Physics 225A: General Relativity

Lecturer: McGreevy

Last updated: 2014/01/13, 17:08:04

0.1 Introductory remarks

I will begin with some comments about my goals for this course.

General relativity has a bad reputation. It is a classical field theory, conceptually of the same status as Maxwell's theory of electricity and magnetism. It can be described by an action principle – a functional of the dynamical variables, whose variation produces well-posed equations of motion. When supplied with appropriate boundary conditions (I'm including initial values in that term), these equations have solutions. Just like in the case of E&M, these solutions aren't always easy to find analytically, but in cases with a lot of symmetry they are.

A small wrinkle in this conceptually simple description of the theory is the nature of the field in question. The dynamical variables can be organized as a collection of fields with two spacetime indices: $g_{\mu\nu}(x)$. It is extremely useful to think of this field as the *metric tensor* determining the distances between points in spacetime. This leads to two problems which we'll work hard to surmount:
1) It makes the theory seem really exotic and fancy and unfamiliar and different from E&M.
2) It makes it harder than usual to construct the theory in such a way that it doesn't depend on what coordinates we like to use. [1]

We'll begin by looking at some motivation for the identification above, which leads immediately to some (qualitative) physical consequences. Then we'll go back and develop the necessary ingredients for constructing the theory for real, along the way reminding everyone about special relativity.

0.2 Conventions and acknowledgement

The speed of light is $c = 1$. $\hbar$ will not appear very often but when it does it will be in units where $\hbar = 1$. Sometime later, we may work in units of mass where $8\pi G_N = 1$.

We will use mostly-plus signature, where the Minkowski line element is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}\vec{x}^2.$$

In this class, as in physics and in life, time is the weird one.

I highly recommend writing a note on the cover page of any GR book you own indicating which signature it uses.

[1] To dispense right away with a common misconception: all the theories of physics you've been using so far have had this property of general covariance. It's not a special property of gravity that even people who label points differently should still get the same answers for physics questions.

The convention (attributed to Einstein) that repeated indices are summed is always in effect, unless otherwise noted.

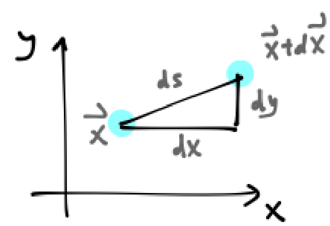
I will reserve τ for the proper time and will use weird symbols like $\mathfrak{s}$ (it's a gothic 's' (`\mathfrak{s}`)!) for arbitrary worldline parameters.

Please tell me if you find typos or errors or violations of the rules above.

Note that the section numbers below do not correspond to lecture numbers. I'll mark the end of each lecture as we get there.

I would like to acknowledge that this course owes a lot to the excellent teachers from whom I learned the subject, Robert Brandenberger and Hirosi Ooguri.

1 Gravity is the curvature of spacetime



Let's begin with Pythagoras:

$$ds = \sqrt{dx^2 + dy^2 + dz^2}$$

This is the distance between the points with cartesian coordinates

$$\vec{x} = (x, y, z) \quad \text{and} \quad \vec{x} + d\vec{x} = (x + dx, y + dy, z + dz)$$

in flat space. This is the world of Euclidean geometry. Square roots are annoying so we'll often think instead about the square of this distance:

$$ds^2 = dx^2 + dy^2 + dz^2 \equiv d\vec{x}^2 \; . \tag{1}$$

Some consequences of this equation which you know are: the sum of interior angles of a triangle is π, the sum of the interior angles of a quadrilateral is 2π.

Similarly, the squared 'distance' between events in flat space<u>time</u> is

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

Physics in flat spacetime is the content of special relativity. (I promise to give a little more review than that soon.)

Equivalence principles

Newton's second law: $\vec{F} = m_i \vec{a}$. $m_i \equiv$ inertial mass. (The i here is not an index but is to emphasize that this is the *i*nertial mass.) A priori, this quantity m_i has nothing to do with gravity.

Newton's law of gravity: $\vec{F}_g = -m_g \vec{\nabla}\phi$. $m_g \equiv$ gravitational mass. ϕ is the gravitational potential. Its important property for now is that it's independent of m_g.

It's worth pausing for a moment here to emphasize that this is an amazingly successful physical theory which successfully describes the motion of apples, moons, planets, stars, clusters of stars, galaxies, clusters of galaxies... A mysterious observation is that $m_i = m_g$ as far as anyone has been able to tell. This observation is left completely unexplained by this description of the physics.

Experimental tests of $m_i = m_g$:

Galileo, Newton (1686): If $m_i = m_g$, Newton's equation reduces to $\vec{a} = -\vec{\nabla}\phi$, independent of m. Roll objects of different inertial masses (ball of iron, ball of wood) down a slope; observe same final velocity.

Eötvös 1889: Torsion balance. Same idea, better experimental setup (null experiment):

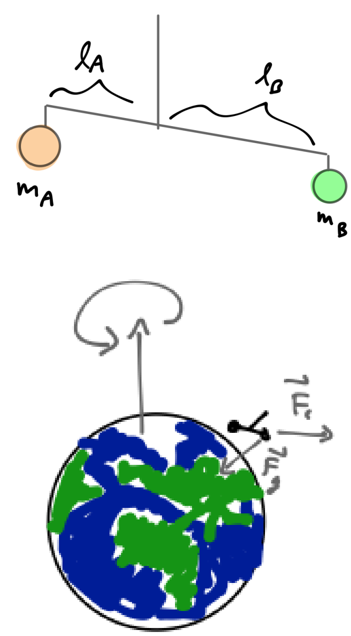


Figure 1: The Eötvös experiment.

Require that the rod is horizontal: $m_A^g \ell_A = m_B^g \ell_B$. Torque due to earth's rotation (centripetal force):

$$T = \ell_A \tilde{g} m_A^g \left(\frac{m_A^i}{m_A^g} - \frac{m_B^i}{m_B^g} \right).$$

$\tilde{g}$: centripetal acceleration.

Results:
Eotvos: $\frac{m_g}{m_i} = 1 \pm 10^{-9}$.
Dicke (1964): 1 ± 10^{-11}.
Adelberger (1990): 1 ± 10^{-12}.
Various current satellite missions hope to do better.

Exercise: What is the optimal latitude for performing this experiment?

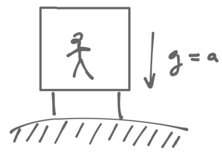
Q: doesn't this show that the *ratio is the same for different things*, not that it is always one?
A: Yes. But: if the ratio were the same for every material, but different than one, we could simply redefine the strength of gravity G_N by a factor (the square root of the ratio) to make it one.

We enshrine this observation as a foundation for further development (known to Galileo):
Weak Equivalence Principle: $m_i = m_g$ for any object.

A consequence of this observation is that we cannot distinguish (by watching trajectories of particles obeying Newton's laws) between the following two circumstances:

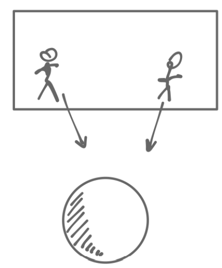
1) constant acceleration

and 2) a gravitational field:



(Note that we are assuming the box you're in is small compared to variations in the field.

Else: we can detect the variations in the field by tidal forces :



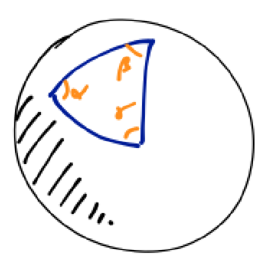
)

Einstein's (or strong) **Equivalence Principle**: In a small region of spacetime, the laws of physics satisfy special relativity – that is, they are invariant under the Poincaré group (we'll review below what this is!). In particular, in such a region, it is impossible to detect the existence of a gravitational field.

Q: how is it stronger? it is just the version that incorporates special relativity, rather than Newtonian mechanics. Hence, it had to wait for Einstein. I would now like to argue that

This implies that gravity is curvature of spacetime

[Zee V.1] Paths of commercial airline flights are curved. (An objective measurement: sum of angles of triangle.)



Is there a force which pushes the airplanes off of a straight path and onto this curved path? If you want. A better description of their paths is that they are 'straight lines' ($\equiv$ geodesics) in curved space. They are straight lines in the sense that the paths are as short as possible (fuel is expensive). An objective sense in which such a space (in which these are the straight lines) is curved is that the sum of interior angles of a triangle is different from (bigger, in this case) than π.

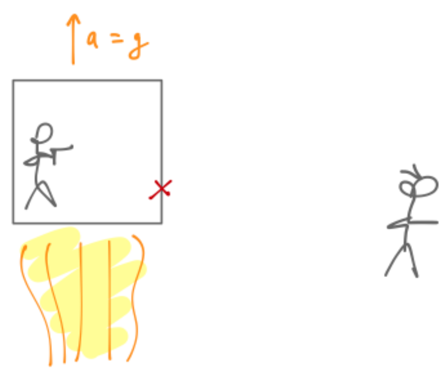
Similarly: it can usefully be said that there is no gravitational force. Rather, we'll find it useful to say that particles follow (the closest thing they can find to) straight lines (again, geodesics) in curved spacetime. To explain why this is a good idea, let's look at some consequences of the equivalence principle (in either form).

[Zee §V.2] preview of predictions: two of the most striking predictions of the theory we are going to study follow (at least qualitatively) directly from the equivalence principle. That is, we can derive these qualitative results from thought experiments. Further, from these two results we may conclude that *gravity is the curvature of spacetime*.

Here is the basic thought experiment setup, using which we'll discuss four different protocols. Put a person in a box in deep space, no planets or stars around and accelerate it uniformly in a direction we will call 'up' with a magnitude g. According to the EEP, the person experiences this in the same way as earthly gravity.

You can see into the box. The person has a laser gun and some detectors. We'll have to points of view on each experiment, and we can learn by comparing them and demanding that everyone agrees about results that can be compared.

1) Bending of light by gravity.

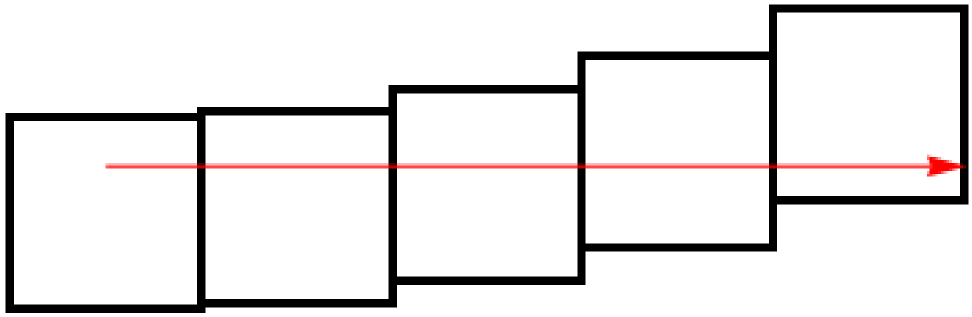


Thought Experiment **1a:** To see this effect, suppose the person fires a laser beam at the wall of the box in a direction transverse to the acceleration. Standing outside the box, you see the photons move in a straight line. While the photons are travelling, the box moves a bit and the spot hits the wall lower than where the beam was fired.

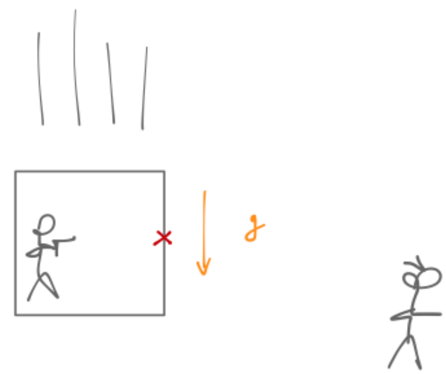
Everyone agrees on where the spot on the wall is. From the point of view of the person in the box, the light moved in a parabola, and he could just as well conclude that it bent because of gravity. If it bent differently when the person was in his room at home, he could distinguish constant acceleration from gravity, violating the EEP. Note that we haven't said quantitatively how much the light bends; that would require incorporating more special relativity than we have so far. And in fact, it bends by different amounts in Newtonian gravity and in Einstein gravity, by a factor of two[2].

Here's a figure that shows the path of the laser beam and the successive heights of the box:

[2]This factor of two is the red one in (33).



1g: A second thought experiment gives the same result: this time, drop the guy in the box in a gravitational field. He will experience free-fall: no gravity. His laser beam hits a target on the wall right at the height he shoots it.

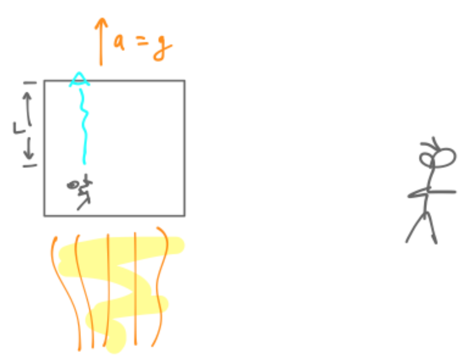


On the other hand, you see him falling. During the time the beam is in the air, the box falls. In the absence of gravity you would say that the beam would have hit higher on the wall. In order to account for the fact that it hit the target, *you* must conclude that gravity bent the light.

If we further demand that a light ray always moves in a straight line, locally (and this is a consequence of the EEP, since it's the case in special relativity), then we must conclude that the existence of a gravitational field means that space is curved. (That is: a triangle whose sides are locally straight has the sum of the internal angles different from π.)

2) Gravitational redshift.

2a: Perhaps the setup has made it clear that we should also try to shoot the laser gun at the ceiling of the box, and see what we get. Put a detector on the ceiling; these detectors can tell the frequency of the light.



From the outside, we see the detector accelerating away from the source: when the beam gets to the detector, the detector is moving faster than when the light was emitted. The Doppler effect says that the frequency is redshifted. From the inside, the victim sees only a gravitational field and concludes that light gets redshifted as it climbs out of the gravitational potential well in which he resides.

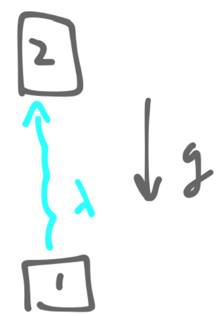
This one we can do quantitatively: The time of flight of the photon is $\Delta t = h/c$, where h is the height of the box. During this time, the box gains velocity $\Delta v = g\Delta t = gh/c$. If we suppose a small acceleration, $gh/c \ll c$, we don't need the fancy SR Doppler formula (for which see Zee §1.3), rather:

$$\frac{\omega_{\text{detector}} - \omega_{\text{source}}}{\omega_{\text{source}}} = \frac{\Delta v}{c} = \frac{gh/c}{c} = \frac{gh}{c^2} = -\frac{\phi_{\text{detector}} - \phi_{\text{source}}}{c^2}$$

Here $\Delta\phi$ is the change in gravitational potential between top and bottom of the box.

This effect of a photon's wavelength changing as it climbs out of the gravitational potential of the Earth has been observed experimentally [Pound-Rebka 1960].
More generally:

$$\frac{\Delta\lambda}{\lambda} = -\int_{\boxed{1}}^{\boxed{2}} \frac{\vec{g}(x)\cdot \mathrm{d}\vec{x}}{c^2} = \frac{\Delta\phi}{c^2}.$$

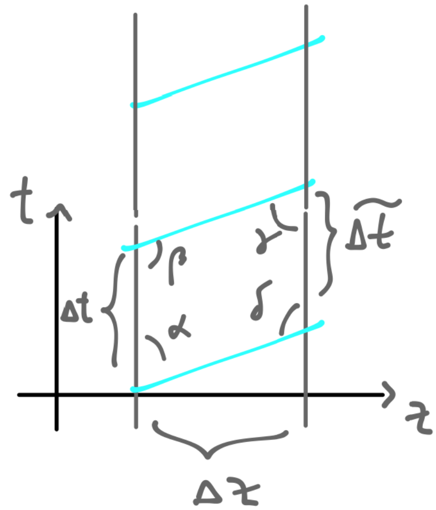


Thought experiment **2g:** Consider also (what Zee calls) the 'dropped' experiment for this case. The guy in the box is in free-fall. Clearly the detector measures the same frequency of the laser he shoots. From the point of view of the outside person, the detector is accelerating towards the source, which would produce a blueshift. The outside person concludes that the gravitational field must cancel this blueshift!

How can gravity change the frequency of light? The frequency of light means you sit there and count the number of wavecrests that pass you in unit time. Obviously gravity doesn't affect the integers. We conclude that gravity affects the flow of time.

Notice that in each thought experiment, both observers agree about the results of measurements (the location of the laser spot on the box, the presence or absence of a frequency shift). They disagree about what physics should be used to explain them! It is the fact that the EEP relates non-gravitational physics (an accelerated frame) to gravitational physics that allows us (rather, Einstein) to base a theory of gravity on it.

Actually: once we agree that locally physics is relativistically invariant, a Lorentz boost relates the redshift to the bending.



Here's a situation which involves both time and space at once: we can measure time by the intervals between crests of a lightwave: $\Delta t = \lambda/c$ where λ is the wavelength of the light. Suppose we try to make a flying clock – *i.e.* send the light ray to someone else – in a gravitational field, as in the figure. According to the EEP, the lightrays are parallel (this is the statement that the speed of light is constant), which means that $\alpha + \beta = \pi$ and $\gamma + \delta = \pi$. (And indeed, in flat spacetime, we would have $\alpha + \beta + \gamma + \delta = 2\pi$). We might also want to declare $\Delta t = \Delta\tilde{t}$ – that is: we demand that the person at $z + \Delta z$ use our clock to measure time steps. On the other hand, the time between the crests seen by an observer shifted by Δz in the direction of the acceleration is:

$$\Delta\tilde{t} = \frac{\lambda}{c}\left(1 + \frac{a\Delta z}{c^2}\right) > \Delta t \ .$$

Something has to give.

We conclude that spacetime is curved.

Combining these two ingredients we conclude that gravity is the curvature of spacetime. This cool-sounding statement has the scientific virtue that it explains the equality of inertial mass and gravitational mass. Our job in this class is to make this statement precise (*e.g.* how do we quantify the curvature?, what determines it?) and to understand some its dramatic physical consequences (*e.g.* there are trapdoors that you can't come back from).

[End of Lecture 1]

Conceptual context and framework of GR

Place of GR in physics:

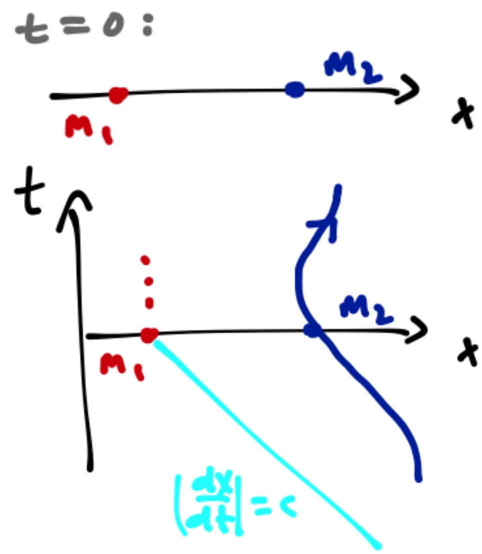
Classical, Newtonian dynamics with Newtonian gravity

$\subset$ special relativity + Newtonian gravity (?)

$\subset$ GR

In the first two items above, there was action at a distance: Newtonian gravity is not consistent with causality in SR, which means that information travels at the speed of light or slower.

[Illustration: given two point masses sitting at r_1, r_2, Newton says the gravitational force from 1 on 2 has magnitude $F_G = G\frac{m_1 m_2}{|r_{12}|^2}$. Now suppose they move: given the history of the motion of m_2 find the force on m_1 at a given time. If particle 2 is moving on some prescribed wiggly trajectory, how does particle 1 know what is $r_{12}(t)$?]



So, once we accept special relativity, we must fix our theory of gravity.

What is GR? A theory of spacetime, and a theory for the motion of matter in spacetime. It can be useful to think that GR has two parts [Wheeler]:

1. spacetime tells matter how to move (equivalence principle)
2. matter tells spacetime how to curve (Einstein's equations).

Some of you have noticed that we haven't yet discussed the second point. We'll see that both of these parts are implemented by the same action principle.

2 Euclidean geometry and special relativity

Special relativity is summarized well in this document – dripping with hindsight.

2.1 Euclidean geometry and Newton's laws

Consider Newton's law

$$\vec{F} = m\ddot{\vec{r}} \, . \tag{2}$$

[Notation: this equation means the same thing as $F^i = m\ddot{r}^i \equiv m\partial_t^2 x^i$. Here $i = 1, 2, 3$ and $r^1 \equiv x, r^2 \equiv y, r^3 \equiv z$. I will probably also use $r^i \equiv x^i$.] And let's consider again the example of a gravitational attraction between two particles, $F = F_G$, so Newton is on both sides. We've already chosen some cartesian coordinates for the space in which the particles are living. In order to say what is F_G, we need a notion of the *distance* between the particles at the positions $\vec{r}_a, \vec{r}_b$. You will not be shocked if I appeal to Pythagoras here:

$$r_{ab}^2 = \sum_{i=1,2,3} ((x_a - x_b)^i)^2 \equiv (x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2 \quad .$$

$$\text{Notation:} \quad r_{ab}^2 \equiv (x_a - x_b)^i (x_a - x_b)^i \; .$$

In terms of this distance the magnitude of the gravitational force is $||\vec{F}_G|| = G\frac{m_a m_b}{r_{ab}^2}$.

Note for the future: It will become increasingly useful to think of the distance r_{ab}^2 as made by adding together the lengths of lots of little line segments:

$$\mathrm{d}s^2 = \mathrm{d}x^i \mathrm{d}x^i$$

$$r_{ab} = \int \mathrm{d}s = \int \mathrm{d}\mathfrak{s} \sqrt{\frac{\mathrm{d}x^i}{\mathrm{d}\mathfrak{s}} \frac{\mathrm{d}x^i}{\mathrm{d}\mathfrak{s}}}.$$

For a straight line, this agrees with our previous expression because we can parametrize the line as $x^i(\mathfrak{s}) = (x_a - x_b)\mathfrak{s}$, with $\mathfrak{s} \in (0, 1)$. I mention this now to emphasize the role in the discussion of the line element (aka the metric) $\mathrm{d}s^2$.

Symmetries of Newton's Laws

What are the symmetries of Newton's Laws? The equations (2) are form-invariant under the substitutions

$$\tilde{x}^i = x^i + a^i, \tilde{t} = t + a^0$$

for some constants a^i, a^0 – the equations look the same in terms of the $\tilde{x}^\mu$. These changes of variables are time and space translations – the invariance of Newton's laws says that there is no special origin of coordinates.

Newton's laws are also invariant under rotations :

$$\tilde{x}^i = R^i_j x^j \quad R^T R = \mathbb{1}. \tag{3}$$

Why this condition (the condition is pronounced 'R is orthogonal' or '$R \in O(3)$)? It preserves the length of a (infinitesimal) line segment:

$$\mathrm{d}s^2 = \mathrm{d}\tilde{x}^i \mathrm{d}\tilde{x}^j \delta_{ij} = (R\mathrm{d}\vec{x})^T \cdot (R\mathrm{d}\vec{x}) = \mathrm{d}x^i \mathrm{d}x^i. \tag{4}$$

And e.g. this distance appears in $||\vec{F}_G|| = G\frac{m_1 m_2}{r_{12}^2}$. If we performed a map where $R^T R \neq \mathbb{1}$, the RHS of Newton's law would change form.

Let me be more explicit about (4), for those of you who want to practice keeping track of upper and lower indices:

$$\mathrm{d}s^2 = \mathrm{d}\tilde{x}^i \mathrm{d}\tilde{x}^j \delta_{ij} = R^i_l \mathrm{d}x^l R^j_k \mathrm{d}x^k \delta_{ij} = \mathrm{d}x^l \mathrm{d}x^k \left(R_{ik} R^i_l\right)$$

Here I defined $R_{ik} \equiv \delta_{ij} R^j_k$ – I used the 'metric' δ_{ij} to lower the index. (This is completely innocuous at this point.) But using $R_{ik} = \left(R^T\right)_{ki}$ we can make this look like matrix multiplication again:

$$\mathrm{d}s^2 = \mathrm{d}x^l \mathrm{d}x^k \left(\left(R^T\right)_{ki} R^i_l\right)$$

– so the condition that the line element is preserved is

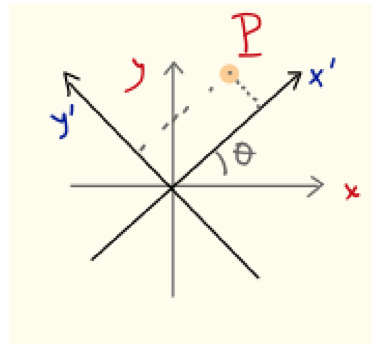
$$\left(R^T\right)_{ki} R^i_l = \delta_{kl}.$$

Comments on rotations

$$\text{ROTATION:} \qquad R^i_k \delta_{ij} R^j_m = \delta_{km} \quad . \tag{5}$$

Focus for a moment on the case of two dimensions. We can parametrize 2d rotations in terms of trig functions. Think of this as solving the equations (5).

We can label the coordinates of a point P in $\mathbb{R}^n$ ($n = 2$ in the figure) by its components along any set of axes we like. They will be related by:



$$x'_i = R_i^j x_j \quad \text{where} \quad R_i^j = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}_i^j = \langle j'|i\rangle$$

is the matrix of overlaps between elements of the primed and unprimed bases. So: using $\mathbb{1} = \sum_j |j'\rangle\langle j'|$, any vector P in $\mathbb{R}^n$ is

$$|P\rangle = \sum_i P^i |i\rangle = \sum_i P^i \left(\sum_j |j'\rangle\langle j'| \right) |i\rangle = \sum_j P^i R_i^j |j'\rangle \ .$$

In three dimensions, we can make an arbitrary rotation by composing rotations about the coordinate axes, each of which looks like a 2d rotation, with an extra identity bit, *e.g.*:

$$(R_z)_i^j = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}_i^j$$

While we're at it, let's define tensors:
Definition: A tensor is a thing that transforms like a tensor.

(You may complain about this definition now, but you will come to see its wisdom.)

By 'transforms', (for now) I mean how it behaves when you rotate your axes, as above. And by 'transforms like a tensor' I mean that all of the indices get an R stuck on to them. Like x^i:

$$x^i \mapsto \tilde{x}^i \equiv R_j^i x^j$$

And like $\frac{\partial}{\partial x^i}$: (use the chain rule)

$$\partial_i \mapsto \tilde{\partial}_i = \frac{\partial x^j}{\partial \tilde{x}^i} \tilde{\partial}_j = (R^{-1})_i^j \partial_{x^j} = (R^t)_i^j \partial_{x^j} \ . \tag{6}$$

A more complicated example would be an object with two indices:

$$T^{ij} \mapsto \tilde{T}^{ij} = R_k^i R_l^j T^{kl}.$$

We could distinguish between 'contravariant' and 'covariant' indices (*i.e.* upper and lower) according to which of the two previous behaviors we see. But for now, this doesn't make much difference – actually it doesn't make any difference at all because of the orthogonality property of a rotation. For Lorentz transformations (and for general coordinate changes) the story will be different.

Clouds on the horizon: Another symmetry of Newton's law is the *Galilean boost*:

$$\tilde{x}^i = x^i + v^i t, \tilde{t} = t. \tag{7}$$

Newton's laws are form-invariant under this transformation, too. Notice that there is a rather trivial sense in which (7) preserves the length of the infinitesimal interval:

$$\mathrm{d}s^2 = \mathrm{d}\tilde{x}^i \mathrm{d}\tilde{x}^j \delta_{ij} = \mathrm{d}x^i \mathrm{d}x^i$$

since time simply does not appear – it's a different kind of thing.

The preceding symmetry transformations comprise the Galilei group: it has ten generators (time translations, 3 spatial translations, 3 rotations and 3 Galilean boosts). It's rightfully called that given how vigorously Galileo emphasized the point that physics looks the same in coordinate systems related by (7). If you haven't read his diatribe on this with the butterflies flying indifferently in every direction, do so at your earliest convenience; it is hilarious. An excerpt is here.

2.2 Maxwell versus Galileo and Newton: Maxwell wins

Let's rediscover the structure of the Lorentz group in the historical way: via the fact that Maxwell's equations are not invariant under (7), but rather have Lorentz symmetry.

Maxwell said ...

$$\begin{aligned} \vec{\nabla} \times \vec{E} + \frac{1}{c}\partial_t \vec{B} &= 0, \qquad \vec{\nabla} \cdot \vec{B} = 0 \\ \vec{\nabla} \times \vec{B} - \frac{1}{c}\partial_t \vec{E} &= \tfrac{4\pi}{c}\vec{J}, \qquad \vec{\nabla} \cdot \vec{E} = 4\pi\rho \end{aligned} \tag{8}$$

.... and there was light. If you like, these equations are empirical facts. Combining Ampere and Faraday, one learns that (*e.g.*̃in vacuum)

$$\left(\partial_t^2 - c^2 \nabla^2\right) \vec{E} = 0$$

– the solutions are waves moving at the speed c, which is a constant appearing in (8) (which is measured by doing experiments with capacitors and stuff).

Maxwell's equations are not inv't under Gal boosts, which change the speed of light.

They *are* invariant under the Poincaré symmetry. Number of generators is the same as Gal: 1 time translation, 3 spatial translations, 3 rotations and 3 (Lorentz!) boosts. The last 3 are the ones which distinguish Lorentz from Galileo, but before we get there, we need to grapple with the fact that we are now studying a field theory.

A more explicit expression of Maxwell's equations is:

$$\begin{aligned} \epsilon^{ijk}\partial_j E_k + \frac{1}{c}\partial_t B^i &= 0, \qquad \partial^i B_i = 0 \\ \epsilon^{ijk}\partial_j B_k - \frac{1}{c}\partial_t E^i &= \tfrac{4\pi}{c} J^i, \qquad \partial_i E^i = 4\pi\rho \end{aligned} \tag{9}$$

Here in writing out the curl: $\left(\vec{\nabla}\times\vec{E}\right)_i = \epsilon_{ijk}\partial_j E_k$ we've introduced the useful Levi-Civita symbol, ϵ_{ijk}. It is a totally anti-symmetric object with $\epsilon_{123} = 1$. It is a "pseudo-tensor": the annoying label 'pseudo' is not a weakening qualifier, but rather an additional bit of information about the way it transforms under rotations that include a parity transformation (*i.e.* those which map a right-handed frame (like xyz) to a left-handed frame (like yxz), and therefore have $\det R = -1$. For those of you who like names, such transformations are in $O(3)$ but not $SO(3)$.) As you'll show on the homework, it transforms like

$$\epsilon_{ijk} \mapsto \tilde{\epsilon}_{ijk} = \epsilon_{\ell mn} R_i^\ell R_j^m R_k^n = (\det R)\,\epsilon_{ijk}.$$

$$R^T R = \mathbb{1} \implies (\det R)^2 = 1 \implies \det R = \pm 1$$

If R preserves a right-handed coordinate frame, $\det R = 1$.

Notice by the way that so far I have not attributed any meaning to upper or lower indices. And we can get away with this when our indices are spatial indices and we are only thinking about rotations because of (6).

Comment about tensor fields

Here is a comment which may head off some confusions about the first problem set. The objects $\vec{E}(x,t)$ and $\vec{B}(x,t)$ are vectors (actually $\vec{B}$ is a pseudovector) at each point in space-time, that is – they are vector *fields*. We've discussed above the rotation properties of vectors and other tensors; now we have to grapple with transforming a vector at each point in space, while at the same time rotating the space.

The rotation is a relabelling $\tilde{x}^i = R_{ij}x^j$, with $R_{ij}R_{kj} = \delta_{ik}$ so that lengths are preserved. As always when thinking about symmetries, it's easy to get mixed up about active and passive transformations. The important thing to keep in mind is that we are just *relabelling* the points (and the axes), and the values of the fields at the points are not changed by this relabelling. So a scalar field (a field with no indices) transforms as

$$\tilde{\phi}(\tilde{x}) = \phi(x).$$

Notice that $\tilde{\phi}$ is a different function of its argument from ϕ; it differs in exactly such a way as to undo the relabelling. So it's NOT true that $\tilde{\phi}(x) \stackrel{?}{=} \phi(x)$, NOR is it true that $\tilde{\phi}(\tilde{x}) \stackrel{?}{=} \phi(Rx)$ which would say the same thing, since R is invertible.

A vector field is an arrow at each point in space; when we rotate our labels, we change our accounting of the *components* of the vector at each point, but must ensure that we don't change the vector itself. So a vector field transforms like

$$\tilde{E}^i(\tilde{x}) = R_{ij}E^j(x).$$

For a clear discussion of this simple but slippery business[3] take a look at page 46 of Zee's book.

The statement that $\vec{B}$ is a pseudovector means that it gets an extra minus sign for parity-reversing rotations:

$$\tilde{B}^i(\tilde{x}) = \det R R_{ij}B^j(x).$$

To make the Poincaré invariance manifest, let's rewrite Maxwell (8) in better notation:

$$\epsilon^{\mu\cdots}\partial_\cdot F_{\cdot\cdot} = 0, \qquad \eta^{\cdot\cdot}\partial_\cdot F_{\mu\cdot} = 4\pi j_\nu.$$

[4] Again this is 4 + 4 equations; let's show that they are the same as (8). Writing them in such a compact way requires some notation (which was designed for this job, so don't be too impressed yet[5]).

In terms of $F_{ij} \equiv \epsilon_{ijk}B^k$ (note that $F_{ij} = -F_{ji}$),

$$\begin{aligned} \partial_j E_k - \partial_k E_j + \frac{1}{c}\partial_t F_{jk} &= 0, \qquad \epsilon^{ijk}\partial_i F_{jk} = 0 \\ \partial_i F_{ij} - \frac{1}{c}\partial_t E_i &= \tfrac{4\pi}{c}J_i, \qquad \partial_i E^i = 4\pi\rho \end{aligned} \tag{10}$$

Introduce $x^0 = ct$. Let $F_{i0} = -F_{0i} = E_i$. $F_{00} = 0$. (Note: F has lower indices and they mean something.) With this notation, the first set of equations can be written as

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0$$

$$\text{better notation:} \quad \epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0.$$

In this notation, the second set is

$$\eta^{\nu\rho}\partial_\nu F_{\mu\rho} = 4\pi j_\mu \tag{11}$$

[3] Thanks to Michael Gartner for reminding me that this is so simple that it's actually quite difficult.

[4] A comment about notation: here and in many places below I will refuse to assign names to dummy indices when they are not required. The ·s indicate the presence of indices which need to be contracted. If you must, imagine that I have given them names, but written them in a font which is too small to see.

[5] See Feynman vol II §25.6 for a sobering comment on this point.

where we've packaged the data about the charges into a collection of four objects:

$$j_0 \equiv -c\rho, j_i \equiv J_i \ .$$

(It is tempting to call this a four-vector, but that is a statement about its transformation laws which remains to be seen. Spoilers: it is in fact a 4-vector.)

Here the quantity

$$\eta^{\nu\rho} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{\nu\rho}$$

makes its first appearance in our discussion. At the moment is is just a notational device to get the correct relative sign between the $\partial_t E$ and the $\vec{\nabla} E$ terms in the second Maxwell equation.

The statement that the particle current is conserved:

$$0 = \frac{c\partial\rho}{\partial\,(ct)} + \vec{\nabla}\cdot\vec{J} \quad \text{now looks like} \quad 0 = -\partial_t j_0 + \partial_i j_i = \eta_{\mu\nu}\frac{\partial}{\partial x^\mu} j_\nu \equiv \partial^\mu j_\mu \equiv \partial_\mu j^\mu. \qquad (12)$$

This was our first meaningful raising of an index. Consistency of Maxwell's equations requires the equation :

$$0 = \partial_\mu \partial_\nu F^{\mu\nu} \quad .$$

It is called the 'Bianchi identity' – 'identity' because it is identically true by antisymmetry of derivatives (as long as F is smooth).

Symmetries of Maxwell equations

Consider the substitution

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu_\nu x^\nu,$$

under which

$$\partial_\mu \mapsto \tilde{\partial}_\mu, \text{ with } \partial_\mu = \Lambda^\nu_\mu \tilde{\partial}_\nu \ .$$

At the moment Λ is a general 4×4 matrix of constants.

If we declare that everything in sight transforms like a *tensor* under these transformations – that is *if*:

$$F_{\mu\nu}(x) \mapsto \tilde{F}_{\mu\nu}(\tilde{x}) \ , \quad F_{\mu\nu}(x) = \Lambda^\rho_\mu \Lambda^\sigma_\nu \tilde{F}_{\rho\sigma}(\tilde{x})$$

$$j_\mu(x) \mapsto \tilde{j}_\mu(\tilde{x}) \ , j_\mu(x) = \Lambda^\nu_\mu \tilde{j}_\nu(\tilde{x})$$

then Maxwell's equations in the new variables are

$$(\det\Lambda)\Lambda^\mu_\kappa \epsilon^{\kappa\nu\rho\sigma}\tilde{\partial}_\nu \tilde{F}_{\rho\sigma} = 0$$

$$\eta^{\nu\rho}\Lambda^\kappa_\nu \tilde{\partial}_\kappa \Lambda^\sigma_\mu \Lambda^\lambda_\rho \tilde{F}_{\sigma\lambda} = 4\pi \Lambda^\aleph_\mu J_\aleph.$$

[6] Assume Λ is invertible, so we can strip off a Λ from this equation to get:

$$\eta^{\nu\rho}\Lambda^{\kappa}_{\nu}\Lambda^{\lambda}_{\rho}\tilde{\partial}_{\kappa}\tilde{F}_{\sigma\lambda} = 4\pi J_{\sigma}.$$

This looks like the original equation (11) in the new variables if Λ satisfies

$$\boxed{\Lambda^{\kappa}_{\nu}\eta^{\nu\rho}\Lambda^{\lambda}_{\rho} = \eta^{\kappa\lambda}} \tag{13}$$

– this is a condition on the collection of numbers Λ:

$$\Lambda^T\eta\Lambda = \eta$$

which is pronounced 'Λ is in $O(3,1)$' (the numbers are the numbers of $+1$s and -1s in the quadratic form η). Note the very strong analogy with (5); the only difference is that the transformation preserves η rather than δ. The ordinary rotation is a special case of Λ with $\Lambda^0_{\mu} = 0, \Lambda^{\mu}_0 = 0$. The point of the notation we've constructed here is to push this analogy in front of our faces.

Are these the right transformation laws for $F_{\mu\nu}$? I will say two things in their defense. First, the rule $F_{\mu\nu} = \Lambda^{\rho}_{\mu}\Lambda^{\sigma}_{\nu}\tilde{F}_{\rho\sigma}$ follows immediately if the vector potential A_μ is a vector field and

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ .$$

Perhaps more convincingly, one can derive pieces of these transformation rules for $\vec{E}$ and $\vec{B}$ by considering the transformations of the charge and current densities that source them. This is described well in the E&M book by Purcell and I will not repeat it.

You might worry further that the transformation laws of j^μ are determined by already by the transformation laws of x^μ – *e.g.* consider the case that the charges are point particles – we don't get to pick the transformation law. We'll see below that the rule above is correct – j^μ really is a 4-vector.

A word of nomenclature: the Lorentz group is the group of Λ^{μ}_{ν} above, isomorphic to $SO(3,1)$. It is a subgroup of the Poincaré group, which includes also spacetime translations, $x^\mu \mapsto \Lambda^{\mu}_{\nu}x^\nu + a^\nu$.

Lorentz transformations of charge and current density

Here is a simple way to explain the Lorentz transformation law for the current. Consider a bunch of charge Q in a (small) volume V moving with velocity $\vec{u}$. The charge and current density are

$$\rho = Q/V, \quad \vec{J} = \rho\vec{u}.$$

[6] Two comments: (1) Already we're running out of letters! Notice that the only meaningful index on the BHS of this equation is μ – all the others are dummies. (2) Notice that the derivative on the RHS is acting only on the $\tilde{F}$ – everything else is constants.

In the rest frame, $\vec{u}_0 = 0, \rho_0 \equiv \frac{Q}{V_0}$. The charge is a scalar, but the volume is contracted in the direction of motion: $V = \frac{1}{\gamma}V_0 = \sqrt{1 - u^2/c^2}V_0$

$$\implies \quad \boxed{\rho = \rho_0\gamma, \vec{J} = \rho_0\gamma\vec{u}}.$$

But this means exactly that

$$J^\mu = (\rho_0\gamma c, \rho_0\gamma\vec{u})^\mu$$

is a 4-vector.

(In lecture, I punted the discussion in the following paragraph until we construct worldline actions.) We're going to have to think about currents made of individual particles, and we'll do this using an action in the next section. But let's think in a little more detail about the form of the current four-vector for a single point particle: Consider a point particle with trajectory $\vec{x}_0(t)$ and charge e. The charge density and current density are only nonzero where the particle is:

$$\rho(t, \vec{x}) = e\delta^{(3)}(\vec{x} - \vec{x}_0(t)) \tag{14}$$

$$\vec{j}(t, \vec{x}) = e\delta^{(3)}(\vec{x} - \vec{x}_0(t))\frac{\mathrm{d}\vec{x}_0}{\mathrm{d}t} \tag{15}$$

$$x_0^\mu(t) \equiv (ct, \vec{x}_0(t))^\mu$$

transforms like

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu_\nu x^\nu$$

$$j^\mu(x) = e\delta^{(3)}(\vec{x} - \vec{x}_0(t))\frac{\mathrm{d}x_0^\mu(t)}{\mathrm{d}t} = e\int_{-\infty}^{\infty} \mathrm{d}t'\delta^{(4)}(x - x_0(t'))\frac{\mathrm{d}x_0^\mu}{\mathrm{d}t'}(t')$$

using $\delta^{(4)}(x - x_0(t')) = \delta^{(3)}(\vec{x} - \vec{x}_0(t'))\delta(x^0 - ct')$. Since we can choose whatever dummy integration variable we want,

$$j^\mu = e\int_{-\infty}^{\infty} \mathrm{d}s\delta^{(4)}(x - x_0(s))\frac{\mathrm{d}x_0^\mu}{\mathrm{d}s}(s) \tag{16}$$

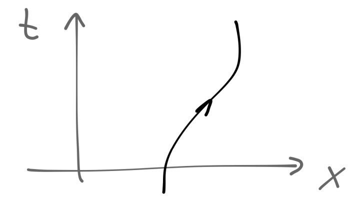
is manifestly a Lorentz 4-vector – it transforms the same way as x^μ.

2.3 Minkowski spacetime

Let's collect the pieces here. We have discovered Minkowski spacetime, the stage on which special relativity happens. This spacetime has convenient global coordinates $x^\mu = (ct, x^i)^\mu$. $\mu = 0, 1, 2, 3$, $i = 1, 2, 3$ or x, y, z.

Our labels on points in this space change under a Lorentz transformation by $x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu_\nu x^\nu$. It's not so weird; we just have to get used to the fact that our time coordinate is a provincial notion of slow-moving creatures such as ourselves.

The trajectory of a particle is a curve in this spacetime. We can describe this trajectory (the worldline) by a parametrized path $\mathfrak{s} \to x^\mu(\mathfrak{s})$. (Note that there is some ambiguity in the choice of parameter along the worldline. For example, you could use time measured by a watch carried by the particle. Or you could use the time measured on your own watch while you sit at rest at $x = 0$.)



Raising and lowering

When we discussed rotations in space, we defined vectors $v^i \to R^i_j v^j$ (like x^i) and co-vectors $\partial_i \to R^j_i \partial_j$ (like ∂_i) (elements of the dual vector space), but they actually transform the same way because of the $O(3)$ property, and it didn't matter. In spacetime, this matters, because of the sign on the time direction. On the other hand we can use $\eta^{\mu\nu}$ to raise Lorentz indices, that is, to turn a vector into a co-vector. So *e.g.* given a covector v_μ, we can make a vector $v^\mu = \eta^{\mu\nu} v_\nu$.

What about the other direction? The inverse matrix is denoted

$$\eta_{\nu\rho} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\nu\rho}$$

– in fact it's the same matrix. Notice that it has lower indices. They satisfy $\eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu_\nu$. So we can use $\eta_{\mu\nu}$ to lower Lorentz indices.

Matrix notation and the inverse of a Lorentz transformation:

During lecture there have been a lot of questions about how to think about the Lorentz condition

$$\Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\lambda = \eta_{\rho\lambda}$$

as a matrix equation. Let us endeavor to remove this confusion more effectively than I did during lecture.

Consider again the rotation condition, and this time let's keep track of the indices on R in $x^i \to R^i_j x^j$. The condition that this preserves the euclidean metric δ_{ij} is:

$$\delta_{ij} \mathrm{d}\tilde{x}^i \mathrm{d}\tilde{x}^j = R^i_k \mathrm{d}x^k R^j_l \mathrm{d}x^l \delta_{ij} \stackrel{!}{=} \delta_{ij} \mathrm{d}x^i \mathrm{d}x^j, \qquad \Leftrightarrow \qquad R^i_k \delta_{ij} R^j_l = \delta_{kl}.$$

Now multiply this equation on the BHS by the inverse of R, $(R^{-1})^l_{\ m}$ which satisfies $R^j_l \, (R^{-1})^l_{\ m} = \delta^j_m$ (and sum over l!):

$$R^i_k \delta_{ij} R^j_l (R^{-1})^l_m \quad = \quad \delta_{kl} (R^{-1})^l_m$$

$$
\begin{aligned}
R^i_k \delta_{im} &= \delta_{kl}(R^{-1})^l_m, \\
\delta^{kl} R^i_k \delta_{im} &= (R^{-1})^l_m,
\end{aligned} \tag{17}
$$

This is an equation for the inverse of R in terms of R. The LHS here is what we mean by R^T if we keep track of up and down.

The same thing for Lorentz transformations (with Λ^{-1} defined to satisfy $\Lambda^\nu_\sigma \left(\Lambda^{-1}\right)^\sigma_\rho = \delta^\nu_\rho$) gives:

$$
\begin{aligned}
\Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\sigma &= \eta_{\rho\sigma} \\
\Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\sigma \left(\Lambda^{-1}\right)^\sigma_\rho &= \eta_{\rho\sigma} \left(\Lambda^{-1}\right)^\sigma_\rho \\
\Lambda^\mu_\rho \eta_{\mu\rho} &= \eta_{\rho\sigma} \left(\Lambda^{-1}\right)^\sigma_\rho \\
\eta^{\rho\sigma} \Lambda^\mu_\rho \eta_{\mu\rho} &= \left(\Lambda^{-1}\right)^\sigma_\rho
\end{aligned} \tag{18}
$$

This is an expression for the inverse of a Lorentz transformation in terms of Λ itself and the Minkowski metric. This reduces to the expression above in the case when Λ is a rotation, which doesn't mix in the time coordinate, and involves some extra minus signs when it does.

Proper length.

It will be useful to introduce a notion of the proper length of the path $x^\mu(\mathfrak{s})$ with $\mathfrak{s} \in [\mathfrak{s}_1, \mathfrak{s}_2]$.

First, if v and w are two 4-vectors – meaning that they transform under Lorentz transformations like

$$
v \mapsto \tilde{v} = \Lambda v, \quad i.e. \;\; v^\mu \mapsto \tilde{v}^\mu = \Lambda^\mu_\nu v^\nu = w_\mu v^\mu = w^\mu v_\mu .
$$

– then we can make a number out of them by

$$
v \cdot w \equiv -v^0 w^0 + \vec{v} \cdot \vec{w} = \eta_{\mu\nu} w^\mu w^\nu
$$

with (again)

$$
\eta^{\nu\rho} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{\nu\rho} .
$$

This number has the virtue that it is invariant under Lorentz transformations by the defining property (13). A special case is the proper length of a 4-vector $\parallel v \parallel^2 \equiv v \cdot v = \eta_{\mu\nu} v^\mu v^\nu$.

A good example of a 4-vector whose proper length we might want to study is the tangent vector to a particle worldline:

$$
\frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} .
$$

Tangent vectors to trajectories of particles that move slower than light have a negative proper length-squared (with our signature convention). For example, a particle which just sits at $x = 0$ and we can take $t(\mathfrak{s}) = \mathfrak{s}, x(\mathfrak{s}) = 0$, so we have

$$\| \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} \|^2 = -\frac{\mathrm{d}t}{\mathrm{d}t}^2 = -1 < 0 \ .$$

(Notice that changing the parametrization $\mathfrak{s}$ will rescale this quantity, but will not change its sign.) Exercise: Show that by a Lorentz boost you cannot change the sign of this quantity.

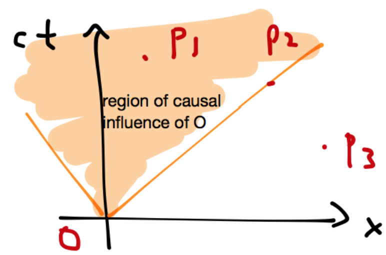
Light moves at the speed of light. A light ray headed in the x direction satisfies $x = ct$, can be parametrized as $t(\mathfrak{s}) = \mathfrak{s}, x(\mathfrak{s}) = c\mathfrak{s}$. So the proper length of a segment of its path (proportional to the proper length of a tangent vector) is

$$\mathrm{d}s^2|_{lightray} = -c^2\mathrm{d}t^2 + \mathrm{d}x^2 = 0$$

(factors of c briefly restored). Rotating the direction in space doesn't change this fact. Proper time does not pass along a light ray.

More generally, the proper distance-squared between the points labelled x^μ and $x^\mu + \mathrm{d}x^\mu$ (compare the Euclidean metric (1)) is

$$\mathrm{d}s^2 \equiv \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -\mathrm{d}t^2 + \mathrm{d}\vec{x}^2 \ .$$



Consider the proper distances between the origin O and the points $P_{1,2,3}$ in the figure.

$$ds^2_{OP_1} < 0 \quad \text{time-like separated}$$

These are points which could both be on the path of a massive particle.

$$ds^2_{OP_2} = 0 \quad \text{light-like separated}$$

$$ds^2_{OP_3} > 0 \quad \text{space-like separated}$$

The set of light-like separated points is called the light-cone at O, and it bounds the *region of causal infuence of* O, the set of points that a massive particle could in principle reach from O if only it tries hard enough.

The proper length of a finite segment of worldline is obtained by adding up the (absolute values of the) lengths of the tangent vectors:

$$\Delta s = \int_{\mathfrak{s}_1}^{\mathfrak{s}_2} \mathrm{d}\mathfrak{s} \sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}\frac{\mathrm{d}x^\nu}{\mathrm{d}\mathfrak{s}}} = \int \mathrm{d}s \ .$$

The last equation is a useful shorthand notation.

Symmetries

The Galilean boost (7) does not preserve the form of the Minkowski line element. Consider 1d for simplicity:

$$\mathrm{d}\tilde{t}^2 - \mathrm{d}\tilde{\vec{x}}^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2 - 2\mathrm{d}\vec{x}\cdot\vec{v}\mathrm{d}t - v^2\mathrm{d}t^2 \neq \mathrm{d}t^2 - \mathrm{d}x^2.$$

It does not even preserve the form of the lightcone.

Lorentz boosts instead:

$$\tilde{x}^\mu = \Lambda^\mu_\nu x^\nu.$$

$$\text{ROTATION:} \quad R^i_k \delta_{ij} R^j_m = \delta_{km} \quad .$$

$$\text{BOOST:} \quad \Lambda^\mu_\rho \eta_{\mu\nu} \Lambda^\nu_\lambda = \eta_{\rho\lambda} \quad .$$

[End of Lecture 2]

Unpacking the Lorentz boost

Consider $D = 1 + 1$ spacetime dimensions[7]. Just as we can parameterize a rotation in two dimensions in terms of trig functions because of the identity $\cos^2\theta + \sin^2\theta = 1$, we can parametrize a $D = 1+1$ boost in terms of hyperbolic trig functions, with $\cosh^2\Upsilon - \sinh^2\Upsilon = 1$.

The $D = 1 + 1$ Minkowski metric is $\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}_{\mu\nu}$.

The condition that a boost Λ^μ_ν preserve the Minkowski line element is

$$\Lambda^\rho_\mu \eta^{\mu\nu} \Lambda^\sigma_\nu = \eta^{\rho\sigma} \quad . \tag{19}$$

A solution of this equation is of the form

$$\Lambda^\mu_\nu = \begin{pmatrix} \cosh\Upsilon & \sinh\Upsilon \\ \sinh\Upsilon & \cosh\Upsilon \end{pmatrix}^\mu_\nu .$$

from which (19) follows by the hyperbolig trig identity.

(The quantity parameterizing the boost Υ is called the rapidity. As you can see from the top row of the equation $\tilde{x}^\mu = \Lambda^\mu_\nu x^\nu$, the rapidity is related to the velocity of the new frame by $\tanh\Upsilon = \frac{u}{c}$.)

[7]Notice the notation: I will try to be consistent about writing the number of dimensions of spacetime in this format; if I just say 3 dimensions, I mean space dimensions.

A useful and memorable form of the above transformation matrix between frames with relative x-velocity u (now with the y and z directions going along for the ride) is:

$$\Lambda = \begin{pmatrix} \gamma & \frac{u}{c}\gamma & 0 & 0 \\ \frac{u}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \gamma \equiv \frac{1}{\sqrt{1-u^2/c^2}} (= \cosh \Upsilon). \tag{20}$$

So in particular,

$$x = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \mapsto \tilde{x} = \Lambda x = \begin{pmatrix} \gamma\left(ct + \frac{u}{c}x\right) \\ \gamma\left(x + \frac{u}{c}ct\right) \\ y \\ z \end{pmatrix} \quad .$$

Review of relativistic dynamics of a particle

So far we've introduced this machinery for doing Lorentz transformations, but haven't said anything about using it to study the dynamics of particles. We'll rederive this from an action principle soon, but let's remind ourselves.

Notice that there are two different things we might mean by velocity.

- The *coordinate velocity* in a particular inertial frame is

$$\vec{u} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \mapsto \tilde{\vec{u}} = \frac{\vec{u} - \vec{v}}{1 - \frac{uv}{c^2}}$$

  You can verify this transformation law using the Lorentz transformation above.

- The *proper velocity* is defined as follows. The proper time along an interval of particle trajectory is defined as $\mathrm{d}\tau$ in:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + \mathrm{d}\vec{x}^2 \equiv -c^2\mathrm{d}\tau^2$$

  – the minus sign makes $\mathrm{d}\tau^2$ positive. Notice that

$$\left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^2 = 1 - u^2/c^2 \implies \frac{\mathrm{d}\tau}{\mathrm{d}t} = \sqrt{1 - u^2/c^2} = 1/\gamma_u.$$

  The proper velocity is then

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}s} = \gamma_u \vec{u}.$$

  Since $\mathrm{d}x^\mu = (c\mathrm{d}t, \mathrm{d}\vec{x})^\mu$ is a 4-vector, so is the proper velocity, $\frac{1}{\mathrm{d}\tau}\mathrm{d}x^\mu = (\gamma c, \gamma \vec{u})^\mu$.

To paraphrase David Griffiths, if you're on an airplane, the coordinate velocity in the rest frame of the ground is the one you care about if you want to know whether you'll have time to go running when you arrive; the proper velocity is the one you care about if you are calculating when you'll have to go to the bathroom.

So objects of the form $a_0(\gamma c, \gamma \vec{u})^\mu$ (where a_0 is a scalar quantity) are 4-vectors. [8] A useful notion of 4-momentum is

$$p^\mu = m_0 \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = (m_0 \gamma c, m_0 \gamma \vec{v})^\mu$$

which is a 4-vector. If this 4-momentum is conserved in one frame $\frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} = 0$ then $\frac{\mathrm{d}\tilde{p}^\mu}{\mathrm{d}\tau} = 0$ in another frame. (This is not true of m_0 times the coordinate velocity.) And its time component is $p^0 = m_0 \gamma c$, that is, $E = p^0 c = m(v) c^2$. In the rest frame $v = 0$, this is $E_0 = m_0 c^2$.

The relativistic Newton laws are then:

$$\vec{F} = \frac{\mathrm{d}}{\mathrm{d}t}\vec{p} \;\; \text{still}$$

$$\vec{p} = m\vec{v} \;\; \text{still}$$

$$\text{but} \quad m = m(v) = m_0 \gamma.$$

Let's check that energy $E = m(v)c^2$ is conserved according to these laws. A force does the following work on our particle:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \vec{F} \cdot \vec{v} = \frac{\mathrm{d}}{\mathrm{d}t}(m(v)\vec{v}) \cdot v$$

$$2m \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(m(v)c^2\right) = \vec{v} \cdot \frac{\mathrm{d}}{\mathrm{d}t}(m\vec{v}) \right)$$

$$c^2 2m \frac{\mathrm{d}m}{\mathrm{d}t} = 2m\vec{v} \cdot \frac{\mathrm{d}}{\mathrm{d}t}(m\vec{v})$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m^2 c^2\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(m\vec{v}\right)^2 \implies (mc)^2 = (mv)^2 + \text{const.}$$

$$v = 0 : m^2(v=0)c^2 \equiv m_0 c^2 = \text{const}$$

$$\implies m(v)^2 c^2 = m(v)^2 v^2 + m_0^2 c^2 \implies m(v) = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} = m_0 \gamma_v.$$

[8] We have already seen an example of something of this form in our discussion of the charge current in Maxwell's equations:

$$j^\mu = \rho_0 \left(\gamma c, \gamma \vec{u}\right)^\mu$$

where ρ_0 is the charge density in the rest frame of the charge. So now you believe me that j^μ really is a 4-vector.

Let me emphasize that the reason to go through all this trouble worrying about how things transform (and making a notation which makes it manifest) is because we want to build physical theories with these symmetries. A way to guarantee this is to make actions which are invariant (next). But by construction any object we make out of tensors where all the indices are contracted is Lorentz invariant. (A similar strategy will work for general coordinate invariance later.)

2.4 Non-inertial frames versus gravitational fields

[Landau-Lifshitz volume 2, §82] Now that we understand special relativity, we can make a more concise statement of the EEP:

It says that *locally* spacetime looks like Minkowski spacetime. Let's think a bit about that dangerous word 'locally'.

It is crucial to understand the difference between an actual gravitational field and just using bad (meaning, non-inertial) coordinates for a situation with no field. It is a slippery thing: consider the transformation to a uniformly accelerating frame[9]:

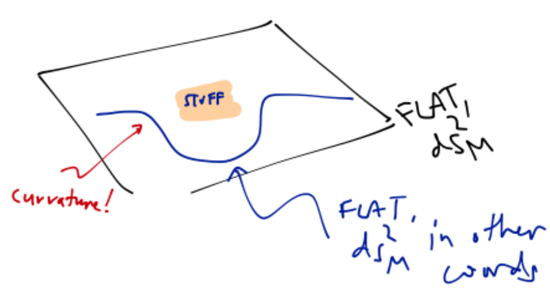
$$\tilde{x} = x - \frac{1}{2}at^2, \ \tilde{t} = t.$$

You could imagine figuring out $\mathrm{d}\tilde{x} = \mathrm{d}x - at\mathrm{d}t, \mathrm{d}\tilde{t} = \mathrm{d}t$ and plugging this into $\mathrm{d}s^2 = \mathrm{d}s^2_{\text{Mink}}$ to derive the line element experienced by someone in Minkowski space using these non-inertial coordinates. This could be a useful thing – *e.g.* you could use it to derive inertial forces (like the centrifugal and coriolis forces; see problem set 3). You'll get something that looks like

$$\mathrm{d}s^2_{\text{Mink}} = \tilde{g}_{\mu\nu}(\tilde{x})\mathrm{d}\tilde{x}^\mu \mathrm{d}\tilde{x}^\nu.$$

It will no longer be a sum of squares; there will be cross terms proportional to $\mathrm{d}\tilde{x}\mathrm{d}\tilde{t} = \mathrm{d}\tilde{t}\mathrm{d}\tilde{x}$, and the coefficients $\tilde{g}_{\mu\nu}$ will depend on the new coordinates. You can get some pretty complicated things this way. But they are still just Minkowski space.

But this happens *everywhere*, even at $x = \infty$. In contrast, a gravitational field from a localized object happens only near the object. The EEP says that *locally*, we can choose coordinates where $\mathrm{d}s^2 \simeq \mathrm{d}s^2_{\text{Mink}}$, but demanding that the coordinates go back to what they were without the object far



[9] A similar statement may be made about a uniformly rotating frame:

$$\tilde{x} = R_{\theta=\omega t}x, \tilde{t} = t$$

where R_θ is a rotation by angle θ.

away forces something to happen in between. That something is *curvature*.

Evidence that there is room for using $g_{\mu\nu}$ as dynamical variables (that is, that not every such collection of functions can be obtained by a coordinate transformation) comes from the following counting: this is a collection of functions; in $D = 3+1$ there are 4 diagonal entries ($\mu = \nu$) and because of the symmetry $\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}x^\nu \mathrm{d}x^\mu$ there are $\frac{4\cdot 3}{2} = 6$ off-diagonal ($\mu \neq \nu$) entries, so 10 altogether. But an *arbitrary* coordinate transformation $x^\mu \to \tilde{x}^\mu(x)$ is only 4 functions. So there is room for something good to happen.

3 Actions

So maybe (I hope) I've convinced you that it's a good idea to describe gravitational interactions by letting the metric on spacetime be a dynamical variable. To figure out how to do this, it will be very useful to be able to construct *action functionals* for the stuff that we want to interact gravitationally. In case you don't remember, the action is a single number associated to every configuration of the system (*i.e.* a functional) whose extremization gives the equations of motion.

As Zee says, physics is where the action is. It is also usually true that the action is where the physics is[10].

3.1 Reminder about Calculus of Variations

We are going to need to think about functionals – things that eat functions and give numbers – and how they vary as we vary their arguments. We'll begin by thinking about functions of one variable, which let's think of as the (1d) position of a particle as a function of time, $x(t)$.

The basic equation of the calculus of variations is:

$$\boxed{\frac{\delta x(t)}{\delta x(s)} = \delta(t-s)}. \tag{21}$$

From this rule and integration by parts we can get everything we need. For example, let's ask how does the potential term in the action $S_V[x] = \int dt V(x(t))$ vary if we vary the path

[10]The exceptions come from the path integral measure in quantum mechanics. A story for a different day.

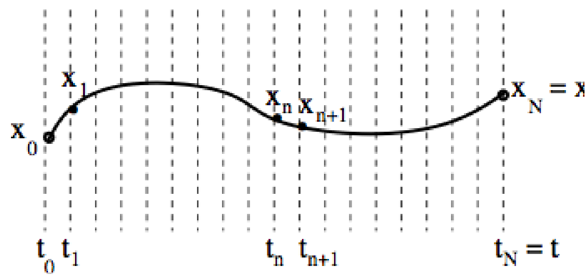
of the particle. Using the chain rule, we have:

$$\delta S_V = \int ds \delta x(s) \frac{\delta \int dt V(x(t))}{\delta x(s)} = \int ds \delta x(s) \int dt \partial_x V(x(t)) \delta(t-s) = \int dt \delta x(t) \partial_x V(x(t)). \tag{22}$$

We could rewrite this information as :

$$\frac{\delta}{\delta x(s)} \int dt V(x(t)) = V'(x(s)).$$

If you are unhappy with thinking of (22) as a use of the chain rule, think of time as taking on a discrete set of values t_n (this is what you have to do to define calculus anyway) and let $x(t_n) \equiv x_n$. Now instead of a functional $S_V[x(t)]$ we just have a function of several variables $S_V(x_n) = \sum_n V(x_n)$. The basic equation of calculus of variations is perhaps more obvious now:



[picture from Herman Verlinde]

$$\frac{\partial x_n}{\partial x_m} = \delta_{nm}$$

and the manipulation we did above is

$$\delta S_V = \sum_m \delta x_m \partial_{x_m} S_V = \sum_m \delta x_m \partial_{x_m} \sum_n V(x_n) = \sum_m \sum_n \delta x_m V'(x_n) \delta_{nm} = \sum_n \delta x_n V'(x_n).$$

What about the kinetic term $S_T[x] \equiv \int dt \frac{1}{2} M \dot{x}^2$? Here we need integration by parts:

$$\frac{\delta}{\delta x(s)} S_T[x] = \frac{2}{2} M \int dt \dot{x}(t) \partial_t \frac{\delta x(t)}{\delta x(s)} = M \int dt \dot{x}(t) \partial_t \delta(t-s) = -M \int dt \ddot{x}(t) \delta(t-s) = -M\ddot{x}(s).$$

Combining the two terms together into $S = S_T - S_V$ we find the equation of motion

$$0 = \frac{\delta}{\delta x(t)} S = -M\ddot{x} - V'$$

i.e. Newton's law.

More generally, you may feel comfortable with lagrangian mechanics: given $L(q, \dot{q})$, the EoM are given by the Euler-Lagrange equations. I can never remember these equations, but they are very easy to derive from the action principle:

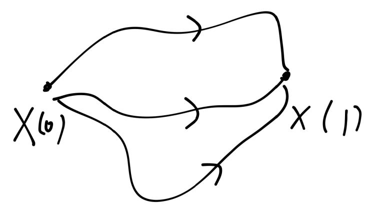
$$0 = \frac{\delta}{\delta q(t)} \underbrace{S[q]}_{= \int \mathrm{d}s L(q(s), \frac{\mathrm{d}}{\mathrm{d}s} q(s))} \overset{\text{chain rule}}{=} \int \mathrm{d}s \left(\frac{\partial L}{\partial q(s)} \frac{\delta q(s)}{\delta q(t)} + \frac{\partial L}{\partial \dot{q}(s)} \underbrace{\frac{\delta \dot{q}(s)}{\delta q(t)}}_{= \frac{\mathrm{d}}{\mathrm{d}s} \delta(s-t)} \right)$$

$$\stackrel{IBP}{=} \quad \int \mathrm{d}s \delta(s-t) \left(\frac{\partial L}{\partial q(s)} - \frac{\mathrm{d}}{\mathrm{d}s} \frac{\partial L}{\partial \dot{q}(s)} \right) = \frac{\partial L}{\partial q(t)} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}(t)}. \tag{23}$$

(Notice at the last step that L means $L(q(t), \dot{q}(t))$.)

Another relevant example: What's the shortest distance between two points in Euclidean space?

The length of a path between the two points (in flat space, where $\mathrm{d}s^2 = \mathrm{d}x^i \mathrm{d}x^i$) is



$$S[x] = \int \mathrm{d}s = \int \sqrt{\dot{x}^2 + \dot{y}^2} \mathrm{d}\mathfrak{s} \equiv \int \sqrt{\dot{x}^i \dot{x}^i}$$

$\dot{x} \equiv \frac{\mathrm{d}x}{\mathrm{d}\mathfrak{s}}$. $\mathfrak{s}$ is an arbitrary parameter. We should consider only paths which go between the given points, *i.e.* that satisfy $x(0) = x_0, x(1) = x_1$.

An extremal path $\underline{x}(\mathfrak{s})$ satisfies

$$0 \stackrel{!}{=} \frac{\delta S}{\delta x^i(\mathfrak{s})}|_{x=\underline{x}} = -\partial_{\mathfrak{s}} \left(\frac{\dot{x}^i}{\sqrt{\dot{x}^2}} \right) \tag{24}$$

This is solved if $0 = \ddot{\underline{x}}$; a solution satisfying the boundary conditions is $x(\mathfrak{s}) = (1-\mathfrak{s})x_0 + \mathfrak{s}x_1$.

In lecture the question arose: are there solutions of (24) which do not have $\ddot{x}^i = 0$? To see that there are not, notice that the parameter $\mathfrak{s}$ is complely arbitrary. If I reparametrize $\mathfrak{s} \mapsto \tilde{\mathfrak{s}}(\mathfrak{s})$, the length $S[x]$ does not change. It will change the lengths of the tangent vectors $T^i = \frac{\mathrm{d}x^i}{\mathrm{d}\mathfrak{s}} = \dot{x}^i$. We can use this freedom to our advantage. Let's choose $\mathfrak{s}$ so that the lengths of the tangent vectors are constant, that is, so that $\dot{x}^i \dot{x}^i$ does not depend on $\mathfrak{s}$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \left(\dot{x}^i \dot{x}^i \right)$$

(Such a parameter is called an *affine parameter*.) One way to achieve this is simply to set the parameter $\mathfrak{s}$ equal to the distance along the worldline.

By making such a choice, the terms where the $\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}}$ hits the $\frac{1}{\sqrt{\dot{x}^2}}$ are zero. Since this choice doesn't change the action, it doesn't change the equations of motion and there are no other solutions where these terms matter. (See Zee p. 125 for further discussion of this point in the same context.)

3.2 Covariant action for a relativistic particle

To understand better the RHS of Maxwell's equation, and to gain perspective on Minkowski space, we'll now study the dynamics of a particle in Minkowski space. This discussion will be useful for several reasons. We'll use this kind of action below to understand the motion of particles in curved space. And it will be another encounter with the menace of coordinate invariance (in $D = 0 + 1$). In fact, it can be thought of as a version of general relativity in $D = 0 + 1$.

$$S[x] = mc^2 \int \mathrm{d}\tau = \int_{-\infty}^{\infty} \mathrm{d}\mathfrak{s} L_0 \ .$$

$$L_0 = -mc^2 \frac{\mathrm{d}\tau}{\mathrm{d}\mathfrak{s}} = -m\sqrt{-\left(\frac{\mathrm{d}x}{\mathrm{d}\mathfrak{s}}\right)^2} = -m\sqrt{-\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}\frac{dx^\nu}{\mathrm{d}\mathfrak{s}}}$$

What is this quantity? It's the (Lorentz-invariant) proper time of the worldline, $-mc^2 \int \mathrm{d}\tau$ ($\mathrm{d}s^2 \equiv -c^2\mathrm{d}\tau^2$ – the sign is the price we pay for our signature convention where time is the weird one.) in units of the mass. Notice that the object S has dimensions of action; the overall normalization doesn't matter in classical physics, which only cares about differences of action, but it is not a coincidence that it has the same units as $\hbar$.

Let $\dot{x}^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}$. The canonical momentum is

$$p_\mu \equiv \frac{\partial L_0}{\partial \dot{x}^\mu} = \frac{mc\eta_{\mu\nu}\dot{x}^\nu}{\sqrt{-\dot{x}^2}}.$$

(Beware restoration of cs.)

A useful book-keeping fact: when we differentiate with respect to a covariant vector $\frac{\partial}{\partial x^\mu}$ (where x^μ is a thing with an upper index) we get a contravariant object – a thing with an lower index.

The components p^μ are not all independent – there is a *constraint*:

$$\eta^{\mu\nu} p_\mu p_\nu = \frac{m^2c^2\dot{x}^2}{-\dot{x}^2} = -m^2c^2.$$

That is:

$$\left(p^0\right)^2 = \vec{p}^2 + m^2c^2 \ .$$

But $p^0 = E/c$ so this is

$$E = \sqrt{\vec{p}^2c^2 + m^2c^4} \ ,$$

the Einstein relation.

Why was there a constraint? It's because we are using more variables than we have a right to: the parameterization $\mathfrak{s}$ along the worldline is arbitrary, and does not affect the physics. [11]

HW: show that an arbitrary change of worldline coordinates $\mathfrak{s} \mapsto \mathfrak{s}(\tilde{\mathfrak{s}})$ preserves S.

[11] Note that I will reserve τ for the proper time and will use weird symbols like $\mathfrak{s}$ for arbitrary worldline parameters.

We should also think about the equations of motion (EoM):

$$0 = \frac{\delta S}{\delta x^\mu(\mathfrak{s})} = -\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \underbrace{\eta_{\mu\nu}\left(\frac{mc\dot{x}^\nu}{\sqrt{-\dot{x}^2}}\right)}_{=\eta_{\mu\nu}p^\nu = p_\mu}$$

Notice that this calculation is formally identical to finding the shortest path in euclidean space; the only difference is that now our path involves the time coordinate, and the metric is Minkowski and not Euclid.

The RHS is (minus) the mass times the covariant acceleration – the relativistic generalization of $-ma$, as promised. This equation expresses the conservation of momentum, a consequence of the translation invariance of the action S.

To make this look more familiar, use time as the worldline parameter: $x^\mu = (c\mathfrak{s}, \vec{x}(\mathfrak{s}))^\mu$. In the NR limit ($v = |\frac{\mathrm{d}\vec{x}}{\mathrm{d}\mathfrak{s}}| \ll c$), the spatial components reduce to $m\ddot{\vec{x}}$, and the time component gives zero. (The other terms are what prevent us from accelerating a particle faster than the speed of light.)

3.3 Covariant action for E&M coupled to a charged particle

Now we're going to couple the particle to E&M. The EM field will tell the particle how to move (it will exert a force), and at the same time, the particle will tell the EM field what to do (the particle represents a charge current). So this is just the kind of thing we'll need to generalize to the gravity case.

3.3.1 Maxwell action in flat spacetime

We've already rewritten Maxwell's equations in a Lorentz-covariant way:

$$\underbrace{\epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0}_{\text{Bianchi identity}}, \qquad \underbrace{\partial^\nu F_{\mu\nu} = 4\pi j_\mu}_{\text{Maxwell's equations}} \quad .$$

The first equation is automatic – the 'Bianchi identity' – if F is somebody's (3+1d) curl:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

in terms of a smooth vector potential A_μ. (Because of this I'll refer to just the equations on the right as Maxwell's equations from now on.) This suggests that it might be a good idea to think of A as the independent dynamical variables, rather than F. On the other hand, changing A by

$$A_\mu \rightsquigarrow A_\mu + \partial_\nu \lambda, \qquad F_{\mu\nu} \rightsquigarrow F_{\mu\nu} + (\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)\,\lambda = F_{\mu\nu} \tag{25}$$

doesn't change the EM field strengths. This is a redundancy of the description in terms of A.[12] [End of Lecture 3]

So we should seek an action for A which

1. is *local*: $S[\text{fields}] = \int d^4x \mathcal{L}(\text{fields and } \partial_\mu\text{fields at } x^\mu)$
2. is invariant under this 'gauge redundancy' (25). This is automatic if the action depends on A only via F.
3. and is invariant under Lorentz symmetry. This is automatic as long as we contract all the indices, using the metric $\eta_{\mu\nu}$, if necessary.

Such an action is

$$S_{EM}[A] = \int d^4x \left(-\frac{1}{16\pi c} F_{\mu\nu}F^{\mu\nu} + \frac{1}{c} A_\mu j^\mu \right) \quad . \tag{26}$$

Indices are raised and lowered with $\eta^{\mu\nu}$. [13] The EoM are:

$$0 = \frac{\delta S_{EM}}{\delta A_\mu(x)} = \frac{1}{4\pi c}\partial_\nu F^{\mu\nu} + \frac{1}{c} j^\mu \tag{27}$$

So by demanding a description where the Bianchi identity was obvious we were led pretty inexorably to the (rest of the) Maxwell equations.

Notice that the action (26) is gauge invariant only if $\partial_\mu j^\mu = 0$ – if the current is conserved. We observed earlier that this was a consistency condition for Maxwell's equations.

Now let's talk about who is j: Really j is whatever else appears in the EoM for the Maxwell field. That is, if there are other terms in the action which depend on A, then when we vary

[12] Some of you may know that quantumly there are some observable quantities made from A which can be nonzero even when F is zero, such as a Bohm-Aharonov phase or Wilson line $e^{\mathbf{i}\oint A}$. These quantities are also gauge invariant.

[13] There are other terms we could add consistent with the above demands. The next one to consider is

$$L_\theta \equiv F_{\mu\nu}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$$

This is a total derivative and does not affect the equations of motion. Other terms we could consider, like:

$$L_8 = \frac{1}{M^4}FFFF$$

(with various ways of contracting the indices) involve more powers of F or more derivatives. Dimensional analysis then forces their coefficient to be a power of some mass scale (M above). Here we appeal to experiment to say that these mass scales are large enough that we can ignore them for low-energy physics. If we lived in $D = 2+1$ dimensions, a term we should consider is the *Chern-Simons* term,

$$S_{CS}[A] = \frac{k}{4\pi}\int \epsilon^{\mu\nu\rho}A_\mu F_{\nu\rho}$$

which depends explicitly on A but is nevertheless gauge invariant.

the action with respect to A, instead of just (27) with $j = 0$, we'll get a source term. (Notice that in principle this source term can depend on A.)

3.3.2 Worldline action for a charged particle

So far the particle we've been studying doesn't care about any Maxwell field. How do we make it care? We add a term to its action which involves the gauge field. The most minimal way to do this, which is rightly called 'minimal coupling' is:

$$S[x] = \int_{-\infty}^{\infty} \mathrm{d}s\; L_0 + q \int_{wl} A$$

Again A is the vector potential: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. I wrote the second term here in a very telegraphic but natural way which requires some explanation. 'wl' stands for worldline. It is a line integral, which we can unpack in two steps:

$$\int_{wl} A = \int_{wl} A_\mu \mathrm{d}x^\mu = \int_{-\infty}^{\infty} \mathrm{d}\mathfrak{s} \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} A_\mu(x(\mathfrak{s})). \tag{28}$$

Each step is just the chain rule. Notice how happily the index of the vector potential fits with the measure $\mathrm{d}x^\mu$ along the worldline. In fact, it is useful to think of the vector potential $A = A_\mu \mathrm{d}x^\mu$ as a *one-form*: a thing that eats 1d paths C and produces numbers via the above pairing, $\int_C A$. A virtue of the compact notation is that it makes it clear that this quantity does not depend on how we parametrize the worldline.

What's the current produced by the particle? That comes right out of our expression for the action:

$$j^\mu(x) = \frac{\delta S_{\text{charges}}}{\delta A_\mu(x)}.$$

Let's think about the resulting expression:

$$j^\mu = e \int_{-\infty}^{\infty} \mathrm{d}\mathfrak{s}\, \delta^{(4)}(x - x_0(\mathfrak{s})) \frac{\mathrm{d}x_0^\mu}{\mathrm{d}\mathfrak{s}}(\mathfrak{s}) \tag{29}$$

Notice that this is manifestly a 4-vector, since x_0^μ is a 4-vector, and therefore so is $\frac{\mathrm{d}x_0^\mu}{\mathrm{d}\mathfrak{s}}$.

Notice also that s is a dummy variable here. We could for example pick $\mathfrak{s} = t$ to make the $\delta(t - t_0(\mathfrak{s}))$ simple. In that case we get

$$j^\mu(x) = e\delta^{(3)}(\vec{x} - \vec{x}_0(t)) \frac{\mathrm{d}x_0^\mu(t)}{\mathrm{d}t}$$

or in components:

$$\begin{aligned} \rho(t, \vec{x}) &= e\delta^{(3)}(\vec{x} - \vec{x}_0(t)) && (30) \\ \vec{j}(t, \vec{x}) &= e\delta^{(3)}(\vec{x} - \vec{x}_0(t)) \frac{\mathrm{d}\vec{x}_0}{\mathrm{d}t}\ . && (31) \end{aligned}$$

Does this current density satisfy the continuity equation – is the stuff conserved? I claim that, yes, $\partial_\mu j^\mu = 0$ – as long as the worldlines never end. You'll show this on the problem set.

Now let's put together all the pieces of action we've developed: With

$$S_{\text{total}}[A, x_0] = -mc \int_{\text{worldline}} \mathrm{d}\mathfrak{s} \sqrt{-\left(\frac{\mathrm{d}x_0}{\mathrm{d}\mathfrak{s}}\right)^2} + \frac{e}{c}\int_{\text{worldline}} A - \frac{1}{16\pi c}\int F^2 \mathrm{d}^4 x \tag{32}$$

we have the EoM

$$0 = \frac{\delta S_{\text{total}}}{\delta A_\mu(x)} \;\; \text{(Maxwell)}. \qquad 0 = \frac{\delta S_{\text{total}}}{\delta x_0^\mu(\mathfrak{s})} \;\; \text{(Lorentz force law)}.$$

The latter is

$$0 = \frac{\delta S_{\text{total}}}{\delta x_0^\mu(\mathfrak{s})} = -\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} p_\mu - \frac{e}{c} F_{\mu\nu} \frac{\mathrm{d}x_0^\nu}{\mathrm{d}\mathfrak{s}}$$

(Warning: the $\partial_\nu A_\mu$ term comes from $A_\nu \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \frac{\delta x^\nu}{\delta x^\mu}$.) The first term we discussed in §3.2. Again use time as the worldline parameter: $x_0^\mu = (c\mathfrak{s}, \vec{x}(\mathfrak{s}))^\mu$. The second term is (for any v) gives

$$e\vec{E} + \frac{e}{c}\vec{B} \times \dot{\vec{x}} \; ,$$

the Lorentz force.

The way this works out is direct foreshadowing of the way Einstein's equations will arise.

3.4 The appearance of the metric tensor

[Zee §IV] So by now you agree that the action for a relativistic free massive particle is proportional to the proper time:

$$S = -m \int \mathrm{d}\tau = -m \int \sqrt{-\eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu}$$

A fruitful question: How do we make this particle interact with an external static potential in a Lorentz invariant way? The non-relativistic limit is

$$S_{NR}[x] = \int \mathrm{d}t \left(\frac{1}{2} m \dot{\vec{x}}^2 - V(x) \right) = S_T - S_V$$

How to relativisticize this? If you think about this for a while you will discover that there are two options for where to put the potential, which I will give names following Zee (chapter IV):

$$\text{Option E:} \quad S_E[x] = -\int \left(m\sqrt{-\eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu} + V(x)\mathrm{d}t \right)$$

S_E reduces to S_{NR} if $\dot{x} \ll c$.

$$\text{Option G:} \quad S_G[x] = -m \int \sqrt{\left(1 + \frac{2V}{m}\right) \mathrm{d}t^2 - \mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x}} \, . \tag{33}$$

S_G reduces to S_{NR} if we assume $\dot{x} \ll c$ *and* $V \ll m$ (the potential is small compared to the particle's rest mass). Note that the factor of 2 is required to cancel the $\frac{1}{2}$ from the Taylor expansion of the sqrt:

$$\sqrt{1 + \text{☼}} \simeq 1 + \frac{1}{2}\text{☼} + \mathcal{O}(\text{☼}^2). \tag{34}$$

Explicitly, let's parametrize the path by lab-frame time t:

$$S_G = -m \int \mathrm{d}t \sqrt{\left(1 + \frac{2V}{m}\right) - \dot{\vec{x}}^2} \, .$$

If $\dot{x} \ll c$ and $V \ll m$ we can use (34) with $\text{☼} \equiv \frac{2V}{m} - \dot{\vec{x}}^2$.

How to make Option E manifestly Lorentz invariant? We have to let the potential transform as a (co-)vector under Lorentz transformations (and under general coordinate transformations):

$$S_E = \int \left(m\sqrt{-\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu} - qA_\mu(x)\mathrm{d}x^\mu \right)$$

$$A_\mu \mathrm{d}x^\mu = -A_0 \mathrm{d}t + \vec{A} \cdot \mathrm{d}\vec{x}. \quad qA_0 = -V, \quad \vec{A} = 0.$$

We've seen this before (in §3.3.2).

And how to relativisticize Option G? The answer is:

$$S_G = -m \int \sqrt{-g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu}$$

with $g_{00} = 1 + \frac{2V}{m}$, $g_{ij} = \delta_{ij}$. Curved spacetime! Now $g_{\mu\nu}$ should transform as a tensor. S_G is proportional to $\int \mathrm{d}\tau$, the proper length in the spacetime with line element $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$.

The Lorentz force law comes from varying S_E wrt x:

$$\frac{\delta S_E[x]}{\delta x_\mu(\mathfrak{s})} = -m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\mathfrak{s}^2} + qF^\mu_\nu \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} \, .$$

What happens if we vary S_G wrt x? Using an affine worldline parameter like in §3.1 we find

$$\frac{\delta S_G[x]}{\delta x_\mu(\mathfrak{s})} = -m\frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\mathfrak{s}^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\mathfrak{s}}\frac{\mathrm{d}x^\rho}{\mathrm{d}\mathfrak{s}} \, .$$

You can figure out what is Γ by making the variation explicitly. If the metric is independent of x (like in Minkowski space), Γ is zero and this is the equation for a straight line.

So by now you are ready to understand the geodesic equation in a general metric $g_{\mu\nu}$. It just follows by varying the action

$$S[x] = -m \int \mathrm{d}\tau = -m \int \mathrm{d}\mathfrak{s} \sqrt{-\dot{x}^\mu \dot{x}^\nu g_{\mu\nu}(x)}$$

just like the action for a relativistic particle, but with the replacement $\eta_{\mu\nu} \to g_{\mu\nu}(x)$. But we'll postpone that discussion for a bit, until after we learn how to think about $g_{\mu\nu}$ a bit better.

A word of advice I wish someone had given me: in any metric where you actually care about the geodesics, you should not start with the equation above, which involves those awful Γs. Just ignore that equation, except for the principle of it. Rather, you should plug the metric into the action (taking advantage of the symmetries), and then vary that simpler action, to get directly the EoM you care about, without ever finding the Γs.

3.5 Toy model of gravity

[Zee p. 119, p. 145] This is another example where the particles tell the field what to do, and the field tells the particles what to do.

Consider the following functional for a scalar field $\Phi(\vec{x})$ in d-dimensional (flat) *space* (no time):

$$E[\Phi] = \int \mathrm{d}^d x \left(\frac{1}{8\pi G} \vec{\nabla}\Phi \cdot \vec{\nabla}\Phi + \rho(x)\Phi(x) \right).$$

What other rotation-invariant functional could we write? Extremizing (in fact, minimizing) this functional gives

$$0 = \frac{\delta E}{\delta \Phi} = -\frac{1}{4\pi G}\nabla^2\Phi + \rho$$

– the equation for the Newtonian potential, given a mass density function $\rho(\vec{x})$. If we are in $d = 3$ space dimensions and $\rho = \delta^3(x)$, this produces the inverse-square force law[14]. In other dimensions, other powers. E is the energy of the gravitational potential.

Notice that we can get the same equation from the action

$$S_G[\Phi] = -\int \mathrm{d}t\mathrm{d}^d x \left(\frac{1}{8\pi G} \vec{\nabla}\Phi \cdot \vec{\nabla}\Phi + \rho(x)\Phi(x) \right) \tag{35}$$

(the sign here is an affection in classical physics). Notice also that this is a *crazy* action for a scalar field, in that it involves no time derivatives at all. That's the action at a distance.

We could do even a bit better by including the dynamics of the mass density ρ itself, by assuming that it's made of particles:

$$S[\Phi, \vec{q}] = \int \mathrm{d}t \left(\frac{1}{2} m\dot{\vec{q}}^2 - m\Phi(q(t)) \right) - \int \mathrm{d}t\mathrm{d}^d x \frac{1}{8\pi G} \vec{\nabla}\Phi \cdot \vec{\nabla}\Phi$$

– the $\rho(x)\Phi(x)$ in (35) is made up by

$$\int \mathrm{d}t\mathrm{d}^3 x \rho(x)\Phi(x) = \int \mathrm{d}t m\Phi(q(t))$$

which is true if

$$\rho(x) = m\delta^3(x - q(t))$$

just the expression you would expect for the mass density from a particle. We showed before that varying this action with respect to q gives Newton's law, $m\vec{a} = \vec{F}_G$. We could add more particles so that they interact with each other via the Newtonian potential.

[End of Lecture 4]

[14] Solve this equation in fourier space: $\Phi_k \equiv \int \mathrm{d}^d x\, e^{-\mathbf{i}\vec{k}\cdot\vec{x}}\Phi(x)$ satisfies $-k^2\Phi_k = a$ for some number a, so $\Phi(x) \sim \int đ^d k \frac{e^{+\mathbf{i}\vec{k}\cdot\vec{x}}}{k^2} \sim \frac{1}{r^{d-2}}$.

4 Stress-energy-momentum tensors, first pass

Gravity couples to mass, but special relativity says that mass and energy are fungible by a change of reference frame. So relativistic gravity couples to energy. So we'd better understand what we mean by energy.

Energy from a Lagrangian system:

$$L = L(q(t), \dot{q}(t)).$$

If $\partial_t L = 0$, energy is conserved. More precisely, the momentum is $p(t) = \frac{\partial L}{\partial \dot{q}(t)}$ and the Hamiltonian is obtained by a Legendre transformation:

$$H = p(t)\dot{q}(t) - L, \qquad \frac{\mathrm{d}H}{\mathrm{d}t} = 0.$$

More generally consider a field theory in $D > 0 + 1$:

$$S[\phi] = \int \mathrm{d}^D x \mathcal{L}(\phi, \partial_\mu \phi)$$

Suppose $\partial_\mu \mathcal{L} = 0$. Then the replacement $\phi(x^\mu) \mapsto \phi(x^\mu + a^\mu) \sim \phi(x) + a^\mu \partial_\mu \phi + ...$ will be a *symmetry*. Rather D symmetries. That means D Noether currents. Noether method gives

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_A)} \partial_\nu \phi_A - \delta^\mu_\nu \mathcal{L}$$

as the conserved currents. (Note the sum over fields. Omit from now on.) Notice that the ν index here is a *label* on these currents, indicating which translation symmetry gives rise to the associated current. The conservation laws are therefore $\partial_\mu T^\mu_\nu = 0$, which follows from the EoM

$$0 = \frac{\delta S}{\delta \phi} = -\partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} + \frac{\partial \mathcal{L}}{\partial \phi}.$$

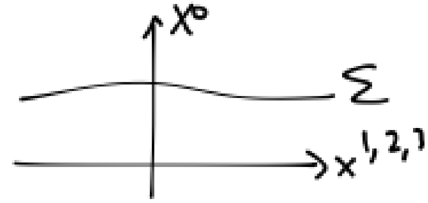
$T^0_0 = \frac{\partial \mathcal{L}}{\partial(\dot{\phi}_A)} \dot{\phi}_A - \mathcal{L} = \mathcal{H}$: energy density. The energy is $H = \int \mathrm{d}^{D-1} x \mathcal{H}$, and it is *constant* in time under the dynamics: $\frac{\mathrm{d}H}{\mathrm{d}t} = 0$.

T^i_0: energy flux in the i direction.

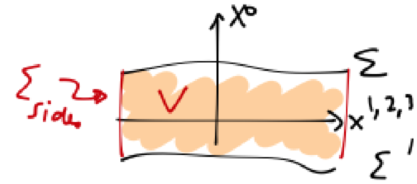
T^0_j: j-momentum density. ($\int_{\text{space}} T^0_j = p_j$)

T^i_j: j-momentum flux in the i direction.

More generally and geometrically, take a *spacelike slice* of spacetime. This means: pick a time coordinate and consider $\Sigma = \{t = t_0\}$, a slice of constant time. Define



$$P_\mu(\Sigma) = \int_\Sigma T^0_\mu \mathrm{d}^3 x \quad .$$



(Later we'll have to worry about the measure on the slice.) A priori this is a function of which Σ we pick, but because of the conservation law, we have

$$\begin{aligned} P_\mu(\Sigma) - P_\mu(\Sigma') &= \int_{\Sigma-\Sigma'} T^0_\mu \mathrm{d}^D x \\ &= \int_{\partial V} T^\alpha_\mu \mathrm{d}S_\alpha \\ &= \int_V \partial_\alpha T^\alpha_\mu \ . \end{aligned} \tag{36}$$

In the penultimate step we assumed that no flux is escaping out the sides: $T^i_\mu = 0$ on Σ_{sides}. In the last step we used Stokes' theorem. (Notice that these steps were not really particular to the energy-momentum currents – we could have applied them to any $\int_\Sigma j^0$.)

So in fact the energy is part of a tensor, like the electromagnetic field strength tensor. This one is not antisymmetric.

Sometimes it is symmetric, and it can be made so. (Notice that symmetric $T^{\mu\nu}$ means that the energy flux is the same as the momentum density, up to signs.) A dirty truth: there is an ambiguity in this construction of the stress tensor. For one thing, notice that classically nobody cares about the normalization of S and hence of T. But worse, we can add terms to S which don't change the EoM and which change T. Given $T^{\mu\nu}$ (raise the index with η),

$$\tilde{T}^{\mu\nu} \equiv T^{\mu\nu} + \partial_\rho \Psi^{\mu\nu\rho}$$

for some antisymmetric object (at least $\Psi^{\mu\nu\rho} = -\Psi^{\rho\nu\mu}$ and $\Psi^{\mu\nu\rho} = -\Psi^{\mu\rho\nu}$) is still conserved (by equality of the mixed partials $\partial_\rho\partial_\nu = +\partial_\nu\partial_\rho$), gives the same conserved charge

$$P^\mu + \int \mathrm{d}^3 x \underbrace{\partial_\rho \Psi^{\mu 0 \rho}}_{=\partial_i \Psi^{\mu 0 i}} = P^\mu + \underbrace{\int_{\partial\Sigma} \mathrm{d}S_i \Psi^{\mu 0 i}}_{=0 \text{ if } \Psi \to 0 \text{ at } \infty} = P^\mu.$$

and it generates the same symmetry transformation.

We will see below (in section 7) that there is a better way to think about $T_{\mu\nu}$ which reproduces the important aspects of the familiar definition[15].

[15]Spoilers: $T_{\mu\nu}(x)$ measures the response of the system upon varying the spacetime metric $g_{\mu\nu}(x)$. That is:

$$T^{\mu\nu}_{\text{matter}}(x) \propto \frac{\delta S_{\text{matter}}}{\delta g_{\mu\nu}(x)}.$$

4.1 Maxwell stress tensor

For now we can use the ambiguity we just encountered to our advantage to make an EM stress tensor which is gauge invariant and symmetric. Without the coupling to charged particles, we have $\mathcal{L}_{EM} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, so

$$\begin{aligned} \tilde{T}^{\mu}_{\nu} &= -\frac{\partial \mathcal{L}}{\partial\left(\partial_{\mu}A_{\rho}\right)}\partial_{\nu}A_{\rho} + \delta^{\mu}_{\nu}\mathcal{L} \\ &= +F^{\mu\rho}\partial_{\nu}A_{\rho} - \frac{1}{4}\delta^{\mu}_{\nu}F_{\rho\sigma}F^{\rho\sigma}. \end{aligned} \tag{37}$$

Notice that this is not gauge invariant:

$$A_{\mu} \rightsquigarrow A_{\mu} + \partial_{\mu}\lambda \implies \tilde{T}^{\mu}_{\nu} \rightsquigarrow \tilde{T}^{\mu}_{\nu} + F^{\mu\rho}\partial_{\nu}\partial_{\rho}\lambda.$$

We can fix this problem by adding another term

$$\partial_{\rho}\Psi^{\mu\nu\rho}_{AS} = -F^{\mu\rho}\partial_{\rho}A^{\nu}$$

(this relation requires the EoM) to create a better T:

$$T^{\mu\nu} = F^{\mu\rho}F^{\nu}_{\rho} - \frac{1}{4}\eta^{\mu\nu}F^{2}.$$

This is clearly gauge invariant since it's made only from F; also it's symmetric in $\mu \leftrightarrow \nu$.

Notice also that it is traceless: $\eta_{\mu\nu}T^{\mu\nu} = T^{\mu}_{\mu} = 0$. This a consequence of the scale invariance of E&M, the fact that photons are massless.

More explicitly for this case

$$\begin{aligned} cT^{00} &= \tfrac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right) && \text{energy density of EM field} \\ c^2T^{0i} &= c\epsilon^{ijk}E_jB_k = c\left(\vec{E}\times\vec{B}\right)^i && \text{Poynting vector} \end{aligned} \tag{38}$$

The latter is both the momentum density and the energy flux, since T is symmetric. If we substitute in a wave solution, we'll see that the energy in the wave is transported in the direction of the wavenumber at the speed of light.

What's the interpretation of T^{ij}?

$$0 = \partial_{\mu}T^{\mu\nu} \overset{\nu=0}{\implies} \frac{1}{c}\partial_t T^{00} + \partial_i T^{0i} = 0$$

The energy inside a box (a region of space) D is

$$E_{\text{box}} = \int_D cT^{00}\mathrm{d}^3x.$$

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\text{box}} = \int_D c\dot{T}^{00}\mathrm{d}^3x = -\int_D c^2\partial_i T^{0i}\mathrm{d}^3x = -c^2\int_{\partial D} T^{0i}\mathrm{d}S_i.$$

So c^2T^{0i} is the energy flux, like we said.

$$0 = \partial_\mu T^{\mu\nu} \overset{\nu=j}{\Longrightarrow} \frac{1}{c}\partial_t T^{i0} + \partial_j T^{ij} = 0$$

$$P^i = \int_D \mathrm{d}^3x T^{i0} \;\text{ is the momentum in the box.}$$

$$\frac{\mathrm{d}P^i}{\mathrm{d}t} = \int_D \mathrm{d}^3x\partial_0 T^{i0} = -\int_D \mathrm{d}^3xc\partial_j T^{ij} = -c\int_{\partial D}\mathrm{d}S_j T^{ij}$$

So cT^{ij} measures the flow of i-momentum in the j direction. Notice that the reason that momentum changes is because of a force imbalance: some stress on the sides of the box – if ∂D were a material box, the sides would be pushed together or apart. Hence, T^{ij} is called the stress tensor. Altogether $T^{\mu\nu}$ is the stress-energy-momentum tensor, but sometimes some of these words get left out.

4.2 Stress tensor for particles

$\partial_\mu T^{\mu\nu}_{EM} = 0$ if F obeys the *vacuum* Maxwell equations.

If we couple our Maxwell field to charged matter as in §3.3 the Maxwell stress tensor is no longer conserved, but rather

$$\partial_\mu T^{\mu\nu}_{EM} = -\frac{1}{c}j^\rho F^\nu_\rho \tag{39}$$

so that $P^\mu_{EM} = \int \mathrm{d}^3x T^{0\mu}_{EM}$ no longer has $\dot{P}^\mu_{EM} = 0$. The EM field can give up energy and momentum to the particles. We can rewrite this equation (39) as

$$\partial_\mu\left(T^{\mu\nu}_{EM} + T^{\mu\nu}_{\text{particle}}\right) = 0$$

that is – we can interpret the nonzero thing on the RHS of (39) as the divergence of the particles' stress tensor. This gives us an expression for $T^{\mu\nu}_{\text{particle}}$!

In more detail: consider *one* charged particle (if many, add the stress tensors) with charge e and mass m. The current is

$$j^\mu(x) = e\int_{-\infty}^{\infty}\mathrm{d}\mathfrak{s}\delta^4(x - x_0(\mathfrak{s}))\frac{\mathrm{d}x_0^\mu}{\mathrm{d}\mathfrak{s}}$$

so

$$\partial_\mu T^{\mu\nu}_{EM} = -\frac{e}{c}\int \mathrm{d}\mathfrak{s}\delta^4(x - x_0(s))\;\underbrace{F^\nu_\rho\frac{\mathrm{d}x_0^\rho}{\mathrm{d}\mathfrak{s}}}_{\text{Lorentz Force}}$$

Using $0 = \frac{\delta S_{\text{particle}}}{\delta x_0} \implies mc^2 \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \left(\dot{x}^\mu / \sqrt{-\dot{x}^2}\right) = eF^\mu_\rho \frac{\mathrm{d}x^\rho}{\mathrm{d}\mathfrak{s}}$ we have

$$F^\nu_\rho \frac{\mathrm{d}x^\rho}{\mathrm{d}\mathfrak{s}} = \frac{mc^2}{e} \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \left(\frac{\dot{x}^\nu}{\sqrt{-\dot{x}_0^2}} \right).$$

$$\begin{aligned}
\implies \partial_\mu T^{\mu\nu}_{EM} &= -mc \int_{-\infty}^{\infty} \mathrm{d}\mathfrak{s} \delta^4(x - x_0(s)) \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \left(\frac{\dot{x}_0^\nu}{\sqrt{-\dot{x}_0^2}} \right) \\
& \qquad \left(\text{Using} \quad \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} \delta^4(x - x_0(s)) = -\frac{\mathrm{d}x_0^\mu}{\mathrm{d}s} \frac{\partial}{\partial x^\mu} \delta^4(x - x_0(\mathfrak{s})) \qquad \right) \\
&= -mc \int \mathrm{d}\mathfrak{s} \frac{\mathrm{d}x_0^\mu}{\mathrm{d}s} \partial_\mu \delta^4(x - x_0(\mathfrak{s})) \frac{\dot{x}_0^\nu}{\sqrt{-\dot{x}_0^2}} \\
&= -\partial_\mu \underbrace{\left(mc \int \mathrm{d}\mathfrak{s} \frac{\dot{x}_0^\mu \dot{x}_0^\nu}{\sqrt{-\dot{x}_0^2}} \delta^4(x - x_0(\mathfrak{s})) \right)}_{T^{\mu\nu}_{\text{particle}}} \qquad (40)
\end{aligned}$$

To recap what we just did, we constructed $T^{\mu\nu}_{\text{particle}}$

$$T^{\mu\nu}_{\text{particle}} = mc \int_{-\infty}^{\infty} \mathrm{d}\mathfrak{s} \delta^4(x - x_0(\mathfrak{s})) \frac{\dot{x}_0^\mu \dot{x}_0^\nu}{\sqrt{-\dot{x}_0^2}}$$

so that

$$\partial_\mu \left(T^{\mu\nu}_{EM} + T^{\mu\nu}_{\text{particle}} \right) = 0$$

on solutions of the EoM $\frac{\delta S_{\text{total}}}{\delta x_0} = 0$ and $\frac{\delta S_{\text{total}}}{\delta A} = 0$, where S_{total} is the action (32).

Notice that the dependence on the charge cancels, so we can use this expression even for neutral particles. (Later we will have a more systematic way of finding this stress tensor.) Also notice that this $T^{\mu\nu}_{\text{particle}} = T^{\nu\mu}_{\text{particle}}$ is symmetric.

4.3 Fluid stress tensor

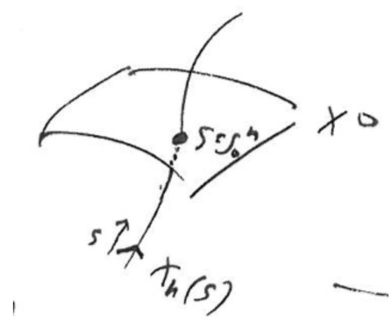
Now consider a collection of particles e_n, m_n with trajectories $x_n^\mu(\mathfrak{s}), n = 0, 1, ...$

$$T^{\mu\nu}_{\text{particles}}(x) = \sum_n m_n c \int \mathrm{d}\mathfrak{s} \delta^4(x - x_n(\mathfrak{s})) \frac{\dot{x}_n^\mu \dot{x}_n^\nu}{\sqrt{-\dot{x}_n^2}}$$

We are going to extract from this a *coarse-grained* description of the stress-energy tensor of these particles, thinking of them as (an approximation to) a continuum fluid. It's a little bit involved.

$$T^{\mu\nu}_{\text{particles}} = \sum_n m_n c \int \mathrm{d}\mathfrak{s} \delta(x^0 - x_n^0(\mathfrak{s})) \delta^3(\vec{x} - \vec{x}_n(\mathfrak{s})) \frac{\dot{x}_n^\mu \dot{x}_n^\nu}{\sqrt{-\dot{x}_n^2}}$$

We want to use the $\delta(x^0)$ to eat the worldline integral:



$$\delta(x^0 - x_n^0(\mathfrak{s})) = \frac{1}{\left|\frac{\mathrm{d}x_n^0(\mathfrak{s})}{\mathrm{d}\mathfrak{s}}\right|} \delta(\mathfrak{s} - \mathfrak{s}_0^n(x^0)) \quad ;$$

here $\mathfrak{s}_0^n(t)$ is the worldline parameter value where particle n is at time $t = x^0$. We will assume that $\frac{\mathrm{d}x_n^0}{\mathrm{d}\mathfrak{s}} > 0$, so the parameters go forwards in time.

$$T^{\mu\nu}_{\text{particles}}(x) = \sum_n m_n c \delta^3(\vec{x} - \vec{x}_n(\mathfrak{s}_0^n)) \frac{\dot{x}_n^\mu \dot{x}_n^\nu}{\dot{x}_n^0 \sqrt{-\dot{x}_n^2}}$$

We would like to write this in terms of the *mass density* function

$$\mu(x) \equiv \sum_n m_n \delta^{(3)} \left(\vec{x} - \vec{x}_n(\mathfrak{s}_0^n)\right) \ .$$

Since T is only nonzero if x is on some trajectory x_n (assume the particles don't collide), we can erase the ns on the xs and write:

$$T^{\mu\nu}_{\text{particles}} = \sum_n m_n \delta^{(3)}(\vec{x} - \vec{x}_n(\mathfrak{s}_0^n)) c \frac{\frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} \frac{\mathrm{d}x^\nu}{\mathrm{d}\mathfrak{s}}}{\dot{x}^0 \sqrt{-\left(\frac{\mathrm{d}x}{\mathrm{d}\mathfrak{s}}\right)^2}}$$

which is

$$\boxed{T^{\mu\nu}_{\text{particles}} = \mu(x) c \frac{\dot{x}^\mu \dot{x}^\nu}{\dot{x}^0 \sqrt{-\left(\dot{x}\right)^2}}}$$

If we parametrize our paths by lab-frame time, this is:

$$T^{\mu\nu}_{\text{particles}} = \mu(x) c \frac{\frac{\mathrm{d}x^\mu}{\mathrm{d}x^0} \frac{\mathrm{d}x^\nu}{\mathrm{d}x^0}}{\sqrt{1 - \left(\frac{\mathrm{d}\vec{x}}{\mathrm{d}x^0}\right)^2}}$$

Next goal: define the fluid 4-velocity field $u^\mu(x)$ (a tricky procedure). 1) Pick a point in spacetime labelled x^μ. 2) Find a particle whose worldline passes through this point, with some arbitrary parametrization. 3) If you can't find one, define $u(x) = 0$. 4) For points x on the worldline of some particle x_n, define:

$$u^\mu(x) = \frac{\dot{x}_n^\mu}{\sqrt{-\dot{x}_n^2}} \quad , \quad \text{where} \quad \dot{x}_n \equiv \frac{\mathrm{d}x_n}{\mathrm{d}\mathfrak{s}} \tag{41}$$

– if there are enough particles, we'll have a $u^\mu(x)$ defined at enough points. This *velocity field* is (a) is a 4-vector at each point and (b) is independent of the parametrization $\mathfrak{s} \to \tilde{\mathfrak{s}}(s)$.

In terms of this velocity field, we have:

$$T^{\mu\nu}_{\text{particles}} = \mu(x) c \frac{u^\mu u^\nu}{u^0}$$

Notice that μ is not a scalar field. Rather

$$\rho \equiv \frac{\mu(x)}{u^0} \quad \text{is a scalar field.}$$

Altogether:

$$\boxed{T^{\mu\nu}_{\text{particles}} = c\rho(x) u^\mu(x) u^\nu(x)} \ . \tag{42}$$

Notice that we can apply the same logic to the current density 4-vector for a collection of particles. The analogous result is:

$$j^\mu = \frac{e}{m} \rho u^\mu \quad .$$

(The annoying factor of $\frac{e}{m}$ replaces the factor of the mass in the mass density with a factor of e appropriate for the charge density. If you need to think about both $T^{\mu\nu}$ and j^μ at the same time it's better to define ρ to be the *number* density.)

So these results go over beautifully to the continuum. If we fix a time coordinate x^0 and use it as our worldline coordinates in (41) then

$$u^0(x) = \frac{1}{\sqrt{1 - \frac{\vec{v}^2(x)}{c^2}}}, \quad \vec{u}(x) = \frac{\vec{v}(x)/c}{\sqrt{1 - \frac{\vec{v}^2(x)}{c^2}}}.$$

If we set the EM forces to zero (say $T^{\mu\nu}_{EM} = 0$) then the particles' stress tensor is conserved $\partial_\mu T^{\mu\nu}_{\text{particles}} = 0$; these equations have names:

$\boxed{\nu = 0}$: $\quad 0 = \partial_t(\rho c) + \vec{\nabla} \cdot (\rho \vec{u})$ — energy conservation

$\boxed{\nu = i}$: $\quad 0 = \partial_t \vec{u} + \left(\vec{u} \cdot \vec{\nabla}\right) \vec{u}$ — Navier-Stokes (w/ no viscosity and no pressure)

[End of Lecture 5]

4.3.1 Pressure

[Weinberg §2.10 (page 47), Landau-Lifshitz$_{+---}$ vol II, §35]

Consider a little cube of our fluid of particles at the point x. The momentum flux through a face of the cube is a force acting on the cube $-T^{ij}\mathrm{d}S_j$ where $\mathrm{d}S_j$ is an area element to a face with normal in the j direction. In the rest frame of this fluid element (where its 4-velocity is $u^\mu = (1, \vec{0})^\mu$), there is no special direction in space and the aforementioned force is

$$T^{ij}(x)\mathrm{d}S_j = p(x)\mathrm{d}S_k\delta^{ki}$$

where p is the *pressure*. [16] In the rest frame of our chunk its momentum is zero, so $T^{i0} = 0$. And the energy density T^{00} is just the rest mass density of the fluid (times c^2): $\rho = \frac{m}{\text{vol}}c^2$. So, we have names for the components of our stress tensor:

$$cT^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}^{\mu\nu} \tag{43}$$

This agrees with our previous expression (42) if we set $u^\mu = (1, \vec{0})^\mu$ *and* $p = 0$. What is the pressure? It's some coarse-grained description of short-ranged interactions between the particles or some other mechanism by which the particles might exert a force on the walls of their container (*e.g.* thermal excitations, Fermi degeneracy pressure...).

A *perfect fluid* more generally is one where we can find its stress tensor for any point in the fluid by simply boosting to the rest frame and using (43). So in the frame where its local velocity vector if $u^\mu(x)$, we get

$$cT^{\mu\nu} = (p + \epsilon)\, u^\mu u^\nu + p\eta^{\mu\nu} \quad . \tag{44}$$

(ϵ is the energy density, which reduces to ρc^2 in the rest frame.) It's Lorentz covariant and has the right form when we set u to the rest frame. Now the conservation law gives a better Navier-Stokes equation:

$$0 = \partial_\mu T^{\mu i} = (p + \epsilon)\, \partial_\mu u^\mu u^i + \partial_\mu p \eta^{\mu i}$$

(that is, $0 = ma - F$).

For an *imperfect fluid*, there are additional terms on the RHS of (44) (the *constitutive relation*) depending on the *gradients* of the velocity field. The terms with the fewest extra derivatives are of the form $T^{\mu\nu}_{\text{viscosities}} \sim \eta\partial^\mu u^\nu$.

[16] The fact that the pressure is the same in all directions for a chunk of fluid at rest is called *Pascal's Law* to intimidate you. It is of course not true for every chunk of stuff: for example a solid at rest still breaks rotation invariance. Let's not worry about that; the gravitational effects of these deviations from rotation invariance are teeny.

In the non-relativistic limit, the momentum density becomes $T^{0i} = (p+\rho)u^i$; quite generally, $(p+\rho)/c^2$ plays the role of the mass density.

The right way to think about this study of fluids (*hydrodynamics*) is as an effective field theory: It is an expansion in small derivatives (of the velocity field u). The 'perfect fluid' assumption is merely the approximation that the leading term in that expansion (no derivatives of u at all) is exact. In that approximation, the expression (44) is the only covariant expression we can write which reduces to the correct expression in the rest frame (43) (which it must, again by the assumption of perfectness).

Notice that the trace of the fluid stress tensor is

$$(cT_{\text{fluid}})^{\mu}_{\mu} = +(p+\epsilon)\underbrace{u^2}_{=-1} + 4p = 3p - \epsilon.$$

Going back to the description in terms of particles this is a sum of terms of the form

$$(cT_{\text{particle}})^{\mu}_{\mu} = mc^2 \int d\mathfrak{s}\,\delta^4(x - x_0(\mathfrak{s}))\frac{\dot{x}^2}{\dot{x}^0\sqrt{-\dot{x}^2}} \overset{\mathfrak{s}=t}{=} -mc^2\sqrt{1-\vec{v}^2/c^2}\,\delta^3(\vec{x} - \vec{x}_0(t)) \leq 0$$

$= 0$ only if $m = 0$. So we learn the equation of state of relativistic particles is $\epsilon = 3p$:

$$T^{\mu\nu}_{\text{radiation}} = \frac{\epsilon}{3c}\left(4u^\mu u^\nu + \eta^{\mu\nu}\right).$$

(In cosmology such a stress tensor is called 'radiation dominated'.) And for massive particles $\epsilon > 3p$. In the NR limit, $\epsilon \simeq \mu_0 c^2 \gg 3p$ and

$$T^{\mu\nu}_{\text{matter}} \simeq \mu_0 c u^\mu u^\nu.$$

5 Differential Geometry Bootcamp

Q: OK, so: why are we talking about manifolds and all this mathy stuff?
A: Because the Einstein equivalence principle says exactly :
spacetime is a semi-Riemannian manifold. (Vaguely, it says "spacetime looks like Minkowski spacetime, locally". Our job now is to make this precise.)

So bear with me.

We are going to make *intrinsic* constructions of manifolds. The surface of the Earth is an example of the kind of space we want to think about. It is clearly embedded in a third dimension, and we can use this extra (*extrinsic*) structure to describe it. But in the case of spacetime itself we don't have the luxury of some embedding space (as far as we know) and so we must learn to do without it.

5.1 How to defend yourself from mathematicians

Here is some self-defense training. There are going to be a lot of definitions. But it will help to provide some context – to situate our discussion within the world of mathematics. And some of them are actually useful.

A pertinent fact about math that I took a long time to learn and keep forgetting:

Principle of mathematical sophistication: the more horrible-sounding adjectives there are in front of the name of a mathematical object, the *simpler* it is.

Here, as in math in general, we will start with something that sounds simple but is actually totally horrible. Then we'll add assumptions that seem like they complicate the story, but actually make things much easier, because they rule out the pathological cases that we don't care about. And best of all, these assumptions come from physics.

Each of the entries in the following list is called a *category*. In moving from one entry to the next in the list, we assume all of the previous properties and more – that is, we are *specifying*, and ruling out pathological behaviors.

1. Topological spaces. (Can't do anything.) it's just a list (a collection of points) with some notion of 'open set'. Horrible things can happen here. The German word for 'not horrible' is *Hausdorff*.
2. Topological manifolds. (Can do topology.) locally looks like $\mathbb{R}^n$. patches glued together with continuous functions. Continuous means that they map open sets to open

sets. But they don't have to be differentiable!

3. Differentiable manifolds, aka smooth spaces. (Can do calculus.) Now we glue patches together with differentiable functions, aka diffeomorphisms.
4. Riemannian manifolds. (Can measure lengths.) All of the above and here's a *metric*.
5. Semi-Riemannian manifolds. (Can measure lengths, but they might be negative.) (More precisely, the metric has signature $(-+++)$, with one timelike coordinate.)

[Wald, Chapter 1, Appendix A] Let's go through that list again in a little bit more detail: A *topological space* X is a set with a *topology* $\mathcal{D}$, a family of open sets $\{\mathcal{U}_\alpha \subset X\}_{\alpha \in A}$ which satisfies

1. $\mathcal{D} \ni X, \emptyset$ (The whole space X is open and so is the empty set.)
2. The intersection of a finite number of open sets is also an open set, *i.e.* if $\mathcal{D} \ni \mathcal{U}_1, \mathcal{U}_2$, then $\mathcal{D} \ni \mathcal{U}_1 \cap \mathcal{U}_2$.
3. An arbitrary union of open sets is an open set, *i.e.* if $\forall a \mathcal{U}_a \in \mathcal{D}$, then $\cup_a \mathcal{U}_a \in \mathcal{D}$ (maybe an infinite number of them).

Def: given $x \in X, \mathcal{U} \in \mathcal{D}$ is a *neighborhood* of x if $x \in \mathcal{U}$.

Here's a decent example: $X = \mathbb{R}$ with $\mathcal{D} = \{(a,b), a < b \in \mathbb{R}\}$, a collection of open intervals.

Here's a pathological example: take any X and let $\mathcal{D} = \{X, \emptyset\}$.

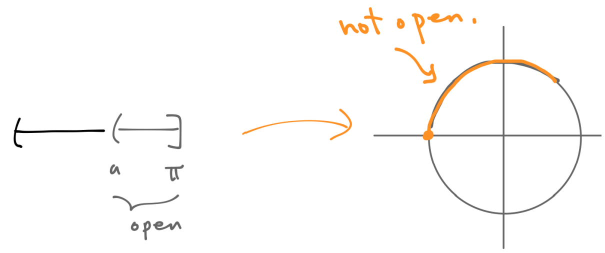
Hausdorff criterion: $\forall x, y \in X, x \neq y$, $\exists \mathcal{U} \in \mathcal{D}$ such that $(s.t.)$ $x \in \mathcal{U}$
and $\exists V \in \mathcal{D}, s.t.\ y \in \mathcal{V}$
and $\mathcal{U} \cap \mathcal{V} = \emptyset$. This is a necessary condition for the topology to be useful – you can separate points – and it excludes the second example above.
(Exercise: is the Hausdorff condition equivalent to $\exists \mathcal{U} \ni x$ s.t. y is not in $\mathcal{U}$?)

Q: when should we consider two topological spaces to be the same?
A: Two topological spaces $X, \tilde{X}$ are equivalent if $\exists f : X \to \tilde{X}$ which maps open sets to open sets. More precisely:
Def: A *homeomorphism* (isomorphism in the category of topological spaces) $(X, \mathcal{D}) \simeq (\tilde{X}, \tilde{\mathcal{D}})$ is a map $f : X \to \tilde{X}$ which is:
one-to-one $(x \neq y \implies f(x) \neq f(y))$
and onto $(\forall \tilde{x} \in \tilde{X}, \exists x \text{ s.t. } f(x) = \tilde{x},\ i.e.\ f(X) = f(\tilde{X}))$
and is *bicontinuous* (*i.e.* satisfies $\forall \mathcal{U} \in \mathcal{D}, f(\mathcal{U}) \in \tilde{\mathcal{D}}$ and $\forall \mathcal{V} \in \tilde{\mathcal{D}}, \exists \mathcal{U} s.t.\ \mathcal{V} = f(\mathcal{U})$) .

This latter condition $(\forall \mathcal{V} \in \tilde{\mathcal{D}}, \exists \mathcal{U} s.t. \mathcal{V} = f(\mathcal{U}))$ says f is *continuous*.

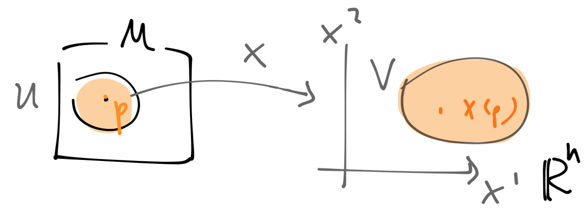
Here's an example of a map that's 1-1 and onto and continuous, but not *bicontinuous*, that is, f^{-1} is not continuous:

$$\begin{array}{rcl} f : I \equiv (-\pi, \pi] & \to & S^1 \\ x & \mapsto & e^{\mathbf{i}x} \end{array}$$



where I has the "intersection topology": open sets in I are the intersections of open sets in $\mathbb{R}$ with I. The set $(a, \pi]$ is open in this topology, and it maps to a set which is not open on S^1.

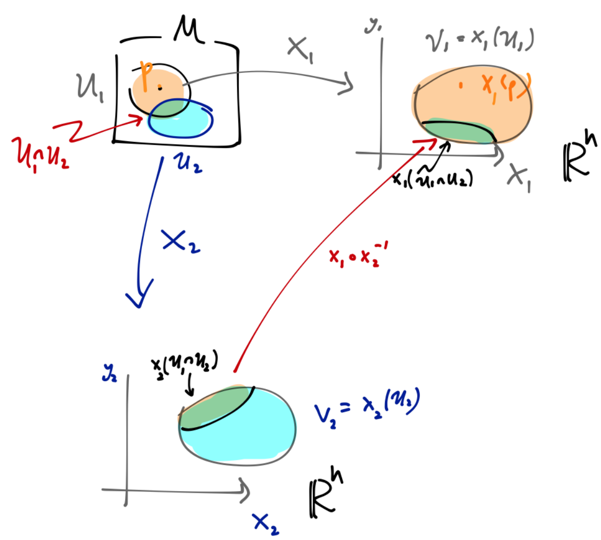
Next step: A *topological manifold* M is what we get if we introduce coordinates on a topological space – M is a topological space which is Hausdorff and includes the following setup (local coordinate systems): $\forall p \in M, \exists \mathcal{U} \in \mathcal{D}$ containing p s.t. $\mathcal{U}$ is homeomorphic to an open set $\mathcal{V} \in \mathbb{R}^n$ – that is,



$$\begin{array}{rcl} x : \mathcal{U} & \to & V \\ p & \mapsto & x(p) = (x_1(p), ...x_n(p)) \in \mathcal{V} \subset \mathbb{R}^n \end{array} \tag{45}$$

is 1-1, onto, bicontinuous, and is called a *local coordinate.*

Here's an efficient way to do this: Pick a subcover $\{\mathcal{U}_\alpha \in \mathcal{D}\}$ s.t. $\cup_\alpha \mathcal{U}_\alpha = M$. The collection $\{\underbrace{(\mathcal{U}_\alpha, x_\alpha)}_{\text{chart}}\}$ is called an *atlas*. This is a good name: it's just like a book where each page is a map of some region of the world (the $\mathbb{R}^n$ in question is the $\mathbb{R}^2$ of which each page is a subset). You also need to know that there's overlaps between the charts and what is the relation between the coordinates, *i.e.* *transition functions*: given an atlas of M suppose $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$: The definition of topological manifold requires that $x_i : \mathcal{U}_i \to \mathcal{V}_i, i = 1, 2$ are homeomorphisms and $\exists x_2^{-1} : \mathcal{V}_2 \to \mathcal{U}_2$, continuous, so that we can construct



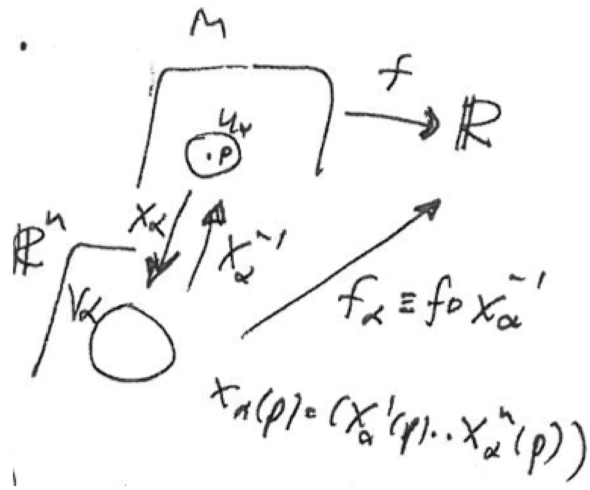
$$x_1 \circ x_2^{-1} : x_2(\underbrace{\mathcal{U}_1 \cap \mathcal{U}_2}_{\subset \mathcal{V}_2 \subset \mathbb{R}^{n'}}) \to \underbrace{x_1(\mathcal{U}_1 \cap \mathcal{U}_2)}_{\subset \mathcal{V}_1 \subset \mathbb{R}^n}$$

which is a homeomorphism. This implies $n' = n$ – connected parts of M have a well-defined dimension.

Proposition: A map $f : M \to \mathbb{R}$ is continuous iff $\forall(\mathcal{U}_\alpha, x_\alpha)$ in an atlas of M

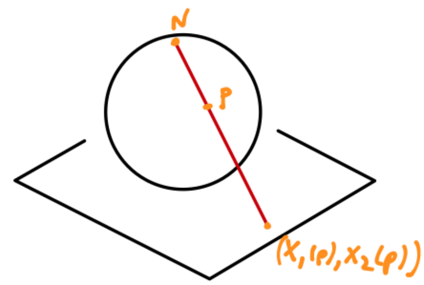
$$f_\alpha \equiv f \circ x_\alpha^{-1} : \underbrace{\mathcal{V}_\alpha}_{\equiv x_\alpha(\mathcal{U}_\alpha)} \to \mathbb{R}$$

is continuous. This f_α is an ordinary function of n real variables.



Next category. Define $\mathcal{C}^r(\mathcal{U})$ to be the space of r-times differentiable functions from $\mathcal{U}$ to $\mathbb{R}$. A $\mathcal{C}^r$ *differentiable manifold* is a topological manifold where the transition functions $x_\alpha \circ x_\beta^{-1}$ are diffeomorphisms (r-times continuously differentiable).

A function $f : M \to \mathbb{R}$ is $\mathcal{C}^r$ iff $f \circ x_\alpha^{-1}$ is $\mathcal{C}^r$.



For example:

- $\mathbb{R}^n$ with $\{\mathbb{R}^n, \text{identity map}\}$.
- S^2 with stereographic projections (see the figure at right and the next hw)
  $x_N : S^2 - \{\text{north pole}\} \to \mathbb{R}^2$
  $x_S : S^2 - \{\text{south pole}\} \to \mathbb{R}^2$.
  (Exercise: write the map explicitly in terms of the embedding in $\mathbb{R}^3$ ($\{(x_1, x_2, x_3) \in \mathbb{R}^3 | \sum_i x_i^2 = 1\} \to \mathbb{R}^2$). Show that the transition function $x_S \circ x_N^{-1} : \mathbb{R}^2_N \to \mathbb{R}^2_S$ is differentiable on the overlap (everything but the poles).)

Q: When are two $\mathcal{C}^r$ differentiable manifolds $M, \tilde{M}$ the same?

Def: A $\mathcal{C}^r$-differentiable map (*diffeomorphism*) $f : M^m \to \tilde{M}^{\tilde{m}}$ is a continuous map s.t. $\forall(x_\alpha, \mathcal{U}_\alpha) \in \mathcal{D}_M$ and $\forall(y_i, \mathcal{V}_i) \in \mathcal{D}_{\tilde{M}}$,

$$y_i \circ f \circ x_\alpha^{-1} : \mathcal{U}_\alpha \to \mathcal{V}_i$$

is $\mathcal{C}^r$-differentiable.

A: When $\exists f : M \to \tilde{M}$ which is a diffeomorphism – (note: they are in particular equivalent as topological spaces.)

$\exists$ Weird examples:
(1) (Friedman, Donaldson): $\{(\mathbb{R}^4, \text{identity})\}$ is a differentiable manifold. But there is another differentiable structure on $\mathbb{R}^4$ which is NOT diffeomorphic to the usual one.

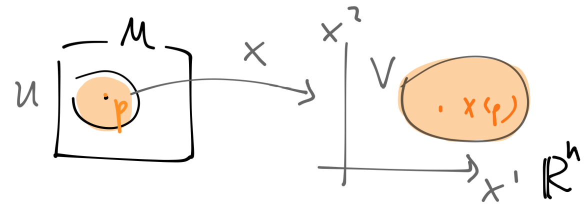
(2) (Milnor) S^7 with (the appropriate generalization) of stereographic projection coordinates is not the only atlas on the topological manifold S^7!

So now we are going to start caring about how objects transform under general changes of coordinates, rather than just rotations and Lorentz transformations. Notice that we did not get all the way through the second pass on our list of categories: we didn't yet introduce the metric structure. There are some other things we can and must do first.

[End of Lecture 6]

5.2 Tangent spaces

This subsection applies already to differentiable manifolds – we're not going to use the metric yet. Let's use the local maps to patches of $\mathbb{R}^n$ to define tangent vectors. (We could define them by embedding our manifold in $\mathbb{R}^{N>n}$; this is ugly and depends on the choice of embedding. The following construction is intrinsic and therefore good.)



Def: A tangent vector is something that eats functions and gives numbers and is a derivation. That is: V is a tangent vector to M at $p \in \mathcal{U}$ if $\forall f : \mathcal{U} \to \mathbb{R}$ (we'll assume f is $\mathcal{C}^\infty$ – we can take as many derivatives as we want), $V : f \mapsto V(f) \in \mathbb{R}$ satisfying ('is a derivation')

1. linear: $V(af + bg) = aV(f) + bV(g)$.
2. Liebniz: $V(fg) = V(f)g + fV(g)$. That is: the product rule.

More succinctly, then: a tangent vector to M at p is a map (to $\mathbb{R}$) on functions in a nbhd of p which is a derivation.

Given a set of coordinates, we find some examples of vector fields: define $(\partial_\mu)_p : f \mapsto \frac{\partial}{\partial x^\mu} f$. This clearly satisfies 1, 2.

We will use the name T_pM for the vector space of tangent vectors to M at p.
Fact: $\{(\partial_\mu)_p\}$ span T_pM.
Why: Taylor expand

$$f(x_1, ...x_n) = f(x(p)) + \frac{\partial f}{\partial x^\mu}(x(p))(x^\mu - x^\mu(p)) + \frac{1}{2}\partial_\mu\partial_\nu f(x(p)(x^\mu - x^\mu(p))(x^\nu - x^\nu(p)) + ...$$

$$(1) + (2) \implies V(1) = 0 \text{ since } V(1) = V(1 \cdot 1) = 2V(1) \implies V(a) = 0 \forall a \in \mathbb{R}.$$

$$\implies V(f) = V(\underbrace{f(x(p))}_{\text{const} \implies 0} + \underbrace{V(\frac{\partial f}{\partial x^\mu}(x(p))(x^\mu - x^\mu(p)))}_{=\frac{\partial f}{\partial x^\mu}(x(p))V(x^\mu - x^\mu(p))} + ...$$

$$V(f) = 0 + \partial_\mu f_p V(x^\mu) + \frac{1}{2}\partial_\mu\partial_\nu f_p \underbrace{V\left((x^\mu - x^\mu(p))(x^\nu - x^\nu(p))\right)}_{=V(x^\mu)(x^\nu - x^\nu(p)) + (x^\mu - x^\mu(p))V(x^\nu)|_p = 0}$$

So:

$$V(f) = \frac{\partial f}{\partial x^\mu}(p)V^\mu \ .$$

Summary: the tangent vector space T_pM is

$$T_pM \ni V : f \mapsto V(f) \in \mathbb{R} \quad \text{with} \quad V(f) = V^\mu \frac{\partial}{\partial x^\mu} f(p)$$

Perhaps you are not used to thinking of vectors as such a map – as a differential operator. This operator is just the directional derivative (in the direction of the thing you usually think of as the vector). It contains the same information.

How does changing coordinates act on such a vector? Simple: it doesn't. That is, $\forall f \in \mathcal{C}^\infty(\mathcal{U})$, the value of $V(f)$ shouldn't change!

But the *components* of the vector in the coordinate basis V^μ certainly do change: under $x^\mu \to \tilde{x}^\mu(x)$,

$$V(f) = V^\mu \left(\frac{\partial f}{\partial x^\mu}\right)_p \stackrel{!}{=} \tilde{V}^\mu \left(\frac{\partial f}{\partial \tilde{x}^\mu}\right)_p = \tilde{V}^\mu \left(\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\right)_p \left(\frac{\partial f}{\partial x^\nu}\right)_p$$

$$\implies \boxed{\tilde{V}^\mu \left(\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\right)_p = V^\nu}$$

To go in the other direction: because

$$\left(\frac{\partial x^\nu}{\partial \tilde{x}^\mu}\right)_p \left(\frac{\partial \tilde{x}^\mu}{\partial x^\rho}\right)_p = \delta^\nu_\rho$$

(or by using the symmetry between tilded and untilded vars) the inverse transformation is

$$V^\mu \left(\frac{\partial \tilde{x}^\nu}{\partial x^\mu}\right)_p = \tilde{V}^\nu$$

To emphasize: as a differential operator, V is coordinate independent:

$$V = V^\mu \frac{\partial}{\partial x^\mu}|_p = \tilde{V}^\mu \frac{\partial}{\partial \tilde{x}^\mu}|_p \ .$$

Fundamental principle of relabelling: Physics shouldn't depend on what coordinates we use to describe it.

Or as Zee says: physics should not depend on the physicist [Zee page 334].

Compare with the previous mention that was made of vectors: tangent vector to a curve, $\mathfrak{s} \mapsto x^\mu(\mathfrak{s})$ we defined to be $V^\mu \equiv \frac{dx^\mu}{d\mathfrak{s}}$. Meaning: these are the components of the tangent vector to the curve in the coordinate basis:

$$V = \frac{dx^\mu}{d\mathfrak{s}} \frac{\partial}{\partial x^\mu}$$

Notice that if we change coordinates on spacetime $x^\mu \to \tilde{x}^\mu$ this object manifestly stays the same.

In fancier language: a differentiable curve in M is a map

$$\varphi : \underbrace{(a, b)}_{\subset \mathbb{R}} \to M$$

with the following properties. If we pick $p \in \mathcal{U} \subset M$ ($\mathcal{U}$ an open set with coordinate x^μ) in the image of φ, and let $\varphi^\mu = x^\mu(\varphi)$ – that is

$$\varphi^\mu(\mathfrak{s} \in (a, b)) = x^\mu(\underbrace{\varphi(\mathfrak{s})}_{\in M}) ,$$

then if we are given a function $f \in \mathcal{C}^\infty(\mathcal{U})$ we can restrict f to $\varphi(a, b)$ and construct the tangent vector

$$\underbrace{\frac{d}{d\mathfrak{s}} f(\phi(t))|_{\mathfrak{s}_0}}_{\in \mathbb{R}} \overset{\text{chain rule}}{=} \frac{\partial f}{\partial x^\mu} \left(\frac{d\varphi^\mu}{d\mathfrak{s}} \right) (t_0)$$

so $V = \frac{d\phi^\mu}{d\mathfrak{s}} \frac{\partial}{\partial x^\mu}$, $\quad i.e.\ V^\mu = \frac{d\phi^\mu}{d\mathfrak{s}}$.

Def of tangent vector *field*: this is just a tangent vector at every point. The name for the collection of tangent spaces at each point is the tangent bundle TM.

$$\begin{aligned} V : M &\to TM \\ p &\mapsto V_p \in T_p M \end{aligned} \tag{46}$$

An example of a vector field on the sphere is obtained by embedding it in $\mathbb{R}^3$ and rotating it: define v_p to be the velocity of the sphere at each point (it vanishes at the poles).

Def of *commutator* (aka *Lie bracket*): consider $u = u^\mu \partial_\mu$. Define a product of vector fields by thinking of them as differential operators and acting twice on the same function: $uvf = u^\mu \partial_\mu (v^\nu \partial_\nu f)$.

$$[v, u] = vu - uv = [v^\mu \partial_\mu, u^\nu \partial_\nu] = v^\mu \partial_\mu (u^\nu \partial_\nu) - u^\nu \partial_\nu (v^\mu \partial_\mu) = (v^\mu (\partial_\mu u^\nu) - u^\mu (\partial_\mu v^\nu)) \partial_\nu$$

also a (contravariant) vector field. So the set of vector fields on M with $[,]$ form a closed algebra. In fact it is a Lie algebra (i.e. the bracket satisfies the Jacobi identity).

A vector field on M produces a *flow*, that is, a one-parameter family of maps $\phi_v(s) : M \to M$ with $\phi(s) \circ \phi(t) = \phi(s+t)$ and each $\phi(s)$ is smooth and 1-1. The map is arranged so that the vector field is tangent to the flow line going through each point, that is, the flow lines are *integral curves* of the vector field. Let x^μ label a point p. Then $\phi_v(s)(p)$ has coordinates $x^\mu(s)$, where

$$\frac{\mathrm{d}}{\mathrm{d}s} x^\mu(s) = v^\mu(x) \tag{47}$$

is an ordinary differential equation for $x^\mu(s)$; the initial condition that $\phi(s=0) = \mathbb{1}$, the identity map, translates to $x^\mu(0) = x^\mu$. Notice that flows and vector fields contain the same information; if we are given $\phi(s)$, we can construct the associated v using (47) in the other direction.

A reason to care about the commutators of vector fields is if you want to compare the results of their flows; flowing by commuting vector fields you end up at the same point no matter which order you do the flows in. (A more precise statement is called the Frobenius Theorem, which I will not discuss.)

5.2.1 Cotangent vectors

equals one-forms. This is something that eats a vector and gives a number, and is linear. That is, the space of cotangent vectors $T_p^\star M$ is the *dual space* (in the sense of linear algebra) to T_pM: the vector space of *linear maps* from T_pM to $\mathbb{R}$.

$$\begin{aligned} \omega : T_pM &\to \mathbb{R} \\ V &\mapsto \omega(V) \end{aligned} \qquad \text{and} \quad \omega(aV + bW) = a\omega(V) + b\omega(W) \quad (\forall a, b \in \mathbb{R})$$

$$\dim T_p^\star M = \dim T_pM$$

since an element is determined by its action on a basis of T_pM.

In a basis, if $V = V^\mu \partial_\mu$, then

$$\omega(V) = V^\mu \omega(\partial_\mu) \equiv V^\mu \omega_\mu.$$

Linear eating is just contraction of the indices.

Example of a one-form: Given a function on M, we can make a one-form as follows. Recall that $v \in T_pM$ maps $f \mapsto v(f)|_p = v^\mu \left(\partial_\mu f\right)_p \in \mathbb{R}$ (the directional derivative). f then defines a cotangent vector by $v \mapsto v(f)$. This is a linear map: $av + bw \mapsto av(f) + bw(f) = (av^\mu + bw^\mu)\, \partial_\mu f|_p$. Denote this cotangent vector $\mathrm{d}f|_p \in T_p^\star M$.

$$\mathrm{d}f|_p \in T_p^\star M \text{ is defined by } \; \mathrm{d}f|_p : M \to \mathbb{R}$$

$$v \mapsto \mathrm{d}f|_p(v) \equiv v(f). \tag{48}$$

Coordinates induce a basis for $T_p^\star M$, just as they do for T_pM. If

$$\begin{aligned} x:\mathcal{U} &\to \mathbb{R}^n \\ p &\mapsto (x^1(p)...x^n(p)) \end{aligned} \tag{49}$$

then each x^μ is a function on M in a neighborhood of p. The associated cotangent vectors are

$$\begin{aligned} \mathrm{d}x^\mu|_p : T_pM &\to \mathbb{R} \\ v &\mapsto \mathrm{d}x^\mu(v) \equiv v(x^\mu) = v^\nu \partial_\nu x^\mu = v^\mu \end{aligned} \tag{50}$$

$$\{(\mathrm{d}x^\mu)_p, \mu = 1..n\}$$

is a basis of $T_p^\star M$ since $\omega = \omega_\mu \mathrm{d}x^\mu|_p$ is an arbitrary 1-form: for any $v \in T_pM$,

$$\omega(v) = \omega_\mu \mathrm{d}x^\mu|_p (v^\nu \partial_\nu) \overset{\text{linearity}}{=} \omega_\mu v^\nu \underbrace{\mathrm{d}x^\mu|_p (\partial_\nu)}_{=\partial_\nu x^\mu = \delta^\mu_\nu} = \omega_\mu v^\mu$$

and any linear functional on T_pM has this form.

And under a coordinate transformation $x^\mu \to \tilde{x}^\mu$,

$$\begin{aligned} \partial_\mu \to \tilde{\partial}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu}\partial_\nu \quad &\implies v = v^\mu \partial_\mu = \tilde{v}^\mu \tilde{\partial}_\mu \quad &\implies \tilde{v}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu \\ \mathrm{d}x^\mu \to \mathrm{d}\tilde{x}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu}\mathrm{d}x^\nu \quad &\implies \omega = \omega_\mu \mathrm{d}x^\mu = \tilde{\omega}_\mu \mathrm{d}\tilde{x}^\mu \quad &\implies \tilde{\omega}_\mu = \frac{\partial x^\nu}{\partial \tilde{x}^\mu}\omega_\nu \end{aligned} \tag{51}$$

$$\text{contravariant:} \quad v \to \tilde{v} \;\; \text{with} \;\; \tilde{v}^\mu(\tilde{x}) = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu(x)$$

$$\text{covariant:} \quad \omega \to \tilde{\omega} \;\; \text{with} \;\; \tilde{\omega}_\mu(\tilde{x}) = \frac{\partial x^\nu}{\partial \tilde{x}^\mu} v_\nu(x)$$

worth noticing:

$$\mathrm{d}x^\mu(v) \equiv v(x^\mu) = v^\nu \partial_\nu x^\mu = v^\mu \implies \tilde{v}^\nu \tilde{\partial}_\nu \tilde{x}^\mu \equiv v(\tilde{x}^\mu) = \mathrm{d}\tilde{x}^\mu(v) = \tilde{v}^\mu.$$

5.2.2 Tensor fields.

A rank $\begin{pmatrix} s \\ r \end{pmatrix}$ tensor is

$$t \in \underbrace{T_p^\star M \otimes .... \otimes T_p^\star M}_{r} \otimes \underbrace{T_pM \otimes .... \otimes T_pM}_{s}$$

In the coordinate bases it is

$$t = t^{\nu_1 \dots \nu_s}_{\mu_1 \dots \mu_r}(x) \mathrm{d}x^{\mu_1} \otimes \dots \otimes \mathrm{d}x^{\mu_r} \otimes \frac{\partial}{\partial x^{\nu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\nu_s}}$$

– it is a thing which eats r vectors and s cotangent vectors at once and gives a number (and it is linear in each feed-hole). (The fancy symbol $\otimes$ is just meant to indicate this linearity in each argument, and will often be omitted.) An example of a rank $\binom{0}{2}$ tensor is a metric – it eats two vectors at a time.

Note: so far "d" is not an operator in that we haven't explained how it acts on anything other than functions, *i.e.*

$$\begin{array}{rcl} \mathrm{d} : \mathcal{C}^\infty(M) & \to & T^\star M \\ f & \mapsto & \mathrm{d}f \end{array} \qquad (52)$$

(and we also haven't explained how to use it as something to integrate) It's not hard to define the exterior derivative on one-forms:

$$\begin{array}{rrcl} \mathrm{d} : & T^\star_p M & \to & \left(T^\star_p M \otimes T^\star_p M\right)_{AS} \\ & \omega = \omega_\mu \mathrm{d}x^\mu & \mapsto & \left(\partial_\mu \omega_\nu - \partial_\nu \omega_\mu\right) \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \end{array} \qquad (53)$$

'AS' stands for antisymmetric. Notice that $\mathrm{d}^2 = 0$ by equality of mixed partials. Notice that the components $(\mathrm{d}\omega)_{\mu\nu} = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu$ are antisymmetric; this allows us to write

$$\mathrm{d}\omega \equiv \left(\partial_\mu \omega_\nu - \partial_\nu \omega_\mu\right) \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

– the wedge product means that the basis one-forms are antisymmetrized. A differential form of rank r can then be expanded as:

$$\omega = \omega_{\mu_1 \cdots \mu_r} \mathrm{d}x^{\mu_1} \wedge \dots \wedge \mathrm{d}x^{\mu_r} \ .$$

Demanding that the $\mathrm{d}^2 = 0$ property continue, we can define the exterior derivative on p-forms of arbitrary rank.

The Lie derivative on vector fields can be defined in a similar way, so far without metric. But curvatures and connections require a metric.

5.2.3 ENTER THE METRIC.

First, let's consider metric on the tangent space $T_p M$ – it's just an inner product between tangent vectors.

$$\begin{array}{rcl} g(\cdot, \cdot) : T_p M \times T_p M & \to & \mathbb{R} \\ (u, v) & \mapsto & g(u, v) \ . \end{array} \qquad (54)$$

It is *bilinear* ($g(au + b\tilde{u}, v) = ag(u, v) + bg(\tilde{u}, v), \forall a, b \in \mathbb{R}$. same for $v \to av + b\tilde{v}$) and *symmetric* ($g(u, v) = g(v, u)$)

In terms of components (by linearity), this is

$$g(u, v) = u^\mu v^\nu g((\partial_\mu)_p, (\partial_\nu)_p) \equiv u^\mu v^\nu g_{\mu\nu}|_p.$$

So symmetry says $g_{\mu\nu}|_p = g_{\nu\mu}|_p$.

Notice that $g|_p$ is a thing that eats two vectors in T_pM and gives a number (and is linear in each argument) – this is exactly the definition of an element of $T_p^\star M \otimes T_p^\star M$ – a rank two covariant tensor. We can expand such a thing in a coordinate basis for each factor:

$$g = g_{\mu\nu} \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \ .$$

We have already seen such a thing and called it $\mathrm{d}s^2$, the line element, omitting the $\otimes$ symbol.

Now demand that $g(u, v)$ is independent of coordinates. This tells us (I hope you are getting used to this stratagem) how $g_{\mu\nu}$ transforms:

$$g(u, v) = u^\mu v^\nu g_{\mu\nu}|_p \stackrel{!}{=} \tilde{u}^\mu \tilde{v}^\nu \tilde{g}_{\mu\nu}|_p$$

This says that g is a rank-2 covariant tensor[17]:

$$g_{\mu\nu}|_p = \left(\frac{\partial \tilde{x}^\rho}{\partial x^\mu}\right)_p \left(\frac{\partial \tilde{x}^\sigma}{\partial x^\nu}\right)_p \tilde{g}_{\rho\sigma}(p).$$

Alternatively we could arrive at this conclusion using linearity:

$$\tilde{g}_{\mu\nu} = g(\tilde{\partial}_\mu|_p, \tilde{\partial}_\nu|_p) = g\left(\frac{\partial x^\rho}{\partial \tilde{x}^\mu}\partial_\rho, \frac{\partial x^\sigma}{\partial \tilde{x}^\nu}\partial_\sigma\right) \stackrel{\text{linearity}}{=} \frac{\partial x^\rho}{\partial \tilde{x}^\mu}\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\mu\nu}$$

Since $g_{\mu\nu}|_p$ is a symmetric matrix, it has real eigenvalues. Their signs determine the *signature* of the metric. Three plus and one minus will be the relevant case for us.

Important remarks:

1. To get a metric on TM from the ones on T_pM for each point p, just erase the p subscripts.

[17] Notice by the way that I am abusing the terminology a bit, using the term 'tensor' for both the coordinate-invariant fancy thing g and for the collection of its components in a coordinate basis $g_{\mu\nu}$, which actually do change under coordinate transformations. If you find this confusing then I suggest the following exercise: make up different names for the two things and go through the lecture notes and correct every appearance of the word tensor. If after doing that you think it's worth making this distinction, come try to convince me.

2. In general, the components of the metric depend on coordinates (in a smooth way, we will assume).

3. Notice that we haven't discussed where to find $g_{\mu\nu}(x)$; until our discussion of Einstein's equations, we will just assume that we've found one in our basement or something.

4. Any expression made of tensors with all indices contracted is *invariant* under general coordinate transformations. This is the point of going through all this trouble.

   Proof: use explicit expressions for transformation laws.

Why $g_{\mu\nu}$ – the metric on spacetime – is symmetric: the EEP. At any point in spacetime x, you have to be able to choose coordinates so that locally it looks like Minkowski spacetime, *i.e.* $g_{\mu\nu}(x) = \eta_{\mu\nu}$, which is symmetric. Under a coordinate change to any general coordinate system,

$$g_{\mu\nu} \mapsto \tilde{g}_{\mu\nu} = \frac{\partial x^\rho}{\partial \tilde{x}^\mu}\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}$$

which is still symmetric. [End of Lecture 7]

Here is a list of uses of the metric: basically we can put it everywhere we had an $\eta_{\mu\nu}$ when we were working in Minkowski space. There are a few places where the η was hidden, *e.g.* because $|\det \eta| = 1$.

1. lengths of vectors: $\| v \|^2 = g_{\mu\nu} v^\mu v^\nu$. The sign of this quantity still distinguishes timelike from spacelike vectors.

2. angles between vectors at the same point: $\frac{g_{\mu\nu} V^\mu W^\nu}{\| V \| \| W \|}$.

3. lengths of curves: add up the lengths of the tangent vectors $L[x] = \int \mathrm{d}\mathfrak{s} \sqrt{|g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}} \frac{\mathrm{d}x^\nu}{\mathrm{d}\mathfrak{s}}|}$. Hence, we can make actions for particles in curved spacetime.

4. inverse metric: if the metric at p is non-degenerate ($\det_{\mu\nu} g_{\mu\nu} \neq 0$) then its inverse as a matrix exists:

   $$g_{\mu\nu} \left(g^{-1}\right)^{\nu\rho} = \delta^\rho_\mu$$

   it is convenient to write this statement as

   $$g_{\mu\nu} g^{\nu\rho} = \delta^\rho_\mu$$

   – that is, we denote the matrix inverse with upper indices and don't need to write the inverse symbol. Notice that this worked for free with $\eta_{\mu\nu}$ and δ_{ij} since as matrices they were their own inverse. Notice that the inverse metric is a tensor field with two contravariant indices.

5. A non-degenerate metric allows us to identify T_pM and $T^\star_pM$

   $$v = v^\mu \partial_\mu \mapsto v^\star \equiv g_{\mu\nu} v^\nu \mathrm{d}x^\mu$$

The map from T_pM to $T_p^\star M$ is

$$v \mapsto \begin{pmatrix} v^\star : & T_pM & \to & \mathbb{R} \\ & u & \mapsto & v^\star(v) \equiv g(v,u) \end{pmatrix}$$

So $v^\star(\cdot) = g(v, \cdot)$. This is what we mean by lowering indices. That is, we don't need to write the stars if we write the indices: given a vector v^μ, we can define $v^\star_\mu \equiv v_\mu$. Similarly, raising indices is just a convenient notation for the map $g^{-1} : T_p^\star M \to T_pM$. And this dropping-stars notation justifies our notation $(g^{-1})^{\mu\nu} \equiv g^{\mu\nu}$ because

$$g^{\mu\nu} = (g^{\star\star})^{\mu\nu} = \left(g^{-1}\right)^{\mu\rho} \left(g^{-1}\right)^{\nu\sigma} g_{\rho\sigma} = \left(g^{-1}\right)^{\mu\rho} \delta^\nu_\sigma = \left(g^{-1}\right)^{\mu\nu}.$$

6. **Volumes.**

   (Preview: we're going to use this to define actions and covariant derivatives.)

   We're going to want to write $S = \int_M \mathrm{d}^D x \mathcal{L}(x)$ where $\mathcal{L}(x)$ is some (scalar) function on x.

   [Here $\mathcal{L}$ will depend on x via various fields, and the point of S is so we can vary it; $\frac{\delta S}{\delta \phi(x)} = 0$ is the EOM.]

   But $\mathrm{d}^D x$ is not coordinate invariant.

   Claim: $\mathrm{d}^D x \sqrt{g}$ is coordinate invariant. Here, by $\sqrt{g}$ I mean $\sqrt{|\det g|}$. I am omitting the abs by convention. If you like I am defining $g \equiv -\det g$. $\det g = \det_{\mu\nu} g_{\mu\nu}$.

   Proof:

$$\tilde{x}^\mu = \tilde{x}^\mu(x^\nu). \quad J^\mu_\nu \equiv \frac{\partial \tilde{x}^\mu}{\partial x^\nu}$$

$$\tilde{g}_{\mu\nu} = \left(J^{-1}\right)^\alpha_\mu \left(J^{-1}\right)^\beta_\nu g_{\alpha\beta}.$$

$$\det \tilde{g} = \det J^{-2} \det g, \qquad \mathrm{d}^D \tilde{x} = \det J \mathrm{d}^D x.$$

   So:

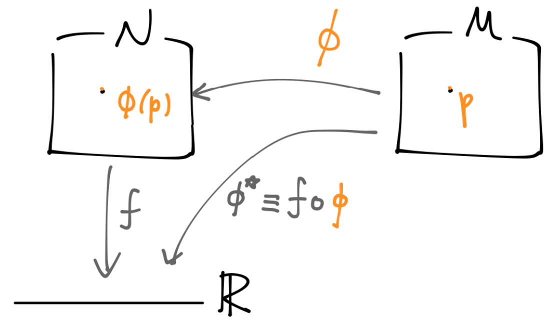
$$\sqrt{g} \mathrm{d}^D x = \sqrt{\tilde{g}} \mathrm{d}^D \tilde{x}.$$

7. Make a stress tensor. Put some stuff (a classical field theory) on a spacetime with a metric. By varying the metric at a point, we can probe the response of the stuff. We'll do this in §7.

8. Derivatives of tensors which are tensors (next).

5.2.4 Mathematical due-diligence postlude: Pullback and pushforward

[Wald Appendix C] A smooth map from one manifold to another $\phi : M \to N$ induces some maps between tensor fields on M and N. In particular, ϕ "pulls back" a function $f : N \to \mathbb{R}$ to $f \circ \phi : M \to \mathbb{R}$. Also, we can make a map ("push forward")



$$\begin{array}{rcl} \phi^\star : T_pM & \to & T_{\phi(p)}N \\ v & \mapsto & \phi^\star(v) \end{array}$$

defined by its action on a function on N:

$$(\phi^\star v)(f) \equiv v(f \circ \phi) \ .$$

In components, this is just the chain rule: if we pick coordinates x^μ on M and y^a on N, then the map ϕ is telling us $y(x)$, and

$$(\phi^\star v)(f) = v^\mu \frac{\partial}{\partial x^\mu}\left(f(y(x))\right) = v^\mu \frac{\partial y^a}{\partial x^\mu}\frac{\partial}{\partial y^a} f.$$

That is:

$$(\phi^\star v)^a = (\phi^\star)^a_\nu\, v^\nu = \frac{\mathrm{d}y^a}{\mathrm{d}x^\nu} v^\nu, \quad i.e. \quad (\phi^\star)^a_\nu = \frac{\mathrm{d}y^a}{\mathrm{d}x^\nu} \ .$$

Notice that $\phi^\star$ is a linear map[18]. This also induces a pullback map on forms (for p in the image of ϕ):

$$\begin{array}{rcl} \phi_\star : T^\star_{\phi(p)}M & \to & T^\star_pM \\ \omega & \mapsto & \phi_\star\omega \end{array}$$

by the natural thing: given $v \in T_pM$,

$$\phi_\star\omega(v) \equiv \omega(\phi^\star v) = \omega_a \frac{\partial y^a}{\partial x^\nu} v^\nu.$$

If the map ϕ is a diffeomorphism (which in particular requires that the dimensions of M and N are the same), then we can do more, since we can also use ϕ^{-1}, and the Jacobian matrix $\frac{\mathrm{d}y^a}{\mathrm{d}x^\nu}$ is invertible and gives a map $\left(\phi^{-1}\right)^\star : V_{\phi(p)} \to V_p$ (where V is some space of tensors). Specifically, given a rank $\begin{pmatrix} r \\ s \end{pmatrix}$ tensor T on M, we can make one on N defined its action on a collection of vectors and forms v, ω:

$$(\phi^\star T)^{b_1\cdots b_r}_{a_1\cdots a_s}(\omega_1)_{b_1}...(\omega_r)_{b_r}(v_1)^{a_1}...(v_s)^{a_s} = T^{\mu_1\cdots\mu_r}_{\nu_1\cdots\nu_s}\left(\left(\phi^{-1}\right)^\star \omega_1\right)_{\mu_1}\cdots(\phi^\star v_1)^{\nu_s} \ .$$

[18] Notice that it makes things clearer to use a different index set for the new coordinate; μ and a run over the same values $0,1,2,3$, but using different names tells us whether we should contract with $\frac{\partial}{\partial x^\mu}$ or $\frac{\partial}{\partial y^a}$ to make a coordinate-invariant quantity.

The way to read this equation is: "stick the matrix $\frac{dy^a}{dx^\nu}$ or its inverse $\frac{dx^\mu}{dy^a}$ wherever you need to in order to get the indices to work out". I will not keep careful track of where the stars and inverses are. In any example, there's only one right place to put them, so there's no need for them.

Recall that *diffeomorphic* was the precise notion of *same* in the category of smooth manifolds. The existence of a diffeomorphism $\phi : M \to N$ means that the two manifolds are just related by a change of coordinates. Classical field theories on these two manifolds therefore produce the same physics. The previous discussion shows how to use this to translate all the tensor fields living on M into tensor fields living on N.

5.2.5 Appendix: Integration of forms over submanifolds

We can describe a submanifold Σ of a manifold M in two very different ways: by giving conditions on coordinates of M which specify where Σ is (such a thing is called a *variety* for some reason), or by *embedding* Σ into M by maps from some auxiliary parameter space. We'll discuss these in order, and how to think about integrals in each case.

1. So for example we can specify $\Sigma_{n-k} \subset M_n$ by giving k (independent, local) conditions

$$\Sigma = \{y_\alpha = 0, \alpha = 1...k\}$$

   the number of such conditions is called the *codimension* of Σ. (In general, these conditions may have to be locally specified in order to be independent. This is a long story.)

   Then, given a $n-k$ form on M, we can define a number by integrating it over the submanifold Σ_{n-k} by

$$\int_{\Sigma_{n-k}} \omega_{n-k} = \int_M \omega_{n-k} \wedge dy^1 \wedge ... \wedge dy^k \delta^k(y)$$

   On the RHS is an integral of an n-form over all of M, which (if M is connected) it must be proportional to the volume form. Notice that this is a direct generalization of our expression for $\int_{\text{worldline}} A$ the minimal coupling term in the worldline action in (28).

2. Alternatively, suppose we are given some maps whose image is Σ:

$$\begin{array}{rcl} x : \mathcal{U} \subset \mathbb{R}^{n-k} & \to & M \\ (\sigma^\alpha)_{\alpha=1..n-k} & \mapsto & (x^a(\sigma))_{a=1..n} \end{array}$$

   This generalizes our notion of a parametrized curve $x^\mu(\mathfrak{s})$. Notice that we haven't said anything about the parametrization. Again these maps may need to be defined locally.

In this case we can define the integral of a form of appropriate rank by

$$\int_{\Sigma_{n-k}} \omega_{n-k} = \int \mathrm{d}\sigma^1 \cdots \mathrm{d}\sigma^{n-k} \epsilon^{\alpha_1 \ldots \alpha_{n-k}} \frac{\partial x^{a_1}}{\partial \sigma^{\alpha_1}} \cdots \frac{\partial x^{a_{n-k}}}{\partial \sigma^{\alpha_{n-k}}} \omega_{a_1 \ldots a_{n-k}} \quad .$$

(If I wanted to be pedantic I would write this as

$$\int_{\Sigma_{n-k}} x_\star \omega_{n-k}$$

where $x_\star$ is the pullback of the form by the map x to a form $\mathcal{U}$.)

Notice that Stokes' theorem is beautifully simple in this notation:

$$\int_{\partial\Sigma} \omega = \int_{\Sigma} \mathrm{d}\omega$$

where $\partial\Sigma$ is the boundary of Σ (and gives zero if Σ has no boundary).

Notice that we didn't use the metric in the above discussion. This implies another possibility for covariantly integrating things to make actions, besides the $\int \sqrt{g}$(scalar) strategy that we used above. If, amongst our fields for which we're constructing an action, are some differential forms, and we can wedge[19] them together to make a form $\Omega = \Omega_{a_1 \cdots a_n} \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ of rank $n = \dim M$, we can integrate

$$\int_M \Omega = \int \mathrm{d}^n x \Omega_{1 \cdots n}$$

and it is coordinate invariant. This is because under a coordinate transformation $\mathrm{d}^n \tilde{x} = \det J \mathrm{d}^n x$, while $\tilde{\Omega}_{1 \cdots n} = \det^{-1} J \Omega_{1 \cdots n}$.

A simple example is in Maxwell theory in four dimensions. In that case $F = \mathrm{d}A$ is a two-form, and we can make a four-form which is gauge invariant by $F \wedge F$. The resulting term

$$\frac{\theta}{16\pi^2} \int_M F \wedge F$$

is generally covariant, independent of the metric. It is the same as $\star F^{\mu\nu} F_{\mu\nu} \propto \vec{E} \cdot \vec{B}$. Such a term which is generally covariant, independent of the metric is said to be *topological*. Another important example arises in gauge theory in $2+1$ dimensions, where $S_{CS}[A] \equiv \frac{k}{4\pi} \int A \wedge F$ is called a *Chern-Simons term*.

---

[19] By 'wedge' I mean the following. We showed earlier how to make a $p+1$ form from a p form by exterior derivative. Given a p-form $\omega = \omega_{a_1 \cdots a_p} \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_p}$ and a q-form $\mu = \mu_{b_1 \cdots b_q} \mathrm{d}x^{b_1} \wedge \ldots \wedge \mathrm{d}x^{b_q}$ we can make a $p+q$-form by antisymmetrizing the product:

$$\omega \wedge \mu \equiv \omega_{a_1 \cdots a_p} \mu_{b_1 \cdots b_q} \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_p} \wedge \mathrm{d}x^{b_1} \wedge \ldots \wedge \mathrm{d}x^{b_q} \ .$$

5.3 Derivatives

Consider $V^\mu(x)$, the components of a vector field in a coordinate basis. We want to differentiate the vector field.

What not to do: What might we mean by $\partial_\nu V^\mu$? Under a change of coordinates $x^\mu \to \tilde{x}^\mu$,

$$\partial_\nu V^\mu \mapsto \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \partial_\alpha \left(\frac{\partial \tilde{x}^\mu}{\partial x^\beta} V^\beta \right) = \tilde{\partial}_\nu x^\alpha \partial_\beta \tilde{x}^\mu \partial_\alpha V^\beta + \left(\tilde{\partial}_\nu x^\alpha \partial_\beta \partial_\alpha \tilde{x}^\mu \right) V^\beta. \tag{55}$$

This is not a tensor! Also from (55) you see why antisymmetrizing is a good idea, and the exterior derivative is perfectly coordinate invariant.

How to take a derivative of a tensor and get a tensor?

For scalar quantities there's no problem. Consider a scalar field living on some R'n manifold, $\phi(x)$. Like temperature of CMB. Or the Higgs condensate. $\partial_\mu \phi$ is perfectly coordinate covariant. It transforms like a vector.

Q: Is there an intrinsic definition of derivatives?

A: At least two.

5.3.1 Lie Derivative

Given a vector field v on M, we know how to use it to take a derivative of a function f on M in a coordinate-invariant way:

$$vf = v^\mu \partial_\mu f \equiv \mathcal{L}_v f$$

– this (directional derivative) is sometimes called the *Lie derivative* of f along v.

The general def of a Lie derivative of a tensor T along a vector field v is as follows. Recall that vector field on M produces a *flow* $\phi_s : M \to M$ (the v will be implicit here). The Lie derivative of T along v is:

$$\mathcal{L}_v T(p) = \lim_{s \to 0} \left(\frac{\phi^\star_{-s} T(\phi_s(p)) - T(p)}{s} \right)$$

Notice that the difference on the RHS is between tensors at the same point p, since $\phi^\star_{-s} : V_{\phi_s(p)} \to V_p$.

For a function, this is just

$$\mathcal{L}_v f(p) = \frac{\mathrm{d}}{\mathrm{d}s} \left(f(\phi_s(p)) \right)|_{s=0} \overset{\text{def of } \phi}{=} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \frac{\partial f}{\partial x^\mu} = v^\mu \frac{\partial f}{\partial x^\mu} = vf.$$

The Lie derivative of a vector field w along a vector field v is

$$\mathcal{L}_v w \equiv [v, w]$$

– it's just the commutator. This is a useful exercise in unpacking the definition of the pullback map. [20]

Given a vector field v, we can also construct explicitly the Lie derivative along v of other tensor fields. To figure out the explicit form, note again from the definition that $\mathcal{L}_v$ is Liebniz, so

$$\underbrace{\mathcal{L}_v(\omega(Y))}_{\text{known}} = (\mathcal{L}_v\omega)\, Y + \omega \underbrace{\mathcal{L}_v Y}_{\text{known}}.$$

One point of Lie derivative is to detect symmetries of the metric. The statement that for some vector field v

$$\mathcal{L}_v g = 0$$

means that the metric doesn't change along the flow by v – it's a symmetry of g. Such a v is called a Killing vector field.

Summary of properties:

$$\text{on a function, it gives:} \quad \mathcal{L}_v f = vf$$

$$\text{on a vector field, it gives:} \quad \mathcal{L}_v w = [v, w]$$

$$\text{it's linear:} \quad \mathcal{L}_v\left(aT_1 + bT_2\right) = a\mathcal{L}_v T_1 + b\mathcal{L}_v T_2$$

$$\text{it's Liebniz:} \quad \mathcal{L}_v\left(T_1 T_2\right) = \left(\mathcal{L}_v T_1\right) T_2 + T_1\left(\mathcal{L}_v T_2\right).$$

Recall that the last two combine to the statement that $\mathcal{L}_v$ is a derivation.

[20] Here we go. Expanding in a coordinate basis for T_pM $v = v^\mu \partial_\mu, w = w^\mu \partial_\mu$, we have

$$\mathcal{L}_v w = \mathcal{L}_v\left(w^\mu \frac{\partial}{\partial x^\mu}\right) = v^\nu \partial_\nu w^\mu \frac{\partial}{\partial x^\mu} + w^\mu \mathcal{L}_v\left(\frac{\partial}{\partial x^\mu}\right)$$

using the Liebniz property. Now

$$\mathcal{L}_v\left(\frac{\partial}{\partial x^\mu}\right) \equiv \lim_{t\to 0} \frac{\phi_v(-t)^\star\left(\frac{\partial}{\partial x^\mu}\right) - \frac{\partial}{\partial x^\mu}}{t} = -\frac{\partial v^\nu}{\partial x^\mu}\frac{\partial}{\partial x^\nu}.$$

Putting these together gives the coordinate expression for the commutator:

$$\mathcal{L}_v w = \left(w^\mu \partial_\mu v^\nu - v^\mu \partial_\mu w^\nu\right)\frac{\partial}{\partial x^\nu} = [v, w].$$

You may like Wald's argument for this; he choose coordinates where the flow parameter t is one of the coords, $x^1 = t$.

5.3.2 Covariant Divergence

[Zee page 326] First we are going to guess a (useful) formula for a covariant divergence that does not involve the horrible Christoffel symbols. We can check that it is indeed covariant. Then we will come back and develop the fancy-pants understanding of why this is a good construction (in terms of parallel transport), and how to extend it to a more general derivative operator.

The basic reason we want to differentiate tensors is so that we can write field equations. Where do field equations come from? (For non-dissipative systems) they come from action principles. Actions are simpler – for each field configuration they give a *number*, usually (always in this class) a real number. Why in the world would we start by writing field equations, which is some horrible tensor-field-valued functional? Further, the action has to be *invariant* under coordinate changes in order for physics not to depend on our choice of coordinates. Whereas the horrible tensor-valued functional merely has to be covariant.

This apparent detour will be useful because we'll use the same kind of logic in constructing the Einstein-Hilbert action for the metric.

So let's try to guess what a coordinate invariant action functional for a scalar field in curved space might be. We want it to be *local* in spacetime, so that we don't get action at a distance. What I mean is that it should be a single integral over spacetime

$$S[\phi] = \int_M L(\phi, \partial_\mu \phi)$$

of a lagrangian density with finitely many derivatives of ϕ. Wait – we're trying to figure out how to take derivatives here. But we already know how to take *one* derivative of ϕ in a covariant way, it's just $\partial_\mu \phi$. So let's just look for a lagrangian density which depends only on $\phi, \partial_\mu \phi$.

So there are two things I should clarify: what I mean by $\int_M$, and what should we use for L? We already found a covariant integration measure above[21]. So more explicitly we have:

$$S[\phi] = \int_M \mathrm{d}^D x \sqrt{g} L(\phi(x), \partial_\mu \phi(x))$$

To make scalars, contract indices.

Just for simplicity, let's suppose $\phi \to \phi + \epsilon$ is a symmetry (like a goldstone boson), so the action can't depend explicitly on ϕ. Then the invariant thing with the fewest derivatives is:

$$S[\phi] = \int_M \mathrm{d}^D x \sqrt{g} \partial_\mu \phi \partial_\nu \phi g^{\mu\nu}$$

[21] There may be other ways to make a covariant integral; for another example, see §5.2.5.

[End of Lecture 8]

What are the equations of motion? If we were in flat space, it would be $\Box\phi = 0$, the wave equation. What is $\Box\phi$? In flat space it is $\partial_\mu \partial^\mu \phi$, a second derivative of ϕ! In curved space it is:

$$-\frac{1}{\sqrt{g}} \frac{\delta S}{\delta \phi(x)} = \frac{1}{\sqrt{g}} \partial_\mu \sqrt{g} g^{\mu\nu} \partial_\nu \phi \equiv \Box \phi. \tag{56}$$

(The factor of $\frac{1}{\sqrt{g}}$ on the LHS is to make it a scalar, rather than a scalar density.)

On HW: Similarly for Maxwell: where you see η put g, and stick a $\sqrt{g}$ in the measure.

So we've learned to differentiate the particular vector field $\partial^\mu \phi = \nabla^\mu \phi = g^{\mu\nu} \nabla_\mu \phi$ in a covariant way, but only in a certain combination. Demanding that

$$\Box\phi \overset{!}{=} \nabla_\mu \nabla^\mu \phi$$

and comparing with our expression (56) suggests that we should define

$$\nabla_\nu V^\mu \overset{?}{=} \frac{1}{\sqrt{g}} \left(\partial_\nu \sqrt{g} V^\mu\right) \tag{57}$$

Is this covariant? In fact, no.

5.3.3 Parallel transport and covariant derivative

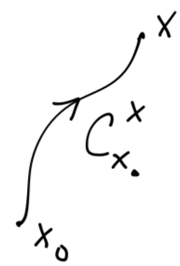
Let's think for a moment about a (quantum) charged particle moving in a background abelian gauge field configuration. Its wavefunction is $\psi(x)$ and transforms under a gauge transformation like

$$A_\mu \to A_\mu + \partial_\mu \lambda, \quad \psi(x) \to e^{\mathbf{i}e\lambda(x)}\psi(x)$$

(in order to preserve the form of the Schrödinger equation). This means that $\partial_\mu \psi$ transforms weirdly under gauge transformations, by an extra term proportional to $(\mathbf{i}e\partial_\mu \lambda)\,\psi(x)$. On the other hand, the *covariant* derivative

$$D_\mu \psi(x) \equiv (\partial_\mu - \mathbf{i}eA_\mu)\,\psi(x)$$

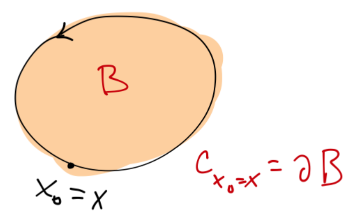
transforms just like ψ.



A *covariantly constant* wavefunction the satisfies $D_\mu \underline{\psi}(x) = 0$. Formally this is solved by the exp of the line integral:

$$\underline{\psi}(x) = e^{\mathbf{i}\int_{C_{x_0}^x} A_\mu(y)dy^\mu}\psi(x_0)$$

where C is a path with endpoints x_0, x. This process of constructing $\underline{\psi}$ is called *parallel transport*; it gives us a gauge-covariant way to compare objects (here, wavefunctions).



Notice that the answer for $\underline{\psi}(x)$ depends on the path if $\mathrm{d}A \neq 0$. In particular, for a *closed* path $(x_0 = x)$, the phase acquired is

$$e^{\mathbf{i}\oint_C A} = e^{\mathbf{i}\int_B F}$$

where $F = F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$ is the field strength and $C = \partial B$. That is: B is a 2d region whose boundary is the closed curve C. Notice that a two-form is something that we can integrate over a two-dimensional subspace (see the appendix 5.2.5 for an amplification of this remark).

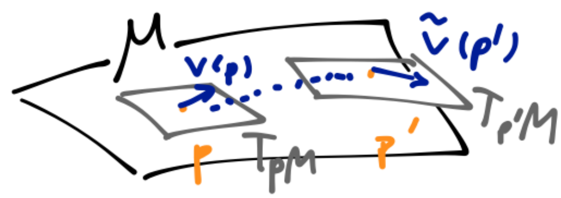
What's a derivative?

$$V^\mu = V^\mu(x, y, z...), \quad \partial_x V^\mu \equiv \lim_{\epsilon\to 0} \frac{1}{\epsilon} \left(V^\mu(x+\epsilon, y, z...) - V^\mu(x, y, fz...) \right)$$

But $V(x) \in T_pM$ and $V(x+\epsilon) \in T_{x^{-1}(x(p)+\epsilon)}M$ are in different spaces! In order to take differences of vectors at nearby points, we need to put them in the same vector space! First we should transport $V|_p$ to $T_{x^{-1}(x(p)+\epsilon)\equiv p'}M$ and then take the difference. To do this, we need some additional structure (in the sidebar above, the additional structure came from the background gauge field A_μ; below (in 5.4.1) we'll see how to get it from the metric).

Parallel transport: Take a vector field V on M. Parallel transport is a linear map

$$\begin{array}{rcl} T_pM & \to & T_{p'}M \\ V^\mu & \mapsto & \tilde{V}^\mu = U^\mu_\nu V^\nu \end{array} \tag{58}$$



with $U^\mu_\nu(\epsilon) \overset{\epsilon\to 0}{\to} \delta^\mu_\nu$ and $U^\mu_\nu(\epsilon)$ differentiable in ϵ, so it can be Taylor expanded:

$$\tilde{V}^\mu = U^\mu_\nu V^\nu = V^\mu \underbrace{-}_{\text{convention}} \underbrace{\epsilon^\rho}_{\substack{\text{pick an } \epsilon \\ \text{in each coord dir}}} \underbrace{\Gamma^\mu_{\rho\nu}}_{\equiv \text{ connection}} V^\nu + \mathcal{O}(\epsilon^2)$$

From this (we haven't shown how to make U or Γ yet) we can make a *covariant derivative* ∇:

$$\underbrace{V^\mu \underbrace{(p+\epsilon)}_{\equiv p'} - \tilde{V}^\mu(p+\epsilon)}_{=\epsilon^\rho \left(\partial_\rho V^\mu + \Gamma^\mu_{\rho\nu} V^\nu \right)} \equiv \epsilon^\rho \nabla_\rho V^\mu$$

In the underbraced step, we used the fact that when $\epsilon = 0$, $V = \tilde{V}$, and I am setting to zero the $\mathcal{O}(\epsilon)^2$ terms.

$$\implies \quad \boxed{\nabla_\rho V^\mu \equiv \partial_\rho V^\mu + \Gamma^\mu_{\rho\nu} V^\nu}.$$

We can demand that this transforms as a tensor (with one covariant and one contravariant index), and that will determine how Γ should transform. We can use this information to define a covariant derivative on any tensor field by further demanding that ∇ is Liebniz (it is already linear, so this makes it a derivation), via *e.g.*

$$\nabla_\mu \left(V^\nu \omega_\nu \right) \overset{!}{=} \left(\nabla_\mu V^\nu \right) \omega_\nu + V^\nu \nabla_\mu \omega_\nu, \quad \text{plus} \quad \nabla_\mu \Bigg(\underbrace{V^\nu \omega_\mu}_{\text{scalar!}} \Bigg) = \partial_\mu \left(V^\nu \omega_\nu \right).$$

After some algebra, this gives:

$$\boxed{\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu \underbrace{-}_{\uparrow} \Gamma^\rho_{\mu\nu} \omega_\rho}.$$

Notice the change of sign in the connection term between contravariant and covariant tensors. Similarly, we can figure out how to get $\nabla_\mu T^{\rho\sigma}$ by using the Liebniz property on $\nabla_\mu (v^\rho u^\sigma) =$ ... and the demand that any such tensor field gets transported the same way. This gives

$$\nabla_\mu T^{\rho\sigma} = \partial_\mu T^{\rho\sigma} + \Gamma^\rho_{\mu\lambda} T^{\lambda\sigma} + \Gamma^\sigma_{\mu\lambda} T^{\rho\lambda} \ .$$

I hope you see the pattern.

Now let's talk about how Γ transforms under coordinate changes[22]. The demand we are really making is that parallel transport is independent of coordinate choice. This is the same as demanding that $\nabla_\mu v^\nu$ are the components of a $\binom{1}{1}$ tensor. It transforms as

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\rho} v^\rho \mapsto \tilde{\partial}_\mu x^\rho \partial_\sigma \partial_\rho \tilde{x}^\nu v^\sigma + \tilde{\partial}_\mu x^\rho \partial_\sigma \tilde{x}^\nu \begin{pmatrix} \partial_\rho v^\sigma \\ +\tilde{\Gamma}^\sigma_{\rho\kappa} \tilde{v}^\kappa \end{pmatrix} \stackrel{!}{=} \tilde{\partial}_\mu x^\rho \partial_\sigma \tilde{x}^\nu \begin{pmatrix} \partial_\rho v^\sigma \\ +\Gamma^\sigma_{\rho\kappa} v^\kappa \end{pmatrix} \quad (59)$$

from which we infer (notice that the top lines are already equal):

$$\tilde{\Gamma}^\mu_{\nu\rho} = \underbrace{\tilde{\partial}_\nu x^\sigma \partial_\kappa \tilde{x}^\mu \tilde{\partial}_\rho x^\delta \Gamma^\kappa_{\sigma\delta}}_{\text{tensor transf}} - \underbrace{\partial_\kappa \partial_\delta \tilde{x}^\mu \tilde{\partial}_\nu x^\kappa \tilde{\partial}_\rho x^\delta}_{\text{extra badness}} \ . \quad (60)$$

The extra badness in the transformation of Γ is precisely designed to cancel the badness from the transformation of the ordinary derivative of the vector (55). [For more discussion of this sea of indices, see Zee page 329]

Q: What does it mean for a vector field to be constant? Consider the vector field $W = \partial_x$ in the plane. Write its components in polar coordinates and compute their partial derivatives. [see Zee page 323 and problem set 5.]

5.4 Curvature and torsion

Given a connection Γ, there are two tensor fields you can make right away.

1) Torsion tensor $\equiv$ the antisymmetric part of the connection.

$$\Gamma^\mu_{[\rho,\sigma]} \equiv \Gamma^\mu_{\rho\sigma} - \Gamma^\mu_{\sigma\rho}.$$

Notice that the bad stuff of the transformation law (60) (the inhomogeneous term) cancels out of this combination:

$$\tilde{\Gamma}^\mu_{[\sigma,\rho]} = \tilde{\partial}_\rho x^\alpha \tilde{\partial}_\sigma x^\beta \partial_\gamma \tilde{x}^\mu \Gamma^\gamma_{[\alpha,\beta]}.$$

Below, we'll choose a connection with no torsion, $\Gamma^\gamma_{[\alpha,\beta]} = 0$. Comments:

[22] Please studiously ignore Wald's discussion of this point which could not be made more confusing.

1. Because the torsion is a tensor, this choice is coordinate-independent.

2. If you want you can undo this assumption. Then you will get a more complicated theory of gravity which is experimentally different from GR and wrong [23].

3. Some effects of this assumption:
   – In the exterior derivative, we can use ordinary derivative or covariant derivative and it doesn't matter:

$$\mathrm{d}\omega = \nabla_{[\mu}\omega_{\nu]}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = \left(\partial_{[\mu}\omega_{\nu]} - \Gamma^\rho_{[\mu,\nu]}\right)\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \ .$$

   In general they are both tensors, but in the torsion-free case, they are equal.

   – For scalars:

$$\nabla_{[\mu}\nabla_{\nu]}f = \nabla_\mu\left(\partial_\nu f\right) - \nabla_\nu\left(\partial_\mu f\right) = \left(\partial_\mu\partial_\nu - \Gamma^\rho_{\mu\nu}\right)f - \left(\partial_\nu\partial_\mu - \Gamma^\rho_{\nu\mu}\right)f = \Gamma^\rho_{[\nu,\mu]}\partial_\rho f$$

   so if $T^\rho_{\nu\mu} \equiv \Gamma^\rho_{[\nu,\mu]} = 0$ then $\nabla_{[\mu}\nabla_{\nu]}f$ – the mixed covariant derivatives commute *on scalars*. I emphasize those last two words because of the next item.

2) Curvature tensor.

Consider more generally what happens if we take the commutator of covariant derivatives, for example, when acting on a one-form. By construction, what we get is also (the components of) a tensor:

$$\begin{aligned}
\nabla_{[\mu}\nabla_{\nu]}\omega_\rho &= \nabla_\mu\nabla_\nu\omega_\rho - (\mu \leftrightarrow \nu) \\
&= \nabla_\mu\left(\partial_\nu\omega_\rho - \Gamma^\kappa_{\nu\rho}\omega_\kappa\right) - (\mu \leftrightarrow \nu) \\
&= \left(\partial_\mu\partial_\nu\omega_\rho - \Gamma^\gamma_{\mu\nu}\partial_\gamma\omega_\rho - \Gamma^\gamma_{\mu\rho}\partial_\nu\omega_\gamma - \Gamma^\kappa_{\nu\rho}\partial_\mu\omega_\kappa + (\text{terms without derivatives of }\omega)_{\mu\nu}\right) \\
&- \left(\partial_\nu\partial_\mu\omega_\rho - \Gamma^\gamma_{\nu\mu}\partial_\gamma\omega_\rho - \Gamma^\gamma_{\nu\rho}\partial_\mu\omega_\gamma - \Gamma^\kappa_{\mu\rho}\partial_\nu\omega_\kappa + (\text{terms without derivatives of }\omega)_{\nu\mu}\right) \\
&\quad (\text{If } T = 0, \text{ derivatives of } \omega \text{ cancel! }) \\
&\equiv R_{\mu\nu\rho}{}^\sigma\omega_\sigma
\end{aligned} \tag{61}$$

(Beware that there is some disagreement about the sign in the definition of the curvature. In particular, Zee (page 341) uses the opposite sign.)

Alternative to explicit computation above: Consider any scalar function f. Then

$$\nabla_{[\mu}\nabla_{\nu]}\left(f\omega_\rho\right) = \underbrace{\left(\nabla_{[\mu}\nabla_{\nu]}f\right)}_{=0 \text{ if no torsion}}\omega_\rho + f\nabla_{[\mu}\nabla_{\nu]}\omega_\rho = f\nabla_{[\mu}\nabla_{\nu]}\omega_\rho$$

[23] Unsubstantiated claim which we should revisit after we understand how to use this machinery to do physics: A nonzero torsion included in the simplest way would violate the equivalence principle, from the contribution of the energy of the gravitational field.

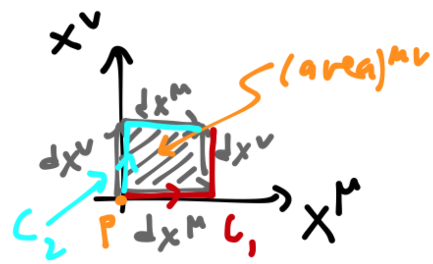
(The terms where one derivative hits each cancel in antisymmetrization.) But this not impossible if $\nabla_{[\mu}\nabla_{\nu]}\omega_\rho$ involves any derivatives on ω, so it must be of the form $\nabla_{[\mu}\nabla_{\nu]}(f\omega_\rho) = fR_{\mu\nu\rho}{}^\sigma\omega_\sigma$.

From the explicit expression above, we have:

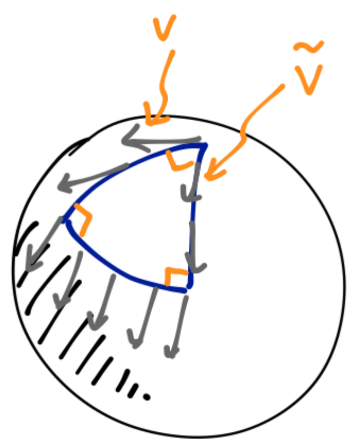
$$\boxed{\begin{aligned} R_{\mu\nu\rho}{}^\sigma &= \partial_\nu\Gamma^\sigma_{\mu\rho} - \partial_\mu\Gamma^\sigma_{\nu\rho} + \Gamma^\kappa_{\mu\rho}\Gamma^\sigma_{\nu\kappa} - \Gamma^\kappa_{\nu\rho}\Gamma^\sigma_{\mu\kappa} \\ &= -\left(\partial_\mu\Gamma^\sigma_{\nu\rho} - \Gamma^\kappa_{\mu\rho}\Gamma^\sigma_{\nu\kappa} - (\mu \leftrightarrow \nu)\right). \end{aligned}}$$

Beware the red sign – don't trust me to get it right[24]; you get the other sign if you define R via commutators on *vectors*, instead of one-forms like we did above. See Zee p. 351 for a mnemonic for R from Γ, though beware of his opposite sign convention.

Note that until we decide to use the metric (Christoffel) connection below, the order of indices matters here.



Geometric interpretation: The curvature measures the path-dependence of the result of parallel transport (just like $F_{\mu\nu}$ in the charged-particle parable at the beginning of the parallel transport discussion 5.3.3). Consider two 2-step paths which connect the same two points: one path goes along two fixed infinitesimal directions dx^μ and then dx^ν (these are not one-forms, just infinitesimal coordinate distances); the other traverses these infinitesimal line segments in the opposite order. Infinitesimally upon parallel transport of v^ρ in the direction dx,



$$v^\rho \mapsto (U_{dx}v)^\rho = \tilde{v}^\rho(p+dx) = v^\rho(p+dx) - dx\nabla_x v^\rho$$

(from the definition of ∇). The second step results in

$$(U_{C_1}v)^\rho = \tilde{v}^\rho(p+dx+dy) = v^\rho(p+dx+dy) + dy dx \nabla_y\nabla_x v^\rho.$$

The difference in the resulting v's is

$$\Delta v^\rho \equiv (U_{C_1}v)^\rho - (U_{C_2}v)^\rho = -dx^\mu dx^\nu[\nabla_\mu, \nabla_\nu]v^\rho = -(\text{area})^{\mu\nu}R^\rho_{\mu\nu\sigma}v^\sigma \quad \text{(no sum on } \mu,\nu!\text{)}.$$

Here $(\text{area})^{\mu\nu} = dx^\mu dx^\nu$ Notice that it is useful to think of the RHS as a matrix acting on v (it is linear after all):

$$\Delta v^\rho = \pm\left((\text{area})^{\mu\nu}R_{\mu\nu}\right)^\rho_\sigma v^\sigma$$

[24]Thanks to Frederick Matsuda for pointing out an important sign error in a previous version of these notes in the second line of the above boxed equation.

Exercise: check this formula using the explicit Riemann tensor on S^2 that you will find on problem set 5, using the diagram at right: begin at the north pole and carry a south-pointing vector with you toward the equator; after an infinitesimal step south $d\theta$, head east so that your longitude changes by $d\varphi$, still carrying your south-pointing vector; then head back north and compare your south-pointing vector to the original one. [End of Lecture 9]

A word of warning about notation (which I failed to heed in a previous version of these notes): when studying tensors with multiple indices which are not symmetrized (such as $R_{\mu\nu\rho}{}^{\sigma}$), it is a good idea to keep track of the order of the indices, even when they are distinguished by their height (*i.e.* σ is a covariant index and hence special in the previous expression). This is so that if you have occasion to lower an index (as in *e.g.* $R_{\mu\nu\rho\lambda} \equiv R_{\mu\nu\rho}{}^{\sigma} g_{\sigma\lambda}$) it is clear which of the now-all-lower indices is the one that began its life as an upper index. If you want to TeX these things, this is accomplished by `R_{\mu\nu\rho}{}^\sigma` rather than `R_{\mu\nu\rho}^\sigma`.

Properties of $R_{\mu\nu\rho}{}^{\sigma}$: (The symmetry properties below are very useful – they reduce the 4^4 components of a general 4-index object down to just at-worst 20 independent components of R)

1. From the definition we see that $R_{\mu\nu\rho}{}^{\sigma} = -R_{\nu\mu\rho}{}^{\sigma}$.

2. If $\Gamma^{\rho}_{[\mu,\nu]} = 0$ then

$$0 = R_{[\mu\nu\rho]}{}^{\sigma} \equiv \begin{matrix} R_{\mu\nu\rho}{}^{\sigma} & +R_{\nu\rho\mu}{}^{\sigma} & +R_{\rho\mu\nu}{}^{\sigma} \\ -R_{\nu\mu\rho}{}^{\sigma} & -R_{\mu\rho\nu}{}^{\sigma} & -R_{\rho\nu\mu}{}^{\sigma} \end{matrix}$$

   That is, R has 'cyclic symmetry'.

   Proof: For any one-form ω,

$$\begin{aligned} \nabla_{[\mu}\nabla_{\nu]}\omega_{\rho} &\overset{\text{no torsion}}{=} R_{\mu\nu\rho}{}^{\sigma}\omega_{\sigma} \\ \implies R_{[\mu\nu\rho]}{}^{\sigma}\omega_{\sigma} &= 2\nabla_{[\mu}\nabla_{\nu}\omega_{\rho]} \end{aligned} \tag{62}$$

   But $\nabla_{[\mu}\omega_{\nu]} = \partial_{[\mu}\omega_{\nu]} + \underbrace{\Gamma^{\rho}_{[\mu,\nu]}}_{=T=0}$ so

$$R_{[\mu\nu\rho]}{}^{\sigma}\omega_{\sigma} = 2\partial_{[\mu}\partial_{\nu}\omega_{\rho]} = 2\begin{pmatrix} \partial_{\mu}\partial_{\nu}\omega_{\rho} & +\partial_{\nu}\partial_{\rho}\omega_{\mu} & +\partial_{\rho}\partial_{\mu}\omega_{\nu} \\ -\partial_{\nu}\partial_{\mu}\omega_{\rho} & -\partial_{\mu}\partial_{\rho}\omega_{\nu} & -\partial_{\rho}\partial_{\nu}\omega_{\mu} \end{pmatrix} = 0 \tag{63}$$

3. Bianchi identity (not just algebraic):

$$\nabla_{[\mu}R_{\nu\rho]\sigma}{}^{\delta} = 0.$$

   Notice its resemblance to $0 = \epsilon^{\mu\nu\rho\sigma}\partial_{\nu}F_{\rho\sigma} = (\star \mathrm{d}F)^{\mu}$, which came from $F = \mathrm{d}A$.

   We will prove this statement later (in §6.0.6) using Riemann normal coordinates (since it's a tensor equation, proving it in some special coordinate system suffices to show that it's a true statement about the tensor), once we know what those are. It also follows simply from the Cartan structure equations (§9.)

5.4.1 Whence the connection? (Christoffel gets it from the metric)

There are many rocks underneath which one can find connections. Any correspondence between tangent spaces at different points (linear, differentiable) does it. A nice one comes from the metric. This is called the Christoffel connection. It is determined by the two reasonable demands:

- torsion-free: $\Gamma^{\rho}_{[\mu\nu]} = 0$.
- 'metric-compatible': $\nabla_\mu g_{\nu\rho} = 0$. The metric is covariantly constant. The wonderful practical consequence of this is: you can raise and lower indices without worrying about whether you did it before or after doing the parallel transport or the covariant derivative.

  Geometric interpretation: This is a very natural condition – it says that parallel transport preserves the (proper) lengths of vectors, as well as angles between them, as follows: v being parallelly transported in the ϵ^μ direction means

  $$0 = v^\mu(x+\epsilon) - \tilde{v}^\mu(x+\epsilon) = \epsilon^\rho \nabla_\rho v^\mu \ .$$

  – the first equality says that the vector $\tilde{v}$ we get by parallel transport at $x+\epsilon$ is the value of the vector field at that point anyway, and the second is the definition of ∇. The change in the angle between v and w is proportional to

  $$\delta\left(v \cdot w\right) = \epsilon^\rho \partial_\rho \left(v \cdot w\right).$$

  But:

  $$\begin{aligned} \partial_\rho(v\cdot w) &= \partial_\rho\left(g_{\mu\nu} v^\mu w^\nu\right) \\ &\overset{\text{It's a scalar}}{=} \nabla_\rho\left(g_{\mu\nu} v^\mu w^\nu\right) \\ &\overset{\text{Liebniz}}{=} \underbrace{\left(\nabla_\rho g_{\mu\nu}\right)}_{=0} v^\mu w^\nu + g_{\mu\nu}\left(\nabla_\rho v^\mu\right) w^\nu + g_{\mu\nu} v^\mu \left(\nabla_\rho w^\nu\right) \end{aligned} \qquad (64)$$

  So we see that $\delta\left(v \cdot w\right) = \epsilon^\rho \partial_\rho \left(v \cdot w\right) = 0$.

Notice also that since the above are both differential conditions at a point, any confusions I might have created about finite parallel transport drop out of our discussion.

Existence and uniqueness of the Christoffel connection. This is a fancy way of saying that we can find an explicit expression for the Christoffel connection Γ in terms of g and its derivatives. Consider the following equations (the same equation three times, with cyclically-relabelled indices)

$$
\begin{array}{rcll}
0 &=& \nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\lambda_{\rho\mu} g_{\lambda\nu} - \Gamma^\lambda_{\rho\nu} g_{\mu\lambda} & \cdot (+1) \\
0 &=& \nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\lambda_{\mu\nu} g_{\lambda\rho} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda} & \cdot (-1) \\
0 &=& \nabla_\nu g_{\rho\mu} = \partial_\nu g_{\rho\mu} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu} - \Gamma^\lambda_{\nu\mu} g_{\rho\lambda} & \cdot (-1) \\
\hline
0 &=& \partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} + 2\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} \ . &
\end{array}
\tag{65}
$$

Using torsion-freedom ($\Gamma^{\cdot}_{\mu\nu} = +\Gamma^{\cdot}_{\nu\mu}$) and symmetry of $g_{\mu\nu} = g_{\nu\mu}$, the like-colored terms eat each other when we add together the equations with the signs indicated at right. Contracting with the inverse metric $g^{\rho\lambda}$, we can use this to find:

$$
\boxed{\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} \left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right)} \ . \tag{66}
$$

5.4.2 Curvature from the Christoffel connection

With the Chistoffel connection, the curvature satisfies one more condition:

$$
(4) \qquad R_{\mu\nu\rho\sigma} \equiv R_{\mu\nu\rho}{}^{\lambda} g_{\sigma\lambda} = -R_{\mu\nu\sigma\rho}.
$$

Notice that it is the *last* index on $R_{\mu\nu\rho\sigma}$ which came from lowering the upper index on $R_{\mu\nu\rho}{}^{\lambda}$. Proof: Recall that curvature is defined by acting with the commutator of covariant derivatives on a tensor $\nabla_{[\mu}\nabla_{\nu]}\omega_\rho = R^\sigma_{\mu\nu\rho}\omega_\sigma$. But the fact that the derivatives of the tensor drop out means that we can use any tensor, including g:

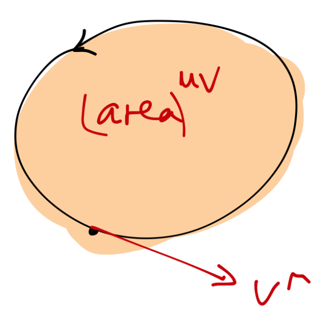
$$
0 = \nabla_{[\mu}\nabla_{\nu]} g_{\rho\sigma} = R_{\mu\nu\rho\sigma} + R_{\mu\nu\sigma\rho}
$$

But the LHS here is zero for the metric-compatible connection.

This property has a geometric interpretation: it says that parallel transport along a closed path (this linear operator is called the *holonomy*)

$$
\Delta v^\rho = \left(\pm \, (\text{area})^{\mu\nu} \, R_{\mu\nu}\right)_\sigma{}^{\rho} v^\sigma
$$

acts upon a vector by not an arbitrary linear operation but by an element of $SO(4)$ or $SO(3,1)$ – the generators are antisymmetric.



Notice that properties (4) and (3) imply

$$(5) \qquad R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu}.$$

While we're at it we can define some other tensors that we can derive from the Riemann curvature (the one with 4 indices):

$$\boxed{\text{Ricci tensor:} \qquad R_{\mu\nu} \equiv R_{\mu\rho\nu}{}^{\rho}}.$$

(Notice that this contraction is a weird one to consider from the point of view of the parallel transport interpretation, but it's still a tensor.)

$$\boxed{\text{Ricci scalar curvature:} \qquad R \equiv R^{\mu}{}_{\mu} = g^{\mu\nu} R_{\mu\nu}}.$$

Notice that the four tensors called R are distinguished by how many indices they have.

$$\boxed{\text{Einstein tensor:} \qquad G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R}.$$

The reason to care about this combination is that it is "covariantly conserved" for any $g_{\mu\nu}$:

$$0 = \nabla_\mu G_\nu{}^\mu = \nabla^\mu G_{\nu\mu}.$$

Proof:

Begin from Bianchi (not yet proved): $\qquad 0 = \nabla_{[\mu} R_{\nu\rho]\sigma}{}^{\delta}$

contract with δ^μ_δ:

$$0 = \nabla_{[\mu} R_{\nu\rho]\sigma}{}^{\mu} = 2 \left(\nabla_\mu R_{\nu\rho\sigma}{}^\mu + \nabla_\nu \underbrace{R_{\rho\mu\sigma}{}^\mu}_{=R_{\rho\sigma}} - \nabla_\rho \underbrace{R_{\nu\mu\sigma}{}^\mu}_{=R_{\nu\sigma}} \right)$$

and now contract with $g^{\nu\sigma}$:

$$0 = \nabla_\mu R_\rho{}^\mu + \nabla_\nu R_\rho{}^\nu - \nabla_\rho R = 2 \nabla_\mu G^\mu{}_\rho \quad .$$

(This is necessary if we're going to set $G_{\mu\nu} \propto T_{\mu\nu}$ with $T_{\mu\nu}$ a (covariantly) conserved stress tensor.)

WRAP-UP OF COVARIANT DERIVATIVE DISCUSSION

So the covariant derivative defines a notion of *parallel* for two vectors at different points.

Properties of ∇.

1. ∇ is a derivation.
2. The metric is covariantly constant: $\nabla_\rho g_{\mu\nu} = 0$.

A bit of notation: occasionally it is useful to use the following notation for derivatives of a tensor (*e.g.*) X_α:

$$\partial_\mu X_\alpha \equiv X_{\alpha,\mu}, \quad \nabla_\mu X_\alpha \equiv X_{\alpha;\mu}.$$

6 Geodesics

We've already discussed several warmup problems for the study of geodesics in curved spacetime (flat space, Minkowski spacetime, and in fact on problem set 4 you already studied the general case), so this should be easy. Consider a parametrized path in a manifold:

$$\begin{array}{rcl} x : \mathbb{R} & \to & M \\ \mathfrak{s} & \mapsto & x^\mu(\mathfrak{s}) \end{array} \tag{67}$$

and let $\dot{x} \equiv \frac{dx}{d\mathfrak{s}}(\mathfrak{s})$. The

$$\text{length of the curve, } \ell = \int_{\mathfrak{s}_0}^{\mathfrak{s}_1} \mathrm{d}\mathfrak{s} \sqrt{|g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu|}$$

is a geometric quantity – it is independent of the choice of coordinates $x^\mu \to \tilde{x}^\mu(x)$ *and* of the parametrization: $\mathfrak{s} \to \tilde{\mathfrak{s}}(\mathfrak{s})$ (assuming $\tilde{\mathfrak{s}}'(\mathfrak{s}) > 0$).

Given a curve x, this produces a number – it is a functional $\ell[x]$. A *geodesic* is a curve for which $\ell[x]$ is stationary under small variations of the curve. This variation (fixing the locations of the endpoints $\delta x^\mu(\mathfrak{s}_0) = 0 = \delta x^\mu(\mathfrak{s}_1)$) is (assuming for the moment a spacelike curve to avoid some annoying absolute values)

$$0 = \frac{\partial S}{\partial x^\mu(\mathfrak{s})} \propto e(\mathfrak{s}) \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{e(\mathfrak{s})} g_{\mu\nu}(x(\mathfrak{s}))\dot{x}^\nu\right) - \frac{1}{2}\left(\partial_\mu g_{\rho\sigma}\right)\dot{x}^\rho \dot{x}^\sigma$$

where $e \equiv \sqrt{g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}$. (Recall that this is the same quantity from problem set 4.) This doesn't look covariant, but we can make it more manifestly so using $\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}} g_{\mu\nu}(x(\mathfrak{s})) = \partial_\rho g_{\mu\nu}\dot{x}^\rho$ and

$$\Gamma_\mu{}^\rho{}_\nu = \frac{1}{2} g^{\rho\sigma}\left(g_{\mu\sigma,\nu} + g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}\right)$$

in terms of which

$$g_{\mu\nu}\Gamma_\rho{}^\mu{}_\sigma \dot{x}^\rho \dot{x}^\sigma = \frac{1}{2}\left(2\partial_\rho g_{\nu\sigma}\dot{x}^\rho \dot{x}^\sigma - \partial_\nu g_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma\right)$$

and finally the geodesic equation is

$$e\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{e}\dot{x}^\mu\right) + \Gamma_\rho{}^\mu{}_\sigma \dot{x}^\rho \dot{x}^\sigma = 0.$$

The geodesic equation is formal and useless, but true. It is (I think) always better to plug in the form of the metric to ds rather than beginning with the geodesic equation. We can improve its utility slightly by taking advantage as before of its reparametrization invariance to make $e(s) = \sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}}$ independent of s. (Notice that now I am using s and not $\mathfrak{s}$ since it is not arbitrary. We can make such a choice since under

$$\mathfrak{s} \to s(\mathfrak{s}), e(\mathfrak{s}) \to e(s) = \sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}} = \sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}\frac{\mathrm{d}x^\nu}{\mathrm{d}\mathfrak{s}}}\frac{\mathrm{d}\mathfrak{s}}{\mathrm{d}s}$$

so we just need to pick $\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s} = \frac{1}{e(\tilde{s})}$. As before, such a parametrization is called an *affine parameter* since any other parameter satisfying this condition is $\tilde{s} = as+b$, with a, b constants (this is an "affine transformation" of s). Notice that the affine parameter is the length so far: $\ell = \int_{s_0}^{s_1} e(s)\mathrm{d}s = (\text{const})(s_1 - s_0)$. Note that for a null curve, 'const' here means 'zero'.

So in an affine parametrization, the geodesic equation is

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\right) + \Gamma_\rho{}^\mu{}_\sigma \frac{\mathrm{d}x^\rho}{\mathrm{d}s}\frac{\mathrm{d}x^\sigma}{\mathrm{d}s} = 0} \quad \text{(Affine geodesic equation).}$$

[25]

IVP: Consider for a moment the geodesic equation as a 2d order nonlinear ODE with independent variable s and dependent variables x^μ, just like the Newton equation $m\frac{\mathrm{d}^2x^\mu}{\mathrm{d}t^2} = F^\mu(x) \overset{e.g.}{=} -\eta^{\mu\nu}\partial_\nu V(x)$. As in that case, if we specify $(x^\mu, \frac{\mathrm{d}x^\mu}{\mathrm{d}s}) \equiv (x_0^\mu, v_0^\mu) \in (M, T_xM)$ at $s = s_0$, then $\exists!$ solution in a neighborhood $s > s_0$. This geodesic may not be extendible to $s = \infty$ – for example it might hit some singularity or boundary. If it *is* so extendible, then the geodesic is called *complete*. If $\forall(x^\mu, v^\mu)$, the resulting geodesic is complete, then the manifold M is *geodesically complete*. [End of Lecture 10]

Some examples:

1. The unit two-sphere S^2: $\mathrm{d}s^2_{S^2} = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2$

$$\implies e = \sqrt{g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} = \sqrt{\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2}.$$

Geodesics satisfy

$$0 = \frac{\delta\ell}{\delta\theta} = -e\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{e}\underbrace{g_{\theta\theta}}_{=1}\dot{\theta}\right) + \underbrace{\frac{1}{2}\partial_\theta \sin^2\theta}_{=\sin\theta\cos\theta}\dot{\varphi}^2$$

$$0 = \frac{\delta\ell}{\delta\varphi} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{e}g_{\varphi\varphi}\dot{\varphi}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\sin^2\theta\dot{\varphi}}{e}\right).$$

With the affine parametrization, we can set $e = 1$ (this is a spacelike manifold, so the constant can't be zero)

$$\implies 0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(\sin^2\theta\dot{\varphi}\right)$$

the quantity in parenthesis is conserved, an integral over the motion. This happened because the metric was independent of φ – there is an *isometry* generated by the vector

[25] A point I under-emphasized in lecture is that this choice of affine parametrization $\dot{e} = 0$ is consistent with the geodesic equation, in the sense that if we make this choice at one value of s, the geodesic equation implies that it continues to be true.

field ∂_φ. (The conserved quantity is the z-component of the angular momentum. In fact there are two more conserved charges from the two other independent rotations, as you'll see on problem set 5. Three conserved charges for a system with two dofs means the motion is completely integrable.)

2. Schwarzchild metric:

$$\mathrm{d}s^2_{\mathrm{Sch}} = -f(r)\mathrm{d}t^2 + \frac{1}{f(r)}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2\right), \quad \text{with} \quad f(r) \equiv 1 - \frac{r_0}{r}$$

and r_0 is a constant. If $r_0 \to 0$, this is Minkowski space. Notice that the 2-sphere from the previous example is what you get at fixed values of r and t, so the metric has the above spherical symmetry. (Foreshadowing remarks: This solves Einstein's equations with $T_{\mu\nu} = 0$ for $r > r_0$; with $r_0 = \frac{2GM}{c^2}$, it describes a spherical object of mass M centered at $r = 0$. You can see that some funny things happen at $r = r_0$ and $r = 0$; those are topics for later and much later, respectively.)

For now, let's consider a particle moving in the $\theta = \pi/2$ plane (in fact this is WLOG by 3d rotational symmetry of the metric).

$$\theta = \frac{\pi}{2} \implies \qquad \ell[t(\mathfrak{s}), r(\mathfrak{s}), \varphi(\mathfrak{s})] = \int_{\mathfrak{s}_0}^{\mathfrak{s}_1} \mathrm{d}\mathfrak{s} \sqrt{f(r(\mathfrak{s}))\dot{t}^2 - \frac{\dot{r}(\mathfrak{s})^2}{f(r(\mathfrak{s}))} - r^2\dot{\varphi}^2}$$

The metric is independent of t and of φ: ∂_t and ∂_φ generate isometries (time translation invariance and azimuthal rotation invariance, respectively). The t and φ variations are going to give us conservation laws[26]:

$$0 = \frac{\delta \ell}{\delta t} \implies f(r)\dot{t} \equiv \epsilon = \text{const, `energy'}$$

$$0 = \frac{\delta \ell}{\delta \varphi} \implies r^2\dot{\varphi} \equiv L = \text{const, angular momentum.}$$

In affine coords

$$-g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = +f(r)\dot{t}^2 - \frac{\dot{r}^2}{f(r)} - r^2\dot{\varphi}^2 = \kappa = \text{const}$$

This is three integrals of motion (ϵ, L, κ) for 3 dofs (t, r, φ), so we can integrate. Eliminate t, φ in κ:

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{r_0}{r}\right)\left(\frac{L^2}{r^2} + \kappa\right) = \frac{1}{2}\epsilon^2$$

[26] In general, if $g_{\mu\nu}$ is independent of x^μ for some μ, then

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(g_{\mu\nu}\dot{x}^\nu\right) - \frac{1}{2}\partial_\mu g_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma = \frac{\mathrm{d}}{\mathrm{d}s}\left(g_{\mu\nu}\dot{x}^\nu\right)$$

so the momentum $p_\mu = g_{\mu\nu}\dot{x}^\nu$ is conserved. As you can see from the S^2 example, not every isometry need be so manifest – $\partial_\mu g_{\nu\rho} = 0$ in some coordinate system is a sufficient condition for ∂_μ to generate an isometry, but not necessary. It is also not a tensor equation. More generally the condition is Killing's equation, $\mathcal{L}_v g = 0$.

This is the Newtonian equation for the conserved energy $E = \frac{1}{2}\epsilon^2 + \text{const}$ with a potential which is

$$V(r) = \frac{1}{2}\left(1 - \frac{r_0}{r}\right)\left(\frac{L^2}{r} + \kappa\right) = \underbrace{\frac{1}{2}\kappa}_{const} \quad \underbrace{-\frac{\kappa r_0}{2r}}_{\text{Newtonian grav potential}} \quad + \quad \underbrace{\frac{1}{2}\frac{L^2}{r^2}}_{\text{centripetal potential}} \quad \underbrace{-\frac{r_0 L^2}{2r^3}}_{\text{new stuff!}}$$

The last term is new – a deviation from the Newtonian potential, and hence a deviation from elliptical orbits! This implies a precession of the perihelion (the point of closest approach), which was observed first for Mercury. If you can't wait to hear more about this, see Appendix 1 of Zee §VI.3, p. 371.

The rest of this section is somewhat miscellaneous: it is a discussion of various topics related to geodesics, which will improve our understanding of the notions of curvature and parallel transport introduced in the previous section, and move us toward contact with physics.

6.0.3 Newtonian limit

[Weinberg, page 77] Reality check: suppose (1) $|\frac{dx^i}{d\lambda}| \ll \frac{dx^0}{d\lambda}$ (slow particles) and (2) $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, |h_{\mu\nu}| \ll 1$ (in a weak field[27]). Let's look at the geodesic equation in this limit. Neglecting $\frac{d\vec{x}}{d\tau}$ in favor of $\frac{dt}{d\tau}$, the affine geodesic equation reduces to:

$$0 = \frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{dt}{d\tau}\right)^2 .$$

Further, let's assume (3) a time-independent (or weakly time-dependent) gravitational field, $\partial_t g_{\mu\nu} = 0$, so

$$\Gamma^\mu_{00} = -\frac{1}{2}g^{\mu\nu}\partial_\nu g_{00} \overset{\text{weak field}}{=} -\frac{1}{2}\eta^{\mu\nu}\partial_\nu h_{00} + \mathcal{O}(h^2)$$

where we used

$$g^{\mu\nu} \simeq \eta^{\mu\nu} - \underbrace{h^{\mu\nu}}_{\equiv \eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}} + \mathcal{O}(h^2).$$

Using again $\dot{g} \sim 0$, $\Gamma^0_{00} \propto \partial_t h_{00} = 0$ and we have

$$0 = \frac{d^2 x^i}{d\tau^2} - \frac{1}{2}\nabla^i h_{00}\left(\frac{dt}{d\tau}\right)^2, \quad \frac{d^2 t}{d\tau^2} = 0$$

– the second equation is the affine parametrization condition in this limit. Dividing the BHS by the constant $\frac{dt}{d\tau}$, we get

$$\frac{d^2\vec{x}}{dt^2} = \frac{1}{2}\vec{\nabla}h_{00} .$$

[27] You could ask me what I mean by $|h_{\mu\nu}| \ll 1$ in a coordinate invariant way. One thing to do to make a scalar is to contract the indices with the Minkowski metric $|h|_1 = \eta^{\mu\nu}h_{\mu\nu}$. Another is $|h|_2 = \sqrt{\eta^{\mu\nu}h_{\nu\rho}\eta^{\rho\sigma}h_{\sigma\mu}}$. We could keep going.

Comparing with Newton's law

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2} = -\vec{\nabla}\phi \ ,$$

the conclusion of this discussion is that the Newtonian potential is

$$\boxed{\phi = -\frac{1}{2}h_{00}.}$$

This confirms the relationship between r_0 and M that I claimed in the above discussion of Schwarzchild. Notice that this factor of 2 is the same one advertised above in (33).

6.0.4 Parallel transport along geodesics

[Zee page 330] A useful comment about the Christoffel connection and geodesics. Notice that both the geodesic equation and the covariant derivative of a vector involve these awful Γ symbols.

Definition: A vector field w is *parallel* with respect to a (parametrized) geodesic $x(\mathfrak{s})$ with tangent vectors $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}$ if

$$\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}}g(v,w) = 0$$

that is, the vf makes the same angle with the tangent vector at each point.

This says

$$\frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}}g(v,w) = \frac{\mathrm{d}}{\mathrm{d}\mathfrak{s}}\left(g_{\mu\nu}v^\mu w^\nu\right) = g_{\mu\nu,\alpha}v^\alpha v^\mu w^\nu + g_{\mu\nu}\frac{\mathrm{d}v^\mu}{\mathrm{d}\mathfrak{s}}w^\nu + g_{\mu\nu}v^\mu\frac{\mathrm{d}w^\nu}{\mathrm{d}\mathfrak{s}} \quad . \tag{68}$$

Let us assume that the parameter $\mathfrak{s}$ is affine. In that case we have

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}\mathfrak{s}} + \Gamma^\mu_{\gamma\delta}v^\gamma v^\delta = 0. \tag{69}$$

(otherwise some extra terms). The Christoffel connection is

$$\Gamma^\mu_{\gamma\delta} = \frac{1}{2}g^{\mu\sigma}\left(g_{\sigma\delta,\gamma} + g_{\gamma\sigma,\delta} - g_{\gamma\delta,\sigma}\right).$$

So we can find a differential condition on the parallel vector field w:

$$D_{\mathfrak{s}}w^\mu \equiv \frac{\mathrm{d}w^\mu}{\mathrm{d}\mathfrak{s}} + \Gamma^\mu_{\beta\gamma}v^\beta w^\gamma = 0 \ \text{ if } w \text{ is parallel along } x^\mu(\mathfrak{s}), v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}. \tag{70}$$

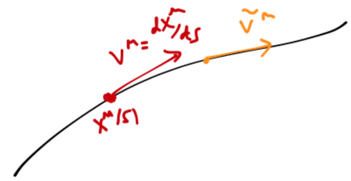
(It's a little bit of work to check that this implies (68).) So the (affine-parametrized) geodesic equation (69) can be summarized as: a (affine-parametrized) geodesic parallel transports its own tangent vectors, $D_s\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = 0$. [28]

[28] Lest this notation cause confusion and fear: on any vector w, $D_s w^\mu \equiv \frac{\mathrm{d}x^\nu}{\mathrm{d}s}\nabla_\nu w^\mu$ is just the 'directional covariant derivative'. The affine geodesic equation can then be written variously as:

$$D_s T^\mu = 0, \qquad T^\nu\nabla_\nu T^\mu = 0$$

with $T^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}s}$.

The parallel condition is easy to express in terms of the (Christoffel, *i.e.* metric-compatible) covariant derivative: Notice that



$$\frac{\mathrm{d}w^\mu}{\mathrm{d}\mathfrak{s}} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}\partial_\nu w^\mu = v^\nu \partial_\nu w^\mu$$

by the chain rule. So the parallel condition (70) is

$$0 = \frac{\mathrm{d}w^\mu}{\mathrm{d}\mathfrak{s}} + \Gamma^\mu_{\beta\gamma} v^\beta w^\gamma = \left(\partial_\beta w^\mu + \Gamma^\mu_{\beta\gamma} w^\gamma\right) v^\beta = v^\beta \nabla_\beta w^\mu \ .$$

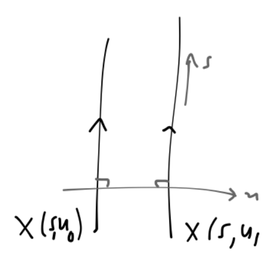
Comments:

- Wald begins his discussion of parallel transport with this equation. It's a useful exercise to follow the chain of relations in both directions. It is in fact that same definition as we gave above, which was:
$$\tilde{v}^\mu = v^\mu - \epsilon^\rho \Gamma_\rho{}^\mu{}_\nu v^\nu$$
since we should set the parameter indicating the direction of transport equal to $\frac{\epsilon^\mu}{\epsilon} = v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\mathfrak{s}}$.
- Note that parallel transport along a geodesic with tangent vector v is *not* the same as the map induced by the flow ϕ_v! For one thing, the latter doesn't depend on the metric.

[End of Lecture 11]

6.0.5 Geodesic deviation and tidal forces

The tabloid-headline title for this subsubsection is "WHEN PARALLEL LINES MEET!" [Wald pp. 46-47] Consider a family of geodesics, parametrized by a family-parameter u, $x^\mu(s,u)$.
For each fixed u,



$$\partial_s^2 x^\mu + \Gamma^\mu_{\rho\sigma} \partial_s x^\rho \partial_s x^\sigma = 0 \ .$$

It is possible to choose the parameters so that the variation in u is orthogonal to the variation in s: $g_{\mu\nu}\partial_s x^\mu \partial_u x^\nu$. Consider two 'nearby' geodesics, $x^\mu(\cdot, u), x^\mu(\cdot, u+\delta u)$. It will be useful to call their relative separation $\xi^\mu \equiv \partial_u x^\mu$.

The distance between nearby geodesics $= \delta u \sqrt{g_{\mu\nu}\partial_u x^\mu \partial_u x^\nu} = \delta u \|\, \xi \,\|$.

Let's compute their *acceleration* toward each other. The vector field ξ^μ is not parallel transported in general; that is, their relative velocity is

$$v^\mu \equiv D_s \xi^\mu = \partial_s \xi^\mu + \Gamma^\mu_{\nu\rho} \partial_s x^\nu \xi^\rho \neq 0$$

(we introduced the notation D_s in (70) – recall that it just means $D_s w^\mu \equiv \frac{dx^\nu}{ds} \nabla_\nu w^\mu$). More interesting is

$$a^\mu \equiv D_s v^\mu = D_s D_s \xi^\mu = \partial_s v^\mu + \Gamma^\mu_{\rho\sigma} \partial_s x^\nu v^\sigma = -R_{\nu\sigma\rho}{}^\mu \partial_u x^\nu \partial_s x^\sigma \partial_s x^\rho .$$

For a derivation of the last step, see Wald p 47. You can see how there are two covariant derivatives involved; the derivation uses only the geodesic equation $T^\mu \nabla_\mu T^\nu = 0$, ($T^\mu \equiv \frac{dx^\mu}{ds}$), and the fact that the Lie bracket $[T, \xi] = 0$.[29]

So this is another interpretation for the curvature: only if the curvature vanishes do initially-parallel geodesics remain parallel; a nonzero curvature measures the extent to which nearby geodesics accelerate towards each other.

6.0.6 Riemann normal coordinates

[Zee page 343-344] The EEP guarantees that we can choose coordinates (*i.e.* go to an appropriate accelerating frame) which get rid of the gravitational force at any given point p (though not necessarily at any other points with the same choice!). From the previous discussion, we see that this means that the Christoffel symbols vanish (at the chosen point) in those coordinates: $\Gamma^\rho_{\mu\nu}(p) = 0$.

I suppose that we must show this is actually possible mathematically; the fact that it's a principle of physics means that it must be, if we are using the right mathematics! In any case it is useful to know how to do it. A more formal definition of Riemann normal coordinates $(\xi^1, ...\xi^n)$ in a neighborhood of a point $p \in M$ is as follows. Given a tangent vector $\xi \in T_p M$, find an affine geodesic $x_\xi(s)$ with the initial condition

$$x_\xi(s = 0) = p, \quad \frac{dx^\mu_\xi}{ds}(s = 0) = \xi^\mu \ .$$

Then define the *exponential map*, $\exp : T_p M \to M$ as

$$\exp(\xi) \equiv x_\xi(s = 1) \in M \quad .$$

[29] For your convenience, I copy Wald (3.3.18) here:

$$\begin{aligned}
a^\mu &= T^c \nabla_c \left(T^b \nabla_b \xi^\mu \right) \overset{[T,\xi]=0}{=} T^c \nabla_c \left(\xi^b \nabla_b T^\mu \right) \\
&\overset{\text{product rule}}{=} (T^c \nabla_c \xi^b) \nabla_b T^\mu + T^c \xi^b \nabla_c \nabla_b T^\mu \\
&\overset{[T,\xi]=0,\text{def of } R}{=} (\xi^c \nabla_c T^b)(\nabla_b T^\mu) + \xi^b T^c \nabla_b \nabla_c T^\mu - R_{cbd}{}^\mu \xi^b T^c T^d \\
&\overset{\text{product rule}}{=} \xi^c \nabla_c (T^b \nabla_b T^\mu) - R_{cbd}{}^\mu \xi^b T^c T^d \overset{\text{geodesic eqn}}{=} -R_{cbd}{}^\mu \xi^b T^c T^d
\end{aligned} \tag{71}$$

(If the manifold M is geodesically complete, the map exp is defined for any tangent vector ξ. Otherwise, we may have to limit ourselves to a subset $\mathcal{V}_p \subset T_pM$ on which $\exp(\xi)$ is nonsingular. Since $T_pM \simeq \mathbb{R}^n$, its subspace $\mathcal{V}$ is an open subset of $\mathbb{R}^n$.) We can then use a set of linearly independent vectors $\{\xi^{(i)} \in \mathcal{V}_p\}_{i=1}^n$ to produce coordinates in the neighborhood $\exp(\mathcal{V}_p)$ of p which is the image of $\mathcal{V}_p$.

On problem set 6 you will show that $\Gamma^k_{ij} = 0$ in this coordinate system. The basic idea is that straight lines in the tangent space (which is the coordinate space here) are mapped to geodesics by these coordinates. But the geodesic equation differs from the equation for a straight line by a term proportional to Γ.

Let's Taylor expand the metric about such a point (p is $x = 0$) in such coordinates:

$$g_{\tau\mu}(x) = \eta_{\tau\mu} + \underbrace{g_{\tau\mu,\nu}}_{=0} x^\nu + \frac{1}{2} g_{\tau\mu,\lambda\sigma} x^\lambda x^\sigma + ...$$

In the first term we have done a (constant) general linear transformation on our normal coordinates to set $g_{\tau\mu}(0) = \eta_{\tau\mu}$. Notice that $g_{\tau\mu,\lambda\sigma}$ is symmetric under interchange of $\tau \leftrightarrow \mu$ or of $\lambda \leftrightarrow \sigma$. Let's plug this into our expression for the Christoffel symbols (66) (that we got by demanding metric-compatibility):

$$\Gamma^\lambda_{\rho\nu} = \frac{1}{2} g^{\lambda\tau} \left(g_{\tau\nu,\rho} + g_{\tau\rho,\nu} - g_{\rho\nu,\tau} \right) = \underbrace{\eta^{\lambda\tau} \left(g_{\tau\nu,\mu\rho} + g_{\tau\rho,\mu\nu} - g_{\rho\nu,\mu\tau} \right)} x^\mu + ... \tag{72}$$

Plugging this into the expression for the R'n curvature we have[30]

$$\begin{aligned} R_{\mu\nu\rho}{}^\lambda(p) &= -\partial_\mu \Gamma^\lambda_{\nu\rho} - (\mu \leftrightarrow \nu) + \underbrace{\Gamma\Gamma}_{=0 \text{ at } p} \\ &= -\eta^{\lambda\tau} \left(g_{\tau\nu,\mu\rho} + g_{\tau\rho,\mu\nu} - g_{\rho\nu,\mu\tau} \right) + \mathcal{O}(x) - (\mu \leftrightarrow \nu) \end{aligned} \tag{73}$$

– at leading order in x, the ∂_μ pulls out the (underbraced) coefficient of x^μ in (72). The middle term is symmetric in $\mu\nu$ drops out of the sum. For one thing, it makes manifest the antisymmetry in the last two indices: $R_{\mu\nu\rho\sigma} = -R_{\mu\nu\sigma\rho}$:

$$R_{\mu\nu\rho\sigma} = -g_{\sigma\nu,\mu\rho} + g_{\rho\nu,\mu\sigma} - (\mu \leftrightarrow \nu) \quad .$$

This is an explicit (tensor) expression for the R'n tensor which can be used to demonstrate various of its properties. (Although we've derived the equation in specially-chosen coordinates, it is a tensor equation, so properties that we find in this coordinate system are just true.)

[Zee page 392, but beware that he puts the lowered index of R first!] Another is the Bianchi identity, which we can see as follows. In general,

$$\nabla_\nu R_{\mu\sigma\lambda}{}^\rho = -\nabla_\nu \left(\partial_\mu \Gamma^\rho_{\sigma\lambda} + \Gamma^\rho_{\kappa\sigma}\Gamma^\kappa_{\mu\lambda} - (\mu \leftrightarrow \sigma) \right) .$$

[30] Beware Zee's notation for the order of the indices on the Riemann tensor. It is different from ours!

In R'n normal coords at p this simplifies to ($R_{\mu\sigma\lambda\rho} \equiv R_{\mu\sigma\lambda}{}^{\alpha} g_{\alpha\rho}$!)

$$\nabla_\nu R_{\mu\sigma\lambda\rho}|_p = \partial_\nu R_{\mu\sigma\lambda\rho}|_p = \left(\partial_\nu\partial_\sigma\Gamma_{\rho|\mu\lambda} - \partial_\nu\partial_\lambda\Gamma_{\rho|\mu\sigma}\right)|_p$$

(here I have denoted $\partial_\nu\partial_\lambda\Gamma_{\rho|\mu\sigma} \equiv g_{\rho\alpha}\Gamma^{\alpha}_{\mu\sigma}$ to make clear which index is the lowered one). Now if we cyclically permute the $\nu\sigma\lambda$ indices and add the three equations (like we did to derive the expression for Γ in terms of g, but all with a + this time) we get

$$\begin{array}{rcc}
\nabla_\nu R_{\mu\sigma\lambda\rho} & = & \left(\partial_\nu\partial_\sigma\Gamma_{\rho|\mu\lambda} - \partial_\nu\partial_\lambda\Gamma_{\rho|\mu\sigma}\right) \\
\nabla_\sigma R_{\mu\lambda\nu\rho} & = & \left(\partial_\sigma\partial_\lambda\Gamma_{\rho|\mu\nu} - \partial_\sigma\partial_\nu\Gamma_{\rho|\mu\lambda}\right) \\
\nabla_\lambda R_{\mu\nu\sigma\rho} & = & \left(\partial_\lambda\partial_\nu\Gamma_{\rho|\mu\sigma} - \partial_\lambda\partial_\sigma\Gamma_{\rho|\mu\nu}\right) \\
\hline
\nabla_\nu R_{\mu\sigma\lambda\rho} + \nabla_\sigma R_{\mu\lambda\nu\rho} + \nabla_\lambda R_{\mu\nu\sigma\rho} & = & 0
\end{array}$$

Like-colored terms cancel. This is the Bianchi identity. The version we wrote earlier

$$\nabla_{[\nu} R_{\mu\sigma]\lambda}{}^{\rho}$$

is related to this one by property (5) $R_{\mu\sigma\lambda\rho} = R_{\lambda\rho\mu\sigma}$.

7 Stress tensors from the metric variation

Matter in curved space. We've already discussed the strategy for taking an action for a classical field in Minkowski space (*e.g.* $S = \int \mathrm{d}^4x \mathcal{L}(\phi, \partial_\mu\phi) = -\int \mathrm{d}^4x \left(\frac{1}{2}\partial_\mu\phi\partial_\nu\phi\eta_{\mu\nu} + V(\phi)\right)$) and turning it into an action in a curved metric $g_{\mu\nu}(x)$:

- Contract tensor indices with $g^{\mu\nu}$ instead of $\eta^{\mu\nu}$.
- Use the covariant measure $\mathrm{d}^4x \to \sqrt{g}\mathrm{d}^4x$.

This will produce an action for ϕ which is independent of coordinate choice, which reduces to the flat space action when $g \to \eta$. Is it unique? No: for example we could add

$$S_R \equiv \int \mathrm{d}^4x\sqrt{g}\left(\xi R\phi^2 + \Upsilon R^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + ...\right)$$

where R is the Ricci scalar curvature... There are many terms, and the equivalence principle doesn't decide for us the values of the coefficients. The dynamics in flat space does *not* uniquely determine the curved-space dynamics. The only consolation I can give you is that by dimensional analysis nearly all of them are suppressed by some mass scale which seems to be quite large.

Stress-energy tensor in curved space. Earlier, we applied the Noether method to translation symmetry to find

$$T^{\mu}_{\text{flat}\,\nu} = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi + \delta^\mu_\nu\mathcal{L}$$

and in flat space $\partial_\mu T^\mu_\nu = 0$ by the EoM. In curved space, we'll find that:

$$T^\mu_{\text{curved}\,\nu} = \frac{1}{\sqrt{g}} \left(T^\mu_{\text{flat}\,\nu} + (\text{improvements})^\mu{}_\nu \right).$$

Using the equations of motion from the curved space action, the conservation law is *covariant* conservation:

$$\nabla_\mu T^\mu_\nu = 0 \ .$$

Here is the better definition I promised (automatically symmetric, automatically has the symmetries of S, differs from the old one only by improvement terms):

$$T^{\mu\nu}(x) \equiv \frac{2}{\sqrt{g}} \frac{\delta S_{\text{matter}}}{\delta g_{\mu\nu}(x)} \quad .$$

Useful facts about derivatives with respect to elements of a matrix (for the next bit we forget that g depends on space):

$$\frac{\partial}{\partial g_{\mu\nu}} \det g = (\det g) g^{\mu\nu} \tag{74}$$

$$\frac{\partial}{\partial g_{\rho\sigma}} g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \tag{75}$$

The second equation, (75), follows directly from varying $g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma$. Note that we are defining

$$\frac{\partial g_{ab}}{\partial g_{cd}} = \delta^c_a \delta^d_b$$

– the fact that the metric is symmetric does not come into it.

Two proofs of (74):

- $$\det g = \frac{1}{4!} \epsilon^{abcd} \epsilon^{mnpq} g_{am} g_{bn} g_{cp} g_{dq}$$

  $$\implies \delta \det g = \underbrace{\frac{4}{4!} \epsilon^{abcd} \epsilon^{mnpq} g_{am} g_{bn} g_{cp}}_{= g^{dq} \det g} \delta g_{dq}$$

  where the underbraced equation follows from the cofactor expression for the inverse:

  $$(A^{-1})_{ij} = (-1)^{i+j} \frac{1}{\det A} \det A(ij)'$$

  where $A(ij)'$ is the matrix obtained by removing the jth row and the ith column of A.

- For a symmetric matrix, (74) can be seen more usefully as follows, using the identity

$$\det g = \exp \operatorname{tr} \log g$$

What's the log of a (symmetric) matrix? One way to define it is to Taylor expand about the identity and use $\log(1+x) = x - \frac{x^2}{2} + ...$ So if $g_{ab} = \eta_{ab} + h_{ab}$ then

$$(\log g)_{ab} = h_{ab} - \frac{1}{2} h_{ac} h^c_b + \frac{1}{3} ... \ .$$

A more robust understanding is as follows: A symmetric matrix can be diagonalized

$$g = U \begin{pmatrix} \lambda_1 & 0 & \cdots \\ 0 & \lambda_2 & \cdots \\ 0 & 0 & \ddots \end{pmatrix} U^{-1}$$

with real eigenvalues λ_i. The log is then just

$$\log g = U \begin{pmatrix} \log \lambda_1 & 0 & \cdots \\ 0 & \log \lambda_2 & \cdots \\ 0 & 0 & \ddots \end{pmatrix} U^{-1}.$$

The det of g is the product of the eigenvalues: $\det g = \prod_i \lambda_i$, and $\operatorname{tr} \log g = \sum_i \log \lambda_i$. So

$$\delta \det g = e^{\operatorname{tr} \log g} \delta \left(\operatorname{tr} \log g \right) = \det g \operatorname{tr}_{ab} \left(\delta \log g \right)_{ab} = \det g \operatorname{tr}_{ab} \left(g^{-1} \delta g \right)_{ab} = \det g g^{ab} \delta g_{ab},$$

as claimed.

Putting back the position-dependence, we have

$$\frac{\delta g_{ab}(x)}{\delta g_{cd}(y)} = \delta^c_a \delta^d_b \delta^4(x-y)$$

[End of Lecture 12]

Let's apply the better definition of stress tensor to a scalar field with action

$$S_{\text{matter}}[\phi, g_{\mu\nu}] = -\frac{1}{2} \int \mathrm{d}^4 x \sqrt{g} \left(g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + m^2 \phi^2 \right)$$

– I have highlighted the metric dependence. In Lorentzian signature, our notation $\sqrt{g}$ hides the fact that

$$\sqrt{g} \equiv \sqrt{- \det g}$$

and therefore

$$\delta \sqrt{g} = -\frac{1}{2} \frac{1}{\sqrt{g}} \delta \det g = -\frac{1}{2\sqrt{g}} g^{ab} \delta g_{ab} \det g = \frac{1}{2} \sqrt{g} g^{ab} \delta g_{ab}.$$

Using this,

$$\frac{2}{\sqrt{g}}\frac{\delta S_{\text{matter}}[\phi, g]}{\delta g_{ab}} = 2\left(\frac{1}{2}g^{ac}g^{bd}\partial_c\phi\partial_d\phi - \frac{1}{4}g^{ab}\left(g^{cd}\partial_c\phi\partial_d\phi + m^2\phi\right)\right)$$

which cleans up to the same expression we found previously:

$$T_{ab} = \partial_a\phi\partial_b\phi - g_{ab}\mathcal{L}\ .$$

Similarly, you found on problem set 5 that E&M in curved space is described by the action

$$S_{EM}[A, g] = -\frac{1}{16\pi}\int \mathrm{d}^4x\sqrt{g}\,\underbrace{g^{ab}g^{cd}F_{ac}F_{bd}}_{\equiv F^2}$$

– notice that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ does not involve any metric dependence.

$$T^{ab}_{EM} = \frac{2}{\sqrt{g}}\frac{\delta}{\delta g_{ab}(x)}S_{EM}[A, g] = \frac{-2}{16\pi}\left(\frac{1}{2}g^{ab}F^2 - F^a{}_cF^{cb}\right) = \frac{1}{16\pi}\left(4F^a{}_cF^{cb} - g^{ab}F^2\right).$$

Finally, consider a point particle

$$S_{pp}[x, g] = -m\int \mathrm{d}\mathfrak{s}\sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}$$

$$T^{\mu\nu}_{pp} = \frac{2}{\sqrt{g}}\frac{\delta}{\delta g_{ab}(x)}S_{pp} = \int \mathrm{d}\mathfrak{s}\delta^4(x - x(\mathfrak{s}))\frac{m\dot{x}^a\dot{x}^b}{\sqrt{-g_{cd}\dot{x}^c\dot{x}^d}}$$

which is what we got before using the trick of coupling to EM and demanding conservation of the total stress tensor.

Persnickety comments:

- Notice that the generally-covariant delta function satisfies:

$$1 = \int \mathrm{d}^4y\sqrt{g(y)}\delta^4(y - x).$$

- Here's how to remember the 2 in the definition: we're only varying $g_{\mu\nu}$ and not its symmetric partner $g_{\nu\mu}$ when we take the derivative $\frac{\delta}{\delta g_{\mu\nu}}$. The 2 is meant to make up for this.

- A warning about the sign [Zee page 380]: Notice that the sign in (74) implies that

$$T_{\mu\nu} = \frac{2}{\sqrt{g}} \frac{\delta S_M}{\delta g_{\mu\nu}} \overset{\text{chain rule}}{=} -g_{\mu\rho} g_{\nu\sigma} \frac{2}{\sqrt{g}} \frac{\delta S_M}{\delta g^{\rho\sigma}}$$

  – it's not just the result of lowering the indices on $\frac{2}{\sqrt{g}} \frac{\delta S_M}{\delta g^{\rho\sigma}}$. The reason is that varying with respect to a component of the *inverse* metric is a different thing: $\delta g^{\rho\sigma} = -g^{\rho\mu} \delta g_{\mu\nu} g^{\nu\sigma}$– different things are held fixed when we do that.

8 Einstein's equation

8.1 Attempt at a 'correspondence principle'

So far we've been imagining that the metric $g_{\mu\nu}(x)$ is given to us by some higher power. Finally, we are ready to ask what determines the form of the metric, that is, we would like to find the equations of motion governing it. What I mean by a 'correspondence principle' is a guess based on the limit we know, namely the non-relativistic one.

In this limit, in a non-relativistic expansion about Minkowski spacetime, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, recall that we found above that the Newtonian potential was $\phi = -\frac{1}{2} h_{tt}$. But the Newtonian potential satisfies the Poisson equation, with the mass density as its source:

$$\nabla^2 \phi = 4\pi G_N \rho \quad . \tag{76}$$

Let's try to covariantize this equation.

$$\begin{aligned} \rho &\mapsto T_{00} \\ \nabla^2 h_{00} &\overset{??}{\mapsto} \nabla^2 g_{00} \text{ ?} \qquad \text{No: } \nabla_\mu g_{\rho\sigma} = 0. \\ \text{How about: } \nabla^2 h_{00} &\overset{??}{\mapsto} R_{00} \text{ ?} \end{aligned} \tag{77}$$

This might be good, since

$$R_{..} \sim \partial\Gamma + \Gamma\Gamma \sim \partial^2 g + \partial g \partial g$$

is a tensor which involves two derivatives of the metric. Demanding that (77) be the tt component of a tensor equation would then give:

$$R_{\mu\nu} \overset{??}{=} 8\pi G_N T_{\mu\nu}$$

A problem with this: the RHS is covariantly conserved, but the LHS is not! [Recall $\partial_\mu (\text{maxwell})^\mu$.]

The step (77) is ambiguous. With much hindsight we can see that we must use the (conveniently-named) Einstein tensor on the LHS instead,

$$G_{\mu\nu} \overset{?(\text{yes})}{=} 8\pi G_N T_{\mu\nu}$$

since we showed above that it satisfies

$$\nabla^\mu G_{\mu\nu} = 0 \qquad \text{(Bianchi)}.$$

So this equation is compatible with covariant conservation of stress energy tensor, and (we'll see later that it) reduces properly to the Newtonian limit[31].

8.2 Action principle

What coordinate invariant local functionals of the metric and its derivatives are there? By local functional again I mean we only get to integrate once. Let's not allow too many derivatives. Better: let's organize the list by how many derivatives there are. It's a short list:

$$S_{\text{gravity}}[g] = \int \mathrm{d}^D x \sqrt{g}\,(a + bR + \ldots)$$

Here I stopped at terms involving two derivatives. If we go to the next term (four derivatives), it's:

$$S_{\text{gravity}}[g] = \int \mathrm{d}^D x \sqrt{g}\left(a + bR + c_1 R^2 + c_2 R^{\mu\nu} R_{\mu\nu} + c_3 R^{\cdots\cdot} R_{\cdots\cdot} + \ldots\right).$$

[32]

Dimensional analysis:

$$[R] = [g^{-1}\partial\Gamma + g^{-1}\Gamma\Gamma]$$

The first term is $[R] = [g^{-1}\partial\left(g^{-1}\partial g\right)]$ $[a] \sim M^4, [b] \sim M^2, [c] \sim M^0$.

While we're at it, let's talk about units. You know that you can set $c = 1$ by measuring lengths in light-years. You can also set $\hbar = 1$ by measuring energy in units of frequency (or momenta in units of inverse length). You can set $G_N = 1$ by measuring masses in units of the Planck mass, M_{Planck}, which as you can see from the above dimensional analysis will go like $\frac{1}{\sqrt{b}} = \frac{1}{\sqrt{16\pi G_N}}$. That is, from $G_N, \hbar, c$ we can make a quantity with the dimensions of mass

$$M_{\text{Planck}} \equiv \sqrt{\frac{\hbar c}{16\pi G_N}}.$$

(People disagree about the factors of 2 and π in this definition.) So we no longer need to choose arbitrary systems of units for anything. $\hbar, c, G_N$ determined what units we should use. For better or worse, $M_{\text{Planck}} \sim 10^{19}$GeV, so natural units are not so useful in most laboratories.

[31] In four dimensions, the Einstein tensor is the only combination of the curvatures involving at most two derivatives of the metric with this property. In higher dimensions there are other terms with these properties.

[32] Note that we are just studying the dynamics of gravity here, and not thinking yet about including other matter, like the EM field. We'll see how to include that very naturally below. The cosmological term (proportional to a) is a special case here.

For no good reason, let's set $a = 0$. (FYI, it is called the 'cosmological constant'.) Further, ignore c because it is suppressed by a large mass scale – if we are interested in smooth spacetimes where $|R| \ll M_{\text{huge}}$, then this is a good idea.

So I'm suggesting that we study

$$S_{\text{EH}}[g] = \frac{1}{16\pi G_N} \int \mathrm{d}^D x \sqrt{g} R \quad . \tag{78}$$

This is called the Einstein-Hilbert action. I have renamed the constant $b \equiv \frac{1}{16\pi G_N}$ to be consistent with convention.

We are going to show that

$$\frac{\delta S_{\text{EH}}}{\delta g^{\mu\nu}(x)} = +\frac{1}{16\pi G_N} \sqrt{g} \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) = +\frac{1}{16\pi G_N} \sqrt{g} G_{\mu\nu} \quad . \tag{79}$$

First of all, we can see by easy considerations that the variation must be of the form

$$R_{\mu\nu} - \alpha g_{\mu\nu} R$$

for some constant α. [Zee page 347, argument that $\alpha \neq \frac{1}{4}$: Zee page 350]

To proceed more systematically, we will collect some useful intermediate steps. Things that are true:

1. The variation of the inverse of a matrix is related to the variation of the matrix by: $\delta g^{\mu\nu} = -g^{\mu\alpha} g^{\nu\beta} \delta_{\alpha\beta}$. If we include the spacetime dependence:
$$\frac{\delta g^{\mu\nu}(x)}{\delta g_{\alpha\beta}(y)} = -g^{\mu\alpha} g^{\nu\beta} \delta^D(x-y)$$
2. $\delta\sqrt{g} = \frac{1}{2}\sqrt{g} g^{\alpha\beta} \delta g_{\alpha\beta}$ Use $\log \det M = \operatorname{tr} \log M$.
$$\frac{\delta\sqrt{g}(x)}{\delta g_{\alpha\beta}(y)} = \frac{1}{2}\sqrt{g} g^{\alpha\beta} \delta^D(x-y)$$
3. $(\delta R_{\mu\nu})\, g^{\mu\nu} = w^\alpha_{;\alpha}$ for some vector field w^α.

Claim:

$$\delta R_{\mu\nu} = \nabla_\rho \left(\delta\Gamma^\rho_{\mu\nu} \right) - \nabla_\mu \left(\delta\Gamma^\rho_{\rho\nu} \right)$$

where

$$\delta\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\nabla_\mu \delta g_{\nu\sigma} + \nabla_\nu \delta g_{\mu\sigma} - \nabla_\sigma \delta g_{\mu\nu} \right)$$

where ∇ here is the covariant derivative with respect to the $g_{\mu\nu}$ about which we are varying. Notice by the way that the extra term in the coordinate transformation of Γ (the badness in (60)) is independent of the metric, so $\delta\Gamma$ *is* a tensor, and the above expression includes the $\Gamma\Gamma$ terms. (Alternatively, we can arrive at the same conclusion using normal coordinates at any point p, in which case, at p

$$R_{\mu\nu}|_p = \Gamma^\alpha_{\mu\nu,\alpha} - \Gamma^\alpha_{\mu\alpha,\nu},$$

and the following steps are the same.)

Notice that this is

$$\left(\delta R_{\mu\nu}\right) g^{\mu\nu} = w^\alpha_{;\alpha} = \frac{1}{\sqrt{g}}\partial_\alpha\left(\sqrt{g}w^\alpha\right) \tag{80}$$

, by our formula for the covariant divergence.

4. $\int \mathrm{d}^D x\sqrt{g}w^\alpha_{;\alpha} = 0$ if w has finite support. This is just IBP and the ordinary Gauss'/Stokes' Thm.

Putting these together:

$$\delta S_{EH} = \frac{1}{16\pi G_N}\int \mathrm{d}^4x\left(-\sqrt{g}\left(R^{\mu\nu} - \frac{1}{2}Rg^{\mu\nu}\right)\delta g_{\mu\nu} + \sqrt{g}g^{\mu\nu}\delta R_{\mu\nu}\right).$$

Neglecting the last term, which is a total derivative (by point 3 above), we have shown (79).

Comments:

- It is not an accident that $\frac{\delta S_{EH}}{\delta g_{\mu\nu}} \propto G^{\mu\nu}$ is covariantly conserved – it is the Noether current for translation invariance coming from the fact that S_{EH} doesn't depend on position. That is: it's the stress tensor for the metric.
- If we added higher-curvature terms, they would in general mess up the well-posedness of the Einstein equations as an initial value problem. There are certain combinations of $R^2, R_{\mu\nu}R^{\mu\nu}, R_{....}R^{....}$ which does not produce 4-derivative terms in the action. They appear in string theory sometimes and are named for Lovelock.
- If our space has a boundary, the boundary term produced by the variation of $R_{\mu\nu}$ is signficant. If you need to worry about this, the keyword is 'Gibbons-Hawking term'.

8.3 Including matter (the RHS of Einstein's equation)

The variation of the action (78) gave just the LHS of Einstein's equations. Given our discussion of the right way to think about the stress tensor in section 7, including matter in Einstein's equation is totally simple. Just add in the action for the matter:

$$S[g, \text{stuff}] = S_{EH}[g] + S_M[g, \text{stuff}].$$

Then the EoM for the metric are

$$
\begin{aligned}
0 = \frac{\delta S}{\delta g_{\mu\nu}(x)} &= \frac{\delta S_{EH}}{\delta g_{\mu\nu}(x)} + \frac{\delta S_M}{\delta g_{\mu\nu}(x)} \\
&= -\frac{1}{16\pi G_N}\sqrt{g} G^{\mu\nu} + \frac{\sqrt{g}}{2} T^{\mu\nu}
\end{aligned}
\tag{81}
$$

which says

$$\boxed{G_{\mu\nu} = 8\pi G_N T_{\mu\nu}},$$

the Einstein equation.

Notice that this is a collection of $n(n+1)/2 = 10$ in $n = 4$ dimensions nonlinear partial differential equations for the metric components, of which there are also a priori $n(n+1)/2 = 10$ components. The Bianchi identity is $n = 4$ ways in which the equations are not independent, which matches the $n = 4$ coordinate transformations we can use to get rid of bits of the metric. Nonlinear means hard because you can't just add solutions to find another solution – no Fourier decomposition for you!

Notice that, if $T_{\mu\nu} = 0$, flat space $g_{\mu\nu} = \eta_{\mu\nu}$ is a solution.

8.3.1 The cosmological constant

[Zee page 356] For example: why did we leave out the constant term in the gravity action? Suppose we put it back; we'll find an example of a stress tensor to put on the RHS of Einstein's equation. Then in the total action above,

$$S_M = -\int \mathrm{d}^4x \sqrt{g}\Lambda$$

where Λ is called the *cosmological constant*. Then the stress tensor on the RHS of Einstein's equation is just

$$T_{\mu\nu} = \Lambda g_{\mu\nu}.$$

Notice that it is covariantly conserved because g is covariantly constant. (Notice also that there is no consequence of our decomposition of S into a matter action and a gravity action, so the ambiguity in where to put the cosmological constant is not a physics ambiguity.)

Notice that if $\Lambda \neq 0$, flat spacetime is not a solution of Einstein's equations. The sign of Λ makes a big difference. (See problem set 8).

What is this Λ? Consider the case when there are scalar fields in the world. Then their stress tensor looks like $T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}L$. If we evaluate this in a configuration of the field which is constant ϕ_0 in spacetime, then we just get

$$T_{\mu\nu} = V(\phi_0) g_{\mu\nu}$$

– so $\Lambda = V(\phi_0)$, the value of the potential energy density. If we want the scalar to just sit there this value should probably be the *minimum* value of the potential. So we see that the additive normalization of our energy now really really matters. If there are multiple fields, the thing we care about is the total energy density. You might guess that the value is determined by the highest energy scale of any physics that we have to include here. So this should include contributions from atomic physics $\sim (\text{meV})^4$, from QCD $\sim (\text{GeV})^4$, from electroweak physics $\sim (\text{TeV})^4$, maybe from Grand Unified Theories $\sim (10^{16}\text{GeV})^4$? and maybe from the Planck scale $\sim (10^{19}\text{GeV})^4$. However, the observed value is 122 orders of magnitude smaller than this, $\Lambda \sim +10^{-47}\text{GeV}^4$.

As you'll see on problem set 8, a positive cosmological constant produces a metric of the form

$$ds^2_{\text{FRW}} = -dt^2 + a(t)^2 dx^i dx^j \delta_{ij}$$

with $a(t) = a_0 e^{Ht}$ and $H \sim \sqrt{G_N \Lambda}$ – space expands exponentially fast. If the value of Λ in our universe were even a little bit bigger, the universe would have expanded so quickly (compared to rates involved in particle interactions) that no galaxies would have formed, and things would have been very different then. Our colossal failure to explain its smallness combined with its crucial environmental role suggest that selection effects are responsible for its value. For more on this depressing but apparently correct line of thought, I recommend this or (more complete but before the observation of nonzero Λ) this, describing Weinberg's 1988 *prediction* of its observed value from this perpective, and also this paper, describing the mechanism (eternal inflation) by which values of Λ can be sampled.

[End of Lecture 13]

Next we will learn some useful math to help us recover from this crushing defeat.

9 Curvature via forms

The following is the way to compute curvatures if you have to do it analytically.

9.0.2 Vielbeins

We begin by introducing a new basis for vector fields and one forms.

Our previous basis for T_pM was a *coordinate* basis: $\frac{\partial}{\partial X^\mu}, \mu = 1, , \dim M$.

$$g\left(\frac{\partial}{\partial X^\mu}, \frac{\partial}{\partial X^\nu}\right) = g_{\mu\nu} \neq \eta_{\mu\nu}$$

– in general this is *not* an orthonomal basis (except in Riemann normal coords about p or in flat space). Their advantage is that

$$[\frac{\partial}{\partial X^\mu}, \frac{\partial}{\partial X^\nu}] = 0.$$

New basis for T_pM: tetrad basis:

$$e_a, a = 1.. \dim M$$

defined to be orthonormal

$$g(e_a, e_b) = \eta_{ab}. \tag{82}$$

(I'll use η to mean δ if we are in a space of euclidean signature. The letters $a, b, c...$ run over the orthonormal-basis elements.) These es are called vielbeins or n-beins or tetrads. Price for orthonormality: $[e_a, e_b] \neq 0$. But so what? See (83).

Expand these vectors in the old basis:

$$e_a \equiv e_a^\mu \frac{\partial}{\partial X^\mu}.$$

Similarly, we can expand the old basis vectors in the new basis:

$$\partial_\mu = e_\mu^a e_a$$

$$\overset{(82)}{\Longrightarrow} \quad g_{\mu\nu} = \eta_{ab} e_\mu^a e_\nu^b \quad .$$

Notice that as matrices, the basis coefficients satisfy

$$\text{“}e_\mu^a = (e_a^\mu)^{-1}\text{”}, i.e. \quad e_\mu^a e_a^\nu = \delta_\mu^\nu, \quad e_\mu^a e_b^\mu = \delta_b^a.$$

This basis for the tangent space produces a dual basis of one-forms:

$$\theta^a, \quad \text{defined by} \quad \theta^a(e_b) = \delta^a_b = \langle \theta^a, e_b \rangle$$

(just like $\mathrm{d}x^\nu \left(\frac{\partial}{\partial X^\mu}\right) = \delta^\nu_\mu$).

In terms of these tetrad-basis one-forms, the metric is just

$$\mathrm{d}s^2 = \theta^a \otimes \theta^b \eta_{ab}.$$

So these θs are easy to determine given an expression for the metric in coordinates.

Notation reminder (wedge product): From two one-forms θ^i, θ^j, we can define a section of $T^\star M \otimes T^\star M$, defined by its action on a pair of vector fields X, Y by

$$\left(\theta^i \wedge \theta^j\right)(X, Y) \equiv \theta^i(X)\theta^j(Y) - \theta^j(X)\theta^i(Y).$$

It's antisymmetric. This is called the wedge product.

Notation reminder (exterior derivative):

$$\mathrm{d}f = \partial_i f \mathrm{d}x^i, \qquad \mathrm{d}\omega = \partial_i \omega_j \mathrm{d}x^i \wedge \mathrm{d}x^j, \quad \mathrm{d}A^{(p)} = c_p \partial_{[i_1} A^{(p)}_{i_2 \dots i_{p+1}} \mathrm{d}x^{i_1} \wedge \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_{p+1}}$$

where c_p is a conventional combinatorial factor. The operation d is designed so that $\mathrm{d}^2 = 0$.

9.0.3 The price of using an orthonormal frame

We defined the Riemann tensor as

$$[\nabla_\mu, \nabla_\nu]\omega_\rho = R_{\mu\nu\rho}{}^\sigma \omega_\sigma$$

– in a coordinate basis for the tangent space. If we wanted to define the same object in a more general basis we need an extra term. For any two vector fields X, Y and one-form ω, let

$$R(X, Y)\omega \equiv \left(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\right) \omega \quad .$$

(for any ω). In this expression $\nabla_X \equiv X^\mu \nabla_\mu$ is the directional covariant derivative along the vector field X. The earlier version of Riemann returns if we set $X = \partial_\mu, Y = \partial_\nu$, in which case the weird extra term goes away. Since this is true for all ω, we don't need to write the ω :

$$R(X, Y) \equiv \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} \quad . \tag{83}$$

The last, new term is required if we want $R(X, Y)$ to be a linear, non-differential operator, satisfying

$$R(X, Y)(f\omega) = f R(X, Y)\omega$$

for any function f.

9.0.4 Connection and curvature in tetrad basis

Now consider

$$\nabla_X e_b$$

This is again a vector field and can be expanded in the new basis:

$$\nabla_X e_b = \omega_b{}^a(X) e_a$$

The $\omega_b{}^a$ are called connection coefficients. They are one-forms – they eat a v.f. (X) to give a number for each a, b. They contain the same data as the Γ^p_{mn}[33]. Notice that

$$0 = \nabla_X \underbrace{\langle e_a, e_b \rangle}_{= g_{ab}} = \omega_a{}^c(X) g_{cb} + \omega_b{}^c(X) g_{ac} = \omega_{ab} + \omega_{ba} \tag{84}$$

using metric-compatibility. This says that ω is antisymmetric in its two indices[34].

Similarly the curvature tensor can be written by acting on the basis vectors, as:

$$R(X,Y)e_b \equiv -\left(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\right) e_b = \Omega_b{}^a(X,Y) e_a. \tag{85}$$

(Note the extra sign because we are acting on a vector here, and not a one-form.) The curvature 2-forms Ω can be further expanded in our basis:

$$\Omega_b{}^a = \frac{1}{2} R_{cdb}{}^a \theta^c \wedge \theta^d \ . \tag{86}$$

In the tetrad basis, we have $R_{cdb}{}^a = -R_{dcb}{}^a$ by construction, since the wedge is antisymmetric.

We can verify that this is consistent by taking the overlap with the tetrad-basis one-forms:

$$\begin{aligned}
\langle \theta^a, R(e_c, e_d) e_b \rangle &\overset{(85)}{=} \langle \theta^a, \underbrace{\Omega_b^f(e_c, e_d)}_{= R_{ghb}{}^f \theta^g \wedge \theta^f (e_b, e_c)} e_f \rangle \\
&\overset{(86)}{=} \frac{1}{2} \underbrace{\langle \theta^a, e_f \rangle}_{= \delta^a_f} \left(R_{cdb}{}^f - R_{dcb}{}^f\right) \\
&= \frac{1}{2} \left(R_{cdb}{}^a + R_{cdb}{}^a\right) = R_{cdb}{}^a.
\end{aligned} \tag{87}$$

Notice that we can also expand the curvature two forms in a coordinate basis for $T_p^\star M$:

$$\Omega_b{}^a = \frac{1}{2} \Omega_{\mu\nu b}{}^a \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

[33]For the explicit connection between the connections see Zee page 603

[34]Notice that I have lowered the *second* index – although these indices are raised and lowered with the innocuous η_{ab}, the antisymmetry of ω means we must be careful to distinguish $\omega_a{}^b$ from $\omega^b{}_a$.

This lets us relate this fancy Ω thing to the components of the Riemann tensor in a coordinate basis:

$$R_{\mu\nu\rho}{}^{\sigma} = e^b_\rho e^\sigma_a \Omega_{\mu\nu b}{}^{a}. \tag{88}$$

Cartan structure equations. Here comes the payoff:

$$\mathrm{d}\theta^a + \omega_b{}^a \wedge \theta^b = 0 \qquad (\text{'}\mathrm{d}\theta + \omega\theta = 0\text{'}) \tag{89}$$

$$\Omega_b{}^a = -\mathrm{d}\omega_b{}^a + \omega_b{}^c \wedge \omega_c{}^a \qquad (\text{'}R = -\mathrm{d}\omega + \omega^2\text{'}). \tag{90}$$

These allow us to calculate ω and Ω respectively, from which we can extract R using (86).

Eqn. (89) follows from torsion freedom of ∇, which can be written as the statement that the following vector field vanishes:

$$\nabla_X Y - \nabla_Y X - [X, Y] = 0. \tag{91}$$

Expand Y in our new basis: $Y = \theta^b(Y)e_b$. Then, from our expansion of $\nabla_X e_b$ in the tetrad basis,

$$\nabla_X Y = X(\theta^b(Y))e_b + \theta^b(Y)\omega_b{}^a(X)e_a$$

The torsion-free condition is then:

$$0 \overset{(91)}{=} (X\theta^a(Y))e_a - (Y\theta^a(X))e_a - \theta^a([X,Y])e_a + \left(\omega_b{}^a(X)\theta^b(Y) - \omega_b{}^a(Y)\theta^b(X)\right) e_a = 0\,. \tag{92}$$

Since the basis vectors are independent, each coefficient of e_a must vanish independently:

$$0 = \underbrace{X\theta^a(Y) - Y\theta^a(X) - \theta^a([X,Y])}_{=\mathrm{d}\theta^a(X,Y)} + \left(\omega_b{}^a \wedge \theta^b\right)(X,Y)$$

The underbraced equation follows from the general statement:

$$X\omega(Y) - Y\omega(X) - \omega([X,Y]) = \mathrm{d}\omega(X,Y)\ , \tag{93}$$

true for any vfs X, Y and one-form ω, which follows *e.g.* by writing it out in local coordinates. (This is sometimes taken as the definition of d.)

Pf of (90): This follows from the definition of the Riemann tensor

$$\begin{aligned}
-\Omega_b{}^a(X,Y)e_a &= -R(X,Y)e_b \\
&= \left(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\right) e_b \\
&= \nabla_X\left(\omega_b{}^a(Y)e_a\right) - \nabla_Y\left(\omega_b{}^a(X)e_a\right) - \omega_b{}^a\left([X,Y]\right)e_a \\
&= X\omega_b{}^c(Y)e_c - Y\omega_b{}^c(X)e_c - \omega_b{}^c\left([X,Y]\right)e_c + \underbrace{\omega_b{}^a(Y)\omega_a{}^c(X)e_c - \omega_b{}^a(X)\omega_a{}^c(Y)e_c}_{=-\omega_b{}^a\wedge\omega_a{}^c(X,Y)e_k} \\
&= \underbrace{\left(X\omega_b{}^c(Y) - Y\omega_b{}^c(X) - \omega_b{}^c\left([X,Y]\right)\right)}_{\overset{(93)}{=}\mathrm{d}\omega_b{}^c(X,Y)} e_c - \omega_b{}^a \wedge \omega_a{}^c(X,Y)e_c \\
&= \left(\mathrm{d}\omega_b{}^c - \omega_b{}^a \wedge \omega_a{}^c\right)(X,Y)e_c
\end{aligned} \tag{94}$$

which proves the second Cartan structure equation. (Beware that some sign conventions differ between the above expressions and ones you will see in other sources.)

Sample application: $\mathbb{R}^2$ in polar coordinates

step 1: demand that the metric is diagonal (δ or η) in the tetrad basis:

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2 = \theta^{\hat{r}}\theta^{\hat{r}} + \theta^{\hat{\varphi}}\theta^{\hat{\varphi}}.$$

Notice that I am using hats to indicate orthonormal-basis indices which are named after coordinate indices. From which we conclude that the tetrad basis is:

$$\theta^{\hat{r}} = \mathrm{d}r, \theta^{\hat{\varphi}} = r\mathrm{d}\varphi.$$

step 2: take derivatives. Since $\mathrm{d}^2 = 0$, we have

$$\mathrm{d}\theta^{\hat{r}} = \mathrm{d}^2 r = 0.$$

On the other hand, from (89) we have

$$\mathrm{d}\theta^{\hat{r}} = -\omega_{\hat{\varphi}}{}^{\hat{r}} \wedge \theta^{\hat{\varphi}}$$

and there is no $\omega_{\hat{r}}{}^{\hat{r}}$ by (84). This means that $\omega_{\hat{\varphi}}{}^{\hat{r}} \wedge \theta^{\hat{\varphi}} = 0$ which means that $\omega_{\hat{\varphi}}{}^{\hat{r}} \propto \theta^{\hat{\varphi}}$. And by explicit computation

$$\mathrm{d}\theta^{\hat{\varphi}} = \mathrm{d}(r\mathrm{d}\varphi) = \mathrm{d}r \wedge \mathrm{d}\varphi$$

while (i) says

$$\mathrm{d}\theta^{\hat{\varphi}} = -\omega_{\hat{r}}{}^{\hat{\varphi}} \wedge \theta^{\hat{r}} = -\omega_{\hat{r}}{}^{\hat{\varphi}} \wedge \mathrm{d}r$$

from which we conclude

$$\omega_{\hat{r}}{}^{\hat{\varphi}} = \mathrm{d}\varphi$$

and all others vanish or are determined from it by symmetry. This is like knowing Γ.

step 3: Use (90) :

$$\Omega_b{}^a = -\mathrm{d}\omega_b{}^a - \omega_c{}^a \wedge \omega_b{}^c = 0$$

In this silly example we conclude that the curvature is zero.

Here is a less trivial example: **Sample application: round S^2 of radius R**

$$\mathrm{d}s^2 = R^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right) = e^{\hat{\theta}}e^{\hat{\theta}} + e^{\hat{\varphi}}e^{\hat{\varphi}}$$

step 1[35]:

$$\implies e^{\hat{\theta}} = R\mathrm{d}\theta, e^{\hat{\varphi}} = R\sin\theta\mathrm{d}\varphi = e^{\hat{\varphi}}_{\varphi}\mathrm{d}\varphi.$$

[35] I've called the orthonormal one-forms e here while I called them θ earlier. Sorry.

step 2:

$$\mathrm{d}e^{\hat\theta} = 0 = -\omega_{\hat\varphi}{}^{\hat\theta} \wedge e^{\hat\varphi} \implies \omega_{\hat\varphi}{}^{\hat\theta} = -\omega_{\hat\theta}{}^{\hat\varphi} \propto e^{\hat\varphi}.$$

$$\mathrm{d}e^{\hat\varphi} = R\cos\theta \mathrm{d}\theta \wedge \mathrm{d}\varphi = \frac{1}{R}\cot\theta e^{\hat\theta} \wedge e^{\hat\varphi} = -\omega_{\hat\theta}^{\hat\varphi} \wedge e^{\hat\theta}.$$

$$\implies \omega_{\hat\theta}{}^{\hat\varphi} = \frac{1}{R}\cot\theta e^{\hat\varphi} = -\omega_{\hat\theta}{}^{\hat\varphi}, \quad \omega_{\hat\theta}{}^{\hat\theta} = 0 = \omega_{\hat\varphi}{}^{\hat\varphi}.$$

step 3:

$$\Omega_{\hat\theta}^{\hat\varphi} = -\mathrm{d}\omega_{\hat\theta}^{\hat\varphi} + \underbrace{\omega_{\hat\varphi}{}^{\hat\theta} \wedge \omega_{\hat\theta}{}^{\hat\varphi}}_{\propto e^{\hat\varphi}\wedge e^{\hat\varphi}=0} = -\mathrm{d}\left(\cos\theta \mathrm{d}\varphi\right) = +\sin\theta \mathrm{d}\theta \wedge \mathrm{d}\varphi \quad .$$

Using (88) we have

$$R_{\theta\varphi\theta}{}^{\varphi} = \underbrace{e_{\theta}^{\hat\theta}}_{=1} \underbrace{e_{\hat\varphi}^{\varphi}}_{1/\sin\theta} \underbrace{\Omega_{\theta\varphi\hat\theta}{}^{\hat\varphi}}_{-\sin\theta} = \frac{1}{\sin\theta}\left(+\sin\theta\right) = 1.$$

and

$$R_{\theta\varphi\varphi}{}^{\theta} = \underbrace{e_{\hat\theta}^{\theta}}_{=1} \underbrace{e_{\varphi}^{\hat\varphi}}_{\sin\theta} \underbrace{\Omega_{\theta\varphi\hat\varphi}{}^{\hat\theta}}_{-\sin\theta} = -\sin^2\theta$$

(and other nonzero components determined by symmetry) which is what we got using Christoffel symbols on problem set 6.

[End of Lecture 14]

10 Linearized Gravity

[I recommend Carroll's discussion of the subject of this chapter.] Now that we know what equation to solve for the dynamical metric, we must look for solutions! In §12 we will discuss the Schwarzchild solution which describes the geometry created by a spherically symmetric mass distribution. Here we'll ask a simpler question about solutions of Einstein's equations which are nearly Minkowski space.

As we did earlier in §6.0.3, let's consider *linearizing* about flat space (which is the only solution of Einstein's equations that we know so far!):

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h| \ll 1, \qquad g^{\mu\nu} = \eta^{\mu\nu} - \underbrace{h^{\mu\nu}}_{\eta^{\mu\alpha}\eta^{\nu\beta}h_{\alpha\beta}} \quad .$$

Everywhere in what follows "$+\mathcal{O}(h^2)$ " is implied. (If this makes you uncomfortable, imagine a little ϵ in front of every $h_{\mu\nu}$.) Indices are raised and lowered with the Minkowski metric η.

Let's plug this into Einstein's equation, $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G_N T_{\mu\nu}$, again keeping only the leading terms in h. First

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}\underbrace{\left(g_{\mu\sigma,\nu} + g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}\right)}_{\mathcal{O}(h)} = \frac{1}{2}\eta^{\rho\sigma}\left(h_{\mu\sigma,\nu} + h_{\sigma\nu,\mu} - h_{\mu\nu,\sigma}\right)$$

$$R_{\mu\nu\rho\sigma} = \partial_\nu\Gamma_{\mu\rho|\sigma} - \partial_\mu\Gamma_{\nu\rho|\sigma} - \underbrace{\Gamma\Gamma}_{=\mathcal{O}(h^2)} = \frac{1}{2}\left(-h_{\nu\sigma,\mu\rho} + h_{\mu\sigma,\nu\rho} + h_{\nu\rho,\mu\sigma} - h_{\mu\rho,\nu\sigma}\right).$$

Let $h \equiv \eta^{\mu\nu}h_{\mu\nu}, \Box \equiv \partial_\mu\partial_\nu\eta^{\mu\nu}$. Then

$$R_{\mu\nu} = \frac{1}{2}\left(\partial_\mu\partial_\nu h + \partial_\mu\partial^\rho h_{\nu\rho} + \partial_\nu\partial^\rho h_{\mu\rho} - \Box h_{\mu\nu}\right)$$

$$\begin{aligned} G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \\ &= -\frac{1}{2}\left(\Box h_{\mu\nu} + \partial_\mu\partial_\rho h^\rho_\nu + \partial_\nu\partial_\rho h^\rho_\mu - \partial_\mu\partial_\nu h - \eta_{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} + \eta_{\mu\nu}\Box h\right) \\ &\overset{\text{Einstein}}{=} 8\pi G_N T_{\mu\nu}. \end{aligned} \tag{95}$$

So this is a 2nd order linear differential equation, of the form

$$(\mathcal{L}h)_{\mu\nu} = \text{sources}_{\mu\nu}.$$

Here $\mathcal{L}$ is a 2nd order Laplace-like operator acting on symmetric rank-two tensor fields ($\mathcal{L}$ is for 'linear' or 'Lichnerowicz').

Let's compare to Maxwell's equations (which are linear in A so no need to linearize):

$$\partial^\mu F_{\mu\nu} = \Box A_\nu - \partial_\nu \partial^\mu A_\mu \overset{\text{Maxwell}}{=} 4\pi j_\nu .$$

Again the LHS is of the form $\mathcal{L}A$ with $\mathcal{L}$ a 2d order linear differential operator acting on a tensor field, this time of rank 1. Notice that a gauge transformation $A_\mu \to A_\mu + \partial_\mu \lambda$ preserves exactly this combination

$$\Box A_\nu - \partial_\nu \partial^\mu A_\mu \to \Box A_\nu - \partial_\nu \partial^\mu A_\mu + \Box \partial_\nu \lambda - \partial_\nu \partial^\mu \partial_\mu \lambda = \Box A_\nu - \partial_\nu \partial^\mu A_\mu .$$

Similarly, we could have arrived at the above linearized Einstein equation (95) by *demanding* coordinate invariance. That is, under (recall pset 7)

$$x^\mu \to x^\mu + \epsilon^\mu(x), \quad g_{\mu\nu} \to g_{\mu\nu} - \left(\nabla_\mu \epsilon_\nu + \nabla_\nu \epsilon_\mu\right) \quad \Longrightarrow h_{\mu\nu} \to h_{\mu\nu} - \left(\underbrace{\partial_\mu}_{=\nabla \text{ for } g=\eta} \epsilon_\nu + \partial_\nu \epsilon_\mu \right)$$

Under this replacement, $\mathcal{L}h$ in (95) is invariant. So we could have begun the course with the equation $\Box h_{\mu\nu} = 0$ (generalizing the Poisson equation for the Newtonian potential) and added stuff to cancel the variations. This particle-physics-like point of view is taken in for example the book by Weinberg and Feynman's lectures on gravitation. From that point, you can use this as the kinetic term for a free field with two indices and try to add interactions which preserve this invariance. The geometric approach we've taken is conceptually more slippery perhaps, but much less hairy algebraically.

From our point of view now, the coordinate invariance is an annoying redundancy. We can use it to make our equation simpler. The same approach works well for Maxwell's equations, where we could choose various gauge conditions, such as 'temporal gauge' $A_0 = 0$ (not Lorentz invariant but good for getting rid of minus signs) or 'Lorentz gauge' $\partial_\mu A^\mu = 0$. We'll do the latter. This condition does not completely fix the redundancy since

$$A_\mu \to A_\mu + \partial_\mu \lambda \implies \partial_\mu A^\mu \to \partial_\mu A^\mu + \partial_\mu \partial^\mu \lambda$$

so λ with $\Box \lambda = 0$ preserve our gauge condition (this ambiguity is removed by a choice of initial condition).

The GR analog of Lorentz gauge is "de Donder gauge":

$$\partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h = 0. \tag{96}$$

A linearized coordinate transformation

$$h_{\mu\nu} \to h_{\mu\nu} - \left(\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu\right)$$

preserves the gauge condition (96) if

$$0 = \Box \epsilon_\nu + \partial^\mu \partial_\nu \epsilon_\mu - \frac{1}{2} 2 \partial_\nu \partial^\mu \epsilon_\mu = \Box \epsilon_\nu.$$

Side remark. Note that (96) is the *linearized* definition of de Donder gauge. The nonlinear definition is

$$0 = g^{\mu\nu} \Gamma^\rho_{\mu\nu}. \tag{97}$$

You can check that this reduces to (96) if you linearize about Minkowski space. Note that neither is a tensor equation, which is what we want in order to fix a property of our coordinate system. Under a nonlinear coordinate transformation $x^\mu \to \tilde{x}^\mu(x)$, (97) changes by

$$0 = g^{\mu\nu} \Gamma^\rho_{\mu\nu} \to 0 = \tilde{g}^{\mu\nu} \tilde{\Gamma}^\rho_{\mu\nu} = g^{\mu\nu} \Gamma^\sigma_{\mu\nu} \partial_\sigma \tilde{x}^\rho - \partial_\mu \partial_\nu \tilde{x}^\rho g^{\mu\nu}.$$

Requiring that (97) be true both before and after the transformation then restricts our choice of coordinate transformations to those $\tilde{x}(x)$ which satisfy

$$0 = \partial_\mu \partial_\nu \tilde{x}^\rho g^{\mu\nu} = \nabla^\mu \nabla_\mu \tilde{x}^\rho = \Box \tilde{x}^\rho(x)$$

– just like in the linearized case, the condition is that the transformation is harmonic.

- Note that the scalar box is

$$\nabla^\mu \nabla_\mu = g^{\mu\nu} \partial_\mu \partial_\nu + \underbrace{g^{\mu\nu} \Gamma^\rho_{\mu\nu}}_{\overset{(97)}{=} 0} \partial_\rho = g^{\mu\nu} \partial_\mu \partial_\nu.$$

- Notice that in nonlinear de Donder gauge,

$$\nabla^\mu \omega_\mu = g^{\mu\nu} \left(\partial_\nu \omega_\mu - \Gamma^\rho_{\mu\nu} \omega_\rho \right) = \partial^\mu \omega_\mu$$

- Notice that we have treated $\tilde{x}^\rho(x)$ as a scalar quantity here – they are just functions.
- $g^{\mu\nu} \Gamma^\rho_{\mu\nu} = 0$ is n conditions.
- If you find yourself in some coordinate system such that $g^{\mu\nu} \Gamma^\rho_{\mu\nu} \neq 0$, you can pick the n functions $\tilde{x}^\rho(x)$ such that

$$0 = \tilde{g}^{\mu\nu} \tilde{\Gamma}^\rho_{\mu\nu} = g^{\mu\nu} \Gamma^\sigma_{\mu\nu} \partial_\sigma x^\rho - \partial_\mu \partial_\nu \tilde{x}^\rho g^{\mu\nu}, \forall \rho = 1..n$$

Back to linearized gravity. Imposing (linear) de Donder gauge $\partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h = 0$,

$$G_{\mu\nu} = -\frac{1}{2} \left(\Box h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \Box h \right) \overset{\text{Einstein}}{=} 8\pi G_N T_{\mu\nu}. \tag{98}$$

In particular, in vacuum, $T_{\mu\nu} = 0$. Take the trace of both sides of (98), which gives $\Box h = 0$ (for $n > 2$). Plugging this back in gives

$$\boxed{\Box h_{\mu\nu} = 0}$$

– gravitational waves !! Because the wave equation has nontrivial solutions, the metric in GR has a life of its own, even away from any sources, just like the electromagnetic field. Restoring dimensionful quantities, you see that gravitational waves move at the speed of light. This is completely new compared to the Newton gravity theory where the gravitational field is completely determined by the mass distribution.

To understand these waves better, it is useful to define $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$ in terms of which de Donder gauge is $\partial^\mu \bar{h}_{\mu\nu} = 0$ and the Einstein equation (not in vacuum) is

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu}$$

which is just n^2 copies of the Poisson equation.

(Notice that this bar operation (called 'trace reversal') squares to one in four dimensions. $R_{\mu\nu} = G_{\mu\nu} - \frac{g_{\mu\nu}}{2}G^\rho_\rho = \Box\bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\Box\bar{h} = \Box h_{\mu\nu}$.)

10.0.5 Newtonian limit.

Newton says: 1) $\rho = T_{00} \gg T_{0i}, T_{ij}$ – rest energy dominates. A nonrelativistic source is very massive compared to its kinetic energy.

2) Stationary: no x^0-dependence: $\Box = -\partial_t^2 + \Delta = \Delta$.

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu} \quad \overset{\text{NR limit}}{\rightsquigarrow} \quad \begin{cases} \Delta \bar{h}_{00} &= -16\pi G_N \rho \\ \Delta \bar{h}_{i0} &= 0 \\ \Delta \bar{h}_{ij} &= 0 \end{cases}$$

The latter two equations are easy to solve (with suitable boundary conditions): $\bar{h}_{i0} = 0 = \bar{h}_{ij}$, which says

$$h_{ij} = \frac{1}{2}\delta_{ij}h, h_{i0} = 0, h = h^\mu_\mu = -h_{00} + h_{ii} = -h_{00} + \frac{3}{2}h \implies h_{00} = \frac{1}{2}h$$

It remains to determine h:

$$\bar{h}_{00} = h_{00} - \frac{1}{2}\eta_{00}h = h_{00} + \frac{1}{2}h = \frac{1}{2}h + \frac{1}{2}h = h.$$

$$\implies \Delta \bar{h}_{00} = -16\pi G_N \rho = \Delta h$$

Comparing to Newton's equation $\Delta\phi = 4\pi G_N \rho$, we have $h = -4\phi, h_{00} = -2\phi, g_{ij} = -2\delta_{ij}\phi$. Our solution is: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ which says

$$\begin{cases} g_{00} &= -1 + h_{00} = -(1+2\phi) \\ g_{ij} &= \delta_{ij} + h_{ij} = (1-2\phi)\delta_{ij} \\ g_{i0} &= 0 \end{cases} .$$

or more compactly

$$\mathrm{d}s^2_\phi = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -(1+2\phi(\vec{x}))\mathrm{d}t^2 + (1-2\phi(\vec{x}))\mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x}$$

$$\text{with} \qquad \Delta\phi(\vec{x}) = +4\pi G_N \rho(\vec{x}) \qquad \text{(Poisson).} \tag{99}$$

Notice: (1) There's a nontrivial spatial component. So, although $\phi(x)$ satisfies the same equation as in the Newtonian gravity, it produces twice as much bending of light in Einstein's theory!

(2) As it must, since it solves the same equation, the Schwarzschild metric

$$\mathrm{d}s^2 = -\left(1 - \frac{2G_N M}{r'}\right)\mathrm{d}t^2 + \frac{1}{1-\frac{2G_N M}{r'}}\mathrm{d}r'^2 + r'^2 \mathrm{d}s^2_{S^2}$$

is of this form to leading order in h, *i.e.* in $G_N M$ with $\phi = -\frac{G_N M}{r}$. This is because $\mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x} = \mathrm{d}r^2 + r^2\mathrm{d}s^2_{S^2}$ where r' and r differ by terms which are subleading in $G_N M$. (The relation is $\left(1 - \frac{2G_N M}{r'}\right) r^2 \equiv r'^2$.)

10.0.6 Gravitational radiation.

The existence of propagating wave solutions is something new and exciting that doesn't happen in the Newtonian theory. This should be like the moment when you combined Ampere's and Faraday's laws and found that there was light.

Put back the time derivatives:

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu}, \quad \Box = -\partial_t^2 + \Delta.$$

Even without sources, $T_{\mu\nu} = 0$, this has nontrivial solutions $h_{\mu\nu}(t,\vec{x})$. To solve $\Box\bar{h}_{\mu\nu} = 0$, Fourier transform[36]:

$$\bar{h}_{\mu\nu}(x) = \int đ^4 k C_{\mu\nu}(k) e^{\mathbf{i}k\cdot x}$$

($k \cdot x \equiv k_\mu x^\mu$). Since we have linearized, we are studying a linear equation and can study it mode-by-mode, so pick $C_{\mu\nu}(k') = C_{\mu\nu}\delta(k-k')$ for some k. So

$$0 = \Box\bar{h}_{\mu\nu} = -k^2 C_{\mu\nu} e^{\mathbf{i}k\cdot x} \tag{100}$$

[36] $đ^4 k \equiv \frac{\mathrm{d}^4 k}{(2\pi)^4}$, in direct analogy with $\hbar \equiv \frac{h}{2\pi}$

requires $k^2 = 0$ for a nontrivial solution, and we see that $k^\mu = (\omega, \vec{k})^\mu$ means $|\omega| = |\vec{k}|$, *i.e.* the waves propagate at the speed of light.

We arrived at the wave equation (100) by requiring de Donder gauge, and in momentum space, this condition is

$$\partial^\mu \bar{h}_{\mu\nu} = 0 \implies \boxed{k^\mu C_{\mu\nu}(k) = 0}.$$

There is also a residual gauge ambiguity (like in Lorentz-gauge E&M) – the de Donder gauge is preserved by $x^\mu \to x^\mu + \epsilon^\mu(x)$ if $\Box \epsilon^\mu = 0$. This acts on h by

$$h_{\mu\nu} \to h_{\mu\nu} - (\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu)$$

$$\implies \bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} - (\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu - \eta_{\mu\nu}\partial^\rho \epsilon_\rho)$$

If $\epsilon_\mu(x) = \epsilon^0_\mu e^{\mathbf{i}k\cdot x}$ (everything is linear so we can study one momentum mode at a time) this acts on the Fourier components by

$$\implies C_{\mu\nu} \to C_{\mu\nu} - \mathbf{i}\left(k_\mu \epsilon^0_\nu + k_\nu \epsilon^0_\mu - \eta_{\mu\nu} k^\rho \epsilon^0_\rho\right)$$

– these two choices of C describe the *same* gravitational wave in (infinitesimally) different coordinates.

Altogether, the general solution for each wavevector k^μ satisfying $k^2 = 0$ is:

$$\begin{aligned} \bar{h}_{\mu\nu}(x) &= C_{\mu\nu} e^{\mathbf{i}k\cdot x}, \\ &\text{with} \quad C_{\mu\nu} = C_{\nu\mu}, \quad k^\mu C_{\mu\nu} = 0, \\ &\text{and} \quad C_{\mu\nu} \simeq C_{\mu\nu} + \mathbf{i}\left(k_\mu \epsilon_\nu + k_\nu \epsilon_\mu - \eta_{\mu\nu} k\cdot\epsilon\right) \end{aligned} \tag{101}$$

for any constant ϵ^μ.

How many components? That is – how many polarization states are there?

- $C_{\mu\nu} = C_{\nu\mu}$ symmetric 4-index object is 10.
- gauge condition $k^\mu C_{\mu\nu} = 0$ is 4 equations $10 - 4 = 6$.
- residual gauge equivalence ϵ^μ removes 4 more: $10 - 4 - 4 = 2$.

(The analogous counting for Maxwell in Lorentz gauge: four possible polarizations $A_\mu(k)$ is $n = 4$, $k^\mu A_\mu(k) = 0$ is one constraint $n - 1 = 4 - 1 = 3$, and the residual gauge transf is $A_\mu \to A_\mu + k_\mu \epsilon$ removes one more: $n - 2 = 4 - 2 = 2$ – same answer, different counting. In other dimensions, these numbers aren't the same. The photon always has $n-2$ polarizations. The graviton has $\frac{n(n+2)}{2} - n - n = \frac{1}{2}n(n-3)$. Notice that this is *zero* in three dimensions.)

Explicitly, for any four-vector $k_\mu = (\omega, \vec{k})_\mu$ with $\omega = \pm|\vec{k}|$ pick axes so that $\vec{k} = (0, 0, \omega)$. Then we can use our choice of ϵ and the condition $k^\mu C_{\mu\nu} = 0$ to write[37]

$$C_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_+ & C_\times & 0 \\ 0 & C_\times & -C_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{\mu\nu} \tag{102}$$

with C_+ and $C_\times$ parametrizing the independent polarizations (time is the first row of the matrix). (The names will be explained below.) A gravitational wave in this form is said to be in "tranvserse, traceless" gauge; notice that with this choice $\bar{h}_{\mu\nu} = h_{\mu\nu}$, since they differ by a term proportional to the trace. [End of Lecture 15]

Comment on problem set 7 number 1 and problem set 8 number 3: this is a definition: $\tilde{\phi}(\tilde{x}) = \phi(x)$ for the coordinate transformation of a scalar field. This is a statement of invariance: $\tilde{\phi}(x) = \phi(x)$. For example, if $\tilde{x} = x + \epsilon$, then this says $\tilde{\phi}(x) = \phi(x - \epsilon)$, that is: ϕ is constant. Similarly, $\tilde{g}_{\mu\nu}(x) = g_{\mu\nu}(x)$ is an invariance statement. To test for invariance, you have to compare the tensors at different points related by the alleged symmetry transformation.

10.1 Gravitational wave antennae

Two questions we should address: (1) what do you feel if a gravitational wave of the form (101), (102) passes by you? (2) How are gravitational waves produced?

10.1.1 Receiver

(1) A single particle doesn't care (a proof of this is the construction of *Fermi normal coordinates* which set the Christoffel symbols to zero along a whole geodesic; see Zee p. 557). The interesting thing is the resulting tidal forces between particles, as expressed in the geodesic deviation equation. We (mostly) derived this earlier (in §6.0.5) and found that the relative acceleration was determined by the Riemann tensor:

$$D_\tau^2 \xi^\mu = R_{\nu\rho\sigma}{}^\mu \xi^\nu \partial_\tau x^\rho \partial_\tau x^\sigma$$

where $\xi^\mu = \partial_s x^\mu(\tau, s)$ is the deviation vector describing the separation vector between two nearby geodesics. Recall also that $D_\tau \equiv \partial_\tau x^\mu \nabla_\mu$. Let's consider a collection of slow-moving particles in our gravitational wave. Slow-moving means $\partial_\tau x^\rho \simeq (1, \vec{0})^\rho$, and $\tau \sim t$, so the geodesic deviation equation becomes

$$\partial_t^2 \xi^\mu = R_{\nu tt}{}^\mu \xi^\nu$$

[37] In words: we can use our $n = 4$ components of ϵ_μ^0 to make $C_{\mu\nu}$ traceless and and to make $C_{0i} = 0$.

which in the wave solution

$$R_{\nu tt}{}^{\mu} = \frac{1}{2}\partial_t^2 h_\nu^\mu$$

which gives

$$\partial_t^2 \xi^\mu = \frac{1}{2}\xi^\nu \partial_t^2 h_\nu^\mu .$$

We can solve this equation given some initial separations $\xi(0)$. Notice that if the wave travels in the x^3 direction, the transverse property means that ξ^3 doesn't care about it, so we can study particles distributed in the $x - y$ plane. First consider a pure C_+ wave; then the deviation equation (evaluated for simplicity at $z = 0$) is

$$\partial_t^2 \xi^i(t) = \frac{1}{2}\xi^i(t)(-1)^i C_+ \partial_t^2 e^{\mathbf{i}\omega t}, \quad i = 1, 2$$

which is solved *to leading order in* C by

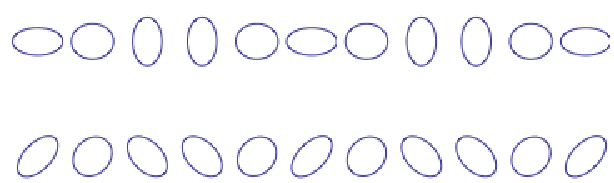
$$\xi^i(t) = \xi^i(0)\left(1 + \frac{1}{2}(-1)^i C_+ e^{\mathbf{i}\omega t}\right), \quad i = 1, 2$$

Similarly the $C_\times$ wave produces

$$\partial_t^2 \xi^i(t) = \frac{1}{2}\xi^j(t)|\epsilon_{ij}|C_\times \partial_t^2 e^{\mathbf{i}\omega t}, \quad i = 1, 2$$

which is solved *to leading order in* C by

$$\xi^i(t) = \xi^i(0) + \frac{1}{2}|\epsilon_{ij}|\xi^j(0) C_\times e^{\mathbf{i}\omega t}, \quad i = 1, 2$$



We can illustrate this behavior by thinking about particles distributed on a ring in the xy plane, so $(x, y)(0) = (\cos\theta, \sin\theta)$. When one of the two gravitational waves above hits them, to leading order in C, their subsequent separations from a particle at the origin are given by $(x, y)(t) = \xi^i(t)$ with the formulae above. The results of the two polarizations are illustrated respectively in the figures at right (and by the animations on the lecture notes page). Hence the names. (So more precisely: these are polar plots of the separations from the origin of the particles initially at $(\cos\theta, \sin\theta)$. The problem is actually completely translation-invariant in x, y until we place test-particles.)

You can also add the two polarizations out of phase to make something that goes around in a circle, like circular polarizations of light. This is illustrated in the rightmost animation on the lecture notes page.

Notice that these waves have *spin two* in the sense that a rotational by π (*e.g.* in the transverse plane illustrated here) returns them to themselves. (More generally, recall that a spin s excitation is returned to itself after a rotation by $2\pi/s$. For example, an E&M wave is described by a polarization *vector*, which takes 2π to go all the way around.)

10.1.2 Transmitter

(2) Sources. A necessary condition for a source of a gravitational wave is a region of spacetime with $T_{\mu\nu} \neq 0$. Then we'll solve

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu} \tag{103}$$

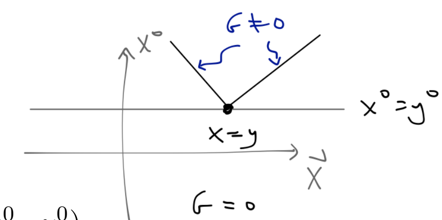
where $\Box = -\partial_t^2 + \vec{\nabla}^2$, which is more or less 10 copies of the familiar equation

$$\Box \phi(x) = \rho(x) \ .$$

We solve this (linear) equation using Green's functions:

$$\Box G(x, y) = \delta^4(x - y)$$

with retarded boundary conditions that $G(x, y) = 0$ if $x^0 < y^0$. (In the E&M context, this is sometimes called Lienard-Wiechert potential.) The solution (which you can find by Fourier transform) is



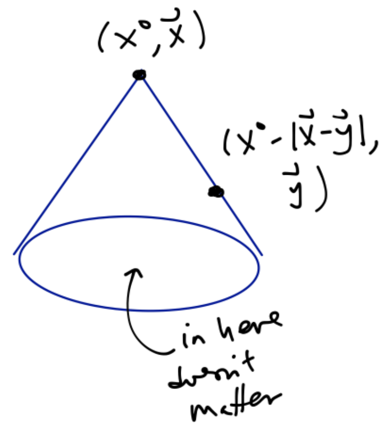
$$G(x, y) \overset{\text{transl. inv.}}{=} G(x-y) = -\frac{1}{4\pi|\vec{x} - \vec{y}|}\delta(x^0-y^0-|\vec{x}-\vec{y}|)\theta(x^0-y^0).$$

It is only supported on the future lightcone of the source event at $y^0, \vec{y}$. (This statement is called *Huygens' principle*, and it is in fact false in even spatial dimensions, where *e.g.* lightrays would propagate also *inside* the lightcone.)

So to solve (103) we make the source by adding together delta functions and find:

$$\begin{aligned} \bar{h}_{\mu\nu}(x) &= -16\pi G_N \int G(x - y) T_{\mu\nu}(y) \mathrm{d}^4 y \\ &= 4G_N \int \frac{T_{\mu\nu}(y^0 = x^0 - |\vec{x} - \vec{y}|, \vec{y})}{|\vec{x} - \vec{y}|} \mathrm{d}^3 y \end{aligned} \tag{104}$$

This is *a* causal solution; we can add to it a solution of the homogeneous equation. By Huygens' principle, the response at x doesn't care about the stuff *inside* the past lightcone.

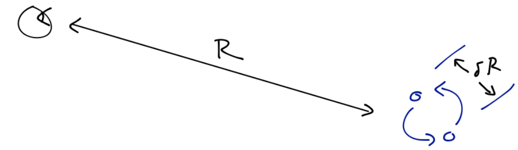


Now let's Fourier decompose the source, and consider one frequency at a time:

$$T_{\mu\nu}(x^0, \vec{x}) = T_{\mu\nu}(x^0 + \frac{2\pi}{\omega}, \vec{x}) = e^{\mathbf{i}\omega x^0} \mathbb{T}_{\mu\nu}(\vec{x}).$$

The response to this periodic source is

$$\bar{\mathfrak{h}}_{\mu\nu}(x) = 4G_N \left(\int \mathrm{d}^3 y \mathbb{T}_{\mu\nu}(\vec{y}) \frac{e^{\mathbf{i}\omega(x^0 - |\vec{x} - \vec{y}|)}}{|\vec{x} - \vec{y}|} \right)$$



Let's also assume the spatial extent of the source is much smaller than the distance to us (and in particular, the source is localized so we can IBP with impunity). For example, imagine you are on Earth and asking about gravitational waves created by a binary pulsar of size δR, a distance $R \gg \delta R$ away. Here $R \equiv |\vec{x} - \vec{y}|$, the distance to the center of mass (CoM) of the source. We will also assume $\omega \delta R \ll 1$. This allows a multipole expansion, the leading term of which is

$$\bar{\mathfrak{h}}_{\mu\nu}(x) \simeq 4G_N \frac{1}{R} e^{\mathbf{i}\omega(x^0 - R)} \int \mathrm{d}^3 y \mathbb{T}_{\mu\nu}(\vec{y})$$

and $\mathfrak{h}_{00} = 2\phi = \frac{2G_N M}{R}$, $M = \int \mathbb{T}_{00}$. To distinguish the propagating wave from this bit which is present even in the Newtonian theory, look at $\mathfrak{h}_{ij}$, for which we will need $\int \mathrm{d}^3 y \mathbb{T}_{ij}$. The corrections to this approximation are down by additional powers of $\frac{G_N \mu}{R}$ where μ represents the various multipole moments of the mass distribution.

Trick (virial theorem, aka IBP): Then conservation of stress-energy says $T_{\mu\nu}(x) = \mathbb{T}_{\mu\nu}(\vec{x}) e^{\mathbf{i}\omega x^0}$ satisfies $\partial^\mu T_{\mu\nu}(x) = 0$ (to leading order in h), so:

$$0 = -\partial_t T_{0\nu} + \partial_i T_{i\nu} = e^{\mathbf{i}\omega x^0} \left(-\mathbf{i}\omega \mathbb{T}_{0\nu} + \partial_i \mathbb{T}_{i\nu} \right) \implies \mathbb{T}_{0\nu} = -\frac{\mathbf{i}}{\omega} \partial_i \mathbb{T}_{i\nu}$$

Using this twice gives

$$\implies \boxed{\rho = \mathbb{T}_{00} = -\frac{1}{\omega^2} \partial^i \partial^j \mathbb{T}_{ij}}$$

Put the CoM of the source at $\vec{y} = 0$. Notice that the very fact that we can do this by adding a constant to our coordinates is the statement that the dipole moment of the mass distribution is not going to be meaningful as a source for gravitational waves. Then

$$-\frac{1}{2}\omega^2 \int \mathrm{d}^3 y y^i y^j \rho(\vec{y}) \equiv -\frac{\omega^2}{6} q^{ij} = \frac{1}{2} \int y^i y^j \partial^k \partial^l \mathbb{T}_{kl}(\vec{y}) \mathrm{d}^3 y \overset{\text{IBP } 2\times}{=} +\frac{3}{2} \int \mathbb{T}_{ij}(\vec{y}) \mathrm{d}^3 y$$

is (proportional to) the quadrupole moment for $\omega \neq 0$. There is no boundary term if the source is localized.

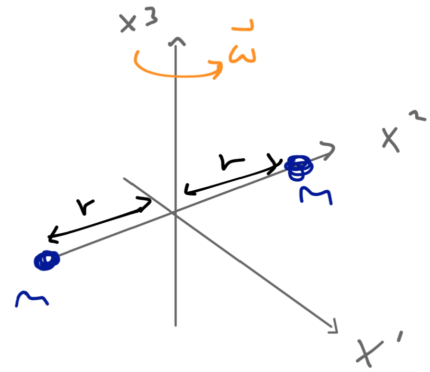
$$\implies \bar{\mathfrak{h}}_{ij}(x) \simeq \frac{4G_N}{R} e^{\mathbf{i}\omega(x^0 - R)} \int \mathbb{T}_{ij}(\vec{y}) \mathrm{d}^3 y = -\frac{2G_N \omega^2}{3} \frac{e^{\mathbf{i}\omega(x^0 - R)}}{R} q_{ij}.$$

Back in the time domain, this is

$$\boxed{\bar{\mathfrak{h}}_{ij}(x) = \frac{2G_N}{3R} \frac{\mathrm{d}^2}{\mathrm{d}t^2} q_{ij}(t - R)}, \quad q_{ij}(y^0) \equiv 3 \int y_i y_j \rho(y^0, \vec{y}) \mathrm{d}^3 y \tag{105}$$

This is small for many reasons, one of which is that the ∂_t brings down powers of $\omega \delta R \ll 1$.

Notice that the GR wave has no dipole moment contribution, since there's no negative gravitational charge. A dipole requires polarization. So the quadrupole approximation, given above, is the leading contribution.



For example, let's consider the quadrupole moment of a binary star system, rotating about the $\hat{z}$ axis. Treat the two stars as point masses, and assume their masses are equal. Its equations of motion relate the rotation frequency ω to the mass and size: $\omega = \sqrt{\frac{G_N M}{4r^3}}$, where $2r$ is the separation between the two point masses.

The resulting signal along the z-axis is:

$$\bar{h}_{ij} = \frac{8G_N M}{R}\omega^2 r^2 \begin{pmatrix} -\cos 2\omega t & -\sin 2\omega t & 0 \\ -\sin 2\omega t & \cos 2\omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}$$

where $t \equiv x^0 - R$ is the retarded time. This is a superposition of the two polarizations C_+ and $C_\times$, out of phase by $\pi/2$; we shouldn't be too surprised that sources moving in a circle produce a circularly polarized wave.

(Note that we did a potentially illegal thing here: the circular orbit of the stars around each other requires non-linear (albeit Newtonian, even Keplerian) gravity; on the other hand, we've treated the gravitational radiation that is produced by the resulting stress tensor only to first order in its amplitude. It is OK.)

10.2 The gravitational field carries energy and momentum

Gravitational waves carry energy. How is this consistent with the idea that we can locally eliminate the gravitational field, by the EEP? Well, the fact that we can't eliminate a gravitational wave by a coordinate change (recall that we used those up and were still left with two nontrivial polarization states for each (null) wavevector) already means that we are not in trouble with the EEP. The thing we can eliminate (locally) is a (locally) *uniform* gravitational field.

[End of Lecture 16]

But the definition of energy is quite tricky. The following discussion applies to the case where the spacetime is *asymptotically flat* – this means that far away, it looks like Minkowski space. (This notion needs to be made more precise.) In this case, we can define a time coordinate (a nice inertial frame at infinity), and define an energy associated with that time coordinate. Here's how the association works.

First define the energy-momentum tensor $T_{\mu\nu}$ for the matter in a fixed geometry $g_{\mu\nu}(x)$ (with $g_{\mu\nu}(x) \to \eta_{\mu\nu}$ as $x \to \infty$) by the $\frac{\delta S}{\delta g_{\mu\nu}}$ procedure from section 7. By problem set 7, we have $\nabla^\mu T_{\mu\nu} = 0$. Notice that this is *not* $\partial^\mu T_{\mu\nu} = 0$. Recall that in flat spacetime (*i.e.* in special relativity, as in (36)), we have

$$P^\mu = \int_{\Sigma_t} T^{0\mu} \mathrm{d}^3\vec{x}$$

where Σ_t is a slice of spacetime at fixed time, *i.e.* what someone could consider 'space'.

$$\implies \partial_t P^\mu = \int_{\Sigma_t} \partial_t T^{0\mu} \mathrm{d}^3 x \overset{\text{IF } \partial_\mu T^{\mu\nu}=0}{=} + \int_{\Sigma_t} \partial_i T^{i\mu} \mathrm{d}^3 x = \oint_{\partial\Sigma_t = \emptyset} \mathrm{d}s_i T^{i\mu} = 0$$

– any change in the amount of conserved quantity in the region is accounted for by flux through the boundary of the region. In curved spacetime, $g_{\mu\nu}(x)$, instead we have

$$\partial_t P^\mu \overset{?}{=} \int \mathrm{d}^3 x \left(\partial_i T^{i\mu} - \Gamma^\mu_{i\alpha} T^{i\alpha} - \Gamma^i_{i\alpha} T^{\mu\alpha} \right)$$

The first term can be Stokesed, but the Christoffel terms can't. Also we haven't used the covariant integration measure. What about the more ideologically-correct object[38]

$$P^\mu \equiv \int_{\Sigma_t} \mathrm{d}^3 x \sqrt{g} T^{0\mu}$$

where Σ_t is a slice of spacetime at fixed time t? Now we have

$$\partial_\mu \left(\sqrt{g} T^\mu_\nu \right) \overset{\nabla_\mu T^{\mu\nu}=0}{=} \frac{1}{2} (\partial_\nu g_{\rho\sigma}) \sqrt{g} T^{\rho\sigma} \overset{\text{Einstein}}{=} \frac{1}{16\pi G_N} (\partial_\nu g_{\rho\sigma}) \sqrt{g} \left(R^{\rho\sigma} - \frac{1}{2} R g^{\rho\sigma} \right) .$$

[38] Notice that this is related to the construction you studied on problem set 8. If $\xi = \partial_t$ is a Killing vector field then in fact the Christoffel terms go away and the energy is conserved. In the case of a general solution of the linearized equations, ∂_t is a Killing vector field in the background Minkowski space, but not of the wave solution. The point is that energy can be exchanged between the gravitational field and matter.

(Notice that this does not contradict our claims (near (80)) about being about to integrate by parts with covariant derivatives because we are violating the assumption there that all the indices are contracted.) This expression *is* somebody's (ordinary) divergence:

$$\partial_\mu \left(\sqrt{g} T^\mu_\nu\right) \qquad\qquad = -\partial_\mu \left(\sqrt{g} t^\mu_\nu\right)$$

where

$$t^{\mu\nu} \equiv \frac{1}{16\pi G_N} \left((2\Gamma\Gamma - \Gamma\Gamma - \Gamma\Gamma)(gg - gg) + g^{\mu\rho} g^{\sigma\kappa} (\Gamma\Gamma + \Gamma\Gamma - ...) + (\mu \leftrightarrow \nu) + gg (\Gamma\Gamma - \Gamma\Gamma) \right) \tag{106}$$

is (a sketch of) the Landau-Lifshitz "quasi-tensor"[39]. So we have

$$\partial_\mu \left(\sqrt{g} \left(T^\mu_\nu + t^\mu_\nu\right)\right) = 0$$

and can try to interpret the second term as the EM "tensor" for the gravitational field. But it can't be a tensor because we could choose Riemann normal coordinates at p and set $\Gamma|_p = 0 \implies t^{\mu\nu}|_p = 0$. So the gravitational energy is NOT a local concept, as expected by the EEP.

If gravity is weak, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, in de Donder gauge ($\partial^\mu \bar{h}_{\mu\nu} = 0$), this expression simplifies dramatically:

$$t_{\mu\nu} = \frac{1}{32\pi G_N} \partial_\mu h_{\rho\sigma} \partial_\nu h^{\rho\sigma} \tag{107}$$

Notice that this is manifestly $\mathcal{O}(h^2)$, so we can consistently ignore the back-reaction of the metric on the gravitational wave at first order in h.

[Wald page 84] Here's another way to think about this: the linearized Einstein equation is

$$G^{(1)}_{\mu\nu}[h_{\rho\sigma}] = T^{(1)}_{\mu\nu}$$

where the superscript indicates the order in the expansion in h. But this linearized solution is *not* a solution to the next order; rather:

$$0 \neq G^{(2)}_{\mu\nu}[h_{\rho\sigma}] = \frac{1}{2} h^{\rho\sigma} \partial_\mu \partial_\nu h_{\rho\sigma} + ...$$

where ... is lots of other terms, all second order in h and which are what you get if I tell you where the indices go in (106) and you plug in the linearized solution to second order. So through $\mathcal{O}(h^2)$, the Einstein equation is

$$0 = (G_{\mu\nu} - 8\pi G_N T_{\mu\nu}) = 0 + s\left(G^{(1)}_{\mu\nu}[h] - 8\pi G_N T^{(1)}_{\mu\nu}\right) + s^2 \left(G^{(1)}_{\mu\nu}[h^{(2)}] + \underbrace{G^{(2)}_{\mu\nu}[h] - 8\pi G_N T^{(2)}_{\mu\nu}}_{= -8\pi G_N T^{\text{total}}_{\mu\nu}} \right)$$

---

[39] Two confessions: This term 'quasi-tensor' is not standard. Wald and Landau-Lifshitz themselves call it a 'pseudotensor', but we've already used that term for tensors that get a minus sign under improper rotations. Also, the thing whose divergence is the expression above differs fro Landau and Lifshitz' object by various terms that don't change the total energy, as explained a bit more in (108) below. The actual expression for the LL quasi-tensor is equation (96.8) of volume II of Landau-Lifshitz. It has many indices!

where $h_{(2)}$ is the second-order correction to $g_{\mu\nu} = \eta_{\mu\nu} + sh_{\mu\nu} + s^2 h^{(2)}_{\mu\nu} + ...$ (and we have so far assumed that $T_{\mu\nu} = 0$ in our linear study). So you see that we can interpret the second order equation as again Einstein's equation but where the 1st order metric contributes to the stress tensor at the next order:

$$T^{\text{total}}_{\mu\nu} = T^{(2)}_{\mu\nu} - \frac{1}{8\pi G_N} G^{(2)}_{\mu\nu}[h] \ .$$

So this is great, except for two problems:
(1) The expression $t_{\mu\nu}$ is not actually gauge invariant. Under $h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\xi_{\nu)}$, it changes. Awful but inevitable. You knew this already from the dramatic simplification in de Donder gauge.
(2) The replacement

$$t_{\mu\nu} \to t_{\mu\nu} + \partial^\rho \partial^\sigma U_{\mu\nu\rho\sigma} \tag{108}$$

where U is local in h, quadratic in h, and satisfies symmetries similar to the Riemann tensor, $U_{abcd} = U_{[ac]bd} = U_{bdac} = U_{ac[bd]}$ doesn't change anything above.
Consolations.
(1) A virtue of the choice made by Landau-Lifshitz is that it only depends on first derivatives of g and is symmetric. They claim that it is the unique choice with these two properties. The reason we care that it is symmetric is that this makes the angular momentum, in the form

$$M^{\mu\nu} = \int_{\Sigma_t} \sqrt{g} \left(x^\mu T^{\nu t}_{\text{total}} - x^\nu T^{\mu t}_{\text{total}} \right)$$

conserved.

(2) Another consolation is that the LL quasi-tensor *is* a tensor under *linear* coordinate transformations, in particular Lorentz transformations which preserve the Minkowski metric about which we are expanding.

(3) A final consolation is that the total energy is invariant under ambiguities (1) and (2) (if the coordinate transformation ξ preserves the asymptotics). Supposing still that we have an asymptotically-flat spacetime, so that we have some canonical coordinates at ∞, with a well-defined spacelike hypersurface at each t, then

$$P^\mu_\Sigma = \int_\Sigma \left(T^{\mu 0} + t^{\mu 0} \right) \sqrt{g} \mathrm{d}^3 x$$

satisfies $\frac{\mathrm{d}}{\mathrm{d}t} P^\mu = 0$ as usual:

$$\frac{\mathrm{d}}{\mathrm{d}t} P^\mu = \frac{1}{\Delta t} \left(\int_{\Sigma_{t+\Delta t} - \Sigma_t} \sqrt{g} \mathrm{d}^3 x \left(T^{\mu 0} + t^{\mu 0} \right) \right) = \frac{1}{\Delta t} \int_t^{t+\Delta t} \mathrm{d}t \mathrm{d}^3 x \partial_\mu \left(\sqrt{g} \left(T^{\mu\nu} + t^{\mu\nu} \right) \right)$$

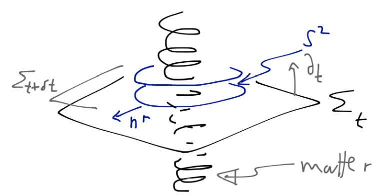
Notice that this requires the matter to be localized (not extending to infinity), so that T vanishes at ∞ and the Stokesing is justified. And in fact a more precise definition of *asymptotically flat* is that t vanishes fast enough at ∞ that we may IBP here with impunity. [40]

P^μ is constant means in particular that $\Delta E_{\text{total}} = 0$ under time evolution:

$$\Delta E_{\text{matter}} = -\Delta E_{\text{grav}} = -\Delta \left(\int t^{00} \sqrt{g} \mathrm{d}^3 x \right).$$

Compute the total gravitational energy inside a sphere, between times t and $t + \Delta t$, by studying the *flux* through the sphere:

$$\Delta E_{\text{grav}} = \Delta t \int_{S^2} t^{0\mu} n_\mu \mathrm{d}^2 x$$



Using (107) and (105), this gives

$$|\Delta E_{\text{matter}}| = \Delta t \frac{G}{45} \left(\frac{\mathrm{d}^3 Q_{ij}}{\mathrm{d}t^3} \frac{\mathrm{d}^3 Q_{ij}}{\mathrm{d}t^3} \right) |_{t-R}, \quad Q_{ij} = q_{ij} - \frac{1}{3} \delta_{ij} q$$

where the $|_{t-R}$ is a reminder to wait until the signal arrives, *i.e.* to use the retarded time. This is the gravitational analog of the Larmor formula in the dipole approximation, $P = \frac{2}{3} \frac{1}{c^3} (\ddot{p})^2$.

For the example of the binary star (two stars of mass M separated by $2r$, hence, with $F_G = Mv^2/r \implies \omega = \sqrt{\frac{G_N M}{4r^3}}$)

$$P = \frac{\Delta E}{\Delta t} = \frac{2}{5} \frac{G_N^4 M^5}{r^5}$$

– that factor of G_N^4 hurts. Using this formula, you can compute the energy loss and hence the rate of change of the frequency as the stars lose energy and spiral in toward each other. Despite the many powers of small quantities, the predictions of this formula were confirmed by observations of Hulse and Taylor.

[40] Warning about the general Stokes' theorem:

$$\int_V \sqrt{g} \nabla_\mu W^\mu = \int_{\partial V} \mathrm{d}S_\mu W^\mu$$

where the measure on the boundary involves a choice of *normal* to the boundary.

10.2.1 Exact gravitational wave solutions exist.

We have found wave solutions in perturbation theory about flat space (to leading order). There are also exact nonlinear solutions describing plane-fronted gravitational waves. Consider Minkowski space:

$$\mathrm{d}s_0^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^1\mathrm{d}y^1 + \mathrm{d}y^2\mathrm{d}y^2$$

and introduce lightcone coordinates $u \equiv t + x, v \equiv t - x$ so that

$$\mathrm{d}s_0^2 = -\mathrm{d}u\mathrm{d}v + \sum_{i=1,2} \mathrm{d}y^i\mathrm{d}y^i.$$

Consider the following deformation:

$$\mathrm{d}s_0^2 \to \mathrm{d}s^2 = -\mathrm{d}u\mathrm{d}v + \sum_{i=1,2} \mathrm{d}y^i\mathrm{d}y^i + F(u, y^i)\mathrm{d}u^2. \tag{109}$$

Here the function F is independent of v, but not necessarily small in any sense. Plug into the vacuum einstein equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} = 8\pi G_N T_{\mu\nu} = 0$$

the trace of which implies that $R = 0$ and hence $R_{\mu\nu} = 0$. I claim that with the ansatz (110)

$$R_{\mu\nu} = 0 \qquad \Leftrightarrow \qquad \sum_i \frac{\partial^2}{\partial y^i \partial y^i} F(u, y^1, y^2) = 0. \tag{110}$$

(!) (110) is solved by

$$F(u, y^i) = \sum_{i,j=1,2} h_{ij}(u) y^i y^j \quad \text{ if } \quad \sum_i h_{ii}(u) = 0 \ .$$

so

$$(h)_{ij} = \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & -h_{11} \end{pmatrix}$$

which become the two linear polarizations $C_+, C_\times$ in the weak-field limit. This is called a 'plane-fronted wave'. It is a bit shocking that the dependence on u is completely unconstrained; this is like for solutions of the 2d wave equation where you get to pick the initial wave profile.

[End of Lecture 17]

11 Time evolution

11.1 Initial value formulation

[Zee, §VI.6, p. 400] Recall our rough counting of degrees of freedom in GR. Let's make this a little more precise, and in particular, let's imagine trying to teach a computer to solve Einstein's equations. To do, so let's back up to the beginning:

Particles. Recall Newton's law for a particle $m\ddot{q} = F(q, \dot{q})$. The fact that this equation is second order in time means you need to specify two initial conditions for each coordinate to determine a solution. Think of these as the initial position and initial momentum of the particle. So, given $q(t), \dot{q}(t)$ at one time, we can determine their values at the next time step $q(t + \mathrm{d}t), \dot{q}(t + \mathrm{d}t)$ by *e.g.*

$$\begin{cases} q(t + \mathrm{d}t) = q(t) + \mathrm{d}t\dot{q}(t) \\ \dot{q}(t + \mathrm{d}t) = \dot{q}(t) + \mathrm{d}t\frac{F(q(t),\dot{q}(t))}{m} \end{cases}$$

and you can do this again and again. Of course this can be optimized in many ways, but here the point is the principle.

Fields. The above can be done also for many particles – just decorate with indices as appropraite:

$$m_{\alpha\beta}\ddot{q}_\alpha = F_\beta\left(\{q, \dot{q}\}\right) \ .$$

A scalar field is just a collection of such variables where the label α includes spatial coordinates. No problem.

We could even be thinking about such fields in curved space. Notice that if we are beginning with a covariant description, we have to pick a time coordinate along which to evolve.

Gauge fields. A small wrinkle arises if one is studying a system in a description involving gauge redundancy. So consider E&M again: the four equations

$$\partial_\mu F^{\mu\nu} = 4\pi j^\nu$$

should determine $A_\mu(t + \mathrm{d}t, x)$ given $A_\mu(t, x), \partial_t A_\mu(t, x), j_\mu(t, x)$ right? In fact, they had better not, because nothing can fix the gauge for us – that is, A_μ and $\tilde{A}_\mu \equiv A_\mu + \partial_\mu\lambda$ have to give the same physics. So really we have only $4 - 1 = 3$ variables to solve for.

Happily, the $\nu = t$ equation is *not* an equation specifying the time evolution of anything. In more familiar langauge, it is the Gauss law constraint:

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho$$

which notice does not contain second time derivatives of A_μ. Rather, it is a *constraint* on possible initial data – the charge density at time t determines the electric field at time t in the familiar way. This means we have fewer equations (3) than we thought (4). But we also have fewer degrees of freedom (3) because the A_μ are determined by physics only up to a gauge transformation. So initial data can be specified by $\vec{A}$ and its field momentum $\vec{E} = \frac{\partial L}{\partial \dot{\vec{A}}}$, subject to Gauss' Law. Notice that if our initial data solves the Gauss law, the time-evolved configuration does, too because

$$\partial_t \left(\vec{\nabla} \cdot \vec{E} - 4\pi\rho \right) = \vec{\nabla} \cdot \underbrace{\left(\partial_t \vec{E}\right)}_{\overset{\text{Ampere}}{=} \vec{\nabla} \times \vec{B} - 4\pi \vec{J}} - 4\pi \underbrace{\partial_t \rho}_{\overset{\text{continuity eqn}}{=} -\vec{\nabla}\cdot\vec{J}} = 0.$$

So before we teach our computer to time-evolve our configuration of E&M fields, we have to teach it to integrate the Gauss law at on our initial time slice.

Gravity. Now we have what seems like 10 equations

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} = 8\pi G_N T^{\mu\nu}$$

for what seems like 10 objects $\partial_t^2 g_{\mu\nu}$ required to determine the metric components $g_{\mu\nu}(t + \mathrm{d}t, x)$ and its time derivatives $\partial_t g_{\mu\nu}(t + \mathrm{d}t, x)$ at the next time step in terms of their present values and the stress tensor $T^{\mu\nu}(x)$.

Again, the Einstein equations *don't* determine this many quantities. By general covariance of the Einstein equations, four objects must not be determined, since we can make the four replacements $x^\mu \to \tilde{x}^\mu(x)$.

And in fact, the four equations with a time index

$$G^{t\nu} = 16\pi G_N T^{t\nu}$$

are not evolution equations, but constraints, in that:
Claim: $G^{t\nu}$ does not contain $\partial_t^2 g_{\mu\nu}$.
Proof: (a) write it out explicitly and check.
(b) Use $\nabla_\mu G^{\mu\nu} = 0$ to derive a contradiction – if there were second time derivatives in $G^{t\nu}$, there would be third time derivatives in the quantity

$$\partial_t G^{t\nu} = -\partial_i G^{i\nu} + \left(\Gamma^{\cdot}_{\cdot\cdot} G^{\cdot\cdot}\right)^\nu$$

which you can see from the RHS of the identity there just aren't. (Note that it *does* contain terms like $(\partial_t g_{\cdot\cdot})^2$, but this is useless for determining $\partial_t^2 g_{\cdot\cdot}$.)

Comments:

1. In a initial-value formulation, we are forced to distinguish the time coordinate. So we might as well choose a gauge (like $A_0 = 0$ gauge) where we get rid of $g_{t\mu}$, components

of the metric with a time component. So we need to specify g_{ij} on the initial-time slice (ij are spatial indices), and its time derivative $\partial_t g_{ij}$. A more covariant expression for the latter is the *extrinsic curvature*, K_{ij}, of the spatial slice Σ_t. I claim without proof that it is the field momentum of the gravitational field, $K_{ij} = \frac{\partial L}{\partial \dot{g}_{ij}}$. More explicitly, a definition of the extrinsic curvature is

$$K_{ij} = \nabla_i \xi_j$$

where ξ_j is a unit normal to the surface Σ_t, *i.e.* $\xi = \frac{1}{\|\partial_t\|}\partial_t$. If the metric takes the form $\mathrm{d}s^2 = -g_{tt}\mathrm{d}t^2 + \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j$, then $\xi = \frac{1}{\sqrt{g_{tt}}}\partial_t$ and

$$K_{ij} = \Gamma^t_{ij} g_{tt} \frac{1}{\sqrt{g_{tt}}} = \frac{1}{2}\frac{1}{\sqrt{g_{tt}}}\partial_t g_{ij}$$

so it is indeed basically the time derivative of the spatial metric.

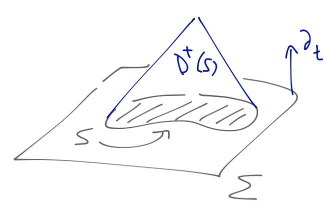
The extrinsic curvature can also be defined in terms of the Lie derivative,

$$K_{ij} = \frac{1}{2}\mathcal{L}_t g_{ij} \ .$$

2. If we only know the initial data on a patch $S \subset \Sigma_t$ of the initial time slice, then we can only time evolve into $D^+(S)$, a region of spacetime which we define next.

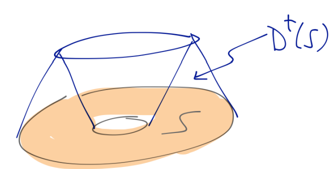
11.2 Causal structure of spacetime.

We are now in possession of a very interesting system of coupled differential equations. One part of this system is: given $g_{\mu\nu}$, solve the matter evolution (which of course will produce a stress tensor and thereby a new metric and so on). Suppose there exists a spacelike hypersurface Σ and some time coordinate orthogonal to Σ. Imagine you can solve for the matter field evolution given $g_{\mu\nu}$. How much of the future region can we control by controlling stuff in a region $S \subset \Sigma$?



This is called the "future domain of dependence" of S,

$$D^+(S) \equiv \begin{array}{c}\text{a region of spacetime such that } \forall p \in D^+(S), \\ \text{every past-moving timelike or null (inextendible) curve} \\ \text{starting at } p \text{ passes through } S.\end{array}$$



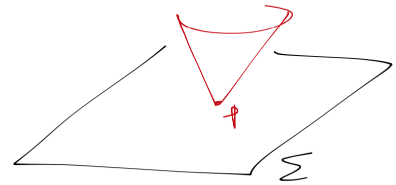
Note that if S has holes, you'll get some weird-looking stuff.

No point in the complement $\Sigma - S$ influences stuff in $D^+(S)$. The boundary of this region

$$H^+(S) \equiv \partial D^+(S) - S$$

is called the "future Cauchy horizon" of S; these points are just barely in the region of influence of S. We can similarly define $D^-(S)$ (and thereby $H^-(S)$) as the set of points which can possibly influence goings-on in S by the transmission of massive *or* massless particles: $D^-(S)$ is the stuff you can see from S.

Note that if in solving the GR Cauchy problem starting at Σ we encounter a singularity at some spacetime point p we must remove its region of influence (any $D^+(S)$ for S which contains p) from our considerations since if we don't know what goes on at p we can't figure out what happens in its domain of influence.



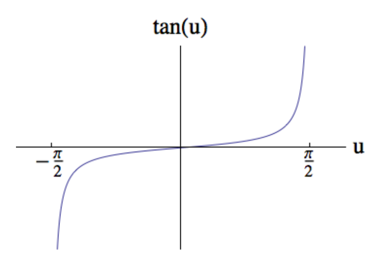
Penrose/conformal diagrams. Our next job will be to think about black holes. But which solutions of Einstein's equations should be considered black holes depends on the boundary conditions – on the asymptotics of the spacetime in which the black hole finds itself. In particular, we will discuss black holes in asymptotically flat spacetime. It will be useful first to characterize this in a coordinate invariant way. What we mean is that at ∞, the spacetime looks like $\mathbb{R}^{3,1}$ at its coordinate infinity. What is this? The usual rectilinear coordinates are not good coordinates there. In trying to appreciate the asymptotics of such a spacetime it really helps if you can draw the whole thing on a piece of paper. This is the point of a *Penrose diagram*, which is constructed by finding a set of coordinates whose range is finite and in which light goes on 45° lines (at the price of some horrible overall prefactor in the metric).

Consider $1+1$ dimensions for a moment:

$$\begin{aligned} \mathrm{d}s^2_{1,1} &= -\mathrm{d}t^2 + \mathrm{d}x^2, \quad -\infty < t, x < \infty \\ &= -\mathrm{d}u\mathrm{d}v, \quad u \equiv t+x, v = t-x, -\infty < u, v < \infty \end{aligned} \tag{111}$$

Now let us shrink the range of the lightcone coordinates to a finite region:

$$u = \tan\tilde{u}, \quad v = \tan\tilde{v}, \qquad -\pi/2 \leq \tilde{u}, \tilde{v} \leq \pi/2.$$

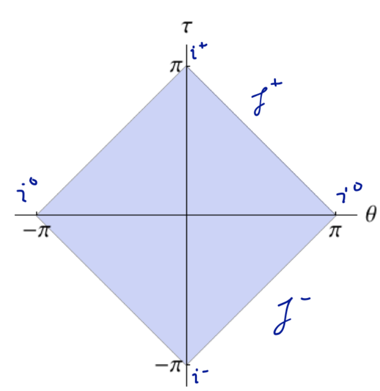


$u = \infty$ is $\tilde{u} = -\pi/2$. The metric is

$$\mathrm{d}s^2_{1,1} = -\frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}\mathrm{d}\tilde{u}\mathrm{d}\tilde{v}.$$

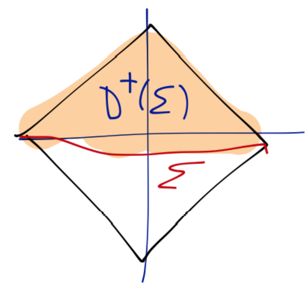
Now if we define $\tilde{u} = (\tau+\theta)/2, \quad \tilde{v} = (\tau-\theta)/2$

$$\mathrm{d}s^2_{1,1} = \frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}\left(-\mathrm{d}\tau^2 + \mathrm{d}\theta^2\right)$$



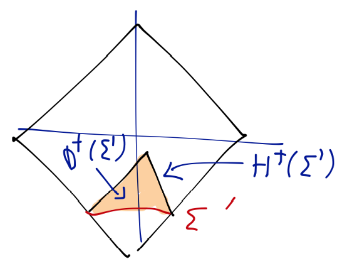
The range of these coordinates is as indicated in the figure. Notice that the new metric is the same as the original flat metric except for the overall factor $\Omega(u,v) = \frac{1}{\cos^2\tilde{u}\cos^2\tilde{v}}$. Erasing that factor does not change angles between vectors – the metrics are conformal to each other. The lines of constant $\tilde{v}$ are also lines of constant v, which means they are lightlike – the conformal transformation (which gets rid of $\Omega(u,v)$) preserves lightcones. That diagram is called a *Penrose (Newmann-Carter conformal) diagram.* Its defining property is that light moves on $45°$ lines in such a diagram. Given some subspace S of a spacelike hypersurface, this makes it easy to determine what is its domain of dependence $D^+(S)$.

Examples. $D^+(\Sigma)$ is the whole future of the spacetime. But if you specify data on Σ', you don't know what happens after you hit the Cauchy horizon $H^+(\Sigma')$.

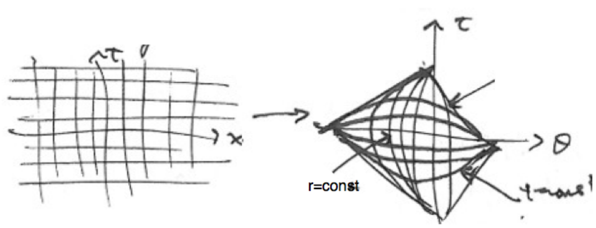


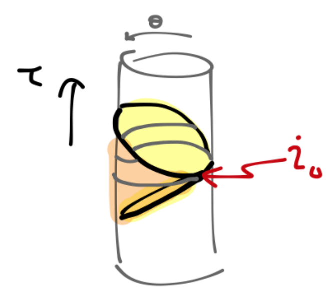
Def: a spacetime M is *asymptotically flat* if it has the same conformal structure as Minkowski space at ∞.

Since angles are preserved, lines of constant τ and constant θ are perpendicular.



Notice that the points at 'spacelike infinity' i_0, at $\theta = \pi, -\pi$, ee The locus of points where lightrays all end up is called $\mathscr{I}^+$, pronounced "scri plus" (I think it's short for 'script i').



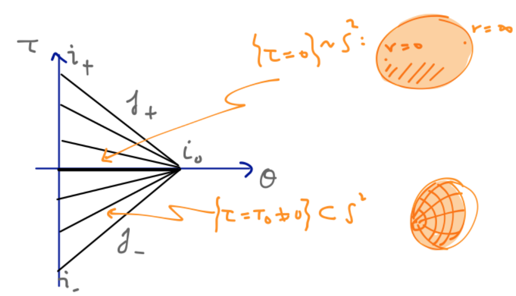


The procedure for constructing the Penrose diagram for $D = 2+1$ and $D = 3+1$ Minkowski space is similar.

$$\begin{aligned}
\mathrm{d}s^2_{2+1} &= -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 \\
&= -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2, \qquad 0 \le r \le \infty, -\infty \le t \le \infty, \underbrace{0 \le \varphi < 2\pi}_{OK} \\
u \equiv t + r, v \equiv t - r \quad & \\
&= -\mathrm{d}u\mathrm{d}v + \frac{1}{4}(u - v)^2\mathrm{d}\varphi^2 \\
u \equiv \tan\tilde{u}, v \equiv \tan\tilde{v} \quad & \\
&= \frac{1}{4\cos^2\tilde{u}\cos^2\tilde{v}}\left(-\mathrm{d}\tilde{u}\mathrm{d}\tilde{v} + \frac{1}{4}\sin^2(\tilde{u} - \tilde{v})\mathrm{d}\varphi^2\right) \\
\tilde{u} \equiv \frac{1}{2}(\tau + \theta), \tilde{v} \equiv \frac{1}{2}(\tau - \theta) \quad & \\
&= \frac{1}{4\cos^2\tilde{u}\cos^2\tilde{v}}\left(-\mathrm{d}\tau^2 + \underbrace{\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2}_{=\mathrm{d}s^2_{S^2}}\right) \qquad (112)
\end{aligned}$$

$$t \pm r = \tan\left(\frac{\tau \pm \theta}{2}\right), \tau \pm \theta \in (-\pi, \pi), \theta \in (0, \pi).$$

We can embed this in $\mathbb{R} \times S^2$. The locus $\tau = 0$ (a line in the figure, since we aren't drawing the φ direction) is actually a whole S^2. The locus of constant $\tau = \tau_0 \neq 0$ is part of an S^2: $\{-\pi < \tau_0 \pm \theta < \pi, 0 < \theta < \pi, \varphi \in [0, 2\pi)\}$ (just like in the $D = 1+1$ case, the locus $\tau = \tau_0 \neq 0$ was part of the circle S^1).



In 3+1d:

$$\begin{aligned}
\mathrm{d}s^2_{3+1} &= -\mathrm{d}t^2 + \mathrm{d}\vec{x}^2 \\
&= -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right), \\
& \qquad 0 \le r \le \infty, -\infty \le t \le \infty, \underbrace{0 \le \theta \le \pi, 0 \le \varphi < 2\pi}_{OK} \\
t \pm r \equiv \tan\left(\frac{\tau \pm \rho}{2}\right) \quad & \\
&= \frac{1}{4\cos^2\frac{\tau+\rho}{2}\cos^2\frac{\tau-\rho}{2}}\left(-\mathrm{d}\tau^2 + \underbrace{\mathrm{d}\rho^2 + \sin^2\rho\mathrm{d}s^2_{S^2}}_{=\mathrm{d}s^2_{S^3}}\right) \qquad (113)
\end{aligned}$$

so we can embed this one in $\mathbb{R} \times S^3$. Without the conformal factor, this is called the "Einstein static universe" and is a(n unstable) solution of $G_{\mu\nu} = \Lambda T_{\mu\nu}$ with $\Lambda > 0$.

[End of Lecture 18]

12 Schwarzschild black hole solution

Schwarzschild ('black sign'), who discovered black holes even before Einstein figured out the correct field equations (!) is someone we can add to our list of appropriately-named physicists (Poynting, Killing, Cutkowski, Cubitt, D'Eath and Payne, ...).

12.1 Birkhoff theorem on spherically-symmetric vacuum solutions

We will study solutions of the vacuum Einstein's equations $G_{\mu\nu} = 0$. By trace-reversal, this implies $R_{\mu\nu} = 0$ – the spacetime is Ricci flat. The general solution is certainly not known (even in euclidean signature, it is not known; in that case, solutions include *Calabi-Yau manifolds* which are useful in the context of string compactification). We will make two more demands: we will demand that the spacetime be asymptotically Minkowski. And we will demand spherical symmetry.

The solution we will find will also be *static*; that is a conclusion, not an assumption. This means that the unique spherical solution we will find (given asymptotically-flat boundary conditions) amounts to a GR generalization of Newton's 20-year theorem on the gravitational effects outside a spherical mass distribution (it is called Birkhoff's theorem). A useful consequence is: in a process of spherically-symmetric gravitational collapse, the solution we have described is the correct solution at all times outside the region where the matter density is nonzero. (The analogous thing happens in Maxwell theory: spherical symmetry implies that the vector potential is time-independent and the only solution is the Coulomb field.)

So we could start from the spherical and static ansatz

$$\mathrm{d}s^2 = r^2 \underbrace{\left(\mathrm{d}s^2_{S^2}\right)}_{\equiv \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2} + e^{2b(r)}\mathrm{d}r^2 - e^{2a(r)}\mathrm{d}t^2$$

and just plug it into $R_{\mu\nu} = 0$, but the following is more instructive.

Symmetries of spacetime. What do we mean by a symmetry of spacetime (without a priori coordinates)? Given some metric $g_{\mu\nu}(x)$, its change at x under the flow generated by a vector field ξ is (problem set 8)

$$\delta g_{\mu\nu}(x) = \mathcal{L}_\xi g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = -\left(\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu\right)$$

and if this vanishes then ξ is said to be a Killing vector field (Kvf) in that spacetime. Further claim: if $\xi^{(1)}$ and $\xi^{(2)}$ are Kvfs then so is their Lie bracket $[\xi^{(1)}, \xi^{(2)}]$. So the Kvfs of a spacetime form a Lie algebra; exponentiating their action produces a Lie group whose elements are the associated flow maps.

Spherical symmetry means that the manifold admits a set of three Kvfs $\xi^{(a)}, a = 1, 2, 3$

satisfying the so(3) Lie algebra:

$$[\xi^{(a)}, \xi^{(b)}] = \epsilon^{abc}\xi^{(c)}$$

that we found on the 2-sphere. By the Frobenius theorem (see Appendix B of Wald), the existence of such Kvfs means that we can *foliate* the manifold by the orbits of the Kvfs, which in this case are S^2s. If the manifold is 4-dimensional, this means that the metric must be of the form:

$$\mathrm{d}s^2 = \mathrm{d}s^2_{S^2} r^2(a,b) + \mathrm{d}\tilde{s}^2(a,b)$$

– the coordinates a, b parametrize a two-dimensional family of two-spheres. In particular, there can be no $\mathrm{d}a\mathrm{d}\theta$ cross-terms; that would require an $SO(3)$-invariant vector field v on the S^2 out of which to build $\mathrm{d}a\,(v_\theta \mathrm{d}\theta + v_\varphi \mathrm{d}\varphi)$. Such a v does not exist.

Inexorable logic. From here it's downhill to the Schwarzschild solution.

• First, choose $r(a,b) = r$ itself to be one of the two other coordinates. We are *defining* the coordinate r to be the radius of the spherical sections (at fixed t). (Here "radius" means the r in the area $= 4\pi r^2$.) You could worry that it's not monotonic or all the spheres are the same size, but this is not a real concern, since we will demand that asymptotically the solution approach Minkowski space, where r is the radial polar coordinate. So the most general metric now is

$$\mathrm{d}s^2 = r^2 \mathrm{d}s^2_{S^2} + g_{aa}(r,a)\mathrm{d}a^2 + 2g_{ar}(r,a)\mathrm{d}r\mathrm{d}a + g_{rr}(r,a)\mathrm{d}r^2 \ .$$

(Notice that a must be the time coordinate.)

• We can get rid of the cross term by choosing $t(a,r)$ appropriately.

Let's argue backwards. Suppose we start with a metric of the diagonal form:

$$\mathrm{d}\tilde{s}^2 = -e^{-2\alpha(r,t)}\mathrm{d}t^2 + e^{2\beta(r,t)}\mathrm{d}r^2$$

and change coordinates by $t = t(a,r)$. Then

$$\mathrm{d}\tilde{s}^2 = -e^{-2\alpha(r,t)}\left(\partial_a t\right)^2 \mathrm{d}a^2 - 2e^{2\alpha(r,t)}\partial_a t \partial_r t \mathrm{d}a\mathrm{d}r + \left(e^{2\beta(r,t)} - e^{2\alpha(r,t)}\left(\partial_r t\right)^2\right)\mathrm{d}r^2$$

So we can use the three freedoms $\alpha(r,t), \beta(r,t), t(a,r)$ to reproduce the previous $g_{rr}(r,a), g_{ar}(r,a), g_{aa}(r,a)$.

$$\implies \qquad \mathrm{d}s^2 = r^2\mathrm{d}s^2_{S^2} - e^{-2\alpha(r,t)}\mathrm{d}t^2 + e^{2\beta(r,t)}\mathrm{d}r^2$$

• So far we have only used spherical symmetry and our ability to adapt our coordinates to meet our needs. Now, finally, we impose $R_{\mu\nu} = 0$ on our choice of $\alpha(r,t), \beta(r,t)$. There isn't really any good reason not to employ a computer to determine $\Gamma^\rho_{\mu\nu}[\alpha,\beta]$ and thereby $R_{\mu\nu\rho}{}^\sigma[\alpha,\beta]$ and thereby $R_{\mu\nu}[\alpha,\beta]$. A choice selection of components of the latter are:

$$R_{tr} = \frac{2}{r}\dot{\beta} \qquad (114)$$

$$R_{\theta\theta} = e^{-2\beta(r,t)} \left(r(\beta' - \alpha') - 1 \right) + 1 \tag{115}$$

where $\dot{} \equiv \partial_t, ' \equiv \partial_r$. So (114) immediately tells us that away from points where the sphere shrinks to $r = 0$,

$$\dot{\beta} = 0 \implies \beta = \beta(r).$$

And then ∂_t(115) implies

$$0 = e^{-2\beta} \left(r \partial_t \partial_r \alpha \right) \implies 0 = \partial_r \partial_t \alpha \implies \alpha(r,t) = \alpha(r) + \gamma(t).$$

- We can absorb this $\gamma(t)$ into a redefinition of our time coordinate:

$$\mathrm{d}s^2 = -e^{2\alpha(r,t)} \mathrm{d}t^2 + \cdots = -e^{2\alpha(r)} \underbrace{e^{2\gamma(t)} \mathrm{d}t^2}_{\equiv d\tilde{t}^2} + \ldots$$

$$\implies \frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = e^{\gamma(t)} \implies \tilde{t}(t) = \int^t \mathrm{d}t' e^{\gamma(t')} \ .$$

Now drop the tildes and we are home: the most general spherically symmetric solution of $R_{\mu\nu} = 0$ has the form:

$$\mathrm{d}s^2 = r^2 \left(\mathrm{d}\theta^2 + \sin^2 \theta \mathrm{d}\varphi^2 \right) + e^{2\beta(r)} \mathrm{d}r^2 - e^{2\alpha(r)} \mathrm{d}t^2$$

So a spherically-symmetric solution of $R_{\mu\nu} = 0$ is also *static*. This means that there exist coordinates with $\partial_t g_{\mu\nu} = 0$ *and* no cross-terms: $g_{\mu t} = 0$. (A spacetime is merely *stationary* if $g_{\mu t} \neq 0$ in coordinates where the metric is time-independent; the Kerr solution for a rotating black hole is an example of this. Fancier concise statement: stationary means merely that there exists a timelike Kvf; static means stationary and foliated by spacelike submanifolds normal to the Kvf.)

The demand of asymptotic flatness imposes the boundary conditions

$$\alpha(r) \stackrel{r \to \infty}{\to} 0, \beta(r) \stackrel{r \to \infty}{\to} 0 \ .$$

The remaining Ricci-flat equations are usefully organized into

$$0 = R^r_r - R^t_t = \frac{2}{r} \left(\alpha' + \beta' \right) \qquad \implies \partial_r \left(\alpha + \beta \right) = 0 \quad \implies \alpha(r) = -\beta(r) + \underbrace{c}_{\text{absorb into } t \to e^{-c} t} \tag{116}$$

So $\alpha(r) = -\beta(r)$. The final contentful equation is:

$$\begin{aligned} 0 \quad &= \quad R_{\theta\theta} = e^{-2\beta} \left(r \left(\beta' - \alpha' \right) - 1 \right) + 1 \\ &\stackrel{(116)}{=} \quad e^{2\alpha} \left(r(-2\alpha') - 1 \right) + 1 = -\partial_r \left(r e^{2\alpha} \right) + 1 \\ &\qquad \implies r e^{2\alpha} = r - r_0 \end{aligned} \tag{117}$$

where r_0 is an integration constant.

$$e^{2\alpha(r)} = 1 - \frac{r_0}{r}.$$

You can check that this also solves $R_{tt} = 0$ and $R_{rr} = 0$ independently[41].

$$\boxed{\mathrm{d}s^2_{\mathrm{sch}} = -\left(1 - \frac{r_0}{r}\right)\mathrm{d}t^2 + \frac{1}{1-\frac{r_0}{r}}\mathrm{d}r^2 + r^2\mathrm{d}s^2_{S^2}}$$

is the Schwarzschild metric. When $r_0 \to 0$, this is just Minkowski space. As we've seen, the constant r_0 is related to the mass of the source $r_0 = 2G_N M$ by matching to the Newtonian limit.

12.1.1 Appendix: tetrad method assuming spherical static ansatz

Assumptions: we'll look for a solution of Einstein's equations in vacuum $T_{\mu\nu} = 0$ (for $r \neq 0$) which is spherically symmetric and static. Ansatz:

$$\mathrm{d}s^2 = r^2 \underbrace{\left(\mathrm{d}s^2_{S^2}\right)}_{=\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2} + e^{2b(r)}\mathrm{d}r^2 - e^{2a(r)}\mathrm{d}t^2.$$

In fact, this ansatz is completely general given the symmetry assumptions, as should be clear from the earlier discussion.

We want to find the functions $a(r), b(r)$ by solving Einstein's equations. We're going to need $R_{\mu\nu}, R$; let's construct them using the tetrad formalism, Eqs (89), (90).

---

1. find ON basis:

$$\theta^{\hat{t}} = e^{a(r)}\mathrm{d}t, \theta^{\hat{\theta}} = r\mathrm{d}\theta, \theta^{\hat{\varphi}} = r\sin\theta\mathrm{d}\varphi, \theta^{\hat{r}} = e^{b(r)}\mathrm{d}r.$$

2. Calculate connection one-forms ${\omega_b}^a$ using the first Cartan structure equation (89):

$$\mathrm{d}\theta^a = -{\omega_b}^a \wedge \theta^b \ .$$

$$\mathrm{d}\theta^{\hat{t}} = a'(r)e^a\mathrm{d}r \wedge \mathrm{d}t = a'e^a e^{-b}e^{-a}\theta^{\hat{r}} \wedge \theta^{\hat{t}} = -a'e^{-b}\theta^{\hat{t}} \wedge \theta^{\hat{r}}$$

[41] For completeness, their form in the static ansatz is

$$\begin{aligned} R_{tt} &= e^{2(\alpha-\beta)}\left(\alpha'' + (\alpha')^2 - \alpha'\beta' + \frac{2}{r}\alpha'\right) \\ R_{rr} &= -\alpha'' - (\alpha')^2 + \alpha'\beta' + \frac{2}{r}\beta'. \end{aligned} \tag{118}$$

$$\implies \omega_{\hat{r}}{}^{\hat{t}} = a' e^{-b}\theta^{\hat{t}}, \quad \omega_{\hat{\theta}}{}^{\hat{t}} \wedge \theta^{\hat{\theta}} = 0, \omega_{\hat{\varphi}}{}^{\hat{t}} \wedge \theta^{\hat{\varphi}} = 0.$$

$$\mathrm{d}\theta^{\hat{\theta}} = \mathrm{d}r \wedge \mathrm{d}\theta = \frac{1}{r}e^{-b}\theta^{\hat{r}} \wedge \theta^{\hat{\theta}} = -\frac{1}{r}e^{-b}\theta^{\hat{\theta}} \wedge \theta^{\hat{r}} \qquad \implies \omega_{\hat{r}}{}^{\hat{\theta}} = -\frac{e^{-b}}{r}\theta^{\hat{\theta}}$$

$$\mathrm{d}\theta^{\hat{\varphi}} = \sin\theta \mathrm{d}r \wedge \mathrm{d}\varphi + r\cos\theta \mathrm{d}\theta \wedge \mathrm{d}\varphi = \frac{\sin\theta}{r}e^{-b}\theta^{\hat{r}} \wedge \theta^{\hat{\theta}} + \frac{\cot\theta}{r}\theta^{\hat{\theta}} \wedge \theta^{\hat{\varphi}}$$

$$\implies \omega_{\hat{r}}{}^{\hat{\varphi}} = \frac{e^{-b}}{r}\theta^{\hat{\varphi}}, \omega_{\hat{\theta}}{}^{\hat{\varphi}} = \frac{1}{r}\cot\theta\theta^{\hat{\varphi}}$$

3. Calculate the curvature two-forms with the second Cartan structure equation (90):

$$\Omega_b{}^a = -\mathrm{d}\omega_b{}^a + \omega_b{}^c \wedge \omega_c{}^a$$

$$\Omega_{\hat{r}}{}^{\hat{t}} = \mathrm{d}\left(a' e^{-b}\theta^{\hat{t}}\right) + 0$$

– the $\omega \wedge \omega$ terms all vanish either by antisymmetry of forms or by vanishing of ωs.

$$\Omega_{\hat{r}}{}^{\hat{t}} = \left(a'' - a'b' + a'^2\right) e^{-2b}\theta^{\hat{r}} \wedge \theta^{\hat{t}}.$$

$$\Omega_{\hat{\theta}}{}^{\hat{t}} = \omega_{\hat{\theta}}{}^{\hat{r}} \wedge \omega_{\hat{r}}{}^{\hat{t}} = -\frac{a'}{r}e^{-2b}\theta^{\hat{t}} \wedge \theta^{\hat{\theta}}$$

$$\Omega_{\hat{\varphi}}{}^{\hat{t}} = -\frac{a'}{r}e^{-2b}\theta^{\hat{t}} \wedge \theta^{\hat{\varphi}}$$

$$\Omega_{\hat{\theta}}{}^{\hat{r}} = -\mathrm{d}\left(\frac{1}{r}e^{-b}\right)\theta^{\hat{\theta}} - \frac{1}{r}e^{-b}\mathrm{d}\theta^{\hat{\theta}} + \omega_{\hat{\theta}}{}^{\hat{\varphi}} \wedge \omega_{\hat{\varphi}}{}^{\hat{r}}$$

$$= \frac{b'}{r}e^{-2b}\theta^{\hat{r}} \wedge \theta^{\hat{\theta}}.$$

$$\Omega^{\hat{r}}_{\hat{\varphi}} = \frac{b'}{r}e^{-2b}\theta^{\hat{r}} \wedge \theta^{\hat{\theta}}$$

$$\Omega^{\hat{\theta}}_{\hat{\varphi}} = (1 - e^{-2b})r^{-2}\theta^{\hat{\theta}} \wedge \theta^{\hat{\varphi}}.$$

Other nonzero components are determined by $\Omega_b{}^a = -\Omega_a{}^b$.

4. Extract the Riemann tensor using

$$\Omega_b{}^a = R_{bcd}{}^a\theta^c \wedge \theta^d$$

$$\implies R^t_{rtr} = \left(a'' - a'b' + a'^2\right) e^{-2b}$$

Einstein equations: $G^{\mu}_{\nu} = 0$.

$$0 = G^t_t - G^r_r = -2(b' + a')r^{-1}e^{-2b}$$

Boundary conditions: $g_{\mu\nu} \to \eta_{\mu\nu}$ as $r \to \infty$ – we are looking for an *asymptotically flat* solution[42]. In our ansatz, this means $a(r) \to 0, b(r) \to 0$ as $r \to \infty$. So

$$0 = a' + b' \implies a(r) = -b(r).$$

$$0 = G^r_r = 2a'r^{-1}e^{2a} - r^{-2}(1 - e^{2a}).$$
$$\implies 1 = e^{2a}(2a'r + 1) = (re^{2a})'$$
$$\implies re^{2a} = r - r_0$$

where r_0 is a well-named integration constant.

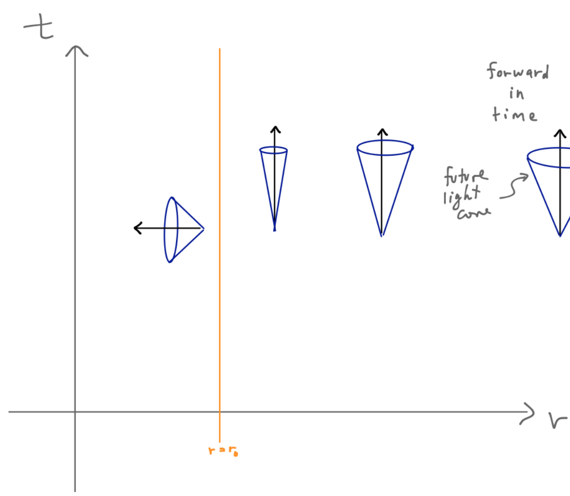
$$e^{2a} = 1 - \frac{r_0}{r}, e^{2b} = \frac{1}{1 - \frac{r_0}{r}}$$

$$\boxed{\mathrm{d}s^2_{\text{sch}} = -\left(1 - \frac{r_0}{r}\right)\mathrm{d}t^2 + \frac{1}{1 - \frac{r_0}{r}}\mathrm{d}r^2 + r^2\mathrm{d}s^2_{S^2}}$$

is the Schwarzschild metric.

12.2 Properties of the Schwarzschild solution

Causal structure



$r = r_0$ is called the *Schwarzschild radius*. Consider the path of light rays with no angular momentum ($\mathrm{d}\theta = \mathrm{d}\varphi = 0$; this simplification doesn't affect the conclusion). As $r \to r_0$ from above,

$$0 = \mathrm{d}s^2 \implies \mathrm{d}t = \frac{1}{1 - \frac{r_0}{r}}\mathrm{d}r$$

(we choose the positive root because we want the forward light cone) – $\mathrm{d}t$ gets huge. It

[42] You might be unhappy with this specification, and insist on a manifestly coordinate-independent definition of what it means for the spacetime to be asymptotically flat. The real demand we are making is that the spacetime has the same *causal structure* as Minkowski spacetime at ∞. This is a kettle of fish.

takes longer and longer to move a fixed radial coordinate distance dr. The lightcones close up; as we pass through $r = r_0$, they flip over: for $r < r_0$, t becomes spacelike and r becomes timelike! Nothing can escape $r < r_0$. Hence the name 'black hole'.

Plugging in numbers

We have already seen that in the Newtonian limit, $g_{tt} = -(1 + 2\phi(r))$, so if this metric results from a collection of mass centered at $r = 0$, we have the Newtonian potential

$$\phi = -\frac{G_N M}{r} \quad \Longrightarrow r_0 = 2G_N M.$$

So as the object gets heavier, the Schwarzschild radius grows (linearly in $D = 3 + 1$ dimensions).

The sun has $M_{\text{sun}} \sim 10^{57}\text{GeV}$ and $R_{\text{sun}} \sim 10^6\text{km}$. The Schwarzschild radius for this mass is

$$r_0(s) = 2G_N M_{\text{sun}} \sim 2(10^{-38}\text{GeV}^{-2})(10^{57}\text{GeV}) \sim 10^{19}\text{GeV}^{-1} \sim 10^5\text{cm} \sim 1\text{km}.$$

So the mass density of a solar mass black hole is $\rho_{BH} \sim 10^{18}\rho_{\text{sun}} \sim 1\text{GeV}^4$. Nuclear density is $\rho_{\text{nuc}} \sim \frac{1\text{GeV}}{1\text{fermi}^3} \sim 10^{-3}\text{GeV}^4$. On the other hand, it's not clear that "density" is the right way to think about what's inside a black hole.

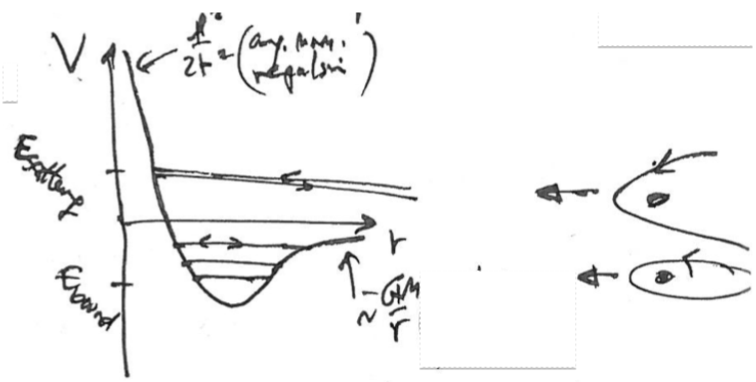
12.2.1 Light-bending and perihelion shift.

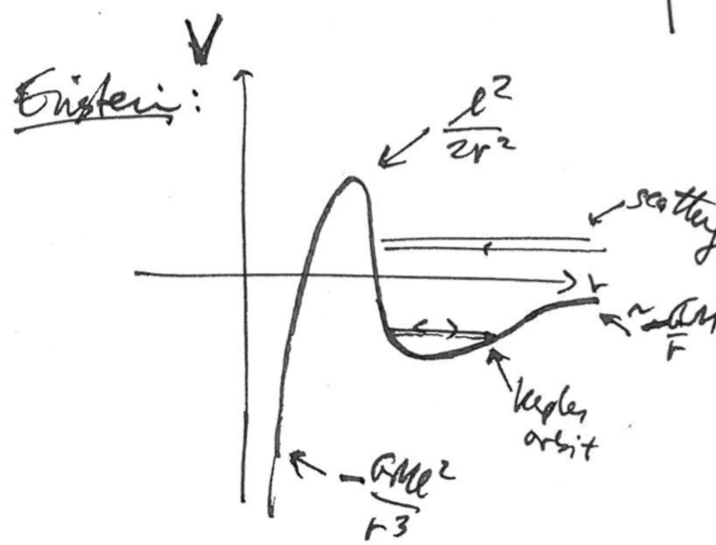
[Zee §VII.1] Recall that we already discussed geodesics in the Schwarzschild geometry (at $r > r_0$) in §6, and found that they are governed by an ordinary mechanics problem with the potential

$$V(r) = -\frac{\kappa G_N M}{r} + \frac{1}{2}\frac{L^2}{r^2} - \frac{G_N M L^2}{r^3}.$$

Here L is the angular momentum about $r = 0$. κ is speed of the affine parameter: more precisely, $\kappa = -g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$, and we should set $\kappa = 1$ to match the Newtonian energy conservation equation $\frac{p^2}{2m} + V = E$ for a massive particle. I leave κ in this expression because it is zero if our probe particle is massless. WLOG we are studying orbits in the equatorial $\theta = \pi/2$ plane. The last term is the only difference from Newtonian gravity. The following figures apply to the case of a massive probe particle ($\kappa \neq 0$)

Newton:





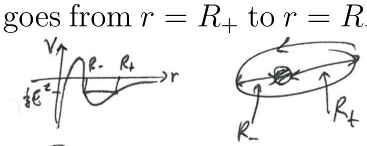
The fact that orbits in the Newtonian potential are closed ellipses is special; the r^{-3} perturbation makes it not so. Eliminating s from the equations

$$\frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 + V(r) = \frac{1}{2}\epsilon^2, \quad r^2\frac{\mathrm{d}\varphi}{\mathrm{d}s} = L$$

(recall that ϵ is the conserved energy along the worldline) gives

$$\frac{1}{2}\frac{L^2}{r^4}\left(\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 + V(r) = \frac{1}{2}\epsilon^2 \implies \varphi(r) = \pm\int^r \mathrm{d}\tilde{r}\frac{L}{\tilde{r}^2\sqrt{\epsilon^2 - 2V(\tilde{r})}}. \tag{119}$$

A closed orbit (this requires $L^2 > 12G_N^2M^2$) has $\Delta\varphi = \pi$ when r goes from $r = R_+$ to $r = R_-$

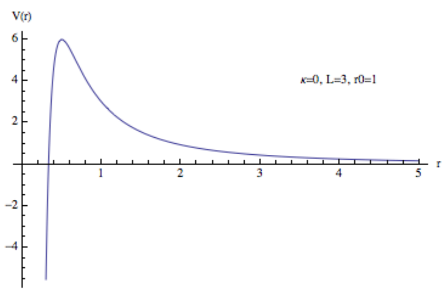


once, where $R_\pm$ are the perihelion and ...farthest-from-helion:

The perihelion shift is[43]

$$\Delta\varphi = 2\left(\varphi(R_+) - \varphi(R_-)\right) - 2\pi \simeq 3\pi G_N M \left(\frac{1}{R_+} + \frac{1}{R_-}\right) = \frac{6\pi G_N M}{(1-e^2)},$$

where $R_\pm = 1 \pm e$, e is the eccentricity.



Next, **Light-bending:** There are only two small differences from the above discussion. First is that for massless particles such as light, the terms with κ are absent. The second is that this implies only scattering (unbound) orbits. With $M = 0$, a geodesic

is a straight line:
For nonzero M, the $\frac{1}{r^3}$ term is attractive; the amount

of bending can be quantified by the angle $\Delta\varphi$ like this:
The formula above for the angle-change (between $R = R_0$ an $R = \infty$ now) gives:

$$\Delta\varphi = 2\int_{R_0}^{\infty} \frac{L dr}{r^2\sqrt{E^2 - 2V(r)}} - \pi \simeq \frac{4G_N M}{R_0}.$$

Two comments: (1) as in the Newton theory, the equations simplify a bit in terms of $u \equiv 1/r$. (2) if we're just interested in the leading correction, we can just use the leading term in perturbation theory.

[Zee page 372] In terms of $u = \frac{1}{r}$, the GR version of Kepler's law (119) is

$$\epsilon^2 = L^2\left((u')^2 + u^2\right) - r_0 u + 1 - r_0 L^2 u^3 \tag{120}$$

To proceed from here we can observe that ignoring the GR correction (the last term on the right), this is the equation for the energy of a harmonic oscillator. More explicitly, we can differentiate (120) with respect to φ and divide by $u' \equiv \partial_\varphi u$ and L^2 to get

$$u'' + u = \frac{G_N M}{L^2} + 3G_N M u^2$$

[43] The Newtonian result here relies on the integral

$$\int_{u_-}^{u_+} \frac{\mathrm{d}u}{\sqrt{(u-u_-)(u_+-u)}} = \pi$$

which arises by the subsitution $u = \frac{1}{r}$.

The last term is the GR correction. (The first term on the RHS is absent for massless particles.) The Newtonian Kepler orbit is

$$u_0 = \frac{G_N M}{L^2}\left(1 + e\cos\varphi\right) = \frac{1}{P}(1 + e\cos\varphi)$$

where P is the perihelion distance and e is the eccentricity of the orbit. Perturbation theory develops by setting $u = u_0 + u_1$ and linearizing in u_1, which gives:

$$u_1'' + u_1 = \frac{3G_N M}{P^2}\left(1 + 2e\cos\varphi + e^2\cos^2\varphi\right) \ .$$

The constant sources just perturb the radius of the orbit; we also don't care about periodic terms in u_1, since they don't cause precession. The interesting bit is the not periodic bit:

$$u_1 = \frac{3eG_N M^3}{L^4}\varphi\sin\varphi + \ldots$$

which grows as you orbit more times; altogether

$$u(\varphi) \simeq \frac{1}{P}\left(1 + e\cos\varphi + \frac{3eG_N M^3 P}{L^4}\varphi\sin\varphi\right) \simeq \frac{1}{P}\left(1 + e\cos\left(\varphi(1+\alpha)\right)\right)$$

with $\alpha = -3\frac{G_N M^3 P}{L^4}$. This returns to $u(0)$ when $\Delta\varphi = \frac{2\pi}{1+\alpha}$. $\alpha \neq 0$ means an *advance* of the perihelion.

12.2.2 The event horizon and beyond

At $r = r_0$, $g_{tt} = 0$ and $g_{rr} = \infty$. This makes it hard to use this metric to do physics there (*e.g.* it makes it hard to compute the Christoffel symbols). And it has physical consequences: for example, light emitted from $r = r_1$ with finite frequency ω and observed at radial position r is redshifted to[44]

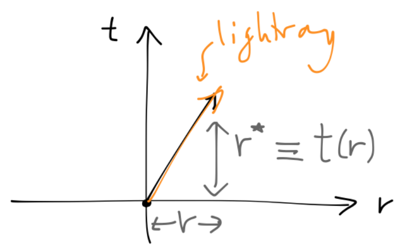
$$\frac{\omega(r)}{\omega} = \sqrt{\frac{g_{tt}(r_1)}{g_{tt}(r)}} \overset{r_1 \to r_0}{\to} 0.$$

But in fact the thing at $r = r_0$ is only a coordinate singularity. Evidence for this is that R and $R_{\mu\nu}R^{\mu\nu}$ and every other curvature invariant you can think of is finite there. To see how to go past this point, send geodesics in and see what happens. In fact it's just like the horizon in Rindler space that you are studying on problem set 9. [45]

In fact, it is possible to construct new coordinates (following Kruskal) which do just fine at the horizon. In these coordinates, the metric looks like

$$ds^2_{\text{sch}} = f^2(T, X)\left(-dT^2 + dX^2\right) + r^2(T, X)ds^2_{S^2}.$$

The form of this metric guarantees that the lightcones in the TX plane do not do any tipping over, and in fact are always at 45°. The idea for finding Kruskal coordinates is just to follow the lightrays.



Define a new radial coordinate $r^\star$ such that for a lightray with zero angular momentum, $t = \pm r^\star$. We have

$$0 = ds^2 \implies \frac{dt}{dr} = \pm e^{-2\alpha(r)} = \pm \frac{r}{r - r_0} \quad (121)$$

$$t(r) \overset{\text{def of } r^\star}{=} \pm r^\star + \text{const} \quad \overset{\text{integrate (121)}}{\Longrightarrow} \quad r^\star = r + r_0 \log\left(\frac{r}{r_0} - 1\right)$$

$$\left(\quad \text{A check:} \quad \frac{dt}{dr} = \frac{dt}{dr^\star}\frac{dr^\star}{dr} = \pm \frac{dr^\star}{dr} = \pm \frac{r}{r - r_0} \quad \right)$$

($r^\star$ is called a 'tortoise' coordinate.)

From here it's just like what we did for Minkowski space in the last section. Now introduce lightcone coords

$$u \equiv r - r^\star, v = t + r^\star$$

[44] We discussed the Newtonian limit of this formula at the very beginning of the course. See Zee page 303 for a derivation and the requisite assumptions.

[45] Here's one way we know that $r = 0$ is *not* a coordinate singularity:

$$R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} = \frac{12r_0^2}{r^6}.$$

This *scalar* quantity blows up there. No amount of changing coordinates is going to change that conclusion.

$$-\mathrm{d}u\mathrm{d}v = -\mathrm{d}t^2 + (\mathrm{d}r^\star)^2 = -\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{\left(1-\frac{r_0}{r}\right)^2}$$

$$\implies \mathrm{d}s^2|_{\mathrm{d}\theta=\mathrm{d}\varphi=0} = -\left(1-\frac{r_0}{r}\right)\mathrm{d}u\mathrm{d}v$$

Now we can let u, v run from $-\infty$ to ∞ with impunity. r is a function $r(u,v)$ defined by:

$$r^\star(u,v) = \frac{1}{2}(v-u) = r + r_0 \log\left(\frac{r}{r_0} - 1\right)$$

$$\implies e^{\frac{v-u}{2r_0}} e^{-r/r_0} = \frac{r}{r_0} - 1 \implies 1 - \frac{r_0}{r} = \frac{r_0}{r} e^{-r/r_0} e^{\frac{v-u}{2r_0}}.$$

The final step is to get rid of the singular prefactor g_{uv}:

$$dU = \frac{1}{2r_0} e^{-\frac{u}{2r_0}} \mathrm{d}u, \qquad dV = \frac{1}{2r_0} e^{-\frac{v}{2r_0}} \mathrm{d}v,$$

$$\implies \mathrm{d}s^2 = -\frac{4r_0^3}{r} e^{-\frac{r}{r_0}} \mathrm{d}U\mathrm{d}V.$$

(And $U = -e^{-\frac{u}{2r_0}}, V = e^{\frac{v}{2r_0}}$.) Now just undo these lightcone coordinates $T = \frac{1}{2}(V+U), X = \frac{1}{2}(V-U)$, and put back the sphere:

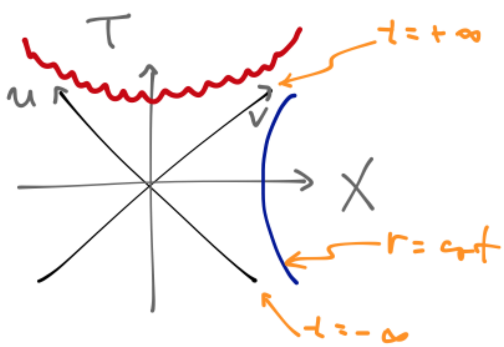
$$\mathrm{d}s^2_{\text{sch}} = \frac{4r_0^3}{r} e^{-\frac{r}{r_0}} \left(-\mathrm{d}T^2 + \mathrm{d}X^2\right) - r^2(T,X)\mathrm{d}s^2_{S^2}.$$

The inverse coordinate map is

$$t = \frac{1}{2}(v+u) = r_0 \log\left(\frac{X+T}{X-T}\right),$$

$$r^\star = \frac{1}{2}(v-u) = r_0 \log\left(X^2 - T^2\right) \quad (122)$$

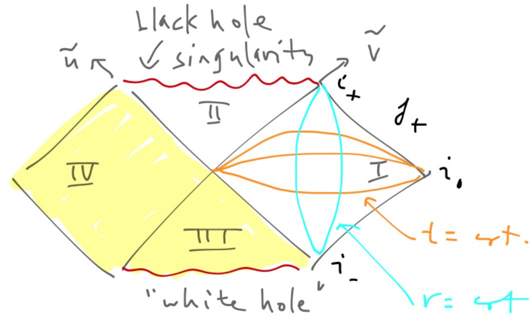
$$\implies X^2 - T^2 = \left(\frac{r}{r_0} - 1\right) e^{\frac{r}{r_0}}.$$



A locus of constant r in these (Kruskal) coordinates is $X^2 - T^2 = \text{const}$. The horizon is $r = r_0$, $X^2 - T^2 = 0$. The real singularity is $r = 0$: $X^2 - T^2 = -1$. Outside the horizon, $X^2 - T^2 > 0$. The region that we could access in the Schwarzschild coordinates also had $U = T - X < 0$ – it was just like one wedge of Rindler space.

Notice that the horizon is at $t = \infty$ – it takes infinitely long in the Schwarzschild time for something to reach the horizon. This leads to an arbitrarily large disagreement with the watch of someone falling in.

From the Kruskal coordinates we can construct the Penrose diagram for Schwarzchild and develop a better appreciation of its asymptotics. Just one more change of coordinates to make the range finite:



$$\tilde{U} = \tan^{-1} U, \tilde{V} = \tan^{-1} V.$$

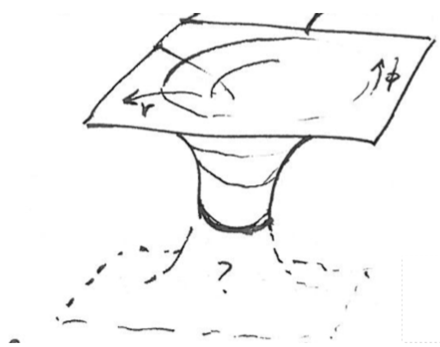
$$\implies \mathrm{d}s^2_{\text{sch}} = \Omega^2 \mathrm{d}\tilde{U}\mathrm{d}\tilde{V} + r^2(\tilde{U}, \tilde{V})\mathrm{d}s^2_{S^2}$$

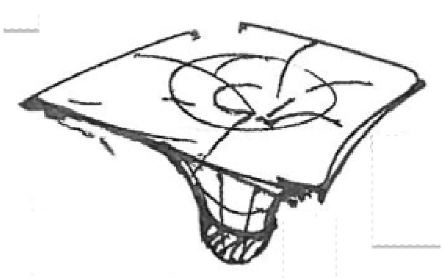
$\tilde{U}, \tilde{V} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

This picture makes it clear that what happens in region II stays in region II: once something enters this region, its whole future D^+(something), lies in region II.

The region with $\tilde{V} < 0$ is completely unexpected based on Schwarzschild coordinates. This is an artifact of our demanding an *eternal* black hole – a static solution that has *always* been sitting there. It leads to the weird phenomenon of the "white hole" (region III) from which *everything* escapes. A black hole which forms by collapse of some stuff does not have this region.

Here is a picture of the extended geometry of the eternal black hole at fixed time t: This picture is made by demanding that the induced metric from the embedding be the correct metric (at $\theta = \pi/2$) $\mathrm{d}x^i \mathrm{d}x^i|_{z=z(r)} = \left(1 - \frac{r_0}{r}\right)^{-1} \mathrm{d}r^2 + r^2 \mathrm{d}\varphi^2$ (which is solved by $z^2 = 4r_0(r - r_0)$). z is a good coordinate near $r = 0$. Notice that this 'throat' is not something that any actual object can pass through – only *spacelike* curves go through it.



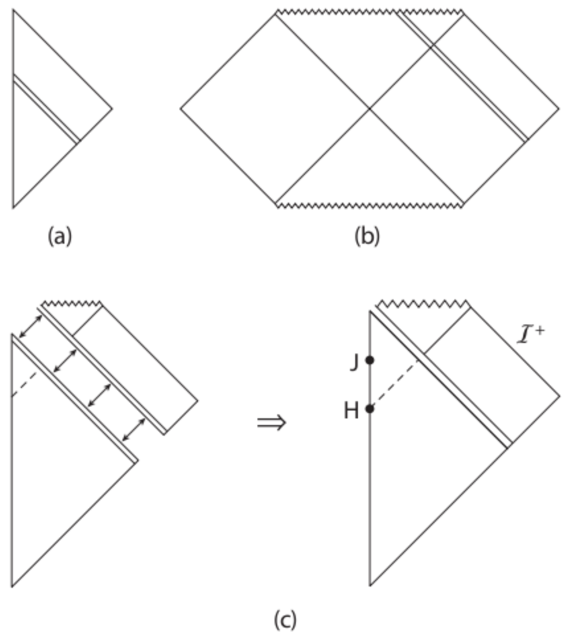


In the solution which describes a compact source, say the sun, the analogous picture is the one at left. where for $r <$ radius of star ($> r_0$), we have to solve Einstein's equations with the star-stuff as source.

We conclude with two cartoons.

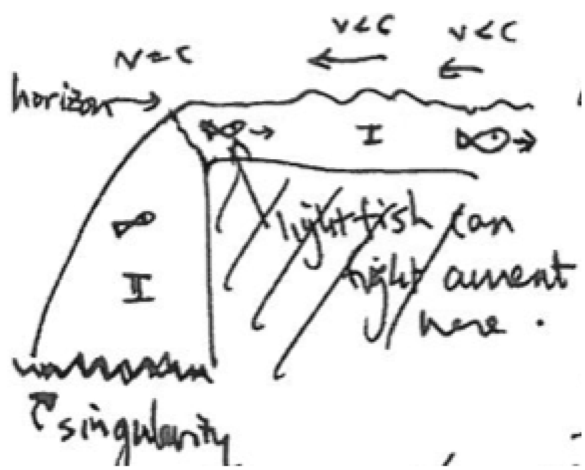
Here is (the Penrose diagram for) a spacetime with a black hole which could more plausibly exist: It is a black hole formed by a spherical shell of photons imploding. Outside the shell, there is no matter and spherical symmetry, so our Birkhoff argument shows that the solution is Schwarzchild, with the mass equal to the total energy of the shell of photons. *Inside* the shell, the same argument applies, but there is no mass inside so the solution is just Minkowski spacetime.

Notice that this shows that the event horizon is not a local concept – *e.g.* you need to know whether you are inside a shell of photons collapsing toward you in order to know whether or not you are right now behind the horizon of a black hole. Could be so.



[from Zee]

Here is a useful cartoon of a black hole, due to Unruh: It is just a waterfall; observers are fish. The small fish in region (I) don't notice the 'tidal forces'. The horizon is the point beyond which the velocity of the current is faster than c, the maximum velocity of fish. The fish don't experience anything special happening when they pass that point, the moment at which all hope of escape is lost. The singularity is of course the rocks at the bottom of the waterfall.



[End of Lecture 19]