{d^2x^\alpha}{d\tau^2} = -\Gamma^\alpha_{\beta\gamma}u^\beta u^\gamma \,. \tag{0.323}$$

The rough idea is that we want to subtract the acceleration of one observer from the other one, like we did when calculating the proper acceleration in Schwarzschild. The difficulty is that these are vectors defined at different points in spacetime. This isn't an issue in flat space, where we are used to moving vectors around in the plane, keeping their orientation and magnitude fixed. But how do we move vectors around on, say, a potato? The vectors are really defined in the tangent space as we discussed before, and the tangent space is different at each point. The correct mathematical concept for moving vectors around on potatos (or more general manifolds) is that of parallel transport. Let's ignore the details for a moment and Taylor expand the acceleration at $x_0^\alpha + \Delta^\alpha$ (the location of one observer) around x_0^α (the location of the parallel observer):

$$\left.\frac{d^2x^\alpha}{d\tau^2}\right|_{x^\alpha = x_0^\alpha + \Delta^\alpha} = \left.\frac{d^2x^\alpha}{d\tau^2}\right|_{x^\alpha = x_0^\alpha} + \Delta^\beta C^\alpha_\beta(x_0) + O(\Delta^2)\,. \tag{0.324}$$

Comparing to the geodesic equation we can guess that C^α_β should be something that depends on the 4-velocity u^β and the Christoffel symbol. In fact the result is

$$a^\alpha_{\text{relative}} = \Delta^\beta C^\alpha_\beta(x_0) = R_{\gamma\beta\delta}{}^\alpha u^\gamma \Delta^\beta u^\delta + O(\Delta^2)\,. \tag{0.325}$$

This relative acceleration is an honest "contravariant" 4-vector (contravariant means the index is upstairs, while covariant means the index is downstairs) which transforms just like u^α. We will get to why the raised index of $R_{\gamma\beta\delta}{}^\alpha$ is off to the right, but it's basically so we know that the 4th index is the one that's raised (in previous cases it didn't matter for us which index was raised). This object is known as the Riemann curvature tensor, and it is pretty gnarly:

$$R_{\gamma\beta\delta}{}^\alpha = \partial_\beta \Gamma^\alpha_{\gamma\delta} - \partial_\gamma \Gamma^\alpha_{\beta\delta} + \Gamma^\mu_{\delta\gamma}\Gamma^\alpha_{\beta\mu} - \Gamma^\mu_{\delta\beta}\Gamma^\alpha_{\gamma\mu}\,. \tag{0.326}$$

Remember since $\Gamma^\alpha_{\beta\gamma}$ involves first derivatives of the metric, $R_{\gamma\beta\delta}{}^\alpha$ involves second derivatives. This is to be expected from our parameter-counting discussion where we tried to fix the metric to be Minkowski in the neighborhood of a point. We were able to fix the metric at a point and fix the first derivatives to vanish via appropriate coordinate transformations, but we did not have enough freedom from the coordinate transformations to also fix the second derivatives to zero. This suggests that they are physical, and indeed, they precisely constitute the Riemann curvature tensor.

Another way to think about the Riemann curvature tensor is that it measures how much the orientation of tangent vectors changes as you transport them around a closed path on the geometry. (The rule for how to transport them is known as "parallel transport" but it amounts to what you would do intuitively, see e.g. the figure here.)

The Riemann curvature tensor has a set of symmetries, which is simplest to discuss in the object

$$R_{\alpha\beta\gamma\delta} := R_{\alpha\beta\gamma}{}^\lambda g_{\lambda\delta}\,. \tag{0.327}$$

The symmetries are as follows:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}\,, \qquad R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma}\,, \qquad R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}\,. \tag{0.328}$$

In words, the tensor (with all indices down) is antisymmetric in the first two indices, antisymmetric in the last two indices, and symmetric under swapping the first two indices with the last two indices. We also have the (1st) Bianchi identity

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} = 0\,. \tag{0.329}$$

These symmetries reduce the number of independent components from 4^4 to 20, which is precisely what we got from our counting argument several sections ago. There is a particularly useful "contraction" (meaning take independent indices and make them the same to sum) of the Riemann curvature tensor:

$$R_{\beta\gamma} \equiv R_{\beta\alpha\gamma}{}^{\alpha} = R_{\beta\alpha\gamma\lambda}g^{\lambda\alpha}\,. \tag{0.330}$$

This is known as the Ricci curvature tensor. It is the only interesting contraction of the Riemann tensor. Another contraction one might try doesn't give anything interesting:

$$R_{\beta\gamma\alpha}{}^{\alpha} = R_{\beta\gamma\alpha\lambda}g^{\lambda\alpha} = -R_{\beta\gamma\lambda\alpha}g^{\lambda\alpha} = -R_{\beta\gamma\alpha\lambda}g^{\alpha\lambda} \implies R_{\beta\gamma\alpha}{}^{\alpha} = 0\,, \tag{0.331}$$

where the minus sign in the second equality appears due to antisymmetry, and then we relabel the dummy indices in the third equality $\alpha \leftrightarrow \lambda$ to end with the expression equaling the negative of itself (and therefore it equals zero). Contracting the first and fourth indices also doesn't give anything new:

$$R_{\alpha\gamma\delta}{}^{\alpha} = -R_{\gamma\alpha\delta}{}^{\alpha} = -R_{\gamma\delta} \tag{0.332}$$

due to the antisymmetry property of the first two indices.

The Ricci tensor is symmetric in its indices:

$$R_{\alpha\beta} = R_{\alpha\gamma\beta}{}^{\gamma} = R_{\alpha\gamma\beta\delta}g^{\gamma\delta} = R_{\beta\delta\alpha\gamma}g^{\gamma\delta} = R_{\beta\delta\alpha}{}^{\delta} = R_{\beta\alpha}\,. \tag{0.333}$$

We can further contract the Ricci tensor to obtain the Ricci scalar

$$R = R^{\alpha}_{\alpha} = R_{\alpha\beta}g^{\alpha\beta}\,. \tag{0.334}$$

This is an **invariant** measure of the curvature, meaning it is the same regardless of which coordinate system you use to compute it (tensors on the other hand depend on coordinates). It is the simplest invariant, but not the only one, for example you could also consider

$$R_{\alpha\beta}R^{\alpha\beta} = R_{\alpha\beta}R_{\gamma\delta}g^{\gamma\alpha}g^{\delta\beta}\,. \tag{0.335}$$

In 4 dimensions, the Ricci tensor has $4 \times 5/2$ independent components and the Ricci scalar only has 1 component, so they do not have complete information about the curvature. For this we need the full Riemann tensor, with its 20 independent components. In particular, $R_{\alpha\beta\gamma\delta} = 0$ implies that we have zero curvature and therefore spacetime is flat, but $R = 0$ or even $R_{\alpha\beta} = 0$ does not imply this.

Einstein-Hilbert action

We are finally ready to write down Einstein's equations without any sources:

$$\boxed{R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 0\,.} \tag{0.336}$$

Without sources means that there is no mass or energy anywhere in the spacetime. The object on the left-hand-side is called the Einstein tensor, but it is a simple combination of the Ricci tensor we introduced above and the metric tensor.

If we introduce mass or energy, then it comes with its own tensor known as the stress-energy tensor $T_{\mu\nu}$, and the equations generalize to

$$\boxed{R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 8\pi G T_{\mu\nu}\,.} \tag{0.337}$$

How can we understand these equations? A familiar way to derive equations of motion is to start from a Lagrangian or its time integral, the action, and use the variational principle. Let's recall a few actions you may have seen in your studies. We have discussed the action for free particles

$$S = -m \int d\tau\,, \tag{0.338}$$

which integrates over the worldline of the particle. This gives the dynamics of the particle. The action for spacetime should give the *dynamics of spacetime itself*, so should be an integral over the 4d spacetime. This is not so different than, for example, the Maxwell action, or the action for any classical field theory. Those fields, like the metric, are defined over space and time, and therefore the action is an integral over spacetime. (In Maxwell theory it is $S = \int d^4x F_{\mu\nu}F^{\mu\nu}$.)

But what should the integrand (or Lagrangian density) be for general relativity? Actions should share the symmetries of our theory. For example, the Maxwell action is invariant under Lorentz transformations. We know this because the integrand is a Lorentz scalar. What are the symmetries of general relativity? Or in other words, what operation leaves the physics unchanged? The key one is coordinate transformations! These change the form of the metric and all the other tensors we've been playing with, but the physics should not change. So the action of general relativity should be invariant under coordinate transformations. This means it should be defined in terms of invariant geometric quantities. Integrating a constant $\int d^4x k$ doesn't work, since the coordinate transformation $x \to ax$ changes the answer. But if we instead integrate to get the 4-volume of spacetime

$$V_4 = \int \sqrt{-\det(g)}\, d^4x \equiv \int \sqrt{-g}\, d^4x\,, \tag{0.339}$$

then the resulting object is invariant under a change of coordinates! However d^4x changes,

$\sqrt{-g}$ will change in a compensating way.[13] This is a 4-dimensional version of the integral you might do to calculate the 2-dimensional area of a sphere. It doesn't matter whether you do it in spherical coordinates or Cartesian coordinates, the answer will always be $4\pi R^2$, since it is an invariant geometric quantity. Requiring our action to be invariant under coordinate changes means we can write our action as

$$S = \int \mathcal{L}\sqrt{-g}\, d^4x\,, \tag{0.340}$$

where the Lagrangian density $\mathcal{L}$ is a scalar, i.e. invariant under coordinate changes. Another important requirement for our action is that the physics be local. For example, whatever $\mathcal{L}$ is a function of, it should be a function of those objects evaluated at *the same spacetime point*. This is the case in, for example, the Maxwell action above.

——————————————— END LEC 15———————————————

We have many scalar functions we can make from these at the same spacetime point. The simplest option is the Ricci scalar R, but any function $f(R)$ is valid, as is $R_{\alpha\beta}R^{\alpha\beta}$, and more. The Lagrangian density could involve all of these. It shouldn't involve inverse powers like $1/R$ though because these would diverge for flat space, where $R_{\alpha\beta\gamma\delta} = 0$.

Let's say we consider the most general action, expanded in a series of powers of contractions of $R_{\alpha\beta\gamma\delta}$:

$$\mathcal{L} = A + BR + CR^2 + DR_{\alpha\beta}R^{\alpha\beta} + ER_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} + \cdots\,. \tag{0.341}$$

In the limit where spacetime is weakly curved, terms with the least powers of $R_{\alpha\beta\gamma\delta}$ dominate. The least power is zero, so that would give

$$S^{(0)} = \int A\sqrt{-g}d^4x\,. \tag{0.342}$$

The equations of motion that follow from the variational principle are

$$\delta S^{(0)} = 0 \implies \sqrt{-g}g^{\alpha\beta} = 0\,. \tag{0.343}$$

These are not interesting equations of motion: the only solution is the "zero" metric. Let's instead look at the next term:

$$S^{(1)} = B\int d^4x\sqrt{-g}\,R\,. \tag{0.344}$$

[13] To be more precise, let's consider a 1d metric $ds^2 = dx^2$ and the change of coordinates $x = 2y$. We consider the expression $\int_a^b dx\,k$. This means that if we want to write it in y coordinates we would have to write $\int_{a/2}^{b/2} dy\,k$, which isn't equal to the original expression. Of course, performing the coordinate change $x = 2y$ to the original expression gives $\int_{a/2}^{b/2} 2dy\,k$, which is correct, but that is what inputting $\sqrt{g}$ is meant to capture. In that case we have $ds^2 = dx^2 = 4dy^2$, so $\int_a^b dx\sqrt{g_{xx}} = \int_{a/2}^{b/2} dy\sqrt{g_{yy}} = \int_{a/2}^{b/2} 2dy$.

This action has sensible Euler-Lagrange equations:

$$\delta S^{(1)} = 0 \implies R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = 0 \,. \tag{0.345}$$

These are the Einstein equations without source (also known as the vacuum Einstein equations) we wrote down before! It also has second derivatives of the dynamical field (the metric), as good equations of motion often do (e.g. $F = m\ddot{x}$). Notice also that B dropped out of the equations because it was an overall factor in front of the action, and such factors don't affect the equations of motion. This action is known as the Einstein-Hilbert action, and is written with normalization

$$S_{\mathrm{EH}} = \frac{1}{16\pi G} \int d^4x \sqrt{-g}\, R \,. \tag{0.346}$$

Flat Minkowski space has $R_{\alpha\beta} = 0 = R$ and so is a solution of the vacuum Einstein equations. The Schwarzschild black hole is also a solution of these equations, although in that case $R_{\alpha\beta}$ and R are nonzero, but the Einstein tensor $R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}$ is zero.

Adding a cosmological constant source

We argued that the action $S^{(0)}$ led to uninteresting equations of motion. But now that we have a more interesting action, we can dress it up with this constant. Let's pick a normalization $A = -\Lambda/(8\pi G)$ and consider the action

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g}\, (R - 2\Lambda) \,. \tag{0.347}$$

This leads to equations of motion

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} + \Lambda g_{\alpha\beta} = 0 \,. \tag{0.348}$$

These are the vacuum Einstein equations with a **cosmological constant** Λ. Notice as before that G drops out since it is an overall constant in the action. If $\Lambda \neq 0$, then flat Minkowski space is no longer a solution. Because Λ is a parameter in the action, it parameterizes different theories. Einstein set $\Lambda = 0$ initially to obtain flat Minkowski space as a solution.

The universe we live in is not Minkowski space. The "dark energy" era we are entering is best modeled by $\Lambda > 0$ and positive curvature. The simplest solution to the equations above with $\Lambda > 0$ is de Sitter space:

$$ds^2 = -dt^2 + e^{2Ht}(dx^2 + dy^2 + dz^2) \,. \tag{0.349}$$

The time-dependence indicates that the (potentially infinite) space is expanding exponentially. In fact, the expansion is accelerating! We can see this simply by computing the second derivative of the proper distance between two point:

$$\frac{d^2}{dt^2} \Delta\sigma = H^2 e^{Ht} \Delta x \,. \tag{0.350}$$

This accelerated expansion was discovered in 1998 by measuring redshifts from supernovae. The parameter H is known as the Hubble constant (we'll get to this when we cover cosmology more completely soon), and our late-time universe satisfies

$$H = \frac{1}{14 \text{ billion years}} \,. \tag{0.351}$$

14 billion years is known as the Hubble time, and it is the amount of time for the universe to "stretch" by a factor of e (since e^{Ht} for $t = 1/H$ gives e).

Adding matter sources

The cosmological constant term in the action is a very special sort of source. But we can have much more general matter sources. What are the rules for supplementing the action with these?

First let's recall the type of equation we are after: (curvature) = (source). We have seen that the curvature comes in the form of the Einstein tensor, a rank-2 symmetric tensor. So the source must also be a rank-2 symmetric tensor. And it should encode information about the energy of the objects. Since the Einstein tensor is local (i.e. a function of spacetime coordinates as opposed to some integral over space and time), the source also has to be local. So it should really encode energy density. In relativity, energy comes with momentum in the momentum 4-vector, so the energy density should come along with the momentum density:

$$\frac{E}{L_x L_y L_z} \implies \frac{p^i}{L_x L_y L_z} \,. \tag{0.352}$$

But we also know, even more fundamentally, that in relativity space and time should be treated on equal footing, so we should probably also have terms like

$$\frac{p^i}{L_x L_y t} = \frac{F^i}{L_x L_y} \,. \tag{0.353}$$

For $i = z$ this is like a pressure (transverse force per unit area), whereas for $i = x$ or $i = y$ this is like a stress. All of these objects get packaged together into the stress energy tensor $T_{\alpha\beta}$, a rank-2 symmetric tensor. T_{tt} is the energy density, T_{ti} is the momentum density, and T_{ij} contains the stress and pressure terms (pressure terms if $i = j$ and stress terms otherwise).

But to see that this indeed appears as a source in Einstein's equation, we need to supplement our action:

$$S_{\text{tot}} = S_{\text{EH}} + S_{\text{matter}} \,. \tag{0.354}$$

The total action governs both the dynamics of the spacetime metric and matter. We still need that S_{matter} is coordinate invariant, so the actions you are used to will have to be supplemented by a factor of $\sqrt{-g}$, e.g.

$$S_{\text{Maxwell}} = \int d^4x F_{\mu\nu} F^{\mu\nu} \to \int d^4x \sqrt{-g} F_{\mu\nu} F^{\mu\nu}. \tag{0.355}$$

This minimal modification is known as "minimal coupling," and it does two things. As already stated, it ensures the full action is coordinate invariant. But it also makes sure that spacetime and matter "talk" to one another.[14] Otherwise, if you add two distinct actions depending on distinct fields to one another, then the equations of motion will just split into distinct equations of motion instead of a mixed equation like (curvature) = (source) that we are looking for.

Now we implement our variational principle:

$$\delta S_{\mathrm{EH}} = -\frac{1}{16\pi G}\int d^4x\sqrt{-g}\,\delta g_{\alpha\beta}(R^{\alpha\beta} - \frac{1}{2}Rg^{\alpha\beta})\,, \quad \delta S_{\mathrm{matter}} \equiv -\frac{1}{2}\int d^4x\sqrt{-g}\,\delta g_{\alpha\beta}T^{\alpha\beta}. \tag{0.356}$$

Requiring $\delta S_{\mathrm{tot}} = \delta S_{\mathrm{EH}} + \delta S_{\mathrm{matter}} = 0$ for any $\delta g_{\alpha\beta}$ gives

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = 8\pi G T_{\alpha\beta}\,. \tag{0.357}$$

As promised, this is Einstein's equation with a source! We can also consider a nonzero cosmological constant, and the equation is sometimes written as

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = 8\pi G T_{\alpha\beta}\,. \tag{0.358}$$

But from the perspective of sources, a cosmological constant Λ can just be thought of as a term in the matter action. It is just some other source for spacetime curvature (albeit quite a special one!) For example, it doesn't contain any derivatives of the metric, just like other matter actions you might consider (and unlike terms like R^2). By moving the cosmological constant term to the right-hand-side of the equation, we can obtain its stress tensor

$$T^{(\Lambda)}_{\alpha\beta} = -\frac{\Lambda}{8\pi G}g_{\alpha\beta}\,. \tag{0.359}$$

We see that it is just a special form of matter with a uniquely simple stress-energy tensor.

———————————— END LEC 16————————————

FLRW cosmology

In this section we will discuss FLRW cosmology, named after the Soviet Friedmann, the American Robertson, the Englishman Walker, and the Belgian priest Lemaitre. Let's give the big picture first. These geometries are meant to describe an expanding universe. They assume that the *spatial* geometry is isotropic (spherically symmetric around any point in space, i.e. the same any direction you look) and homogeneous (the same at every point in space). These are not the same notion; the geometry $\mathbb{R}\times S^2$ is homogeneous but not isotropic, and the cone is isotropic at the apex but not homogeneous (since the apex is a special point). If the space is isotropic *everywhere* then it is homogeneous, and if it is isotropic around a

[14] $F_{\mu\nu}$ and $g_{\mu\nu}$ also talk to one another through the contraction $F_{\mu\nu}F^{\mu\nu}$.

given point and homogeneous, then it will be isotropic around every point. The assumption of isotropy and homogeneity comes, historically, from the Copernican principle, which says that the universe is basically the same everywhere. We now have plenty of observational evidence for this, but it is to be understood as a statement about *large* distance scales. Clearly the universe is different at the surface of the Sun than it is in interstellar space.

The simplest example of an FLRW geometry can be written as

$$ds^2 = -dt^2 + a(t)^2(dx^2 + dy^2 + dz^2)\,. \tag{0.360}$$

The function $a(t)$ is called the scale factor and determines the expansion of space. Notice space here is infinite, $\mathbb{R}^3$; at any coordinate time T the geometry is still $\mathbb{R}^3$:

$$d\sigma^2 = a(T)^2(dx^2 + dy^2 + dz^2). \tag{0.361}$$

If $a(t)$ grows in time this space is "expanding." This is what it means for infinitely big space to expand! Points at fixed coordinates (x, y, z) can still get further apart. For example, let's compute the proper distance between two points $(0, 0, 0)$ and $(\Delta x, 0, 0)$ at time t_1 and then at time t_2:

$$\int d\sigma = a(t_1)\Delta x \longrightarrow a(t_2)\Delta x\,. \tag{0.362}$$

This geometry is a very good approximation to our universe *on large distance scales*. Our universe is clearly not homogeneous and isotropic; the planets, galaxies, clusters of galaxies etc. make up localized sources of stress-energy, where the geometry will deviate from parts of the universe that have gas or are mostly vacuum. But if you blur your eyes and average the stress-energy over long distance scales, then there is roughly just as much in any direction you look.

Let's also see how this geometry implies a cosmological redshift. The idea is very simple. Without the expansion of the universe, i.e. $a(t) = 1$, the wavelength λ at some time t_1 would be the same as the wavelength at a later time t_2. But the expansion of space implies the wavelength gets stretched out. So at t_1 the wavelength is $\lambda_1 = a(t_1)\lambda$ and at time t_2 it is $\lambda_2 = a(t_2)\lambda$. This means we have

$$\frac{\lambda_2}{\lambda_1} = \frac{a(t_2)}{a(t_1)} \implies \frac{\omega_2}{\omega_1} = \frac{a(t_1)}{a(t_2)}\,. \tag{0.363}$$

So in an expanding universe $(a(t_2) > a(t_1))$ the frequency decreases in tme. This is the cosmological redshift. It is used as a measure of distance or time because faraway photons take a long time to get to us, and in this intervening time the universe is expanding and its wavelength is therefore growing. This leads to the definition of the redshift z as

$$1 + z = \frac{\lambda_{\text{today}}}{\lambda_{\text{emitted}}} = \frac{\omega_{\text{emitted}}}{\omega_{\text{today}}} = \frac{a(t_{\text{today}})}{a(t_{\text{emitted}})}\,. \tag{0.364}$$

Objects with large redshift are far away in space, and the light from them is therefore from very early in time. The furthest galaxies have $z \sim 11$, which means its light is from ~ 13 billion years ago. This places them at around 32 billion light years away.[15]

[15]If the universe is only 13.8 billion years old, how are there objects further than 13.8 billion light years

Solution to Einstein's equations

With Einstein's equations (with sources!) in hand, and some experience with various metric tensors, we can now see how the above geometries arise as a solution to the equations. As discussed above, we assume the spatial geometry is homogeneous and isotropic. This means the space is maximally symmetric. There are three maximally symmetric spaces: flat space $\mathbb{R}^3$ which has zero spatial curvature, a sphere S^3 which has constant positive curvature, and hyperbolic space H^3 which has constant negative curvature (or discrete quotients thereof). These cases are sometimes called flat, closed, and open, respectively. So we can write our general FLRW cosmology as

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \right), \qquad k = \begin{cases} +1\,, & S^3 \\ 0\,, & \mathbb{R}^3 \\ -1\,, & H^3 \end{cases} \tag{0.365}$$

A more illuminating form of the metric is as follows

$$ds^2 = -dt^2 + a(t)^2 \left(d\chi^2 + \begin{Bmatrix} \sin^2\chi \\ \chi^2 \\ \sinh^2\chi \end{Bmatrix} d\Omega^2 \right) \qquad \begin{Bmatrix} \text{closed} \\ \text{flat} \\ \text{open} \end{Bmatrix} \tag{0.366}$$

This cosmology is not a solution of the vacuum Einstein's equations. It needs a source, which we will take to be a "perfect fluid." These are fluids which have a stress tensor given by

$$T_{\mu\nu} = (\rho + p) u_\mu u_\nu + p g_{\mu\nu} \tag{0.367}$$

where ρ and p are the mass density and pressure, respectively, in the rest frame of the fluid. If we evaluate this in the rest frame where $u^\mu = (1, 0, 0, 0)$ we find

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & & & \\ 0 & & g_{ij}p & \\ 0 & & & \end{pmatrix} \tag{0.368}$$

The Einstein field equations are given by (restoring factors of G)

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}\,. \tag{0.369}$$

These are generally ten equations (since the objects are symmetric 4×4 matrices). But due to the homogeneity and isotropy assumption, we only end up with two independent equations. The first is given by the $\mu\nu = 00$ equation

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3p) \tag{0.370}$$

away? It is because space is continuously expanding. Saying the universe is 13.8 billion years old only means that light from the furthest objects has traveled 13.8 billion light years to get to us, but in the meantime space has expanded, placing the object even further from us!

and the second is given by the $\mu\nu = ij$ equations

$$\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2} = 4\pi G(\rho - p)\,. \tag{0.371}$$

These can be simplified into the two Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}\,, \qquad \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \tag{0.372}$$

Often just the first of these is called the Friedmann equation.

So the FLRW cosmology (0.365) - (0.366) is a solution of Einstein's equations as long as the scale factor $a(t)$ satisfy the equations above. Notice we found this solution in a sort of funny way, picking an ansatz for what the solution should look like, then specifying a source $T_{\mu\nu}$, then looking at the equations. In a totally general case all we will know is $T_{\mu\nu}$ and we would have to go out and solve for $g_{\mu\nu}$. This is too difficult in general, so using symmetries to constrain the form of the geometry is essential for analytic treatments.

Phases of our universe

To simplify things we can choose an equation of state for our perfect fluid, which is a relationship between the pressure p and energy density ρ. We can choose the simple equation of state

$$p = w\rho\,. \tag{0.373}$$

We can now take a simple combination of the two Friedmann equations as follows. Write the left-hand-side of the second equation as $\frac{\ddot{a}}{a} = \frac{d}{dt}\frac{\dot{a}}{a} + \frac{\dot{a}^2}{a^2}$ and plug in $\dot{a}/a$ from the first Friedmann equation. This gives

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}\,. \tag{0.374}$$

This equation can be written as $d\rho/\rho = -3(1+w)da/a$ and integrated to obtain

$$\rho \propto a^{-3(1+w)}\,. \tag{0.375}$$

This tells us how the energy density in matter changes as a function of time (and the equation of state of the fluid, i.e. the parameter w). Reasonable energy conditions enforce $|w| \leq 1$.

Important examples of fluids which dominated our universe at some point in time are the vacuum, radiation and matter. It is a little strange to refer to the vacuum as a fluid, but it has an energy density, pressure, and equation of state like a fluid. In fact it has the equation of state $p_\Lambda = -\rho_\Lambda$, which leads to $\rho_\Lambda \propto a^0$. So this energy density does not decrease at all. It is believed that this fluid dominated what is called the inflationary phase of our universe, near its birth, and that we are entering a similar phase now. Looking back at our section on Einstein's equations, we see that this just corresponds to the cosmological constant source discussed there (notice that the stress tensor $T^{(\Lambda)}_{\alpha\beta} = -\frac{\Lambda}{8\pi G} g_{\alpha\beta}$ from (0.359) is the same as the perfect fluid stress tensor (0.368) with $p = -\rho$).

Radiation refers to anything moving with relativistic speeds. This has equation of state $p_R = \frac{1}{3}\rho_R$, so we find $\rho_R \propto a^{-4}$. Our universe's radiation-dominated phase began after inflation ended and lasted until around 47,000 years after the Big Bang. Photons and neutrinos made up the bulk of the radiation in this phase.

Matter refers to nonrelativistic particles that do not collide, which means they will have zero pressure, $w = 0$. This matter therefore has an energy density that falls off as $\rho \propto a^{-3}$. This phase began after the radiation-dominated phase ended and lasted until around 10 billion years after the Big Bang.

Determining what dominates the energy density of our universe depends, of course, on what's around, and also on how quickly the energy density dilutes. Notice that a cosmological constant is the only thing that doesn't dilute, so if it is present, it will *eventually* dominate. That seems to be what is happening in our universe.

In the case of a flat universe, $k = 0$, the first Friedmann equation can be integrated to find a simple power-law behavior for the scale factor

$$\left(\frac{\dot{a}}{a}\right)^2 \propto a^{-3(1+w)} \implies a^{\frac{1+3w}{2}} da \propto dt \implies a \propto t^{\frac{2}{3+3w}} . \tag{0.376}$$

So for matter domination we have $a \propto t^{2/3}$ and radiation domination has $a \propto t^{1/2}$. Vacuum energy domination $w = -1$ looks like a singular case, but we should revisit the original differential equation $\dot{a}^2/a^2 \propto$ const. which integrates to $a \propto e^{Ht}$, where we have defined the Hubble constant $H \equiv \frac{\Lambda}{3}$. This is a special case of the more general definition $H = \dot{a}/a$.

Using this general definition of the Hubble parameter, we can introduce a few more definitions that are useful in cosmology. We have the density parameter

$$\Omega = \frac{8\pi G}{3H^2}\rho = \frac{\rho}{\rho_{\text{crit}}} , \qquad \rho_{\text{crit}} \equiv \frac{3H^2}{8\pi G} . \tag{0.377}$$

The quantity ρ_{crit} is called the critical density because the first Friedmann equation can be written as

$$\Omega - 1 = \frac{k}{H^2 a^2} . \tag{0.378}$$

In particular this means that

$$\rho < \rho_{\text{crit}} \iff \Omega < 1 \iff k < 0 \iff \text{open} \tag{0.379}$$
$$\rho = \rho_{\text{crit}} \iff \Omega = 1 \iff k = 0 \iff \text{flat} \tag{0.380}$$
$$\rho > \rho_{\text{crit}} \iff \Omega > 1 \iff k > 0 \iff \text{closed} \tag{0.381}$$
$$\tag{0.382}$$

The Hubble parameter $H = \dot{a}/a$ today is known as the Hubble constant, H_0. As of writing, there is currently an interesting discrepancy in the measurement of the Hubble constant. Measurements from the cosmic microwave background give $H_0 \approx 67$ km/sec/Mpc, where an Mpc is 3×10^{24} cm. Measurements of the velocity of stars, however, gives a value of 73 km/sec/Mpc.

Acceleration and event horizons

From the expanding phases of our universe that we have discussed above, only one of them has an expansion that is *accelerating*. This is the vacuum-dominated aka cosmological constant dominated aka dark-energy dominated phase. Whether the universe is accelerating is determined by $\ddot{a}$; acceleration corresponds to $\ddot{a} > 0$ and deceleration corresponds to $\ddot{a} < 0$.

Example 29: For the case of flat FRW, what values of w in the equation of state correspond to accelerating universes? Since we have

$$a(t) \propto t^{\frac{2}{3(1+w)}} \tag{0.383}$$

with positive proportionality constant, we calculate

$$\ddot{a} \propto -\frac{2(1+3w)}{9(1+w)^2} t^{\frac{2}{3(1+w)}-2} > 0 \implies (1+3w) < 0 \implies -1 \leq w < -1/3\,. \tag{0.384}$$

We bounded w below by -1 from our general bound $|w| \leq 1$ stemming from picking a "reasonable" source of energy.

In an accelerating universe, there will be spacetime events that will be forever out of reach. This is the notion of an event horizon, a concept we met while studying black holes, which restricts how much of the universe we can observe even if we live for an infinitely long time. This distinct from the notion of a particle horizon which is often discussed in cosmology. The latter is a function of time and just has to do with how much of our universe we have seen *up to that point in time*. Particles that are not in our particle horizon may enter into our horizon at a later time. In contrast, an event horizon is a *global* property of the spacetime, i.e. it has to do with the entire spacetime geometry, and its existence implies that there are particles that will forever be outside of our reach.

Example 30: A constant source of confusion in popular science (and among physicists!) is that if space expands fast enough then light from sufficiently far away will not make it to us. Here, "fast enough" is meant to just refer to a velocity. But the key point that keeps light from reaching us is **acceleration** – the derivative of the expansion rate – not just the expansion rate.

To see this in a simplified setting, consider Martin Gardner's ant on a rubber rope puzzle. The rubber rope is a finite length L. The ant starts off at the left end of the rope, $x = 0$, traveling at some speed v toward the right end $x = L$. Say $v = 1$ cm/s. Once the ant starts moving, the rope begins expanding uniformly, with its overall size growing at a rate $V \gg v$. Say $V = 10$ cm/s. So the right end of the rope is getting 10 cm further every second, and the poor ant is only traveling at 1 cm/s. Will the ant ever reach the end of the rope?

Since the rope is expanding very quickly, it seems like the ant's not going to make it. But the key point is that the rope is expanding *uniformly*. This means that, even though the right end of the rope travels 10 cm/s, that total 10 cm increase is evenly distributed along the rope. In particular, some of that 10 cm will be *behind* the ant (if the ant is at position x then $x/L\times$ 10 cm of the increased length will be behind the ant. This effect means that, as long as the expansion of the rope is not accelerating, the ant will **always** reach the end of the rope. The expansion of the universe is the same: as long as it is not accelerating, light from arbitrarily far away will eventually reach us.

Big Bang singularity

The FLRW cosmologies we have been studying have an $a(t)$ that grows in time. This means if we rewind the clock, $a(t)$ decreases and space shrinks. If this continues indefinitely, space gets contracted to a point. This is the Big Bang singularity.[16] It is a true singularity: you can compute for yourself the curvature invariant of an FLRW cosmology with positive power-law scale factor $a(t) \sim t^n$ with $n > 0$ and see that it diverges as $t \to 0$. Part of the reason many of us are obsessed with black holes and trying to understand the curvature singularity inside black holes is because it is similar to the Big Bang curvature singularity.

———————————— END LEC 17————————————

Gravitational waves

We would be remiss not to discuss one of the most exciting observational developments in general relativity in recent times: the observation of gravitational waves by LIGO!

Solving Einstein's equations are hard. To this day there are only a handful of analytic solutions to the equations, and much of the analysis of the equations is numerical. But there is one limit, like in many differential equations, where the solution is straightforward. This is the limit of a weak source, which we expect to lead to a small change in the metric away from the flat metric (which is the solution with no source). In particular, this means we can consider an expansion around the flat metric

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \qquad h_{\mu\nu} \ll 1 \tag{0.385}$$

and work to linear order in $h_{\mu\nu}$. Einstein's equations in this limit become, when we pick an appropriate gauge (known as *harmonic* gauge or *Lorenz* gauge) and leave out technical subtleties,

$$-\partial^\alpha \partial_\alpha \bar{h}_{\mu\nu} = \left(\partial_t^2 - \vec{\nabla}\cdot\vec{\nabla}\right)\tilde{h}_{\mu\nu} = 8\pi G T_{\mu\nu}\,. \tag{0.386}$$

where $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu}$ is some close cousin of $h_{\mu\nu}$ we defined above. The gauge is $\partial_\mu \bar{h}^{\mu\lambda} = 0$. This is analogous to the Lorenz gauge in E&M $\partial_\mu A^\mu = 0$. In our case it is

[16] Some unfortunate double-use of language. The Big Bang was a term coined before inflation was theorized. So it referred to the beginning of the universe. Nowadays, sometimes people mean the beginning of the universe when they say Big Bang, but often they mean "reheating," the phase immediately after inflation where lots of radiation was produced and came to dominate the energy density of the universe.

four conditions (one for each value of λ) which reduces the ten components of $h_{\mu\nu}$ to six components.

This looks like a wave equation, and indeed has solutions that look like waves. This is not disputed. But historically there was a lot of confusion around whether these solutions to Einstein's equations were actually novel. As a silly example, consider Minkowski spacetime

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{0.387}$$

and consider the coordinate transformation

$$x = \tilde{x} + h \sin(t - z)\,, \qquad h \ll 1\,. \tag{0.388}$$

Working to **linear order** in h, the Minkowski metric becomes

$$ds^2 = -dt^2 + d\tilde{x}^2 + dy^2 + dz^2 + 2h\cos(t - z)(dt\, d\tilde{x} - d\tilde{x}\, dz)\,. \tag{0.389}$$

This metric sort of looks like a wave, but it is clearly Minkowski space masquerading as one.

This was such a confusing issue that Einstein and Nathan Rosen initially dismissed the reality of the waves. When an anonymous referee (Robertson the R of FLRW) explained that they are actually real, Einstein threw a hissyfit and wrote to the editor of the journal

Dear Sir, We (Mr. Rosen and I) had sent you our manuscript for publication and had not authorized you to show it to specialists before it is printed. I see no reason to address the in any case erroneous comments of your anonymous expert. On the basis of this incident I prefer to publish the paper elsewhere.

He never published with Physical Review again. To be fair, peer review as a mainstream concept was still somewhat new at the time. Einstein submitted the paper to the Journal of the Franklin Institute, modifying the second half of the paper to now argue that gravitational waves were real (and thanked Robertson in the acknowledgments, presumably because he explained to Einstein the error in person).

An important way people understood that these gravitational waves are novel and real is that they carry energy. One of the original popular arguments for this is due to Feynman, who explained that the motion of particles as a gravitational wave crosses them can lead them to rub against each other and produce heat due to friction (this is the sticky bead argument, which he proposed in a conference in 1957).

The Lorenz or harmonic gauge reduced the seeming ten independent components of $h_{\mu\nu}$ to six independent components. It turns out there is further gauge freedom we have. We can fix to what is called "transverse traceless" gauge. To see what this means, let's actually write down the solution to the equations of motion:

$$ds^2 = -dt^2 + [1 + f(t - z)]dx^2 + [1 - f(t - z)]dy^2 + dz^2\,. \tag{0.390}$$

The function f can be anything, as long as it is small $|f(t-z)| \ll 1$. For example we can pick a Gaussian wave $f(t-z) = a\exp[-(t-z)^2/\sigma^2]$ or a sinusoidal wave $f(t-z) = a\sin[\omega(t-z)]$, as long as a is small.

The "transverse" means that the part of the metric that is changing (the x and y directions in this case) is different from the direction the wave is moving (the z direction in this case, due to the function $f(t-z)$). The "traceless" part means $h^\mu_\mu = 0$. Notice that working to linear order in $h_{\mu\nu}$ means we are allowed to use the Minkowski metric to take the trace, $h^\mu_\mu \equiv \eta^{\mu\nu} h_{\mu\nu}$. The traceless condition rules out e.g.

$$ds^2 = -dt^2 + (1 + 2h\sin(t-z))dx^2 + dy^2 + dz^2\,, \tag{0.391}$$

which is transverse. We can change coordinates in (0.391) as $\tilde{x} = x(1 + h\sin(t-z))$ to get

$$ds^2 = -dt^2 + d\tilde{x}^2 + dy^2 + dz^2 - 2h\tilde{x}\cos(t-z)(dt\,d\tilde{x} + d\tilde{x}\,dz)\,. \tag{0.392}$$

This is now traceless but it is no longer transverse.

Our solution (0.390) is a lot like a light wave, because light waves propagate in a direction perpendicular to the ones where the electric field and magnetic field are changing. So it is transverse. It also has E and B fields in opposite directions, similar to how the spatial metric has varying components in the x and y directions. The solution seems to imply a time-dependent change to the x and y positions of particles. Let's see this now.

We want to see how a gravitational wave affects something we can measure. A single particle won't be good enough for this, since in a freely falling frame they won't feel anything happen. So we will consider two particles, particle A at the origin and particle B at position $\vec{x}_B$. We always need at least two particles to measure spacetime curvature. They both begin at rest $u^\mu_{(A)} = u^\mu_{(B)} = (1, \vec{0})$ and we consider the wave (0.390) passing by. Our task is now to solve the geodesic equation, which we will do to first order in the amplitude of the wave. The geodesic equation for either particle is

$$\frac{du^\mu}{d\tau} = -\Gamma^\mu_{\nu\gamma} u^\nu u^\gamma\,. \tag{0.393}$$

Notice that it is sufficient to use Γ to order h and therefore u^γ to order h^0 to get the right-hand-side to order h (if we used Γ to order h^0 we would have the flat-space Christoffel symbol, which would vanish). Thus we have

$$\frac{d^2x^i}{d\tau^2} = -\Gamma^i_{00}\,. \tag{0.394}$$

Evaluating Γ^i_{00} for (0.390) shows that it vanishes. So we have

$$\frac{d^2x^i}{d\tau^2} = 0\,. \tag{0.395}$$

Since the particles begin with zero velocity, and we find that they have zero acceleration, then their **coordinate** positions remain unchanged!

The physical distance is what we actually care about, not the coordinate distance (recall the analogous discussion in the case of FLRW cosmology). To see this, we calculate an example. Let's put the second particle at $(L,0,0)$. The proper distance between them is given by

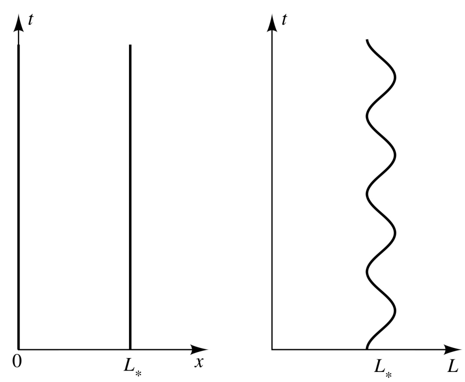
$$\int d\sigma = \int_0^L dx[1+f(t)]^{1/2} \approx L\left[1+\frac{1}{2}f(t)\right] \implies \frac{\Delta L(t)}{L} = \frac{1}{2}f(t)\,. \tag{0.396}$$

where in the final expression we wrote a fractional change in the proper distance. Notice that if we separated a different pair of particles by distance L in the y direction, we would find

$$\int d\sigma \approx L\left[1-\frac{1}{2}f(t)\right] \implies \frac{\Delta L(t)}{L} = -\frac{1}{2}f(t)\,. \tag{0.397}$$

So the two lengths are out of phase. When the x direction stretches, the y direction squeezes, and vice versa.

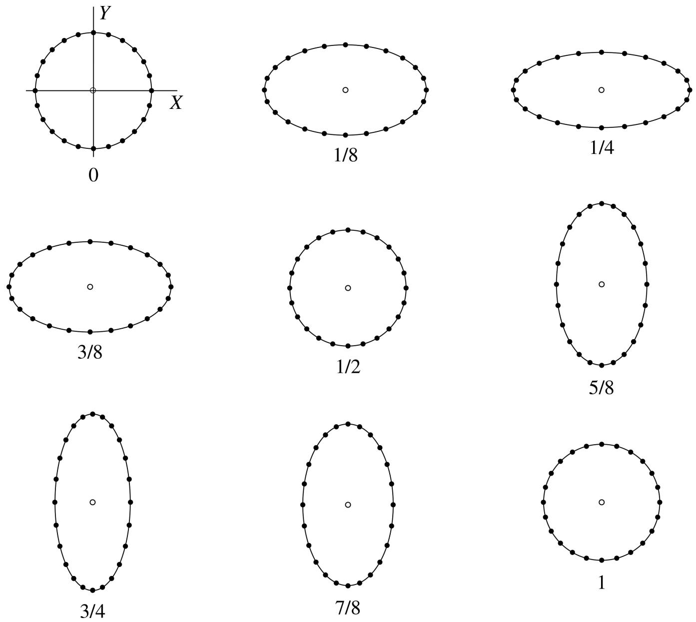
Let's pick a particular function, a sinusoidal wave $f(t-z) = a\sin[\omega(t-z)]$. Then the total length oscillates around an equilibrium position, with amplitude 1/2 the amplitude of the wave! A picture is below, with the image on the left being the case with no gravity wave and the image on the right the case with a gravity wave.



We can generalize this to see what happens to a ring of test masses in the $z=0$ plane. To calculate this it is simplest to pick new coordinates

$$X = \left(1+\frac{1}{2}a\sin\omega t\right)x\,, \qquad Y = \left(1-\frac{1}{2}a\sin\omega t\right)y \tag{0.398}$$

which leads to the metric in the $x-y$ plane being written as dX^2+dY^2. The physical distance between test masses can be calculated as their $X(t)$ and $Y(t)$ positions, which have time-dependence according to (0.398). The behavior of the particles therefore looks like the below, where we pick $a=0.8$ and the time is represented below each figure as a fraction of a period:



The above is one of the two polarizations that a gravitational wave can take, called the + (plus) polarization, since the ring oscillates in the vertical and horizontal directions. The other independent polarization is the $\times$ (cross) polarization, because the motion of the ring of masses is simply rotated by 45 degrees.

An important fact about linearized equations of motion is that they are *linear*! This means that we can take multiple solutions to the equations, add them together, and get a new solution. This means that a general gravitational wave moving in the z direction can be written as a superposition of the two polarizations as

$$h_{\alpha\beta}(t,z) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & f_+(t-z) & f_\times(t-z) & 0 \\ 0 & f_\times(t-z) & -f_+(t-z) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{0.399}$$

for two different functions $f_+(t-z)$ and $f_\times(t-z)$. The most general solution to the linearized equations are superpositions of the above in different directions of propagation.

These changes in length are precisely what gravitational wave detectors like LIGO-Virgo look for. The basic idea is similar to the Michelson interferometer. There are two long, perpendicular arms of equal length running off from a beam splitter. A laser is shined through the beam splitter and splits into two pieces which go down the long arms and reflect back. If there is a difference in the lengths of the two arms (which a gravitational wave would cause, as in our ellipses above), then there will be an interference pattern when the beams recombine.

Energy in gravitational waves and general relativity

It is a somewhat surprising fact, when you first face it, that there is no local notion of energy in general relativity. But it follows pretty immediately from Emmy Noether's connection between symmetries of spacetime and conserved quantities (like energy). For example, in the Schwarzschild spacetime we had time-translation and ϕ-translation symmetries, which led to conserved energy and angular momentum. But general relativity is a **theory of spacetime**, and there are no symmetries which characterize **all** spacetimes.[17]

We should still be able to define an approximate notion of energy for a gravitational wave. One way to do this is to average over many wavelengths, but not so many such that we become sensitive to the curvature scale of the background geometry (which can be more general than the flat $\eta_{\mu\nu}$ we have been working with so far). Hartle discusses this. We expect the energy, as usual, to be proportional to the square of the amplitude. To end up with an object with the dimensions of an energy density, the answer should therefore be $\epsilon \sim \omega^2 a^2$. Whether gravity waves carry energy was controversial for quite a while, and one of the best original arguments showing that it does is Feynman's sticky bead argument.

Sources of gravitational waves

We now solve the linearized Einstein's equations (0.386) (linearized around flat space $\eta_{\mu\nu}$). This can be solved by the Green's function method, where we first replace the source $T_{\mu\nu}(x)$ by a delta-function source, and then integrate up the answer to that differential equation to get the general result. This results in

$$\bar{h}_{\mu\nu}(t,\vec{x}) = 4\int d^3x' \, \frac{T_{\mu\nu}(t_{\rm ret},\vec{x}')}{|\vec{x}-\vec{x}'|}\,, \qquad t_{\rm ret} \equiv t - |\vec{x}-\vec{x}'|. \tag{0.400}$$

Let's assume our source $T_{\mu\nu}$ has finite extent, with characteristic scale $R_{\rm source}$, and place ourselves far from it $r \gg R_{\rm source}$. We will also assume the velocities of the source are slow, $\lambda = 2\pi/\omega \gg R_{\rm source}$, where ω is the characteristic frequency of variation of the source. We can approximate our solution as

$$\bar{h}_{\mu\nu}(t,\vec{x}) \longrightarrow \frac{4}{r}\int d^3x' \, T_{\mu\nu}(t-r,\vec{x}') \tag{0.401}$$

In flat spacetime the stress-energy tensor is conserved, $\partial_\mu T^{\mu\nu} = 0$. We can write this as

$$\frac{\partial T^{tt}}{\partial t} = -\frac{\partial T^{kt}}{\partial x^k} \implies \frac{\partial^2 T^{tt}}{\partial t^2} = -\frac{\partial}{\partial t}\frac{\partial T^{tk}}{\partial x^k} = \frac{\partial}{\partial x^k}\frac{\partial T^{tk}}{\partial t} = \frac{\partial^2 T^{kl}}{\partial x^k \partial x^l}\,, \tag{0.402}$$

where we used $T^{tk} = T^{kt}$ and $\partial_\mu T^{\mu\nu}$ again. We can multiply both sides of the final equation by $x^i x^j$ and integrate over space to get

$$\int d^3x \, \frac{\partial^2 T^{tt}}{\partial t^2} x^i x^j = \int d^3x \, x^i x^j \frac{\partial^2 T^{kl}}{\partial x^k \partial x^l} \implies \int d^3x \, T^{ij}(x) = \frac{1}{2}\frac{d^2}{dt^2}\int d^3x \, x^i x^j T^{tt}(x)\,, \tag{0.403}$$

[17] In more careful treatments, we often fix some asymptotic boundary conditions for the metric and consider all solutions of the theory with those boundary conditions. In such cases we can define conserved quantities, because now all spacetimes have the same asymptotic structure, which can have e.g. time translation symmetry, or other symmetries. Such conserved quantities are *maximally nonlocal* because they are defined as integrals over all of spacetime.

where we integrated by parts to get the second form of the equation. For sources with low velocities we can take $T^{\alpha\beta} = \mu u^\alpha u^\beta$ where μ is the rest-mass density. This just means that the energy is mostly given by the rest mass, without any special relativistic enhancement by γ. Defining the second mass moment

$$I^{ij}(t) \equiv \int d^3x\, \mu(t, \vec{x}) x^i x^j \tag{0.404}$$

means we can write the metric perturbation as

$$\bar{h}^{ij}(t, \vec{x}) \longrightarrow \frac{2}{r} \ddot{I}^{ij}(t - r)\,, \tag{0.405}$$

i.e. the gravitational wave is sourced by the second time derivative of the second mass moment. The second mass moment is closely related to the quadrupole moment. So we see that gravitational waves are sourced by a time-dependent quadrupole, in contrast with electromagnetic waves which only need a time-dependent dipole. A nice table is given in Hartle comparing the two cases, which is reproduced below.

	Linearized gravitation ($c = G = 1$)	Electromagnetism ($c = 1$)
Field equation	Einstein equation with $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$	Maxwell's equations
Basic potentials	Linearized metric perturbations $h_{\alpha\beta}(x)$	Vector and scalar potentials $\big(\Phi(x), \vec{A}(x)\big)$
Sources	Stress-energy $T_{\alpha\beta}$	Charge and current $(\rho_{\text{elec}}, \vec{J})$
Lorentz gauge	$\dfrac{\partial \bar{h}^{\alpha\beta}}{\partial x^\alpha} = 0$	$\dfrac{\partial \Phi}{\partial t} + \vec{\nabla} \cdot \vec{A} = 0$
Wave equation with source	$\Box \bar{h}_{ij} = -16\pi T_{ij}$	$\Box \vec{A} = -\mu_0 \vec{J}$
General solution	$\bar{h}^{ij} = 4 \int d^3x' \dfrac{[T^{ij}]_{\text{ret}}}{\lvert\vec{x} - \vec{x}'\rvert}$	$\vec{A} = \dfrac{\mu_0}{4\pi} \int d^3x' \dfrac{[\vec{J}]_{\text{ret}}}{\lvert\vec{x} - \vec{x}'\rvert}$
Large r, long-wavelength approximation	$\bar{h}^{ij} = \dfrac{2[\ddot{I}^{ij}]_{\text{ret}}}{r}$ $I^{ij} = \int d^3x\, \mu x^i x^j$	$\vec{A} = \dfrac{\mu_0}{4\pi} \dfrac{[\dot{\vec{p}}]_{\text{ret}}}{r}$ $\vec{p} = \int d^3x\, \rho_{\text{elec}} \vec{x}$
Time-averaged radiated power	$\dfrac{dE}{dt} = \dfrac{1}{5} \langle \dddot{I}_{ij} \dddot{I}^{ij} \rangle$	$\dfrac{dE}{dt} = \dfrac{\mu_0}{6\pi} \langle \ddot{\vec{p}}^{\,2} \rangle$

The angle brackets, $\langle \cdot \rangle$, denote an average over a time longer than the characteristic period of the source. The equations in the linearized gravitation column are in $c = G = 1$ geometrized units. The electromagnetic equations are in SI units with $c = 1$, where $\mu_0 \equiv 4\pi \times 10^{-7}$. To convert this column to Gaussian units with $c = 1$, replace μ_0 by 4π.