PHYS 364 (General Relativity) Notes

Rio Weil

This document was typeset on March 6, 2026

Introduction:
This is a set of lecture notes taken from UChicago's PHYS 364 (General Relativity) taught by Dam Son.

Contents

1 Course Overview, Special Relativity

1.1 Course Overview

In this course we study the theory of gravity, as first developed by Einstein - widely considered to be one of the most beautiful theories of physics (and Son agrees).

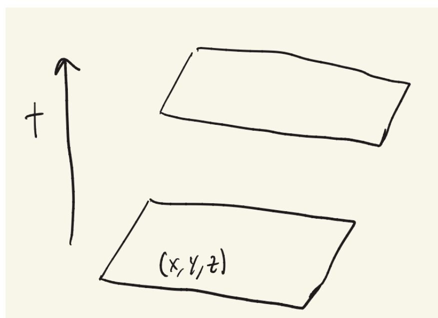
The course will use two books, that by Wald and that by Carroll. Wald is a bit more mathematically sophisticated, Carroll is a bit more modern. We'll use both as we go through.

In the first lecture, we will review special relativity (SR), and discuss the notions of vectors and tensors in SR.

1.2 Special Relativity

1.2.1 Newtonian Mechanics

We usually don't think about the structure of spacetime in Newtonian mechanics, but we can! In this setting, time is absolute, and we can think of coordinates (x, y, z) which evolve under a time parameter t.



Newtonian mechanics possesses Galilean invariance:

$$t \to t' = t \tag{1.1}$$

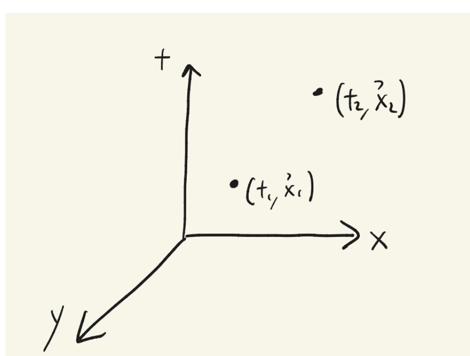
$$\mathbf{x} \to \mathbf{x}' = \mathbf{x} - \mathbf{v}t \tag{1.2}$$

1.2.2 Spacetime intervals

However, Maxwell's equations are *not* invariant under Galilean transformations. This introduces the need for special relativity, where space and time are united as spacetime (drawn as the 2+1D diagram below). In SR, we can consider the spacetime interval between two events as:

$$\Delta s^2 = -c^2(\Delta t)^2 + (\Delta \mathbf{x})^2 = -c^2(t_2 - t_1) - (\mathbf{x}_2 - \mathbf{x}_1)^2 \tag{1.3}$$

and the notion of invariance in SR is that this spacetime interval is independent of the frame.



In this course we use natural units with $c = 1$, so we can rewrite:

$$\Delta s^2 = -(\Delta x^0)^2 + (\Delta \mathbf{x})^2 \tag{1.4}$$

with $x^0 = ct = t$ and $\mathbf{x} = (x^1, x^2, x^3)$.

We use the convention that greek letters take on indices $\mu, \nu = 0, 1, 2, 3$ and latin indices take on indices $i, j = 1, 2, 3$.

1.2.3 Metrics

We also introduce the notion of a metric - in GR this is a crucial object. In SR we can avoid discussing it, but to make the transition to GR a bit smoother let us see how it manifests in this simpler setting. We have the Minkowski metric:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.5}$$

wherein the spacetime interval can be written as:

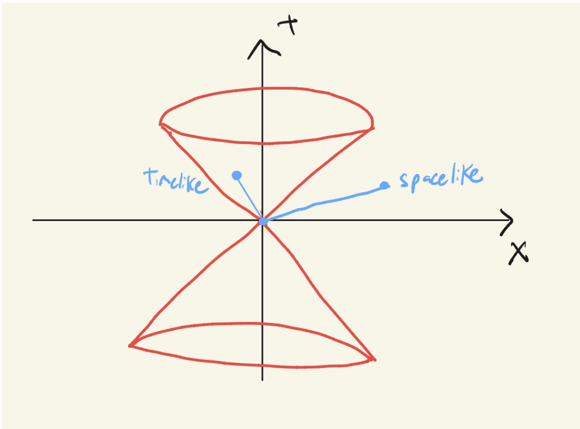
$$\Delta s^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \tag{1.6}$$

note the placement of the indices! Will be important to keep track of. Also, note that we use Einstein summation convention, where an identical lower and upper index are paired over, i.e. in full gory detail the above is written as:

$$\Delta s^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \tag{1.7}$$

1.2.4 Spacetime diagrams

We can also consider spacetime diagrams:



Where the lines 45 degrees from both axes correspond to events connected by a photon/a message travelling at the speed of light. We have a future and past lightcone, which correspond to the events that can be influenced by the origin/can influence the event at the origin respectively. These are timelike separated ($\Delta s^2 < 0$) with the event from the origin. The events outside of the lightcone cannot influence, nor influenced by, the event at the origin. These are spacelike separated ($\Delta s^2 > 0$) with the event from the origin.

1.2.5 Proper Time

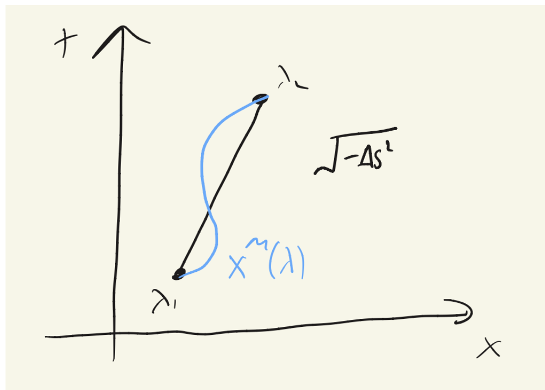
If an observer moves between events 1, 2 without acceleration, we can consider the time that passes on their clock to be the proper time:

$$\Delta\tau = \sqrt{-\Delta s^2} \tag{1.8}$$

We can consider some other trajectory $x^\mu(\lambda)$ of a (generically, accelerating) observer. Their experienced proper time is then:

$$\Delta\tau = \int d\lambda \sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{\lambda}} \tag{1.9}$$

in the case of zero acceleration, this formula reduces to the first, but this one is more general.



1.3 Lorentz Transformations and the Lorentz Group

In Minkowski spacetime, we have the spacetime invariant/interval:

$$ds^2 = -\eta_{\mu\nu}dx^\mu dx^\nu \tag{1.10}$$

We can consider coordinate transformations that preserve this. One would be a coordinate transformation by a constant a^μ (can be a space or time translation, or both):

$$x^\mu \to x^\mu = x^\mu + a^\mu \tag{1.11}$$

We can also consider Lorentz transformations:

$$x^\mu \to x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu \tag{1.12}$$

where $\Lambda^{\mu'}{}_{\mu}$ is a 4×4 matrix. Let us see what restrictions the invariance conditions place on this matrix:

$$\Delta s^2 = -\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu \to -\eta_{\mu'\nu'}\Delta x^{\mu'}\Delta x^{\nu'} = -\eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\mu}\Delta x^\mu \Lambda^{\nu'}{}_{\nu}\Delta x^\nu \tag{1.13}$$

so then the invariance condition is:

$$\eta_{\mu\nu} = \Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}\eta_{\mu'\nu'} \tag{1.14}$$

or in matrix form:

$$\eta = \Lambda^T \eta \Lambda \tag{1.15}$$

the group of all such transformations is the Lorentz group, written as $O(3,1)$.

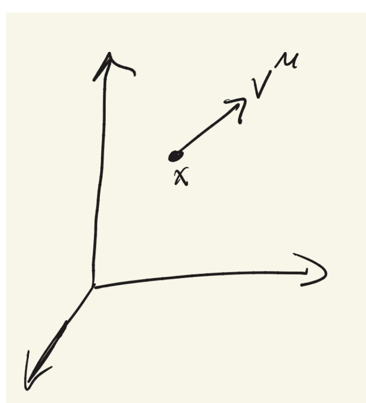
Let us consider some examples of Lorentz transformations:

1. Rotations, where time is untouched, and we rotate around one of the three spatial axes (below, a z-rotation) by angle ϕ
$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & \sin\phi & 0 \\ 0 & -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.16}$$
2. Boosts, where we boost along a spatial direction (mixing time/space) with rapidity η:
$$\Lambda = \begin{pmatrix} \cosh\eta & \sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.17}$$
3. Reflections; $x^i \to -x^i$. These correspond to transformations with $\det\Lambda = -1$. These are mathematically permitted, as we can see from the Lorentz group condition that:
$$\det\eta = \det(\Lambda^T\eta\Lambda) \implies -1 = (\det\Lambda)^2 \implies \det\Lambda = \pm 1 \tag{1.18}$$
We restrict/do not consider transformations of this type (i.e. we keep the orientation/chirality of space fixed), by restricting to those with $\det\Lambda = 1$.
4. There is one more type, where $t \to -t$ and $x^i \to -x^i$ (the latter condition to enforce $\det\Lambda = 1$). This cannot be simply connected to rotations/boosts (as time is inverted). So, we place the further condition/restriction that:
$$\Lambda^0_{\ 0} > 0 \tag{1.19}$$

1.4 Vectors

1.4.1 Transformation Rules of Vectors

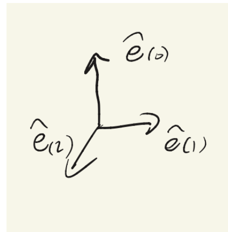
We can consider a vector at a point in spacetime $x^\mu(\lambda)$ (with trajectory parametrized by λ). For example, the velocity vector:
$$U^\mu = \frac{dx^\mu}{d\lambda} \tag{1.20}$$



Under lorentz transformations, vectors transform as:
$$\boxed{U^\mu \to U^{\mu'} = \Lambda^{\mu'}_{\ \mu} U^\mu} \tag{1.21}$$

We can consider basis vectors $\hat{e}_{(\mu)}$, wherein we can expand a vector v as:

$$U = U^{\mu}\hat{e}_{(\mu)} \tag{1.22}$$



If we look at the transformation rule, we have:

$$U^{\mu}\hat{e}_{(\mu)} \to \Lambda^{\mu'}{}_{\mu}U^{\mu}\hat{e}_{(\mu')} \tag{1.23}$$

so then:

$$\hat{e}_{(\mu)} = \Lambda^{\mu'}{}_{\mu}\hat{e}_{(\mu)} \tag{1.24}$$

or:

$$\boxed{\hat{e}_{(\mu')} = (\Lambda^{-1})^{\mu}{}_{\mu'}\hat{e}_{(\mu)}} \tag{1.25}$$

Where Λ^{-1} satisfies:

$$(\Lambda^{-1})^{\mu}{}_{\mu'}\Lambda^{\mu'}{}_{\nu} = \delta^{\mu}_{\nu} \tag{1.26}$$

$$\Lambda^{\mu'}{}_{\mu}(\Lambda^{-1})^{\mu}{}_{\nu} = \delta^{\mu'}_{\nu'} \tag{1.27}$$

1.4.2 Dual Vectors

As you already know, vectors v live in a vector space V. We can now consider dual vectors ω, which map from vectors into (real) numbers:

$$\begin{aligned} \omega \quad : \quad V &\longrightarrow \mathbb{R} \\ v &\longmapsto \omega(v) \end{aligned} \tag{1.28}$$

You can convince yourself that the space of all of these linear maps itself forms a vector space.

In components, we have:

$$\omega_{\mu} : v^{\mu} \to \omega_{\mu}v^{\mu} \tag{1.29}$$

We remark that vectors/dual vectors exist independent of any coordinate system. However, their components will depend on the choice of coordinate system.

Under Lorentz transformations:

$$\omega_{\mu'}v^{\mu'} = \omega_{\mu}v^{\mu} \tag{1.30}$$

which gives us the transformation law:

$$\boxed{\omega_{\mu'} = (\Lambda^{-1})^{\mu}{}_{\mu'}\omega_{\mu}} \tag{1.31}$$

1.4.3 Example: Derivative

As an example, let us consider a function (scalar field) $\phi(x)$. Now, consider a derivative of this function:

$$\partial_\mu \phi = \frac{\partial \phi}{\partial x^\mu} \tag{1.32}$$

Then:

$$\partial_{\mu'} \phi = \frac{\partial \phi}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial \phi}{\partial x^\mu} \tag{1.33}$$

But if we recall that:

$$x^{\mu'} = \Lambda^{\mu'}{}_{\mu} x^\mu \tag{1.34}$$

$$x^\mu = (\Lambda^{-1})^{\mu}{}_{\mu'} x^{\mu'} \tag{1.35}$$

so:

$$\partial_{\mu'} \phi = (\Lambda^{-1})^{\mu}{}_{\mu'} \partial_\mu \phi \tag{1.36}$$

so (at any given point) the derivative of a function transforms like a dual vector.

1.5 Tensors

1.5.1 Tensors and Transformation Rules

Let us define the space of all vectors at a point p to be T_p, and the space of all dual vectors at that point to be T_p^*.

A (k,l) tensor T is defined as the multilinear (i.e. it is linear with respect to any of its arguments) map:

$$\begin{aligned} T \quad : \quad & \underbrace{T_p^* \times T_p^* \times \ldots \times T_p^*}_{k \text{ times}} \times \underbrace{T_p \times T_p \times \ldots \times T_p}_{l \text{ times}} && \longrightarrow \quad \mathbb{R} \\ & (\omega_1^{(1)}, \omega^{(2)}, \ldots, \omega^{(k)}, v^{(1)}, \ldots, v^{(l)}) && \longmapsto \quad T(\omega_1^{(1)}, \omega^{(2)}, \ldots, \omega^{(k)}, v^{(1)}, \ldots, v^{(l)}) \end{aligned} \tag{1.37}$$

In components:

$$T(\omega_1^{(1)}, \omega^{(2)}, \ldots, \omega^{(k)}, v^{(1)}, \ldots, v^{(l)}) = T^{\mu_1\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\nu_l} \omega^{(1)}_{\mu_1} \omega^{(2)}_{\mu_2} \ldots \omega^{(k)}_{\mu_k} v^{(1)\nu_1} v^{(2)\nu_2} \ldots v^{(l)\nu_l} \tag{1.38}$$

Tensors transform as:

$$T \to T^{\mu_1'\ldots\mu_k'}{}_{\nu_1'\ldots\nu_k'} = \Lambda^{\mu_1'}{}_{\mu_1} \ldots \Lambda^{\mu_k'}{}_{\mu_k} (\Lambda^{-1})^{\nu_1}{}_{\nu_1} \ldots (\Lambda^{-1})^{\nu_l}{}_{\nu_l'} T^{\mu_1\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\nu_l} \tag{1.39}$$

In addition to addition $(T_1 + T_2)$ and scalar multiplication (λT) or tensors, we can also consider a tensor product between a (k,l) tensor T and a (m,n) tensor S which is the $(k+m, l+n)$ tensor denoted by $T \otimes S$.

1.5.2 Examples: Metric and Kronecker Delta, Levi-Civita

We claim that $\eta_{\mu\nu}$ is a tensor. To see this, look at the transformation:

$$\eta_{\mu\nu} \overset{?}{\to} \eta_{\mu'\nu'} = (\Lambda^{-1})^{\mu}{}_{\mu'} (\Lambda^{-1})^{\nu}{}_{\nu'} \eta_{\mu\nu} \tag{1.40}$$

now, we know that:

$$\eta' = (\Lambda^{-1})^T \eta \Lambda^{-1} \tag{1.41}$$

for Lorentz, we have:

$$\Lambda^T \eta \Lambda = \eta \tag{1.42}$$

so then:

$$\eta' = (\Lambda^{-1})^T \eta \Lambda^{-1} = (\Lambda^{-1})^T \Lambda^T \eta \Lambda \Lambda^{-1} = \eta \tag{1.43}$$

Thus, we have shown $\eta_{\mu\nu}$ to be a $(0,2)$ tensor. Specifically:

$$(v_1, v_2) \to \eta_{\mu\nu} v_1^\mu v_2^\nu \tag{1.44}$$

We alo have the inverse metric:

$$\eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.45}$$

which is a $(2,0)$ tensor. In Minkowski spacetime they coincide, this will not be true generically curved spacetime.

We can also see that $\delta^\mu_{\ \nu}$ is a tensor. In matrix form:

$$\delta^\mu_{\ \nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.46}$$

We know that the kronecker delta maps:

$$(\omega, v) \to \omega(v) \equiv \omega_\mu v^\mu \tag{1.47}$$

Finally, we consider the Levi-Civita symbol:

$$\epsilon^{\mu\nu\lambda\rho} = \begin{cases} +1 & (\mu\nu\lambda\rho) \text{ is an even permutation of (0123)} \\ -1 & (\mu\nu\lambda\rho) \text{ is an odd permutation of (0123)} \\ 0 & \text{otherwise} \end{cases} \tag{1.48}$$

Let us check that this is indeed a tensor, by lorentz transforming into a different frame and checking that the components are left invariant.

$$\epsilon^{\mu'\nu'\lambda'\rho'} = \Lambda^{\mu'}_{\ \mu} \Lambda^{\nu'}_{\ \nu} \Lambda^{\lambda'}_{\ \lambda} \Lambda^{\rho'}_{\ \rho} \epsilon^{\mu\nu\lambda\rho} \tag{1.49}$$

Note that ϵ is totally antisymmetric by construction, and indeed even after the transformation it retains this property. Thus, the transformed tensor must be proportional to the original (as it is is the only totally antisymmetric tensor). If we then look at $\epsilon^{0'1'2'3'}$, we find (recalling that we get terms with alternating signs):

$$\epsilon^{0'1'2'3'} = \det(\Lambda) = +1. \tag{1.50}$$

and hence the proportionality becomes equality.

Next time, we finish our discussion of SR by introducing the Maxwell equations and the energy-momentum tensor. Afterwards, we begin our discussion of Riemannian geometry.

2 Electromagnetism and the Stress-Energy Tensor

Last time, we defined the notions of vectors, dual vectors, and tensors (all in the context of Minkowski space).

Today, we discuss how to manipulate and construct tensors. We will then remind ourselves of the notions of the electromagnetic field tensor and stress energy tensors. We will then discuss fluids and classical field theory.

2.1 Tensor operations

We defined the notion of a (k, l) tensor, with components:

$$T^{\mu_1 \ldots \mu_k}{}_{\nu_1 \ldots \nu_l} \tag{2.1}$$

which obeys the transformation law:

$$T^{\mu'_1 \ldots \mu'_k}{}_{\nu'_1 \ldots \nu'_l} = \Lambda^{\mu'_1}{}_{\mu_1} \ldots \Lambda^{\mu'_k}{}_{\mu_k} (\Lambda^{-1})^{\nu_1}{}_{\nu'_1} \ldots (\Lambda^{-1})^{\nu_l}{}_{\nu_l} T^{\mu_1 \ldots \mu_k}{}_{\nu_1 \ldots \nu_l} \tag{2.2}$$

2.1.1 Tensor contractions

We can form a $(k-1, l-1)$ tensor from a (k, l) tensor as follows:

$$T^{\mu\nu\rho}{}_{\sigma\tau} \to S^{\mu\rho}{}_{\sigma} = T^{\mu\nu\rho}{}_{\sigma\nu} \tag{2.3}$$

where the sum over ν is implied via the Einstein summation convention. You can check that this indeed is a $(k-1, l-1)$ tensor/has the correct transformation properties.

2.1.2 Lower/Raising Indices, Scalar product

By using the metric tensor $\eta_{\mu\nu}, \eta^{\mu\nu}$, we can lower and raise indices. For example:

$$\omega_\mu \to \omega^\mu = \eta^{\mu\nu}\omega_\nu \tag{2.4}$$

If we were being pedantic, we would split the above into two steps:

1. First, form the tensor product:
$$(\eta \otimes \omega)^{\mu\nu}{}_{\alpha} = \eta^{\mu\nu}\omega_\alpha \tag{2.5}$$
2. Then, contract by setting $\alpha = \nu$ and summing over.

We can lower indices in an identical way.

We can define the scalar product of two vectors A^μ, B^μ by:

$$AB = \eta_{\mu\nu} A^\mu B^\nu = A_\mu B^\mu = A^\mu B_\mu \tag{2.6}$$

2.1.3 Symmetrization and Antisymmetrization

If I have a tensor $T_{\mu_1 \ldots \mu_n \rho}{}^{\sigma}$, we can symmetrize over the indices $\mu_1 \ldots \mu_n$:

$$T_{(\mu_1 \ldots \mu_n)\rho}{}^{\sigma} = \frac{1}{n!}(T_{\mu_1 \ldots \mu_n \rho}{}^{\sigma} + \text{all permutations of } (\mu_1 \ldots \mu_n)) \tag{2.7}$$

We can also antisymmetrize over these indices:

$$T_{[\mu_1 \ldots \mu_n]\rho}{}^{\sigma} = \frac{1}{n!}(T_{\mu_1 \ldots \mu_n \rho}{}^{\sigma} + (-1)^{\text{parity of permutation}} \text{all permutations of } (\mu_1 \ldots \mu_n)) \tag{2.8}$$

As examples:

$$T_{(\mu_1\mu_2)} = \frac{1}{2}(T_{\mu_1\mu_2} + T_{\mu_2\mu_1}) \tag{2.9}$$

$$T_{[\mu_1\mu_2]} = \frac{1}{2}(T_{\mu_1\mu_2} - T_{\mu_2\mu_1}) \tag{2.10}$$

$$T_{(\mu_1\mu_2\mu_3)} = \frac{1}{6}(T_{\mu_1\mu_2\mu_3} + T_{\mu_1\mu_3\mu_2} + T_{\mu_2\mu_1\mu_3} + T_{\mu_2\mu_3\mu_1} + T_{\mu_3\mu_1\mu_2} + T_{\mu_3\mu_2\mu_1}) \tag{2.11}$$

$$T_{[\mu_1\mu_2\mu_3]} = \frac{1}{6}(T_{\mu_1\mu_2\mu_3} - T_{\mu_1\mu_3\mu_2} - T_{\mu_2\mu_1\mu_3} + T_{\mu_2\mu_3\mu_1} + T_{\mu_3\mu_1\mu_2} - T_{\mu_3\mu_2\mu_1}) \tag{2.12}$$

We notice that the rank-2 tensor can be written as a sum of a symmetric and antisymmetric tensor:

$$T_{\mu_1\mu_2} = T_{(\mu_1\mu_2)} + T_{[\mu_1\mu_2]} \tag{2.13}$$

Note also that:

$$X^{(\mu\nu)}Y_{\mu\nu} = X^{(\mu\nu)}Y_{(\mu\nu)} \tag{2.14}$$

and so:

$$X^{(\mu\nu)}Y_{[\mu\nu]} = 0 \tag{2.15}$$

which will be a very useful identity.

2.1.4 Trace

We can trace the trace via:

$$X^{\mu}_{\ \mu} \equiv X^{\mu\nu}\eta_{\nu\mu} \tag{2.16}$$

We can look at the trace of η, for example:

$$\eta_{\mu\nu}\eta^{\nu\mu} = 4 \tag{2.17}$$

note that this is not the naive trace one would expect from looking at the matrix representation of $\eta^{\mu\nu}$ (which has a -1), but indeed it is the correct expression/thing that comes out of the manifestly Lorentz invariant way of taking the trace.

2.1.5 Derivatives

If we have the tensor:

$$T^{\mu_1\dots\mu_k}_{\qquad \nu_1\dots\nu_l}(x) \tag{2.18}$$

Then:

$$\frac{\partial}{\partial x^{\alpha}}T^{\mu_1\dots\mu_k}_{\qquad \nu_1\dots\nu_l}(x) = \tilde{T}^{\mu_1\dots\mu_k}_{\qquad \nu_1\dots\nu_l\alpha} \tag{2.19}$$

2.2 Electromagnetism

2.2.1 Electromagnetic Field Tensor and Maxwell Equations

Your first exposure to electromagnetism treated the **E** and **B** fields separately. But (as you may have seen in a more advanced course), we can combine them into a single electromagnetic field tensor:

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_1 & -B_2 \\ E_1 & -B_1 & 0 & B_3 \\ E_3 & B_2 & B_3 & 0 \end{pmatrix} \tag{2.20}$$

Note that this tensor is antisymmetric:

$$F_{\mu\nu} = -F_{\nu\mu} \tag{2.21}$$

And that $F_{\mu\nu}$ transforms as a tensor under Lorentz transformations, and this implies the familiar transformations of the $\mathbf{E}, \mathbf{B}$ fields under Lorentz transformations that you have studied.

The celebrated Maxwell equations:

$$\nabla \cdot \mathbf{E} = \rho \tag{2.22}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.23}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.24}$$

$$\nabla \times \mathbf{B} = \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \tag{2.25}$$

can be written in a manifestly Lorentz covariant manner:

$$\partial_\mu F^{\nu\mu} = J^\nu \tag{2.26}$$

with $J^\mu = (\rho, \mathbf{J})$. When $\nu = 0$, we get Eq. (2.22) and when $\nu = 1, 2, 3$ we get Eq, (2.25). We also have the antisymmetrized equation:

$$\partial_{[\mu} F_{\nu\lambda]} = 0 \tag{2.27}$$

which yields Eqs. (2.23) and (2.24).

2.2.2 Particle in EM field

A particle's trajectory is given by the worldline $x^\mu(\lambda)$ parameterized by λ. The velocity of the particle is given by:

$$U^\mu = \frac{dx^\mu}{d\lambda}. \tag{2.28}$$

One can choose λ such that $U^\mu U_\mu = -1$; this choice corresponds to choosing λ to be the proper time.

The energy-momentum of the particle is given by:

$$p^\mu = mU^\mu \tag{2.29}$$

The equation of motion of the particle (again taking $\lambda = \tau$) is given by:

$$\frac{dp^\mu}{d\tau} = qU^\lambda T_\lambda^{\ \mu} \tag{2.30}$$

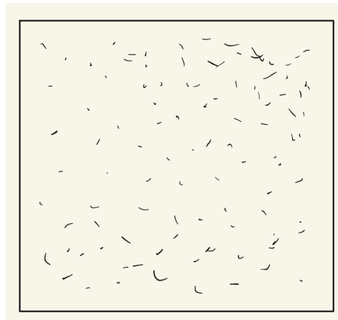
The RHS corresponds to the Lorentz force:

$$U^0 F_0^{\ i} \implies q\mathbf{E} \tag{2.31}$$

$$U^j F_j^{\ i} \implies q(\mathbf{v} \times \mathbf{B}) \tag{2.32}$$

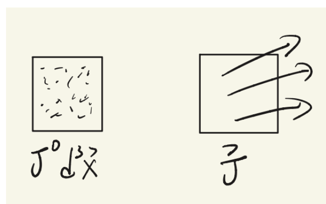
2.3 The Stress-Energy Tensor

2.3.1 Conservation Equations



Consider a medium filled with particles. We can consider a number current $J^\mu(x)$. Note that if all the particles carry a (uniform) charge, then this corresponds to an electrical current as well.

The number of particles in a volume $d^3\mathbf{x}$ is given by $J^0 d^3\mathbf{x}$, and $\mathbf{J}$ corresponds to the flux.



We then have the conservation law:

$$\boxed{\partial_\mu J^\mu = \frac{\partial J^0}{\partial t} + \nabla \cdot \mathbf{J} = 0} \tag{2.33}$$

In Newtonian gravity, the source of gravity is mass. In general relativity, the source is the relativistic version of the mass distribution, known as the stress energy tensor. Let us construct this (symbolized by $T^{\mu\nu}$). What properties do we want it to have?

If we look at a volume $d^3\mathbf{x}$, the total energy is given by:

$$T^{00} d^3 x \tag{2.34}$$

with the first 0 corresponding to the energy and the second 0 to the density.

If we look at the total momentum in the same volume, we have:

$$T^{i0} d^3\mathbf{x} \tag{2.35}$$

with the i denoting the momentum in the ith direction and the 0 to the density.

The energy flux in the x^ith direction is given by:

$$T^{0i} \tag{2.36}$$

So the conservation of energy can be phrased as:

$$\boxed{\partial_0 T^{00} + \partial_i T^{0i} = 0} \tag{2.37}$$

And the flux of momentum p^i in the direction x^j is given by:

$$T^{ij} \tag{2.38}$$

So the conservation of momentum can be written as:

$$\boxed{\partial_0 T^{i0} + \partial_j T^{ij} = 0} \tag{2.39}$$

Writing the above conservation equations in Lorentz covariant form, we get:

$$\boxed{\partial_\mu T^{\mu\nu} = 0} \tag{2.40}$$

The total energy of the system is given by:

$$E = \int d^3\mathbf{x} T^{00}(\mathbf{x}) \tag{2.41}$$

and the total momentum by:

$$p^i = \int d^3\mathbf{x} T^{i0}(\mathbf{x}) \tag{2.42}$$

2.3.2 Stress-Energy Tensor for Kientic Theory

Let us now derive the expression for $T^{\mu\nu}$, for a gas of particles in kinetic theory. They are parametrized by a probability distribution $f(t, \mathbf{x}, \mathbf{p})$ (which we leave unspecified) over phase space.

The 00 component is given by:

$$T^{00}(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} p^0 f(x, \mathbf{p}) \tag{2.43}$$

$$T^{i0} = \int \frac{d^3\mathbf{p}}{(2\pi)^3} p^i f(x, \mathbf{p}) \tag{2.44}$$

The above two formulas are not Lorentz covariant; lets fix this. Note that:

$$p^2 = -(p^0)^2 + \mathbf{p}^2 \implies p_0 = \sqrt{\mathbf{p}^2 + m^2} \tag{2.45}$$

So then:

$$\int d^3\mathbf{p} = \int d^4p 2p^0 \delta(p^2 + m^2)\Theta(p^0) \tag{2.46}$$

which is the on-shell condition for the mass and momentum (the $\Theta(p^0)$ is there to enforce that the mass is positive, and the $2p^0$ for the correct normalization of the integration measure). Thus our formula for $T^{00}(x)$ becomes:

$$T^{00}(x) = \int \frac{d^3p}{(2\pi)^3} 2\delta(p^2 + m^2)\Theta(p^0)p^0 p^0 f \tag{2.47}$$

Now, $\int \frac{d^3p}{(2\pi)^3} 2\delta(p^2 + m^2)\Theta(p^0)$ is Lorentz invariant, and we may package this as:

$$T^{00}(x) = \int \tilde{d}^4 p p^0 p^0 f(x, p) \tag{2.48}$$

where we have introduced an arbitrary dependence of f on the four-vector/including p^0, which is handled by the delta function in the integration measure. Similarly:

$$T^{i0}(x) = \int \tilde{d}^4 p p^0 p^i f(x, p) \tag{2.49}$$

and note that, more generically:

$$T^{\mu\nu} = \int \tilde{d}^4 p p^\mu p^\nu f(x, p) \tag{2.50}$$

with:

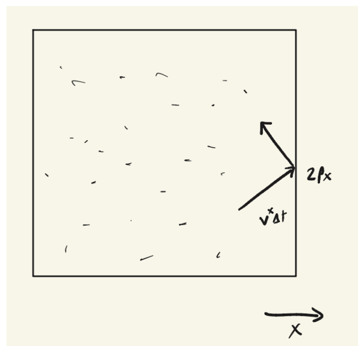
$$\tilde{d}^4 p = \frac{d^3 p}{(2\pi)^3} 2\delta(p^2 + m^2)\Theta(p^0) \tag{2.51}$$

Note that the stress-energy tensor is symmetric, with $T^{\mu\nu} = T^{\nu\mu}$ (and this is generally true of physical systems). This has the implication that (as one example) the momentum density is equal to the energy flux. Further, $T^{ij} = T^{ji}$ (the p^i flux in the x^jth direction is equal to the p^j flux in the x^ith direction).

Let's look at the physical meaning of T^{xx}. From our formula:

$$T^{xx} = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{p^x p^x}{p^0} f = \int \frac{d^3\mathbf{p}}{(2\pi)^3} v^x p^x f \tag{2.52}$$

We have a gas of particles that collides with a wall (perpendicular to the x direction).



The momentum transfer of a collision is p^x and the number of particles colliding the wall in unit time is $v^x \Delta\tau$. The pressure applied to the wall is defined as the momentum transfer divided by (time $\times$ unit area). So as a formula:

$$P_x = \int_{p^x>0} \frac{d^3 p}{(2\pi)^3} 2p^x v^x f \tag{2.53}$$

exactly the entry of the stress-energy tensor! (up to a factor of 2, which is handled by the $p^x > 0$ condition).

Thus the 00 component of T is the energy density and the spatial diagonal components are the pressure.

2.3.3 Perfect Fluids

Perfect fluids are those for which the stress-energy tensor is diagonal (in the rest frame of the fluid, ie. $U^\mu = (1, \mathbf{0})$):

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{2.54}$$

We can write down the stress-energy tensor in an arbitrary flame by considering an arbitrary velocity U^μ of the gas, wherein things generalize to:

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + p\eta^{\mu\nu} \tag{2.55}$$

One such example of a perfect fluid is "dust", where $p = 0$. In the rest frame of the dust, there is only energy density, no pressure. The general formula thus has the form:

$$T^{\mu\nu} = pU^\mu U^\nu \tag{2.56}$$

Next class, we'll study classical fields in generality. We will look at the stress energy tensor of classical fields, and specifically for the EM field:

$$T^{00} = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) \tag{2.57}$$

$$T^{0i} = T^{i0} = (\mathbf{E} \times \mathbf{B})_i \tag{2.58}$$

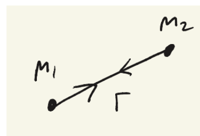
3 Introduction to Gravity, Differential Geometry I

Due to time constraints, we skip over the review of classical field theory - we may come back to it later, or you may want to read about it in your own time.

3.1 Newtonian Gravity

Newton's law of gravitation tells us that masses attract, with magnitude:

$$|\mathbf{F}| = \frac{Gm_1m_2}{r^2} \tag{3.1}$$



Consider a test mass m. The equation of motion of it is given by:

$$\boxed{m\ddot{\mathbf{x}} = -m\boldsymbol{\nabla}\Phi} \tag{3.2}$$

where Φ is the gravitational potential (c.f. $\mathbf{F} = q\mathbf{E} = -q\boldsymbol{\nabla}\phi$ in EM).
The potential satisfies the Poisson equation:

$$\boxed{\boldsymbol{\nabla}^2\Phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x})} \tag{3.3}$$

Again, compare to the EM case with $\boldsymbol{\nabla}^2\phi = -4\pi\rho_e(\mathbf{x})$. Note the difference in sign, which encodes that the gravitational force is attractive.

3.2 The Equivalence Principle

How do we modify these equations to be consistent with relativity? This is what Einstein accomplished (from 1905-1915) - his construction was quite different from all other physical theories. Here we see that Einstein working alone was a positive thing in some sense. The theory of GR will rely on a very small number of assumptions.

One such assumption was the *principle of equivalence*. In mechanics, if we consider a particle in a gravitational field (and under influence of other forces), we have:

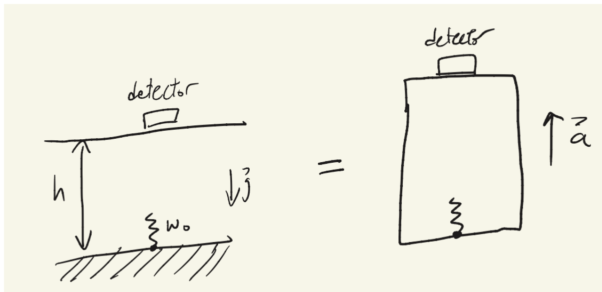
$$m\ddot{\mathbf{x}} = \mathbf{F}_{\text{other particles}} + m\mathbf{g}. \tag{3.4}$$

Say we move to a coordinate system that moves with acceleration $\mathbf{A}$. Then, the equation of motion acquires a term corresponding to the fictitious force:

$$m\ddot{\mathbf{x}} = \mathbf{F} + m\mathbf{g} - m\mathbf{A} \tag{3.5}$$

Now, we can take $\mathbf{A} = \mathbf{g}$, at which point the gravitational force vanishes! This is the classic thought experiment Einstein considered, where in a falling elevator you don't feel gravity. This thought experiment suggests a kind of equivalence between frames, and that we should be able to use a more general coordinate system.

Let's consider another thought experiment (in fact, one that has been done). We consider a detector raised from a source of a light beam with frequency ω_0, by a height h. We can solve this using the equivalence principle. Let us solve this problem in space, inside an elevator with an upwards acceleration $\mathbf{a} = -\mathbf{g}$ (wherein we would feel the same gravitational force as we would on Earth).



At time $t_0 = 0$, both the photon source and detector have velocity $\mathbf{v} = \mathbf{v}_0 = \mathbf{0}$. At this time, a photon with speed c is emitted from the source. At time $t_1 \approx \frac{h}{c}$, the photon reaches the detector. At this time, the velocity of the detector is $\frac{gh}{c} \ll c$. Now, the frequency of the photon as seen by the detector is decreased, due to the Doppler effect:

$$\omega = \omega_0 \left(1 - \frac{gh}{c}\frac{1}{c}\right) = \omega_0 \left(1 - \frac{gh}{c^2}\right) \tag{3.6}$$

Now, let us recall that our definition of our time comes from frequency (of Cesium atoms) - this thought experiment thus suggests that time flows differently in the presence of a gravitational field! Thus, to handle gravity relativistically, we need to treat spacetime in a new way - in a way that can handle the fact that two clocks that do not appear to be moving w.r.t. each other can nonetheless tick at different rates. This motivated Einstein to think about differential/Riemannian geometry.

With this introduction out of the way, let us spend the next two lectures discussing basic notions from differential geometry:

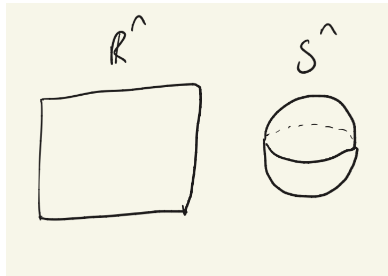
- Manifolds
  - Vectors, Dual vectors
  - Tensors, Tensor densities
  - Metric Tensor
- Parallel Transport
- Covariant Derivative

3.3 Manifolds

We can both define a manifold very precisely, and also have a geometric/intuitive definition - we will introduce both, let's start with a picture and some examples.

- n-dimensional Euclidean space $\mathbb{R}^n$
- The n-sphere satisfying:

$$(x^1)^2 + \ldots + (x^{n+1})^2 = 1 \tag{3.7}$$



The key point of the second example is that locally, the sphere looks flat/Euclidean. To make this intuition more precise, we define:

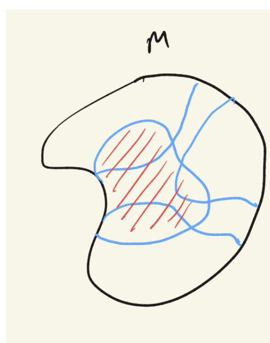
Definition: Manifolds

A manifold M is a set of points and a collection of subsets θ_α, such that:

- Any $p \in M$ belongs to at least one θ_α
- On each subset θ_α, we can define a coordinate system x^μ
- If $\theta_\alpha, \theta_\beta$ intersect, then the coordinates x^μ, y^μ can be related by differentiable maps:

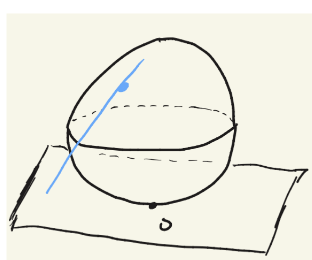
$$y^\mu = f(x^\mu) \tag{3.8}$$

$$x^\mu = g(y^\mu) \tag{3.9}$$



This definition is quite abstract and general, and it is this way because for (many) manifolds a single coordinate system is not enough to cover it.

In the case of $\mathbb{R}^n$, indeed a single coordinate system/subset suffices. For the case of S^n, a single coordinate system no longer suffices. Taking the two-sphere S^2 as an example, we can map it to a plane $\mathbb{R}^2$ via stereographic projection:



which works as a one-to-one - *except* for the north pole. We can then consider another analogous projection map that works for every point except for the south pole, and this gives us our 2 coordinate systems/subsets that cover the entire space (with $\theta_1 = S^2 \setminus \{(0,0,1)\}$ and $\theta_2 = S^2 \setminus \{(1,0,0)\}$).

Consider a function $M \to \mathbb{R}$. We have a coordinate representation of this map (around a point p), with:

$$x^\mu \to f(x^\mu) \tag{3.10}$$

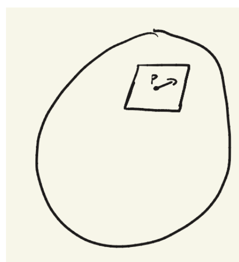
for $\mu = 1, \ldots, n$. There are different ways to represent the same function, in a different set of coordinates $x^\mu, x^{\mu'}$:

$$\phi(x^\mu) \to \phi'(x^{\mu'}) = \phi(x^\mu(x^{\mu'})) \tag{3.11}$$

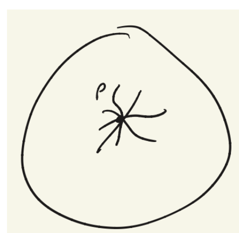
and such functions are "scalar fields".

3.4 Tangent Vectors

We can consider a manifold, and intuitively the tangent vector at p is the vector tangential to the manifold at p.



This intuitive definition requires an embedding into a lower dimensional space, so is unsatisfying for a mathematician. So, let us make it more formal. We can consider curves $x^\mu = x^\mu(\lambda)$ that run through p.



Definition: Tangent vectors

Consider $p \in M$, and curves $x^\mu = x^\mu(\lambda)$ that run through p. We can define the derivative of a scalar function ϕ along the curve as:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\phi(x^\mu(\lambda)) = \frac{dx^\mu}{d\lambda}\frac{\partial\phi}{\partial x^\mu} \tag{3.12}$$

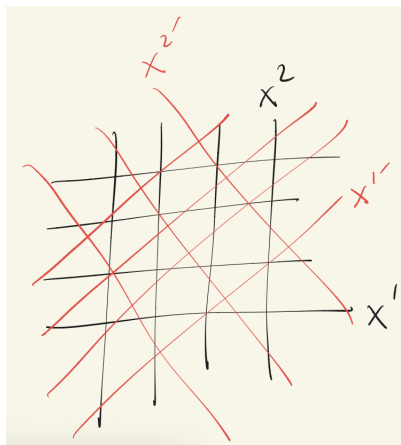
Then, we can define the tangent vectors as all differential operators:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} : V = V^\mu \frac{\partial}{\partial x^\mu} \tag{3.13}$$

with:

$$V^\mu = \frac{dx^\mu(\lambda)}{d\lambda}. \tag{3.14}$$

These tangent vectors form a vector space.



Let us consider a coordinate transformation of the tangent vectors:

$$V = V^\mu(x)\frac{\partial}{\partial x^\mu}\phi(x) = V^{\mu'}(x')\frac{\partial}{\partial x^{\mu'}}\phi(x') \tag{3.15}$$

Now using the chain rule:

$$V^\mu(x)\frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial\phi}{\partial x^{\mu'}} = V^{\mu'}(x^{\mu'})\frac{\partial}{\partial x^{\mu'}}\phi \tag{3.16}$$

with:

$$V^{\mu'}(x^{\mu'}) = \frac{\partial x^{\mu'}}{\partial x^\mu}V^\mu(x) \tag{3.17}$$

if we consider the special case of a Lorentz transformation:

$$x^\mu \to x^{\mu'} = \Lambda^{\mu'}_{\ \mu}x^\mu \tag{3.18}$$

then:

$$V^{\mu'} = \Lambda^{\mu'}_{\ \mu}V^\mu \tag{3.19}$$

3.5 Dual Vectors

Definition: Dual Vectors

For any coordinate system with basis vectors $e_\mu = \frac{\partial}{\partial x^\mu}$, we have the tangent space at p:

$$T_p = V^\mu \frac{\partial}{\partial x^\mu} = V^\mu e_\mu \tag{3.20}$$

the dual vectors are then linear maps ω_μ from $T_p \to \mathbb{R}$:

$$\begin{array}{cccccc} T_p^* & : & T_p & \longrightarrow & \mathbb{R} \\ & & V^\mu & \longmapsto & \omega(v) = \omega_\mu V^\mu \end{array} \tag{3.21}$$

We can see how the dual vectors act on vectors via a scalar product.
we can write $\omega = \omega_\mu e^\mu$ with e^μ the dual basis vectors.

Under a coordinate change, the dual vectors transform as:

$$\omega_{\mu'}(x^{\mu'}) = \frac{\partial x^\mu}{\partial x^{\mu'}} \omega_\mu(x) \tag{3.22}$$

The basis vectors in T_p^* are e^μ, which we can denote as dx^μ. and they transform as:

$$e^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} e^\mu \tag{3.23}$$

so:

$$\omega'_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \omega_\mu \tag{3.24}$$

with:

$$\omega = \omega_\mu e^\mu (= \omega_\mu dx^\mu) = \omega_{\mu'} e^{\mu'} \tag{3.25}$$

The duality is clear if we write:

$$dx^\mu \left(\frac{\partial}{\partial x^\nu} \right) = \delta^\mu_\nu \tag{3.26}$$

3.6 Tensors and the Metric Tensor

Definition: Tensors

A (k, l) tensor T is a multilinear map:

$$T : \underbrace{T_p^* \times \ldots \times T_p^*}_{k} \times \underbrace{T_p \times \ldots \times T_p}_{l} \to \mathbb{R}. \tag{3.27}$$

It has coordinate representation $T^{\mu_1 \ldots \mu_k}_{\nu_1 \ldots \nu_l}$. It has transformation law:

$$\boxed{T^{\mu'_1 \ldots \mu'_k}_{\nu'_1 \ldots \nu'_l}(x') = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}} T^{\mu_1 \ldots \mu_k}_{\nu_1 \ldots \nu_k}}(x). \tag{3.28}$$

In a more elementary treatment of differential geometry, one would take the above transformation law as the definition of the tensor. But our definition in terms of tangent spaces is coordinate free.

The metric (0, 2) tensor is the $g_{\mu\nu}(x)$, with:

$$g = g_{\mu\nu}(x) dx^\mu \otimes dx^\nu \tag{3.29}$$

where dx^μ are, as before, the basis vectors in T_p^*. This is the mathematical definition, but the intuition is that the metric tensor encodes distances.

As a concrete example, consider a sphere, where the points on the surface are parametrized by the polar angle θ and the azimuthal angle ϕ, so $x^\mu = \begin{pmatrix} \theta \\ \phi \end{pmatrix}$. The kinetic energy of a free particle on the sphere is given by:

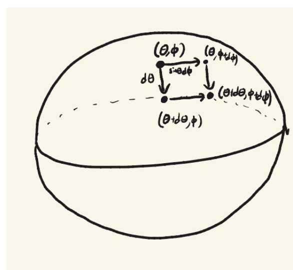
$$\mathcal{L} = \frac{m}{2}\left(\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2\right) \tag{3.30}$$

the "line element" is given by:

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{3.31}$$

so then the metric tensor on the sphere is:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}. \tag{3.32}$$



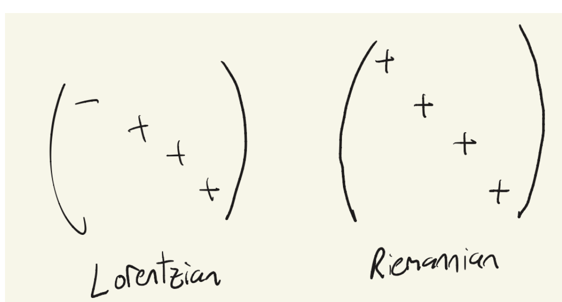
More generally, we require of the metric tensor (beyond the fact that it is a tensor) that it is symmetric:

$$g_{\mu\nu} = g_{\nu\mu} \tag{3.33}$$

that it is not degenerate:

$$\det g_{\mu\nu} \neq 0 \tag{3.34}$$

or alternatively that it has no 0 eigenvalues. We will consider the case where we have one negative eigenvalue and 3 positive eigenvalues - the $(-+++)$ signature (a so-called Lorentzian manifold). Compare this to the $(++++)$ case which is a Riemannian manifold.



The metric tensor transforms as:

$$g_{\mu'\nu'}(x') = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} g_{\mu\nu}(x) \tag{3.35}$$

We can raise and lower indices using the metric tensor:

$$A^{\mu} \to g_{\mu\nu} A^{\nu} = A_{\mu} \tag{3.36}$$

The inverse metric is the (2, 0) tensor $g^{\mu\nu}$, which can be used to raise indices:

$$A_{\mu} \to g^{\mu\nu} A_{\nu} = A^{\mu} \tag{3.37}$$

3.7 Tensor Densities

We recall the Levi-Civita symbol:

$$\tilde{\epsilon}^{\mu\nu\lambda\rho} = \begin{cases} +1 & (\mu\nu\lambda\rho) \text{is an even permutation} \\ -1 & (\mu\nu\lambda\rho) \text{is an odd permutation} \\ 0 & \text{otherwise} \end{cases} \tag{3.38}$$

Actually, we can show that this is not generally a tensor:

$$\tilde{\epsilon}^{\mu'\nu'\lambda'\rho'} = \frac{\partial x^{\mu'}}{\partial x^{\mu}} \cdots \frac{\partial x^{\rho'}}{\partial x^{\rho}} \epsilon^{\mu\nu\lambda\rho} = \det|\frac{\partial x^{\mu'}}{\partial x^{\mu}}| \tag{3.39}$$

So:

$$\tilde{\epsilon}^{\mu\nu\lambda\rho} \to \tilde{\epsilon}^{\mu'\nu'\lambda'\rho'} |\det \left(\frac{\partial x^{\mu'}}{\partial x^{\mu}} \right) |^{-1} \frac{\partial x^{\mu'}}{\partial x^{\mu}} \cdots \frac{\partial x^{\rho'}}{\partial x^{\rho}} \epsilon^{\mu\nu\lambda\rho} \tag{3.40}$$

and since the transformation law is "off" by the $|\det \left(\frac{\partial x^{\mu'}}{\partial x^{\mu}} \right) |^{-1}$ term, we say that the Levi-Civita symbol is a *tensor density*, rather than a tensor. More generally:

Definition: Tensor Densities

If T is a tensor, then if there is a metric $g_{\mu\nu}$, then $\sqrt{|\det g_{\mu\nu}|} T$ is a tensor density.

Sometimes, we write $\det g_{\mu\nu} = g$ as shorthand.

Next class we discuss curvature, and then we have enough of the mathematical technology to discuss the Einstein equations.

4 Differential Geometry II

Last lecture, we discussed the principle of equivalence, and then established general notions of differential geometry - manifolds, vectors, and tensors. We introduced the (0, 2) metric tensor $g_{\mu\nu}(x)$, which is a map $V_1 \times V_2 \to \mathbb{R}$ which component wise acts as:

$$(V_1, V_2) \to g_{\mu\nu} V_1^{\mu} V_2^{\nu} \tag{4.1}$$

This metric allows us to look at a given point of a manifold, and define a flat geometry around that point.

We also studied the notion of a *tensor density*, which transform as $A \to \sqrt{-g}A$. one example being the Levi-Cevita symbol $\tilde{\epsilon}^{\mu\nu\lambda\rho}$. We can make this into a tensor by dividing by $\sqrt{-g}$, i.e.:

$$\frac{\epsilon^{\mu\nu\lambda\rho}}{\sqrt{-g}} \tag{4.2}$$

is a genuine tensor. Note the negative sign as $g = \det g_{\mu\nu} < 0$ in our $(-+++)$ convention.

Concrete example: Flat space in spherical coordinates, we have:

$$ds^2 = -dt^2 + d\mathbf{x}^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{4.3}$$

then $g_{tt} = 1, g_{rr} = -1, g_{\theta\theta} = r^2, g_{\phi\phi} = r^2 \sin^2\theta$ and $\det g_{\mu\nu} = -r^4 \sin^2\theta$.

4.1 Derivatives, Parallel Transport, Connections

Today, we ask how we can take the derivative of a tensor field.

If we have a scalar field $\phi(x)$, then we have the derivative:

$$\frac{\partial\phi}{\partial x^{\mu}} = \partial_{\mu}\phi \tag{4.4}$$

which is a dual vector.

We can now ask if the derivative of $V^{\mu}(x)$ is a tensor:

$$\partial_{\nu} V^{\mu}(x) = \frac{\partial V^{\mu}}{\partial x^{\nu}} \tag{4.5}$$

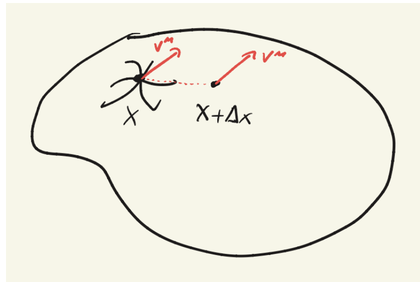
if we consider a coordinate transformation:

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^{\mu}} V^{\mu} = \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial}{\partial x^{\nu}} \left(\frac{\partial x^{\mu'}}{\partial x^{\mu}} V^{\mu} \right) = \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial x^{\mu'}}{\partial x^{\mu}} \frac{\partial V^{\mu}}{\partial x^{\nu}} + \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial^2 x^{\mu'}}{\partial x^{\nu} \partial x^{\mu}} V^{\mu} \tag{4.6}$$

if we could limit ourselves to the first term, we would have a tensor, but the second term spoils it. This problem was not there for a Lorentz transformation because that was a linear transformation (so the second derivative vanished), but in general this fails.

There is indeed a fundamental reason why this fails. When we defined a vector, we were considering a tangent vector (vector on the tangent space at the surface of the manifold). By taking a derivative, we are comparing vectors at *two different* points, x and $x + \Delta x$ - but of course the vector spaces there are going to generically be different! Thus we require the idea of parallel transport, where we consider $V^{\mu}(x)$ at x and transport to $x + \Delta x$ where we have $V^{\mu}_{\text{transport}}(x + \Delta x)$. Then in taking the derivative we can compare these two, by looking at the difference:

$$V^{\mu}(x + \Delta x) - V^{\mu}_{\text{transport}}(x + \Delta x) \tag{4.7}$$



Note that this transport is not unique, for a generic manifold. Let's consider some conditions that we would like. First, we want the transport to be linear in $V^\mu, \Delta x$:

$$V^\mu_{\text{transport}}(x+\Delta x) = V^\mu + \underbrace{\Delta V^\mu}_{O(V^\mu),\, O(\Delta x)} \tag{4.8}$$

The most general such map is:

$$V^\mu_{\text{transport}}(x+\Delta x) = V^\mu - \Gamma^\mu_{\nu\lambda}(x) V^\nu \Delta x^\lambda \tag{4.9}$$

where $\Gamma^\mu_{\nu\lambda}$ is known as the connection.

With this form, let us consider the derivative:

$$V^\mu(x+\Delta x) - \left[V^\mu(x) - \Gamma^\mu_{\lambda\nu} V^\nu \Delta x^\lambda\right] = \partial_\lambda V^\mu \Delta x^\lambda + \Gamma^\mu_{\lambda\nu} V^\nu \Delta x^\lambda = \Delta x^\lambda \boldsymbol{\nabla}_\lambda V^\mu \tag{4.10}$$

where we have introduced the covariant derivative:

$$\boldsymbol{\nabla}_\lambda = \partial_\lambda V^\mu + \Gamma^\mu_{\lambda\nu} V^\nu \tag{4.11}$$

Positing that the covariant derivative transforms like a tensor, we can deduce how the connections transform under coordinate transformations. Let's do this.

$$\boldsymbol{\nabla}_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{4.12}$$

In a new coordinate system:

$$\boldsymbol{\nabla}_{\mu'} V^{\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \partial_\mu V^\nu \tag{4.13}$$

as we saw already, with the first term of Eq. (4.12) only, the object does not transform like a tensor (we have the second derivative term). But, we will posit that the second/connection term removes this and makes sure the covariant derivative indeed transforms like a tensor. This will constrain the connection.

$$\boldsymbol{\nabla}_\mu V^\nu \to \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial V^\mu}{\partial x^\nu} + \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\mu} V^\mu + \Gamma^{\mu'}_{\nu'\lambda'} V^{\lambda'} \tag{4.14}$$

with:

$$\Gamma^{\mu'}_{\nu'\lambda'} V^{\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \left(\partial_\mu V^\mu + \Gamma^\nu_{\mu\lambda} V^\lambda\right) \tag{4.15}$$

Comparing the two expressions, we obtain the transformation law for the connection:

$$\Gamma^{\mu'}_{\nu'\lambda'} = \frac{\partial x^{\mu'}}{\partial x^{\nu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \Gamma^\nu_{\mu\lambda} - \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\mu'}}{\partial x^\nu \partial x^\lambda} \tag{4.16}$$

it is clear that Γ is not a tensor - and of course it shouldn't be, because we needed to add this term to "cancel out" the non-tensorial/second derivative part of the transformation of $\partial_\mu V^\nu$ to get a tensor.

Note that the connection is not something that comes naturally with our space, but is something we impose.

Note however that the difference between two connections is indeed a tensor:

$$(\Gamma^{\mu'}_{\nu'\lambda'} - \hat{\Gamma}^{\mu'}_{\lambda'\nu'}) \to \frac{\partial x^{\mu'}}{\partial x^{\mu}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \frac{\partial x^{\lambda}}{\partial x^{\lambda'}} \left(\Gamma - \hat{\Gamma}\right)^{\mu}_{\nu\lambda} \tag{4.17}$$

This is equivalent to the fact that the difference of two covariant derivatives (two tensors) is also a tensor:

$$\boldsymbol{\nabla}_{\mu} V^{\nu} - \hat{\boldsymbol{\nabla}}_{\mu} V^{\nu} = \underbrace{\left(\Gamma - \hat{\Gamma}\right)^{\nu}_{\mu\lambda}}_{\text{tensor}} V^{\lambda} \tag{4.18}$$

Note that the antisymmetric part of the connection is indeed a tensor:

$$T^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\nu\lambda} - \Gamma^{\mu}_{\lambda\nu} = 2\Gamma^{\mu}_{[\nu\lambda]} \tag{4.19}$$

which you can see immediately from the transformation law.

$T^{\mu}_{\nu\lambda}$ is known as a torsion tensor. A torsionless connection is that where $\Gamma^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\lambda\nu}$ and so the torsion vanishes.

4.2 Covariant Derivatives of Dual Vectors and Tensors

We've established the notion of the covariant derivative of a vector - let us also establish the notion of the covariant derivative of a dual vector and tensor. For a dual vector, we require that:

$$\boldsymbol{\nabla}_{\mu}(\omega_{\nu} V^{\nu}) = \partial_{\mu}(\omega_{\nu} V^{\nu}) \tag{4.20}$$

because the covariant derivative of the scalar should reduce to the regular derivative. So, looking at the LHS:

$$(\boldsymbol{\nabla}_{\mu}\omega_{\nu})V^{\nu} + \omega_{\nu}(\partial_{\mu} V^{\mu} + \Gamma^{\nu}_{\mu\lambda} V^{\lambda}) \tag{4.21}$$

and looking at the RHS:

$$\partial_{\mu}(\omega_{\nu} V^{\nu}) = (\partial_{\mu}\omega_{\nu})V^{\nu} + \omega_{\nu}\partial_{\mu} V^{\nu} \tag{4.22}$$

The $\omega_{\nu}\partial_{\mu} V^{\nu}$ cancels, and so:

$$\left(\boldsymbol{\nabla}_{\mu}\omega_{\nu} V^{\nu} + \omega_{\lambda}\Gamma^{\lambda}_{\mu\nu}\right) V^{\nu} = \partial_{\mu}\omega_{\nu} V^{\nu} \tag{4.23}$$

which implies:

$$\boldsymbol{\nabla}_{\mu}\omega_{\nu} = \partial_{\mu}\omega_{\nu} - \Gamma^{\lambda}_{\mu\nu}\omega_{\lambda} \tag{4.24}$$

Now, for a general tensor $T^{\mu}{}_{\nu\lambda}$ (in this case a 3-index tensor, but the logic applies more general ones), we require that again the covariant derivative of a scalar reduces to an ordinary derivative, so:

$$\boldsymbol{\nabla}_{\alpha}(T^{\mu}{}_{\nu\lambda} V_1^{\nu} V_2^{\lambda} \omega_{\mu}) = \partial_{\alpha}\left(T^{\mu}{}_{\nu\lambda} V_1^{\nu} V_2^{\lambda} \omega_{\mu}\right) \tag{4.25}$$

Thus:

$$\boldsymbol{\nabla}_{\alpha} T^{\mu}{}_{\nu\lambda} = \partial_{\alpha} T^{\mu}{}_{\nu\lambda} + \Gamma^{\mu}_{\alpha\beta} T^{\beta}{}_{\nu\lambda} - \Gamma^{\beta}_{\alpha\nu} T^{\mu}{}_{\beta\lambda} - \Gamma^{\beta}_{\alpha\lambda} T^{\mu}{}_{\nu\beta} \tag{4.26}$$

and its clear to see how this generalizes to a tensor with any number of indices.

One can show that summing over a pair of upper and lower indices commutes with a covariant derivative:

$$\boldsymbol{\nabla}_{\alpha}(T^{\mu\nu}\omega_{\nu}) = \boldsymbol{\nabla}_{\alpha} S^{\mu} \tag{4.27}$$

also note that:

$$\boldsymbol{\nabla}_{\alpha}\delta^{\mu}_{\nu} = 0 \tag{4.28}$$

4.3 The Christoffel Connection

Clearly there is something missing with this general notion of the connection - there is some sense when we do transport that there is a "preferred direction". Let us thus prove the following:

Theorem: Christoffel Connection

Let M be a manifold with metric $g_{\mu\nu}$. Then, there is a unique connection that is:

1. Torsionless, i.e. $\Gamma^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\lambda\nu}$
2. Compatible with the metric, i.e. $\nabla_{\mu} g_{\alpha\beta} = 0$.

Using these two conditions, let's find the (unique) connection. Taking the covariant derivative of the metric:

$$\nabla_{\rho} g_{\mu\nu} = \partial_{\rho} g_{\mu\nu} - \Gamma^{\lambda}_{\rho\mu} g_{\lambda\nu} - \Gamma^{\lambda}_{\rho\nu} g_{\mu\lambda} = 0 \tag{4.29}$$

Now, doing the same thing while exchanging indices (the fact that the covariant derivative is 0 remains true under such exchanges):

$$\nabla_{\mu} g_{\rho\nu} = \partial_{\mu} g_{\rho\nu} - \Gamma^{\lambda}_{\mu\rho} g_{\lambda\nu} - \Gamma^{\lambda}_{\mu\nu} g_{\rho\lambda} = 0 \tag{4.30}$$

$$\nabla_{\nu} g_{\rho\mu} = \partial_{\nu} g_{\rho\mu} - \Gamma^{\lambda}_{\nu\rho} g_{\lambda\mu} - \Gamma^{\lambda}_{\nu\mu} g_{\rho\lambda} = 0 \tag{4.31}$$

Now adding the second/third equations to the negative of the first, we find that (using that the connection is torsionless and the metric is symmetric), we are left with the terms:

$$\partial_{\mu} g_{\nu\rho} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu} = 2\Gamma^{\lambda}_{\mu\nu} g_{\rho\lambda} \tag{4.32}$$

Now, if we multiply both sides by $g^{\rho\sigma}$, the sum over the ρs in the metric gives us a kronecker delta, and so we end up with:

$$\boxed{\Gamma^{\sigma}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}(\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\rho\mu} - \partial_{\rho} g_{\mu\nu})} \tag{4.33}$$

The components of this special connection are known as the *Christoffel symbols*.

Note that by construction:

$$\nabla_{\mu} g_{\alpha\beta} = 0 \tag{4.34}$$

and we can also show that:

$$\nabla_{\mu} g^{\alpha\beta} = 0 \tag{4.35}$$

from:

$$0 = \nabla_{\mu} \delta^{\beta}_{\alpha} = \nabla_{\mu}(g_{\alpha\gamma} g^{\gamma\beta}) = g_{\alpha\gamma} \nabla_{\mu} g^{\gamma\beta} \implies \nabla_{\mu} g^{\gamma\beta} = 0 \tag{4.36}$$

This is a very nice feature, since the operation of raising/lowering indices commutes with the covariant derivative:

$$\nabla_{\alpha}(g_{\mu\nu} V^{\nu}) = g_{\mu\nu} \nabla_{\alpha} V^{\nu} \tag{4.37}$$

$$\nabla_{\alpha}(g^{\mu\nu} \omega_{\nu}) = g^{\mu\nu} \nabla_{\alpha} \omega_{\lambda} \tag{4.38}$$

Let us consider examples of the Christoffel connection, specifically the 2-D plane. In cartesian, we have:

$$ds^2 = dx^2 + dy^2 \implies g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \implies \Gamma^{\mu}_{\nu\lambda} = 0. \tag{4.39}$$

Trivial! In polar coordinates, we have:

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 \implies g_{\mu\nu} = \begin{pmatrix} g_{\rho\rho} & g_{\rho\phi} \\ g_{\phi\rho} & g_{\phi\phi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix} \tag{4.40}$$

we also have the inverse metric:

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\rho^2} \end{pmatrix} \tag{4.41}$$

we then have the six Christoffel symbols, which here we can compute directly from the formula - in practice there is an easier way, which we will introduce later:

$$\Gamma^{\rho}_{\rho\rho} = \frac{1}{2} g^{\rho\rho} \left(\partial_\rho g_{\rho\rho} + \partial_\rho g_{\rho\rho} - \partial_\rho g_{\rho\rho} \right) = 0 \tag{4.42}$$

the derivatives are zero since $g_{\rho\rho}$ is constant. We also find that:

$$\Gamma^{\rho}_{\rho\phi} = \Gamma^{\phi}_{\rho\phi} = 0. \tag{4.43}$$

Computing a nonzero one:

$$\Gamma^{\rho}_{\phi\phi} = \frac{1}{2} g^{\rho\rho} \left(2\partial_\phi g_{\rho\phi} - \partial_\rho g_{\phi\phi} \right) = \rho \tag{4.44}$$

the other two symbols are left as an exercise. We find that, even in flat space, the Christoffel symbols can be nonzero!

We end the lecture by deriving a clean formula for the divergence of a vector:

$$\boldsymbol{\nabla}_\mu V^\mu = \partial_\mu V^\mu + \Gamma^{\mu}_{\mu\nu} V^\lambda \tag{4.45}$$

Looking at the connection:

$$\Gamma^{\mu}_{\mu\nu} = \frac{1}{2} g^{\mu\lambda} \left(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu} \right) = \frac{1}{2} g^{\mu\lambda} \partial_\nu g_{\mu\lambda} \tag{4.46}$$

Consider a matrix A_{ij} and its determinant $\det A$. Under an infinitesimal change:

$$A_{ij} \to A_{ij} + \det A_{ij} \tag{4.47}$$

Then the change in the determinant is:

$$\delta(\det A) = \det A \cdot (A^{-1})^{ij} \delta A_{ij} \tag{4.48}$$

Now, if I take the derivative of the determinant of the metric, I get:

$$\partial_\alpha |\det g_{\mu\nu}| = \det g_{\mu\nu} g^{\mu\nu} \partial_\alpha g_{\mu\nu} \tag{4.49}$$

or in shorthand:

$$\partial_\alpha g = g g^{\mu\nu} \partial_\alpha g_{\mu\nu} \tag{4.50}$$

and so our connection term is:

$$\Gamma^{\mu}_{\mu\nu} = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g}) \tag{4.51}$$

and so our divergence has the form:

$$\boxed{\boldsymbol{\nabla}_\mu V^\mu = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} V^\mu)} \tag{4.52}$$

This formula will be very useful when we start thinking about conserved currents in spacetime, which satisfy the continuity equation:

$$\partial_\mu J^\mu = 0 \tag{4.53}$$

which we can write as:

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}J^\mu) = 0 \tag{4.54}$$

which (ignoring the prefactor):

$$\partial_0(\sqrt{-g}J^0) + \boldsymbol{\nabla}\cdot(\sqrt{-g}\mathbf{J}) = 0 \tag{4.55}$$

and thus we can say $\sqrt{-g}J^0$ is the charge density and $\sqrt{-g}\mathbf{J}$ the current density. $\sqrt{-g}J^\mu$ is a tensor density. We have the total charge:

$$Q = \int d^3x\sqrt{-g}J^0 \tag{4.56}$$

Next time, we will see the Christoffel symbol appear in the motion of a particle acting under a gravitational field (in curved space).

5 Particle Motion, Locally Inertial Coordinates

Last time we discussed parallel transport - because vectors are defined at a point (elements of a tangent space), in order to define a derivative in such a way that the derivative of a vector is indeed a tensor, we derived a covariant derivative, wherein:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{5.1}$$

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda. \tag{5.2}$$

If we imposed that the connection $\Gamma^\nu_{\mu\lambda}$ was metric compatible:

$$\nabla_\mu g_{\alpha\beta} = 0 \tag{5.3}$$

and torsion free:

$$\Gamma^\lambda_{[\mu\nu]} = 0 \tag{5.4}$$

then it is unique, and has the form:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) \tag{5.5}$$

known as the Christoffel connection.

Today, we discuss:

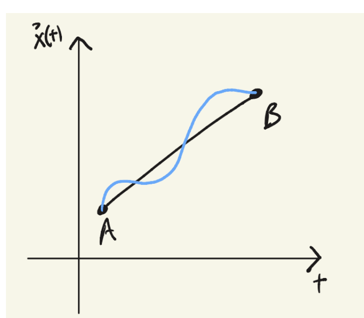
1. The motion of a particle in a gravitational field.
2. The motion in the limit of a weak field.
3. Locally inertial coordinates
4. Curvature (note - did not actually get to this today).

5.1 Particle in Gravitational Field

Assume that the metric is known:

$$g_{\mu\nu}(x) \tag{5.6}$$

In the case of flat/Minkowski space with $g_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$, we know that the particle will move in a straight line.



Indeed, we can solve for the motion via the extremization of the action:

$$S = -m \int dt \sqrt{1 - \dot{x}^2} \tag{5.7}$$

Notice in the non-relativistic limit of $|\dot{\mathbf{x}}^2| \ll 1$ we have (via Taylor expansion):

$$S \approx -m \int dt (1 - \frac{\dot{\mathbf{x}}^2}{2}) = \int dt (-m + \underbrace{\frac{m\dot{\mathbf{x}}^2}{2}}_{\text{KE}}) \tag{5.8}$$

note however that the original action we wrote is fully Lorentz invariant/compatible with relativity - parametrizing $x^\mu = x^\mu(\lambda)$, we have:

$$S = -m \int d\lambda \sqrt{-\dot{x}^\mu \dot{x}_\mu} \tag{5.9}$$

where we have chosen $\lambda = x^0 = t$, and $\dot{x}^\mu = \frac{dx^\mu}{d\lambda}$. The principle of least action then can be used to find the (straight line) trajectories.

To generalize this to curved space, we define:

$$S = -m \int d\lambda \sqrt{-g_{\mu\nu}(x(\lambda)) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \tag{5.10}$$

Let us briefly look at the symmetries of this action. First, note that it is invariant under coordinate transformations:

$$x^\mu \to x^{\mu'} = x^{\mu'}(x^\mu) \tag{5.11}$$

we have:

$$g_{\mu\nu}(x) \to g_{\mu'\nu'}(x') = \frac{\partial x^\mu}{\partial x^{\nu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu}(x) \tag{5.12}$$

and so what is under the square root stays invariant.

Another symmetry is reparametrization - if we have a parametrization $x^\mu(\lambda)$, then by taking:

$$\lambda \to \lambda' = \lambda'(\lambda) \tag{5.13}$$

we have:

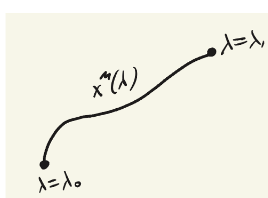
$$\frac{dx^\mu}{d\lambda'} = \frac{\partial \lambda}{\partial \lambda'} \frac{dx'}{d\lambda} \tag{5.14}$$

so:

$$\sqrt{\cdots} \to \frac{\partial \lambda}{\partial \lambda'} \sqrt{\cdots} \tag{5.15}$$

which is cancelled out by the change in the integration measure:

$$d\lambda \to d\lambda' = \frac{\partial \lambda'}{\partial \lambda} d\lambda \tag{5.16}$$



Now, I want to take the action and derive the equation of motion, by finding the geodesic/extrema of the proper time:

$$\frac{\delta S}{\delta x^\mu(\lambda)} = 0 \tag{5.17}$$

Note however there will be multiple solutions, corresponding to different choices of λ. We will make a particular choice.

Instead of extremizing the action directly, we will use a "Polyakov trick", which is a technique also used in string theory (see e.g. Polinchski Ch.1). We extremize the related action:

$$S = +\frac{m}{2}\int d\lambda \left[\frac{1}{\lambda(\eta(\lambda))} g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - \eta(\lambda)\right] = S[x^\mu(\lambda, \eta(\lambda))] \tag{5.18}$$

$\eta(\lambda)$ is a dynamical variable we minimize over, so let us solve:

$$\frac{\delta S}{\delta \eta(\lambda)} = -m\left[-\frac{1}{\eta^2}\dot{x}^2 - 1\right] = 0 \tag{5.19}$$

which is solved by:

$$\eta = -\sqrt{\dot{x}^2} \tag{5.20}$$

so plugging this back into the action:

$$S = -m\int d\lambda \sqrt{-\dot{x}^2} \tag{5.21}$$

thus we see that the two actions are equivalent in terms of producing the same equations of motions. But, the Polyakov action is more convenient, because it does not have a square root.

So let's try to extremize the Polyakov action, w.r.t $x^\mu(\lambda)$:

$$-\frac{\delta S}{\delta x^\mu(\lambda)} = \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0 \tag{5.22}$$

which is the familiar formula from classical mechanics, with:

$$S = \int d\lambda L(x^\mu, \dot{x}^\mu) \tag{5.23}$$

thus we get:

$$m\frac{d}{d\lambda}\left(\frac{1}{\eta}g_{\mu\nu}\frac{dx^\nu}{d\lambda}\right) - \frac{m}{2\eta}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}\frac{dx^\alpha}{d\lambda}\frac{d\dot{x}^\beta}{d\lambda} = 0 \tag{5.24}$$

Under reparametrization $\lambda \to \lambda' = \lambda'(\lambda)$ we have:

$$\frac{dx^\mu}{d\lambda'} = \frac{\partial\lambda}{\partial\lambda'}\frac{dx^\mu}{d\lambda} \tag{5.25}$$

with:

$$\eta'(\lambda') = \frac{\partial\lambda}{\partial\lambda'}\eta(\lambda) \tag{5.26}$$

Thus there is a parameter redundancy - let us thus choose λ such that $\eta = 1$. This choice corresponds to $\dot{x}^2 = -1$, or λ being the proper time.

With this, our equation of extremizing the action becomes:

$$g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2} + \frac{\partial g_{\mu\nu}}{\partial x^\alpha}\frac{dx^\alpha}{d\tau}\frac{dx^\nu}{d\tau} - \partial_\mu g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0 \tag{5.27}$$

Now, let us multiply this equation by $g^{\rho\mu}$, noting that $g^{\rho\mu}g_{\mu\nu} = \delta^\rho_\nu$ so:

$$\frac{d^2x^\rho}{d\tau^2} + g^{\rho\mu}(\partial_\alpha g_{\mu\beta} - \frac{1}{2}\partial_\mu g_{\alpha\beta})\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{5.28}$$

Since the metric is symmetric, we can write the term in brackets in its symmetric form, so:

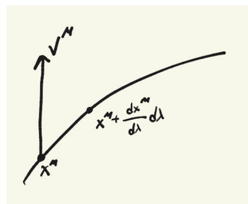
$$\frac{d^2x^\rho}{d\tau^2} + \frac{1}{2}g^{\rho\mu}(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta})\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{5.29}$$

And then we just notice that $\frac{1}{2}g^{\rho\mu}(\partial_\alpha g_{\mu\beta} + \partial_\beta g_{\mu\alpha} - \partial_\mu g_{\alpha\beta})$ is just the Christoffel connection, so:

$$\boxed{\frac{d^2x^\rho}{d\tau} + \Gamma^\rho_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0} \tag{5.30}$$

This is the geodesic equation. Let us interpret it geometrically. Suppose we have a curve and a vector at a point. Then, if we parallel transport this vector along the curve, we can think of the derivative along the curve satisfying:

$$\frac{dx^\nu}{d\lambda}\boldsymbol{\nabla}_\nu V^\mu = 0 \tag{5.31}$$



If we write $V^\mu = \frac{dx^\mu}{d\tau}$ as the velocity vector, the geodesic equation becomes:

$$\frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta}V^\alpha V^\beta = 0 \tag{5.32}$$

which we can further rewrite as:

$$V^\alpha \boldsymbol{\nabla}_\alpha V^\mu = 0 \tag{5.33}$$

i.e. if a particle moves in curved space, its velocity is parallel transported along its trajectory.

5.2 Nonrelativistic particle in Weak, Time-Independent Gravitational Field

Consider a nonrelativistic particle $|\dot{\mathbf{x}}| \ll 1$, with:

$$\dot{\mathbf{v}}^2 = |\dot{\mathbf{x}}| = O(\epsilon) \tag{5.34}$$

in a weak gravitational field, so:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{5.35}$$

with $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ and $|h_{\mu\nu}| \ll 1$ or $|h_{\mu\nu}| = O(\epsilon)$.

Our geodesic equations then tell us that (for the spatial components:)

$$\frac{d^2x^i}{d\tau^2} + \Gamma^i_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta \tag{5.36}$$

the Christoffel symbol is $\sim \partial h = O(\epsilon)$. Looking at the velocity terms, we have:

$$\dot{x}^0 = \frac{dt}{d\tau} \approx \sqrt{1 - v^2} \approx 1 \tag{5.37}$$

$$\dot{x}^i \sim v \sim \sqrt{\epsilon} \tag{5.38}$$

so the dominant contributions come from the 00/time-time terms, and so approximately the geodesic equation reads:

$$\ddot{x}^i + \Gamma^i_{00} = 0 \tag{5.39}$$

which expanding out the Christoffel symbol:

$$\Gamma^i_{00} = \frac{1}{2} g^{ik}(2\partial_0 g_{0k} - \partial_k g_{00}) \tag{5.40}$$

Now, using the time-independent assumption, the first term cancels, and to leading order $g^{ik} = \delta^{ik}$ so:

$$\Gamma^i_{00} \approx -\frac{1}{2}\partial_i h_{00} \tag{5.41}$$

and the EOM becomes:

$$\ddot{\mathbf{x}} - \frac{1}{2}\boldsymbol{\nabla} h_{00} = 0 \tag{5.42}$$

where we can view the second term as the $\mathbf{F}/m$ gravitational force term.

So defining the gravitational force as:

$$\mathbf{F}_{\text{grav}} = -m\boldsymbol{\nabla}\Phi \tag{5.43}$$

we find:

$$h_{00} = -2\Phi \tag{5.44}$$

Thus, in a weak gravitational field, the metric takes the approximate form:

$$\boxed{ds^2 = -(1+2\Phi)dt^2 + d\mathbf{x}^2} \tag{5.45}$$

If we look at the particle action:

$$S = -m\int d\lambda \sqrt{g_{\mu\nu}\dot{x}^\mu \dot{x}_\mu} \overset{\lambda = x^0 = t}{=} -m\int dt\sqrt{(1+2\Phi) - \dot{x}^2} \tag{5.46}$$

thus we find the weak-field correction to the action. Both $2\Phi, \dot{x}^2$ are $O(\epsilon)$, so we can expand:

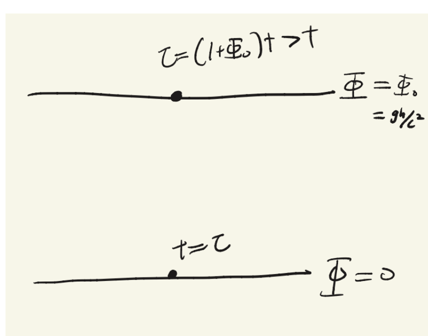
$$S = -m\int dt(1+\Phi - \frac{\dot{x}^2}{2}) = \int dt(-m + \underbrace{\frac{m\dot{x}^2}{2}}_{KE} - \underbrace{m\Phi}_{PE}) \tag{5.47}$$

Thus we see how the familiar notions of kinetic/potential energy in nonrelativistic classical mechanics is recovered!

Note that Eq. (5.45) implies that clocks at different heights flow at different rates; to see this, fix $\mathbf{x}$ so $\dot{x} = 0$. Then:

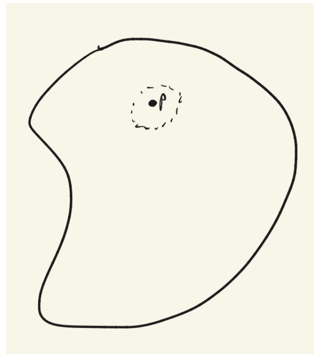
$$d\tau^2 = (1+2\Phi)dT^2 \implies \tau = (1+\Phi)t \tag{5.48}$$

Thus as we increase height, the gravitational potential Φ increases and so does the proper time.



5.3 Locally inertial coordinates

Let us start with our desiredata. We have an arbitrarily manifold, and choose a point p s.t. $x^\mu = 0$.



Expanding the metric in the vicinity of p using a taylor series, we have:

$$g_{\mu\nu}(x) = g_{\mu\nu}(0) + x^\rho \underbrace{\partial_\rho g_{\mu\nu}}_{\Gamma} + O(x^2) + \ldots \tag{5.49}$$

Now consider $x^\mu \to x^{\mu'}(x^\mu)$ s.t.:

$$g_{\mu'\nu'}(0) = \eta_{\mu'\nu'} \tag{5.50}$$

These are locally inertial coordinates, with:

$$\partial_{\rho'} g_{\mu'\nu'}\Big|_{x'=0} = 0 \implies \Gamma^\mu_{\alpha\beta} = 0 \tag{5.51}$$

i.e. (up to first order derivatives) the space locally looks Minkowski.

Without explicitly constructing it, let us show that such a coordinate system exists. We have under a coordinate transformation that:

$$g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} \tag{5.52}$$

and expanding out the coordinate transformation in a Taylor series:

$$x^{\mu'}(x^\mu) = \frac{\partial x^{\mu'}}{\partial x^\mu} x^\mu + \frac{1}{2} \frac{\partial^2 x^{\mu'}}{\partial x^\alpha \partial x^\beta} x^\alpha x^\beta + \ldots \tag{5.53}$$

Note that $\frac{\partial x^{\mu'}}{\partial x^\mu}$ has 16 coefficients, while $g_{\mu\nu}$ has 10 independent coeffients (4×4 symmetric matrix has $\frac{5 \cdot 4}{2} = 10$ independent entries). Since we have 16 coefficients worth of freedom, we can certainly find 10 independent coefficients, and moreover we have 6 remaining degrees of freedom, which corresponds to the Lorentz symmetries (3 translations, 3 boosts).

In your head, you can imagine that when we take the derivative of g:

$$\partial g \to \frac{\partial x'}{\partial x} \frac{\partial x'}{\partial x} \partial_g + \frac{\partial^2 x'}{\partial x \partial x} + \ldots \tag{5.54}$$

We have to make sure that we have enough freedom to fix $\partial_{\rho'} g_{\mu'\nu'} = 0$ for each entry, which corresponds to 40 equations to specify. If we look at the second derivative $\frac{\partial^2 x'}{\partial x^\alpha \partial x^\beta}$, we indeed get 40 equations worth of freedom to specify things (being careful to not double count as second derivatives are symmetric under interchange of order).

6 Curvature

6.1 Review

Today, we will discuss the central objects of differential geometry that play a crucial role in general relativity - namely, curvature.

Let's recall what we've looked at already.

- We looked at the notion of parallel transport, and defined a covariant derivative (which transforms like a tensor, while the bare derivative does not):

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\lambda} v^\lambda \tag{6.1}$$

- We found that there is a unique connection $\Gamma^\nu_{\mu\lambda}$ that is torsion-free and metric-compatible, and this was the Christoffel connection.

6.2 Lie Derivatives & Exterior Derivatives

In an absence of a connection, we can define derivatives in such a way that they transform like tensors. Namely, we can consider the Lie derivative between two vectors X^μ, Y^μ:

$$X : X^\mu \frac{\partial}{\partial x^\mu} \tag{6.2}$$

$$Y : Y^\mu \frac{\partial}{\partial x^\mu} \tag{6.3}$$

Then, consider the commutator between the two operators:

$$[X, Y] = X^\nu \frac{\partial}{\partial x^\nu}(Y^\mu \frac{\partial}{\partial x^\mu}) - (X \leftrightarrow Y) = X^\nu \frac{\partial Y^\mu}{\partial x^\nu}\frac{\partial}{\partial x^\mu} + X^\nu Y^\mu \frac{\partial^2}{\partial x^\nu \partial x^\mu} - (X \leftrightarrow Y) \tag{6.4}$$

by the symmetry of the order of derivatives the second derivative term cancels, and so:

$$[X, Y] = \underbrace{\left(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial x^\mu}{\partial x^\nu} \right)}_{[X,Y]^\mu} \frac{\partial}{\partial x^\mu} \tag{6.5}$$

it is obvious from this definition that the Lie derivative is a vector, and further our derivation made no reference to a connection, so it can be defined in a very general setting.

Another construction for a derivative is as follows. Consider a dual vector ω_μ and therein the $(0, 2)$ tensor:

$$\omega_{\mu\nu} = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu \tag{6.6}$$

note that the individual terms is not a tensor, but the sum is. We can convince ourselves that this is indeed a tensor by replacing the bare derivatives with covariant derivatives:

$$\nabla_\mu \omega_\nu - \nabla_\nu \omega_\mu = (\partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu}\omega_\lambda) - (\mu \leftrightarrow \nu) = \partial_\mu \omega_\nu - \partial_\nu \omega_\mu \tag{6.7}$$

where the connection term vanishes due to the torsion-free assumption. Thus we have checked that $\omega_{\mu\nu}$ is indeed a tensor. This is known as an *exterior derivative*.

Note that for a general antisymmetric tensor:

$$F_{\mu\nu} = -F_{\nu\mu} \tag{6.8}$$

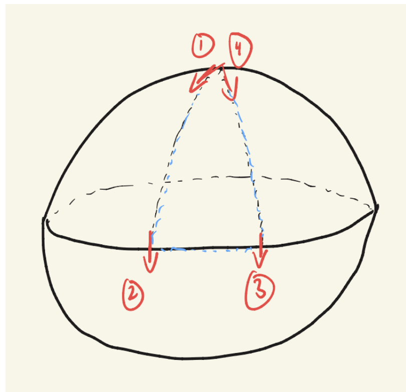
we can construct the exterior derivative:

$$\nabla_{[\lambda} F_{\mu\nu]} = \partial_{[\lambda} F_{\mu\nu]} \tag{6.9}$$

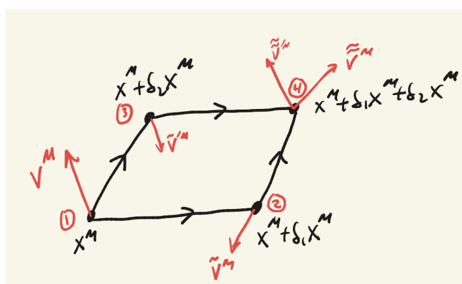
6.3 Riemann Tensor

If we are trying to characterize the non-flatness of space, we *cannot* use the Christoffel symbol $\Gamma^{\lambda}_{\mu\nu}$. Indeed, we saw that even for polar coordinates in flat 2D space that we had a nontrivial connection.

Instead, consider the following example. Take a vector on the sphere, and parallel transport it (1 to 4, along the blue path) on the figure below:



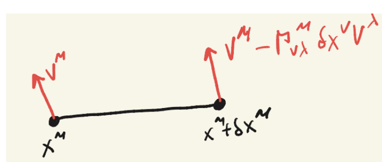
then, we can see that after the sequence of parallel transports that the vector is now pointed in a different direction! Thus, to understand curvature it becomes crucial to understand how vectors transform/are transported. In particular, consider two different contours:



And let us consider what the difference in the final parallel transported vector V^{μ} between the two contours.

Recall that going from $x^{\mu} \to x^{\mu} + \delta x^{\mu}$ a vector is parallel transported as:

$$V^{\mu} \to V^{\mu} - \Gamma^{\mu}_{\nu\lambda} \delta x^{\nu} V^{\lambda} \tag{6.10}$$



Alright, now let us compute the parallel transported vector along the two trajectories. First, going along the lower trajectory from $1 \to 2$, we have:

$$\tilde{V}^{\mu} = V^{\mu} - \Gamma^{\mu}_{\nu\lambda}(x + \frac{\delta_1 x}{2})\delta_1 x^{\nu} V^{\lambda} + O(\delta x^3) \tag{6.11}$$

Actually this is an approximate expression. To do this exactly, we integrate the connection along the trajectory, but here we simply evaluate it at the midway point (from $1 \to 2$) and discard the higher-order errors.

Now, taking this vector from $2 \to 4$, we get:

$$\tilde{\tilde{V}}^\mu = \tilde{V}^\mu - \Gamma^\mu_{\rho\sigma}(x + \delta_1 x + \frac{1}{2}\delta_2 x)\delta x_2^\rho \tilde{V}^\sigma \tag{6.12}$$

and substituing in our expression for $\tilde{V}$:

$$\tilde{\tilde{V}}^\mu = V^\mu - \Gamma^\mu_{\nu\lambda}(x + \frac{\delta_1 x}{2})\delta_1 x^\nu V^\lambda - \Gamma^\mu_{\rho\sigma}(x + \delta_1 x + \frac{1}{2}\delta_2 x)\delta x_2^\rho V^\sigma + \Gamma^\mu_{\rho\sigma}(x)\delta_2 x^\rho \Gamma^\sigma_{\nu\lambda}(x)\delta_1 x^\nu V^\lambda + O((\delta x)^3) \tag{6.13}$$

Note that in the last term we can just approximate the arguments of the Γs as x because the term is already quadratic in δx, so the error in changing the argument is order $O((\delta x)^3)$.

It is easy to convince yourself that going along the other contour, we only need exchange $1 \leftrightarrow 2$. We also interchange $\nu \leftrightarrow \rho, \lambda \leftrightarrow \sigma$ (we are free to call the dummy indices whatever we like):

$$\tilde{\tilde{V}}'^\mu = V^\mu - \Gamma^\mu_{\rho\sigma}(x + \frac{\delta_2 x}{2})\delta_2 x^\rho V^\sigma - \Gamma^\mu_{\nu\lambda}\delta_1 x^\nu V^\lambda + \Gamma^\mu_{\nu\lambda}\delta_1 x^\nu \Gamma^\lambda_{\rho\sigma}\delta_2 x^\rho V^\lambda + O((\delta x)^3) \tag{6.14}$$

Now, let's look at the difference!

$$\begin{aligned} \tilde{\tilde{V}}^\mu - \tilde{\tilde{V}}'^\mu = &\left(\Gamma^\mu_{\nu\lambda}(x + \delta_2 x + \frac{\delta_1 x}{2}) - \Gamma^\mu_{\nu\lambda}(x + \frac{\delta_1 x}{2})\right)\delta_1 x^\nu V^\lambda \\ &- \left(\Gamma^\mu_{\rho\sigma}(x + \frac{\delta_2 x}{2}) - \Gamma^\mu_{\rho\sigma}(x + \delta_1 x + \frac{\delta_2 x}{2})\right)\delta_2 x^\rho V^\sigma \\ &+ \left(\Gamma^\mu_{\rho\sigma}\Gamma^\sigma_{\nu\lambda} - \Gamma^\mu_{\nu\sigma}\Gamma^\sigma_{\rho\lambda}\right)\delta_1 x^\rho \delta_2 x^\nu V^\lambda \end{aligned}$$

There is some nontrivial probability that this formula is correct. Writing the differences in connection in terms of derivatives (and fixing some index errors), we get:

$$\begin{aligned} \tilde{\tilde{V}}^\mu - \tilde{\tilde{V}}'^\mu = &\,\partial_\rho \Gamma^\mu_{\nu\lambda}\delta_1 x^\nu \delta_2 x^\rho V^\lambda \\ &- \partial_\nu \Gamma^\mu_{\rho\sigma}\delta_1 x^\nu \delta_2 x^\rho V^\sigma \\ &+ \left(\Gamma^\mu_{\rho\sigma}\Gamma^\sigma_{\nu\lambda} - \Gamma^\mu_{\nu\sigma}\Gamma^\sigma_{\rho\lambda}\right)\delta_1 x^\nu \delta_2 x^\rho V^\lambda \end{aligned}$$

which has an even better probability to be correct. We can package this all as:

$$\boxed{\tilde{\tilde{V}}^\mu - \tilde{\tilde{V}}'^\mu = R^\mu_{\ \lambda\rho\nu}\delta_1 x^\nu \delta_2 x^\rho V^\lambda} \tag{6.15}$$

where:

$$\boxed{R^\mu_{\ \lambda\rho\nu} = \partial_\rho \Gamma^\mu_{\nu\lambda} - \partial_\nu \Gamma^\mu_{\rho\lambda} + \Gamma^\mu_{\rho\sigma}\Gamma^\sigma_{\nu\lambda} - \Gamma^\mu_{\nu\sigma}\Gamma^\sigma_{\rho\lambda}}. \tag{6.16}$$

Thus we find the vector difference of parallel transport along two different contours (to quadratic order) is proportional to the product of the δxs, the original vector, and the coefficient of proportionality is the Riemann, or curvature tensor $R^\mu_{\ \lambda\rho\nu}$, which encodes the curvature of space. This is a coordinate free expression, and at this point does not specifically refer to a given metric.

6.4 Properties of the Riemann Tensor

We can check the following fact - that the Riemann tensor checks the non-commutativity of covariant derivatives. For a scalar function, we have:

$$[\boldsymbol{\nabla}_\mu, \boldsymbol{\nabla}_\nu]f = \boldsymbol{\nabla}_\mu \boldsymbol{\nabla}_\nu f - \boldsymbol{\nabla}_\nu \boldsymbol{\nabla}_\mu f = \underbrace{(\boldsymbol{\nabla}_\mu)(\partial_\nu f)}_{\partial_\mu \partial_\nu f - \Gamma^\lambda_{\mu\nu}\partial_\lambda f} - (\mu \leftrightarrow \nu) = 0 \tag{6.17}$$

where in the last equality we use the symmetry of derivatives and the torsion-free assumption on the connection. Meanwhile, if we look at the commutator acting on a vector:

$$[\nabla_\mu, \nabla_\nu]V^\lambda = R^\lambda{}_{\rho\mu\nu}V^\rho \tag{6.18}$$

Using this, we can show:

$$[\nabla_\mu, \nabla_\nu]\omega_\lambda = -R^\rho{}_{\lambda\mu\nu}\omega_\rho \tag{6.19}$$

which follows from:

$$[\nabla_\mu, \nabla_\nu](\omega_\lambda V^\lambda) = 0 \tag{6.20}$$

as the commutator acting on a scalar function vanishes.

Some properties of $R^\lambda{}_{\rho\mu\nu}$:

1. $\boxed{R^\rho{}_{\lambda\mu\nu} = -R^\rho{}_{\lambda\nu\mu}}$. This follows from $[\nabla_\mu, \nabla_\nu] \sim R^\cdot{}_{\cdot\mu\nu}$.

2. $\boxed{R^\rho{}_{[\lambda\mu\nu]} = 0}$. This can be seen from $[\nabla_{[\mu}, \nabla_\nu]\omega_{\lambda]} \sim \nabla_{[\mu}\nabla_\nu\omega_{\lambda]} = \nabla_{[\mu}\partial_\nu\omega_{\lambda]} = \partial_{[\mu}\partial_\nu\omega_{\lambda]} = 0$.

3. If we restrict to the case that the connection is metric compatible, then we obtain further properties. Using the metric, we can lower and index and obtain the object $R_{\rho\lambda\mu\nu}$. We can then show that:

$$\boxed{R_{\rho\lambda\mu\nu} = -R_{\lambda\rho\mu\nu}} \tag{6.21}$$

which follows from:

$$[\nabla_\mu, \nabla_\nu]g_{\alpha\beta} = 0 \tag{6.22}$$

- one covariant derivative of the metric vanishes, so of course two will! Expanding out the commutator, we get:

$$-R^\gamma{}_{\alpha\mu\nu}g_{\gamma\beta} - R^\gamma{}_{\beta\mu\nu}g_{\alpha\gamma} = -R_{\beta\alpha\mu\nu} - R_{\alpha\beta\mu\nu} \tag{6.23}$$

and since this is equal to zero, we get the antisymmetry property.

4. There is another nontrivial property, namely the symmetry:

$$\boxed{R_{\rho\lambda\mu\nu} = R_{\mu\nu\rho\lambda}} \tag{6.24}$$

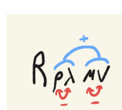
which we will show using local inertial coordinates. Let us specify a point x, and pick a coordinate system such that $\gamma_{\mu\nu}(x) = \eta_{\mu\nu}$ and $\partial_\lambda g_{\mu\nu} = 0, \Gamma^\lambda_{\mu\nu} = 0$. I.e. we pick an arbitrary point on our manifold, and choose the coordinate system so things look (locally) Minkowski. Evaluating the Riemann tensor, we have:

$$R \sim \partial\Gamma - \partial\Gamma + \Gamma\Gamma - \Gamma\Gamma \tag{6.25}$$

where the third/fourth term vanish as Γ vanishes. But the derivative does not vanish (but can be computed simply from $\Gamma \sim g\partial g$), and after the calculation is finished we find that in this local inertial coordinate system:

$$R_{\sigma\rho\mu\nu} = \frac{1}{2}(\partial_\mu\partial_\rho g_{\sigma\nu} - \partial_\mu\partial_\sigma g_{\nu\rho} - \partial_\nu\partial_\rho g_{\sigma\mu} + \partial_\nu\partial_\sigma g_{\nu\rho}) \tag{6.26}$$

by which the symmetry property can be checked. Summarizing the properties thus far:



5. The last property we want to show is the Bianchi identity. This one we can again show without a metric:

$$\boxed{\nabla_{[\lambda} R^{\rho}_{\ |\sigma|\mu\nu]} = 0} \tag{6.27}$$

To see this, consider two sequences by which we carry out the antisymmetrization:

$$\nabla_{[\lambda} \nabla_{\mu} \nabla_{\nu]} \omega_\alpha = \nabla_{[\lambda} (\nabla_{[\mu} \nabla_{\nu]]}) = \nabla_\lambda (R^{\beta}_{\ \alpha\mu\nu} \omega_\alpha) \tag{6.28}$$

$$\nabla_{[\lambda} \nabla_{\mu} \nabla_{\nu]} \omega_\alpha = \nabla_{[[\lambda} \nabla_{\mu]} \nabla_{\nu]} \omega_\alpha = R(\nabla_\alpha) \tag{6.29}$$

6. What are the number of independent components of $R_{\alpha\beta\mu\nu}$? First, view R as a symmetric matrix in $\alpha\beta, \mu\nu$. Then, if we look at $(\alpha\beta)$, we have 6 possible values (as there would be in a 4×4 antisymmetric matrix). Then, we only need ask how many independent components are in a 6×6 symmetric matrix, which is:

$$\frac{6 \cdot 7}{2} = 21 \tag{6.30}$$

seems like a lot, but is a lot less than the 4^4 components we have without symmetry! Actually, there is one more condition, namely property (2) of:

$$R_{\rho[\lambda\mu\nu]} = 0 \tag{6.31}$$

which places one more independent constraint that we have not yet enforced: namely:

$$R_{[\rho\lambda\mu\nu]} = 0 \tag{6.32}$$

so we have $\boxed{20}$ independent components (in 4 spatial dimensions). In $d = 3$ spacetime dimensions, This reduces to the number of components of a 3×3 symmetric matrix, i.e. 6 independent components. In $d = 2$ there is only one independent component, R_{0101}.

Next time, we look at the Ricci tensor, scalar, and the Einstein tensors. Einstein was confused for a year about the difference between these two, but using the knowledge we have in hindsight, we will quickly pave the road to understanding GR.

7 Ricci, Relative Velocity/Acceleration, and Einstein's Equation

Last time, we discussed the Riemann tensor:

$$[\nabla_\mu, \nabla_\nu]V^\lambda = R^\lambda_{\rho\mu\nu} V^\rho \tag{7.1}$$

and its symmetries:

$$R_{\alpha[\beta\mu\nu]} = 0 \tag{7.2}$$

$$\nabla_{[\gamma} R_{\alpha\beta]\mu\nu} = 0 \tag{7.3}$$

Today, we will go through:

- The Ricci tensor, Ricci scalar, and Weyl tensor.
- Geodesic derivation.
- Motivating the Einstein equation.

7.1 Ricci Tensor, Ricci Scalar, Weyl Tensor

We define the Ricci tensor as:

$$\boxed{R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu}} \tag{7.4}$$

still called R because Ricci and Riemann both had names starting with R. We can also define the Ricci scalar:

$$R = g^{\mu\nu} R_{\mu\nu} = R^\mu{}_\mu \tag{7.5}$$

Note that the symmetry of the Riemann tensor implies the symmetry of the Ricci tensor:

$$\boxed{R_{\alpha\mu\beta\nu} = R_{\beta\nu\alpha\mu} \implies R_{\mu\nu} = R_{\nu\mu}} \tag{7.6}$$

It thus has $\frac{5 \cdot 4}{2} = 10$ independent components (c.f. the Riemann tensor, which has 20 independent components).

We can also define the Weyl tensor:

$$\boxed{C_{\rho\sigma\mu\nu} = R_{\rho\sigma\mu\nu} - \frac{2}{n-2}\left(g_{\rho\mu} R_{\nu\sigma} - g_{\rho\nu} R_{\mu\sigma} - g_{\sigma\mu} R_{\nu\rho} + g_{\sigma\nu}\right) + \frac{1}{(n-1)(n-2)}(g_{\rho\mu} g_{\nu\sigma} - g_{\rho\nu} g_{\mu\sigma})R} \tag{7.7}$$

C satisfies the symmetries of the Riemann tensor:

$$C_{\rho\sigma\mu\nu} = -C_{\sigma\rho\mu\nu} = -C_{\rho\sigma\nu\mu} \tag{7.8}$$

Symmetric under interchange of the pairs:

$$C_{\rho\sigma\mu\nu} = +C_{\mu\nu\rho\sigma} \tag{7.9}$$

And:

$$C_{\alpha[\beta\mu\nu]} = 0 \tag{7.10}$$

However, a new property is that it is traceless:

$$C^{\rho}_{\sigma\rho\nu} = 0 \tag{7.11}$$

Recalling that $R^{\rho}{}_{\sigma\rho\nu} = R_{\sigma\nu}$, the Weyl tensor can be interpreted as the Riemann tensor with the Ricci part taken out.

The Weyl tensor has a simple transformation rule. If we take the metric and multiply it by a function of space:

$$g_{\mu\nu}(x) \to \Omega(x) g_{\mu\nu}(x) \tag{7.12}$$

Then:

$$C_{\mu\nu\rho\sigma} \to \Omega^{c}(x) C_{\mu\nu\rho\sigma} \tag{7.13}$$

with c a constant (left to the reader to determine). The Riemann and Ricci tensor do not have such a simple property.

7.2 Bianchi Identity for the Ricci Tensor

Writing out the Bianchi identity for the Riemann tensor:

$$\nabla_{\lambda} R_{\rho\sigma\mu\nu} + \nabla_{\mu} R_{\rho\sigma\nu\lambda} + \nabla_{\nu} R_{\rho\sigma\lambda\mu} = 0 \tag{7.14}$$

Let us multiply this equation by two copies of the inverse metric:

$$g^{\sigma\nu} g^{\rho\mu} [\nabla_{\lambda} R_{\rho\sigma\mu\nu} + \nabla_{\mu} R_{\rho\sigma\nu\lambda} + \nabla_{\nu} R_{\rho\sigma\lambda\mu}] = 0 \tag{7.15}$$

Now, since we know that the covariant derivative of the metric is zero, we can push the metric into the derivatives. This gives (recalling the definition of the Ricci tensor and scalar, as well as the antisymmetry properties of the Riemann tensor):

$$\nabla_{\lambda} R - \nabla^{\mu} R_{\mu\lambda} - \nabla^{\nu} R_{\nu\lambda} = 0 \tag{7.16}$$

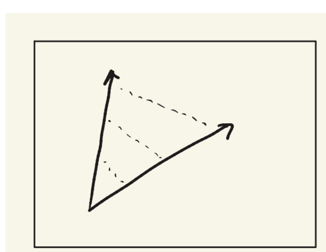
where $\nabla^{\mu} = g^{\mu\nu} \nabla_{\nu}$.

Thus, we have the Bianchi identity for the Ricci tensor:

$$\boxed{\nabla^{\mu} R_{\mu\nu} - \frac{1}{2} \partial_{\nu} R = 0}. \tag{7.17}$$

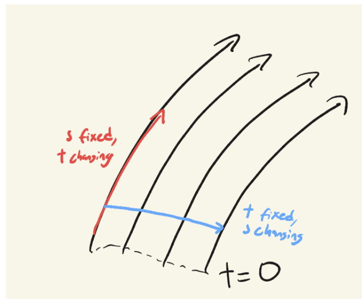
7.3 Geodesics and Relative Velocity/Acceleration

One property of a flat space - if two particles travel at constant velocity, the distance between them grows linearly.



Thus, the relative acceleration is zero. This is not true in curved space! A very quick counterexample is looking at a sphere, where the distance between the two particles grows, then shrinks, than grows... so there is clearly a nonzero relative acceleration.

Let us make this notion of relative velocity/acceleration precise.What we will now consider is a family of geodesics, corresponding to particles at different velocities at $t = 0$.



Specifically, consider a one-parameter of geodesics $\gamma_s(t)$, where if we fix s and increase t we travel along a given geodesic and if we fix t and change s we go in between different geodesics. So, we can consider trajectories $x^\mu(s,t)$, which for fixed s is a geodesic.

Now, we can define two vectors:

$$T = \frac{\partial}{\partial t} = \frac{\partial x^\mu}{\partial t}\frac{\partial}{\partial x^\mu} = T^\mu \frac{\partial}{\partial x^\mu} \tag{7.18}$$

$$S = \frac{\partial}{\partial s} = \frac{\partial x^\mu}{\partial s}\frac{\partial}{\partial x^\mu} = S^\mu \frac{\partial}{\partial x^\mu}. \tag{7.19}$$

And we can define the relative velocity:

$$\boxed{v^\mu = (T^\nu \nabla_\nu) S^\mu} \tag{7.20}$$

and the relative acceleration:

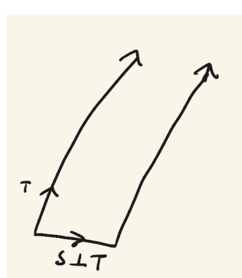
$$\boxed{a^\mu = (T^\nu \nabla_\nu) V^\mu} \tag{7.21}$$

we choose t to be the proper time, such that:

$$T^\mu T_\mu = -1. \tag{7.22}$$

We also specify the initial condition that $S \perp T$ at $t = 0$, i.e.:

$$T \cdot S = T^\mu S_\mu = 0 \tag{7.23}$$



Claim: This remains 0 at all subsequent times t. For this, we require one easy mathematical fact, namely that T, S have zero lie derivative:

$$[T, S] = 0. \tag{7.24}$$

This follows as t, s are independent. This then implies:

$$T^\nu \nabla_\nu S^\mu - S^\nu \nabla_\nu T^\mu = 0 \tag{7.25}$$

where we are free to replace the regular derivative with the covariant derivative inside the Lie derivative. Thus:

$$T^\nu \nabla_\nu S^\mu = S^\nu \nabla_\nu T^\mu \tag{7.26}$$

So if we look at:

$$\frac{\partial}{\partial t}(T \cdot S) = T^\mu \nabla_\mu (T^\nu S_\nu) = T^\mu (\nabla_\mu T^\nu) S_\nu + T^\mu T^\nu (\nabla_\mu S_\nu) \tag{7.27}$$

For the first term, since the particle moves along the geodesic:

$$T^\mu (\nabla_\mu T^\nu) = 0 \tag{7.28}$$

For the second term, we can use Eq. (7.26) to obtain:

$$T^\mu T^\nu (\nabla_\mu S_\nu) = T^\nu S^\mu \partial_\mu T_\nu = \frac{1}{2} S^\mu \nabla_\mu (T^\nu T_\nu) = 0 \tag{7.29}$$

because we have chosen $T^\nu T_\nu = -1$, the derivative which is clearly a constant. Thus:

$$\frac{\partial}{\partial t}(T \cdot S) = 0 \implies T \cdot S = 0 \text{ for all time} \tag{7.30}$$

With this in hand, let us derive the formula for the relative acceleration:

$$\begin{aligned}
a^\mu &= T^\rho \nabla_\rho V^\mu \\
&= T^\rho \nabla_\rho (T^\sigma \nabla_\sigma S^\mu) \\
&= T^\rho \nabla_\rho (S^\sigma \nabla_\sigma T^\mu) \\
&= T^\rho \nabla_\rho S^\sigma \nabla_\sigma T^\mu + T^\rho S^\sigma \nabla_\rho \nabla_\sigma T^\mu \\
&= S^\rho \nabla_\rho T^\sigma \nabla_\sigma T^\mu + T^\rho S^\sigma (\nabla_\sigma \nabla_\rho T^\mu + R^\mu{}_{\nu\rho\sigma} T^\nu) \\
&= S^\rho \nabla_\rho T^\sigma \nabla_\sigma T^\mu + S^\sigma (\nabla_\sigma (T^\rho \nabla_\rho T^\mu) - \nabla_\sigma T^\rho \nabla_\rho T^\mu)
\end{aligned}$$

$T^\rho \nabla_\rho T^\mu$ vanishes because we have a geodesic, and the two $\nabla_\sigma T^\rho \nabla_\rho T^\mu$ terms cancel, so we are left with:

$$\boxed{a^\mu = R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma} \tag{7.31}$$

i.e. the acceleration only depends on the Riemann tensor!

7.4 Einstein's Equation

With this technology developed, let us motivate/build up to Einstein's Equation for GR.

Recall what the challenge of writing down a relativistic theory was - we need to write down relativistic versions of Newton's second law:

$$m\ddot{\mathbf{x}} = -m\nabla\Phi. \tag{7.32}$$

The relativistic version of the above is that a particle moves along a geodesic, and that a weak gravitational field responds like:

$$g_{00} = -(1 + 2\Phi) \tag{7.33}$$

for $\Phi \ll 1$. In a previous lecture, we derived the geodesics for a non-relativistic particle in a weak field, and showed that Newton's equation was reproduced.

The geodesic equation, as we have seen/derived, is:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{7.34}$$

and this reproduces Newton's second law when $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $h_{\mu\nu} \ll 1$ (weak field) and $\ddot{\mathbf{x}} \ll 1$ (non-relativistic).

Ok, so we're set there! The other equation to reproduce is that for the gravitational potential:

$$\nabla^2\Phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x}) \tag{7.35}$$

or the integral equation:

$$\Phi(\mathbf{x}) = -\int d\mathbf{y}\frac{G\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \tag{7.36}$$

So our task left to do is to generalize Eq. (7.35) that is relativistic, and reduces to to the Poisson equation when matter is nonrelativistic and gravity is weak ($\Phi \ll 1, g_{\mu\nu} \approx \eta_{\mu\nu}$).

Let's think about a perfect fluid:

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + Pg^{\mu\nu} \tag{7.37}$$

In the non-relativistic limit, we have that $u^\mu \approx (1, \mathbf{0})$, wherein:

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} \tag{7.38}$$

i.e.:

$$T^{00} = \rho \text{ (energy density)} \tag{7.39}$$

$$T^{ij} = P\delta^{ij} \text{ (pressure)} \tag{7.40}$$

But if I consider a piece of perfect fluid, the energy density is dominated by the mass density of the fluid, thus:

$$\rho \gg P \tag{7.41}$$

Thus $T^{00}(\mathbf{x}) = \rho^{00}(\mathbf{x})$ dominates.

The only relativistically covariant candidate that replace the mass density in the non-relativistic case is the mass-energy tensor. Thus, we expect this to enter our equation:

$$? = \kappa T^{\mu\nu} \tag{7.42}$$

with $\kappa \sim G$. Poisson equation should be obtained in the $\mu = \nu = 0$ NR limit... looking at the LHS of the Poisson equation, this motivates the LHS of our mystery equation with something like:

$$\nabla^\lambda \nabla_\lambda g_{\mu\nu} \tag{7.43}$$

Of course this cannot be the case because the covariant derivative of g vanishes. But, we recall that the Christoffel symbol $\Gamma^\mu_{\nu\lambda}$ contains ∂g terms, and that the Riemann tensor $R_{\mu\nu\sigma\rho}$ contains $\partial\partial g$ terms... so this is our candidate! Except we have too many indices on the Riemann tensor, which leads us to consider the Ricci tensor $R_{\mu\nu}$ instead. This gives us our first proposal for the equation that we want:

$$R_{\mu\nu} = \kappa T_{\mu\nu} \tag{7.44}$$

Does this give the correct equations in the weak gravity + nonrelativistic matter limit?

Let's check. Take:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{7.45}$$

with $h_{\mu\nu} \ll 1$. then:

$$R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_\mu \partial_\sigma h_{\rho\sigma} - \partial_\mu \partial_\rho h_{\nu\sigma} - \partial_\nu \partial_\sigma h_{\rho\mu} + \partial_\nu \partial_\rho h_{\mu\sigma}) + O(h^2) \tag{7.46}$$

where we have substituted in our $g_{\mu\nu}$ expression into the expression for the Riemann tensor (also note that the inverse metric has the form):

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} + O(h^2) \tag{7.47}$$

Now, if we look at the Ricci tensor:

$$R_{\sigma\nu} = g^{\rho\mu} R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_\sigma \partial^\mu h_{\mu\nu} - \partial^2 h_{\nu\sigma} - \partial_\nu \partial_\sigma h^\mu_{\ \mu} + \partial_\nu \partial^\nu h_{\mu\sigma}) \tag{7.48}$$

Let's suppose everything is static in time so $\partial_0 = 0$. Looking at the time-time component:

$$R_{00} = -\frac{1}{2}\nabla^2 h_{00} = \nabla^2 \Phi \tag{7.49}$$

so our proposal seems to work (in the sense that it reproduces the Poisson equation), if we take $\kappa = 4\pi G$!

Well.... unfortunately it doesn't quite work. Recall that the stress-energy tensor is conserved:

$$\nabla_\nu T^{\mu\nu} = 0 \tag{7.50}$$

But if we now take the derivative of both sides of our proposed equation:

$$\nabla^\nu R_{\mu\nu} = \kappa \nabla^\mu T_{\mu\nu} = 0 \tag{7.51}$$

We can then use the Bianchi identity on the LHS to get:

$$\frac{1}{2}\nabla_\mu R = 0 \implies \partial_\mu R = 0 \tag{7.52}$$

but then this implies that:

$$R = \text{const.} \tag{7.53}$$

and hence:

$$T^\mu_{\ \mu} = \text{const.} \tag{7.54}$$

but of course the mass density of our universe is not constant/homogenous! So our equation has a problem.

Ok, time for a second version!! After months of thinking, Einstein comes up with:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu} \tag{7.55}$$

Now, if I take the derivative w.r.t. ν on both sides, I get:

$$\nabla^\nu R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \nabla^\nu R = \kappa \nabla^\nu T_{\mu\nu} = 0 \tag{7.56}$$

but the LHS is forced to be zero already from the Bianchi identity! So we do not get an unphysical constraint on $\partial_\mu R$ or $T^\mu_{\ \mu}$ as a result of conservation. If we look at the nonrelativistic limit:

$$R_{00} = \nabla^2 \Phi \tag{7.57}$$

$$R = g^{\mu\nu} R_{\mu\nu} = -\boldsymbol{\nabla}^2 \Phi \tag{7.58}$$

Thus:

$$G_{00} = R_{00} + \frac{1}{2} g_{00} \boldsymbol{\nabla}^{\nu} R = \frac{1}{2} \boldsymbol{\nabla}^2 \Phi \tag{7.59}$$

Thus with $\kappa = 8\pi G$ we obtain the Poisson equation! It turns out that Hilbert also obtained this equation, a few days before Einstein, using the least action principle. But, we conclude with the celebrated equation:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{7.60}$$

8 Einstein's Equation II

8.1 Einstein's Equation from the action principle

Last time, we saw the Einstein equation motivated by requiring that the equation be second order in the derivative of the metric, and in the Newtonian limit it reduced to the Poisson equation. This is how Einstein came up with the result. However, there is a second way of deriving the same equation, due to Hilbert (a few days before Einstein). This fits into a more general framework of physics as we know it now - most physical systems can be described by some action principle (and if not, the systems are a bit more exotic, e.g. being constrained to move on a sphere - in fundamental physics such constraints are a bit more rare). Thus we ask what kind of action we may write down to get the Einstein equation out.

Before writing down the action for the gravitational field + matter, let us recall what the action for electromagnetism looks like:

$$S = S_{\text{em}}[A_\mu] + S_{\text{matter}}[\phi, A_\mu] \tag{8.1}$$

where A_μ is the gauge field and ϕ are the degrees of freedom of matter. Looking at the first/free term:

$$S_{\text{em}}[A_\mu] = \frac{1}{4}\int d^4x F_{\mu\nu}F^{\mu\nu} \tag{8.2}$$

with:

$$F_{\mu\nu} - \partial_\mu A_\nu - \partial_\nu A_\mu \tag{8.3}$$

wherein if we vary the action w.r.t A_μ we obtain:

$$0 = \frac{\delta S}{\delta A_\mu} = \frac{\delta S_{em}[A_\mu]}{A_\mu} + \frac{\delta S_{\text{matter}}[\phi, A_\mu]}{\delta A_\mu} = 0 \tag{8.4}$$

The first term gives $\partial_\mu F^{\mu\nu}$ and the second gives the current j^μ, leading to the Maxwell equation:

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{8.5}$$

So, let's write down a similar form for the gravitational action:

$$S = S_{\text{grav}}[g_{\mu\nu}] + S_{\text{matter}}[\phi, g_{\mu\nu}] \tag{8.6}$$

8.2 Gravitational term

The first term should be something of the form:

$$\int d^4x\sqrt{-g(x)}\mathcal{L}(g_{\mu\nu}, \partial g_{\mu\nu}) \tag{8.7}$$

We want the above to be invariant under coordinate transformations, and we want $\mathcal{L}(g_{\mu\nu}, \partial g_{\mu\nu})$ to be a scalar. Thus, Hilbert made the guess that:

$$S_{\text{grav}} \propto S_H = \int d^4x\sqrt{-g}R \tag{8.8}$$

with R the Ricci scalar. However, there is a problem here - as we discussed last lecture, R contains second derivatives of g:

$$R \sim \partial\Gamma + \Gamma\Gamma + \ldots \sim \partial(\partial g) + (\partial g)^2 + \ldots \tag{8.9}$$

which to seems to be incompatible with our usual notion in classical mechanics that $\mathcal{L}$ only depends on terms up to first derivatives of g. The saving realization is that we can integrate by parts:

$$f(q)\ddot{q} \to \partial_t(f\dot{q}) - f'(q)\dot{q}^2 \tag{8.10}$$

So - up to a total derivative, we can write:

$$S_H = \int d^4x \sqrt{-g}\mathcal{L}(g_{\mu\nu}, \partial_\lambda g_{\mu\nu}) \tag{8.11}$$

For more detail, see Landau & Lifschiz' classical theory of fields.

If we then vary the Hilbert action w.r.t. the metric, we get:

$$\frac{\delta S_H}{\delta g_{\mu\nu}} = 0 \implies R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0 \tag{8.12}$$

which are the Einstein equations in empty space! (i.e. without matter).

Let us see how this is derived. First, let us consider how to take the functional derivative $\frac{\delta S}{\delta g^{\mu\nu}}$. Consider a variation:

$$g^{\mu\nu} \to g^{\mu\nu} + \delta g^{\mu\nu} \tag{8.13}$$

for a small variation. Then, since:

$$S_H = \int d^4x \sqrt{-g}g^{\mu\nu}R_{\mu\nu} \tag{8.14}$$

we get the variation:

$$\delta S_H = \int d^4x \left(\underbrace{\sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu}}_{\delta S_1} + \underbrace{\sqrt{-g}R_{\mu\nu}\delta g^{\mu\nu}}_{\delta S_2} + \underbrace{R\delta\sqrt{-g}}_{\delta S_3} \right) \tag{8.15}$$

The second term already is in the form we like, as it is $\propto \delta g^{\mu\nu}$:

$$\delta S_2 = \int d^4x \sqrt{-g}R_{\mu\nu}\delta g^{\mu\nu} \tag{8.16}$$

Next, recall when we vary matrices:

$$M_{ij} \to M_{ij} + \delta M_{ij} \tag{8.17}$$

since $\det M = \exp(\mathrm{Tr}\ln M)$, we find that $M \to M + \delta M$ results in:

$$\delta \det M = \exp(\mathrm{Tr}\ln M + \mathrm{Tr}(M^{-1}\delta M)) = \det M \mathrm{Tr}(M^{-1}\delta M) \tag{8.18}$$

Thus looking at the third term, for $M = g^{\mu\nu}$, we have $g = \det g = \frac{1}{\det M}$ and so:

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu} \tag{8.19}$$

Thus:

$$\delta S_3 = \int d^4x \left[-\frac{1}{2}R\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu} \right] \tag{8.20}$$

Note if we have lower indices then $g_{\alpha\beta} \to g_{\alpha\beta} - g_{\alpha\mu}g_{\beta\nu}\delta g^{\mu\nu}$ and so this term would be $\int d^4x \left[\frac{1}{2}R\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}\right]$ instead.

Ok, those are two of the terms, now we have the last one... For this, recall the definition of the Riemann tensor:

$$R^\rho_{\mu\lambda\nu} = \partial_\lambda \Gamma^\rho_{\nu\mu} + \Gamma^\rho_{\lambda\sigma}\Gamma^\sigma_{\nu\mu} - (\lambda \leftrightarrow \nu) \tag{8.21}$$

So under $g^{\mu\nu} \to g^{\mu\nu} + \delta g^{\mu\nu}$ we have to figure out how Γ changes... the direct calculation is quite cumbersome. But, we can use some tricks. The Christoffel symbol changes as:

$$\Gamma^\rho_{\nu\mu} \to \Gamma^\rho_{\nu\mu} + \delta\Gamma^\rho_{\nu\mu} \tag{8.22}$$

but the Christoffel symbol is a connection (one that is torsion free and metric compatible). We discussed how connections are not tensors, but differences between them are - hence $\delta\Gamma^{\rho}_{\nu\mu}$ is indeed a tensor!

Thus, because this $\delta\Gamma$ is indeed a tensor, it makes sense to write things in terms of covariant derivatives (w.r.t. Γ):

$$\begin{aligned}\delta R^{\rho}_{\mu\lambda\nu} &= \partial_\lambda \delta\Gamma^{\rho}_{\nu\mu} + \Gamma^{\rho}_{\lambda\sigma}\delta\Gamma^{\sigma}_{\nu\mu} + \Gamma^{\sigma}_{\nu\mu}\delta\Gamma^{\rho}_{\lambda\sigma} - (\lambda \leftrightarrow \nu) \\ &= \underbrace{\boldsymbol{\nabla}_\lambda \delta\Gamma^{\rho}_{\nu\mu}}_{\partial_\lambda\Gamma^{\rho}_{\nu\mu}+\Gamma^{\rho}_{\lambda\sigma}\delta\Gamma^{\sigma}_{\nu\mu}-\Gamma^{\sigma}_{\lambda\nu}\delta\Gamma^{\rho}_{\sigma\mu}-\Gamma^{\sigma}_{\lambda\mu}\delta\Gamma^{\rho}_{\nu\sigma}} \qquad - \boldsymbol{\nabla}_\nu \delta\Gamma^{\rho}_{\lambda\mu} + ??? + (\lambda \leftrightarrow \nu)\end{aligned}$$

the terms in red cancel, the terms in blue cancel with the $\lambda \leftrightarrow \nu$ terms due to the torsion-free ness, as does the last term. After the dust settles, we end up with:

$$\delta R_{\mu\nu} = \boldsymbol{\nabla}_\lambda \delta\Gamma^{\lambda}_{\mu\nu} - \boldsymbol{\nabla}_\nu \delta\Gamma^{\lambda}_{\lambda\mu} \tag{8.23}$$

and so:

$$\delta S_1 = \int d^4x \sqrt{-g} \left[\boldsymbol{\nabla}_\lambda (g^{\mu\nu}\delta\Gamma^{\lambda}_{\mu\nu}) - \boldsymbol{\nabla}_\nu (g^{\mu\nu}\delta\Gamma^{\lambda}_{\lambda\mu}) \right] \tag{8.24}$$

where we have taken the metric inside of the covariant derivative using that the connection is metric compatible. We can rewrite this as:

$$\delta S_1 = \int d^4x \sqrt{-g} \boldsymbol{\nabla}_\alpha A^\alpha \tag{8.25}$$

with:

$$A^\alpha = g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\nu} - g^{\mu\alpha}\delta\Gamma^{\lambda}_{\lambda\mu} \tag{8.26}$$

but then:

$$\delta S_1 = \int d^4x \sqrt{-g} \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} A^\mu) = \int d^4x \partial_\mu (\sqrt{-g} A^\mu) \tag{8.27}$$

so if we assume that the variation vanishes at the boundaries of spatial/temporal infinity, then $\delta S_1 = 0$ (integral of a total derivative). Therefore, δS_H can just be written as the contribution of the second/third terms:

$$\delta S_H = \int d^4x (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \delta g^{\mu\nu} \tag{8.28}$$

the term appearing in brackets is precisely the LHS of Einstein's equation! So, the RHS of the Einstein's equation should come from varying the matter part. Also, of course at this point there is some unknown coefficient κ:

$$\delta S_H = \kappa \int d^4x (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \delta g^{\mu\nu} \tag{8.29}$$

as this does not change the EOM, but this will be fixed when we consider the matter term.

8.3 Matter term

Consider the action for the matter sector $S[\phi, g_{\mu\nu}]$, e.g. for a real scalar field in flat space:

$$S = \frac{1}{2} \int d^4x \left(-\eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2 \right) \tag{8.30}$$

The terms in the Lagrangian can be more familiarly written as:

$$\underbrace{\dot{\phi}^2}_{\text{KE}} - \underbrace{(\boldsymbol{\nabla}\phi)^2 - m^2\phi^2}_{\text{PE}} \tag{8.31}$$

How do we write this action in a way that is manifestly coordinate transformation invariant? We can replace the integration measure, and also promote the flat metric to the general one:

$$S_{\text{matter}}[\phi, g^{\mu\nu}] = \frac{1}{2}\int d^4x\sqrt{-g}\left(-g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2\right) \tag{8.32}$$

We again vary $g^{\mu\nu} \to g^{\mu\nu} + \delta g^{\mu\nu}$. We then get δS_{matter} from two parts. The first part comes from inside the brackets, and the second term from the $\sqrt{-g}$:

$$\delta S_{\text{matter}} = -\frac{1}{2}\int d^4x\sqrt{-g}\partial_\mu\phi\partial_\nu\phi\delta g^{\mu\nu} - \frac{1}{2}\int d^4x(-g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - m^2\phi^2)(-\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}) = -\frac{1}{2}\int d^4x\sqrt{-g}T_{\mu\nu}\delta g_{\mu\nu} \tag{8.33}$$

Where we have defined:

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi + g_{\mu\nu}\mathcal{L} \tag{8.34}$$

let's check that this indeed coincides with the stress-energy tensor. Looking at T_{00} in flat spacetime, we see that:

$$T_{00} = \dot{\phi}^2 - \mathcal{L} = \frac{\dot{\phi}^2}{2} + \frac{(\boldsymbol{\nabla}\phi)^2}{2} + \frac{m^2}{2}\phi^2 \tag{8.35}$$

which is the sum of kinetic/potential energy, as we'd expect.

We can also look/compare to what we get out of Noether's theorem:

$$\tilde{T}_{\mu\nu} = \partial_\mu\phi\frac{\partial\mathcal{L}}{\partial(\partial\partial^\nu\phi)} + g_{\mu\nu}\mathcal{L} \tag{8.36}$$

So the Euler-Lagrange equation gives:

$$\partial_\nu\tilde{T}^{\mu\nu} = 0 \tag{8.37}$$

which of course is the expected conservation law of the stress-energy tensor.

So, we see that by taking the physical field and varying the metric/spacetime we get the stress energy tensor out! Note that by the symmetry of $g^{\mu\nu}$ we naturally get a naturally symmetric stress tensor $T_{\mu\nu}$ - note that $\tilde{T}_{\mu\nu}$ as we get out of Noether's theorem does *not* have the same guarantee, with generically:

$$\tilde{T}_{\mu\nu} \neq \tilde{T}_{\nu\mu} \tag{8.38}$$

however, it turns out to be the case that $T, \tilde{T}$ do turn out to coincide (up to a total derivative). We will use the one we got from varying the matter action, as it is symmetric and gauge invariant.

8.4 Deriving the Conservation Law

Note that we got:

$$S[\phi, g^{\mu\nu}] \to S[\phi, g^{\mu\nu} + \delta g^{\mu\nu}] = S[\phi, g^{\mu\nu}] - \frac{1}{2}\int d^4x\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu} \tag{8.39}$$

equivalently:

$$S[\phi, g_{\mu\nu} + \delta g_{\mu\nu}] = S[\phi, g_{\mu\nu}] + \frac{1}{2}\int d^4x\sqrt{-g}T^{\mu\nu}\delta g_{\mu\nu} \tag{8.40}$$

From this, let us show that we can derive the EOM:

$$\boldsymbol{\nabla}_\mu T^{\mu\nu} = 0 \tag{8.41}$$

Recall that we have coordinate invariance of the action, i.e. $x^\mu \to x^{\mu'} = x^{\mu'}(x^\mu)$ then:

$$S[\phi', g'_{\mu\nu}] = S[\phi, g_{\mu\nu}] \tag{8.42}$$

We now consider an infinitesimal transformation/diffeomorephism, $x^{\mu'} = x^\mu + \xi^\mu$. We then have:

$$\phi(x) \to \phi'(x') = \phi'(x+\xi) = \phi(x) \tag{8.43}$$

so:

$$\phi'(x) = \phi(x-\xi) = \phi(x) \underbrace{-\xi^\lambda \partial_\lambda \phi}_{\delta\phi} \tag{8.44}$$

Looking at the transformation of the metric:

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x) \tag{8.45}$$

Now using $x^\mu = x'^\mu - \xi^\mu$, we can write:

$$g'_{\mu\nu}(x') = g'_{\mu\nu}(x+\xi)(\delta^\alpha_\mu - \partial_\mu \xi^\alpha)(\delta^\beta_\nu - \partial_\nu \xi^\beta) g_{\alpha\beta} = g_{\mu\nu} - \partial_\mu \xi^\alpha g_{\alpha\nu} - \partial_\nu \xi^\beta g_{\mu\beta} \tag{8.46}$$

Therefore:

$$g'_{\mu\nu}(x) = g_{\mu\nu} - \partial_\mu \xi^\alpha g_{\alpha\nu} - \partial_\nu \xi^\beta g_{\mu\beta} - \xi^\lambda \partial_\lambda g_{\mu\nu} \tag{8.47}$$

As an exercise, this is equal to:

$$g'_{\mu\nu}(x) = g_{\mu\nu}(x) \underbrace{-\boldsymbol{\nabla}_\mu \xi_\nu - \boldsymbol{\nabla}_\nu \xi_\mu}_{\delta g_{\mu\nu}} \tag{8.48}$$

Therefore:

$$S[\phi, g_{\mu\nu}] \to S[\phi + \delta\phi, g_{\mu\nu} + \delta g_{\mu\nu}] = S[\phi, g_{\mu\nu}] + \int d^4x \left[\frac{\delta S}{\delta\phi(x)}\delta\phi(x) + \frac{\delta S}{\delta g_{\mu\nu}(x)}\delta g_{\mu\nu}(x)\right] \tag{8.49}$$

The first term vanishes since the matter fields satisfy the equation of motion. The second term is proportional to $T^{\mu\nu}$, and in particular:

$$\delta S = \int d^4x \sqrt{-g} T^{\mu\nu} (\boldsymbol{\nabla}_\mu \xi_\nu + \boldsymbol{\nabla}_\nu \xi_\mu) = 0 \tag{8.50}$$

We can use the symmetry of $T^{\mu\nu}$ to combine the two derms, and then use integration by parts to obtain:

$$\int d^4x \sqrt{-g} (\boldsymbol{\nabla}_\mu T^{\mu\nu}) \xi^\nu = 0 \tag{8.51}$$

Thus:

$$\boxed{\boldsymbol{\nabla}_\mu T^{\mu\nu} = 0} \tag{8.52}$$

8.5 Combining the Parts and Outlook

Putting the pieces together, we find:

$$S = \kappa \int d^4x \sqrt{-g} R + S_{\text{matter}}[\phi, g^{\mu\nu}] \tag{8.53}$$

taking the functional derivative and setting it to zero:

$$\frac{\delta S}{\delta g^{\mu\nu}} = 0 \implies \kappa\sqrt{-g}(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) - \frac{1}{2}\sqrt{-g} T_{\mu\nu} = 0 \tag{8.54}$$

from which we obtain:

$$\kappa = \frac{1}{16\pi G} \tag{8.55}$$

to get Einstein's equations (equivalently, go through the game of last class of making sure this gives the correct result in the Newtonian limit).

We now know how to write down the EOM of the gravitational field - we know how matter behaves. In the next part of the course, we try to understand applications of/solutions to Einstein's theory of gravity. There are three things we will want to discuss:

- The Schwarzschild Solution
- The Cosmological Solution
- Gravitational Radiation

In Carroll things are introduced in this order, in Wald the reverse sequence is done. Son is still figuring out which order he likes better... see you next time to find out.

9 Weak Gravitational Fields

9.1 Operational Definition of Weak Fields

Today we discuss one consequence of Einstein's equations that is applicable in the limit of weak gravitational field.

What do we mean when we say weak fields? One possible definition is that the metric is very close to Minkowski/flat:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{9.1}$$

where $h_{\mu\nu} \ll 1$.

One problem with this equation is that this is a frame dependent statement... so we may think about another criterion. Instead of the metric, which coordinate-system dependent, we can look at the curvature tensor (contracted to the Ricci scalar, so that it is manifestly coordinate system independent)

$$R_{\mu\nu\lambda\rho} \to R \tag{9.2}$$

So we want R to be small. But small compared to what? This is a perenniel question of theoretical physics. Well, let's look at R:

$$R \sim \underbrace{\partial\Gamma}_{\partial^2 g} + \Gamma\Gamma \sim \frac{O(n)}{L^2} \tag{9.3}$$

so our criterion becomes:

$$RL^2 \ll 1 \tag{9.4}$$

where R is again, the Ricci scalar, and L is the length scale of the system.

Q: Ricci scalar does not contain all the information about the Riemann tensor. So can we guarantee that R small means weak field? A: No... but we can construct all possible geometric objects $R_{\mu\nu}R^{\mu\nu}$, $C_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}$ and find them to be small, then we can guarantee.

So, let's say:

$$h_{\mu\nu} \sim O(\epsilon) \tag{9.5}$$

This is true in the solar system. We have already derived this, but we will see again that:

$$g_{00} = -(1 + 2\Phi) \tag{9.6}$$

with Φ the gravitational potential, and thus for an object of mass m the gravitational potential energy is $m\Phi$. Thus, the $\Phi \ll 1$ limit is that in which $m\Phi \ll mc^2$.

Note that this is a good assumption in many cases, but is violated in some situations:

- In the vicinity of black holes
- In the vicinity of a neutron stars
- Universe at large scales

9.2 Transformations of Metric Deformations

Let us consider diffeomorphism invariance at the weak field level $h_{\mu\nu} \sim O(\epsilon) \ll 1$. Consider a coordinate transformation:

$$x^\mu \to x^{\mu'} = x^\mu + \underbrace{\xi^\mu}_{O(\epsilon)} \tag{9.7}$$

The metric transforms as:

$$g_{\mu\nu} \to \underbrace{g_{\mu\nu}}_{\eta_{\mu\nu}+g_{\mu\nu}} - \underbrace{\xi^\lambda \partial_\lambda g_{\mu\nu}}_{\xi^\lambda \partial_\lambda \eta_{\mu\nu} + O(\epsilon^2)} - \underbrace{g_{\lambda\nu}\partial_\mu \xi^\lambda}_{\approx \eta_{\lambda\nu}\partial_\mu \xi^\lambda} - \underbrace{g_{\mu\lambda}\partial_\nu \xi^\lambda}_{\approx \eta_{\mu\nu}\partial_\nu \xi^\lambda} + O(\epsilon^2) \tag{9.8}$$

So, in this limit we can treat $\eta_{\mu\nu}$ as the metric and $h_{\mu\nu}$ as a field on flat spacetime. So, let us use $\eta_{\mu\nu}$ to lower indices:

$$g_{\mu\nu} \to \eta_{\mu\nu} + h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{9.9}$$

So under this coordinate transformation, the field $h_{\mu\nu}$ undergoes a gauge transformation:

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{9.10}$$

analogous to the gauge transformations in EM:

$$A_\mu \to A_\mu + \partial_\mu \alpha \tag{9.11}$$

9.3 Einstein's Equation in the Weak Field Limit

Recall that we found:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{9.12}$$

Recall the formula for the Ricci tensor:

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\mu \Gamma^\lambda_{\nu\lambda} + O(\Gamma^2) \tag{9.13}$$

where we neglect the $O(\Gamma^2)$ terms as they are $O(\epsilon^2)$. To $O(\epsilon^2)$, we can write:

$$\begin{aligned}
\Gamma^\lambda_{\mu\nu} &= \frac{1}{2} g^{\lambda\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) \\
&\approx \frac{1}{2} \eta^{\lambda\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) \\
&= \frac{1}{2} (\partial_\mu h^\lambda_\nu + \partial_\nu h^\lambda_\mu - \partial^\lambda h_{\mu\nu})
\end{aligned}$$

The Einstein tensor $G_{\mu\nu}$ then to $O(\epsilon^2)$ becomes the following:

$$G_{\mu\nu} = \frac{1}{2} \left[\partial_\mu \partial^\lambda h_{\nu\lambda} + \partial_\nu \partial^\lambda h_{\mu\lambda} - \partial^2 h_{\mu\nu} - \partial_\mu \partial_\nu h - \eta_{\mu\nu} (\partial^\alpha \partial^\beta h_{\alpha\beta} - \partial^2 h) \right] = 8\pi G T_{\mu\nu} \tag{9.14}$$

Note that $\partial^2 = \partial^\lambda \partial_\lambda = \Box$ and $h = h^\mu_{\ \mu}$. The nice part about this equation is that it is linear in h, so we can solve it - though how to do so is not yet clear. Let us define:

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h \tag{9.15}$$

then:

$$\bar{h} = h - 2h = -h \tag{9.16}$$

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \bar{h} \tag{9.17}$$

You can then check that going to $\bar{h}$ simplifies things:

$$G_{\mu\nu} = \frac{1}{2} \left[\partial_\mu \partial^\lambda \bar{h}_{\nu\lambda} + \partial_\nu \partial^\lambda \bar{h}_{\mu\lambda} - \frac{1}{2} \partial^2 \bar{h}_{\mu\nu} \right] = 8\pi G T_{\mu\nu} \tag{9.18}$$

Note that- just as the Maxwell equations are invariant under gauge transformations, so here are the equations about the gauge transformations:

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{9.19}$$

At this point we should *not* want our equation to have a unique solution (because we can always generate another via a gauge transformation). Indeed, we have 4 gauge redundancies, coming from ξ_μ for $\mu = 0, 1, 2, 3$. Let us gauge fix to get a specific solution. We get the Hilbert gauge by fixing:

$$\partial_\mu \bar{h}^{\mu\nu} = 0 \tag{9.20}$$

It is analogous to the Lorentz gauge of EM, where:

$$\partial_\mu A^\mu = 0. \tag{9.21}$$

In this gauge the equation reduces to:

$$\partial^2 \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}. \tag{9.22}$$

Note that prior to imposing this gauge choice, we can do the gauge transformation Eq. (9.19) for any arbitrary choice of function $\xi_\mu(x)$. However, there is still a residual gauge symmetry. Consider:

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{9.23}$$

$$h \to h - 2\partial_\mu \xi^\mu \tag{9.24}$$

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu + (\partial_\lambda \xi^\lambda)\eta_{\mu\nu} \tag{9.25}$$

$$\partial^\mu \bar{h}_{\mu\nu} \to \partial^\mu \bar{h}_{\mu\nu} - \partial^2 \xi_\nu - \partial_\nu(\partial^\mu \xi_\mu) + \partial_\nu(\partial^\mu \xi_\mu) \tag{9.26}$$

i.e. the residual gauge freedom is for functions satisfying:

$$\partial^2 \xi_\nu = 0 \tag{9.27}$$

the wave equation.

9.4 The Newtonian Limit

In the Newtonian limit, the stress-energy tensor takes the form:

$$T^{\mu\nu} = \text{diag}(\rho c^2, 0, 0, 0) \tag{9.28}$$

So for the 00 component we have:

$$\partial^2 \bar{h}_{00} = (-\partial_t^2 + \nabla^2)\bar{h}_{00} = -16\pi G\rho \tag{9.29}$$

and since T vanishes for all other components:

$$\bar{h}_{0i} = \bar{h}_{ij} = 0 \tag{9.30}$$

Consider the static case of $\bar{h}_{00} = -4\Phi(x)$ (no dependence on time). Then we just get the Poisson equation/Newtonian result:

$$\nabla^2 \Phi = 4\pi G\rho \implies \Phi(x) = -G \int d\mathbf{y} \frac{\rho(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} \tag{9.31}$$

So this is the correct solution! We then find:

$$\bar{h} = \eta^{\mu\nu} \bar{h}_{\mu\nu} = +4\Phi \tag{9.32}$$

and solving for h:

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\bar{h} \implies h_{\mu\nu} = \text{diag}(-2\Phi, -2\Phi, -2\Phi, -2\Phi) \tag{9.33}$$

Thus we obtain the metric:

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)d\mathbf{x}^2 \tag{9.34}$$

the 00-piece we knew before, but here we also get the spatial component!

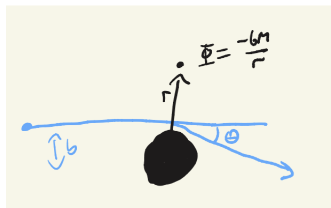
9.5 Deflection of Light

This was an early prediction/test of GR. Consider being at a radius r from a star of mass M (with $r > R$ the radius of the star).

The metric takes the form:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 + \frac{2GM}{r}\right) (dx^2 + dy^2 + dz^2) \tag{9.35}$$

We can then consider a scattering experiment, where we throw in a particle with impact parameter b and then measure the scattering angle θ.



Though we have not solved for the geodesics of massless particles, we will solve for the geodesic of a massive particle and then take the $m \to 0$ limit. Recall the geodesic equation:

$$\frac{d^2 x^\mu}{d\tau^2} = \Gamma^\mu_{\nu\lambda} \frac{dx^\nu}{d\tau} \frac{dx^\lambda}{d\tau} \tag{9.36}$$

where the velocity is $\frac{dx^\mu}{d\tau} = u^\mu$ where we choose the parameter τ to be the proper time such that $u^2 = -1$. We can get the 4-momentum from $p^\mu = mu^\mu$. Let us multiply both sides of the geodesic equation by m, so that:

$$\frac{dp^\mu}{d\tau} + \Gamma^\mu_{\nu\lambda} p^\nu u^\lambda = 0 \tag{9.37}$$

We may then write:

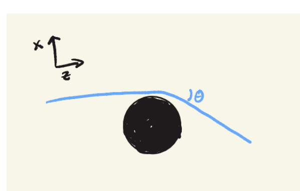
$$dp^\mu = -\Gamma^\mu_{\nu\lambda} p^\nu u^\lambda d\tau = -\Gamma^\mu_{\nu\lambda} p^\nu dx^\lambda \tag{9.38}$$

So, this equation tells me how the momentum of a particle changes along a geodesic. No m appears here, so we can take the $m \to 0$ limit and apply it to photons.

Let us assume that the deflection angle θ is very small:

$$\theta \ll 1 \tag{9.39}$$

this will allow us to solve for the momentum change in a very simple way. We choose coordinates:



Now, at $t = -\infty$, set $p_x = 0$. The change in the x-momentum is then:

$$\Delta p^x = \Delta p^1 = -\int_{-\infty}^{\infty} \Gamma^1_{\nu\lambda} p^\nu dx^\lambda \tag{9.40}$$

with dx^λ along the trajectory of the particle. We may approximate dx^λ as being along a straight line (the unperturbed trajectory if the star was not there), as the error we would make would be $O(\theta^2)$. Thus the above becomes:

$$\Delta p^1 = -\int (\Gamma^1_{\nu 0} p^\nu dt + \Gamma^1_{\nu 3} \partial^\nu dz) = -\int dt (\Gamma^1_{00} p^0 + \Gamma^1_{30} p^3) + \int dz (\Gamma^1_{03} p^0 + \Gamma^1_{33} p^3) \tag{9.41}$$

In the non-relativistic case ($v \ll 1$), we have $p^0 = m$, and we only need consider one term:

$$\Delta p_x = -m \int dt \underbrace{\Gamma^1_{00}}_{-\frac{1}{2}\partial_1 h_{00}} = -m \int dt \partial_1 \Phi = + \int dt F_x \tag{9.42}$$

This is not very surprising - just the integral of the force acting on the (massive) particle for all time!

But, now let's consider the case of light. Now all terms are equally important. Let's take $p^0 = p^3 = p$. Then:

$$\Delta p_x = -p \int dt \left(\Gamma^1_{00} + \Gamma^1_{03} + \Gamma^1_{30} + \Gamma^1_{33} \right) \tag{9.43}$$

If we recall the weak-field limit Christoffel symbol:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2}(\partial_\mu h^\lambda_\nu - \partial_\nu h^\lambda_\mu - \partial^\lambda h_{\mu\nu}) \tag{9.44}$$

Thus we can see that only the diagonal $(00, 33)$ elements enter:

$$\Gamma^1_{00} = -\frac{1}{2}\partial_1 h_{00} = +\partial_1 \Phi \tag{9.45}$$

$$\Gamma^1_{33} = -\frac{1}{2}\partial_1 h_{33} = \partial_1 \Phi \tag{9.46}$$

Therefore:

$$\Delta p_x = -2p \int dt \partial_1 \Phi \tag{9.47}$$

If we had naively assumed that the photon is just a massive particle traveling at the speed of light, we would have missed a factor of 2! Which general relativity gets correct.

Taking:

$$\Phi = -\frac{GM}{\sqrt{x^2 + z^2}} \tag{9.48}$$

we find:

$$\partial_x \Phi = \frac{GMx}{(x^2 + z^2)^{3/2}} \tag{9.49}$$

Thus:

$$\theta = \frac{\Delta p_z}{z} = -GM \int_{-\infty}^{\infty} dz \frac{R}{(R^2 + z^2)^{3/2}} = c\frac{GM}{R} \tag{9.50}$$

with c determined from the integration. We can check that this is small:

$$\frac{GM}{R} = \Phi \ll 1. \tag{9.51}$$

Note that we cheated along the way multiple times (taking the $m \to 0$ limit, approximating the trajectory as a straight line), but this is the correct result.

We will return to this question a little bit later, when we consider a source of strong gravitational fields.

9.6 Intro to Gravitational Waves

With the remaining time, let us begin the discussion of gravitational waves. We consider the sourceless case (where the stress-energy tensor vanishes):

$$\partial^2 \bar{h}_{\mu\nu} = 0 \tag{9.52}$$

Which admits the wave solutions:

$$\bar{h}_{\mu\nu} = \epsilon_{\mu\nu} e^{ik_\sigma x^\sigma} \tag{9.53}$$

with $k_\mu k^\mu = 0$. Note that the Hilbert gauge condition $\partial_\mu \bar{h}^{\mu\nu} = 0$ implies that:

$$k_\mu \epsilon^{\mu\nu} = 0 \tag{9.54}$$

$\epsilon_{\mu\nu}$ being a symmetric 4×4 tensor has 10 coefficients, but the gauge condition reduces this number to 6.

We also discussed earlier in the lecture how there is a residual gauge condition, wherein ξ^μ can satisfy $\partial^2 \xi^\mu = 0$. Thus, such residual gauge parameters are also wavelike:

$$\xi^\mu = \zeta^\mu e^{ik_\sigma x^\sigma} \tag{9.55}$$

Thus under the gauge transformation:

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{9.56}$$

We have:

$$\epsilon^{\mu\nu} \to \epsilon^{\mu\nu} - i(\zeta^\mu k^\nu + \zeta^\nu k^\mu) \tag{9.57}$$

And these are physically equivalent. This reduces the number of parameters to specify from 6 to 2, which are the two physical polarizations of gravitational waves.

10 Weak Gravitational Fields II - Waves and Radiation

Today we continue our discussion about linearized gravity and gravitational waves/radiation. Time permitting, we will also discuss the Lense-Thirring effect (did not get to in class, so will be one of the homeworks, probably).

10.1 Review

In the weak gravitational field limit, we can write:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{10.1}$$

with $h_{\mu\nu} \ll 1$.

The gauge invariance condition is that:

$$h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \tag{10.2}$$

It is more convenient to work with:

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h \tag{10.3}$$

with $h = h^\mu_{\ \mu}$. In the Hilbert gauge:

$$\partial_\mu \bar{h}^{\mu\nu} = 0 \tag{10.4}$$

we get the wave equation:

$$\partial^2 \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} \tag{10.5}$$

Last time we looked at the non-relativistic limit and briefly touched on the gravitational wave equation.

10.2 Gravitational Waves

Consider free space $T_{\mu\nu}$. Then, we get wavelike solutions of the form:

$$\bar{h}_{\mu\nu}(t, \mathbf{x}) = \mathrm{Re} e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} \epsilon_{\mu\nu} \tag{10.6}$$

with the dispersion:

$$\Box \bar{h}_{\mu\nu} = 0 \implies \omega^2 - \mathbf{k}^2 = 0 \tag{10.7}$$

The gauge condition imposes the condition:

$$k^\mu \epsilon_{\mu\nu} = 0 \tag{10.8}$$

Consider a wavevector polarized in the z direction:

$$k^\mu = (\omega, \mathbf{k}) = (k, 0, 0, k) \tag{10.9}$$

Let us go to lightcone coordinates:

$$k^\pm = \frac{k^0 \pm k^3}{\sqrt{2}} \tag{10.10}$$

so:

$$k^+ = \sqrt{2}k \implies k_- \neq 0 \tag{10.11}$$

$$k^- = k^2 = 0 \implies k_+ = k_a = 0 \tag{10.12}$$

and analogously:

$$x^\pm = \frac{x^0 \pm x^3}{\sqrt{2}} \tag{10.13}$$

wherein the metric becomes:

$$ds^2 = -dt^2 + dz^2 + dx^2 + dy^2 = -2dx^+ dx^- + dx^2 + dy^2 \tag{10.14}$$

so we have the metric:

$$g_{+-} = g_{-+} = -1 \tag{10.15}$$

$$g_{11} = g_{22} = 1 \tag{10.16}$$

The Hilbert gauge condition gives us:

$$\partial^\mu \bar{h}_{\mu\nu} = 0 \implies k^\mu \epsilon_{\mu\nu} = 0 \implies \epsilon_{+\nu} = 0 \tag{10.17}$$

thus (for $a, b \in \{1, 2\}$) we only need consider the elements $\epsilon_{--}, \epsilon_{-a}, \epsilon_{ab}$.

Recall that we have the residual gauge transformations (Hilbert gauge does not fix everything):

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \frac{1}{2}\eta_{\mu\nu}(\partial\xi) \tag{10.18}$$

The Hilbert gauge condition $\partial^\mu \bar{h}_{\mu\nu} = 0$ gives us that:

$$\partial^\mu \partial_\mu \xi_\nu = 0 \tag{10.19}$$

Thus we get wavelike solutions for the residual gauge:

$$\xi_\nu(t, \mathbf{x}) = e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} \xi_\nu(\omega, \mathbf{k}) \tag{10.20}$$

And such residual gauge transformations on the polarization ϵ looks like:

$$\epsilon_{\mu\nu} \to \epsilon_{\mu\nu} + i(k_\mu \xi_\nu + k_\nu \xi_\mu - \eta_{\mu\nu} k^\lambda \xi_\lambda) \tag{10.21}$$

The nonzero components we identified transform like:

$$\epsilon_{--} \to \epsilon_{--} + 2ik_- \xi_- \tag{10.22}$$

$$\epsilon_{-a} \to \epsilon_{-a} + ik_- \xi_a \tag{10.23}$$

$$\epsilon_{ab} \to \epsilon_{ab} - i\eta_{ab} \tag{10.24}$$

Thus, there are parameters ξ we can play with - using the residual gauge condition, we can choose a gauge in which $\epsilon_{--}, \epsilon_{-a}$ vanish, and ϵ_{ab} is traceless.

10.3 Polarizations of Gravitational Waves

Thus, we have:

$$\bar{h}_{\mu\nu} \sim \epsilon_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \epsilon_{xx} & \epsilon_{xy} & 0 \\ 0 & \epsilon_{yx} & \epsilon_{yy} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim h_{\mu\nu} \tag{10.25}$$

with $\epsilon_{xx} + \epsilon_{yy} = 0$. This is analogous with the transverse polarization of EM, where $A_0 = A_z = 0$, and $E_x = \dot{A}_x, E_y = \dot{A}_y$.

Suppose that $\epsilon_{xx} = -\epsilon_{yy} = \epsilon$ and $\epsilon_{xy} = 0$. Then, with the perturbation, the metric looks like:

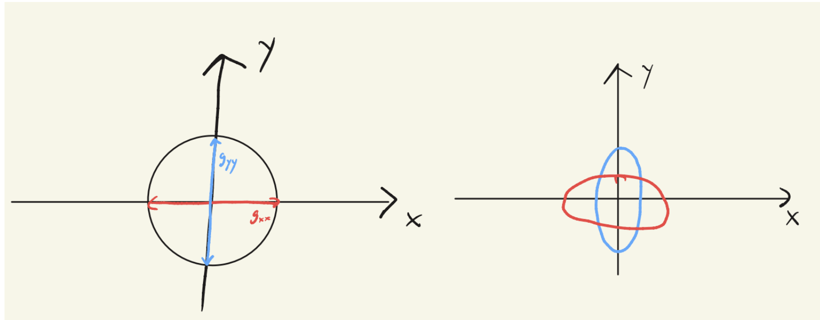
$$g_{xx} = 1 + \epsilon \cos(\omega t - kz) \tag{10.26}$$

$$g_{yy} = 1 - \epsilon \cos(\omega t - kz) \tag{10.27}$$

and:

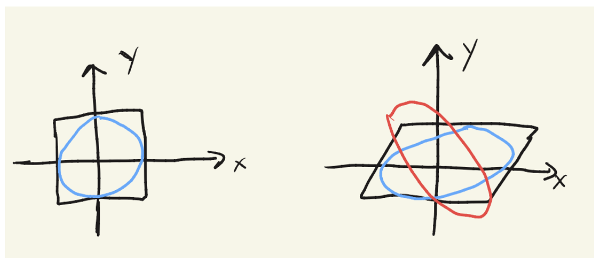
$$-g_{00} = g_{zz} = 1 \tag{10.28}$$

This oscillations correspond to a circle oscillating between ellipses, as depicted below:



These are gravitational waves polarized in the xx/yy direction. To make this precise we would compute the geodesic of a test particle placed on the circle.

The other type is gravitational waves polarized in the xy direction, with $\epsilon_{xx} = \epsilon_{yy} = 0$ and $\epsilon_{xy} \neq 0$, wherein:



Like we can have the circular polarizations $A_\pm = A_x \pm iA_y$ in EM, we can also get circular polarizations of gravitational waves, as well:

$$\epsilon_{xx} - \epsilon_{yy} + 2i\epsilon_{xy} = 2(\epsilon_{xx} + i\epsilon_{xy}) \tag{10.29}$$

$$\epsilon_{xx} - \epsilon_{yy} - 2i\epsilon_{xy} = 2(\epsilon_{xx} - i\epsilon_{xy}) \tag{10.30}$$

What happens in the case when we have matter? then we solve:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{10.31}$$

where:

$$ds^2 = -(1+2\Phi)dt^2 + 2w_i dt dx^i + (\eta_{ij} + s_{ij})dx^i dx^j \tag{10.32}$$

We linearize equations for Φ, w_i, sij, and choose a gauge where $\partial_i w^i = 0, \partial_i s^{ij} = 0$. We then get:

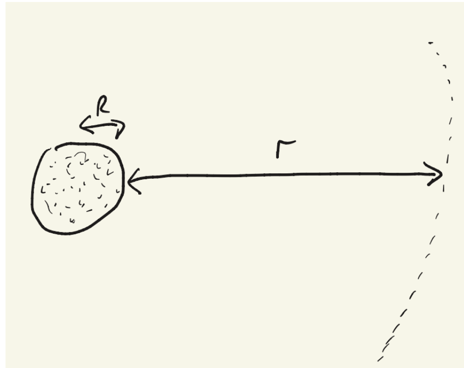
$$\nabla^2 \Phi = T_{00} \tag{10.33}$$

we then find that Φ is not dynamical, i.e. the time dependence of matter fully determines the time dependence of Φ. The same is true of w_i.

10.4 Loss of Energy due to Radiation

In EM, we have a formula that relates the loss of energy due to radiation due to the changes in the dipole moment of a compact collection of charge. We now derive such a formula for gravitational radiation.

We consider some confined collection of mass, and look at the gravitational radiation some large distance $r \gg R$ away:



We have the equation of motion:

$$\Box\bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu} \tag{10.34}$$

which we can then solve via Green's functions:

$$\bar{h}_{\mu\nu}(t,\mathbf{x}) = 4G\int d\mathbf{y}\frac{T_{\mu\nu}(t-\langle\mathbf{x}-\mathbf{y}\rangle,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \tag{10.35}$$

Note that in a flat background we find that the gravitational waves travel at the speed of light, and - like we found in electromagnetic radiation - our answer depends on the retarded time (accounting for the time for the gravitational wave signal to reach the observer).

If we ignore the finite size of the object (i.e we assume that the entire mass is located around $\mathbf{y} = \mathbf{0}$ - this is where the $r \gg R$ assumption kicks in) the above simplifies to:

$$\bar{h}_{\mu\nu}(t,\mathbf{x}) = \frac{4G}{|x|}\int d\mathbf{y}T_{\mu\nu}(t-|\mathbf{x}|,\mathbf{y}) \tag{10.36}$$

Let us consider h^{TT}_{ab}, i.e. the transverse and traceless contribution, wherein:

$$h^{TT}_{ab} = \frac{4G}{r}\int dy\mathbf{y}T^{tt}_{ab}(t-r,\mathbf{y}) \tag{10.37}$$

this is the component which contains information about the energy flux.

Let us see how $\int d\mathbf{y}T_{ij}(\mathbf{y})$ is related to the quadrapole moment. We have the equations:

$$\partial_0 T^{00} + \partial_i T^{0i} = 0 \tag{10.38}$$

$$\partial_0 T^{0i} + \partial_j T^{ij} = 0 \tag{10.39}$$

so:

$$\partial_0 T^{00}(y) = \partial_i\partial_j T^{ij}(y) \tag{10.40}$$

Thus if we integrate against $y^k y^l$, we find:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int d\mathbf{y}y^k y^l T^{00} = \int d\mathbf{y}y^k y^l \partial_i\partial_j T^{ij} \tag{10.41}$$

integrating the LHS by parts (using the boundary condition that the matter distribution vanishes outside of the source, allowing us to discard the boundary terms):

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int d\mathbf{y}y^k y^l T^{00} = -\int d\mathbf{y}(\delta^k_i y^l + \delta^l_i y^k)\partial_j T^{ij} \tag{10.42}$$

and again:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int d\mathbf{y} y^k y^l T^{00} = \int d\mathbf{y}(\delta_i^k\delta_j^l + \delta_i^l\delta_j^k)T^{ij} = 2\int d\mathbf{y} T^{kl} \tag{10.43}$$

Thus we find:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} I^{kl} = 2\int d\mathbf{y} T^{kl} \tag{10.44}$$

with I^{kl} the quadrapole moment. Thus, returning back to our expression for h_{ab}^{TT}:

$$h_{ab}^{TT} = \frac{2G}{r}\ddot{J}_{ab}(t-r,\mathbf{y}) \tag{10.45}$$

with:

$$J_{ab} = I_{ab} - \frac{1}{2}\delta_{ab} I_{cc} \tag{10.46}$$

Aside: his can be made more precise in the quantum theory, where the gravitational field is mediated via the spin-2 graviton.

What does this expression tell us? If a star oscillates spherically symmetrically, then it cannot emit gravitational radiation. Additionally, for a source with a nonzero dipole moment but a zero quadrapole moment, we similarly have zero radiation. The leading term is quadrapolar (though like in EM, you are free to look at higher order terms that also contribute)

10.5 Going beyond the linear regime

Some hackiness/heuristic arguments about to be said here, be warned!

Consider the Einstein equation:

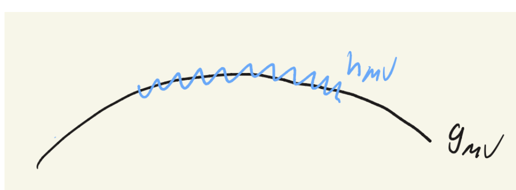
$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 8\pi T_{\mu\nu} \tag{10.47}$$

where $T_{\mu\nu}$ is matter. But therein our discussion is confusing. We said the gravitational wave carries energy-momentum.... but such waves are just fluctuations in the metric g, which should only enter the LHS, not the RHS.

Ok, let's try to make things a bit clearer what we were doing. We decomposed the metric into:

$$g_{\mu\nu} = \tilde{g}_{\mu\nu} + h_{\mu\nu} \tag{10.48}$$

where the $\tilde{g}_{\mu\nu}$ is the long-range background, and $h_{\mu\nu}$ is the wavelike, short range perturbation.



We worked in the linear level, but if we go to the quadratic level, we find:

$$\Box h_{\mu\nu} + O(h^2) = T_{\mu\nu} \tag{10.49}$$

Unlike $h_{\mu\nu}$, h^2 has a nonzero, smooth average over long distances (think cos vs. $\cos^2$). We can treat this short-wavelength part like matter, bringing it to the RHS. In detail, this is discussed in Landau-Lifschiz, as well as Misner/Thorne/Wheeler. Here, we go argue less rigorously, just to give some intuition. Limit h_{ab} to a transverse and traceless wave. Then, we get the "effective action for h_{ab}":

$$S = \int d^4x \left[-\frac{1}{64\pi G}\left(\dot{h}_{ab}^2 - (\partial_z h_{ab})^2\right) + \frac{1}{2}h_{ab}T^{ab}\right] \tag{10.50}$$

where the coefficient is obtained from the wave equation.

The energy in this scenario is:

$$\mathcal{E} = \frac{1}{64\pi G}\left\langle \dot{h}_{ab} + (\partial_z h_{ab})^2 \right\rangle = \frac{1}{32\pi G}\left\langle \dot{h}_{ab}^2 \right\rangle \tag{10.51}$$

Which using our previous expression for h:

$$\mathcal{E} = \frac{G}{8\pi}\frac{1}{r^2}\left\langle \dddot{J}^2_{ab} \right\rangle \tag{10.52}$$

Though engineers apparently feud over whether the quadrapole moment contains the trace or not, we will subtract it out.

$$J_{ij} = \int dy (y^i y^j - \frac{1}{3}\delta^{ij} y^2)\rho(y) \tag{10.53}$$

so J_{ij} has 5 independent components (traceless tensor of second rank). But, with our assumptions, the gravitational wave only depends on the fluctuations of two independent components, $xx(= -yy)$ and xy. If one averages over all directions, all components, so we sum over all angles (and use that only 2 components contribute), and we get:

$$\frac{d\mathcal{E}}{dt} = \frac{G}{8\pi}\frac{4\pi r^2}{r^2}\frac{2}{5}\left\langle \dddot{J}^2_{ij} \right\rangle = \frac{G}{5}\left\langle \dddot{J}^2_{ij} \right\rangle \tag{10.54}$$

Or if we restore factors of light:

$$\frac{d\mathcal{E}}{dt} = \frac{G}{5c^5}\left\langle \dddot{J}^2_{ij} \right\rangle \tag{10.55}$$

so we get two sources of smallness, one from G and one from $\frac{1}{c^5}$.

11 Schwarzchild Solution

Today (and next lecture) we discuss the Schwarzchild solution of the Einstein equation. It is a solution corresponding to a spherically symmetric metric outside of a stellar body, and we will go beyond the linearized approximation we have been exploring the last few lectures. We will discuss:

- Geodesics in the Schwarzchild geometry
- Killing vector fields

11.1 Spherically symmetric metric

We have some spherically symmetric massive object, and we want to look for solutions outside of this massive object, assuming outside of the massive object we have the vacuum. Thus, we solve the vacuum Einstein equation:

$$R_{\mu\nu} = 0. \tag{11.1}$$

Because the setup is spherically symmetric, it will be convenient to work in spherical coordinates. In spherical coordinates, the flat space metric reads:

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{11.2}$$

with θ the polar angle and ϕ the azimuthal angle.

If we want to preserve spherical symmetry but allow for space to not be flat, the most general spherically symmetric and static metric looks like:

$$ds^2 = -f(r)dt^2 + h(r)dr^2 + \tilde{h}(r)(d\theta^2 + \sin^2\theta d\phi^2) \tag{11.3}$$

We have a freedom of reparametrization:

$$r, \theta, \phi \to r'(r), \theta, \phi \tag{11.4}$$

which we can invoke to fix $\tilde{h}(r) = r^2$. Moreover, we can let $f(r) = e^{2\alpha(r)}$ and $h(r) = e^{2\beta(r)}$. Thus the metric looks like:

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\beta(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{11.5}$$

11.2 The Schwarzschild Solution

Given this metric, nonzero components of the Ricci tensor are:

$$R_{tt} = e^{2(\alpha-\beta)}(\alpha'' + \alpha'^2 - \alpha'\beta' + \frac{2}{r}\alpha') = 0 \tag{11.6}$$

$$R_{rr} = -\alpha'' - \alpha'^2 + \alpha'\beta' + \frac{2}{r}\beta' = 0 \tag{11.7}$$

$$R_{\theta\theta} = e^{-2\beta}(r(\beta' - \alpha') - 1) + 1 = 0 \tag{11.8}$$

$$R_{\phi\phi} = \sin^2\theta R_{\theta\theta} = 0 \tag{11.9}$$

Looking at the first two equations, we can ignore the positive $e^{2(\alpha-\beta)}$ factor and then sum up the two equations, which yields:

$$\frac{2}{r}(\alpha' + \beta') = 0 \tag{11.10}$$

which implies:

$$\alpha + \beta = \text{const.} \tag{11.11}$$

Using the rescaling freedom of time $t \to \lambda t$, we can rescale $\alpha \to \alpha + \text{const.}$ which we can use to enforce:

$$\alpha + \beta = 0 \tag{11.12}$$

Thus there is only one function left to determine, by solving the $\theta\theta$ equation:

$$e^{-2\beta}(r(\beta' - \alpha' - 1) - 1) + 1 = 0 \tag{11.13}$$

After solving (exercise), we find:

$$ds^2 = -\left(1 - \frac{r_s}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{r_s}{r}} + r^2 \underbrace{(d\theta^2 + \sin^2\theta d\phi^2)}_{d\Omega^2} \tag{11.14}$$

Note that this solution only makes sense when $r > r_s$ - we will later discuss what happens when $r \to r_s$.

Note that in the limit of $r \to \infty$ we recover the Newtonian theory; there we have:

$$-(1 - \frac{r_s}{r}) = -(1 + 2\Phi) \tag{11.15}$$

with:

$$\Phi = -\frac{r_s}{2r} = -\frac{GM}{r} \tag{11.16}$$

thus $r_s = 2GM$.

This we have found a nontrivial solution to the Einstein equations! This was found by K. Schwarzschild in 1916, very shortly after Einstein discovered general relativity. He then died a few months later on the Eastern front of WW1.

Note that at $r = r_s$ we have the horizon and $r = 0$ we have a singularity. Notably, $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho}$ is finite at the horizon (so it is possible to continue spacetime beyond the horizon) but $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho} \to \infty$ at the singularity.

We now want consider trajectory of bodies in this geometry - can be orbiting planets, or massive/massless particles that come in from spatial infinity and then is influenced at the star/black hole at the origin.

Recall the geodesic equation:

$$\frac{d^2x^\sigma}{d\lambda^2} + \Gamma^\sigma_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0 \tag{11.17}$$

where the particle (whether massive or massless)'s wordline is parameterized by λ.

11.3 Symmetries and Killing Vectors

Before diving into this, let us explore some symmetries of the Schwarzschild solution. We have the time and angle translation symmetries:

$$t \to t + a \tag{11.18}$$

$$\phi \to \phi + a \tag{11.19}$$

It will be convenient to package these symmetries in terms of Killing vectors.

In the time case, we think about a vector $\frac{\partial}{\partial t}$ which corresponds to a vector $\xi_1^\mu = (1,0,0,0)$. In the angle case we have the vector $\frac{\partial}{\partial \phi}$ which corresponds to the vector $\xi_2^\mu = (0,0,0,1)$.

In the language of Killing vectors, coordinate transformations look like:

$$x^\mu \to x^{\mu'} = x^\mu + \underbrace{\xi^\mu(x)}_{\text{Killing}} \tag{11.20}$$

the accompanying metric transformation looks like:

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x) = g_{\mu\nu}(x) - \nabla_\mu \xi_\nu - \nabla_\nu \xi_\mu \tag{11.21}$$

and thus the condition for the coordinate transformation to be a symmetry is:

$$\boxed{\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0} \tag{11.22}$$

For the current solution, we have the dual vectors:

$$\xi_\mu^1 = (-\left(1 - \frac{2GM}{r}\right), 0, 0, 0) \tag{11.23}$$

$$\xi_\mu^2 = (0,0,0,r^2 \sin^2\theta) \tag{11.24}$$

11.4 Symmetries and Conservation Laws in Newtonian Mechanics

Let us connect the symmetries to conservation laws. Think back to Newtonian mechanics, where we have a Lagrangian:

$$\mathcal{L} = \frac{1}{2}\dot{\mathbf{r}}^2 - V(r) = \frac{1}{2}(\dot{\mathbf{r}}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) - V(r) \tag{11.25}$$

Let us fix $\theta = \pi/2$ so the $\dot{\theta}$ and $\sin\theta$ terms vanish, so:

$$\mathcal{L} = \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) - V(r) \tag{11.26}$$

The Euler Lagrange equation then yields:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{d}{dt}(2r^2\dot{\phi}) = 0 \tag{11.27}$$

which tells us there is the conserved quantity of angular momentum $L = r^2\dot{\phi}$.

Note we can also write down the energy:

$$E = \frac{\dot{r}^2}{r} + \frac{1}{2}\dot{r}^2\dot{\phi}^2 + V(r) = \frac{\dot{r}^2}{2} + \underbrace{V(r) + \frac{L^2}{2r^2}}_{V_{\text{eff}}(r)} \tag{11.28}$$

which is also conserved, and can be used to simplify the equations of motion.

11.5 Conserved Quantities

For a Killing vector $\xi^\mu(x)$, we have that:

$$\xi_\mu(x)\frac{dx^\mu}{d\lambda} = \text{const.} \tag{11.29}$$

Let us prove this conservation law:

$$\frac{d}{d\lambda}(\xi_\mu(x)\frac{dx^\mu}{d\lambda}) = \frac{dx^\nu}{d\lambda}\partial_\nu\xi_\mu\frac{dx^\mu}{d\lambda} + \xi_\mu\frac{d^2x^\mu}{d\lambda^2} \tag{11.30}$$

Which recalling the defnition of the covariant derivative $\nabla_\alpha\xi_\mu = \partial_\alpha\xi_\mu - \Gamma^\mu{}_{\alpha\beta}\xi_\mu$, we then get:

$$\frac{d}{d\lambda}(\xi_\mu(x)\frac{dx^\mu}{d\lambda}) == \partial_\alpha\xi_\beta\frac{dx^\alpha}{dx\lambda}\frac{dx^\beta}{d\lambda} - \xi_\mu\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \nabla_\alpha\xi_\beta\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \frac{1}{2}(\nabla_\alpha\xi_\beta + \nabla_\beta\xi_\alpha)\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \tag{11.31}$$

where in the last equality we symmetrize. The condition on Killing vector tells us that this vanishes and so:

$$\frac{d}{d\lambda}(\xi_\mu(x)\frac{dx^\mu}{d\lambda}) = 0 \tag{11.32}$$

which proves that $\xi_\mu(x)\frac{dx^\mu}{d\lambda}$ is indeed a conserved quantity.

11.6 Conservation Laws and Equations of Motion for Schwarzschild

So, let's see what conservation laws we can derive for the Schwarzschild solution. For the killing vector $\xi^\mu_{(\phi)} = (0,0,0,1)$ and $\xi^{(\phi)}_\mu = (0,0,0,r^2)$ (taking $\theta = \pi/2$) we have that:

$$\boxed{r^2\frac{d\phi}{d\lambda} = L = \text{const.}} \tag{11.33}$$

i.e. the familiar conservation law for angular momentum!

For the Killing vector:

$$\xi^{(t)}_\mu = (-\left(1 - \frac{2GM}{r}\right), 0, 0, 0) \tag{11.34}$$

we get the conservation law:

$$\boxed{(1 - \frac{2GM}{r})(\frac{dt}{d\lambda}) = E = \text{const.}} \tag{11.35}$$

which is a kind of energy conservation (we will see soon...)

In addition to these two constants of motion for Schwarzschild, we have another constant of motion which is universal for geodesics, wherein the geodesic equation + metric compatibility implies that:

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = -\epsilon \tag{11.36}$$

with ϵ constant. $\epsilon \neq 0$ for massive particles, and $\epsilon = 0$ for massless particles (null trajectories). Expanding out the RHS:

$$\boxed{-\left(1 - \frac{2GM}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 + \frac{1}{1 - \frac{2GM}{r}}\left(\frac{dr}{d\lambda}\right)^2 + r^2\left(\frac{d\phi}{d\lambda}\right)^2 = -\epsilon} \tag{11.37}$$

Multiplying this equation by $1 - \frac{2GM}{r}$, we get:

$$-E^2 + \left(\frac{dr}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{r}\right)\frac{L^2}{r^2} = -\epsilon\left(1 - \frac{2GM}{r}\right) \tag{11.38}$$

Which can be rewritten as:

$$\frac{1}{2}\left(\frac{dr}{d\lambda}\right)^2 + V_{\text{eff}}(r) = \mathcal{E} = \frac{E^2}{2} \tag{11.39}$$

with the effective potential:

$$V_{\text{eff}}(r) = \frac{\epsilon}{2}\left(1 - \frac{2GM}{r}\right) + \frac{1}{2}\left(1 - \frac{2GM}{r}\right)\frac{L^2}{r^2} \tag{11.40}$$

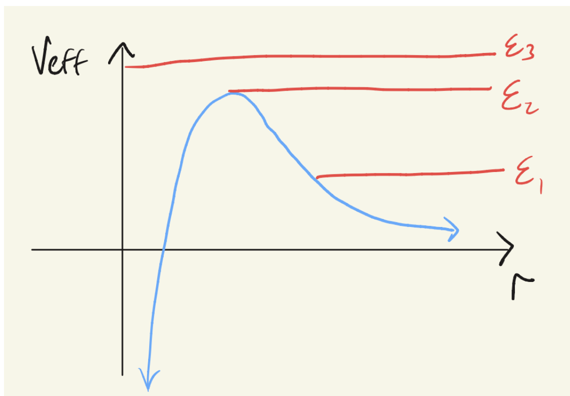
so, at least via mathematical coincidence, our equations of motion look like a statement about the total energy of the system!

11.6.1 Solutions: Massless case

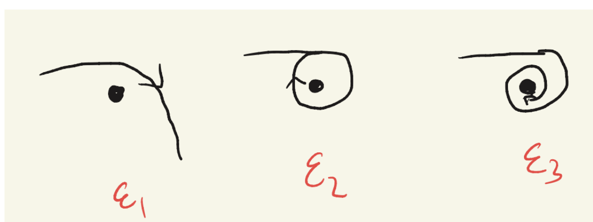
Let's look at solutions in the photon/massless case where $\epsilon = 0$, so then:

$$V_{\text{eff}}(r) = \frac{L^2}{2r^2} - \frac{GML^3}{r^3} \tag{11.41}$$

The potential looks like:



And the three cases of different photon energies correspond to a deflection ($\mathcal{E}_1$), a critical circular orbit ($\mathcal{E}_2$), and spiralling inwards ($\mathcal{E}_3$).



Let us compute the radius of the critical orbit:

$$V'_{\text{eff}}(r) = \frac{L^2}{r^3} - \frac{3GML^2}{r^4} = 0 \tag{11.42}$$

from which we find:

$$r = 3GM = 3r_s > r_s \tag{11.43}$$

so there is an (unstable) circular orbit outside of the black hole/central mass.

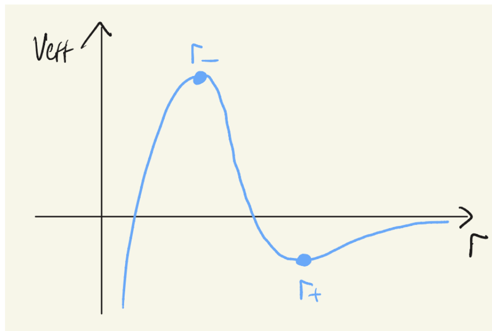
We can also calculate the impact parameter of an incoming photon - this is left as a homework exercise.

11.6.2 Solutions: Massive case

Now consider a massive particle, and take λ to be the proper time and $\epsilon = 1$. The effective potential looks like:

$$V_{\text{eff}}(r) = -\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} \tag{11.44}$$

The potential looks like (for sufficiently high L):

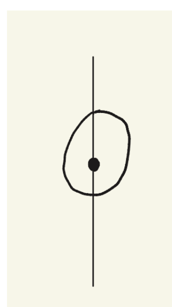


Wherein there are two critical orbits:

$$r_{\pm} = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM} \tag{11.45}$$

Assuming that $L > L_{\text{crit}} = \sqrt{12}GM$.

To finish the lecture, let us calculate the precession in the perehelion of Mercury. This was one of the first calculations done by Einstein. It was astronomically observed that there would be this precession, which could not be explained by the presence of known planets (it was conjectured that some other unobserved planet may exist that was causing this effect, but now we know that the more correct theory of general relativity predicts this precession).



We consider orbits that are nearly circular/small eccentricity of the (elliptical) orbit. Thus we can consider fluctuations around the stable circular orbit at $r = r_+$. The effective potential looks like:

$$V_{\text{eff}}(r) = -\frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3} \tag{11.46}$$

and assuming a minimum at $r = R \gg r_s$, we have:

$$V'_{\text{eff}}(R) = 0 \tag{11.47}$$

Calculating the angular momentum there:

$$L^2 = \frac{GM}{R^2}\frac{1}{1 - \frac{3GM}{R}} \approx \frac{GM}{R^2}\left(1 + \frac{3GM}{R}\right) \tag{11.48}$$

and hence:

$$V''_{\text{eff}}(R) = \frac{GM}{R^3}\left(1 - \frac{3GM}{R}\right) \tag{11.49}$$

The angular velocity is given by:

$$r^2\dot{\phi} = L \implies \dot{\phi} = \frac{L}{R^2} \tag{11.50}$$

Looking at the period, we have:

$$T_{\text{period}} = \frac{2\pi}{\sqrt{V''}} \approx \frac{2\pi R^{3/2}}{\sqrt{GM}}\left(1 + \frac{3}{2}\frac{GM}{R}\right) \tag{11.51}$$

Then we find the precession:

$$\Delta\Phi = T_{\text{period}}\frac{K}{R^2} = 2\pi\left(1 + \frac{3GM}{R}\right) \neq 2\pi \tag{11.52}$$

12 Stars in Equilibrium

12.1 Review of the solution

Today, we discuss stars in equilibrium. Last time, we discussed the solution of the (spherically symmetric) Einstein equation in free space, i.e. outside of the star/maassive body at the origin. Today, we discuss the solutions inside of the star, where the effects of matter cannot be ignored.

Recall from last time that, in the presence of spherical symmetry, the metric looked like:

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\beta(r)}dr^2 + r^2\underbrace{(d\theta^2 + \sin^2\theta d\phi^2)}_{d\Omega^2} \tag{12.1}$$

where the two unknown functions to find were $\alpha(r), \beta(r)$. We then found that the nonzero components of the Ricci tensor to be:

$$R_{tt} = e^{2(\alpha-\beta)}(\alpha'' + \alpha'^2 - \alpha'\beta' + \frac{2}{r}\alpha') \tag{12.2}$$

$$R_{rr} = -\alpha'' - \alpha'^2 + \alpha'\beta' + \frac{2}{r}\beta' \tag{12.3}$$

$$R_{\theta\theta} = e^{-2\beta}\Big[r(\beta' - \alpha') - 1\Big] + 1 \tag{12.4}$$

$$R_{\phi\phi} = \sin^2\theta R_{\theta\theta} \tag{12.5}$$

Last lecture, we. looked at vacuum, in which case $R_{\mu\nu} = 0$ and so we found the Schwartzchild solution.

12.2 Solving the Einstein equations in the presence of matter

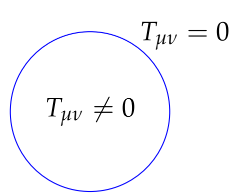
However, now we have matter! We thus also compute:

$$R = -2e^{-2\beta}(\alpha'' + \alpha'^2 - \alpha'\beta' + \frac{2}{r}(\alpha' - \beta') + \frac{1}{r^2}(1 - e^{2\beta})) \tag{12.6}$$

And now we solve the Einstein equation:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} \tag{12.7}$$

for a setting where the stress tensor is nonzero inside the star, and zero inside:



We assume that inside of the star we have a fluid, so:

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu} \tag{12.8}$$

we assume the fluid is at rest, so:

$$u^\mu = (u^0, \mathbf{0}) \tag{12.9}$$

normalized such that:

$$g_{\mu\nu}u^\mu u^\nu = -1 \tag{12.10}$$

Therein, we take:

$$u^0 = e^{-\alpha} \tag{12.11}$$

s.t.:

$$u^\mu = (e^{-\alpha}, 0) \tag{12.12}$$

$$u_\mu = (-e^{-\alpha}, 0) \tag{12.13}$$

The stress tensor then takes the form:

$$T^\mu_{\ \nu} = (\rho + P)u^\mu u_\nu - P\delta^\mu_\nu = \text{diag}(\underbrace{-\rho}_{t}, \underbrace{P}_{r}, \underbrace{P}_{\theta}, \underbrace{P}_{\phi}) \tag{12.14}$$

We assume that the matter is at zero temperature $T = 0$, in which case the equation of state of the matter is specified by a single parameter, either the chemical potential μ or the density ρ. Recall from your thermodynamics/stat mech course that this then determines other thermodynamic quantities, e.g. $P = P(\mu, T = 0) = P(\mu)$. We assume that there is a conserved charge/quantity, namely baryon number. This is what will allow us to say that the equation of state is specified by the baryon density (or the conjugate variable μ).

Looking at the tt equation:

$$R^t_{\ t} - \frac{1}{2}\delta^t_t R = T^t_{\ t} \implies \frac{1}{r^2}e^{-2\beta}(2r\beta' - 1 + e^{2\beta}) = 8\pi G\rho \tag{12.15}$$

The rr equation is:

$$\frac{1}{r}e^{-2\beta}(2r\alpha' + 1 - 2e^{2\beta}) = 8\pi GP \tag{12.16}$$

Note that $R_{\mu\nu} = 0$ for $\mu \neq \nu$ as $T_{\mu\nu}$ is diagonal.

Multiplying the top equation by r^2, we find:

$$(-re^{-2\beta})' + 1 = 8\pi G r^2 \rho \tag{12.17}$$

and further:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r(1 - e^{-2\beta})\right) = 8\pi G r^2 \rho \tag{12.18}$$

Defining the parameter:

$$m(r) \equiv \frac{1}{2G} r(1 - e^{-2\beta}) \tag{12.19}$$

We then find the tt equation becomes:

$$m'(r) = 4\pi r^2 \rho(r) \tag{12.20}$$

and the rr component of the metric becomes:

$$g_{rr} = e^{2\beta} = \frac{1}{1 - \frac{2Gm(r)}{r}} \tag{12.21}$$

Note that this makes explicit the comparison with the Schwartzchild solution, where:

$$g_{rr} = \frac{1}{1 - \frac{2GM}{r}} \tag{12.22}$$

wherein we note that $m(r) = M$ if $r > R$ to agree with the Schwartzchild solution.

Ok, let's solve the tt equation by integrating both sides:

$$m(r) = m_0 + \int_0^r dr' 4\pi r'^2 \rho(r') \tag{12.23}$$

Requiring that $m \to 0$ when $r \to 0$ (and so $g_{rr} \to 1$), we require that $m_0 = 0$. Else we get a singularity at $r = 0$, which we do not expect for a regular star.

The equation for $m(r)$ is very suggestive in flat space - this is just the total mass/energy inside of a sphere of radius r. But this interpretation is not correct in curved spacetime, as instead of $dV = 4\pi r^2 dr$ of Euclidean space we have $dV = \sqrt{g} d^3x = 4\pi r^2 e^{\beta} dr$.

12.3 Solving in the Newtonian and Non-Newtonian limit

In terms of $m(r)$, the α equation becomes:

$$\alpha' = G\frac{m(r) + 4\pi r^3 P(r)}{r(r - 2Gm(r))} \tag{12.24}$$

Let's check that this equation makes sense in the Newtonian limit, where $P(r) \ll \rho(r)$ (pressure is very small compared to energy density). Further since $m = \int r^2 dr \rho \sim r^3 \rho$, we have $r^3 \rho \ll m(r)$ and $Gm(r) \ll r$, and the α' equation reduces to:

$$\alpha' = G\frac{m(r)}{r^2} = F_r^{\text{grav}} = \frac{\partial \Phi(r)}{\partial r} \tag{12.25}$$

Thus ignoring the integration constant, $\alpha = \Phi$, and:

$$g_{tt} = -e^{2\alpha} = -(1 + 2\Phi) \tag{12.26}$$

which is exactly the Newtonian result we found previously!

But the α' equation of course is fully general - let us look at it in the general case. We have two equations (tt, rr), instead of using the angular equations as extra data, let us use the conservation equation for the stress tensor:

$$\boldsymbol{\nabla}_\mu T^\mu_\nu = 0 \implies \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g} T^\mu_\nu) - \Gamma^\lambda_{\mu\nu} T^\mu_\lambda = 0 \tag{12.27}$$

If we look at the $\nu = r$ component (noting that $\sqrt{-g} = \sqrt{r^4 e^{2\alpha\beta} \sin^2\theta}$), we find:

$$\frac{1}{r^2 e^{\alpha+\beta}}\frac{\partial}{\partial r}(r^2 e^{\alpha+\beta} P) - \underbrace{\Gamma^t_{tr}}_{\alpha'}\underbrace{T^t_t}_{-\rho} - \underbrace{\Gamma^r_{rr}}_{\beta'}\underbrace{T^r_r}_{P} - \underbrace{\Gamma^\theta_{\theta r}}_{\frac{1}{r}}\underbrace{T^\theta_\theta}_{P} - \underbrace{\Gamma^\phi_{\phi r}}_{\frac{1}{r}}\underbrace{T^\phi_\phi}_{P} = 0 \tag{12.28}$$

After some algebra, this reduces to the delightfully simple equation:

$$\alpha' = -\frac{P'}{\rho + P} \tag{12.29}$$

Recall from stat mech that:

$$dE = TdS - PdV + \mu dN \tag{12.30}$$

Then with $\rho = E/V$, $m = N/V$, $s = S/V$ we find:

$$d\rho = Tds + \mu dn \tag{12.31}$$

but we work in the $T = 0$ limit so:

$$d\rho = \mu dn \tag{12.32}$$

now using the non-differential form:

$$E - TS + PV - \mu N = 0 \implies P = \mu n - \rho \tag{12.33}$$

$$dP = n d\mu \tag{12.34}$$

Thus we can write our α' equation as:

$$\alpha' = -\frac{n\mu'}{\mu n} = -\frac{\mu'}{\mu} \tag{12.35}$$

wherein integrating yields:

$$\alpha + \ln \mu = \text{const.} \tag{12.36}$$

or equivalently:

$$e^{\alpha}\mu = \sqrt{-g_{tt}}\mu = \text{const.} \tag{12.37}$$

Thus, as we go to larger r, α increases while μ decreases such that the above product remains fixed.

This equation also has a nice non-relativistic limit, where:

$$g_{tt} = 1 + 2\Phi \tag{12.38}$$

$$\mu = m + \mu_{NR} \tag{12.39}$$

(where $\mu_{NR} \ll m$). Then the equation becomes:

$$\sqrt{1+2\Phi}(m + \mu_{NR}) = \text{const.} \implies \mu_{NR} + \underbrace{m\Phi}_{U(r)} = \text{const.} \tag{12.40}$$

which is the usual stat mech relation, where an equilibrium stat mech system in a field has $\mu + U(r)$ fixed. You may have seen this in the form of the barometric formula, which determines pressure as a function of height.

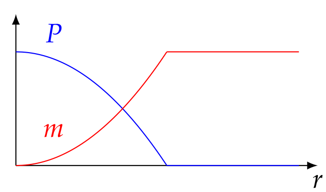
Ok, back to the general case! Combining everything, we have the equations:

$$\frac{dP}{dr} = -\frac{(\rho + P)\left[Gm + 4\pi G r^3 P\right]}{r(r - 2GM)} \tag{12.41}$$

$$\frac{dm}{dr} = 4\pi r^2 \rho \tag{12.42}$$

$$P = P(\rho) \tag{12.43}$$

We can think about solving these equations numerically, starting at $r = 0$ where $m(r = 0)$ and $P(r = 0) = P_c$. At some point, $P(r = R) = 0$ (at the radius of the star) and from then on onwards $m(r \geq R)$ is constant.

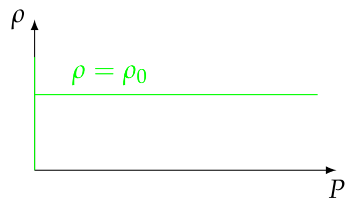


12.4 Incompressible Case

In order to solve this equation precisely, we need to know the equation of state. Let us solve this for the unrealistic case of:

$$\rho = \begin{cases} \rho_0 & r < R \\ 0 & r > R \end{cases} \tag{12.44}$$

noting that ρ is a constant w.r.t. pressure, i.e. the matter is completely incompressible.



Then, we find simply that:

$$m(r) = \frac{4\pi r^3}{3}\rho_0 \tag{12.45}$$

and the pressure equation becomes:

$$\frac{dP}{dr} = -(P+\rho_0)G\frac{\frac{4\pi}{3}\rho_0 r^3 + 4\pi r^3 P}{r(r - \frac{8\pi G}{3}\rho_0 r^3)} \tag{12.46}$$

This can be solved via separation of variables, noting that:

$$\frac{dP}{dr} = -(P+\rho_0)(3P+\rho_0)\frac{4\pi G r}{3 - 8\pi G\rho_0 r^2} \tag{12.47}$$

Then:

$$\int \frac{dP}{(P+\rho_0)(3P+\rho_0)} = -\int \frac{4\pi G r}{3 - 8\pi G\rho_0 r^2} dr \tag{12.48}$$

Solving the integrals, we get[1]:

$$\frac{1}{2\rho_0}\ln\left(\frac{3P+\rho_0}{P+\rho_0}\right) = \frac{1}{4\rho_0}\ln(3 - 8\pi G\rho_0 r^2) + \text{const.} \tag{12.49}$$

with the constant fixed by $P(r=R) = 0$. Exponentiating and applying this boundary condition, we get:

$$\left(\frac{3P(r)+\rho_0}{P(r)+\rho_0}\right)^2 = \frac{3 - 8\pi G\rho_0 r^2}{3 - 8\pi G\rho_0 R^2} \tag{12.50}$$

We can further massage this expression to get the relations that would appear in textbooks. But in this form, we already get curious relations looking at various points.

At $r = R$, note that $1 = 1$ and the equation tells us nothing. Note that at $r = 0$ we find that:

$$\frac{3}{3 - 8\pi G\rho_0 R^2} < 9 \tag{12.51}$$

so then:

$$G\rho_0 R^2 < \frac{1}{3\pi} \tag{12.52}$$

i.e. the star cannot be larger than a certain critical radius - we have the constraint:

$$\frac{GM}{R} < \frac{4}{9} \tag{12.53}$$

where $M = \frac{4\pi}{3}\rho_0 R^3$.

[1] Son: "I did this calculation on a plane, so any mistakes come from there."

One way to think about this - if we keep expanding R, the volume and mass keeps increasing (as R^3, v.s. the denominator going as R) so this is eventually violated, and the pressure at the center of the star reaches infinity - we cannot have a constant density object with an arbitrarily large diameter that is at equilibrium. Note that this is not the case in Newtonian gravity, where:

$$\frac{dP}{dr} = -\frac{Gm(r)\rho(r)}{r^2} = \rho(r)\boldsymbol{\nabla}\Phi \tag{12.54}$$

I.e. the equation of hydrostatic equilibrium/force balance. We can derive it easily by considering a small volume of matter, where the hydrostatic force from the pressure is balanced by the gravitational force. Solving this equation, we find that there is no regime in which the pressure at the center of the star becomes infinite.

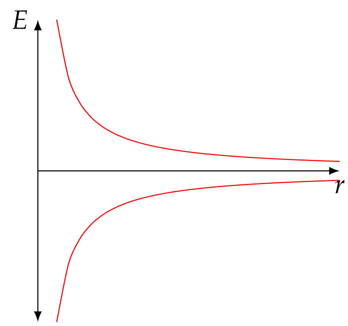
What is the physical interpretation of this? We have found that a sufficiently heavy, $T = 0$ star will, at the end of its life, collapse into a black hole. Our derivation looked at a particular equation of state (where matter is incompressible) but physically it is easy to intuit that if matter is compressible then it only becomes easier for the star to collapse.

We also said that in Newtonian gravity if the star is incompressible then it can be in equilibrium for any mass. But for a realistic equation of state, we will find it also collapses even in the Newtonian limit - this is the physics of white dwarfs. The matter in a white dwarf can be thought as a collection of ions and electrons. At very high densities, the electrons become relativistic.

How to see that this configuration cannot be stable at arbitrary masses? If the pressure is due to the gas of relativistic electrons, then I can show that there is no equilibrium minimum of the energy. The energy of the star comes in two parts, gravitational and the energy of a relativistic gas of electrons:

$$E = -\frac{GM^2}{R} + N_{el}\epsilon_F = -\frac{GM^2}{R} + N_{el}p_F c \tag{12.55}$$

With ϵ_F the Fermi energy. Then note that (as can be derived in quantum stat mech) $p_F \sim n^{1/3}$ and $n \sim \frac{N_{el}}{R^3}$. Hence both terms are of order R, and hence depending on the competition of the two coefficients, either the star wants to expand or the star wants to collapse, so we do get an unstable solution.



Note that white dwarfs are for $M < 1.4M_{\text{sun}}$. Understanding neutron stars is a bit more complicated, and we now know neutron star physics to kick in at 2-3 solar masses.

13 Schwarzchild Black Holes

The main things we will discuss the fully extended Schwarzchild solution, describing an eternal black hole.

13.1 The event horizon and the Rindler metric

We have previously derived the Schwarzchild metric:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{2GM}{r}} + r^2 d\Omega \tag{13.1}$$

at $r = 0$ we have a singularity, and another apparent divergence at $r = r_H = 2GM$. But if we compute $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho}$ we will find that this is is actually finite at r_H - generally, we will find today that spacetime at the horizon is not singular. Further, we will discuss how to extend spacetime beyond the black hole horizon.

If we have a star with radius $R > r_H$ then we need not worry about this complication (as the Schwarzchild solution applies in the vacuum/outside of the star). But if we have a stellar body with $R < r_H$ then the point $r = r_H$ outside of the body becomes a relevant point in spacetime to consider! So let us do it. Consider a point near the horizon:

$$r = r_H + \underbrace{(r - r_H)}_{\ll r_H} \tag{13.2}$$

the angular part of the metric becomes $\approx r_H^2 d\Omega^2$, which is regular/nonsingular.

The term in the temporal/radial part of the metric is, to linear order:

$$1 - \frac{2GM}{r} = 1 - \frac{r_H}{r} \approx \frac{r - r_H}{r_H} + O((r - r_H)^2) \tag{13.3}$$

So the metric becomes (leaving off the angular part/focusing on the 2D part):

$$ds^2 = -\frac{r - r_H}{r_H} dt^2 + \frac{r_H}{r - r_H} dr^2 \tag{13.4}$$

Making the coordinate transformation $r - r_H = \rho^2$, we have $dr = 2\rho d\rho$ and soi:

$$ds^2 = -\frac{1}{r_H}\rho^2 dt^2 + 4r_H d\rho^2 \tag{13.5}$$

By rescaling t, ρ we obtain the simpler form:

$$\boxed{ds^2 = -\rho^2 dt^2 + d\rho^2} \tag{13.6}$$

the above is the *Rindler metric*, describing Rindler space.

13.2 Light-cone coordinates and Geodesics

Now, there is a simple coordinate transformation that takes us into the plane. To motivate this, let us study the light cones, which are the worldlines of objects that study with velocity 1/the speed of light. For such worldlines, we have $ds^2 = 0$, so then $d\rho = \pm\rho dt$ and hence $d\ln\rho = \pm dt$. The solutions to this are:

$$t \pm \ln\rho = \text{const.} \tag{13.7}$$

Thus, we transform into:

$$u = t - \ln \rho, \quad v = t + \ln \rho \tag{13.8}$$

So:

$$t = \frac{u+v}{2}, \quad \rho = e^{\frac{v-u}{2}} \tag{13.9}$$

Thus the metric becomes:

$$ds^2 = -\rho^2(dt^2 - (d\ln\rho)^2) = -\rho^2 d(t + \ln\rho)d(t - \ln\rho) = -e^{v-u}dudv \tag{13.10}$$

the original space had $-\infty < t < \infty$ and $0 < \rho < \infty$. In the new coordinates, we have $-\infty < u, v < \infty$. But, we claim that this range of the coordinates is not complete. Why? We can imagine that we are drawing out a geodesic, but then we hit the "end " of spacetime (i.e. the range of the coordinates), but the geodesics can continue past the end! To make this more concrete, let's look at light-like geodesics in Rindler space.

Consider V^μ the tangent vector and an affine parametrization $V^\nu \nabla_\nu V^\mu = 0$. Further, recall that if we have a Killing vector ξ_μ, then:

$$\xi_\mu \frac{dx^\mu}{d\lambda} = \text{const.} \tag{13.11}$$

so:

$$\xi_\mu V^\mu = \text{const.} \tag{13.12}$$

Let us look at geodesics where u is constant, and v changes. I.e.:

$$u = t - \ln \rho = C \tag{13.13}$$

Thus v grows (as both $t, \ln\rho$ grow). We can then ask what the Killing vector we have is - it is $\xi = \frac{\partial}{\partial t}$! Then, we find that:

$$\xi^\mu = (\underbrace{1}_{\xi^t}, \underbrace{0}_{\xi^\rho}) \tag{13.14}$$

so:

$$\xi_\mu = (-\rho^2, 0) \tag{13.15}$$

Thus for the geodesics, we have the conserved quantity:

$$E = \rho^2 \frac{dt}{d\lambda} = \text{const.} \tag{13.16}$$

This has the physical meaning of the energy, as it comes from the invariance of the metric under time translations.

In terms of u, v we have:

$$e^{v-u}\left(\frac{\partial u}{\partial \lambda} + \frac{\partial v}{\partial \lambda}\right) = \text{const.} \tag{13.17}$$

Again, we study the outgoing geodesics, with $u = \text{const.}$ and $v = t + \ln \rho$. Therein:

$$e^v \frac{\mathrm{d}v}{\mathrm{d}\lambda} = \text{const.} \tag{13.18}$$

Thus:

$$d\lambda = e^v dv \implies \lambda = e^v + \text{const.} \tag{13.19}$$

But then (taking the constant to be 0), we find that $0 < \lambda < \infty$ - but λ is a parameter of our geodesic, and should be allowed to go negative! So, we have to think about coordinate ranges beyond $-\infty < v < \infty$.

Making the same argument for ingoing trajectories, we have:

$$e^{-u} \frac{\mathrm{d}u}{\mathrm{d}\lambda} = \text{const.} \tag{13.20}$$

so:

$$-d(e^{-u}) = d\lambda \implies \lambda = -e^{-u} + \text{const.} \tag{13.21}$$

13.3 Extending Rindler Space

So, motivated by this, let us take the new system of coordinates:

$$\boxed{V = e^v, U = -e^{-u}} \tag{13.22}$$

then:

$$dV = e^v dv, \quad dU = e^{-u} du \tag{13.23}$$

Then the Rindler metric becomes:

$$ds^2 = -dUdV \tag{13.24}$$

Let us make things even more simple, by transforming:

$$T = \frac{U+V}{2}, \quad X = \frac{V-U}{2} \tag{13.25}$$

then:

$$V = T + X, \quad U = T - X \tag{13.26}$$

so then:

$$\boxed{ds^2 = -dT^2 + dX^2} \tag{13.27}$$

just the Minkowski metric in T, X!!

So we have jumped through the hoops of:

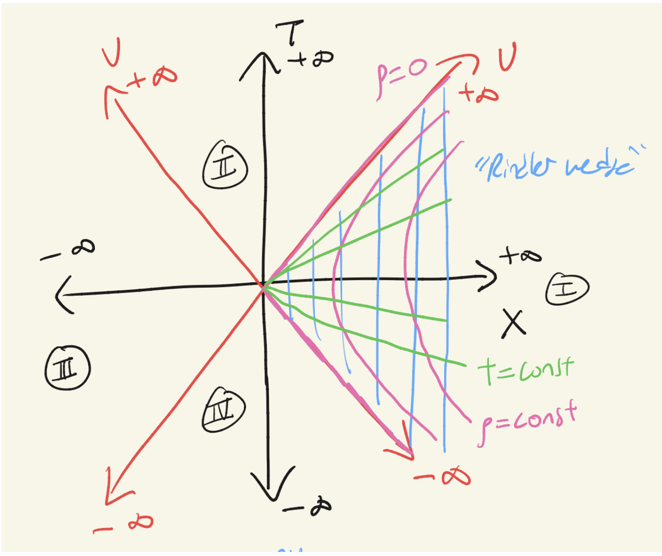
$$(t, \rho) \to (u, v) \to (U, V) \to (T, X) \tag{13.28}$$

Explicitly, the transformation we have done is:

$$\rho = e^{\frac{v-u}{2}} = \sqrt{-VU} = \sqrt{X^2 - T^2} \tag{13.29}$$

$$t = \frac{v+u}{2} = \frac{1}{2}\ln V - \frac{1}{2}\ln(-U) = \frac{1}{2}\ln\frac{X+T}{X-T} = \operatorname{arctanh}\frac{T}{X} \tag{13.30}$$

What have we accomplished? The range of $-\infty < u, v < \infty$ is only part of the range of U, V, i.e. $0 < V < \infty$, $-\infty < U < 0$ (only one quarter of the range)! But with our new set of coordinates we have tried to free things. Explicitly, we can sketch:



(U, V) and (T, X) give us access to the whole of Minkowski space, while the original sets of coordinates only gave us access to one quarter of the space (blue) in the "Rindler Wedge". In the Rindler wedge, lines of $\rho = $ const. (pink) are hyperbolae, with $\rho = 0$ being the edge of space. $t = $ const. (green) lines start at the origin and go outwards towards infinity in the wedge.

One curious remark - the original Rindler space has a Killing vector, an invariance under time translation $t \to t + $ const.. In the new coordinate system, this becomes an invariance under boosts!

Note also the causality structure between the regions $I - IV$, which is the familiar causality structure of Minkowski.

13.4 Transforming the Schwarzchild Metric

We have our Schwarzchild metric:

$$ds^2 = -\left(1 - \frac{2GM}{R}\right)\left[dt^2 - \frac{dr^2}{(1 - \frac{2GM}{r})^2}\right] + r^2 d\Omega^2 \tag{13.31}$$

we drop the angular term for it is regular, and let us define:

$$dr_* = \frac{dr}{1 - \frac{2GM}{r}} \tag{13.32}$$

so then:

$$r_* = r + 2GM \ln(\frac{r}{2GM} - 1) \tag{13.33}$$

We then obtain:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right)\left(dt^2 - dr_*^2\right) \tag{13.34}$$

Now introducing u, v as in the Rindler case, we have:

$$u = t - r_*, v = t + r_* \tag{13.35}$$

wherein the metric can be written as:

$$ds^2 = -\frac{2GM}{r} e^{-r/2M} e^{\frac{v-u}{4M}} du dv \tag{13.36}$$

The prefactor $-\frac{r_H}{r} e^{-r/2M}$ is regular (and hence uninteresting) as $r \to r_H = 2GM$.For the remaining part, let us make an analogous coordinate transformation to what we did before:

$$U = -e^{-u/4M}, \quad V = e^{v/4M} \tag{13.37}$$

wherein:

$$ds^2 = \frac{32(GM)^3 e^{-r/2GM}}{r} dU dV \tag{13.38}$$

and then finally going to:

$$T = \frac{V + U}{2}, \quad X = \frac{V - U}{2} \tag{13.39}$$

the metric is brought to the form:

$$ds^2 = \frac{32(GM)^3 e^{-r/2GM}}{r}(-dT^2 + dX^2) \tag{13.40}$$

this is not a completely flat metric due to the prefactor (which must be understood as some function of r_*).

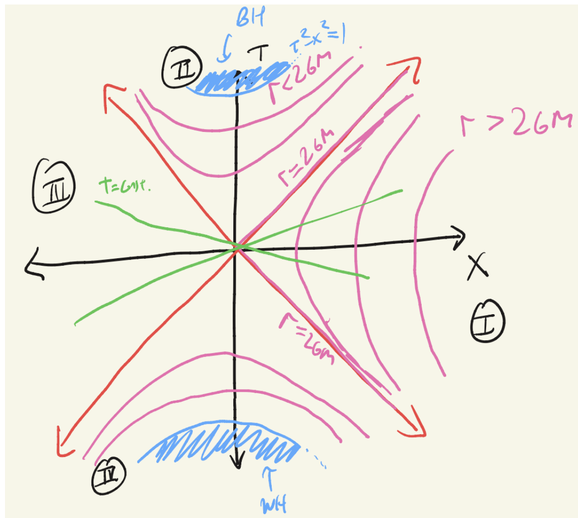
13.5 Black and White Holes

Noting that:

$$X^2 - T^2 = -UV = -e^{-\frac{v-u}{4GM}} = e^{\frac{r_*}{2GM}} = e^{\frac{r}{2GM}}(\frac{r}{2GM} - 1) \tag{13.41}$$

$$\ln \frac{X+T}{X-T} = \ln \frac{V}{-U} = \frac{1}{4M}(v+u) = \frac{t}{2M} \tag{13.42}$$

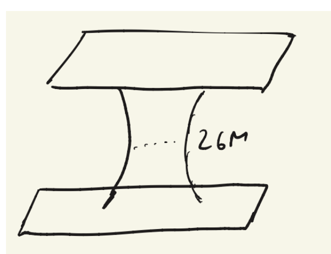
We can sketch:



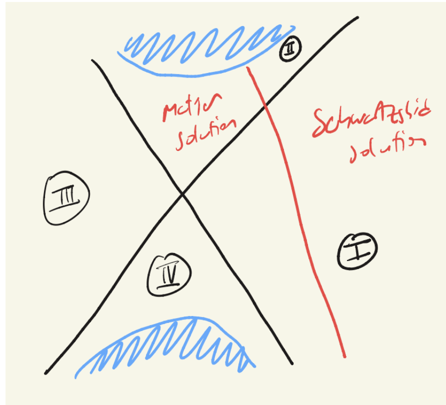
Where r = const. $> 2GM$ solutions are hyperbolae in region I, r = const. $= 2GM$ correspond to the horizon. We have valid trajectories in regions II/IV as well (corresponding to $r < 2GM$), with the $T^2 - X^2 = 1$ lines denoting black/white holes. Lines corresond to constant time trajectories (going through regions III/I).

Causality wise - events in region II are in the future, and cannot influence anything outside of the BH horizon (formal notion of something falling into a black hole not being able to influence things outside).

This static solution predicts a wormhole between a white and black hole:



But if we think about a star that collapses and becomes a black hole, only part of the solution can be used - the cloud of matter/the matter distribution (red line below) defines a radius outside of which the Schwarzchild solution is valid, and inside of it we must consider the solutions with matter. For a collapsing star, there is no well defined meaning/access to the white hole (in region IV), only outside of the black hole horizon in I and inside the horizon in II.

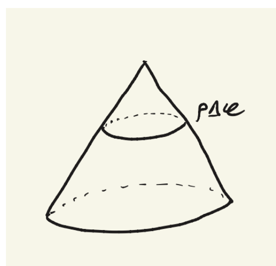


13.6 Black hole temperature

Looking at the Rindler metric again, we have:

$$ds^2 = -\rho^2 dt^2 + d\rho^2 = d\rho^2 + \rho^2 d\phi^2 \tag{13.43}$$

ϕ has periodicity $\phi \sim \phi + \Delta\phi$, and this defines circles with perimeter $\rho\Delta\phi$ on a cone.



Note that we can consider a correspondence $t \to i\phi$, with it imaginary time. The periodicity $\phi \sim \phi + 2\pi$ gives a periodicity $it \sim it + \beta$, and using the time-temperature correspondence $e^{-iHt} \leftrightarrow e^{-\beta H}$ is actually how we can find the temperature $\frac{1}{T} = \beta$ of a black hole.

Next time, we discuss cosmology/solutions to Einstein's equation for the whole universe.

14 GR for the Universe

Today, we will start to look at general relativity applied to the universe. We start by looking at the modification of the Einstein equations in the presence of a cosmological constant.

14.1 EFT and the Cosmological Constant

Recall that we derived the Einstein equations from the Einstein-Hilbert action:

$$S = \frac{1}{16\pi G}\int d^4x\sqrt{-g}R \tag{14.1}$$

where $R = R^{\mu}_{\ \mu}$. The motivation we had for this at the time was that the Lagrangian density must be a scalar, and the simplest such term one could construct from the Riemann tensor was the Ricci scalar R.

However, this is not the only scalar that one could add to the action, we could for example consider at quadratic order:

$$S = \frac{1}{16\pi G}\int d^4x\sqrt{-g}\left(R + c_1R^2 + c_2R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho} + c_3R_{\mu\nu}R^{\mu\nu}\right) \tag{14.2}$$

from the modern point of view, the lens of so called "effective field theory" (a very successful way of looking at QFT in flat space!), there is an in principle infinite number of terms one could include in the Lagrangian, and they are classified in terms of their dimension. What this implies for us here is that the relative constants c_1, c_2, c_3 are dimensionful parameters. Specifically, since $R \sim \partial^2 g_{\mu\nu} \sim \frac{1}{L^2}$, it follows that c_i must have dimension of L^2 in order to cancel out the extra power of $\frac{1}{L^2}$ coming from the second power of R.

In the proceeding argument, each of the steps are things that are generally accepted, but it is good to question them, since we don't have a full understanding of quantum gravity! We believe[2] that the natural length scale is the Planck scale l_P, so $c_i \sim l_P^2$ with $l_P \sim G^{\#}\hbar^{\#}c^{\#}$ and $l_P \approx 10^{-33}$cm - this tells us why we might be able to ignore these terms in most cases, because we generally work at length scales $L > l_P$.

Note however that this argument tells us that there is a term that we have missed! Namely, a constant:

$$S = \frac{1}{16\pi G}\int d^4x\sqrt{-g}\left(R + c_1R^2 + c_2R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho} + c_3R_{\mu\nu}R^{\mu\nu} - 2\Lambda\right) \tag{14.3}$$

Dimension counting tells us that this constant has dimension L^0, i.e. is order 1! Since terms become less important in higher orders of R, this tells us that the dynamics of this constant term should be more important than the R term we already have.

Λ is the cosmological constant, and the "naturalness" argument above suggests that this should be large (M_P^4). Indeed, this is one of the deepest rooted mysteries in physics, where the prediction of Λ from the QFT vacuum energy and that measured in cosmology differ by 120 orders of magnitude!!

Nevertheless, let us press on and see what happens with the inclusion of this term. We have the action:

$$S = \int d^4x\frac{\sqrt{-g}}{16\pi G}(R - 2\Lambda) + S_{\text{matter}} \tag{14.4}$$

then looking fo the equations of motion:

$$\frac{\delta S}{\delta g^{\mu\nu}} = 0 \tag{14.5}$$

we find:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = 8\pi G T^{\text{matter}}_{\mu\nu} \tag{14.6}$$

[2] Based on a theoretical bias coming from QFT/EFT, where one expands in powers of the field, ad the coefficients are based on the ultraviolet cutoff of the theory, and hence suppressed...

where the LHS comes from the free part of the action and the RHS comes from the matter part of the action. So, we can see that the cosmological constant modifies the LHS of the Einstein equations.

But, we can have a second perspective on this cosmological constant, wherein we treat the Λ term as an extra term in the matter part of the action. Then we may write:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G(T_{\mu\nu}^{\text{matter}} - \frac{g_{\mu\nu}}{8\pi G}\Lambda). \tag{14.7}$$

In the QFT view, we treat the $\frac{g_{\mu\nu}}{8\pi G}\Lambda$ term as the vacuum energy(-momentum tensor):

$$T_{\mu\nu}^{\text{vacuum}} = \frac{\Lambda}{8\pi G}g_{\mu\nu} \tag{14.8}$$

where:

$$\rho_{\text{vacuum}} = T_{00}^{\text{vacuum}} = \frac{\Lambda}{8\pi G} \tag{14.9}$$

i.e. in the absence of matter there is still a nonzero energy density.

Usually, in physics we don't care about where we set the zero point of the energy - e.g. in the thermodynamics, we don't care about the absolute value of the energy, only the energy changes ΔE. But in gravity it does matter, because any amount of mass-energy gravitates. So it is relevant to ask whether ρ_{vacuum} is zero or nonzero. And in fact QFT predicts it to be very large (coming from the zero-point energy modes of the field), while in the real world we measure it to be small. We may return to questions about the magnitude of this constant later.

14.2 Possibilities for a Symmetric Universe and the FLRW Metric

We will take it as a given (and as has been confirmed by experiment) that the universe is homogenous (every point in the universe is equivalent) and isotropic (from a given point, all directions look equivalent). Thus, we will look for metrics/solutions which look like $\sim$ time $\otimes$ symmetric space. We will take as our definition of symmetric space that:

$$R_{\mu\nu\lambda\rho} = c(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) \tag{14.10}$$

i.e. the only things that we have our disposal is the metric, and in order to make the Riemann tensor inherit the symmetries of the metric + satisfy:

$$R_{\mu\nu\lambda\rho} = R_{\nu\mu\lambda\rho} \tag{14.11}$$

$$R_{\mu\nu\lambda\rho} = R_{\lambda\rho\mu\nu} \tag{14.12}$$

while only making it up out of the metric, this is the only thing we could write down. Wald goes through the formal argument - if we view $R_{\mu\nu\lambda\rho}$ as a map, since space has no special directions, all eigenvalues of the tensor have to be the same, and so we get this form. c being constant in space is a consequence of the uniformity of the universe.

In 3D, there are 3 different types of symmetric spaces possible.

1. Flat Euclidean, with metric:

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\Omega^2 = dr^2 r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{14.13}$$

2. The sphere, wherein we have the constraint (τ here is again not time, but fictitious - we use it to embed the spacelike surface in a 4D Lorentzian space):

$$\tau^2 + x^2 + y^2 + z^2 = R^2 \tag{14.14}$$

we can parametrize this as:

$$\tau = R\cos\chi \tag{14.15}$$

$$x = R\sin\chi\cos\theta \tag{14.16}$$

$$y = R\sin\chi\sin\theta\cos\phi \tag{14.17}$$

$$z = R\sin\chi\sin\theta\sin\phi \tag{14.18}$$

wherein we find the metric/line element:

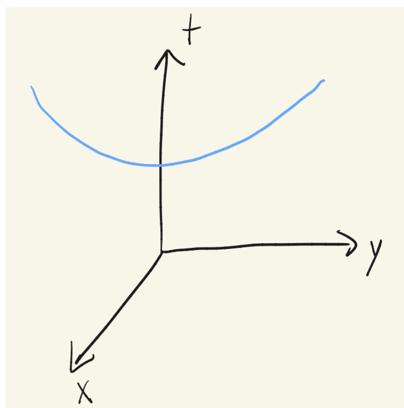
$$ds^2 = d\tau^2 + dx^2 + dy^2 + dz^2 = R^2[d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\phi^2)] \tag{14.19}$$

if we change coordinates to $\sin\chi = r$ (wherein $\cos\chi d\chi = dr$), we get:

$$ds^2 = R^2\left[\frac{dr^2}{1-r^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{14.20}$$

where $0 < r < \infty$ and $0 \leq \chi < \infty$.

3. The final possibility is hyperbolic. If we consider a hyperbola:



We can see that on such a surface, we obtain the constraint (t here is again not time, but fictitious)

$$t^2 - x^2 - y^2 - z^2 = R^2 \tag{14.21}$$

where we can again consider a parametrization:

$$t = R\cosh\chi \tag{14.22}$$

$$x = R\sinh\chi\cos\theta \tag{14.23}$$

$$y = R\sinh\chi\sin\theta\cos\phi \tag{14.24}$$

$$z = R\sinh\chi\sin\theta\sin\phi \tag{14.25}$$

where we can then write:

$$ds^2 = dx^2 + dy^2 + dz^2 - dr^2 = R^2(d\chi + \sinh^2\chi d\Omega^2) \tag{14.26}$$

where again defining $r = \sinh\chi$ we may write:

$$ds^2 = R^2(\frac{dr^2}{1+r^2} + r^2 d\Omega^2) \tag{14.27}$$

where $0 < r < 1$.

The constant in the Riemann tensor can be found to be:

$$c = \frac{R}{D(D-1)} \tag{14.28}$$

Note that the distinction between the three cases is that in flat space, R, on the sphere, $R > 0$, and for the hyperbola, $R < 0$. We can unify all three into the single equation for the spatial part of the metric:

$$ds^2 = R^2 \left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \right) \tag{14.29}$$

or in the time dependent case:

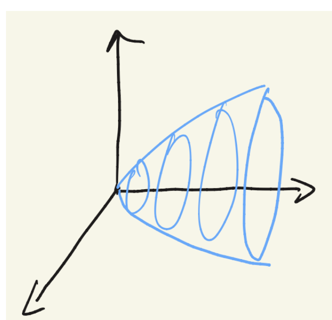
$$ds^2 = -dt^2 + R^2(t) \left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \right] \tag{14.30}$$

with $k = 0$ for flat, $k = 1$ for a sphere, and $k = -1$ for a hyperbola. Note that R is only fixed when the universe is a sphere (radius of the universe) or hyperbolic (wherein it is the curvature of the hyperbola). When space is flat, it is simply a free parameter we are free to factor out.

It is common to take $a(t) = \frac{R(t)}{R_0}$ where R_0 has dimensions of length (physically interpreted as the current size of the universe), making $a(t)$ dimensionless, wherein, we can write:

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1-\kappa r^2} + r^2 d\Omega^2 \right] \tag{14.31}$$

where $\kappa = k/R_0^2$, and $\kappa = 0$ is flat, $\kappa > 0$ for a sphere, and $\kappa < 0$ for a hyperbola. r now represents the comoving radial distance. $a(t)$ is the scale factor of the universe, which is normalized such that $a(t_0) = 1$. We an now think about the universe, e.g. expanding:



Note that the universe in this model is homogenous/isotropic in space, but it is *not* homogenous/isotropic in time.

This metric we have derived is known as the Friedmann-Lemaître-Robertson-Walker metric.

14.3 The Friedmann Equations

The Einstein equations read:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{14.32}$$

where:

$$T^{\mu\nu} = (p + P) u^\mu u^\nu + P g^{\mu\nu} \tag{14.33}$$

and $u^\mu = (1, 0, 0, 0)$ the velocity of the fluid, i.e. we assume the matter of the universe is at rest. For now, let us take this as an ansatz - then we can compute the geodesics of a particle in such a universe, finding that they come to rest (and showing the ansatz to be self-consistent).

We get two equations from this from looking at the different components (and doing some manipulations):

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3P) \tag{14.34}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} \tag{14.35}$$

these are known as the Friedmann equations.

Note that the conservation of energy-momentum:

$$\nabla_\mu T^\mu_\nu = 0 \tag{14.36}$$

gives us the third equation:

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \tag{14.37}$$

14.3.1 Pressureless Matter

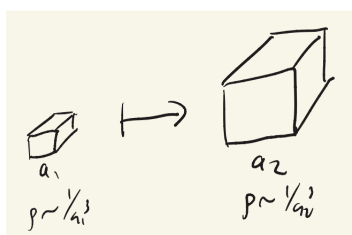
If we think about the limit of $P = 0$ ("pressureless matter") we find that:

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}\rho = 0 \tag{14.38}$$

which we can solve to find:

$$\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a} \implies \rho = \frac{\text{const.}}{a^3} \tag{14.39}$$

This is very intuitive; if the scale factor of space changes, the density scales up/down as the cube of the scale factor as space contracts/expands.



A quick, not quite correct, but intuitive argument to think about what happens to the energy during this process. If we go from a at time t to the $a + \delta a$ at $t + \delta t$, then the change of energy coming from the volume change looks like:

$$\Delta E = -P\Delta V \implies \Delta(\rho a^3) = -P\frac{3\delta a}{a}a^3 \implies \frac{\partial \rho}{\partial t} = -3(\rho + P)\frac{\dot{a}}{a} \tag{14.40}$$

i.e. the universe takes energy to expand.

14.3.2 Radiation

Another type of matter we can consider is radiation, the simplest kind to consider being a gas of photons. For this, we know that:

$$P = \frac{\rho}{3}. \tag{14.41}$$

Let us derive this. If we have a gas of photons, or anything made up out of the EM field, then the energy-momentum tensor must be traceless:

$$T^{\mu}_{\ \mu} = 0 \tag{14.42}$$

we could either compute the energy-momentum tensor for the EM field directly to see this, or for a more abstract/general argument let us look at the Maxwell action in curved space:

$$S = -\frac{1}{4}\int d^4x\sqrt{-g}g^{\mu\nu}g^{\alpha\beta}F_{\mu\alpha}F_{\nu\beta} \tag{14.43}$$

Note that this action is invariant under a rescaling of the matrix:

$$g_{\mu\nu}(x) \to e^{\alpha(x)}g_{\mu\nu}(x) \tag{14.44}$$

as the determinant factor cancels out the contributions from the two metrics. So, we find that (expanding out to first order):

$$S[g_{\mu\nu} + \alpha(x)g_{\mu\nu}] = S[g_{\mu\nu}] + O(\alpha^2) \tag{14.45}$$

but we know that the change of the action is given by the stress tensor, so:

$$S[g_{\mu\nu}] + \frac{1}{2}\int d^4x\sqrt{-g}\alpha(x)g_{\mu\nu}T^{\mu\nu} = S[g_{\mu\nu}] + O(\alpha^2) \tag{14.46}$$

And taking the trace of both sides we must conclude that the trace of $T^{\mu\nu}$ is zero. Thus:

$$T^{\mu}_{\ \mu} = -\rho + 3P = 0 \implies P = \frac{\rho}{3} \tag{14.47}$$

From this equation, if we assume that:

$$P = w\rho \tag{14.48}$$

the conservation equation becomes:

$$\frac{\partial\rho}{\partial t} = 3\frac{\dot{a}}{a}\rho(1+w) \tag{14.49}$$

And thus:

$$\rho(t) \sim a^{-3(1+w)}(t) \tag{14.50}$$

thus for $w = 0$ (pressureless) we have:

$$\rho \sim a^{-3} \tag{14.51}$$

and for photons we have:

$$\rho \sim a^{-4} \tag{14.52}$$

14.3.3 Vacuum Energy

For vacuum energy, we have $T^{\mu\nu} = -\Lambda g^{\mu\nu}$ and thus $p + P = 0$, so this corresponds to the $w = -1$ case. We then find that:

$$\rho \sim \text{const.} \tag{14.53}$$

so our label of "cosmological constant" is apt - it does not change with time/scale factor!

15 GR for the Universe II

Today we discuss the universe, described by the FLRW metric (the "FLRW universe"). We discuss its properties and its dynamics (Einstein's equation).

Recall the FLRW metric:

$$ds^2 = -dt^2 + R^2(t) \underbrace{\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right]}_{\gamma_{ij}dx^i dx^j} \tag{15.1}$$

where $k = 1$ is the sphere (closed/finite volume), $k = 0$ is a flat universe, and $k = -1$ is hyperbolic (open).

15.1 Photons in the FLRW Universe

Let us study a photon propagating in a universe described by this metric, and how its frequency changes with time. This will be characterized by the $R(t)$ scale parameter.

The Hubble constant is given by:

$$H(t) = \frac{\dot{R}(t)}{R(t)} \sim \frac{1}{\text{time}} \tag{15.2}$$

which describes the rate of expansion of the universe. Consider some $t = t_0$ with $H > 0$.

To study this problem, we play a slight mathematical trick; we parameterize things in terms of the "conformal time" $\eta = \eta(t)$ (which can be inverted to give $t = t(\eta)$), such that the metric in terms of the conformal time looks like:

$$ds^2 = R^2(\eta)\left[-d\eta^2 + \gamma_{ij}dx^i dx^j\right] \tag{15.3}$$

i.e. we want:

$$dt^2 = R^2 d\eta \implies d\eta = \frac{dt}{R(t)} \tag{15.4}$$

so then η is given by the integral:

$$\eta = \int^t \frac{dt'}{R(t')} \tag{15.5}$$

defined up to an arbitrary integration constant.

Let us consider the photon - this is a solution to the Maxwell equations, a quanta of the electromagnetic field. If we write the Maxwell action:

$$S = -\frac{1}{4}\int d^4x \sqrt{-g} g^{\mu\nu} g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} \tag{15.6}$$

which we note is invariant under conformal transformations:

$$g(x) \to g'(x) = e^{2\omega(x)} g(x) \tag{15.7}$$

i.e. the Maxwell action is Weyl invariant. This motivates the conformal time as a good parameterization.

In the flat case of $k = 0$, we have:

$$ds^2 = -R^2(\eta)\left[-d\eta^2 + dx^2\right] \implies A_\mu \sim e_\mu e^{-i\omega\eta + i\mathbf{q}\cdot\mathbf{x}} \tag{15.8}$$

with e_μ the polarization. Let us consider the case where $\omega \gg H$, i.e. the universe does not expand very much over one period of oscillation of the EM wave. At time t, recalling that η is not the true time but the conformal time, we have:

$$d\eta \sim \frac{dt}{R(t)} \tag{15.9}$$

so near $t = t_0$, we have approximately:

$$\eta - \eta_0 = \frac{t - t_0}{R(t_0)} + \dots \tag{15.10}$$

thus:

$$e^{-i\omega\eta} \sim e^{-i\frac{\omega}{R(t_0)}(t-t_0)} \tag{15.11}$$

the real frequency of the photon is thus identified as $\frac{\omega}{R(t_0)}$. Thus, we recognize that the frequency of the photon changes inversely with the scale factor of the universe.

So, if we have a source and detector, the frequency detected is smaller than when the photon was emitted, due to the universe changing in scale. More explicitly, if the photon is emitted at time t_1 with frequency ω, at time t_0 we detect it to have frequency:

$$\frac{R(t_1)}{R(t_0)}\omega. \tag{15.12}$$

This is the phenomenon of redshift.

Also, we need to recall that the $\mathbf{x}$ is not quite the measured distance, specifically:

$$e^{i\mathbf{q}\cdot\mathbf{x}} = e^{i\mathbf{q}\cdot\mathbf{X}/R} \tag{15.13}$$

with $x = X/R$. So, the wavelength of the photon increases as the scale of the universe increases.

15.2 Particles in the FLRW Universe

Let us consider a particle moving in the FLRW universe:

$$\frac{du^\mu}{ds} + \Gamma^\mu_{\nu\lambda} u^\nu u^\lambda = 0 \tag{15.14}$$

wherein the four-velocity is $u^\mu = \frac{dx^\mu}{ds}$.

Let us simplify by considering a non-relativistic particle, with $s \approx t + O(v^2)$, and $u^\mu = (1 + O(v^2), v^i)$, with $v^i \ll 1$.

Let's evaluate the dynamics of the spatial components:

$$\frac{du^i}{dt} + \underbrace{\Gamma^i_{00} u^0 u^0}_{=0} + 2\Gamma^i_{0k} \underbrace{u^0}_{1} u^k \tag{15.15}$$

where the first term vanishing can be seen from:

$$\Gamma^i_{00} = \frac{1}{2} g^{ik} (\partial_0 g_{0k} + \partial_0 g_{0k} - \partial_k g_{00}) \tag{15.16}$$

The second term looks like:

$$g^{ij} (\partial_0 \underbrace{g_{jk}}_{R^2\delta_{jk}} + \partial_k \underbrace{g_{0j}}_{0} - \partial_j \underbrace{g_{0k}}_{0}) u^k \tag{15.17}$$

And so the equations become (defining the rescaled $U^i = Ru^i$):

$$\frac{dU^i}{dt} = \dot{R}u^i + R \underbrace{\dot{u}^i}_{=-\frac{2\dot{R}}{R}u^i} = -\dot{R}u^i = -\frac{\dot{R}}{R} U^i = -HU^i \tag{15.18}$$

so particles moving an expanding universe experience a "friction force", with coefficient given by the Hubble constant. This is sometimes called the "Hubble friction".

We can solve these equations to get:

$$U^i(t) \sim \frac{1}{R(t)} \text{Const.} \tag{15.19}$$

15.3 Gases in the FLRW Universe

Imagine that we hav a thermal gas with a Maxwell distribution of the velocities:

$$p(\mathbf{v}) \sim e^{-\frac{mv^2}{2T}} \tag{15.20}$$

Since the velocity decreases as the universe expands, we can identify the temperature of the gas scaling as:

$$T \sim \frac{1}{R^{1/2}} \tag{15.21}$$

If I have an ultra-relativistic gas, then we can consider the Fermi-Dirac or Bose-Einstein distribution:

$$f_{\pm}(p) = \frac{1}{e^{|\mathbf{p}|/T} \pm 1} \tag{15.22}$$

wherein the decrease in momentum results in $T \sim \frac{1}{R}$.

15.4 Einstein's Equations in the FLRW Universe

Consider now the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{15.23}$$

where the Ricci scalar R unfortunately coincides with the notation of the scale factor R. "This is a long-standing problem in the literature that has not been solved." - Son. The energy-momentum tensor takes the form:

$$T_{\mu\nu} = (\rho + P) u_\mu u_\nu + P g_{\mu\nu} \tag{15.24}$$

The 00 components look like:

$$3\left(\frac{\dot{R}^2}{R^2} + \frac{\kappa}{R^2}\right) = 8\pi G \rho \tag{15.25}$$

The ij equations give:

$$2\frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{\kappa}{R^2} = -8\pi G \rho \tag{15.26}$$

The conservation law gives:

$$\nabla_\mu T^{\mu\nu} = 0 \implies \dot{\rho} + 3\frac{\dot{R}}{R}(\rho + P) = 0 \tag{15.27}$$

note that the second equation can be obtained by combining the first equation with the conservation law equation (which makes sense as the Einstein equation is self-consistent with the conservation laws).

The first equation can be rewritten as:

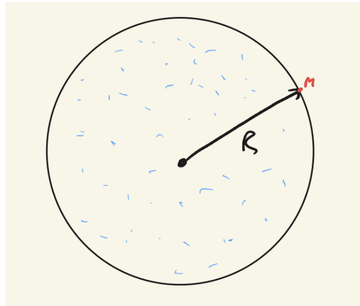
$$H^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{R^2} \tag{15.28}$$

and by taking linear combinations of the equations we get the second equation:

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}(\rho + 3P) \tag{15.29}$$

These are known as the Friedmann equations.

The second equation has a nice physical interpretation in the Newtonian limit. Suppose the universe is flat, and consider a region of size R filled with matter of density ρ and pressure P. Let us assume that this matter is nonrelativistic, so $P \ll \rho$. Further, let us assume that there is no matter outside of this region (note that this is not a correct model of the universe, because it assumes the universe is a finite region that is expanding...).



The gravitational force coming from this matter on a test mass m on the boundary looks like:

$$F = -\frac{GMm}{R^2} \tag{15.30}$$

and with $M = \frac{4\pi}{3}R^3\rho$ looks like:

$$-\frac{4\pi}{3}GR\rho m \tag{15.31}$$

so putting this into Newton's law:

$$m\ddot{R} = -\frac{4\pi}{3}GR\rho m \tag{15.32}$$

and so (even with this incorrect, but physically intuitive model) we recover the equation from the Einstein equations:

$$\frac{\ddot{R}}{R} = -\frac{4\pi}{3}G\rho \tag{15.33}$$

Note that this derivation does not capture the pressure term in the equation, as this is a manifestly relativistic term.

15.5 Solutions to the Friedmann equations

Let's study some solutions to the Friedmann equations.

15.5.1 Matter-dominated universe

Take $k = 0$ (flat) and non-relativistic ($P = 0$), in which case the conservation equation becomes:

$$\dot{\rho} + 3\frac{\dot{R}}{R}\rho = 0 \implies \rho \sim \frac{1}{R^3} \tag{15.34}$$

And the first Friedmann equation looks like:

$$\left(\frac{\dot{R}}{R}\right)^2 \sim \frac{1}{R^3} \tag{15.35}$$

and so $R^{1/2}\dot{R}\tilde{}1$, which implies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(R^{3/2}\right) = 1 \tag{15.36}$$

Thus we find:

$$R(t) \sim t^{2/3}. \tag{15.37}$$

So the Hubble constant looks like:

$$H(t) = \frac{\dot{R}}{R} = \frac{2}{3t} \tag{15.38}$$

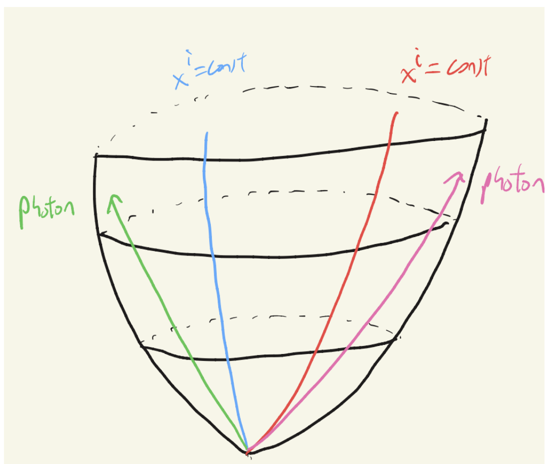
and so the age of the universe and hubble constant in such a universe are related by:

$$t_0 = \frac{2}{3H_0} \sim 10^{10}\text{years} \tag{15.39}$$

wherein H_0 is measured via experiment. But this is contradicted by other findings, namely we have found astronomical objects whose age surpasses this time.

15.5.2 Cosmic Particle Horizon

Consider the following picture of the universe, where we consider observers on x^i = const. trajectories. In such a universe, observers get further apart in time due to expansion. The notion of the cosmic particle horizon is obtained via the following - supposing the observers released a photon at the very beginning of the universe, what is the horizon/distance to which those photons can reach?



This problem is conveniently analyzed in terms of the conformal time. Consider a signal emmited at time $t_{in} \to \eta_{in}$. Then the distance reached looks like:

$$|\mathbf{x}| = \eta - \eta_0 = \int_{t_{in}}^{t_0} \frac{dt}{R(t)} \tag{15.40}$$

Then (rescaling to get the actual distance by including a factor of $R(t_0)$ and taking $t_{in} = 0$):

$$l_{\text{p. horizon}} = R(t_0) \int_0^{t_0} \frac{dt}{R(t)} \tag{15.41}$$

In the matter dominated universe, we know the form of $R(t)$ exactly, so let us evaluate:

$$l_{\text{p. horizon}} = t_0^{2/3} \int_0^{t_0} \frac{dt}{t^{2/3}} = t_0^{2/3} 3t^{1/3}\Big|_0^{t_0} = 3t_0 = \frac{2}{H_0} \tag{15.42}$$

15.5.3 Radiation-dominated universe

Now, let us assume a flat $(k = 0)$ universe so filled with radiation with $T^{\mu}_{\ \mu} = 0$ i.e. $P = \frac{1}{3}\rho$. This could physically correspond to a universe filled with photons, or relativstic neutrinos, etc.

Then, the conservation equation becomes:

$$\dot{\rho} + 3\frac{R}{R}\left(\underbrace{\rho + P}_{\frac{4}{3}\rho}\right) \implies \dot{\rho} + 4\frac{\dot{R}}{R}\rho = 0 \implies \rho \sim \frac{1}{R^4} \tag{15.43}$$

and then from the Friedmann equation we get:

$$\dot{R}R \sim R^2 \implies R \sim t^{1/2} \tag{15.44}$$

and if we compute the cosmic particle horizon for this universe we find:

$$l_{\text{p. horizon}} = \frac{1}{H_0} \tag{15.45}$$

15.5.4 Vacuum-dominated universe

In this universe the stress-energy tensor looks like:

$$T_{\mu\nu} = -\Gamma g_{\mu\nu} \tag{15.46}$$

which implies that $p + P = 0$ and so $p = -P$. Hence the conservation equation reads:

$$\dot{\rho} + \frac{\dot{R}}{R}(\rho + P) = 0 \implies \dot{\rho} = 0 \implies \rho = \text{const.} \tag{15.47}$$

so then from the first Friedmann equation we get:

$$\dot{R}R = H^2 = \text{const.} \implies R(t) \sim e^{Ht} \tag{15.48}$$

so the universe expands exponentially. Such a universe has no beginning/point where it had zero size (as even in the $t \to -\infty$ limit $R(t)$ is still finite). Such a universe is known as "de Sitter".

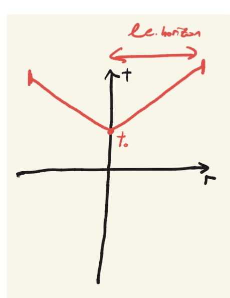
What does the particle horizon look like in this universe?

$$l_{\text{p. horizon}} = R(t_0)\int_{-\infty}^{t_0} = e^{Ht_0}\int_{-\infty}^{t_0} dt e^{-Ht_0} = \infty \tag{15.49}$$

i.e. there is no horizon, so a signal released at $t = -\infty$ can reach anyone in the universe.

15.5.5 Cosmic Event Horizon

There is a different notion of a horizon, known as the cosmic event horizon. This is defined as follows - assuming that all observers in the universe stay at rest - we send a signal, and ask what observers can this signal reach.



This is in some sense the inverse of the particle horizon problem. If I release a photon at time $t = t_0$ and consider where it could be in the future at $t = t$, we have the sphere of the photon looking like:

$$|\mathbf{x}| = \eta - \eta_0 = \int_{t_0}^{t} \frac{dt'}{R(t')} \tag{15.50}$$

Since we are interested in the distance as measured now, we scale by $R(t_0) = R_0$ and consider the reachable sphere as we take $t \to \infty$:

$$l_{e.\text{horizon}} = R_0|\mathbf{x}| = R_0 \int_{t_0}^{\infty} \frac{dt'}{R(t')} \tag{15.51}$$

for the de Sitter/vacuum dominated universe, we find:

$$l_{e.\text{horizon}} = e^{Ht_0} \int_{t_0}^{\infty} \frac{dt'}{e^{Ht'}} = \frac{1}{H_0} \tag{15.52}$$

Next time (last lecture!) We discuss more sophisticated solutions to the Einstein equations, and also discuss problems with the big bang model.

16 GR for the Universe III

Let us recall what we learned last time - we studied the Friedmann equations:

$$H^2 = \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{R^2} \tag{16.1}$$

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}(\rho + 3P) \tag{16.2}$$

Note that the second equation follows from the first equation + the conservation law:

$$\dot{\rho} + \frac{3\dot{R}}{R}(\rho + P) = 0. \tag{16.3}$$

We discussed 3 examples of universes last time - a matter-dominated universe, a radiation-dominated universe, and a vacuum-dominated universe. Let's consider a more general case.

16.1 General Single-Component FLDW universes

Consider the slightly more general case:

$$P = w\rho \tag{16.4}$$

which includes all of the prior cases, i.e. with $w = 0$ we have NR matter, with $w = \frac{1}{3}$ we have radiation, and $w = -1$ we have vacuum energy. This is a generally popular parametrization used in cosmology for the equation state of matter. E.g. discussing what the w value for dark energy is...

In this case, the continuity equation is simple:

$$\dot{\rho} + \frac{3\dot{R}}{R}(1 + w)\rho = 0 \tag{16.5}$$

So then:

$$\rho \sim \frac{1}{R^{3(1+w)}} \tag{16.6}$$

where if $w = 0$ we have $\rho \sim R^{-3}$, if $w = 1/3$ we have $\rho sim R^{-4}$, and if $w = -1$ we have $\rho \sim 1$. If we assume $k = 0$ (flat) and solve the first equation, we get:

$$\frac{\dot{R}}{R} \sim \frac{1}{R^{3(1+w)/2}} \implies R \sim t^{\alpha} \tag{16.7}$$

where:

$$\alpha = \frac{2}{3}\frac{1}{1 + w} \tag{16.8}$$

where if $w = 0$ we have $R \sim t^{2/3}$, if $w = 1/3$ we have $R \sim t^{1/2}$ and $w = -1$ is a special case (we cannot literally take the formula) where we find that $R \sim e^{Ht}$.

Suppose I had $w = -1/3$. Then, $\ddot{R} = 0$ so the universe expands linearly, with $R \sim t$. Last time we discussed a Newtonian model for the universe as a blob of expanding matter, allowing us to derive the second Friedmann equation (without the P term). If I push that model beyond the regime where it is physically relevant - if I take $w = -1/3$ then there is no gravitational force that acts on a galaxy moving from the center, i.e. the gravity from ρ is balanced by the negative pressure/gravitational repulsion coming from P.

16.2 Recollapse

Consider again the first Friedmann equation:

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi}{3}G\rho - \frac{k}{R^2} \tag{16.9}$$

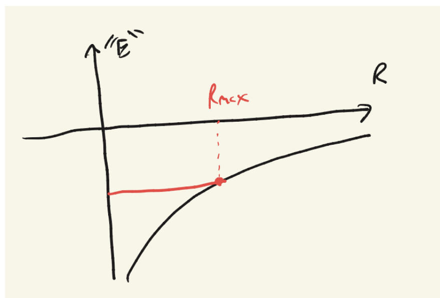
Suppose the universe is filled with NR matter so $\rho \sim \frac{1}{R^3}$. Instead of $k = 0$, let's take $k = +1$ (the universe is closed/is S^3). Then let us write:

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{R_{\text{max}}}{R^3} - \frac{1}{R^2} \tag{16.10}$$

And further rewrite:

$$\dot{R}^2 - \frac{R_{\text{max}}}{R} = -1 \tag{16.11}$$

Which we can interpret as a particle with energy:



And we will see that the universe in this model will expand to an R_{max} and then collapse back.

To make this statement precise, let us solve these equations by going into conformal time:

$$d\eta = \frac{dt}{R(t)} \implies \frac{d}{dt} = \frac{1}{R}\frac{d}{d\eta} \tag{16.12}$$

The equation to solve then becomes:

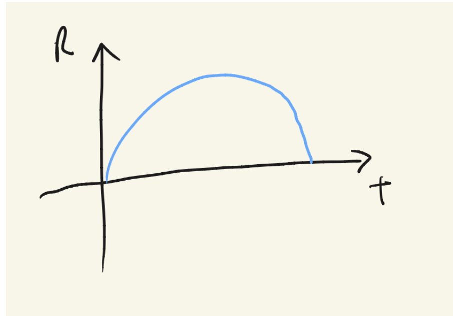
$$\frac{1}{R^2}\left(\frac{\partial R}{\partial \eta}\right)^2 = \frac{R_{\text{max}}}{R} - 1 \implies \frac{\partial R}{\partial \eta} = \pm\sqrt{R_{\text{max}}R - R^2} \tag{16.13}$$

Wherein we find:

$$R(\eta) = \frac{R_{\text{max}}}{2}(1 - \cos\eta) \tag{16.14}$$

$$t(\eta) - = \frac{R_{\text{max}}}{2}(\eta - \sin\eta) \tag{16.15}$$

$R(t)$ cannot be written analytically, but we do have a parameteric dependence, which allows us to numerically solve/sketch. As $\eta \to 0$, $R \sim \eta^2$ and $t \sim \eta^3$, and we can fill out the rest to find:



Note that this is exactly the curve that a point on a rolling disc traces out!

How can we connect this to physics? we can measure ρ, H via observation. We can then see whether ρ is larger than the critical density:

$$\rho \overset{?}{>} \rho_{\text{crit}} = \frac{3}{8\pi G} H_0^2 \tag{16.16}$$

If ρ is larger than criticality, then we know that $k = 1$ and so we must have this solution. If $\rho = \rho_{\text{crit}}$ then we conclude that $k = 0$/flat and if $\rho < \rho_{\text{crit}}$ then we can deduce $k = -1$. Of course we cannot measure the density of the universe to perfect accuracy, so we can only be certain it is flat to a certain degree of confidence.

16.3 Current composition of the universe

If we study the Friedmann equation as it applies to our universe, we have:

$$H^2 = \frac{8\pi G}{3} \left(\rho_m + \rho_{\text{rad}} + \rho_\Lambda + \rho_{\text{curv}}\right) \tag{16.17}$$

with ρ_m corresponding to matter, both baryonic matter and dark matter - it has $w = 1$. ρ_{rad} is radiation and we have $w = 1/3$. ρ_Λ is the cosmological constant term, which we posit to have $w = -1$. We also add a cheat term ρ_{curv} with $w = -1/3$ (just lumping in the k/R^2 term)- there is no such matter associated with this curvature of the spacetime, but this term is 0 if the universe is flat, < 0 if $k = 1$ and > 0 if $k = -1$.

The critical density is:

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G} \tag{16.18}$$

Which is satisfied if:

$$\rho_m + \rho_{\text{rad}} + \rho_\Lambda + \rho_{\text{curv}} = \rho_{\text{crit}} \tag{16.19}$$

which dividing out by ρ_{crit} on both sides and defining $\Omega_i = \frac{\rho_i}{\rho_{\text{crit}}}$:

$$\Omega_m + \Omega_{\text{rad}} + \Omega_\Lambda + \Omega_{\text{curv}} = 1 \tag{16.20}$$

it seems to be the case that:

$$\Omega_{\text{rad}} \leq 10^{-4} \tag{16.21}$$

$$\Omega_m \sim 0.3 \tag{16.22}$$

$$\Omega_\Lambda \sim 0.7 \tag{16.23}$$

$$|\Omega_{\text{curv}}| < 0.02 \tag{16.24}$$

It appears that in our universe, the cosmological constant term dominates. I could try to imagine how the relative weight of the Ωs change with time. It is not always the case that $\Omega_m < \Omega_\Lambda$ - the reason is that the different components of the universe have different dependences on time/current universe size:

$$\rho_m \sim R^{-3} \tag{16.25}$$

$$\rho_{\text{rad}} \sim R^{-4} \tag{16.26}$$

$$\rho_\Lambda \sim R^0 \tag{16.27}$$

$$\rho_{\text{curv}} \sim R^{-2} \tag{16.28}$$

Let us introduce another standard astronomical parametrization:

$$\frac{R_0}{R(t)} = 1 + z(t) \tag{16.29}$$

where $z(t)$ is the redshift - the factor that the frequency of light is reduced when the light comes from a source to us. Instead of saying "the source came from N years ago" astronomers say "this light came from a source at redshift z". In this parameterization, note:

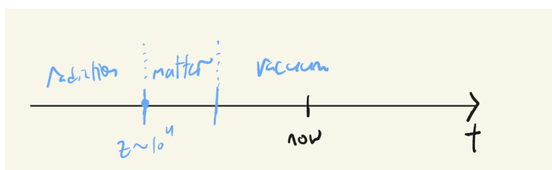
$$\rho_m \sim (1+z)^3 \tag{16.30}$$

$$\rho_{\text{rad}} \sim (1+z)^4 \tag{16.31}$$

$$\rho_\Lambda \sim 1 \tag{16.32}$$

$$\rho_{\text{curv}} \sim (1+z)^2 \tag{16.33}$$

We can see that if we increase z, then (not so long ago) $\rho_m > \rho_\Lambda$ because ρ_m has the cubic dependence. If we look at the timeline of components that dominated the universe:



16.4 The Age of the Universe

Let us consider a simple universe with only $\rho_m \neq 0, \rho_\Lambda \neq 0$, and we will determine the age of the universe using $H_0, \Omega_\Lambda, \Omega_m$. The first Friedmann equation reads:

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\left(\rho_m^0 \frac{R_0^3}{R^3(t)} + \rho_\Lambda\right) = \frac{8\pi G}{3}\rho_{\text{crit}}\left(\Omega_m \frac{R_0^3}{R^3} + \Omega_\Lambda\right) \tag{16.34}$$

So then:

$$\left(\frac{\dot{R}}{R}\right)^2 = H_0^2\left[\Omega_m \left(\frac{R_0}{R}\right)^3 + \Omega_\Lambda\right] \tag{16.35}$$

which we can solve via separation of variables, or a simple way is to use mathematica. We then obtain:

$$R(t) = R_0 \left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{1/3} \left[\sinh(\frac{3}{2}\sqrt{\Omega_\Lambda} H_0 t)\right]^{2/3} \tag{16.36}$$

Notice that as $t \to \infty$:

$$R(t) \sim e^{\sqrt{\Omega_\Lambda} H_0 t} \tag{16.37}$$

which makes sense - in the far future the universe is dominated by a single (vacuum) component, and so sees exponential expansion rate.

The age of the universe t_0 is when $R(t_0) = R_0$, so we get:

$$t_0 = \frac{1}{H_0} \frac{2}{3\sqrt{\Omega_\Lambda}} \operatorname{arcsinh} \sqrt{\frac{\Omega_\Lambda}{\Omega_m}} \tag{16.38}$$

note that in the limit where $\Omega_m \to 1, \Omega_\Lambda \to 0$ that this becomes $t_0 = \frac{2}{3H_0}$, as we derived last lecture. With the current known values of H_0, this appears to be too short. When we put in $\Omega_\Lambda \sim 0.7$ and $\Omega_m \sim 0.3$ we find that:

$$t_0 \sim \frac{1}{H_0} \tag{16.39}$$

which is 50% larger than the matter dominated model. So the presence of the vacuum energy of the universe increased its age.

16.5 Measuring densities

How do we measure densities like Ω_m, Ω_Λ? We look at the relationship between apparent luminosity and redshift. We study "standard candles" which should emit the same amount of energy in the same unit time. We can then measure the apparent luminosity/flux and compare this to redshift.

Assuming $k = -1$ and a homogenous/isotropic universe, we have the metric:

$$ds^2 = -dt^2 + R(t)^2 \left[d\chi^2 + \sinh^2 \chi (d\theta^2 + \sin^2 \theta d\phi^2) \right] \tag{16.40}$$

Take some time t_1 ($\chi = 0$) as the time that a source emitted light. The radius of the sphere (that the light can reach) is then:

$$\chi = \int_{t_1}^{t_0} \frac{dt}{R(t)} \tag{16.41}$$

We want to rewrite this as an integral over z. Recall that:

$$z = \frac{R_0}{R(t)} - 1 \tag{16.42}$$

Then:

$$ds = -\frac{R_0 \dot{R}}{R^2(t)} dt \implies dt = \frac{R^2}{R_0 \dot{R}} dz \tag{16.43}$$

the integral becomes:

$$\chi = \int_0^z dz' \frac{R}{R_0 \dot{R}} \tag{16.44}$$

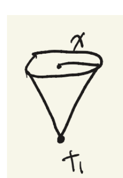
Looking back at the Friemann equaion:

$$\left(\dot{R} R \right)^2 = H_0 (\Omega_m (1+z)^3 + \Omega_\Lambda + \Omega_{\text{curv}} (1+z)^2) \tag{16.45}$$

so then:

$$\chi = \frac{1}{R_0 H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_m (1+z')^3 + \Omega_\Lambda + \Omega_{\text{curv}} (1+z')^2}} \tag{16.46}$$

which tells us the size/radius of the sphere of radiated light.



Now, measuring the energy flux, we have:

$$J \sim \frac{L}{S(z)} \tag{16.47}$$

with:

$$S(z) = 4\pi R_0^2 \sinh^2 \chi(z) \tag{16.48}$$

with $\chi(z)$ determined via the integral.

To be precise, there is also a prefactor:

$$J \sim \frac{L}{S(z)} \frac{1}{(1+z)} \tag{16.49}$$

because the photons that travel to us are redshifted by factor $1+z$. Actually, there is another factor! The rate of photon emission is also hit by a factor $1+z$, so the fully correct expression is:

$$J = \frac{L}{S(z)} \frac{1}{(1+z)^2} \tag{16.50}$$

So, if we know the luminosity and z, then we can compare the measurement of the energy flux J, we can compare it with the above formula for $S(z)$ to solve for the Ωs.

For simplicity, let us consider the case where $z \ll 1$. Then:

$$\chi \approx \frac{1}{R_0 H_0} \int_0^z dz' \frac{1}{\sqrt{1 + 3(\Omega_m + 2\Omega_{\text{curv}})z}} \approx \frac{1}{R_0 H_0} (z - \frac{1}{4}(3\Omega_m + 2\Omega_{\text{curv}})z^2 + O(z^3)) \tag{16.51}$$

so then with:

$$\sinh \chi = \chi + O(\chi^3) \tag{16.52}$$

we find:

$$S(z) = \frac{4\pi}{H_0^2}(z - \frac{1}{4}(3\Omega_m + 2\Omega_{\text{curv}})z^2)^2 \tag{16.53}$$

so to leading order, the size of the sphere grows as $\sim z$. There is a correction at quadratic order, and if we can measure it very accurately, we can deduce the densities $\Omega_m, \Omega_{\text{curv}}$, and further we can look at higher orders of z we get further constraints... This is what Son would do as a theorist (in practice they just fit all orders of z at once, probably).

This is roughly speaking what the model of the universe is right now, forming the standard "Big Bang" model.

16.6 Problems with the Big Bang Model

We end our discussion with some weirdness/problems with the Big Bang model.

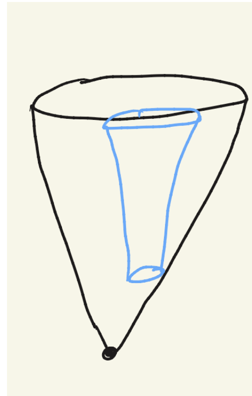
1. *The Horizon problem.* What we see right now is a universe that has a finite particle horizon, which we recall from the last lecture is a patch of the universe that may have come from the same source. It scales as $R_H \sim \frac{1}{H}$. In this region, events can be causally connected. Outside of it, there must be causal disconnection/events must originate from different patches of the early universe. In the past, the horizon size looks like (assuming $R(t) \sim t^\alpha$, with $\alpha = 2/3$ for matter dominated, $\alpha = 1/2$ for radiation dominated):

$$R_H(t_1) = \frac{1}{H(t_1)} \sim t_1 \sim R^{1/\alpha}(t_1) = \frac{1}{(1+z_1)^{1/\alpha}} \frac{1}{H_0} \tag{16.54}$$

with $R = \frac{R_0}{1+z}$. Then, the scale factor takes this region to a new region of size:

$$R_H(t_1) \cdot (1+z_1) \sim \frac{1}{(1+z_1)^{\frac{1}{\alpha}-1}} \frac{1}{H_0} \tag{16.55}$$

Since $\alpha < 1$, this size is much less than R_H^{now}.



What is the meaning of this? The early universe consists of many patches that could not have any causal connection to each other. The question then becomes - why did these patches need to know that they needed to have the same/uniform temperature (given the uniformity of temperature of universe today) even though there is no causal connection between these regions?

2. *The Flatness problem.* When we write:

$$\rho = \rho_m + \rho_\Lambda + \rho_{\text{rad}} + \rho_{\text{curv}} \tag{16.56}$$

the curvature term has $|\Omega_{\text{curv}}| < 0.02$. The surprising thing happens when we extrapolate back in the past - we said that $\rho_m \sim (1+z)^3$, $\rho_{\text{rad}} \sim (1+z)^4$, and $\rho_{\text{curv}} \sim (1+z)^2$. So the relative weight of the curvature term gets smaller and smaller at larger z. The big bang model appears to work well up to $z \sim 10^8 - 10^9$, where we had big bang nucleosynthesis. At this point in the universe, the radiation wins out by a large factor: $\rho_{\text{curv}}/\rho_{\text{rad}} \sim (10^8)^{-2} \cdot 10^2 \sim 10^{-13}$. So the universe at this time was *incredibly* close to being flat, as we ascertain by the current curvature of the universe - essentially there appears to be the fact that in order to see the degree of flatness that we see today, the universe would have needed to be extremely uniform/fine tuned at early times.

Note that this isn't a self-consistently problem. It is solved by the theory of cosmic inflation (occuring right after the big bang), where any nonzero curvature of the universe was stretched out to be negligible. This also is a solution of the horizon problem, where the exponential expansion of space due to inflation allows for distant regions (which shouldn't naively have causal connection) of the universe to appear to uniform in temperature.