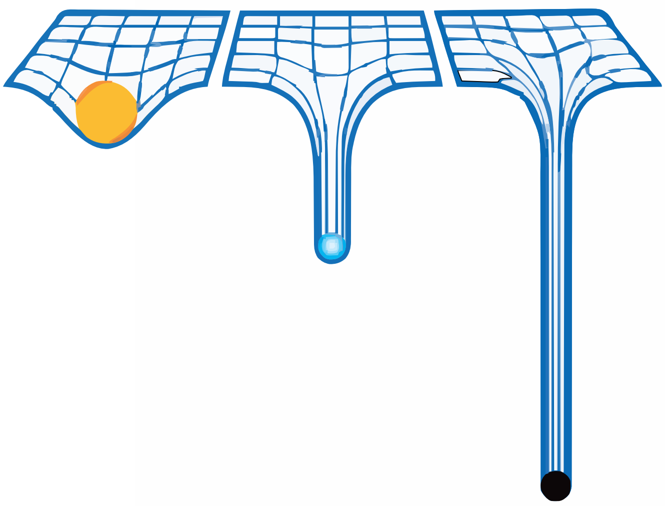
General Theory Of Relativity and Cosmology

Lecture Notes

Instructor: Prof. Golam Mortuza Hossain



Sagnik Seth

Dept of Physical Sciences
IISER Kolkata

Contents

Lecture 01: Introduction

In the physical sense, the word 'relativity' can be used to relate the measurement of one observer with that of another, for the same object/quantity. So using relativity, we can characterise the difference in the observations of two observers in two different frame of reference.

Note that, physical laws should not matter based on who is observing (that is, physical laws are *universal*), hence relativity becomes very important when dealing with the resolution of conflicts.

Let us consider two reference frames, and let v be the relative velocity between the frames.

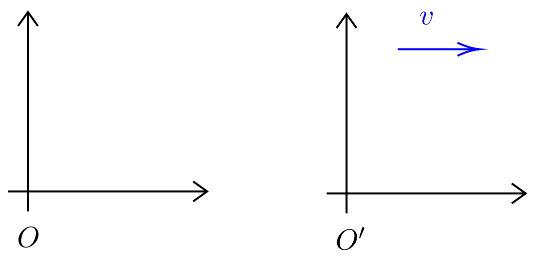
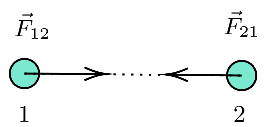


Figure 1: Two frames of reference with a relative velocity v

Fig. 1 shows two reference frames. The moving frame is denoted by O' (henceforth, primes will generally denote moving frames). If the relative velocity v between the two frames remain constant, then it falls under the domain of *Special Relativity* and if unfortunately (or fortunately) it doesn't, then it comes under *General Relativity*.

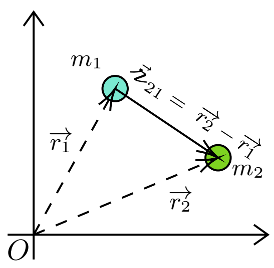
The theory of relativity dates back to the inception of Newtonian mechanics and hence, let us review the laws in brief (without being much technical 😅):

- **The First Law** :≈ velocity of a body remains constant unless acted by an external, unbalanced force.
- **The Second Law** :≈ $\vec{F} = m\vec{a}$ where $\vec{F}$ is the external force and $\vec{a}$ is the acceleration.
- **The Third Law** :≈ Every action has an equal and opposite reaction. Note that for this, two bodies are needed.



From the figure above, we will have $\vec{F}_{21} = -\vec{F}_{12}$

- **The Gravitational Law** :≈



In the above situation, if we consider for mass m_1, we have:

$$\vec{F}_{12} = G\frac{m_1 m_2}{\mathscr{r}_{21}^2}\hat{\mathscr{r}}_{21}$$

An important observation: in the second law, if we take the force to be zero, then apparently $\vec{a} = \frac{\mathrm{d}v}{\mathrm{d}t} = 0 \implies \vec{v} = \mathrm{const.}$ which is the statement of the first law. However, the second law does not imply the first law, since the first law defines what an *inertial frame* is and the statement of the second law applies only in case of inertial frames. So, a better formulation of the second and third law can be like, "In a frame where the first law is valid, blah blah blah..."

Lecture 02: Galilean Stuffs

In the previous lecture, we saw how Newton's laws define an inertial frame and how we can rephrase the second and third law in terms of inertial frames, to avoid confusion. It happens that, there are reference frames where the first law doesn't hold true.

Imagine a person inside a darkened car, isolated from the outside world, with a ball hanging from the roof of the car. The car suddenly starts accelerating and (as intuition speaks) the ball starts moving. The person claims "The ball has moved!" 💀.

Note that from the perspective of the person, no force has been applied to the ball, yet it moved, which is a direct violation of the second law. Thus, we can call this frame *non-inertial*.

Now, consider a person inside an ill-fated lift, tragically falling 🥲. Everything in the lift frame is accelerating together under gravity and hence 'weightlessness' (free-fall) occurs. If somehow a coin is also there in the lift, it simply floats. And if the person pushes the coin, it moves with an almost constant velocity. Thus, it is a very good approximation to an *inertial frame* (although, extremely tragic for the person!).

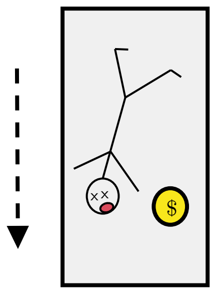


Figure 2: A sad guy floating with a coin in a falling lift

We had always wanted the physical laws to stay the same. This is actually the statement of Galilean relativity:

"Laws of physics (nature) must be the same in all inertial frames of reference"

By physical laws, we imply Newton's laws of mechanics and the law of gravitation. To have a more mathematical (and less philosophical) aspect to the above statement, we define the *Galilean transformation*.

2.1. Galilean Transformation

Let an onject be at point P and consider two reference frames O and O', which coincided at $t = 0$ but have moved apart since, along the common x-axis. (Henceforth, reference frames will simply be termed *observers*).

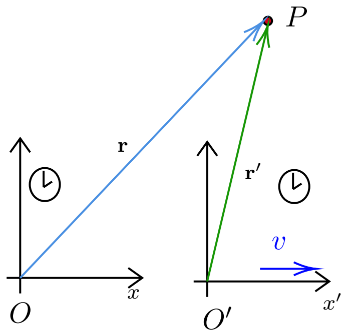


Figure 3: Observing same point from two different frames

Now, the distance between the origins of two frames, at time t, will be $R = vt$. Then accordingly, we can write the relation between the coordinates of the two observers as follows:

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned} \qquad \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

An important assumption: The Galilean transformation assumes a *universal time*, that is, time elapsed on clocks in both frames are the same [1].

Now, it is not necessary for the observers to move along the x axis only. Hence, generalising the transformation to arbitrary direction, we obtain:

$$t' = t \qquad \mathbf{r}' = \mathbf{r} - \mathbf{v}t$$

Now, let us look how the physical laws react under the Galilean transformation.

- **Law of Gravitation:**

  Consider the situation in the diagram below:

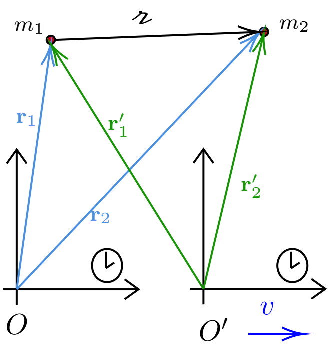


Figure 4: Gravitational law from two different frames

According to Galilean transformation, we have:

$$\begin{aligned} \boldsymbol{\mathscr{r}} &= \mathbf{r}_2 - \mathbf{r}_1 \\ &= (\mathbf{r}'_2 + \mathbf{v}t) - (\mathbf{r}'_1 + \mathbf{v}t) \\ &= \mathbf{r}'_2 - \mathbf{r}'_1 \\ &= \boldsymbol{\mathscr{r}}' \end{aligned}$$

[1] This was perhaps a fair assumption since time, atleast in the philosophical aspect, has always seemed to be *superior* and all-encompassing, second to none.

Thus, we see that the vector $\boldsymbol{\mathscr{r}}$ is unchanged and since this is the only vector appearing in Newton's law of gravitation, we have:

$$\mathbf{F}_{12} = G\frac{m_1 m_2}{\mathscr{r}^2}\ \hat{\boldsymbol{\mathscr{r}}} = G\frac{m_1 m_2}{\mathscr{r}'^2}\ \hat{\boldsymbol{\mathscr{r}}}' = \mathbf{F}'_{12}$$

We see that the expression of the force does not change between frames, thus indicating a *universality*.

- **Second Law:**

  Using the second law, we can write

  $$\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$$

  Now, using Galilean transformation, we obtain:

  $$\mathbf{r}(t) = \mathbf{r}'(t') + \mathbf{v}t \implies \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t} \times \frac{\mathrm{d}t'}{\mathrm{d}t} + \mathbf{v} \implies m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = m\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t^2}$$

  Thus, we see that in general, force will be the same for observers in different inertial frames. Note that we assumed the *time universality*, that is, $\frac{\mathrm{d}t'}{\mathrm{d}t} = 1$

- **Third Law:**

  Note that since the force expression for gravity is symmetric in m_1 and m_2, the law of gravitation automatically incorporates the third law. Newton knew only about gravity and perhaps, third law would not have been necessary if he had relied only on gravity, however, he postulated that for all other forces, third law holds.

2.2. A Big Blow to Galileo

Let us consider that in Fig. 3, point P is also moving with some velocity $\mathbf{u}'$ and $\mathbf{u}$ in primed and unprimed frame respectively. Then, we can write:

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{r}' + \mathbf{v}t\big) = \frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t} + \mathbf{v} = \mathbf{u}' + \mathbf{v} \implies \mathbf{u} = \mathbf{u}' + \mathbf{v} \qquad : \text{velocity addition formula}$$

However, observations, specially the Michelson-Morley experiment, were made in different frames but found the speed of light (henceforth denoted as c) to be a constant frame-independent quantity, which contradicted the velocity addition formula when applied to speed of light.

Since the velocity addition formula depended entirely on the Galilean transformation, to resolve this conflict, we have to change the transformation rule in its entirety.

Lecture 03: Lorentz Transformation

We saw previously that Galilean relativity failed terribly when speed of light was considered. This needed a change in the transformation rule. The new transformation rule is called the *Lorentz transformation* and is given by (for two frames relatively moving along the common x axis):

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma\Big(t - \frac{vx}{c^2}\Big) \end{aligned} \qquad \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

where $\gamma := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ is called the *Lorentz factor*. Using this transformation, we can derive a new rule for vector addition:

$$u' = \frac{dx'}{dt'} = \frac{dx'/dt}{dt'/dt} = \frac{\gamma(dx/dt - v)}{\gamma\left(1 - \frac{v}{c^2} dx/dt\right)} = \frac{u - v}{1 - u\frac{v}{c^2}}$$

If we use the inverse Lorentz transformation, we will obtain:

$$u = \frac{u' + v}{1 + u\frac{v}{c^2}}$$

Now, if we put $u' = c$, then we have

$$u = \frac{c + v}{1 + \frac{cv}{c^2}} = c = u'$$

Hence the conflict is successfully resolved. However, in changing the rule of transformation, we had put a great risk to the already established laws, since their invariance is now not guaranted.

3.1. What happens to physical laws?

- Since there is mixing between x and t in the transformation, it is clear that:

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} \neq \frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t'^2}$$

  Thus, the second law is out of the picture! 💀

- Imagine two balls falling, however the distance between them does not remain same. Then, at some instant, one of the ball conveys its position to the other ball and since the information takes a finite amount of time to reach the other balls, by that time, the first ball has fallen by by some distance and hence the forces $\mathbf{F}_{12}$ and $\mathbf{F}_{21}$ are not equal and opposite. Thus, the third law is out of the picture too! 😵

- Since in general, $r'^2 = (x_2' - x_1')^2 + (y_2' - y_1')^2 + (z_2' - z_1')^2 \neq r^2$, the law of gravitation also falls apart 🤯!

Hence, all the established laws were destroyed by the mere change of a transformation. We desperately needed another theory to reconcile with the physical laws. This led to the formulation of special and general relativity.

It is convenient to consider *infinitesimal transformation*, like:

$$\mathbf{r}_2 = \mathbf{r}_1 + d\mathbf{r} \implies (\mathbf{r}_2 - \mathbf{r}_1)^2 = d\mathbf{r} \cdot d\mathbf{r} = dr^2$$

Note that under Galilean transformation, dr^2 remain unchanged (invariant), however, under Lorentz transformation, it no longer remains constant. So, what is the invariant quantity under Lorentz transformation?

It turns out that if we define something like:

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

then, ds^2 is invariant under Lorentz transformation.

Lemma 1 (Invariance of ds):

$$ds'^2 = ds^2$$

Proof. We use brute force to calculate the quantitites:

$$\begin{aligned}
ds'^2 &= -c^2 dt'^2 + dx'^2 + dy'^2 + dz'^2 \\
&= -c^2\gamma^2\left(dt - \frac{vdx}{c^2}\right)^2 + \gamma^2(dx - vdt)^2 + dy^2 + dz^2 \\
&= -\gamma^2 c^2\left(dt^2 + \frac{v^2dx^2}{c^4} - \cancel{2\frac{dxdtv}{c^2}}\right) + \gamma^2(dx^2 + v^2dt^2 - \cancel{2vdxdt}) + dy^2 + dz^2 \\
&= -\gamma^2c^2dt^2 - \gamma^2v^2\frac{dx^2}{c^2} + \gamma^2dx^2 + \gamma^2v^2dt^2 + dy^2 + dz^2 \\
&= -c^2\cancel{\gamma^2}dt^2\cancel{(1 - v^2/c^2)} + \cancel{\gamma^2}dx^2\cancel{(1 - v^2/c^2)} + dy^2 + dz^2 \\
&= -c^2dt^2 + dx^2 + dy^2 + dz^2 \\
&= ds^2
\end{aligned}$$

3.2. Classification of Intervals

Unlike Galilean transformation where $dr^2 > 0$ always, ds^2 need not be positive. In fact, it can take any value and accordingly we define the intervals. If the interval between two spacetime events is such that:

- $c^2dt^2 > dr^2 \implies ds^2 < 0$: time-like interval
- $c^2dt^2 < dr^2 \implies ds^2 > 0$: space-like interval
- $c^2dt^2 = dr^2 \implies ds^2 = 0$: light-like/null interval

Note that for light, $\frac{\mathrm{dr}}{\mathrm{dt}} = \pm c \implies dr^2 = c^2dt^2 \implies ds^2 = 0$. Hence, we call the interval light-like. Trajectory of light is a null trajectory. However, all massive objects move along a time-like trajectory.

Lecture 04: Vectors and Stuffs

Let us consider rotation about the z-axis by an angle θ, in which case, the x and y components transform as:

$$\begin{aligned}
x' &= x\cos\theta + y\sin\theta \\
y' &= -x\sin\theta + y\cos\theta
\end{aligned}$$

Then, we can write the infinitesimal change in the components in the following form:

$$\begin{pmatrix} dx' \\ dy' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix}$$

Note that, this can also be written as:

$$\begin{pmatrix} dx' \\ dy' \end{pmatrix} = \begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix}$$

The matrix in this case is the *Jacobian matrix* and this is a more general structure of coordinate transformations and hence any transformation can be written like this, even the Galilean or the Lorentz transformations.

This shows that transformations are just multiplication of a matrix with a vector. To this extent, we define a *vector* [1]$\bar{x}$ with the following components:

$$\bar{\mathbf{x}} = (x^0, x^1, x^2, x^3)^T$$

[1]This is often called a *four* vector

Here $x^0 = ct$ and $x^1 = x, x^2 = y, x^3 = z$. Note that since t and x are not dimensionally same, we took $x^0 = ct$ to make it in terms of ditance. We multiplied by c (and not any other velocity) because c is postulated to be a universal constant.

We can express any infinitesimal transformation as:

$$\begin{pmatrix} dx'^0 \\ dx'^1 \\ dx'^2 \\ dx'^3 \end{pmatrix} = \begin{pmatrix} \frac{\partial x'^0}{\partial x^0} & \cdots & \frac{\partial x'^0}{\partial x^3} \\ \vdots & \ddots & \vdots \\ \frac{\partial x'^3}{\partial x^0} & \cdots & \frac{\partial x'^3}{\partial x^3} \end{pmatrix} \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix}$$

In a more compact way, we can write in the vector-matrix form or in the component form using Einstein's summation convention:

$$\mathbf{d\bar{x}}' = \Lambda \mathbf{d\bar{x}} \iff dx'^\mu = \Lambda^\mu{}_\nu dx^\nu$$

where $\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}$ is an element of the *matrix* Λ.

Summation convention: We drop the summation sign if twice repeated indices are there on the same side of the equation.

Now, let us define what a vector is:

Definition 1 (Vector):
If a mathematical object transforms under a coordinate transformation like a differential $\mathbf{d\bar{x}}$, that is, $\mathbf{d\bar{x}} \to \mathbf{d\bar{x}}' = \Lambda\ \mathbf{d\bar{x}}$, then it is called a *contravariant vector* or a *contravector* or a *tensor of rank* $(1,0)$

Let us show that under the Galilean transformation, velocity, as defined by $\mathbf{u} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ is a vector. We will define everything in contravariant and four-vector notation. For this, we consider the vector defined as before:

$$\mathbf{d\bar{x}} = (ct, x, y, z)^T \equiv (x^0, x^1, x^2, x^3)^T$$

The Galilean transformation is defined by:

$$\lambda \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ -{}^{v^1}\!/c & 1 & 0 & 0 \\ -{}^{v^2}\!/c & 0 & 1 & 0 \\ -{}^{v^3}\!/c & 0 & 0 & 1 \end{pmatrix} \qquad \iff \qquad \begin{aligned} x'^0 &= x^0 \\ x'^1 &= x^1 - ({}^{v^1}\!/c)x^0 \\ x'^2 &= x^2 - ({}^{v^2}\!/c)x^0 \\ x'^3 &= x^3 - ({}^{v^3}\!/c)x^0 \end{aligned}$$

We can then define the four velocity as[1]:

$$u^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}t} = (c, u^1, u^2, u^3)^T$$

Now, we have:

$$\begin{aligned} & dx'^\mu &&= \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu \\ \implies\ & u'^\mu \cancel{dt'} &&= \frac{\partial x'^\mu}{\partial x^\nu} u^i \cancel{dt} \qquad (\text{as } dt = dt') \\ \implies\ & u'^\mu &&= \frac{\partial x'^\mu}{\partial x^\nu} u^\nu \end{aligned}$$

Thus, we see that the components of the velocity vector transform similar to the differential displacement and hence, velocity is indeed a vector under Galilean transformation. In a similar way, we can show that

[1] Note that four velocity is defined as derivative with respect to something called as *proper time*, however, for Galilean transformations, proper time coincides with the coordinate time and hence this definition is okay.

acceleration $\mathbf{a} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}$ is also a vector under Galilean transformation.

Now, let us turn to Lorentz transformation. Things become bad here! For example, consider the same definition of velocity $\mathbf{u} = \frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}t}$. However, here t does not remain invariant under coordinate transformation. Then we will have:

$$u'^{\mu} = \frac{\mathrm{d}x'^{\mu}}{\mathrm{d}t'} = \frac{\Lambda^{\mu}{}_{\nu}x^{\nu}}{\Lambda^{0}{}_{\alpha}dx^{\alpha}} \neq \frac{\mathrm{d}x^{\mu}}{\mathrm{d}t}$$

Thus, this definition fails to transform similar to a differential displacement and hence, the way it is defined presently, velocity is not a vector 😭.

Lecture 05: Some more Vectors and Stuffs

Previously, we saw how a vector was defined in terms of the transformation of the differential displacements. In here, we will consider another situation.

First, note that, using a given transformation,

$$dx'^{\mu} = \Lambda^{\mu}{}_{\nu}dx^{\nu}$$

we can define the inverse transformation,

$$dx^{\nu} = \left(\Lambda^{-1}\right)^{\nu}{}_{\mu}dx'^{\mu}$$

The above statement is valid only if the transformation matrix is invertible (in most case, transformations like translations, rotations, etc. are invertible).
Let us now see what the inverse transformation matrix looks like. For that, consider:

$$\begin{aligned} dx'^{\mu} &= \frac{\partial x'^{\mu}}{\partial x^{\nu}}dx^{\nu} \\ \implies \frac{\partial x^{\beta}}{\partial x'^{\mu}}dx'^{\mu} &= \underbrace{\frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial x'^{\mu}}{\partial x^{\nu}}}_{\delta^{\beta}{}_{\nu}}dx^{\nu} \\ \implies \qquad dx^{\beta} &= \frac{\partial x^{\beta}}{\partial x'^{\mu}}dx'^{\mu} \end{aligned}$$

Thus we see that:

$$\left(\Lambda^{-1}\right)^{\nu}{}_{\mu} \equiv \frac{\partial x^{\nu}}{\partial x'^{\mu}}$$

5.1. Gradient Vector

In the Cartesian coordinates, we all know:

$$\boldsymbol{\nabla} = \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z}$$

In line with this, we will define another quantity (often called *four gradient*) such that:

$$\nabla_{\mu} := \frac{\partial}{\partial x^{\mu}}$$

A few points on this:

1. Note that we wrote the index in the down position for ∇_{μ}. This is done since we will see that the four gradient does not follow the transformation rule of a vector, hence it is a different entity. Since in vectors we put the indices upwards, to differentiate it from a vector, we put it down.

2. In spherical or cylindrical or any other coordinates, the *ordinary* gradient is not just the partial derivatives of the coordinates. There are some obnoxious prefactors involved also. Irrespective of that, we define the four gradient as just the derivative of the coordinates ('coz later we will see that those prefactors are just some jugglery with the *metric*).

The four gradient is better denoted by ∂_μ and thus, we have:

$$\bar{\boldsymbol{\partial}} \equiv \begin{pmatrix}\partial_0 & \partial_1 & \partial_2 & \partial_3\end{pmatrix} = \begin{pmatrix}\frac{1}{c}\frac{\partial}{\partial t} & \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\end{pmatrix}$$

Under a coordinate transformation $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$, we have using chain rule:

$$\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} \implies \partial'_\mu = \left(\Lambda^{-1}\right)^\nu{}_\mu \frac{\partial}{\partial x^\nu}$$

The four gradient apparently transform accordingly as $\bar{\boldsymbol{\partial}}' = \Lambda^{-1}\boldsymbol{\partial}$ which is contrary to that of the vector.

Definition 2:
Covector If a mathematical object transforms under a coordinate transformation like the derivative $\bar{\boldsymbol{\partial}}$, then we call it a *covariant vector* or *covector* or a *tensor of rank (0,1)*

Actually vectors and covectors are defined based on how the basis vectors of the space transform. Vectors transform in a way *contrary* to the basis vectors and hence are called *contravariant* while covectors transform in a way similar to how the basis transforms and hence called *covariant*.[1]

5.2. Tensors

Consider the product,

$$\boldsymbol{\omega} = \bar{\mathbf{x}}\bar{\mathbf{x}}^T \equiv \begin{pmatrix}dx^0\\dx^1\\dx^2\\dx^3\end{pmatrix}\begin{pmatrix}dx^0 & dx^1 & dx^2 & dx^3\end{pmatrix} = \begin{pmatrix}dx^0dx^0 & \dots & dx^0dx^3\\ \vdots & \ddots & \vdots\\ dx^3dx^0 & \dots & dx^3dx^3\end{pmatrix}$$

Note that, defined like this, $\boldsymbol{\omega}$ represents a square matrix and thus, has two indices. This operation is often called *outer product* and is denoted as

$$\boldsymbol{\omega} = \bar{\mathbf{x}} \otimes \bar{\mathbf{x}} \qquad \omega^{\mu\nu} = dx^i dx^j$$

Now, let us see how it transforms:

$$\begin{aligned}\omega^{\mu\nu} &= dx'^\mu dx'^\nu\\ &= \left(\frac{\partial x'^\mu}{\partial x^\alpha}dx^\alpha\right)\left(\frac{\partial x'^\nu}{\partial x^\beta}dx^\beta\right)\\ &= \left(\frac{\partial x'^\mu}{\partial x^\alpha}\right)\left(\frac{\partial x'^\nu}{\partial x^\beta}\right)dx^\alpha dx^\beta\end{aligned}$$

This is yet another kind of transformation, different from either vector and covector.

[1] In the elegant language of differential geometry, vectors are actually the *tangent vectors* belonging to the *tangent space* of the space (*manifold*) while covectors belong to the *dual space* of the *tangent space*.

Definition 3 (Tensor):
If a mathematical object transforms, under a coordinate transformation, like $\underbrace{\mathbf{d\bar{x}} \otimes \ldots \otimes \mathbf{d\bar{x}}}_{m \text{ times}}$ and like $\underbrace{\bar{\boldsymbol{\partial}} \otimes \ldots \otimes \bar{\boldsymbol{\partial}}}_{n \text{ times}}$, then it is called a *tensor of rank (m,n)*

A prime example would be the *moment of inertia tensor* I^{ij} or the *electromagnetic tensor* $F_{\mu\nu}$.
We note that tensors have upstairs indices for each contravariant components and downstairs indices for each covariant component [1].

Definition 4 (Scalar):
If a mathematical object does not transform under a coordinate transformation, then it is called a *scalar* or a *tensor of rank (0,0)*.

Few example of scalars would be the invariant displacement $dr^2 = dx^2 + dy^2 + dz^2$ under Galilean transformation and the invariant spacetime interval $ds^2 = -c^2dt^2 + dr^2$ under Lorentz transformation.

5.3. Inner Product

Previously we saw the *outer product* and we expect an *inner product* too. Multiplying a vector with the transpose changed it to a square matrix. Here, we do the opposite. Consider the scalar dr^2 under Galilean transformation, which is written as:

$$dr^2 = \mathbf{dr} \cdot \mathbf{dr} = \mathbf{dr}^{\mathrm{T}}\mathbf{dr} = \begin{pmatrix} dx^1 & dx^2 & dx^3 \end{pmatrix} \begin{pmatrix} dx^1 \\ dx^2 \\ dx^3 \end{pmatrix}$$

However, notice the apparent problem if we generalise this to Lorentz transformation too. We have:

$$ds^2 = -c^2dt^2 + dr^2$$

This cannot be simply written as $\mathbf{d\bar{x}}^{\mathrm{T}}\mathbf{d\bar{x}}$, all due to that *innocent* minus sign in front of $c^2dt^2 \equiv (dx^0)^2$ which was necessary to distinguish spatial and temporal components.

How do we generalise this notion of scalar as some matrix product?

For that, let us define another matrix which will take care of the minus sign:

$$[g] \equiv \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Using this, we can check that:

$$ds^2 = \begin{pmatrix} dx^0 & dx^1 & dx^2 & dx^3 \end{pmatrix} \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} dx^0 \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} = \mathbf{d\bar{x}}^{\mathrm{T}} \; g \; \mathbf{d\bar{x}}$$

This g is called the *metric tensor* and in component form, the above can be written as:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

[1] We can easily change upstairs to downstairs indices by the *metric* and hence it is better to write tensors as $T^{\mu}{}_{\nu}$ and not T^{μ}_{ν} so as to keep a vacant place for the indices to move up/down.

Note that, for Galilean transformation, the metric tensor is just the *identity* matrix,

$$[g]_{\mathrm{GT}} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$$

Lecture 06: Metric is also a Matrix

Previously, we saw that to get a scalar from a contra-vector, we had to use an additional matrix which led to the notion of the *metric tensor*. Let us analyse this object further...

6.1. Transformation of the metric

For this, we will consider the *sacred* rule that space-time interval should remain invariant under coordinate change. Thus, we have:

$$\begin{aligned} ds^2 &= ds'^2 \\ \implies \quad g_{\mu\nu} dx^\mu dx^\nu &= g'_{\alpha\beta} dx'^\alpha dx'^\beta \\ \implies g_{\mu\nu}\left(\frac{\partial x^\mu}{\partial x'^\alpha} dx'^\alpha\right)\left(\frac{\partial x^\nu}{\partial x'^\beta} dx'^\beta\right) &= g'_{\alpha\beta} dx'^\alpha dx'^\beta \\ \implies g_{\mu\nu}\left(\frac{\partial x^\mu}{\partial x'^\alpha}\right)\left(\frac{\partial x^\nu}{\partial x'^\beta}\right) dx'^\alpha dx'^\beta &= g'_{\alpha\beta} dx'^\alpha dx'^\beta \end{aligned}$$

And now by the eternal rule for arbitrariness (of the different displacements), we conclude that:

$$\boxed{g'_{\alpha\beta} = g_{\mu\nu}\left(\frac{\partial x^\mu}{\partial x'^\alpha}\right)\left(\frac{\partial x^\nu}{\partial x'^\beta}\right)}$$

Thus, we see that the metric tensor transforms as a rank (0,2) tensor and hence, we were not wrong about calling this a *tensor*!

Few Properties of the metric:

- Note the following :

  $$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{\nu\mu} dx^\nu dx^\mu = g_{\nu\mu} dx^\mu dx^\nu$$

  In the first step, we interchanged $\mu \leftrightarrow \nu$ since these are dummy indices and then we interchanged the placement of $dx^\mu \leftrightarrow dx^\nu$ since these are just components and hence commutes. Then, from this, we have $g_{\mu\nu} = g_{\nu\mu}$ which leads us to the fact that:

  Metric tensor is symmetric

- If we have g to be non-singular, that is, $\det(g) \neq 0$, then we can define an inverse metric such that:

  $$g g^{-1} = g^{-1} g = \mathbb{1}$$

  In the component form, we can write

  $$g_{\mu\nu}(g^{-1})^{\nu\alpha} = \delta^\alpha{}_\mu$$

  We will henceforth use $g^{\nu\alpha}$ instead of $(g^{-1})^{\nu\alpha}$ for the inverse metric, since this is more sleek visually and also, it does not confuse us.

Note that, by the same logic as we showed that metric is a tensor, we can also show that inverse metric is a rank (2,0) tensor.

6.2. Metric is a map!

The metric (and its inverse too) is a map which takes a vector (covector) and changes it to a covector (vector) according as:

$$A_\mu \overset{g}{\mapsto} A^\nu \qquad A_\mu \quad = g_{\mu\nu} A^\nu$$
$$A^\mu \xmapsto{g^{-1}} A_\nu \qquad A^\mu = g^{\mu\nu} A_\nu$$

Using this, the norm of a vector can be written as:

$$\left\|\bar{\mathbf{A}}\right\| = \bar{\mathbf{A}} \cdot \bar{\mathbf{A}} = A_\mu A^\mu = A^\mu A_\mu$$

Lecture 07: Metric and STR

Consider the following transformation:

$$(r, \theta) \mapsto (x, y) \qquad x = r\cos\theta \quad y = r\sin\theta$$

Let us calculate the transformation matrix for this transformation. Note that, here we have $\mathbf{x} \equiv (r, \theta)$ and $\mathbf{x}' \equiv (x, y)$. We calculate Λ from its definition as:

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \qquad \Lambda \equiv \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Once we have obtained the transformation matrix, we now move on to calculate the metrix tensor in polar coordinates. For that, note, for Cartesian coordinate,

$$ds^2 = dx^2 + dy^2 \implies g \equiv \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$$

Then, according to this question, we already have the knowledge of $g'_{\mu\nu}$ and we have to transform it to $g\alpha\beta$ and we already know the transformation rule. Also, since g' is diagonal, we need to only consider the diagonal components. Thus, all of our work is already done!

$$g_{\alpha\beta} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta g'_{\mu\nu}$$

- $g_{rr} = \Lambda^x{}_r \Lambda^x{}_r g'_{xx} + \Lambda^y{}_r \Lambda^y{}_r g'_{yy} = \cos^2\theta + \sin^2\theta = 1$
- $g_{\theta\theta} = \Lambda^x{}_\theta \Lambda^x{}_\theta g'_{xx} + \Lambda^y{}_\theta \Lambda^y{}_\theta g'_{yy} = r^2(\sin^2\theta + \cos^2\theta) = r^2$
- $g_{r\theta} = \Lambda^x{}_r \Lambda^x{}_\theta g'_{xx} + \Lambda^y{}_r \Lambda^y{}_\theta g'_{yy} = -r\cos\theta\sin\theta + r\sin\theta\cos\theta = 0$ (hence this metric is diagonal too!)

Thus, from this, we obtain

$$g \equiv \begin{pmatrix} 1 & \\ & r^2 \end{pmatrix}$$

From this, the invariant distance becomes our know result:

$$ds^2 = g_{ij} x^i x^j = dr^2 + r^2 (d\theta)^2$$

This is a weirdly powerful technique to find the metric in any coordinate system. And once we have found the metric, then calculations of stuff like gradient and Laplacians become very easy.

7.1. Proper Time

Time no longer remains a scalar in special theory of relativity. However, we do need some notion of invariant time which can be useful to define some quantities. At present, we only know two scalars in STR, that is, c and ds^2. Note that, dimensionally, there is only one way to construct some 'time' using these two:

$$d\tau^2 := -\frac{ds^2}{c^2}$$

We define $d\tau$ to be the *proper time* which is a scalar, since it is constructed out of two scalars. The minus sign in the definition results from the convention of the $(-,+,+,+)$ metric that we had chosen.

Consider two events $\epsilon_1 = (ct_1, x, y, z)$ and $\epsilon_2 = (ct_2, x, y, z)$, which are happening at the same spatial location. Then, since $dr^2 = 0$, for these two events,

$$ds^2 = -c^2 dt^2 \implies dt = d\tau$$

Hence the time interval between two events happening at the same spatial location is the proper time. Note that, using this notion of time (which is a scalar) we can define velocities and momentum.

7.2. Looking back at STR

Let us now define some quantities:

- **four-position:** denoting the position of an object in space-time

$$\bar{\mathbf{x}} = (x^0, x^1, x^2, x^3)^{\mathrm{T}}$$

- **four-velocity:** denoting a velocity of the object in space-time

$$\bar{\mathbf{u}} = \frac{\mathrm{d}\bar{x}}{\mathrm{d}\tau} = \begin{pmatrix} \frac{\mathrm{d}(ct)}{\mathrm{d}\tau} \\ \frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} \end{pmatrix}$$

  Note that, under this defintion, $\bar{\mathbf{u}}$ becomes a tensor of rank $(1,0)$.

  Now, consider two reference frames; the moving frame moves with a velocity $\mathbf{v}$ with respect to the lab frame. Consider an observer at rest in the moving frame. Then, in the observer's frame, since spatial coordinates remain same, we have $ds'^2 = -c^2 d\tau^2$ while for the lab frame, we have $ds^2 = -cdt^2 + dx^2 + dy^2 + dz^2 = -c^2dt^2 + dr^2$. Now, since the observer is at rest, they move with the same velocity $\mathbf{v}$ with respect to the lab frame and hence $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}$. Hence, by the requirement that $ds'^2 = ds$ we have:

$$\begin{aligned} d\tau^2 &= -\frac{ds^2}{c^2} \\ &= -\frac{1}{c^2}\left(-c^2dt^2 + dr^2\right) \\ &= dt^2\left[1 - \frac{1}{c^2}\left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)^2\right] \end{aligned} \tag{1}$$

  From this we obtain that $d\tau = {}^{dt}/_{\gamma}$ where $\gamma := \frac{1}{\sqrt{1-(v^2/c^2)}}$. Substituting this in the above expression for four-velocity, we obtain:

$$u^\mu \equiv \gamma(c, {}^{d\mathbf{x}}/_{dt})^{\mathrm{T}} = \gamma(c, \mathbf{v})^{\mathrm{T}}$$

- **four-momentum:** momentum of an object in space-time

$$\bar{\mathbf{p}} = m_o\bar{\mathbf{u}} = m_o\gamma(c, \mathbf{v})^{\mathrm{T}}$$

Note that m_o as used here is simply the *mass*. The term rest-mass is often used, however, it often confuses us since this implies the existence of some *relativisitc mass*.

Earlier, we used to define the relativistic mass as γm_o which kept the expression for spatial components of four-momentum and three momentum similar (that is, mass times velocity).

However, in recent times, the concept of relativistic mass is being done away with and we instead now change the expression of momentum when considering STR. Thus, we have:

$$\bar{\mathbf{p}} = (\gamma m_o c, \mathbf{p})$$

where $\mathbf{p} = \gamma m_o \mathbf{v}$ is in someway a new definition of 3-momentum.

7.2.1. Metric in STR

We have the invariant displacement as:

$$ds^2 = -c^2 dt^2 + dr^2$$

From this, we infer that the metric $g = \mathrm{diag}(-1, 1, 1, 1)$. This metric is also termed as the *Minkowski metric* and we will henceforth denote this special metric by η. Using this metric, we have:

$$\|\bar{\mathbf{u}}\|^2 = \eta_{\mu\nu} u^\mu u^\nu = -\gamma^2 c^2 + \gamma^2 \|\mathbf{v}\|^2 = \gamma^2 c^2 \left(\frac{\|v\|^2}{c} - 1 \right) = \cancel{\gamma^2} c^2 \times \frac{-1}{\cancel{\gamma^2}} = -c^2$$

Thus, we obtain that the squared-norm of the four-velocity is always $-c^2$ 🤯, which is kinda cool since even without doing anything, we technically *move* at the speed of light!!

In a similar way, we obtain $\|\bar{\mathbf{p}}\|^2 = m_o^2 \|\bar{u}\|^2 = -m_o^2 c^2$

Lecture 08: May the Force be with you!

From our previous definition of the 4-momentum, we can now define the 4-force as its derivative with respect to proper time. The 4-force is generally denoted by $\bar{\mathbf{K}}$. Thus,

$$\bar{\mathbf{K}} := \frac{\mathrm{d}\bar{\mathbf{p}}}{\mathrm{d}\tau}$$

Now, if we consider another frame, then we have:

$$\bar{\mathbf{K}}' = \frac{\mathrm{d}\bar{\mathbf{p}}'}{\mathrm{d}\tau'} = \frac{\mathrm{d}\tau}{\mathrm{d}\tau'} \frac{\mathrm{d}(\Lambda \bar{\mathbf{p}})}{\mathrm{d}\tau} = 1 \times \Lambda \frac{\mathrm{d}\bar{\mathbf{p}}}{\mathrm{d}\tau} = \Lambda \bar{\mathbf{K}}$$

This follows the rule of vector transformation and is thus a 4-vector. Moreover, note that, if $\bar{\mathbf{K}} = 0$ then $\bar{\mathbf{p}}$ is conserved. So, in the relativisitc case too, *momentum is conserved in absence of force.*

8.1. Emergence of Energy

Note that when defining the 4-momentum, we obtained:

$$\bar{\mathbf{p}} = \gamma(m_o c, m_o \mathbf{v})$$

While the spatial components do have a clear Newtonian limit in case of $v << c$ $(\gamma \to 1)$, it is not apriori clear what $p^0 = \gamma m_o c$ represent. Let us try to find out this quantity in the limit $v << c$!

$$\begin{aligned} p^0 &= \frac{m_o c}{\sqrt{1 - {}^{v^2}\!/\!{}_{c^2}}} \\ &= m_o c (1 - {}^{v^2}\!/\!{}_{c^2})^{-1/2} \\ &= m_o c \left[1 + \frac{v^2}{2c^2} + \mathcal{O}({}^{v^4}\!/\!{}_{c^4}) \right] \qquad \text{(from Taylor expansion)} \end{aligned}$$

From the above thing, we obtain:

$$p^0 c = m_o c + \frac{1}{2} m_o v^2 + \mathcal{O}(v^4/c^4)$$

The second term clearly resembles our familiar kinetic energy while we do not have a clear idea about the first term and the higher order terms. However, this quantity $p^0 c$ seems to be a right fit to be called the *energy* (since it *does* contain one of the familiar energies). Let us then define the total energy of a relativistic particle as:

$$E := p^0 c$$

Earlier we had seen that the norm $\|\bar{\mathbf{p}}\| = -m_o^2 c^2$. Using this we have:

$$\begin{aligned}
\bar{\mathbf{p}} \cdot \bar{\mathbf{p}} &= -m_o^2 c^2 \\
\implies g_{\mu\nu} p^\mu p^\nu &= -m_o^2 c^2 \\
\implies -(p^0)^2 + \|\mathbf{p}\|^2 &= -m_o^2 c^2 \\
\implies -E^2/c^2 + \|\mathbf{p}\|^2 &= -m_o^2 c^2 \\
\implies m_o^2 c^4 + \|\mathbf{p}\|^2 c^2 &= E^2 \\
\implies \sqrt{m_o^2 c^4 + \|\mathbf{p}\|^2 c^2} &= E \qquad \text{(ignoring negative solution)}
\end{aligned}$$

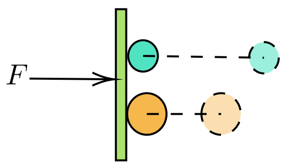
In the rest frame of the particle, the 3-momentum vanish, and hence we get the celebrated:

$$E = m_o c^2$$

Since the total energy of the particle $E = \sqrt{m_o^2 c^4 + \|\mathbf{p}\|^2 c^2}$, the kinetic energy can be obtained by subtracting the *rest energy term*, that is, $\mathrm{K.E} = E - m_o c^2$

8.2. The many faces of Mass

Let us go back to the second law. Suppose there is a *heavy* ball and a *light* ball [1]. Then we subject them to the same force. Acceleration of both bodies happend, however, we of course expect the heavy ball to accelerate less than the light ball.



From this, we can say that there exists some innate property of an object which controls the amount of acceleration, even when same force is applied to different objects. This property is the proportionality constant between force and acceleration.
We can relate this to some kind of *inertia* (a fashionable way of saying lethargy). We define the *inertial mass* m^I as:

$$m^I := \frac{|\mathbf{F}|}{|\mathbf{a}|}$$

The inertial mass is a measure of 'heaviness to move'. Now, let us consider the Coulomb's law and Newton's Gravitational Law parallely:

$$\mathbf{F}_e = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{\mathscr{r}^2} \hat{\mathscr{r}} \qquad \mathbf{F}_g = G \frac{m_1 m_2}{\mathscr{r}^2} \hat{\mathscr{r}}$$

Similar to $\mathbf{F}_e$: electrical force, we have $\mathbf{F}_g$: $\text{gravitaional force}$ while $\frac{1}{4\pi\epsilon_0}$: $\text{electrostatic coupling constant}$ is analogous to G : $\text{gravitational coupling constant.}$

[1] Consider that heaviness or lightness is being subjectively determined by the observer's opinion

Then indeed the electrical charge q_i must have the analogous gravitational charge m_i. Now, note that the gravitational charge, henceforth denoted by m^G has no apriori relation to the previous inertial mass m^I which came from the second law and was a property of the object itself.

We found an electrostatic analog of the gravitational law. It would have been nice to find some *magnetic analog* too, so that, we can also find a combined 'force' thing, similar to the Lorentz force $F = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$

Consider the electric due to a charge Q at a point. Then, the force on a charge q kept at the point is given by:

$$\mathbf{F} = q\mathbf{E} = m^I\mathbf{a} \implies \mathbf{a} = \left(\frac{q}{m^I}\right)\mathbf{E}$$

The acceleration depends on the charge-mass ratio. Similarly then in a gravitational field, we can write:

$$\mathbf{F} = m^G\mathbf{g} = m^I\mathbf{a} \implies \mathbf{a} = \left(\frac{m^G}{m^I}\right)\mathbf{g}$$

Till now we had assumed the equality of the two kind of 'masses' and then we found that the acceleration due to gravity was independent of masses, that is, every body has the same acceleration due to gravity. However, when we relax this assumption, we are at a critical stage.

8.2.1. Eötvös Experiment

So, this experiment was done to measure the different between the inertial and gravitational mass. In the experiment, two balls of different 'masses' were hanged from a rod and maintained in equilibrium.

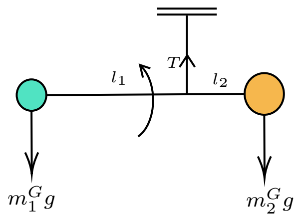


Figure 5: Setup of the Eötvös Experiment

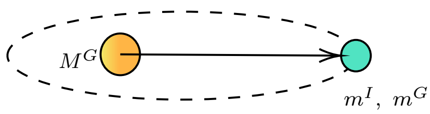
Then, considering the rotation of the Earth also, the torque was measured in the z-axis. From the measurement, we obtained experimentally:

$$\eta = \left| \frac{\frac{m_1^G}{m_1^I} - \frac{m_2^G}{m_2^I}}{\frac{1}{2}\left(\frac{m_1^G}{m_1^I} - \frac{m_2^G}{m_2^I}\right)} \right| \sim 10^{-9}$$

This proved that the ratio between the inertial and gravitational masses for different objects is nearly the same. However, experiments are limited by precision of the instruments. We can never definitely say whether $\frac{m_1^G}{m_1^I} = \frac{m_2^G}{m_2^I}$ by a mere experiment.

Lecture 09: Einstein's thoughts

Consider a planet with inertial mass m^I and gravitational mass m^G, rotating around a star with gravitational mass M^G.



Then, by the force balance condition and using $v = r\omega$, we can write the equation of motion as:

$$\frac{m^I v^2}{r} = G\frac{m^G M^G}{r^2} \implies \omega^2 = \left(\frac{m^G}{m^I}\right)\left(\frac{GM^G}{r^3}\right)$$

What this tells us is that, if indeed the ratio between inertial and gravitational mass is not constant, ω will be different for different objects and hence, two objects launched from the same point, will have their trajectories diverge at some point, which did not comply with the experimental fact.

Principle 1 (Weak Equivalence):
The ratio between gravitational and inertial mass is unity, that is,

$$\frac{m^I}{m^G} = 1$$

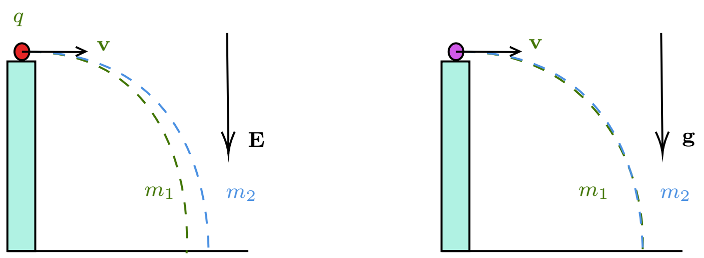
In other words, the principle states that the trajectory of a freely falling object is not affected by its mass. Let us see why!

$$\cancel{m^I}\mathbf{a} = -G\frac{\cancel{m^G}M^G}{\boldsymbol{r}^2}\hat{\boldsymbol{r}} \implies \mathbf{a} = -\frac{GM^G}{\boldsymbol{r}^2}\hat{\boldsymbol{r}}$$

Thus, we see that the acceleration of an object under the gravitational force is independent of its *mass*. In other words, the trajectory is independent on its mass.

9.1. How's gravity different?

Consider the following two cases:



(a) (b)

In case (a), we see two charges q being thrown with same horizontal velocity $\mathbf{v}$, from a building, under an electric field $\mathbf{E}$. We previously saw that the acceleration depends on the charge-mass ratio and hence, depending on the mass of the charges, the trajectories will be different.

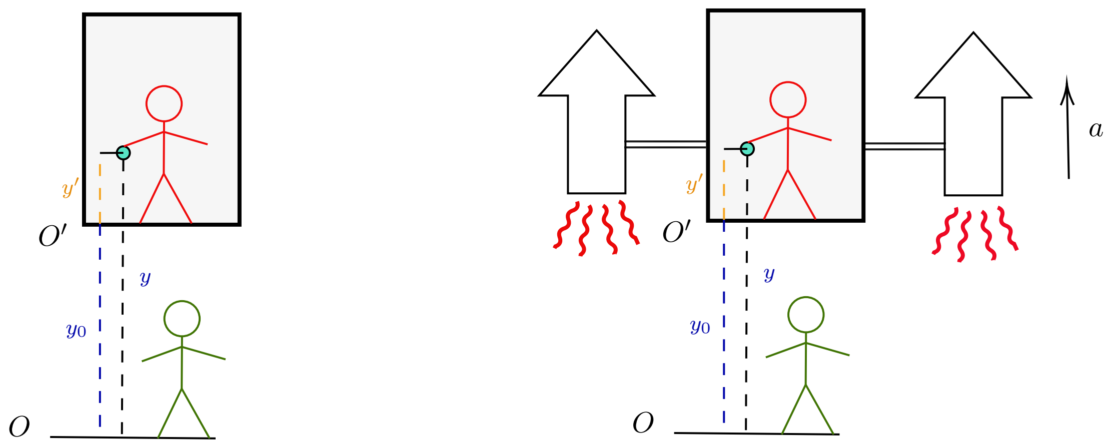
In case (b) however, if two masses m_1 and m_2 are thrown from building with same horizontal velocity $\mathbf{v}$, both of them follow the same tracjectory since acceleration under the force of gravity is independent of masses.

Thus we see an apparent disparity between gravity and other kind of forces (viz. *electromagnetic, strong nuclear, weak nuclear*).

9.2. Gedankenexperiment

We will now analyse a series of experiments in our minds and see what conclusions we can draw from each of them. All of them will consist of a blackened lift with an observer inside it and an 'all-knowing' outside observer who knows everything what is going on inside and outside the lift.

9.2.1. Lift in free space:



Case 1 : Free lift

Case 2 : Lift with rocket

Consider, there is a lift in free space (frame O'). The lift contains an observer with a ball and is shielded from outside, that is, that observer cannot see what is happening outside the lift. There is another observer outside on the ground (frame O), who has, by some divine grace, information both inside and outside the lift. Let the position of the ball from ground be y and the position of the lift from the ground be y_0 and the position of the ball from the lift be y'. Then, the constraint equation gives:

$$y = y' + y_0$$

Case 1: The lift is stationary and hence frame O' is inertial. As the lift is stationary, after the ball is released, no force acts on it and we have:

$$m\ddot{y}' = 0 \qquad\qquad m\ddot{y} = 0$$

This is the trivial case as nothing as such occurs here. The observer in the lift, after releasing the ball, finds it a rest too and infers that the frame is inertial.

Case 2: Suppose now that the lift is attached to a rocket which is producing a constant acceleration a upwards. Then we have $\ddot{y}_0 = a$. Note that, O' is no longer an inertial frame. The observer inside the lift, after releasing the ball, will find it moving. Hence they will infer that their frame is non-inertial and cannot apply NLM 😢.

The outside observer can still apply NLM and since no force acts on the ball inside the lift (force acts only on the lift by the rocket), we have:

$$m\ddot{y} = 0 \implies \ddot{y}' = 0 - a = -a$$

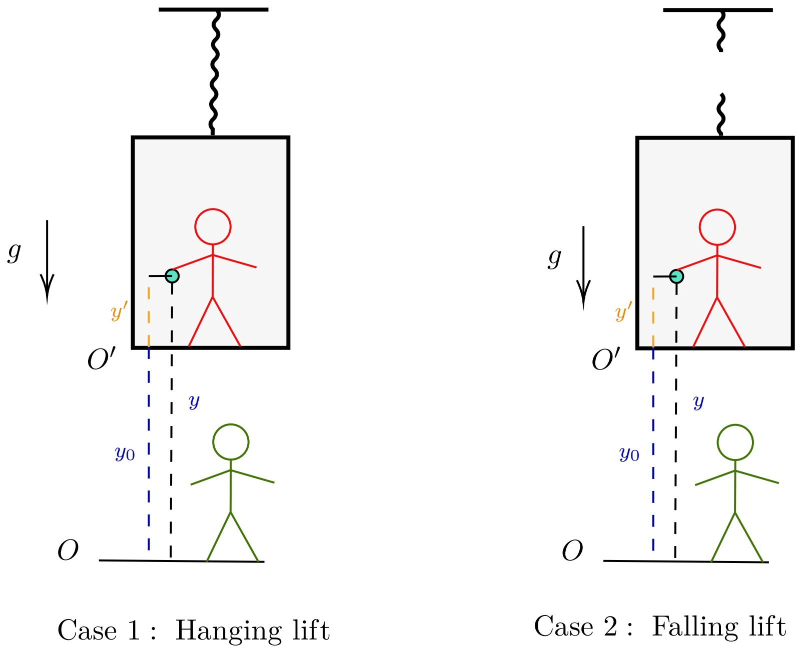
Hence we see that, from the outside frame, the ball will seem to move downwards with an acceleration a. This is what we call the *pseudo-force*. Note that, no force actually acts on the ball.

9.3. Lift under Gravity

Case 1: Now, consider that we turn on a gravitational field $\mathbf{g} = -g\hat{\mathbf{y}}$ and let the lift be hanging from some support. O' is non-inertial (as inside observer, after dropping the ball, will see it move without any force applied) and hence only O frame observer can use NLM. Since from frame O, a force $-mg$ acts on the ball, we have:

$$m\ddot{y} = -mg \implies \ddot{y} = -g$$

The lift is not moving so $\ddot{y}_0 = 0$. Thus, from this, we get $\ddot{y}' = -g$



Case 2: Consider that the thread holding the lift is cut and the lift tragically falls. Then, we have from frame O:

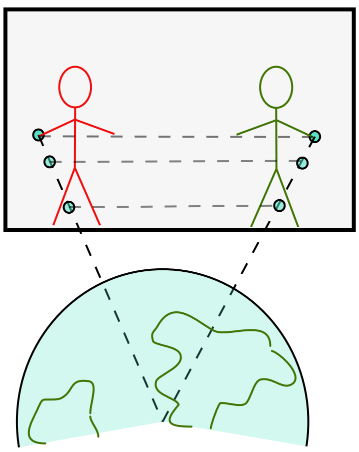
$$m\ddot{y} = -mg$$
$$m_l\ddot{y}_0 = -m_l g$$

Then, from the above two and the constraint equation, we get $\ddot{y}' = -g - (-g) = 0$. The inside observer will then see that the acceleration of the ball with respect to them is zero. When released, the ball remains where it was. The observer then infers that they are in a 'free space'.

Conclusions: For an observer inside the lift,

1. A lift in *constant gravity* is same as an *accelerating lift*.
2. A *freely-falling* lift is the same as an *inertial lift*.

We see that there exist atleast one observer (the freely-falling one) with respect to whom, motion of an object under gravity appears as if it is moving under no force. We can then conclude that the force of gravity cannot be a *vector* as, if it was indeed a vector, it would transform like a vector which would mean that force in any other frame would have been zero[1].



[1] A vector is zero in all frames iff it is zero in one frame (direct consequence of transformation)

Note that, if the force of gravity was *inhomogenous*, that is, varying with distance (which we know it is), then there would be pronounced effect inside the lift. Two balls thrown at a separation d inside the lift, will have their distance change if d is comparable to say the diameter of the earth. Then, the inside observer will no longer think that this is a free-space.

Thus, to eliminate the effect of gravity in a freely-falling frame, one must consider the frame to be *local*, that is, everything should be happening around a point and its neighbourhood only. Only then, the gravitational field can be treated as homogenous and the free-space assumption will be correct.

Lecture 10: Connections are important

In the previous section, we had seen that the concept of being *local* was introduced in order to approximate a frame where the effect of gravity is eliminated. Using such a motivation, we have:

Principle 2 (Einstein's Equivalence Principle):
All freely falling frames under the influence of *gravity* are *locally inertial*, that is, Lorentz frames where special relativity is valid and there is absence of gravity

Consider the case of the falling lift as before. Then, if we write the equation of the ball from the O frame, we have:

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\mathbf{g} \implies \mathbf{r} = \mathbf{r}_i - \frac{1}{2}\mathbf{g}t^2$$

where $\mathbf{r}_i$ is the initial position of the ball. Now, from the O' frame, as no force acts on the ball (note that the observer sees the ball floating in front of them) we have:

$$\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t'^2} = 0 \implies \mathbf{r}' = \mathbf{r}'_i$$

From the above equations we have:

$$\mathbf{r}' = \mathbf{r}'_i = (\mathbf{r}_i - \mathbf{r}_0) = \left(\mathbf{r} + \frac{1}{2}\mathbf{g}t^2\right) - \mathbf{r}_0$$

This essentially describes a *coordinate transformation* where $\mathbf{r} \to \mathbf{r}'$ and $t \to t'$. However, note that a term t^2 appears in the expression and hence, the coordinate transformations are *non-linear*. In general, we can have the coordinate transformation matrices themselves to be dependent on the coordinates, that is, $\Lambda^\mu{}_\nu \equiv \Lambda^\mu{}_\nu(x^\alpha)$

If we allow non-linear coordinate transformations, we can go to a frame where *gravity is absent!*

Thus, we essentially have a recipe to tackle gravity. We first go to a frame where gravity is absent (the equivalence principle guarantees the existence of such a frame), do our calculations in that frame, then transform back to our original coordinate system.

Consider a particle moving freely under a purely gravitational force. Now, suppose O' is such a frame where gravity is absent. Then we have:

$$\frac{\mathrm{d}^2x'^\mu}{\mathrm{d}\tau^2} = 0$$

How do we obtain the equation of motion in frame O? Well, we will just transform to that frame. Since x'^{μ} is dependent on the coordinates of the other frame x^{ν}, we can write,

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x'^{\mu}}{\mathrm{d}\tau}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau}\right) \\
&= \frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\mathrm{d}^2x^{\nu}}{\mathrm{d}\tau^2} + \frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right) \\
&= \frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\mathrm{d}^2x^{\nu}}{\mathrm{d}\tau^2} + \frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau}
\end{aligned}$$

The first term seems to contain the acceleration in the O frame and it would be nice to isolate this term. For that, let us multiply both sides by $\frac{\partial x^{\beta}}{\partial x'^{\mu}}$ and contract using the fact that $\frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial x'^{\mu}}{\partial x^{\nu}} = \delta^{\beta}{}_{\nu}$

$$\begin{aligned}
0 &= \frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\mathrm{d}^2x^{\nu}}{\mathrm{d}\tau^2} + \frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau} \\
&= \delta^{\beta}{}_{\nu}\frac{\mathrm{d}^2x^{\nu}}{\mathrm{d}\tau^2} + \frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau} \\
&= \frac{\mathrm{d}^2x^{\beta}}{\mathrm{d}\tau^2} + \frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}}\frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\tau} \\
&= \frac{\mathrm{d}^2x^{\beta}}{\mathrm{d}\tau^2} + \Gamma^{\beta}{}_{\nu\alpha}u^{\alpha}u^{\nu}
\end{aligned}$$

Here we have defined the following:

$$\Gamma^{\beta}{}_{\nu\alpha} := \frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}} = \left(\Lambda^{-1}\right)^{\beta}{}_{\mu}\frac{\partial}{\partial x^{\nu}}(\Lambda^{\mu}{}_{\alpha})$$

The above entities (for all indices) are termed as *connection coefficients*. The name connection coefficients comes from the fact that they connect neighboring points and tell us how to calculate the rate of change of a vector field from one point to another nearby, even though the coordinate system may itself be changing. The equation of motion thus obtained is:

$$\frac{\mathrm{d}^2x^{\beta}}{\mathrm{d}\tau^2} = -\Gamma^{\beta}{}_{\nu\alpha}u^{\alpha}u^{\nu} \tag{2}$$

The above equation is called the *geodesic equation* and resembles some kind of *pseudo-force* or *non-inertial force*. Note that the connection coefficients in its present form, depends on both $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}}'$ which is kind of problematic since we might not always have knowledge of both the frames.

Since the primed frame is locally inertial, the metric is *Minkowski*, that is,

$$\eta_{\mu\nu} = g'_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}g_{\alpha\beta}$$

Inverting the relation, we have:

$$g_{\mu\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x'^{\beta}}{\partial x^{\nu}}\eta_{\alpha\beta}$$

From the definition of the affine connection, we can see that, since partial derivatives commute, the affine connections are *symmetric* in the lower indices. Thus,

$$\Gamma^{\alpha}{}_{\mu\nu} = \Gamma^{\alpha}{}_{\nu\mu}$$

Now, let us contract the upper index of the affine connection definition which will be used later:

$$\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\Gamma^{\beta}{}_{\nu\alpha} = \frac{\partial x'^{\lambda}}{\partial x^{\beta}}\frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}} = \delta^{\lambda}{}_{\mu}\frac{\partial^2 x'^{\mu}}{\partial x^{\nu}\partial x^{\alpha}} = \frac{\partial^2 x'^{\lambda}}{\partial x^{\nu}\partial x^{\alpha}} = \partial_{\nu}\left(\frac{\partial x'^{\lambda}}{\partial x^{\alpha}}\right)$$

Then, let us differentiate both sides with x^{ρ} and using the above found relation, we obtain:

$$\begin{aligned}
\partial_{\rho} g_{\mu\nu} &= \partial_{\rho}\left(\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\right) \times \frac{\partial x'^{\beta}}{\partial x^{\nu}}\eta_{\alpha\beta} + \partial_{\rho}\left(\frac{\partial x'^{\beta}}{\partial x^{\nu}}\right) \times \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\eta_{\alpha\beta} \\
&= \Gamma^{\sigma}{}_{\rho\mu}\frac{\partial x'^{\alpha}}{\partial x^{\sigma}}\frac{\partial x'^{\beta}}{\partial x^{\nu}}\eta_{\alpha\beta} + \Gamma^{\sigma}{}_{\rho\nu}\frac{\partial x'^{\beta}}{\partial x^{\sigma}}\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\eta_{\alpha\beta} \\
&= \Gamma^{\sigma}{}_{\rho\mu} g_{\sigma\nu} + \Gamma^{\sigma}{}_{\rho\nu} g_{\mu\sigma}
\end{aligned}$$

We thus obtain:

$$\partial_{\rho} g_{\mu\nu} = \Gamma^{\sigma}{}_{\rho\mu} g_{\sigma\nu} + \Gamma^{\sigma}{}_{\rho\nu} g_{\mu\sigma} \tag{3}$$

Since these are true for any indices, we obtain two more equations by the exchange $\rho \leftrightarrow \mu$ and $\rho \leftrightarrow \nu$:

$$\partial_{\mu} g_{\rho\nu} = \Gamma^{\sigma}{}_{\mu\rho} g_{\sigma\nu} + \Gamma^{\sigma}{}_{\mu\nu} g_{\rho\sigma} \tag{4}$$

$$\partial_{\nu} g_{\mu\rho} = \Gamma^{\sigma}{}_{\nu\mu} g_{\sigma\rho} + \Gamma^{\sigma}{}_{\nu\rho} g_{\mu\sigma} \tag{5}$$

Adding Eq. 4 and Eq. 5 and then subtracting Eq. 3, we obtain:

$$\begin{aligned}
\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu} \quad &= \Gamma^{\sigma}{}_{\mu\rho} g_{\sigma\nu} + \Gamma^{\sigma}{}_{\mu\nu} g_{\rho\sigma} + \Gamma^{\sigma}{}_{\nu\mu} g_{\sigma\rho} + \Gamma^{\sigma}{}_{\nu\rho} g_{\mu\sigma} - \Gamma^{\sigma}{}_{\rho\mu} g_{\sigma\nu} - \Gamma^{\sigma}{}_{\rho\nu} g_{\mu\sigma} \\
\implies \qquad \Gamma^{\sigma}{}_{\mu\nu} g_{\rho\sigma} &= \frac{1}{2}(\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu}) \\
\implies \qquad g^{\rho\lambda}\Gamma^{\sigma}{}_{\mu\nu} g_{\rho\sigma} &= \frac{1}{2} g^{\rho\lambda}(\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu}) \\
\implies \qquad \delta^{\lambda}{}_{\sigma}\Gamma^{\sigma}{}_{\mu\nu} &= \frac{1}{2} g^{\rho\lambda}(\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu})
\end{aligned}$$

The red and the green terms cancel first and then we contract the metric on the left to obtain an extremely important identity for the affine connection:

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\rho\lambda}(\partial_{\mu} g_{\rho\nu} + \partial_{\nu} g_{\mu\rho} - \partial_{\rho} g_{\mu\nu}) \tag{6}$$

We obtained an expression for the affine connection in terms of the metric for the space. In this form, these are called the *Christoffel symbols* and denoted by:

$$\begin{Bmatrix} \lambda \\ \mu\nu \end{Bmatrix}$$

NOTE: We can remember the equation like this: In the left hand side, the upper index is λ which comes only in the metric in the right side. The derivatives with $+$ sign are with respect to the lower indices μ and ν and the index of the metric in the derivative is a dummy index ρ and the other lower index not appearing in the derivative. Finally, the derivative with $-$ sign is with respect to the dummy index and the metric contains both the lower indices of the left.

Hence, we see that the equation of motions are known provided the Christoffel symbols are known, which are known if metric is known!

10.1. Action tells a lot!

We will now consider one of the most fundamental quantitites in physics, the *action*, albeit in the *relativistic* sense.

$$\mathcal{S} = \int \mathrm{d}t\ L = \int \mathrm{d}t\ (T - V)$$

Note that, action must be invariant under $3-$rotation in Newtonian mechanics. Since relativisitc case should also converge to the classical action in the Newtonian limit, we want the action to be also invariant.

Till now, we know three invariant (scalar) quantitites: $m, c, d\tau(\text{or } ds)$. Let us construct a quantity with same dimension as of Lagrangian. We then define:

$$\mathcal{S}_{\text{rel}} = \int \mathrm{d}\tau\ L := \mathcal{A} \int \mathrm{d}\tau\ mc^2$$

where we replaced the Lagrangian by $\mathcal{A}mc^2$, $\mathcal{A}$ is just some constant. This is just our guess but let us see what happens in the Newtonian limit! Note that, using Eq. 1, we can write the action as:

$$\begin{aligned}
\mathcal{S}_{\text{rel}} &= \mathcal{A}mc^2 \int \mathrm{d}t\ \sqrt{1 - {v^2}/{c^2}} \\
&= \mathcal{A}mc^2 \int \mathrm{d}t\ (1 - {v^2}/{2c^2} + \mathcal{O}({v^4}/{c^4})) \\
&= -\mathcal{A} \int \mathrm{d}t\ \left(\frac{1}{2}mv^2 - mc^2\right)
\end{aligned}$$

Well, this looks like kind of $T - V$ if we identity the rest energy as some 'potential' energy, provided $\mathcal{A} = -1$. Thus, in the Newtonian limit atleast, this defnition of *relativistic action* makes sense kinda. Hence we have:

$$\mathcal{S}_{\text{rel}} := \int \mathrm{d}\tau\ (-mc^2) \tag{7}$$

Note that, we had previously seen that, $\|\bar{\mathbf{u}}\|^2 = -c^2$ and hence we can replace $-c$ with this quantity in the action expression. Using this we obtain:

$$\mathcal{S}_{\text{rel}} = \int \mathrm{d}\tau\ m\|\bar{\mathbf{u}}\|^2 = \int \mathrm{d}\tau\ m\ g_{\mu\nu}u^\mu u^\nu$$

Now, we know that we can derive the ELEOM from the least action principle ($\delta\mathcal{S} = 0$). We have:

$$\frac{\partial L}{\partial x^\lambda} = m\partial_\lambda g_{\mu\nu}\ u^\mu u^\nu \qquad\qquad \frac{\partial L}{\partial u^\lambda} = 2mg_{\lambda\nu}u^\nu$$

Now, the metric could be a function of the coordinates and hence we have:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial u^\lambda}\right) = 2m \times \left(g_{\lambda\nu}\frac{\mathrm{d}u^\nu}{\mathrm{d}\tau} + \partial_\sigma g_{\lambda\nu}u^\sigma u^\nu\right)$$

Plugging this into the ELEOM, we obtain:

$$2m \times \left(g_{\lambda\nu}\frac{\mathrm{d}u^\nu}{\mathrm{d}\tau} + \partial_\sigma g_{\lambda\nu}u^\sigma u^\nu\right) = m\partial_\lambda g_{\mu\nu}\ u^\mu u^\nu \implies g_{\lambda\nu}\frac{\mathrm{d}u^\nu}{\mathrm{d}\tau} \qquad = \frac{1}{2}(\partial_\lambda g_{\mu\nu}\ u^\mu u^\nu) - \partial_\sigma g_{\lambda\nu}u^\sigma u^\nu$$

Then, contracting the metric on the left side with $g^{\rho\lambda}$, we get:

$$\frac{\mathrm{d}u^\rho}{\mathrm{d}\tau} = g^{\lambda\rho} \times \left(\frac{1}{2}\partial_\lambda g_{\sigma\nu} - \partial_\sigma g_{\lambda\nu}\right)u^\sigma u^\nu$$

Let us manipulate the term in the bracket:

$$\begin{aligned}
&\left(\frac{1}{2}\partial_\lambda g_{\sigma\nu} - \partial_\sigma g_{\lambda\nu}\right)u^\sigma u^\nu \\
&= \frac{1}{2}\partial_\lambda g_{\sigma\nu}u^\sigma u^\nu - \frac{1}{2}(\partial_\sigma g_{\lambda\nu} + \partial_\sigma g_{\nu\lambda})u^\sigma u^\nu \\
&= \frac{1}{2}\partial_\lambda g_{\sigma\nu}u^\sigma u^\nu - \frac{1}{2}\partial_\sigma g_{\lambda\nu}u^\sigma u^\nu - \frac{1}{2}\partial_\sigma g_{\nu\lambda}u^\sigma u^\nu \\
&= \frac{1}{2}\partial_\lambda g_{\sigma\nu}u^\sigma u^\nu - \frac{1}{2}\partial_\sigma g_{\lambda\nu}u^\sigma u^\nu - \frac{1}{2}\partial_\nu g_{\sigma\lambda}u^\nu u^\sigma \\
&= \frac{1}{2}(\partial_\lambda g_{\sigma\nu} - \partial_\sigma g_{\lambda\nu} - \partial_\nu g_{\sigma\lambda})u^\sigma u^\nu
\end{aligned}$$

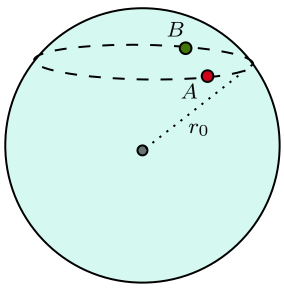
Above, we had used the symmetry of the metric tensor and then we renamed the dummy indices in the sum. We substitute this in the previous equation and obtain:

$$\frac{\mathrm{d}u^\rho}{\mathrm{d}\tau} + \frac{1}{2}g^{\lambda\rho}(\partial_\sigma g_{\lambda\nu} + \partial_\nu g_{\sigma\lambda} - \partial_\lambda g_{\sigma\nu})u^\sigma u^\nu = 0 \implies \frac{\mathrm{d}u^\rho}{\mathrm{d}\tau} + \Gamma^\rho{}_{\sigma\nu}u^\sigma u^\nu = 0$$

This is exactly the *geodesic* equation that we had obtained earlier. Hence, using this definition of the action, we can recover the geodesic equation. Thus, **geodesics are paths of extreme action!**

Lecture 11: Finding Geodesics

Since a particle following the geodesics have no *proper acceleration* in a local inertial frame, the geodesics generally represent the 'straightest' path between two points in a space. Let us take the example for the shortest distance between two cities A and B on the Earth. Since the Earth can be assumed to be a sphere, we just need to find the geodesic equation on $\mathbb{S}^2$ and then solve it.



Intuitively, out of the infinite ways of going between those cities, we think that moving along a line of latitude would be more economical. However, let us check if that is indeed so!

For that, note that the invariant distance on the sphere $\mathbf{S}^2$ of radius r_0 is:

$$ds^2 = r_0^2\mathrm{d}\theta^2 + r_0^2\sin^2\theta\mathrm{d}\phi^2$$

From this, we can easily identity the metric and inverse metric for the space as:

$$[g] \equiv \begin{pmatrix} r_0^2 & 0 \\ 0 & r_0^2\sin^2\theta \end{pmatrix} \qquad [g^{-1}] \equiv \begin{pmatrix} 1/r_0^2 & 0 \\ 0 & 1/r_0^2\sin^2\theta \end{pmatrix}$$

Thus we have,

$$g^{\theta\theta} = \frac{1}{r_0^2} \quad g_{\theta\theta} = r_0^2 \quad g^{\phi\phi} = \frac{1}{r_0^2\sin^2\theta} \quad g_{\phi\phi} = r_0^2\sin^2\theta$$

Now, the geodesic equation contains the Christoffel symbols and let us calculate them one by one. Note that the metric being diagonal is a blessing since then, the apparent summation in the Christoffel symbol expression is not needed.

$$\Gamma^{\theta}{}_{\theta\theta} = \tfrac{1}{2} g^{\theta\theta}\left(\partial_\theta g_{\theta\theta} + \partial_\theta g_{\theta\theta} - \partial_\theta g_{\theta\theta}\right) = 0$$

$$\Gamma^{\theta}{}_{\theta\phi} = \tfrac{1}{2} g^{\theta\theta}\left(\partial_\theta g_{\theta\phi} + \partial_\phi g_{\theta\theta} - \partial_\theta g_{\theta\phi}\right) = 0$$

$$\begin{aligned}\Gamma^{\theta}{}_{\phi\phi} &= \tfrac{1}{2} g^{\theta\theta}\left(\partial_\phi g_{\theta\phi} + \partial_\phi g_{\theta\phi} - \partial_\theta g_{\phi\phi}\right)\\ &= -\tfrac{1}{2}\, r_0^2 \sin\theta\cos\theta \cdot \tfrac{1}{r_0^2}\\ &= -\sin\theta\cos\theta\end{aligned}$$

$$\Gamma^{\phi}{}_{\theta\theta} = \tfrac{1}{2} g^{\phi\phi}\left(\partial_\theta g_{\theta\phi} + \partial_\theta g_{\theta\phi} - \partial_\phi g_{\theta\theta}\right) = 0$$

$$\begin{aligned}\Gamma^{\phi}{}_{\theta\phi} &= \tfrac{1}{2} g^{\phi\phi}\left(\partial_\theta g_{\phi\phi} + \partial_\phi g_{\theta\phi} - \partial_\phi g_{\theta\phi}\right)\\ &= \tfrac{1}{2} \cdot \frac{1}{r_0^2 \sin^2\theta} \cdot \left(2 r_0^2 \sin\theta\cos\theta\right)\\ &= \cot\theta\end{aligned}$$

$$\Gamma^{\phi}{}_{\phi\phi} = \tfrac{1}{2} g^{\phi\phi}\left(\partial_\phi g_{\phi\phi} + \partial_\phi g_{\phi\phi} - \partial_\phi g_{\phi\phi}\right) = 0$$

Thus, we get only two non-zero Christoffel symbols. Since there are two parameters θ and ϕ for the problem, we would obtain two geodesic equations which are:

$$\frac{\mathrm{d}^2 x^\theta}{\mathrm{d}\tau^2} + \Gamma^{\theta}{}_{\phi\phi}\frac{\mathrm{d}x^\phi}{\mathrm{d}\tau}\frac{\mathrm{d}x^\phi}{\mathrm{d}\tau} = 0 \implies \ddot{\theta} - \sin\theta\cos\theta\,\dot{\phi}^2 = 0 \tag{8}$$

$$\frac{\mathrm{d}^2 x^\phi}{\mathrm{d}\tau^2} + \Gamma^{\phi}{}_{\theta\phi}\frac{\mathrm{d}x^\phi}{\mathrm{d}\tau}\frac{\mathrm{d}x^\theta}{\mathrm{d}\tau} + \Gamma^{\phi}{}_{\phi\theta}\frac{\mathrm{d}x^\theta}{\mathrm{d}\tau}\frac{\mathrm{d}x^\phi}{\mathrm{d}\tau} = 0 \implies \ddot{\phi} + 2\cot\theta\,\dot{\theta}\dot{\phi} = 0 \tag{9}$$

Multiplying Eq. 9 with $\sin^2\theta$, we obtain:

$$\sin^2\theta\ddot{\phi} + 2\sin\theta\cos\theta\dot{\phi}\dot{\theta} \implies \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\dot{\phi}\sin^2\theta\right) = 0$$

The above equation implies that $\dot{\phi}\sin^2\theta$ is a constant of motion and let us denote it by $\mathcal{A}$. Thus, we obtain,

$$\dot{\phi} = \frac{\mathcal{A}}{\sin^2\theta}$$

Using the above in Eq. 8, we will get:

$$\ddot{\theta} = \frac{\cos\theta}{\sin^3\theta}\mathcal{A}^2$$

Note that we had obtained the equations for θ and ϕ in terms of τ, however, we want to understand θ in terms of ϕ or vice-versa, which would give us some equation of the trajectory. For this, let us use our beloved chain rule!

$$\begin{aligned}\dot{\theta} &= \frac{\mathrm{d}\theta}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{\mathrm{d}\theta}{\mathrm{d}\phi} \times \frac{\mathcal{A}}{\sin^2\theta}\\ \ddot{\theta} &= \frac{\mathrm{d}\phi}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\frac{\mathcal{A}}{\sin^2\theta}\right) = \frac{\mathcal{A}}{\sin^2\theta}\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\frac{\mathcal{A}}{\sin^2\theta}\right)\end{aligned}$$

Using this in Eq. 8 we have:

$$
\begin{aligned}
\frac{\mathcal{A}}{\sin^2\theta}\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathcal{A}}{\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right) - \frac{\mathcal{A}^2}{\sin^2\theta}\frac{\sin\theta\cos\theta}{\sin^2\theta} &= 0\\
\implies \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathcal{A}}{\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right) - \mathcal{A}\cot\theta &= 0\\
\implies \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{1}{\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right) - \cot\theta &= 0
\end{aligned}
$$

Now note that $\frac{\mathrm{d}}{\mathrm{d}\phi}(\cot\theta) = -\frac{1}{\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\phi}$ and using this, our equation becomes:

$$
-\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathrm{d}}{\mathrm{d}\phi}(\cot\theta)\right) - \cot\theta = 0 \implies \frac{\mathrm{d}^2\cot\theta}{\mathrm{d}\phi^2} = -\cot\theta
$$

The solution of the above can be written as:

$$
\cot\theta = c_1\cos\phi + c_2\sin\phi \qquad \text{for some } c_1, c_2 \tag{10}
$$

So, what does this equation represent? For that, let us consider a plane passing through the origin whose equation is given by:

$$
ax + by + cz = 0
$$

Now consider any sphere of radius r_0 and centred at origin. Any point on the sphere is parameterised by:

$$
x = r_0\sin\theta\cos\phi \qquad y = r_0\sin\theta\sin\phi \qquad z = r_0\cos\theta
$$

The intersection of the plane passing through origin (centre of sphere) and the sphere essentially gives what is know as *great circles* whose equation can be given by simultaneously solving the two equations above. We have,

$$
a(r_0\sin\theta\cos\phi) + b(r_0\sin\theta\sin\phi) + cr_0\cos\theta = 0 \implies (a\cos\phi + b\sin\phi) + c\cot\phi = 0
$$

This gives us the equation of the great circle as,

$$
\cot\theta = d_1\cos\phi + d_2\cos\phi
$$

which is the same form as Eq. 10. Thus, we conclude that the geodesics on the sphere $\mathbb{S}^2$ are the *great circles* and hence it is actually shorter to move along the great circle to reach from city A to city B, rather than moving along the latitudes.

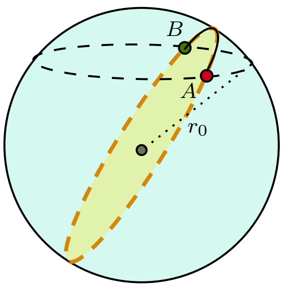
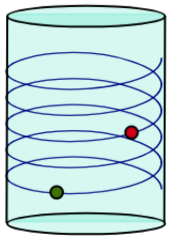


Figure 6: Moving along the great circle is shorter than moving along the latitude.

Note however that, if we travel on the opposite arc of the great circle, it will be the longest route. This is because, geodesics are paths of *extreme* action, not only minimum. In general, we should thus never fly along the latitude (except if it is the equator, since equator is itself a great circle). Now, let us consider a cylinder.



The invariant displacement for the cylinder can be written as

$$ds^2 = r_0^2 \mathrm{d}\theta^2 + \mathrm{d}z^2$$

Note that the metric for this space is,

$$[g] \equiv \begin{pmatrix} r_0^2 & 0 \\ 0 & 1 \end{pmatrix}$$

This is a *constant metric* which does not depend on the parameters θ or z. Then, all the Christoffel symbols will be zero since Christoffel symbols are essentially written in terms of the derivatives of the metric. The geodesic equations thus simplify to:

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}\tau^2} = 0 \qquad \frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} = 0$$

which have the solutions:

$$\theta = c_1\tau + c_2 \qquad z = d_1\tau + d_2 \implies z = d_1\left(\frac{\theta - c_2}{c_1}\right) + d_2 \implies z = m\theta + k$$

Suitably taking $k = 0$, we see that the equation represents a helix with pitch $2\pi m$. Thus, the geodesics on the cylinder are helices and the shortest distance to move from one point to another on the cylinder is along a helix.

NOTE: Similar to the case of the cylinder, for the flat spacetime, the metric is $\eta_{\mu\nu} = \mathrm{diag}(-1, +, +, +)$ which is constant too and hence, the geodesics for the Minkowski space will also satisfy the same equation as of the cylinder:

$$x^\mu = a^\mu + u^\mu\tau$$

where u^μ is the constant four-velocity and a^μ is the initial position (at $\tau = 0$). This shows that geodesics in flat spacetime are straight lines.

Lecture 12: Towards the action for Gravity: Invariant Volumes

Motion under gravity can be determined once we know the metric $g_{\mu\nu}$, however, we want to see how to obtain the equation for the metric. In other words, since equations of motion can be derived from *action*, we need to find an action for gravity!

12.1. Maxwell's action in spacetime

In free space, the *electromagnetic* action can be written as:

$$\mathcal{S}_{\mathrm{EM}} = \int \mathrm{d}^4\bar{\mathbf{x}}\left[-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}\right]$$

Here $F_{\mu\nu}$ is the anti-symmetric electromagnetic tensor and is defined by:

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu \qquad F_{\mu\nu} = -F_{\nu\mu} \tag{11}$$

$A^\mu = (\phi, \mathbf{A})$ is the four-potential, which we will simply call the *electromagnetic field*. From the above defined action, we can obtain all of Maxwell's equations. We can manipulate the action as follows,

$$\begin{aligned}\mathcal{S}_{\text{EM}} &= \frac{1}{4\mu_0}\int \mathrm{d}^4\bar{\mathbf{x}} - [\partial_\mu A_\nu - \partial_\nu A_\mu]F^{\mu\nu} \\ &= 2\times\frac{1}{4\mu_0}\int \mathrm{d}^4\bar{\mathbf{x}}[-\partial_\mu A_\nu F^{\mu\nu}] \qquad \text{(using anti-symmetric property)} \\ &\stackrel{\text{IBP}}{=} \frac{1}{2\mu_0}\int \mathrm{d}^4\bar{\mathbf{x}}[A_\nu\partial_\mu F^{\mu\nu} - \partial_\mu(A_\nu F^{\mu\nu})]\end{aligned}$$

The last integral is a total-derivative which can be converted to the surface integral using Stokes' theorem. We assume that at spatial and temporal infinities, the fields vanish and thus, the integral is zero. Then we obtain,

$$\mathcal{S}_{\text{EM}} = \frac{1}{2\mu_0}\int \mathrm{d}^4\bar{\mathbf{x}}[A_\nu\partial_\mu F^{\mu\nu}] = \frac{1}{2\mu_0}\int \mathrm{d}^4\bar{\mathbf{x}}[A_\nu\partial_\mu(\partial_\mu A_\nu - \partial_\nu A_\mu)]$$

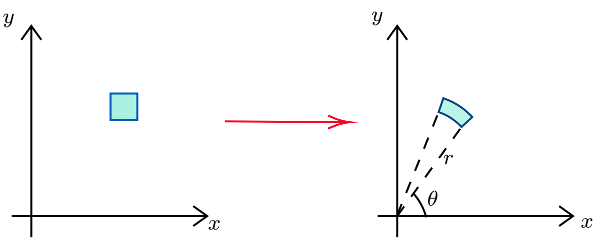
Properties of the Maxwell action:

- $\mathcal{S}_{\text{EM}}$ is a functional of A_μ, that is, it takes a value of A_μ and returns a number.
- $\mathcal{S}_{\text{EM}}$ is a scalar under Lorentz transformation. We can see this since the indices in F are contracted and $\mathrm{d}^4\bar{\mathbf{x}}$ can be shown to be Lorentz-invariant.
- $\mathcal{S}_{\text{EM}}$ has atmost two derivatives of the field A_μ

Einstein took inspiration from this action, however, he tried to include the invariance of this action for general transformations too! He proposed that *action* for the metric should be of the following form:

$$\mathcal{S}_{\text{G}} = \int \mathrm{d}^4\bar{\mathbf{x}} f(g_{\mu\nu}, \partial_\alpha\partial_\beta g_{\mu\nu})$$

The above mentioned form has many subtleties, let us look at them one by one. First note that if the function f has to be a scalar under general coordinate transformations, then the integral measure $\mathrm{d}^4\bar{\mathbf{x}}$ must also be a scalar. However, this is not apriori obvious. As an example, note how the 'volume' element changes when we move from Cartesian to polar coordinates.



In Cartesian coordinates, the volume element (actually the area element since it is 2D) is $\mathrm{d}V = \mathrm{d}x\ \mathrm{d}y$ while in the polar coordinates, the volume elements is $\mathrm{d}V' = \mathrm{d}r\ (r\ \mathrm{d}\theta) = r\ \mathrm{d}\theta\ \mathrm{d}r$ There is an extra factor r apart from the coordinate differential. We denote it by $J(r,\theta)$ which is the *Jacobian* of the transformation. Thus, to keep the volume element invariant, we need to add the effect of the Jacobian too!

12.2. Metric, Transformations and Jacobian

Consider the Leibniz formula for the determinant of a square $n \times n$ matrix. The formula says that for any matrix A, we have [1]:

$$\det(A) = \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \prod_{i=1}^{n} a^i_{\tau(i)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a^i_{\sigma(i)}$$

From this, we can derive another formula for the determinant using the Levi-Civita symbol:

Identity 1:

$$\det(M) = \tilde{\epsilon}^{\mu'_1 \cdots \mu'_n} \tilde{\epsilon}_{\mu_1 \ldots \mu_n} A^{\mu_1}{}_{\mu'_1} A^{\mu_2}{}_{\mu'_2} \ldots A^{\mu_n}{}_{\mu'_n}$$

Let us keep this aside and calculate the volume element for a two-dimensional coordinate system, in primed and unprimed frames. Note that the volume element can be written in the following way:

$$\mathrm{d}V = J(\bar{\mathbf{x}}) \frac{1}{2!} \epsilon_{\alpha\beta} (-1)^p \mathrm{d}x^\alpha \mathrm{d}x^\beta$$
$$\mathrm{d}V = J(\bar{\mathbf{x}}') \frac{1}{2!} \epsilon_{\mu\nu} (-1)^p \mathrm{d}x^\mu \mathrm{d}x^\nu$$

What just happened here! 🥴. First note that p here is the sign of the permutation, that is, if the permutation α, β is odd then $p = -1$, else $p = 0$. This is done to ensure that the quantity is positive. Also note that, J as defined here should always be taken to be positive, to make $\mathrm{d}V$ non-negative. Subtleties aside, if we expand the first expression, we would get:

$$\mathrm{d}V = J(\bar{\mathbf{x}}) \frac{1}{2} \big(\epsilon_{12} (-1)^0 \mathrm{d}x^1 \mathrm{d}x^2 + \epsilon_{21} (-1)^1 \mathrm{d}x^2 \mathrm{d}x^1 \big) = J(\bar{\mathbf{x}}) \mathrm{d}x^1 \mathrm{d}x^2$$

This is indeed the expression for the volume element in the unprimed frame. The above formula is related to differential forms and wedge products which are defined in terms of *alternating-operator* and gives a generalisation of the volume element $\mathrm{d}V = dx^1 \wedge dx^2 \wedge \ldots$ which we will not explore.

Anyways, let us now calculate something. Since we want $\mathrm{d}V = \mathrm{d}V'$, then we demand that:

$$J(\bar{\mathbf{x}}) \frac{1}{2!} \epsilon_{\alpha\beta} (-1)^p \mathrm{d}x^\alpha \mathrm{d}x^\beta = J(\bar{\mathbf{x}}') \frac{1}{2!} \epsilon_{\mu\nu} (-1)^p \mathrm{d}x'^\mu \mathrm{d}x'^\nu$$

Let us take the primed coordinate to be Cartesian and hence, $J(\bar{\mathbf{x}}') = 1$. Then, transforming the coordinates we obtain:

$$J(\bar{\mathbf{x}}) \frac{1}{2!} \epsilon_{\alpha\beta} (-1)^p \mathrm{d}x^\alpha \mathrm{d}x^\beta = \frac{1}{2!} \epsilon_{\mu\nu} (-1)^p \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \mathrm{d}x^\alpha \mathrm{d}x^\beta$$

Now, since the differentials are arbitrary, both quantities are equal. Let us contract both sides with $\epsilon^{\alpha\beta}$ which gives us:

$$J(\bar{\mathbf{x}}) \frac{1}{2!} \epsilon_{\alpha\beta} \epsilon^{\alpha\beta} = \frac{1}{2!} \epsilon_{\mu\nu} \epsilon^{\alpha\beta} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \det(\Lambda^\mu{}_\alpha)$$

In the RHS, we have used the formula that we had written for the determinant previously. In the LHS, note that $\epsilon_{\alpha\beta}\epsilon^{\alpha\beta}$ will be 2 and hence, the final result that we obtained is:

$$J(\bar{\mathbf{x}}) = \det(\Lambda^\mu{}_\alpha)$$

Thus, the Jacobian of the transformation is basically the determinant of the transformation matrix. Using the above found expression, we can write the volume element (now, in arbitrary dimensions) as:

$$\mathrm{d}V = \det(\Lambda^\mu{}_\nu) \prod_\alpha \mathrm{d}x^\alpha$$

[1] You may also look into Sec 6.5.3 here!

Now, let us take the metric. Since the primed frame is *flat*, the metric there is Minkowski, $\eta_{\mu\nu}$. Then, from the transformation rule, we will have:

$$g = \Lambda^{\mathrm{T}} \eta \Lambda$$

Since the determinant of Λ and Λ^{T} are same and $\det(\eta) = -1$, we obtain:

$$\det(g) = -\det(\Lambda)^2 = -J^2 \implies J = \sqrt{|-\mathfrak{g}|}$$

where we have defined $\mathfrak{g} := \det(g)$, the determinant of the metric tensor for the space. Generalising to n dimensions, we have the volume element as:

$$\mathrm{d}V = \mathrm{d}^n\vec{x}\, \sqrt{|-\mathfrak{g}|}$$

As an example, for 2D polar coordinates, we have:

$$g \equiv \mathrm{diag}(1, r^2) \implies \mathfrak{g} = r^2 \implies \mathrm{d}V = \mathrm{d}r\ \mathrm{d}\theta\ \sqrt{r^2} = r\ \mathrm{d}r\ \mathrm{d}\theta$$

And for 3D spherical coordinates we have,

$$g \equiv \mathrm{diag}(1, r^2, r^2\sin^2\theta) \mathfrak{g} = r^4 \sin^2\theta \implies \mathrm{d}V = \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\phi\ \sqrt{r^4\sin^2\theta} = r^2 \sin\theta\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\phi$$

which are indeed the case as we know from our experience! Thus, this gives a nice expression to deal with the volume elements in some arbitrary space with some arbitrary metric.

Lecture 13: Towards the action for Gravity: Covariant Derivatives

Previously, we focussed on making the integral measure invariant in the action expression. Now, we focus on the function $f(g_{\mu\nu}, \partial_\alpha\partial_\beta g_{\mu\nu})$. Since this contains derivatives, let us first see what happens to derivatives of tensors!

Consider a scalar field ϕ[1]. Then, under a coordinate transformation $\bar{\mathbf{x}} \to \bar{\mathbf{x}}'$, we have $\phi(\bar{\mathbf{x}}) \to \phi'(\bar{\mathbf{x}}')$. Since this is a scalar field, we expect $\phi'(\bar{\mathbf{x}}') = \phi(\bar{\mathbf{x}})$. An example of such a field can be the temperature at each point in space and at each time.

Let us see how the derivative of this scalar field changes under a general coordinate transformation.

$$\partial'_\mu \phi'(\bar{\mathbf{x}}') = \left(\Lambda^{-1}\right)^\alpha{}_\mu \partial_\alpha(\phi(\bar{\mathbf{x}}))$$

This transforms as a tensor of rank $(0,1)$ and is thus a valid tensor. Now, if we take a vector instead, let us see what changes.

$$\begin{aligned}\partial'_\mu V'^\rho &= \left(\Lambda^{-1}\right)^\alpha{}_\mu \partial_\alpha\left(\Lambda^\rho{}_\beta V^\beta\right)\\ &= \left(\Lambda^{-1}\right)^\alpha{}_\mu \Lambda^\rho{}_\beta\left(\partial_\alpha V^\beta\right) + \left(\Lambda^{-1}\right)^\alpha{}_\mu V^\beta \partial_\alpha\left(\Lambda^\rho{}_\beta\right)\end{aligned}$$

If the second term had not been there, the quantity would have transformed as a rank $(1,1)$ tensor but since under a general coordinate transformation, Λ is itself a function of the coordinates, the second term is non-zero in general, which spoils the transformation property.

Since this causes a problem, a potential solution could be to define a new kind of derivative operator, say ∇_μ, such that $\nabla_\mu V^\rho$ becomes a tensor of rank $(1,1)$. Thus, we demand that,

$$\nabla'_\mu V'^\nu = \left(\Lambda^{-1}\right)^\alpha{}_\mu \Lambda^\nu{}_\rho (\nabla_\alpha V^\rho)$$

[1] It gives us a single value of some variable for every point in spacetime

Inverting this expression, we obtain:

$$\nabla_\alpha V^\rho = \left(\Lambda^{-1}\right)^\rho{}_\nu \Lambda^\mu{}_\alpha \left(\nabla'_\mu V'^\nu\right) \tag{12}$$

Moving forward with this would be difficult and hence we make a comment on physical grounds, that, in a *local inertial frame*, where spacetime is flat, the new derivative is equivalent to the normal derivative ∂_μ. Hence, if the coordinate system $\{\vec{x}'\}$ is locally inertial, then

$$\nabla'_\mu = \partial'_\mu$$

Using this in Eq. 12, we have:

$$\begin{aligned}
\nabla_\alpha V^\rho &= \left(\Lambda^{-1}\right)^\rho{}_\nu \Lambda^\mu{}_\alpha \left[\partial'_\mu \left(\Lambda^\nu{}_\beta V^\beta\right)\right] \\
&= \left(\Lambda^{-1}\right)^\rho{}_\nu \partial_\alpha \left[\left(\Lambda^\nu{}_\beta V^\beta\right)\right] \\
&= \left(\Lambda^{-1}\right)^\rho{}_\nu \Lambda^\nu{}_\beta \partial_\alpha V^\beta + \left(\Lambda^{-1}\right)^\rho{}_\nu V^\beta \partial_\alpha \left(\Lambda^\nu{}_\beta\right) \\
&= \delta^\rho{}_\beta \partial_\alpha V^\beta + \Gamma^\rho{}_{\alpha\beta} V^\beta \\
&= \partial_\alpha V^\rho + \Gamma^\rho{}_{\alpha\beta} V^\beta
\end{aligned}$$

Thus, on physical grounds we obtained:

$$\nabla_\alpha V^\rho = \partial_\alpha V^\rho + \Gamma^\rho{}_{\alpha\beta} V^\beta \tag{13}$$

The above defines a new derivative, called the *covariant derivative*. If we use this definition, then we can obtain the correct transformation. Note that the extra term that spoiled the transformation, is already incorporated into the covariant derivative definition through the affine connection. Thus, *covariant derivatives behave nicely*!

NOTE: An easy way to remember this is that two of the indices of the affine connection are placed exactly similar as in the LHS while the other index is summed over. And we have a normal derivative on the RHS too.

If we instead start with a covariant vector, then:

$$\begin{aligned}
\nabla_\alpha V_\beta &= \Lambda^\nu{}_\beta \Lambda^\mu{}_\alpha \nabla'_\mu V'_\mu \\
&= \Lambda^\nu{}_\beta \Lambda^\mu{}_\alpha \partial'_\mu V'_\mu \\
&= \Lambda^\nu{}_\beta \Lambda^\mu{}_\alpha \left[\partial'_\mu \left(\left(\Lambda^{-1}\right)^\rho{}_\nu V_\rho\right)\right] \\
&= \Lambda^\nu{}_\beta \partial_\alpha \left[\left(\Lambda^{-1}\right)^\rho{}_\nu V_\rho\right] \\
&= \underbrace{\Lambda^\nu{}_\beta \left(\Lambda^{-1}\right)^\rho{}_\nu}_{\delta^\rho{}_\beta} \partial_\alpha V_\rho + \Lambda^\nu{}_\beta V_\rho \partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\nu \\
&= \partial_\alpha V_\beta + \Lambda^\nu{}_\beta V_\rho \partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\nu
\end{aligned}$$

Now note the following:

$$\begin{aligned}
& & \left(\Lambda^{-1}\right)^\rho{}_\nu \Lambda^\nu{}_\theta &= \delta^\rho{}_\theta \\
\implies & & \left(\partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\nu\right) \Lambda^\nu{}_\theta &= -\left(\partial_\alpha \Lambda^\nu{}_\theta\right)\left(\Lambda^{-1}\right)^\rho{}_\nu \qquad \text{(Liebniz rule)} \\
\implies & & \left(\partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\nu\right) \Lambda^\nu{}_\theta \left(\Lambda^{-1}\right)^\theta{}_\gamma &= -\left(\partial_\alpha \Lambda^\nu{}_\theta\right)\left(\Lambda^{-1}\right)^\rho{}_\nu \left(\Lambda^{-1}\right)^\theta{}_\gamma \\
\implies & & \partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\gamma &= -\left(\partial_\alpha \Lambda^\nu{}_\theta\right)\left(\Lambda^{-1}\right)^\rho{}_\nu \left(\Lambda^{-1}\right)^\theta{}_\gamma
\end{aligned}$$

Substituting this in the above equation, we obtain:

$$\begin{aligned}\nabla_\alpha V_\beta &= \partial_\alpha V_\beta - \Lambda^\nu{}_\beta \partial_\alpha \left(\Lambda^{-1}\right)^\rho{}_\nu V_\rho \\ &= \partial_\alpha V_\beta - \Lambda^\nu{}_\beta V_\rho (\partial_\alpha \Lambda^\gamma{}_\theta) \left(\Lambda^{-1}\right)^\rho{}_\gamma \left(\Lambda^{-1}\right)^\theta{}_\nu \\ &= \partial_\alpha V_\beta - V_\rho (\partial_\alpha \Lambda^\gamma{}_\theta) \left(\Lambda^{-1}\right)^\rho{}_\gamma \delta^\theta{}_\beta \\ &= \partial_\alpha V_\beta - V_\rho (\partial_\alpha \Lambda^\gamma{}_\beta) \left(\Lambda^{-1}\right)^\rho{}_\gamma \\ &= \partial_\alpha V_\beta - \Gamma^\rho{}_{\alpha\beta} V_\rho\end{aligned}$$

Thus, we obtain:

$$\nabla_\alpha V_\beta = \partial_\alpha V_\beta - \Gamma^\rho{}_{\alpha\beta} V_\rho$$

For any general mixed tensor, for each contravariant index, we have '$+$' and for every covariant index, we have '$-$' in the expression for the covariant derivative, for example,

$$\nabla_\mu T^\alpha{}_\beta = \partial_\mu T^\alpha{}_\beta + \Gamma^\alpha{}_{\mu\rho} T^\rho{}_\beta - \Gamma^\rho{}_{\mu\beta} T^\alpha{}_\rho$$

This new definition of derivative gave us a way to compute quantities nicely, so that the transformation rule remains valid. Later, we would also see physical interpretation this quantity.

Lecture 14: Towards the action for Gravity: Riemann Tensor

Covariant derivatives transform a tensor of rank (p, q) to a tensor of rank $(p, q+1)$ as seen before. Since in our proposed form of the action for gravity, there exists two derivatives of the metric, let us calculate and see what the covariant derivative of the metric is. We have,

$$\nabla_\mu (g_{\alpha\beta}) = \partial_\mu g_{\alpha\beta} - \Gamma^\rho{}_{\mu\alpha} g_{\rho\beta} - \Gamma^\sigma{}_{\mu\beta} g_{\alpha\sigma}$$

Now using Eqn. 6 and contracting metric on the right hand side, we get:

$$g_{\rho\beta}\Gamma^\rho{}_{\mu\alpha} = \frac{1}{2}(\partial_\mu g_{\alpha\beta} + \partial_\alpha g_{\mu\beta} - \partial_\beta g_{\alpha\mu})$$

Using this in the equation above, we get:

$$\nabla_\mu (g_{\alpha\beta}) = \partial_\mu g_{\alpha\beta} - \frac{1}{2}(\partial_\mu g_{\alpha\beta} + \partial_\alpha g_{\mu\beta} - \partial_\beta g_{\alpha\mu}) - \frac{1}{2}(\partial_\mu g_{\beta\alpha} + \partial_\beta g_{\mu\alpha} - \partial_\alpha g_{\beta\mu}) = 0$$

We see that the covariant derivative of the metric is zero, hence the metric tensor is said to be *covariantly conserved*.

Show the following things:

- $\Gamma^\alpha{}_{\mu\alpha} = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g})$
- $\nabla_\mu V^\mu = \frac{1}{\sqrt{-\mathfrak{g}}}\partial_\mu(\sqrt{-\mathfrak{g}} V^\mu)$
- Leibniz rule is satisfied by covariant derivatives, that is,

$$\nabla_\nu(\phi V^\mu) = (\nabla_\nu \phi) V^\mu + \phi(\nabla_\nu V^\mu) \qquad \nabla_\nu(W_\lambda V^\mu) = (\nabla_\nu W_\lambda) V^\mu + W_\lambda(\nabla_\nu V^\mu)$$

▷ For this we will prove the following identity first:

Identity 2:
For any diagonalizable matrix M,

$$\mathrm{Tr}(M^{-1}\mathrm{d}M) = \mathrm{d}(\ln\det M)\,,$$

where $\mathrm{d}M$ is the differential of the matrix.

We first calculate the variation of the RHS:

$$\begin{aligned}
\mathrm{d}\ln(\det M) &= \ln\det(M+\mathrm{d}M) - \ln\det(M) \\
&= \ln\left(\frac{\det(M+\mathrm{d}M)}{\det M}\right) \\
&= \ln\left(\det\left(M^{-1}\right)\det(M+\mathrm{d}M)\right) \\
&= \ln\det\left(\mathbb{1} + M^{-1}\mathrm{d}M\right).
\end{aligned}$$

Now suppose that M is diagonalizable, hence, M can be written as, $M = PDP^{-1}$, with D being the diagonal matrix of eigenvalues $\{\lambda_i\}$. Then we have:

$$\begin{aligned}
\det M = \prod_{i=1}^{N}\lambda_i \implies \ln\det M = \sum_i \ln\lambda_i &= \mathrm{Tr}[\mathbb{1}\times\mathrm{diag}(\ln\lambda_1,\ln\lambda_2,\ldots)] \\
&= \mathrm{Tr}\left[P^{-1}P\mathrm{diag}(\ln\lambda_1,\ln\lambda_2,\ldots)\right] \\
&= \mathrm{Tr}\left[P\mathrm{diag}(\ln\lambda_1,\ln\lambda_2,\ldots)P^{-1}\right] \\
&= \mathrm{Tr}\ln M.
\end{aligned}$$

Using the above two facts, we obtain:

$$\begin{aligned}
\mathrm{d}\ln(\det M) = \ln\det\left(\mathbb{1}+M^{-1}\mathrm{d}M\right) &= \mathrm{Tr}\ln\left(\mathbb{1}+M^{-1}\mathrm{d}M\right) \\
&= \mathrm{Tr}\left[M^{-1}\mathrm{d}M + \mathcal{O}(M^{-1}\mathrm{d}MM^{-1}\mathrm{d}M)\right]
\end{aligned}$$

Now, take the case when $M = g^{\mu\nu} \equiv g$ and $\det(g) = \mathfrak{g}$. Applying the above formula, we obtain:

$$\partial_\mu \ln(-\mathfrak{g}) = \mathrm{Tr}(g\partial_\mu g) = (g\partial_\mu g)^\rho{}_\rho = g^{\rho\sigma}\partial_\mu g_{\rho\sigma} \tag{14}$$

Now, from the definition of Christoffel symbol:

$$\Gamma^\alpha{}_{\mu\alpha} = \frac{1}{2}g^{\alpha\sigma}(\partial_\mu g_{\alpha\sigma} + \partial_\alpha g_{\mu\sigma} - \partial_\sigma g_{\mu\alpha}) = \frac{1}{2}g^{\alpha\sigma}\partial_\mu g_{\alpha\sigma}$$

Note that, if we rename $\sigma \to \alpha$, the last term cancels with the first term above and hence we get the simplified result. Substituting this expression in Eqn. 14, we get:

$$\begin{aligned}
\Gamma^\alpha{}_{\mu\alpha} &= \partial_\mu \ln(-\mathfrak{g}) \\
&= \frac{1}{2}\frac{1}{-\mathfrak{g}}\partial_\mu(-\mathfrak{g}) \\
&= \frac{1}{2\sqrt{-\mathfrak{g}}\sqrt{-\mathfrak{g}}}\partial_\mu(-\mathfrak{g}) \\
&= \frac{1}{\sqrt{-\mathfrak{g}}}\partial_\mu(\sqrt{-\mathfrak{g}})
\end{aligned}$$

Therefore we finally proved the identity,

$$\Gamma^\alpha{}_{\mu\alpha} = \frac{1}{\sqrt{-\mathfrak{g}}}\partial_\mu\sqrt{-\mathfrak{g}}.$$

▷ For the second part, we use the first identity.

$$\begin{aligned}
\nabla_\alpha V^\alpha &= \partial_\alpha V^\alpha + \Gamma^\alpha{}_{\mu\alpha}\ V^\mu \\
&= \partial_\alpha V^\alpha + \frac{1}{\sqrt{-\mathfrak{g}}}\partial_\mu \sqrt{-\mathfrak{g}}\ V^\mu \\
&= \partial_\alpha V^\alpha + \frac{1}{\sqrt{-\mathfrak{g}}}\left(\partial_\mu(\sqrt{-\mathfrak{g}}V^\mu) - \sqrt{-\mathfrak{g}}\partial_\mu V^\mu\right) \\
&= \cancel{\partial_\alpha V^\alpha} - \cancel{\partial_\mu V^\mu} + \frac{1}{\sqrt{-\mathfrak{g}}}\partial_\mu(\sqrt{-\mathfrak{g}}V^\mu)
\end{aligned}$$

Changing $\mu \to \alpha$, we have proved the identity.
▷ For the third part, we just try to match the LHS and RHS. First we do it for the scalar and vector.

$$\begin{aligned}
\nabla_\nu(\phi V^\mu) &= \partial_\nu(\phi V^\mu) + \Gamma^\mu{}_{\nu\sigma}(\phi V^\sigma) \\
&= (\partial_\nu \phi)V^\mu + \phi\partial_\nu V^\mu + \phi\Gamma^\mu{}_{\nu\sigma}V^\sigma \\
&= (\nabla_\nu\phi)V^\mu + \phi\nabla_\nu V^\mu
\end{aligned}$$

For scalars, partial and covariant derivatives are the same and hence we obtained the first term like that. Now, let us show for the two contravectors and covectors.

$$\begin{aligned}
\nabla_\mu(V^\alpha W_\beta) &= \partial_\mu(V^\alpha W_\beta) + (\Gamma^\alpha{}_{\mu\sigma}V^\sigma)W_\beta - \left(\Gamma^\rho{}_{\beta\mu}V^\alpha\right)W_\rho \\
&= \partial_\mu(V^\alpha)W_\beta + V^\alpha\partial_\mu(W_\beta) + (\Gamma^\alpha{}_{\mu\sigma}V^\sigma)W_\beta - \left(\Gamma^\rho{}_{\beta\mu}V^\alpha\right)W_\rho \\
&= W_\beta\{\partial_\mu V^\alpha + \Gamma^\alpha{}_{\mu\sigma}V^\sigma\} + V^\alpha\left\{\partial_\mu W_\beta - \left(\Gamma^\rho{}_{\beta\mu}W_\rho\right)\right\} \\
&= W_\beta(\nabla_\mu V^\alpha) + V^\alpha(\nabla_\mu W_\beta)
\end{aligned}$$

Now, let us delve into an involved calculation. The covariant derivative of the metric is already zero, so in the expression for the action, we cannot put any covariant derivatives of the metric. However, we need to somehow obtain an expression with two normal derivatives of the metric. For that, let us just see what happens when we take two covariant derivatives of a contravector. [1]

$$\begin{aligned}
\nabla_\mu(\nabla_\nu V^\rho) &= \partial_\mu(\nabla_\nu V^\rho) + \Gamma^\rho{}_{\mu\sigma}(\nabla_\nu V^\sigma) - \Gamma^\sigma{}_{\mu\nu}(\nabla_\sigma V^\rho) \\
&= \partial_\mu(\partial_\nu V^\rho + \Gamma^\rho{}_{\nu\lambda}V^\lambda) + \Gamma^\rho{}_{\mu\sigma}(\partial_\nu V^\sigma + \Gamma^\sigma{}_{\nu\lambda}V^\lambda) - \Gamma^\sigma{}_{\mu\nu}(\nabla_\sigma V^\rho)
\end{aligned}$$

Let us now calculate the commutator of this, that is, $[\nabla_\mu, \nabla_\nu]\, V^\rho$. To write in short, we will denote the other symmetric term with $\mu \leftrightarrow \nu$.

$$[\nabla_\mu, \nabla_\nu]\, V^\rho = \partial_\mu(\partial_\nu V^\rho + \Gamma^\rho{}_{\nu\lambda}V^\lambda) + \Gamma^\rho{}_{\mu\sigma}(\partial_\nu V^\sigma + \Gamma^\sigma{}_{\nu\lambda}V^\lambda) - \Gamma^\sigma{}_{\mu\nu}(\nabla_\sigma V^\rho) - (\mu \leftrightarrow \nu)$$

Note that the red term is itself symmetric in μ and ν and hence will be cancelled at once from the above expression. Same goes with the blue term. Thus, we are left with:

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu]\, V^\rho &= \partial_\mu\left(\Gamma^\rho{}_{\nu\lambda}V^\lambda\right) + \Gamma^\rho{}_{\mu\sigma}\partial_\nu V^\sigma + \Gamma^\rho{}_{\mu\sigma}\Gamma^\sigma{}_{\nu\lambda}V^\lambda - (\mu \leftrightarrow \nu) \\
&= \partial_\mu(\Gamma^\rho{}_{\nu\lambda})V^\lambda + \left\{\Gamma^\rho{}_{\nu\lambda}\partial_\mu(V^\lambda) + \Gamma^\rho{}_{\mu\sigma}\partial_\nu V^\sigma\right\} + \Gamma^\rho{}_{\mu\sigma}\Gamma^\sigma{}_{\nu\lambda}V^\lambda - (\mu \leftrightarrow \nu)
\end{aligned}$$

Note the term in the bracket. This term, as a whole, is symmetric under the exchange of μ and ν and hence will be cancelled from the commutator. This significantly reduces the number of terms and we obtain,

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu]\, V^\rho &= \partial_\mu(\Gamma^\rho{}_{\nu\lambda})V^\lambda + \Gamma^\rho{}_{\mu\sigma}\Gamma^\sigma{}_{\nu\lambda}V^\lambda - \partial_\nu(\Gamma^\rho{}_{\mu\lambda})V^\lambda - \Gamma^\rho{}_{\nu\sigma}\Gamma^\sigma{}_{\mu\lambda}V^\lambda \\
&= \left\{\partial_\mu(\Gamma^\rho{}_{\nu\lambda}) + \Gamma^\rho{}_{\mu\sigma}\Gamma^\sigma{}_{\nu\lambda} - \partial_\nu(\Gamma^\rho{}_{\mu\lambda}) - \Gamma^\rho{}_{\nu\sigma}\Gamma^\sigma{}_{\mu\lambda}\right\}V^\lambda
\end{aligned}$$

[1] Note that the quantity in the bracket below, on the LHS, is a $(1,1)$ tensor.

The commutator is a proper tensor since this is difference of two tensors and thus, the right hand side must also be a tensor which we define as:

$$R^{\rho}{}_{\lambda\mu\nu} := \partial_{\mu}\Gamma^{\rho}{}_{\nu\lambda} + \Gamma^{\rho}{}_{\mu\sigma}\Gamma^{\sigma}{}_{\nu\lambda} - \partial_{\nu}\Gamma^{\rho}{}_{\mu\lambda} - \Gamma^{\rho}{}_{\nu\sigma}\Gamma^{\sigma}{}_{\mu\lambda} \tag{15}$$

The above object is a tensor of rank $(1,3)$ and is called the *Riemann tensor*. The tensor represents the curvature of the space, which we will discuss later. We can lower all the indices to obtain a $(0,4)$ tensor as,

$$R_{\rho\lambda\mu\nu} = g_{\rho\sigma}R^{\sigma}{}_{\lambda\mu\nu}$$

NOTE: How to remember this horrenduous term? Well, what works for me is, note that the first lower index always goes as the last lower index in the RHS. There are two terms, one containing a derivative and another one containing the product of the Christoffel symbols.

Start putting all the indices on the RHS in the same order as in LHS (and in the same way, upwards or downwards). Then, ∂ will get second lower index, μ Christoffel gets the other two indices. In product of Christoffels, the first Christoffel gets ρ and μ and then all the up indices gets exhausted, so put the other index down in the next Christoffel. Fill the remaining terms with some dummy index (indicating summing over). And in the next other terms, interchange $\mu \leftrightarrow \nu$. As simple as that...

Some fun facts...

- The Riemann tensor has two normal derivatives of the metric (as the Christoffel symbol contains one derivative and we are taking derivative of the symbols).
- For Minkowski space, all Christoffel symbols vanish and thus $R^{\rho}{}_{\lambda\mu\nu}$ is identically zero. This gives a hint of the tensor's connection with curvatures, since Minkowski space is 'flat' with no curvature.
- If we start with a covariant vector, we have:

$$\begin{aligned} [\nabla_{\mu},\nabla_{\nu}]V_{\rho} = [\nabla_{\mu},\nabla_{\nu}](g_{\rho\lambda}V^{\lambda}) = g_{\rho\lambda}[\nabla_{\mu},\nabla_{\nu}]V^{\lambda} &= g_{\rho\lambda}R^{\lambda}{}_{\sigma\mu\nu}V^{\sigma} \\ &= R_{\rho\sigma\mu\nu}V^{\sigma} \\ &= R_{\rho\sigma\mu\nu}g^{\sigma\theta}V_{\theta} \\ &= R_{\rho}{}^{\theta}{}_{\mu\nu}V_{\theta} \end{aligned}$$

- Since the commutator is anti-symmetric with respect to μ and ν, the Riemann tensor is also anti-symmetric with respect to the last two indices by construction, that is,

$$R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$$

- R is also anti-symmetric with respect to the first two indices (which is not too obvious). For that, we use the commutator on the metric and use the property of commutator for covariant vectors:

$$0 = [\nabla_{\mu},\nabla_{\nu}]g_{\alpha\beta} = R_{\alpha}{}^{\sigma}{}_{\mu\nu}g_{\sigma\beta} + R_{\beta}{}^{\sigma}{}_{\mu\nu}g_{\alpha\sigma} = R_{\alpha\beta\mu\nu} + R_{\beta\alpha\mu\nu} \implies R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu}$$

- $R_{\rho\sigma\mu\nu}$ is perhaps a very big ranked tensor with four indices. We can try to contract it perhaps. Note that, if we contract with the first two or last two, it will give zero, since metric is *symmetric* and Riemann tensor is *anti-symmetric* with respect to the first or last two indices. Thus, let us contract with the first and third index. We obtain a tensor called the *Ricci tensor*, that is,

$$R_{\mu\nu} = g^{\alpha\beta}R_{\alpha\mu\beta\nu}$$

- We can contract the Ricci tensor further to get a *Ricci scalar*:

$$R = g^{\mu\nu}R_{\mu\nu}$$

  Note that the Ricci scalar has two derivatives of the metric and is also a scalar, thus satisfying the condition to be a part of the action.

We are finally in a position to sew together all the components of the action. The function f can be a function of the Ricci scalar, that is, $f \equiv f(R)$. We take the simplest possible case where f is linear in R, that is,

$$f \equiv \alpha(R - 2\Lambda)$$

Here α and Λ[1] are some constants. With this, the action becomes:

$$\mathcal{S}_{\mathsf{G}} = \alpha \int \mathrm{d}x^4 \sqrt{-\mathfrak{g}}[R - 2\Lambda]$$

This is almost fine, we just need to find something more about α. Note that action has dimensions of the angular momentum and the metric is dimensionless. R will have dimensions L^{-2} since it has two derivatives of a dimensionless quantity (metric). Thus,

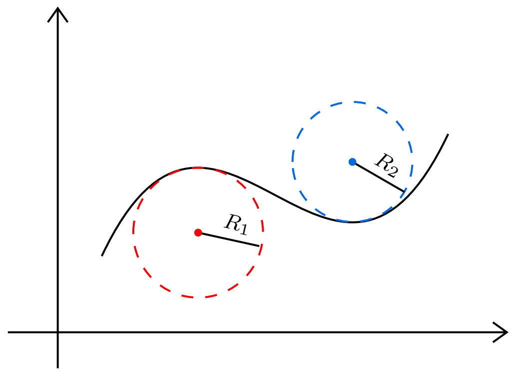
$$\dim \alpha = \frac{\dim \mathcal{S}_{\mathsf{G}}}{[L^4] \dim R} = \frac{ML^2T^{-1}}{L^4 L^{-2}} = \frac{(LT^{-1})^3}{\dim G} = \frac{c^3}{G}$$

So, upto some numerical factor, $\alpha \sim \frac{c^3}{G}$. The final action, which we refer to as the *Einstein-Hilbert action*, is given by:

$$\mathcal{S}_{\mathsf{EH}} = \frac{c^3}{16\pi G} \int \mathrm{d}x^4 \sqrt{-\mathfrak{g}}[R - 2\Lambda] = \frac{c^4}{16\pi G} \int \mathrm{d}t\ \mathrm{d}^3x \sqrt{-\mathfrak{g}}[R - 2\Lambda] \tag{16}$$

Lecture 15: Curvatures

We try to develop a physical interpretation of the Ricci tensor. Note that this contains a *double derivative* of the metric, that is $\partial_\alpha \partial_\beta g_{\mu\nu}$.



Suppose we have some arbitrary smooth curve. Then, at each point of the curve, we can associate a circle whose first and second derivatives match with that of the curve. And there is always a radius associated with such a circle. Thus, to each point on the curve, we can associate a radius r, referred to as the *radius of curvature* of the curve at that point.

Now, for simplicity, let us assume that at a point P on the curve, we have drawn the circle and we shifted our origin to P. Thus, the equation of the circle becomes:

$$x^2 + y^2 = r^2$$

From the equation of the circle, we obtain,

$$y\frac{\mathrm{d}y}{\mathrm{d}x} + x = 0 \implies y\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right) + 1 = 0 \implies y = \frac{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}{-\frac{\mathrm{d}^2y}{\mathrm{d}x^2}}$$

[1] Λ is called the *cosmological constant*. While deriving the equation, Einstein added this since addition of this constant led to a solution which conveyed that the Universe was *static*. However, after Hubble's discovery that the Universe is expanding, this constant was dropped. Then again, the constant was brought back since it was discovered that the Universe was apparently accelerating.

Also, from above, we have

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y}$$

Then we obtain an expression for the radius of the circle:

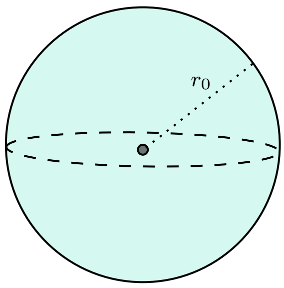
$$\begin{aligned} r = \left|\sqrt{x^2+y^2}\right| &= \left|y\sqrt{1+\left(\frac{x}{y}\right)^2}\right| \\ &= \left|\frac{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}{-\frac{\mathrm{d}^2y}{\mathrm{d}x^2}} \times \sqrt{1+\left(\frac{x}{y}\right)^2}\right| \\ &= \left|\frac{\left[1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{3/2}}{\frac{\mathrm{d}^2y}{\mathrm{d}x^2}}\right| \end{aligned}$$

By choosing an appropriate coordinate system, we can always set the first derivative to zero [1] and we have a result that,

$$r = \left|\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|^{-1}$$

Thus, the second derivative represents the curvature of the curve, since it is intrinsically related to the radius of curvature. And since the Ricci/Riemann tensor contains double derivative of the metric, it somehow represents the curvature of the space. Hence, it is sometimes also called *Riemann curvature tensor*!

Now, for the sake of getting our hands dirty, let us calculate the Ricci scalar of $\mathbb{S}^2$ (sphere), of radius r_0.



For this, we first need to calculate the Riemann tensor, then contract it to get Ricci tensor and then again contract it to get Ricci scalar. Thus, this seems a very involved task and we have to do that carefully!

As calculate before, the metric for the sphere is written as:

$$ds^2 = r_0^2\mathrm{d}\theta^2 + r_0^2\sin^2\theta\mathrm{d}\phi^2$$

From there we get the metric tensor and inverse metric tensor as:

$$g_{\mu\nu} = \mathrm{diag}(r_0^2, r_0^2\sin^2\theta) \qquad g^{\mu\nu} = \mathrm{diag}(1/r_0^2, 1/r_0^2\sin^2\theta)$$

Also, we had earlier found that only two of the Christoffel symbols are non-zero, namely,

$$\Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta \qquad \Gamma^\phi{}_{\theta\phi} = \cot\theta$$

<u>Calculating the Riemann Tensor</u>

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu\Gamma^\rho{}_{\nu\sigma} + \Gamma^\rho{}_{\mu\lambda}\Gamma^\lambda{}_{\nu\sigma} - \partial_\nu\Gamma^\rho{}_{\mu\sigma} - \Gamma^\rho{}_{\nu\lambda}\Gamma^\lambda{}_{\mu\sigma}$$

[1] Note that since first derivative roughly gives us the slope of the tangent, with proper rotation and shifting of the coordinates, we can indeed align our coordinate system along the tangent so that it's slope is zero

Let us take $\rho = \theta, \sigma = \phi, \mu = \theta, \nu = \phi$ first. Then we have:

$$\begin{aligned}
R^{\theta}{}_{\phi\theta\phi} &= \partial_\theta \Gamma^{\theta}{}_{\phi\phi} + \Gamma^{\theta}{}_{\theta\lambda}\Gamma^{\lambda}{}_{\phi\phi} - \partial_\phi \Gamma^{\theta}{}_{\theta\phi} - \Gamma^{\theta}{}_{\phi\lambda}\Gamma^{\lambda}{}_{\theta\phi} \\
&= -\partial_\theta \left(\frac{1}{2}\sin 2\theta\right) + \sin\theta\cos\theta \times \cos\theta \\
&= -\frac{1}{2}\cos 2\theta + \cancel{\sin\theta}\cos\theta \times \frac{\cos\theta}{\cancel{\sin\theta}} \\
&= \sin^2\theta - \cos^2\theta + \cos^2\theta \\
&= \sin^2\theta
\end{aligned}$$

In fact, this is the only non-zero Riemann tensor component. Other components will be zero owing to the anti-symmetric and other nice properties of the Riemann tensor.

Calculating the Ricci Tensor

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}$$

We calculate the components one by one.

$$R_{\theta\theta} = R^{\lambda}{}_{\theta\lambda\theta}$$

Note that we only found out one component of the Riemann tensor. We have to transform all quantitites to that form only. Here, the only non-zero index will be $\lambda = \phi$, thus we get,

$$\begin{aligned}
R_{\theta\theta} &= R^{\phi}{}_{\theta\phi\theta} \\
&= g^{\phi\alpha} R_{\alpha\theta\phi\theta} \\
&= g^{\phi\alpha}(-1)(-1)R_{\theta\alpha\theta\phi} \\
&= g^{\phi\phi} R_{\theta\phi\theta\phi} \qquad (\because g \text{ diagonal, so } \alpha = \phi) \\
&= \frac{1}{r_0^2 \sin^2\theta} g_{\theta\sigma} R^{\sigma}{}_{\phi\theta\phi} \\
&= \frac{1}{r_0^2 \sin^2\theta} r_0^2 R^{\theta}{}_{\phi\theta\phi} \qquad (\because g \text{ diagonal, so } \sigma = \theta) \\
&= 1
\end{aligned}$$

$R_{\theta\phi} = 0$ due to the anti-symmetric property and $R_{\phi\phi} = R^{\lambda}{}_{\phi\lambda\phi} = R^{\theta}{}_{\phi\theta\phi} = \sin^2\theta$. Now, we have obtained all the components of the Ricci tensor and hence, let us move on to the Ricci scalar.

$$R = R^{\mu}{}_{\mu} = g^{\mu\nu} R_{\mu\nu} = \frac{1}{r_0^2} \times 1 + \frac{1}{r_0^2 \sin^2\theta} \times \sin^2\theta = \frac{2}{r_0^2}$$

Thus we obtain that the Ricci scalar for the sphere is $\frac{2}{r_0^2}$. Thus, the Ricci scalar indeed refers to the usual curvature and its inverse represents the radius of curvature!

$$R^{-1} \sim r_0^2$$

Lecture 16: Bianchi Identitites

16.1. Natural Units

Instead of human-defined units, like *metre* and *seconds* are somewhat problematic due since these are not fundamental. We could have easily defined '1 m' to be something else and worked with that. It is better to use some fundamental quantitites, appearing in Nature, to define a system of measurement. In the *natural units*, we use the speed of light and the Planck's constant and put them to be dimensionless and numerically equal to 1, that is,

$$c = 1 \implies 1 \text{ sec} = 3 \times 10^8 \text{ m}$$
$$\hbar = 1 \implies \frac{10^{-34} \text{ kg} - \text{m}^2}{3 \times 10^8 \text{ m}} = 1 \implies 1 \text{ m} \approx 3 \times 10^{42} \text{ kg}^{-1}$$

Note that, since action has the same dimension as $\hbar$, action is actually dimensionless in natural units (which is cool!)

We will see two identitites which will be very useful while deriving the Einstein tensor for the field equations. The starting of the derivations of both is the Jacobi identity, which is a general property satisfied by an *binary operation*.

Identity 3 (Jacobi Identity):
Let $\times$ be a binary operation and x, y, z be arbitrary three elements of a particular mathematical structure. Then the binary operation satisfies the Jacobi Identity if,

$$x \times (y \times z) + z \times (x \times y) + y \times (z \times x) = 0 \iff x \times (y \times z) + \text{cyc.}(x, y, z) = 0$$

The Jacobi identity is satisfied by many binary operations, like the cross-product, the commutator brack, the Poisson bracket, etc[1]. We will consider the Jacobi identity for the commutator of the covariant derivatives here!

16.2. First Bianchi Identity

Let us apply the Jacobi identity on a scalar function ϕ.

$$[\nabla_\rho, [\nabla_\mu, \nabla_\nu]]\,\phi + \text{cyc.}(\rho, \mu, \nu) = 0$$

Now note the following,

$$[\nabla_\mu, \nabla_\nu]\,\phi = \nabla_\mu(\nabla_\nu \phi) - \mu \leftrightarrow \nu = \left[\partial_\mu \partial_\nu \phi - \Gamma^\lambda{}_{\mu\nu} \partial_\lambda \phi\right] - \mu \leftrightarrow \nu$$

Since the term in the bracket is symmetric in μ and ν, it cancels with the second term [2] and hence the commutator acting on a scalar field is zero. With this, the initial equation simplies as,

$$\begin{aligned} 0 &= -\left[\nabla_\mu, \nabla_\nu\right](\nabla_\rho \phi) + \text{cyc.}(\rho, \mu, \nu) \\ &= -R_{\rho\sigma\mu\nu} \partial^\sigma \phi + \text{cyc.}(\rho, \mu, \nu) \\ &= (R_{\sigma\rho\mu\nu} + \text{cyc.}(\rho, \mu, \nu)) \partial^\sigma \phi \end{aligned}$$

Since the scalar field is arbitrary, we have the first identity as:

$$R_{\sigma\rho\mu\nu} + R_{\sigma\nu\rho\mu} + R_{\sigma\mu\nu\rho} = 0$$

Contracting the above with $g^{\lambda\sigma}$, we have:

$$R^\lambda{}_{\rho\mu\nu} + R^\lambda{}_{\nu\rho\mu} + R^\lambda{}_{\mu\nu\rho} = 0$$

[1] All these are particular cases of a broader thing called *Lie Bracket*
[2] Note that in spaces where torsion is there, partial derivatives do not commute and we do not get this result!

With the first Bianchi identity, we can derive a nice property of the Riemann tensor. From the identity, we have:

$$
\begin{aligned}
R_{\rho\sigma\mu\nu} &= -R_{\rho\nu\sigma\mu} - R_{\rho\mu\nu\sigma} \\
&= R_{\nu\rho\sigma\mu} + R_{\mu\rho\nu\sigma} \\
&= -R_{\nu\mu\rho\sigma} - R_{\nu\sigma\mu\rho} - R_{\mu\sigma\rho\nu} - R_{\mu\nu\sigma\rho} \qquad \text{(using Bianchi identity)} \\
&= 2R_{\mu\nu\rho\sigma} + (R_{\sigma\nu\mu\rho} + R_{\sigma\mu\rho\nu}) \\
&= 2R_{\mu\nu\rho\sigma} + (R_{\sigma\nu\mu\rho} + R_{\sigma\mu\rho\nu} + R_{\sigma\rho\nu\mu}) - R_{\sigma\rho\nu\mu} \\
&= 2R_{\mu\nu\rho\sigma} + 0 - R_{\rho\sigma\mu\nu}
\end{aligned}
$$

From this, we obtain that the Riemann tensor is also symmetric under exchange of two pair of indices, that is,

$$R_{(\rho\sigma)(\mu\nu)} = R_{(\mu\nu)(\rho\sigma)}$$

Recall that the Ricci tensor was defined as the contraction of first and third index of the Riemann tensor.

$$R_{\mu\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu} = g^{\beta\alpha} R_{\beta\nu\alpha\mu} = R_{\nu\mu}$$

Hence we obtain that the Ricci tensor is a symmetric tensor.

16.3. Second Bianchi Identity

For this, we apply the Jacobi identity on a vector field V^ρ.

$$
\begin{aligned}
0 &= \nabla_\sigma([\nabla_\mu, \nabla_\nu] V^\rho) - [\nabla_\mu, \nabla_\nu] (\nabla_\sigma V^\rho) + \text{cyc.}(\sigma,\mu,\nu) \\
&= \nabla_\sigma(R^\rho{}_{\lambda\mu\nu} V^\lambda) - R^\rho{}_{\lambda\mu\nu} \nabla_\sigma V^\lambda - R_\sigma{}^\lambda{}_{\mu\nu} \nabla_\lambda V^\rho + \text{cyc.}(\sigma,\mu,\nu) \qquad \text{(applying commutator on } (1,1) \text{ tensor)} \\
&= \Big[\nabla_\sigma(R^\rho{}_{\lambda\mu\nu}) V^\lambda + \cancel{R^\rho{}_{\lambda\mu\nu} \nabla_\sigma(V^\lambda)}\Big] - \cancel{R^\rho{}_{\lambda\mu\nu} \nabla_\sigma V^\lambda} - g^{\beta\lambda} R_{\sigma\beta\mu\nu} \nabla_\lambda V^\rho + \text{cyc.}(\sigma,\mu,\nu) \qquad (\because \text{ Liebniz rule}) \\
&= \nabla_\sigma(R^\rho{}_{\lambda\mu\nu}) V^\lambda + g^{\beta\lambda} R_{\beta\sigma\mu\nu} \nabla_\lambda V^\rho + \text{cyc.}(\sigma,\mu,\nu) \\
&= \Big[\nabla_\sigma(R^\rho{}_{\lambda\mu\nu}) + \text{cyc.}(\sigma,\mu,\nu)\Big] V^\lambda + \Bigg[\cancelto{0}{R^\lambda{}_{\sigma\mu\nu} + \text{cyc.}(\sigma,\mu,\nu)}\Bigg] \nabla_\lambda V^\rho \qquad \text{(first Bianchi identity)} \\
&= (\nabla_\sigma(R_{\rho\lambda\mu\nu}) + \nabla_\nu(R_{\rho\lambda\sigma\mu}) + \nabla_\mu(R_{\rho\lambda\nu\sigma})) V^\lambda
\end{aligned}
$$

Thus, we obtain the second Bianchi identity,

$$\nabla_\sigma(R_{\rho\lambda\mu\nu}) + \nabla_\nu(R_{\rho\lambda\sigma\mu}) + \nabla_\mu(R_{\rho\lambda\nu\sigma}) = 0$$

If we contract the equation with $g^{\rho\mu}$,

$$
\begin{aligned}
0 &= g^{\rho\mu}(\nabla_\sigma(R_{\rho\lambda\mu\nu}) + \nabla_\nu(R_{\rho\lambda\sigma\mu}) + \nabla_\mu(R_{\rho\lambda\nu\sigma})) \\
&= \nabla_\sigma(R^\mu{}_{\lambda\mu\nu}) + \nabla_\nu(R^\mu{}_{\lambda\sigma\mu}) + \nabla^\rho(R_{\rho\lambda\nu\sigma}) \\
&= \nabla_\sigma(R_{\lambda\nu}) - \nabla_\nu(R_{\lambda\sigma}) + \nabla^\rho(R_{\rho\lambda\nu\sigma})
\end{aligned}
$$

Contracting again with $g^{\lambda\sigma}$, we obtain:

$$
\begin{aligned}
0 &= g^{\lambda\sigma}(\nabla_\sigma(R_{\lambda\nu}) - \nabla_\nu(R_{\lambda\sigma}) + \nabla^\rho(R_{\rho\lambda\nu\sigma})) \\
&= \nabla^\lambda R_{\lambda\nu} - \nabla_\nu R + \nabla^\rho R_{\rho\nu} \\
&= \nabla^\lambda R_{\lambda\nu} - \nabla_\nu R + \nabla^\lambda R_{\lambda\nu} \\
&= 2\nabla^\lambda R_{\lambda\nu} - g_{\lambda\nu} \nabla^\lambda R \\
&= \nabla^\lambda \Big(R_{\lambda\nu} - \frac{1}{2} g_{\lambda\nu} R \Big)
\end{aligned}
$$

If we define a new tensor as follows,

$$G_{\lambda\nu} := R_{\lambda\nu} - \frac{1}{2}g_{\lambda\nu}R$$

then, we get the fact that:

$$\nabla^\lambda G_{\lambda\nu} = 0$$

Thus, the tensor $G_{\lambda\nu}$ is covariantly conserved, similar to the metric, as seen before. The tensor is called the *Einstein tensor* and will appear in the field equations to be seen later.

Lecture 17: Einstein Field Equation

In Eq. 16, we had given a form of the Einstein-Hilbert action for gravity. Since the universe consists of matter also, the total action for the Universe is written as:

$$\mathcal{S}[\psi, g] = \mathcal{S}_{\text{EH}}[g] + \mathcal{S}_{\text{M}}[\psi, g]$$

where $\mathcal{S}_{\text{M}}$ denotes the action for any matter field ψ, like electromagnetic field, fermionic and bosonic fields, etc. To obtain the equation of motion for the metric, we have to then vary this action, that is,

$$\delta\mathcal{S} = \frac{1}{16\pi G}\int \mathrm{d}^4x\,\delta\big(\sqrt{-\mathfrak{g}}\; g^{\mu\nu}R_{\mu\nu}\big) + \delta\mathcal{S}_{\text{M}}$$

To find the variation, we need to see some identitites regarding the variation of the metric.

> Show the following things:
>
> - $\delta\mathfrak{g} = \mathfrak{g}\; g^{\mu\nu}\delta g_{\mu\nu} = -\mathfrak{g}\; g_{\mu\nu}\delta g^{\mu\nu}$
> - $\delta(\sqrt{-\mathfrak{g}}) = -\frac{1}{2}\sqrt{-\mathfrak{g}}\; g_{\mu\nu}\delta g^{\mu\nu}$
> - $g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\mu A^\mu$ where $A^\mu = g^{\alpha\beta}\delta\Gamma^\mu{}_{\alpha\beta} - g^{\mu\nu}\Gamma^\rho{}_{\nu\rho}$

▷ To show the first identity, let us use the Jacobi relation for a matrix M that we had derived earlier:

$$\mathrm{Tr}\big(M^{-1}\mathrm{d}M\big) = \mathrm{d}(\ln\det M)$$

Using the formula for the metric g, we obtain the following

$$\mathrm{Tr}\big(g^{-1}\delta g\big) = \frac{1}{\mathfrak{g}}\delta\mathfrak{g} \implies \delta\mathfrak{g} = \mathfrak{g}\big(g^{-1}\delta g\big)^\mu{}_\mu = \mathfrak{g}\; g^{\mu\nu}\delta g_{\nu\mu}$$

Now consider the following obvious identity,

$$g^{-1}g = \mathbb{1}$$

Taking the variation of the above (using product rule) we obtain:

$$\delta g^{-1}\; g + g^{-1}\;\delta g = 0$$

Taking the trace of the above, we obtain the other part of the identity:

$$\mathrm{Tr}\big(g^{-1}\;\delta g\big) = -\,\mathrm{Tr}\big(\delta g^{-1}\; g\big) = -\big(\delta g^{-1}\; g\big)^\mu{}_\mu = -\delta g^{\mu\nu}g_{\nu\mu}$$

▷ For the second identity, using the previous identity we have:

$$\delta\sqrt{-\mathfrak{g}} = -\frac{1}{2\sqrt{-\mathfrak{g}}}\delta\mathfrak{g} = -\frac{1}{2\sqrt{-\mathfrak{g}}}\times\big(-\mathfrak{g}\; g_{\mu\nu}\delta g^{\mu\nu}\big) = -\frac{1}{2}\sqrt{-\mathfrak{g}}\; g_{\mu\nu}\delta g^{\mu\nu}$$

▷ For the third identity, since this contains the Ricci tensor, we will start from its definition as the contracted form of the Riemann tensor:

$$\begin{aligned}\delta R_{\lambda\nu} = \delta R^{\rho}{}_{\lambda\rho\nu} &= \delta\Big(\partial_\rho \Gamma^{\rho}{}_{\nu\lambda} + \Gamma^{\rho}{}_{\rho\sigma}\Gamma^{\sigma}{}_{\nu\lambda} - \partial_\nu \Gamma^{\rho}{}_{\rho\lambda} - \Gamma^{\rho}{}_{\nu\sigma}\Gamma^{\sigma}{}_{\rho\lambda}\Big)\\ &= \partial_\rho \delta\Gamma^{\rho}{}_{\nu\lambda} + \delta(\Gamma^{\rho}{}_{\rho\sigma}\Gamma^{\sigma}{}_{\nu\lambda}) - \partial_\nu \delta\Gamma^{\rho}{}_{\rho\lambda} - \delta(\Gamma^{\rho}{}_{\nu\sigma}\Gamma^{\sigma}{}_{\rho\lambda})\\ &= \partial_\rho \delta\Gamma^{\rho}{}_{\nu\lambda} + \delta\Gamma^{\rho}{}_{\rho\sigma}\Gamma^{\sigma}{}_{\nu\lambda} + \delta\Gamma^{\sigma}{}_{\nu\lambda}\Gamma^{\rho}{}_{\rho\sigma} - \partial_\nu \delta\Gamma^{\rho}{}_{\rho\lambda} - \delta\Gamma^{\rho}{}_{\nu\sigma}\Gamma^{\sigma}{}_{\rho\lambda} - \delta\Gamma^{\sigma}{}_{\rho\lambda}\delta\Gamma^{\rho}{}_{\nu\sigma}\end{aligned}$$

Now, consider the covariant derivative of the variation of the Christoffel symbol which is given as:

$$\nabla_\rho(\delta\Gamma^{\rho}{}_{\nu\lambda}) = \partial_\rho(\delta\Gamma^{\rho}{}_{\nu\lambda}) + \underbrace{\Gamma^{\rho}{}_{\rho\sigma}\delta\Gamma^{\sigma}{}_{\nu\lambda}}_{\text{for } \rho \text{ index}} - \underbrace{\Gamma^{\sigma}{}_{\rho\nu}\delta\Gamma^{\rho}{}_{\sigma\lambda}}_{\text{for } \nu \text{ index}} - \underbrace{\Gamma^{\sigma}{}_{\rho\lambda}\delta\Gamma^{\rho}{}_{\nu\sigma}}_{\text{for } \lambda \text{ index}}$$

$$\nabla_\nu(\delta\Gamma^{\rho}{}_{\rho\lambda}) = \partial_\nu(\delta\Gamma^{\rho}{}_{\rho\lambda}) + \underbrace{\Gamma^{\rho}{}_{\nu\sigma}\delta\Gamma^{\sigma}{}_{\rho\lambda}}_{\text{for } \rho \text{ index}} - \underbrace{\Gamma^{\sigma}{}_{\nu\rho}\delta\Gamma^{\rho}{}_{\sigma\lambda}}_{\text{for } \rho \text{ index}} - \underbrace{\Gamma^{\sigma}{}_{\nu\lambda}\delta\Gamma^{\rho}{}_{\rho\sigma}}_{\text{for } \lambda \text{ index}}$$

Thus, the variation of the Ricci tensor can be written as the difference of the two covariant derivatives:

$$\delta R_{\lambda\nu} = \nabla_\rho(\delta\Gamma^{\rho}{}_{\nu\lambda}) - \nabla_\nu(\delta\Gamma^{\rho}{}_{\rho\lambda})$$

Contracting with the metric $g^{\lambda\nu}$ and noting that the metric is covariantly conserved (i.e, $\nabla_\alpha g^{\mu\nu} = 0$), we have:

$$g^{\lambda\nu}\delta R_{\lambda\nu} = \nabla_\rho(g^{\lambda\nu}\delta\Gamma^{\rho}{}_{\nu\lambda}) - \nabla_\nu(g^{\lambda\nu}\delta\Gamma^{\rho}{}_{\rho\lambda}) = \nabla_\mu(g^{\lambda\nu}\delta\Gamma^{\mu}{}_{\nu\lambda}) - \nabla_\mu(g^{\lambda\mu}\delta\Gamma^{\rho}{}_{\rho\lambda})$$

where we have renamed the dummy indices in the covariant derivative. From this, we finally obtain the identity:

$$g^{\lambda\nu}\delta R_{\lambda\nu} = \nabla_\mu\Big(g^{\lambda\nu}\delta\Gamma^{\mu}{}_{\nu\lambda} - g^{\lambda\mu}\delta\Gamma^{\rho}{}_{\rho\lambda}\Big) = \nabla_\mu A^\mu$$

□

We are finally ready to find the variation of the action to find the equation of motion.

$$\frac{1}{16\pi G}\int \mathrm{d}^4x \Big\{\Big(-\frac{1}{2}\sqrt{-\mathfrak{g}}\; g_{\mu\nu}\delta g^{\mu\nu}\Big)g^{\mu\nu}R_{\mu\nu} + \sqrt{-\mathfrak{g}}\;\delta g^{\mu\nu}R_{\mu\nu} + \sqrt{-\mathfrak{g}}\; g^{\mu\nu}\delta R_{\mu\nu}\Big\} + \delta\mathcal{S}_{\mathsf{M}} = 0$$

The third term in the integral is equal to $\sqrt{-\mathfrak{g}}\; g^{\mu\nu}\delta R_{\mu\nu} = \sqrt{-\mathfrak{g}}\;\nabla_\mu A^\mu = \partial_\mu(\sqrt{-\mathfrak{g}}A^\mu)$ from the identity that we had proved earlier in Lec. 14. This term then becomes a total derivative and vanishes at the boundary. Thus, we are left with the following,

$$\int \mathrm{d}^4x\; \sqrt{-\mathfrak{g}}\Big[-\frac{1}{2}g_{\mu\nu}R + R_{\mu\nu}\Big]\delta g^{\mu\nu} + 16\pi G\delta\mathcal{S}_{\mathsf{M}} = 0$$

Let us define a tensor in the following way using the functional derivative of the action for matter:

$$\mathcal{T}_{\mu\nu} = -\frac{2}{\sqrt{-\mathfrak{g}}}\frac{\delta\mathcal{S}_{\mathsf{M}}}{\delta g^{\mu\nu}} \implies \delta\mathcal{S}_{\mathsf{M}} = \int \mathrm{d}^4x\;\frac{\delta\mathcal{S}_{\mathsf{M}}}{\delta g^{\mu\nu}}\delta g^{\mu\nu} = -\frac{1}{2}\int \mathrm{d}^4x\;\sqrt{-\mathfrak{g}}\;\mathcal{T}_{\mu\nu}\delta g^{\mu\nu}$$

Substituting this above, we obtain:

$$\int \mathrm{d}^4x\; \sqrt{-\mathfrak{g}}\Big[-\frac{1}{2}g_{\mu\nu}R + R_{\mu\nu} - 8\pi G\mathcal{T}_{\mu\nu}\Big]\delta g^{\mu\nu} = 0$$

Since the metric variation is arbitrary, we have the equation of motion as:

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G\mathcal{T}_{\mu\nu}}$$

This equation is called *Einstein's field equation* for material medium. The tensor, which was defined ad-hoc, is the *energy-momentum tensor*. Since $G_{\mu\nu}$ is covariantly conserved, the field equation implies that so is $\mathcal{T}_{\mu\nu}$, that is, $\nabla^\mu\mathcal{T}_{\mu\nu} = 0$. If we also include the *cosmological constant*, then we have the Einstein field equation as:

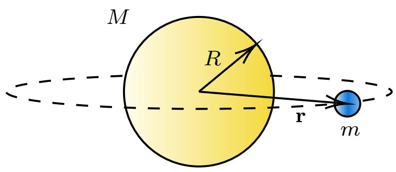
$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G\mathcal{T}_{\mu\nu}$$

The above equation relates the geometry of space to the matter present, in other words, a new description of gravity. We thus hope that Newton's law could be recovered in an appropriate limit from the equation.

17.1. Motion of Planet around the Sun

Consider a planet of mass m orbiting around a star of mass $M \gg m$ and radius R at the origin. In the Newtonian case, the gravitational force and the corresponding potential is given by

$$\mathbf{F} = -\frac{GMm}{r^2}\hat{\mathbf{r}} \qquad V_{\mathrm{N}} = -\frac{GMm}{r}$$



Properties of the Newtonian gravity:

- Since the acceleration has second order derivative in time, the equation has time-reversal symmetry.
- The planets move through a matter free-region in space (for $r > R$) which means that $\mathcal{T}_{\mu\nu}(\mathbf{x}) = 0$ in the exterior part of the star, that is, along the orbit where the planets move.
- Since the potential do not depend on θ and ϕ angles, the potential is spherically-symmetric.
- Since the potential is time-independent, there is no dynamics associated with the field and hence the field is *static*.

Considering the above properties, we must look for a spherically symmetric, time-independent and stationary vacuum solution to the field equation which also enjoys time reversal symmetry. It is kinda obvious that the appropriate coordinates to work with are (t, r, θ, ϕ) which are called the *Schwarzschild coordinates*. In terms of these coordinates, the most general metric is:

$$\begin{aligned} ds^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu &= \left\{g_{tt}\ \mathrm{d}t^2 + 2(g_{tr}\ \mathrm{d}t\ \mathrm{d}r + g_{t\theta}\ \mathrm{d}t\ \mathrm{d}\theta + g_{t\phi}\ \mathrm{d}t\ \mathrm{d}\phi)\right\} \\ &+ \quad \left\{g_{rr}\ \mathrm{d}r^2 + 2(g_{r\theta}\ \mathrm{d}r\ \mathrm{d}\theta + g_{r\phi}\ \mathrm{d}r\ \mathrm{d}\phi)\right\} + \left\{g_{\theta\theta}\ \mathrm{d}\theta^2 + 2g_{\theta\phi}\ \mathrm{d}\theta\mathrm{d}\phi\right\} + g_{\phi\phi}\ \mathrm{d}\phi^2 \end{aligned}$$

Invoking spherical symmetric, we expect that at a particular t_o and fixed radius r_o, the metric should reduce to the metric on $\mathbb{S}^2$ which we had derived earlier. Thus,

$$ds^2\Big|_{r_o,t_o} = r_o^2\ \mathrm{d}\theta^2 + r_o^2 \sin^2\theta\ \mathrm{d}\phi^2$$

From which we can say that the metric elements will be:

$$g_{\theta\phi} = 0 \quad g_{\phi\phi} = r^2\sin^2\theta \quad g_{\theta\theta} = r^2$$

The metric should also be independent of the orientation in θ and ϕ since we expect spherically symmetric solutions. Hence, the metric must remain invariant under $\theta \to -\theta$ and $\phi \to -\phi$, from which we obtain:

$$g_{t\theta} = 0\ g_{r\theta} = 0\ g_{t\phi} = 0\ g_{r\phi} = 0$$

These symmetry considerations significantly reduced the metric. We presently have:

$$ds^2 = g_{tt}\ \mathrm{d}t^2 + 2g_{tr}\ \mathrm{d}t\ \mathrm{d}r + g_{rr}\ \mathrm{d}r^2 + r^2\big(\mathrm{d}\theta^2 + \sin^2\theta\ \mathrm{d}\phi^2\big)$$

From the time-independent consideration, we can say that none of the metric tensor components must depend on time. They should only depend on r (since spherical symmetry is also there). And under time-reversal, the metric should be invariant under $t \to -t$ from which we have $g_{tr} = 0$. Let us denote $g_{tt} = f(r)$ and $g_{rr} = h(r)$. The final expression of the metric, in terms of the two unknowns $f(r)$ and $h(r)$ becomes:

$$\boxed{ds^2 = f(r)\ \mathrm{d}t^2 + h(r)\ \mathrm{d}r^2 + r^2\big(\mathrm{d}\theta^2 + \sin^2\theta\ \mathrm{d}\phi^2\big)} \tag{17}$$

The above metric is the only unique spherically symmetric solution to the vacuum Einstein field equation which is stated in *Birkhoff's Theorem*.

Lecture 18: Some boring calculations

In the previous section, we had found a form of the metric in terms of the unknown functions $h(r)$ and $f(r)$. We thus have diagonal metric and inverse metric:

$$g_{\mu\nu} = \begin{pmatrix} f & & & \\ & h & & \\ & & r^2 & \\ & & & r^2\sin^2\theta \end{pmatrix} \quad \text{and} \quad g^{\mu\nu} = \begin{pmatrix} f^{-1} & & & \\ & h^{-1} & & \\ & & r^{-2} & \\ & & & (r^2\sin^2\theta)^{-1} \end{pmatrix}$$

Since we are interested in finding the *exterior solution*, $T_{\mu\nu}(\mathbf{x}) = 0 \implies G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0$. Let us contract this equation with $g^{\mu\nu}$:

$$\underbrace{g^{\mu\nu}R_{\mu\nu}}_{R} - \frac{1}{2}\underbrace{g^{\mu\nu}g_{\mu\nu}}_{\delta^{\mu}{}_{\mu}=4} R = 0 \implies R = 0 \implies R_{\mu\nu} = 0$$

The goal is now to find the solution to $\boxed{R_{\mu\nu} = 0}$ and the basic starting point is going to be the Christoffel symbols, since the Riemann tensor is written in terms of them. Recall that the Christoffel symbols in terms of the metric were found out to be:

$$\Gamma^{\mu}{}_{\alpha\beta} = \frac{1}{2}g^{\mu\sigma}[g_{\alpha\sigma,\beta} + g_{\sigma\beta,\alpha} - g_{\alpha\beta,\sigma}]$$

We now note a few things:

1. Since metric is diagonal, $\sigma = \mu$ gives the only non-zero value. Since all the symbols are symmetric in lower indices, so we will calculate the value for only one such index pair.
2. If $\alpha = t$ say, then $\Gamma^{\mu}{}_{t\beta} = \frac{1}{2}g^{\mu\mu}\left[g_{t\mu,\beta} + \cancelto{0}{g_{\mu\beta,t}} - g_{t\beta,\mu}\right]$
   The second term is zero due to time independence of metric. If none of μ or β is t, then the Christoffel symbol vanishes since the metric then becomes off-diagonal.
3. If $\alpha = \mu$ then $\Gamma^{\mu}{}_{\mu\beta} = \frac{1}{2}g^{\mu\mu}[g_{\mu\mu,\beta} + \cancel{g_{\mu\beta,\mu}} - \cancel{g_{\mu\beta,\mu}}] = \frac{1}{2}g^{\mu\mu}g_{\mu\mu,\beta} \longrightarrow 0$ if $\beta = t, \phi$. If $\mu \neq \phi$ then for $\beta = \theta$ too, it vanishes.

Using the above will help us to calculate many terms of the Christoffel symbol. Let us start one by one.

▷ $\boxed{\mu = t} \longrightarrow \Gamma^{t}{}_{\alpha\beta} = \frac{1}{2}g^{tt}\left[g_{\alpha t,\beta} + g_{t\beta,\alpha} - \cancelto{0}{g_{\alpha\beta,t}}\right]$

Last term disappears due to time independence. If none of α, β is t, then the terms simply vanish as metric becomes off-diagonal.From this, we can straightaway say,

$$\Gamma^{t}{}_{rr} = \Gamma^{t}{}_{r\theta} = \Gamma^{t}{}_{r\phi} = \Gamma^{t}{}_{\theta\theta} = \Gamma^{t}{}_{\theta\phi} = \Gamma^{t}{}_{\phi\phi} = 0$$

From (3) above, $\Gamma^{t}{}_{tt} = \Gamma^{t}{}_{t\phi} = \Gamma^{t}{}_{t\theta} = 0$. Only remaining symbol is $\Gamma^{t}{}_{tr} = \frac{1}{2} \times f^{-1} \times \frac{\mathrm{d}f}{\mathrm{d}r} = \frac{f'}{2f}$

▷ $\boxed{\mu = r}$

From (3) we can straightaway say that $\Gamma^{r}{}_{rt} = \Gamma^{r}{}_{r\phi} = \Gamma^{r}{}_{r\theta} = 0$. From (2) we can say that $\Gamma^{r}{}_{t\theta} = \Gamma^{r}{}_{t\phi} = 0$. $\Gamma^{r}{}_{\theta\phi} = \frac{1}{2}g^{rr}(g_{r\theta,\phi} + g_{r\phi,\theta} - g_{\theta\phi,r}) = 0$ since off-diagonal. The remaining symbols are:

- $\Gamma^{r}{}_{rr} = \frac{1}{2} \times h^{-1}(h' + h' - h') = \frac{h'}{2h}$
- $\Gamma^{r}{}_{\theta\theta} = \frac{1}{2h} \times (-2r) = -\frac{r}{h}$
- $\Gamma^{r}{}_{\phi\phi} = \frac{1}{2h}\left(-2r\sin^2\theta\right) = -\frac{r}{h}\sin^2\theta$
- $\Gamma^{r}{}_{tt} = \frac{1}{2h} \times (-f) = -\frac{f'}{2h}$

▷ $\boxed{\mu = \theta}$

From (3), we can say straightaway that ${\Gamma^\theta}_{\theta t} = {\Gamma^\theta}_{\theta\phi} = {\Gamma^\theta}_{\theta\theta} = 0$. From (2), we can say ${\Gamma^\theta}_{tr} = {\Gamma^\theta}_{t\phi} = 0$. ${\Gamma^\theta}_{tt} = \frac{1}{2}g^{\theta\theta}[g_{t\theta,t} - g_{tt,\theta}] = 0$ because first term is off-diagonal and last term does not depend on θ. ${\Gamma^\theta}_{r\phi} = 0$ as all terms in the bracket are off-diagonal. ${\Gamma^\theta}_{rr} = 0$ as first two terms will be off-diagonal and last term does not depend on θ. The remaining terms are:

- ${\Gamma^\theta}_{r\theta} = \frac{1}{2}g^{\theta\theta} \times g_{\theta\theta,r} = \frac{1}{2r^2} \times 2r = \frac{1}{r}$
- ${\Gamma^\theta}_{\phi\phi} = \frac{1}{2r^2}\left(-2r^2\sin\theta\cos\theta\right) = -\cos\theta\sin\theta$

▷ $\boxed{\mu = \phi} \longrightarrow {\Gamma^\phi}_{\alpha\beta} = \frac{1}{2}g^{\phi\phi}[g_{\alpha\phi,\beta} + g_{\beta\phi,\alpha} + 0]$

If none of α, β is ϕ then the symbol is zero as element becomes off-diagonal. From (3) we can say ${\Gamma^\phi}_{\phi t} = {\Gamma^\phi}_{\phi\phi} = 0$. The remaining symbols are:

- ${\Gamma^\phi}_{r\phi} = \frac{1}{2r^2\sin^2\theta} \times 2r\sin^2\theta = \frac{1}{r}$
- ${\Gamma^\phi}_{\theta\phi} = \frac{1}{2r^2\sin^2\theta} \times 2r^2\sin\theta\cos\theta = \cot\theta$

The above calculations are now complete. So many of the terms were found out to be zero. These calculations are somewhat boring and are better suited for computers perhaps.

The appendix A contains a Mathematica code to print the Christoffel symbols for different coordinates given any metric. The left index of the final table is the upper index of the Christoffel symbol and the upper index of the table is one of the lower symbol. The table can be divided into four blocks, corresponding to each upper index of the symbol. The other lower symbol is the position of the element in that particular block, like if the element is in the third position, the other index is θ and so on.

Lecture 19: Schwarzschild solution

In the previous section, we calculated the various Christoffel symbols for the metric determining the orbit of planets around a star. We will continue further, towards a solution of the geodesic equation which will tell us about the motion.

As we had found out, the vacuum solutions require $\mathcal{T}_{\mu\nu} = 0$ which implied that $R_{\mu\nu} = 0$. Hence from the Christoffel symbol, we will now find the Ricci tensor and then solve the Einstein field equation. In general, this is an arduous (and boring) task to do it by hand. We will just mention the various non-zero components here [1] and carry on!

$$R_{tt} = -\frac{1}{2h}\left[f'' - \frac{f'h'}{2h} - \frac{(f')^2}{2f} + 2\frac{f'}{r}\right] \tag{18}$$
$$R_{rr} = -\frac{1}{2f}\left[f'' - \frac{(f')^2}{2f} - \frac{f'h'}{2h} - 2\frac{fh'}{rh}\right] \tag{19}$$
$$R_{\theta\theta} = -\frac{r}{h^2 f}\left[hf - \left(\frac{r}{h}\right)^2 + 1\right] \tag{20}$$

Multiplying $2h$ with (18) and multiplying $2f$ with (19) and subtracting them, we get

$$\frac{2h'f}{rh} + \frac{2f'}{r} = 0 \qquad \text{(since from field equation } R_{\mu\nu} = 0\text{)}$$

From this we get

$$h'f + hf' = 0 \implies \frac{\mathrm{d}(hf)}{\mathrm{d}r} = 0 \longrightarrow h(r)f(r) = \mathcal{C}_1$$

[1] The appendix B contains code for symbolically computing the components of the Ricci and Riemann tensor

where $\mathcal{C}_1$ is a constant. Now we introduce some physical argument, that far away from the influence of the star, the metric should transform into the Minkowski metric (in polar coordinates obv)

$$ds^2_{\text{Mink.}} = -\,\mathrm{d}t^2 + \,\mathrm{d}r^2 + r^2\big(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\big)$$

since the effect of gravity becomes negligible as the potential falls off as $\sim r^{-2}$. Thus we have:

$$\left.\begin{aligned}\lim_{r\to\infty} h(r) &= 1\\ \lim_{r\to\infty} f(r) &= -1\end{aligned}\right\} \lim_{r\to\infty} h(r)f(r) = -1 \implies \lim_{r\to\infty} \mathcal{C}_1 = -1 \implies \mathcal{C}_1 = -1 \tag{21}$$

Combining (20) and (21), we get:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r}{h}\right) = 1 \longrightarrow \frac{r}{h} = r - r_s$$

where r_s is an arbitrary constant which will be interpreted as a characterisitc radius (*Schwarzschild radius*) later. Thus, we have essentially solved the problem upto an undetermined constant.

$$h(r) = \frac{r}{r - r_s} \qquad f(r) = \frac{r_s - r}{r}$$

We now write the Christoffel symbols in terms of r_s since we have an expression of $h(r)$ and $f(r)$ now.

$$\begin{aligned}
&\Gamma^t{}_{tr} = \frac{r_s}{2r(r - r_s)} && \Gamma^r{}_{tt} = \frac{r_s(r - r_s)}{2r^3} && \Gamma^r{}_{rr} = \frac{r_s}{2r(r_s - r)}\\
&\Gamma^r{}_{\theta\theta} = r_s - r && \Gamma^r{}_{\phi\phi} = (r_s - r)\sin^2\theta && \Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta\\
&\Gamma^\phi{}_{\theta\phi} = \cot\theta && \Gamma^\phi{}_{\phi r} = \frac{1}{r}
\end{aligned}$$

The motion can now be obtained from the geodesic equation:

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0 \qquad \mu \in \{t, r, \theta, \phi\}$$

▷ Equation for Time

$$\begin{aligned}
&\frac{\mathrm{d}^2t}{\mathrm{d}\tau^2} + \Gamma^t{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0\\
\implies\ &\frac{\mathrm{d}^2t}{\mathrm{d}\tau^2} + 2\frac{r_s}{2r(1 - r_s)}\frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}r}{\mathrm{d}\tau} = 0\\
\implies\ &\frac{\mathrm{d}^2t}{\mathrm{d}\tau^2} + \frac{r_s}{r(1 - r_s)}\frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}r}{\mathrm{d}\tau} = 0
\end{aligned}$$

The above is the equation for time in the given metric. It feels weird since while constructing equation of motions, we had always written quantitites to vary with time but here, time itself varies with proper time. This equation essentially tells us how clock changes with respect to the proper time τ.

Lecture 20: Schwarzschild solution (contd.)

Let us continue our calculation of the geodesic equation. We had already found out the equation for time and the equation for the other coordinates are given as:

▷ $\boxed{\mu = r}$
$$\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} + \frac{r_s}{2r^3}(r - r_s)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \frac{r_s}{2r(r - r_s)}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - (r - r_s)\left(\frac{\mathrm{d}\theta}{\mathrm{d}\tau}\right) - (r - r_s)\sin^2\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2 = 0 \tag{22}$$

▷ $\boxed{\mu = \theta}$
$$\frac{\mathrm{d}^2\theta}{\mathrm{d}\tau^2} + \left(\frac{2}{r}\right)\frac{\mathrm{d}\theta}{\mathrm{d}\tau}\frac{\mathrm{d}r}{\mathrm{d}\tau} - \sin\theta\cos\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2 = 0 \tag{23}$$

▷ $\boxed{\mu = \phi}$
$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\tau^2} + \frac{2}{r}\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\frac{\mathrm{d}r}{\mathrm{d}\tau} + 2\cot\theta\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\frac{\mathrm{d}\theta}{\mathrm{d}\tau} = 0 \tag{24}$$

Now let us do some analysis based on physical argument. So we want our solution to converge to the Newtonian case in some appropriate limit. We know that the Newtonian solution of motion of planet is confined to a plane since angular momentum is conserved for the problem.

Hence, let us also consider a solution restricted to the $\mathrm{x} - \mathrm{y}$ plane, that is $\theta = \frac{\pi}{2}$, say for simplicity. Then Eq. 23 is automatically satisfied since the derivatives vanish and $\cos\theta = 0$ for this specific angle and hence $\theta = \frac{\pi}{2}$ is a valid solution of Eq. 23. Let us now see the other equation for this specific case, starting with Eq. 24,

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\tau^2} + \frac{2}{r}\frac{\mathrm{d}r}{\mathrm{d}\tau}\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = 0 \implies \frac{\mathrm{d}}{\mathrm{d}\tau}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right) = 0 \longrightarrow \boxed{r^2\dot{\phi} = {}^{l}/_{m}}$$

where l is some constant. This gives the interpretation of *angular momentum*, with the subtlety that now, the derivative is with respect to the proper time τ. We thus see that the assumption of planar motion directly handed us the conservation of angular momentum here also.

Let us recall a bit about the Newtonian Lagrangian for the motion of the planet. The general Lagrangian in spherical polar coordinates for the Newtonian case is given by:

$$\mathcal{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right) - V_{\mathsf{N}} \quad \xrightarrow{\theta = \pi/2} \mathcal{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - V_{\mathsf{N}}$$

where V_{N} denotes the Newtonian potential and for this problem, $V_{\mathsf{N}} = -\frac{GM}{r}$. We now find the ELEOM from this:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right) = \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\phi} \longrightarrow \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\phi}) = 0$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{r}}\right) = \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}r} \longrightarrow \ddot{r} = \frac{\partial}{\partial r}\left(\frac{1}{2}mr^2\dot{\phi}^2 - V_{\mathsf{N}}\right) = \frac{\partial}{\partial r}\left(\frac{l^2}{2mr^2} - V_{\mathsf{N}}\right)$$

Thus, we see that the body 'feels' two potentials and we can thus define an *effective potential* as $V_{\mathsf{N}}^{\mathrm{eff}}$ as:

$$\boxed{V_{\mathsf{N}}^{\mathrm{eff}} := V_{\mathsf{N}} + \frac{l^2}{2mr^2} = -\frac{GM}{r} + \frac{l^2}{2mr^2}} \tag{25}$$

Now, let us come back to Eq. 22 and write it in the form below:

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} + \frac{r_s}{2r^2}\left[\left(\frac{r - r_s}{r}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(\frac{r}{r - r_s}\right)\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2\right] - (r - r_s)\frac{l^2}{m^2 r^4} = 0 \tag{26}$$

Now consider the metric in Eq. 17 in the case when $\theta = \pi/2$,

$$ds^2 = -\left(1 - \frac{r_s}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{r_s}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\phi^2$$

Also, recall that in natural units, $\mathrm{d}s^2 = -\mathrm{d}\tau^2$ and rearranging the equation gives us:

$$1 + r^2\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\right)^2 = \frac{r - r_s}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \frac{r}{r - r_s}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2$$

Substituting this in Eq. 22 we get:

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} + \frac{r_s}{2r^2}\left(1 + r^2 \times \frac{l^2}{m^2 r^4}\right) - (r - r_s)\frac{l^2}{m^2 r^4} = 0 \implies \boxed{\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} = \frac{l^2}{m^2 r^3} - \frac{3 r_s l^2}{2m^2 r^4} - \frac{r_s}{2r^2}}$$

From this, we can define an effective potential as ${}^{\mathrm{d}^2 r}/_{\mathrm{d}\tau^2} = -{}^{\partial V_{\mathrm{GR}}^{\mathrm{eff}}}/_{\partial r}$ where,

$$\boxed{V_{\mathrm{GR}}^{\mathrm{eff}} = -\frac{r_s}{2r} + \frac{l^2}{2m^2 r^2}\left(1 - \frac{r_s}{r}\right)} \tag{27}$$

We have to somehow reconcile Eq. 27 and Eq. 25 in an appropriate limit. These look very similar and we can do an identification:

$$r_s = 2GM$$

Also note that if, ${}^{r_s}/_r \ll 1$ then the second term of 27 coincides with the second term of Eq. 25. Hence we conclude that $r \ll r_s$ determines the Newtonian limit. We thus obtain a characteristic length scale r_s which we term as the *Schwarzschild radius*. For a massive object, the solution metric of the Einstein equation becomes:

$$\boxed{ds^2 = -\left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)}$$

And the radial equation becomes in terms of the Schwarzschild radius as:

$$\boxed{\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} = \frac{l^2}{m^2 r^3} - \frac{3GMl^2}{m^2 r^4} - \frac{GM}{r^2}} \tag{28}$$

20.1. GPS and relativity!

GPS (global positioning system) works on the process of *trilateration* where different satellites orbiting the Earth convey their position to the receiver (and thus for each satellite we can draw a sphere of radius equal to the distance between the receiver and the satellite) and based on the intersection of these spheres, we can pinpoint our location.

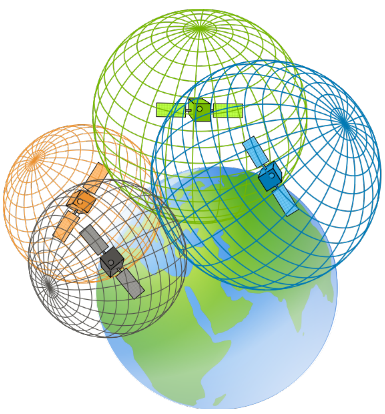


Figure 7: The process of trilateration

The satellites transmit the time of transmission of the radio signal and from there, we indirectly calculate the distance by the difference of the time of transmission and the time of receipt. However, due to the effects of special and general relativity, correction is needed as the clock on Earth and the satellite do not run at the same rate. [1]

Lecture 21: Orbit Equation

In the previous section, we had found the solution to the Einstein metric and in the limit of $GM \to 0$ (i.e, matter-free Universe), the metric reduces to the Minkowski metric. Thus, the Minkowski metric is also a solution of vacuum Einstein field equation.

From Eq. 28, we can see that the term $-\frac{3GMl^2}{2m^2r^4}$ is purely an artifact of general relativity and it is an attractive term (comes with minus sign). For $r \to 0$, this term dominates and can cause problem.

21.1. Orbit of a Planet in GR

Since we are considering spherical polar coordinates, we introduce $u = \frac{1}{r}$, from which we get $\frac{\mathrm{d}u}{\mathrm{d}\phi} = -\frac{1}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\phi}$. The geodesic equations gave us the dynamics of the system, that is, evolution of r and ϕ with τ, however, for finding orbits, we need to find r as a function of ϕ. Then Eq. 28 is written as:

$$\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathrm{d}\phi}{\mathrm{d}\tau}\frac{\mathrm{d}r}{\mathrm{d}\phi}\right) = -\frac{GM}{r^2} + \frac{l^2}{m^2r^3} - \frac{3GMl^2}{m^3r^4}$$

Recall that we had, $\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{l}{mr^2}$ and $\frac{\mathrm{d}r}{\mathrm{d}\phi} = -r^2\frac{\mathrm{d}u}{\mathrm{d}\phi}$, substituting which we obtain:

$$-\frac{l^2u^2}{m^2}\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} = -GMu^2 + \frac{l^2u^3}{m^2} - \frac{3GMl^2}{m^2}u^4$$

Introducing $u_0 = {GMm^2}/{l^2} = {1}/{r_0}$, we get a simplified equation:

$$\boxed{\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = u_0 + 3GMu^2}$$

Let us parameterise this solution by ε such that for $\varepsilon = 0$, we have the Newtonian limit and for $\varepsilon = 1$ we obtain the GR solution. Thus,

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = u_0 + 3GM\varepsilon u^2 \quad \longrightarrow \text{solution:} u = u_{\mathrm{N}} + \varepsilon u_{\mathrm{E}} \tag{29}$$

▷ **Newtonian Solution** $\longrightarrow \varepsilon = 0$

Eq. 29 becomes the simple harmonic oscillators equation:

$$\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}(u - u_0) + (u - u_0) = 0$$

whose solution is given by $u(\phi) = u_0 + \mathcal{A}\cos(\phi - \phi_0)$. We can take $\phi_0 = 0$. We write this in an alternative form as:

$$u_N = \frac{1}{r_N} = \frac{1}{r_0}[1 + e\cos\phi] \tag{30}$$

where $e := \mathcal{A}r_0$ is defined to be the *eccentricity* of the orbit.

▷ $e = 0$: circular orbit ▷ $e \in (0,1)$: elliptic orbit ▷ $e = 1$: parabolic orbit ▷ $e > 1$: hyperbolic orbit

[1] Due to SR, clock in the satellite run slower ($\sim 7\mu s$/day) while due to GR, clock in a lower gravitational field move faster ($\sim 45\mu s$/day), thus there is an error offset of $+38\mu s$/day which is corrected by the GPS.

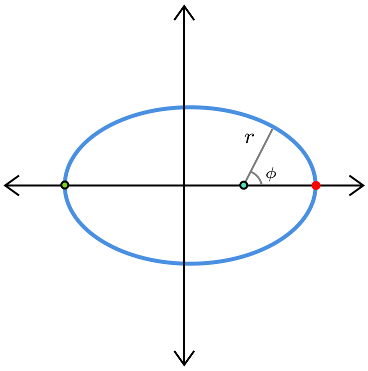


Figure 8: Elliptical orbit of planet showing the position of perihelion in red

.

Consider the case of elliptic orbit, with Sun at one of the focus. We consider two special points, called *perihelion* and *aphelion* as the nearest and farthest point to the Sun respectively. Hence to find their position, we need to extremise $u(\phi)$ with respect to ϕ. Doing this, we obtain:

$$\frac{\mathrm{d}u}{\mathrm{d}\phi} = -\frac{e}{r_0}\sin\phi = 0 \implies \phi = 2n\pi \;\; n = 0, 1, 2\ldots$$

We consider only the perihelion position here. Thus, the Newtonian theory says that perihelion occurs at exactly the same position always. So, after one period of revolution, the planet comes back to the original position (since $\phi = 2n\pi$ all correspond to the same point) however, it was observed that the perihelion shifts gradually.

▷ **GR corrections to the orbit**

Substituting $u = u_{\mathsf{N}} + \varepsilon u_{\mathsf{E}}$ as specified earlier, we have:

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\phi^2} + u_{\mathsf{N}} + \varepsilon u_{\mathsf{E}} = u_0 + 3GM\varepsilon(u_{\mathsf{N}} + \varepsilon u_{\mathsf{E}})^2 \implies u_0 + \varepsilon\frac{\mathrm{d}^2 u_{\mathsf{E}}}{\mathrm{d}\phi^2} + u_{\mathsf{E}} = u_0 + 3GM\varepsilon(u_{\mathsf{N}} + \varepsilon u_{\mathsf{E}})^2$$

Upto $\mathcal{O}(\varepsilon)$, the RHS is just $3GMu_{\mathsf{N}}^2\varepsilon$ and hence we have:

$$\frac{\mathrm{d}^2 u_{\mathsf{E}}}{\mathrm{d}\phi^2} + u_{\mathsf{E}} = 3GMu_{\mathsf{N}}^2 = 3GMu_0^2(1 + e\cos\phi)^2$$

Let us define $u_{\mathsf{E}}^o := 3GMu_0^2$. Now, for many stars, the eccentricity is close to zero, so we can assume $e \ll 1$ which allows us to expand the square term linearly. We get:

$$\frac{\mathrm{d}^2 u_{\mathsf{E}}}{\mathrm{d}\phi^2} + u_{\mathsf{E}} = u_{\mathsf{E}}^o(1 + 2e\cos\phi) \tag{31}$$

Assume an ansatz of $u_E = u_{\mathsf{E}}^o + b\phi\sin\phi$ and note that:

$$u' = b(\phi\cos\phi + \sin\phi)$$
$$u'' = b(2\cos\phi - \phi\sin\phi)$$

Substituting this in Eq. 31, we get:

$$b(2\cos\phi - \phi\sin\phi + \phi\sin\phi) + u_{\mathsf{E}}^o = u_{\mathsf{E}}^o(1 + 2e\cos\phi) \implies \boxed{b = eu_{\mathsf{E}}^o}$$

Then upto $\mathcal{O}(\varepsilon)$ the equation of orbit in GR $(\varepsilon = 1)$ becomes:

$$\frac{1}{r(\phi)} = u_{\mathsf{N}} + u_{\mathsf{E}} = u_0(1 + e\cos\phi) + \left[3GMu_0^2(1 + e\phi\sin\phi)\right] \tag{32}$$

Now, suppose we want to calculate the perihelion for this equation. Ofcourse the perihelion is not going to be some nice expression. However, based on our intuition, we can expect that the shift might not be much from the Newtonian case, hence we assume the perihelion position to be:

$$\phi_n = 2n\pi + \delta\phi$$

Now, we extremise the equation with respect to this ϕ and put the above expression for the corrected perihelion.

$$\frac{\mathrm{d}u}{\mathrm{d}\phi} = -u_0 e \sin\phi + 3GMu_0^2 e(\sin\phi + \phi\cos\phi)\bigg|_{\phi_n} = 0$$

Substituting the value of ϕ_n and noting that $\sin(2n\pi + x) = x$, we get:

$$\sin(\delta\phi)(1 - 3GMu_0) - 3GMu_0(2n\pi + \delta\phi)\cos(\delta\phi) = 0$$

Upto $\mathcal{O}(\delta\phi)$, we obtain $\sin(\delta\phi) \approx \delta\phi$ and $\cos(\delta\phi) \approx 1$. Substituting this above we get:

$$\delta\phi(1 - 6GMu_0) = 6GMu_0 n\pi \implies \delta\phi = \frac{6GMu_0}{(1 - 6GMu_0)} n\pi \simeq \frac{6GM^2m^2}{l^2} n\pi$$

Thus, we see that the shift in the perihelion is by a finite angle given by:

$$\delta\phi = \frac{6GM^2m^2}{l^2} n\pi \longrightarrow \boxed{\Delta\phi = \delta\phi_{n+1} - \delta\phi_n = \frac{6GM^2m^2}{l^2}}$$

So after one period, the planet does not come back to its original position, indicating the orbit must be changing with time. For Mercury, $\Delta\phi \simeq 43''$ per century which is a really, really minute thing[1], however, the effect is noticeable since we have been collecting data for centuries and also with increasing precision of modern instruments, it is evident.

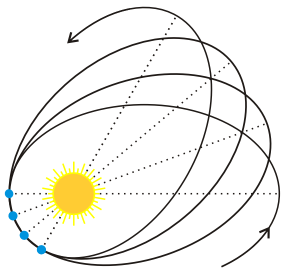


Figure 9: Diagram showing the perihelion shift of a planet

21.2. Black Hole

Note the potential for GR, which has an attractive $\frac{1}{r^3}$ term which dominates for $r \to 0$ and thus, there exists a point of no return, that is, if $r < r_o$, then we cannot come back as $V_{\text{GR}}^{\text{eff}}$ is unbounded from below.

[1] That's what she said 😉

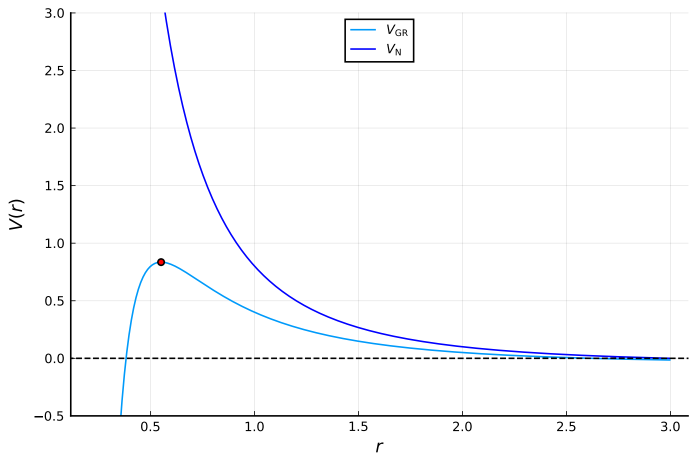


Figure 10: Comparison between Newtonian and relativisitc effective potentials. The red marker denotes the 'point of no return'

We know that in an inertial frame, photons (light) follows a *null trajectory*, that is, $\mathrm{d}s^2 = 0$. So suppose we are not in an inertial frame, however due to equivalence principle, we can always go to a locally inertial frame, calculate $\mathrm{d}s^2$ which will be zero and since $\mathrm{d}s^2$ is a scalar, it will be zero in all other frames.

Suppose a radial motion of the photon in presence of a point mass M at the origin. Then $\mathrm{d}\theta = \mathrm{d}\phi = 0$ and $ds^2 = 0$ and the Schwarzschild metric becomes:

$$0 = \left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 \implies \mathrm{d}t^2 = \frac{r^2}{(r-r_s)^2}\mathrm{d}r^2$$

▷ Time taken for light to go from distance r_o to r_s where $r_o < r_s$

The above relation has two roots, however, since we are interested in $r_o < r_s$ case, we take $\mathrm{d}t = \frac{r}{r_s - r}\mathrm{d}r$. Integrating this, we get:

$$\Delta t = \int_{r_o}^{r_s} \frac{r}{r_s - r}\mathrm{d}r = \int_{r_s}^{r_o}\left(1 - \frac{r_s}{r_s - r}\right)\mathrm{d}r = \left(r + r_s \ln\left(r_s - r\right)\right)\Big|_{r_s}^{r_o} = (r_o - r_s) + r_s\left[\ln(r_s - r_o) - \lim_{x\to r_s}\ln(r_s - x)\right]$$

The red term diverges and as a result, the entire expression of Δt is divergent. In other words, light would take an infinite amount of time to read $r_s = 2GM$ starting from any arbitrary $r_o < r_s$. Hence, from inside the Schwarzschild radius, even light cannot escape!

We conclude that there is a region of radius $r < r_s$ from where even light cannot escape. The surface $r = r_s$ is called the *event horizon* and such an object is called a *black hole*.

Note that blackhole is possible only if $R < r_s$, that is, there exists a limit for the radius of the object. If $R > r_s$, then the solution no longer remain an 'exterior' solution and all the assumptions become invalid since now, $\mathcal{T}_{\mu\nu} \neq 0$ for $r < r_s < R$. If $M = M_\odot$ (where $M_\odot$ denotes the *solar mass*), then blackhole is only possible if,

$$\frac{2GM_\odot}{Rc^2} > 1 \implies R < \frac{2GM_\odot}{c^2} = \frac{2 \times 6.67 \cdot 10^{-11} \times 2 \cdot 10^{30}}{(3 \cdot 10^8)^2} \approx 2964 \text{ m}$$

Thus, if our Sun could be condensed to less than $3\ \mathrm{km}$, then it would have been a black-hole. Every galaxy is known to host a black hole at its centre, around which massive stars are observed to rotate. However, since no light can come out of the black hole, we cannot directly observe it. This calculation, even though seem to be a mathematical artifact, has been confirmed through indirect observation of black hole, by using say interferometry of the surroundings of the event horizon (from where light can escape) or gravitational waves.

Lecture 22: Interior Solution

We will first see a bit more applications of GR and then move on to the interior Schwarzschild solution, where the stress-energy tensor becomes non-zero.

22.1. Time Dilation

Consider two asymptotic observers (that is, farway observers from the star: $r \to \infty$) in two satellites A and B with highly precise and synchronised atomic clocks. Now, let one of them (say B) be fixed and the other observer moves radially towards the mass M and then stops. Then the proper time time for observer A is given as [1]:

$$\mathrm{d}\tau_{\mathsf{A}}^2 = -\left(1 - \frac{2GM}{r_{\mathsf{A}}}\right)\mathrm{d}t^2$$

Note that, since at each point the metric is varying, we cannot equate the proper time at A and B, even though it might be tempting to do so since these are scalars. Also, we have $r \to \infty \implies \mathrm{d}\tau_{\mathsf{A}}^2 = -\mathrm{d}t^2 = \mathrm{d}\tau_{\mathsf{B}}^2$ Thus, we get a relation:

$$\boxed{\frac{\mathrm{d}\tau_{\mathsf{A}}}{\mathrm{d}\tau_{\mathsf{B}}} = \sqrt{1 - \frac{2GM}{r_{\mathsf{A}}}} < 1}$$

The same argument could be done at some other point C and we can thus compare between proper times at any two points. In general, if $r_1 > r_2$ and if a certain proper time interval is measured as $\mathrm{d}\tau_1$ ($\mathrm{d}\tau_2$) at height r_1 (r_2), then $\mathrm{d}\tau_2 < \mathrm{d}\tau_1$.

We thus conclude that *time runs slower near a massive body!* [2]. Close to a black hole ($r \sim r_s$), then time almost freezes as the factor becomes very close to zero (as if, the time never passes). *Black holes can thus make us immortal, however, we will be shredded to pieces near black holes due to high gravitational field!*

22.2. Gravitational redshift

As a direct consequence of time dilation, another pronounced effect is seen. From the time dilation we can find the frequency as inverse of time period, that is, $\nu \sim 1/\Delta\tau$. Suppose a photon is at some distance r and it travels to us at $r \to \infty$. Then,

$$\frac{\nu_\infty}{\nu_r} = \frac{\Delta\tau_r}{\Delta\tau_\infty} = \sqrt{1 - \frac{2GM}{r}}$$

Since $\frac{\nu_\infty}{\nu_r} = \frac{\lambda_r}{\lambda_\infty}$ where λ denotes the wavelength, we find

$$\frac{\lambda_r}{\lambda_\infty} = \sqrt{1 - \frac{2GM}{r}} \implies \frac{\lambda_\infty}{\lambda_r} = \frac{1}{\sqrt{1 - \frac{2GM}{r}}} > 1$$

Thus, the wavelength observed at infinity is longer than the wavelength at r. This is called *red shift* since a the wavelength shifts towards the 'red' end of the spectrum.

22.3. Interior Solution

We now focus on the interior solutiono of the Schwarzschild metric for a spherically symmetric body. In this case, $\mathcal{T}_{\mu\nu} \neq 0$. Recall that the general metric had the form

$$\mathrm{d}s^2 = f(t)\mathrm{d}t^2 + h(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \qquad \text{where } \mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$$

[1] Since we are measuring the time with respect to stationary clocks, thus only the time component remains in the metric
[2] Thus, to slow your ageing, make a fat friend!

We define $f(r) = -e^{2\Phi}$ and $h(r) = e^{2\nu}$ where Φ and ν are some unknown functions of r. Note that the sign for the functions have been arbitrarily chosen. Opposite signs are chosen since space and time should be distinguished somehow and the time part, we saw was negative atleast for the Minkowski case. Hence this choice!

Now, we use the numerical computation (see Appendix C) to find the Christoffel symbols and Einstein tensor. The required non-zero components are:

$$\Gamma^r{}_{tt} = e^{2(\Phi-\nu)}\phi' \qquad \begin{aligned} G_{tt} &= \tfrac{1}{r^2}e^{2(\Phi-\nu)}(-1+2r\nu'+e^{2\nu}) \\ G_{rr} &= \tfrac{1}{r^2}(1-e^{2\nu}+2r\phi') \end{aligned}$$

Recall that the Einstein field equation is,

$$G_{\mu\nu} = 8\pi G\mathcal{T}_{\mu\nu}$$

Now, we will assume a *perfect fluid idealisation* where matter is modelled as a fluid without viscosity. Using this, we can show a specific form of the stress energy tensor.

$$\mathcal{T}_{\mu\nu} = (\rho+P)u_\mu u_\nu + Pg_{\mu\nu} \qquad \begin{aligned} \rho(\mathbf{x}) &: \text{energy density} \\ P(\mathbf{x}) &: \text{pressure} \\ u^\mu(\mathbf{x}) &: \text{four velocity of fluid} \end{aligned}$$

Now, assume a static spacetime, that is, the fluid on average does not move. Hence the four velocity becomes:

$$u^\mu = (u^0, 0, 0, 0)$$

As specified earlier, $g_{\mu\nu}u^\mu u^\nu = -1$, we then have

$$-e^{2\Phi}(u^0)^2 = -1 \implies \boxed{u^0 = e^{-\Phi}} \implies u_0 = -1/u^0 = -e^\Phi$$

Then the components of the stress-energy tensor become:

$$\triangleright\ \mathcal{T}_{00} = (\rho+P)e^{2\Phi} - \underbrace{e^{2\Phi}}_{g_00}P = \rho e^{2\Phi} \qquad \triangleright\ \mathcal{T}_{11} = Pg_{11} = Pe^{2\nu}$$

Then, using the Einstein equation we get:

$$\begin{aligned} G_{tt} = 8\pi G\mathcal{T}_{00} \longrightarrow \quad & \frac{1}{r^2}e^{2(\Phi-\nu)}(-1+2r\nu'+e^{2\nu}) = 8\pi G\rho e^{2\Phi} \\ \implies & (1-e^{-2\nu}+2r\nu' e^{-2\nu}) = 8\pi G\rho r^2 \\ \implies & \boxed{\frac{\mathrm{d}}{\mathrm{d}r}\big(r(1-e^{-2\nu})\big) = 8\pi G\rho r^2} \end{aligned}$$

Now suppose we define a new function $\mathcal{M}(r) := \frac{r(1-e^{-2\nu})}{2G}$. Substituting this above we get:

$$\boxed{\frac{\mathrm{d}\mathcal{M}}{\mathrm{d}r} = 4\pi r^2\rho} \tag{33}$$

This is so much similar like the *mass* of a thin shell. Indeed, if we write

$$e^{-2\nu} = \left[1 - \frac{2G\mathcal{M}(r)}{r}\right]$$

then, at $r \gtrapprox R$, this must be equal to $h(r)$ for the exterior solution case which was equal to $\left(1-\frac{2GM}{r}\right)^{-1}$. Thus, at $r = R$, $\mathcal{M}(r)$ is identified with the mass of the object, that is, $\mathcal{M}(R) = M$. This is okay, since the metric must be continuous at the boundary and the interior and exterior solution must match at $r = R$. Hence, this quantity has some kind of interpretation of mass.

Lecture 23: Interior Solution (contd.)

In the previous section, we started discussing about the interior solutions for Einstein's equation. Let us now start with the equation for r,

$$G_{rr} = 8\pi G \mathcal{T}_{rr} \implies \frac{1}{r^2}\left[1 - e^{2\nu} + 2r\phi'\right] = 8\pi G P e^{2\nu}$$

as u_r component of the four-velocity is zero. Multiplying the above equation by r^2 and $e^{-2\nu}$ we get

$$\underbrace{(e^{-2\nu} - 1)}_{-\frac{2G\mathcal{M}(r)}{r}} + 2r\frac{\mathrm{d}\phi}{\mathrm{d}r}e^{-2\nu} = 8\pi G r^2 \rho$$

Multiplying further by r and simplifying, we have the final expression

$$\boxed{\frac{\mathrm{d}\phi}{\mathrm{d}r} = G\frac{\left[\mathcal{M}(r) + 4\pi r^3 P\right]}{r(r - 2G\mathcal{M}(r))}} \tag{34}$$

Show that for the interior metric, $\nabla_\mu \mathcal{T}^{\mu\nu} = 0$ leads to,

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -(\rho + P)\frac{\mathrm{d}\Phi}{\mathrm{d}r} \tag{35}$$

Proof. Let us first calculate the stress-energy tensor assuming the perfect fluid condition,

$\triangleright\ \mathcal{T}^{rr} = Pe^{-2\nu}$ $\quad\triangleright\ \mathcal{T}^{\theta\theta} = \frac{P}{r^2}$ $\quad\triangleright\ \mathcal{T}^{\phi\phi} = \frac{P}{r^2\sin^2\theta}$ $\quad\triangleright\ \mathcal{T}^{tt} = (P+\rho)e^{-2\Phi} - Pe^{-2\Phi} = \rho e^{-2\Phi}$

Then applying the conservation criterion for the stress-energy tensor, we have

$$\begin{aligned} 0 = \nabla_\mu \mathcal{T}^{\mu r} &= \partial_\mu \mathcal{T}^{\mu r} + \Gamma^r{}_{\mu\sigma}\mathcal{T}^{\mu\sigma} + \Gamma^\mu{}_{\mu\sigma}\mathcal{T}^{\sigma r} \\ &= \partial_r\left(Pe^{-2\nu}\right) + \Gamma^r{}_{rr}\mathcal{T}^{rr} + \Gamma^r{}_{\theta\theta}\mathcal{T}^{\theta\theta} + \Gamma^r{}_{\phi\phi}\mathcal{T}^{\phi\phi} + \Gamma^r{}_{tt}\mathcal{T}^{tt} + \Gamma^\mu{}_{\mu r}\mathcal{T}^{rr} \end{aligned}$$

Expanding the last term we get the following

$$\Gamma^\mu{}_{\mu r} = \Gamma^t{}_{tr} + \Gamma^r{}_{rr} + \Gamma^\theta{}_{\theta r} + \Gamma^\phi{}_{\phi r} = \Phi' + \nu' + \frac{1}{r} + \frac{1}{r} = \Phi' + \nu' + \frac{2}{r}$$

Then, we putting all these together we get

$$\begin{aligned} 0 &= \partial_r\left(Pe^{-2\nu}\right) + \left(P\nu' e^{-2\nu}\right) + \left(-re^{-2\nu}\frac{P}{r^2}\right) + \left(-r\sin^2\theta e^{-2\nu}\frac{P}{r^2\sin^2\theta}\right) + \Phi' e^{2(\Phi-\nu)}\rho e^{-2\Phi} + \left(\Phi' + \nu' + \frac{2}{r}\right)Pe^{-2\nu} \\ &= \left(\frac{\mathrm{d}P}{\mathrm{d}r} - 2\nu' P\right)e^{-2\nu} + \left(P\nu' e^{-2\nu}\right) - \left(e^{-2\nu}\frac{2P}{r}\right) + \Phi' e^{-2\nu}\rho + \left(\Phi' + \nu' + \frac{2}{r}\right)Pe^{-2\nu} \\ &= \left\{\left(\frac{\mathrm{d}P}{\mathrm{d}r} - \cancel{2\nu' P}\right) + \cancel{(P\nu')} - \cancel{\left(\frac{2P}{r}\right)} + \Phi'\rho + \left(\Phi' + \cancel{\nu'} + \cancel{\frac{2}{r}}\right)P\right\}e^{-2\nu} \\ &= \left\{\frac{\mathrm{d}P}{\mathrm{d}r} + \Phi'(\rho + P)\right\}e^{-2\nu} \end{aligned}$$

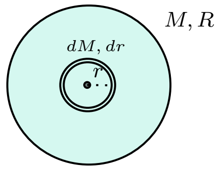
We have used the form of the Christoffel symbols from the code given in the appendix. Eq. 33, 34 and 35 are together called the *Tolman-Oppenheimer-Volkoff* equations which are used to model an isotropic, spherically symmetric body in hydrostatic equilibrium.

23.1. Newtonian Analogy

Consider a spherical mass M and radius R with a thin shell at distance r, which collapses due to gravitational attraction. Then, we can write the *mass-balance* equation as,

$$\frac{\mathrm{d}M(r)}{\mathrm{d}r} = 4\pi r^2 \rho(r)$$

where $\rho(r)$ is the density at distance r. We can also write the hydrostatic equilibrium equation, for which we need to balance the pressure gradient force.



Balancing the pressure, we have the hydrostatic equilibrium,

$$4\pi r^2 (P(r + \mathrm{d}r) - P(r)) = -\frac{GM(r)\mathrm{d}M}{r^2} \implies \frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM(r)\rho}{r^2}$$

Now, from the Newtonian potential we can write

$$\Phi_{\mathsf{N}} = -\frac{GM}{r} \implies \frac{\mathrm{d}\Phi_{\mathsf{N}}}{\mathrm{d}r} = \frac{GM}{r^2}$$

understand what is happening?

23.2. Boundary Matching

We had previously stated also, that the metric should be continuous across the boundary from which we had obtained $\mathcal{M}(R) = M$ where M is the mass of the star. We can also match the component of g_{tt} from which we get

$$e^{2\Phi} = \left(1 - \frac{2GM}{R}\right) \implies \Phi(R) = -\frac{GM}{R} \quad \text{if } \Phi(R) \ll 1$$

Lecture 24: Gravitational Waves

We will now look at a prediction which was purely based on general relativity, having no Newtonian explanation. For that, we have to consider a linearised gravity, that is, some weak perturbation to the flat-spacetime.

Consider a metric which is 'almost Minkowski', parameterised by ε

$$g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon h_{\mu\nu}$$

where $h_{\mu\nu}$ is the perturbation and ε controls the strength of the perturbation and we will consider upto $\mathcal{O}(\varepsilon^1)$. To this end, we will assume that the indices are lowered/raised by the usual Minkowski metric and not by $g^{\mu\nu}$. First we need to find the inverse of the metric and for that, consider an ansatz which is linear in ε

$$g^{\mu\nu} = \eta^{\mu\nu} + \varepsilon \omega^{\mu\nu}$$

Then substituting this in the inverse equations, we obtain

$$\begin{aligned}
\delta^\mu{}_\alpha &= g^{\mu\nu} g_{\nu\alpha} \\
&= \delta^\mu{}_\alpha + \varepsilon(\eta^{\mu\nu} h_{\nu\alpha} + \omega^{\mu\nu}\eta_{\nu\alpha}) + \mathcal{O}(\varepsilon^2) \\
&= \delta^\mu{}_\alpha + \varepsilon(h^\mu{}_\alpha + \omega^\mu{}_\alpha) + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

Thus, upto first order we obtain $h^{\mu}{}_{\alpha} = -\omega^{\mu}{}_{\alpha}$, hence the inverse metric upto the leading order becomes,

$$\boxed{g^{\mu\nu} = \eta^{\mu\nu} - \varepsilon h^{\mu\nu}}$$

We can now proceed to calculate other quantities. Let us start with Christoffel symbols, which takes the form

$$\Gamma^{\mu}{}_{\alpha\beta} = \Gamma^{\mu^{(0)}}{}_{\alpha\beta} + \varepsilon \Gamma^{\mu^{(1)}}{}_{\alpha\beta}$$

where the superscripts denotes the order of expansion. The zeroeth order term (which is purely Minkowski) vanishes, since the Minkowski metric is constant. The first order term is,

$$\Gamma^{\mu^{(1)}}{}_{\alpha\beta} = \frac{1}{2}\eta^{\mu\rho}[\partial_\alpha h_{\rho\beta} + \partial_\beta h_{\alpha\beta} - \partial_\rho h_{\alpha\beta}]$$

Since the zeroeth order term of the Christoffel symbol vanishes, the leading order in Γ is $\mathcal{O}(\varepsilon)$ and hence the Riemann tensor, which contains one of the terms as products of two Christoffel symbols, will become $\mathcal{O}(\varepsilon^2)$ which we do not want. Thus, upto leading order the Riemann tensor becomes:

$$R^{\rho^{(1)}}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho^{(1)}}{}_{\nu\sigma} - \partial_\nu \Gamma^{\rho^{(1)}}{}_{\mu\sigma}$$

since the zeroeth order term vanishes. Similar thing happens with the Ricci tensor and the Ricci scalar. We will now focus on the vacuum solution, for which $\mathcal{T}^{\mu\nu}(\bar{\mathbf{x}}) = 0 \implies G_{\mu\nu} = 0 \implies R_{\mu\nu} = 0$

$$\begin{aligned}
R^{(1)}_{\mu\nu} &= \partial_\rho \Gamma^{\rho^{(1)}}{}_{\nu\nu} - \partial_\nu \Gamma^{\rho^{(1)}}{}_{\mu\rho} \\
&= \frac{1}{2}\partial_\rho\{\eta^{\rho\sigma}[\partial_\mu h_{\sigma\nu} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}]\} - \frac{1}{2}\partial_\nu\{\eta^{\rho\sigma}[\partial_\mu h_{\rho\sigma} + \partial_\rho h_{\sigma\mu} - \partial_\sigma h_{\mu\rho}]\} \\
&= \frac{1}{2}\{\partial_\mu \partial^\sigma h_{\sigma\nu} - \partial_\rho \partial^\rho h_{\mu\nu} - \partial_\mu \partial_\nu (h^\rho{}_\rho) + \partial_\nu \partial^\rho h_{\mu\rho}\} = 0
\end{aligned}$$

Let us define $\Box := \partial_\mu \partial^\mu$ and $\mathfrak{h} := h^\rho{}_\rho$ and using these, we get the linearised Einstein's equation as,

$$\boxed{\partial_\mu \partial^\sigma h_{\sigma\nu} + \partial_\nu \partial^\rho h_{\mu\rho} - \Box h_{\mu\nu} - \partial_\mu \partial_\nu \mathfrak{h} = 0} \tag{36}$$

24.1. Engaging with the gauge

Einstein has always been a simp for Maxwell and held his theory in high regard. We can formulate Maxwell's equations using the anti-symmetric electromagnetic tensor $F_{\mu\nu}$ defined as

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

Suppose were consider a source-free region such that the current $J^\nu = 0$, in which the Maxwell's equation become

$$\partial_\mu F^{\mu\nu} = 0 \implies \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0$$

We can introduce a gauge degree of freedom such that $A_\mu \longrightarrow A_\mu + \partial_\mu \lambda$ which keeps $F_{\mu\nu}$ invariant. A popular gauge choice is the Lorentz gauge such that $\partial_\mu A^\mu = 0$ which gives us $\Box A_\nu = 0$ which has plane-wave solutions,

$$A_\nu(\bar{\mathbf{x}}) = \mathcal{C}_\nu \cos(k \cdot x) + \mathcal{D}_\nu \sin(k \cdot x)$$

where the $\cdot$ represent four-vector inner product. Similar to electromagnetism, we can impose a gauge freedom in the linearised gravity too. A popular choice for the gauge is the *traceless-transverse* gauge which satisfies,

$$\partial^\rho h_{\mu\rho} = 0 \qquad \mathfrak{h} = 0$$

Using this gauge choice, Eq. 36 becomes

$$\Box h_{\mu\nu}(\bar{\mathbf{x}}) = 0$$

which again admits a plane-wave solution for the perturbation $h_{\mu\nu}$, which is given as

$$h_{\mu\nu}(\bar{\mathbf{x}}) = \mathcal{C}_\nu \cos(k \cdot x) + \mathcal{D}_\nu \sin(k \cdot x)$$

We obtained plane-wave solutions as a result of linearised Einstein gravity. Thus, the perturbation to the background Minkowski field propagate as waves through space-time.

To see this more clearly, let us consider two points $P_1(x_1^\mu)$ and $P_2(x_2^\mu)$. The proper distance between these points is defined as,

$$d^2 = g_{\mu\nu}\underbrace{(x_1^\mu - x_2^\mu)}_{\Delta l^\mu}\underbrace{(x_1^\nu - x_2^\nu)}_{\Delta l^\nu}$$

which can be expanded using the linearised gravity as,

$$d^2 = (\eta_{\mu\nu} + h_{\mu\nu})\Delta l^\mu \Delta l^\nu = \eta_{\mu\nu}\Delta l^\mu \Delta l^\nu + h_{\mu\nu}\Delta l^\mu \Delta l^\nu$$

The first part represents the static Minkowski proper distance while the second part denotes the perturbation which is oscillating. Hence, the proper distance between two points in linearised Einstein gravity keeps oscillating which can be detected by modern apparatus which confirmed the existence of gravitational waves.

Lecture 25: Introduction to Cosmology

Cosmology literally means the study of the *cosmos* ('Universe') and is one of the oldest subjects studied by humans. Cosmology, as a philosophy, have been prevalent since prehistoric times. Stars and other heavenly objects have been demonstrated in cave paintings and other evidences of ancient life.

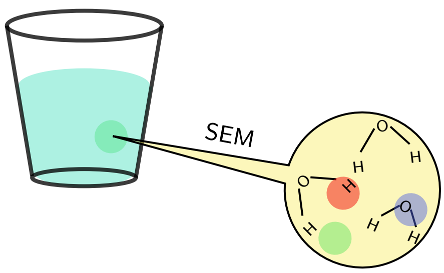
Almost every civilisation have studied cosmology in some form or the other. The concept of Brahmanda goes way back to the Vedic age and describes the creation of the Universe from a 'cosmic egg'.

Keeping aside the philosophical and religious aspects of Cosmology (which might land us in controversis), we will mostly discuss about *physical cosmology*, which is the study of cosmological models, defined with respect to measurable physical quantitites like size, density, etc.

25.1. The scale and the physics

Imagine a glass of water. To measure the density of the water, we can take any mass of water from the glass and then measure the volume. Dividing the mass by the volume we get the density. We find it to be approximately $1\ \mathrm{g/cc}$ and we can fairly estimate that the glass of water is homogenous and at each point the density will be the same. We can take as much small quantity of water for measurement as possible with our everyday apparatus.

However, if we use say some electron microscope to resolve between the molecules of water and in some divine way, use this to measure the density, we would find that the density does not remain fix. In fact, at this length scale $\sim 10^{-10}$ m the concept of density doesn't make much sense.



In the figure, we find that on a coarse-grained scale, the entire glass of water looks homogenous with almost constant density (statistical fluctuations are always there). However, on zooming a particular region upto the atomic scale, we find that the constituents in the red, green and blue regions are totally different and the concept of homogenity becomes invalid.

We note that there is a difference of $8 orders$ between the coarse-grained cm length scale when homogenity is there and the atomic length scale where there is no homogenity.

So one must apply different physical concepts at different scales, even to study the same system 🤯! The same thing happens to our Universe also where based on different length scales, the Universe looks different.

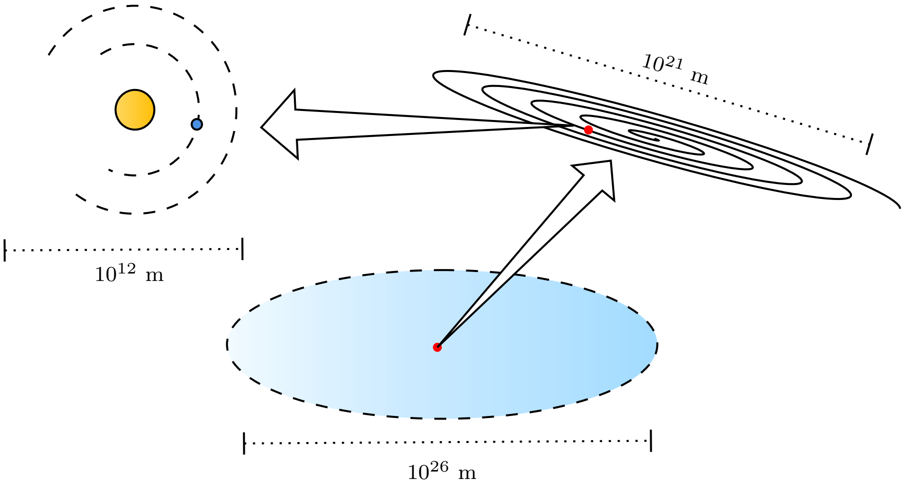


Figure 11: Different length scales of the Universe

The solar system, with a diameter roughly of order 10^12 m, looks fairly inhomogenous to us. In some places, there are planets, while in some other places, some asteroids or something else is present. Moving on to the Milkyway galaxy of size of order 10^21 m, things become a bit smoother but still inhomogenity persists. For the entire observable Universe with estimated size of order 10^26 m, we can safely assume it to be homogenous fluid. In fact there is a difference of more than 14 order between the size of the Solar system (from where all measurements of the Universe are to be done) and the observable Universe. It seems that the Universe is even a better fluid than water!

- ▷ Our Universe at large scales look like a fluid!
- ▷ Only around a small region around us, can we take measurements as we cannot 'physically' access all parts of the observable Universe

Unlike the glass of water, we cannot directly measure the density at two different locations (separated at large lengths) of the Universe, since light (the fastest thing we have) will also take many, many years to reach that point.

Observations from the Earth or around the solar system suggest that the large scale properties of the Universe is independent of the direction of measurement, implying *isotropy*. Hence, the large scale UNiverse is 'isotropic' when viewed from the Earth or around it!

The major question comes to mind: are we (humans) special or some kind of privileged observer? 🤔

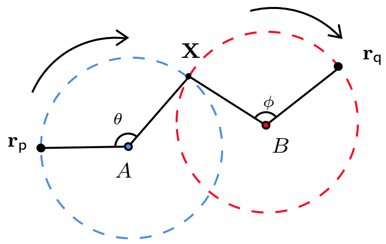
Without delving much into philosophical aspects, it turns out that we simply don't know 😑 (and we might never know).

25.2. Cosmological Principle

It is stated as an axiom that the Earth is not a privileged point 😕 and the Universe would appear to be the same at large scale if viewed from any other point in the space.

Q: Show that if a space appears to be isotropic from just two points in it, then the space is homogenous.

Proof.



Let the space be $\mathcal{M}$ and assume it to be locally Euclidean so that Euclidean metric can be used, and let $A, B \in \mathcal{M}$ such that about A, B the space is isotropic, that is, any function $\rho(r, \theta, \phi) \equiv \rho(r)$ for where r is the distance of any point from A or B. Let r_1, r_2 be given such that $r_1 + r_2 > d > |r_1 - r_2|$ Then effectively we can construct the sets $\mathbb{S}_1^{d-1}$ and $\mathbb{S}_2^{d-1}$ where d is the dimension of the space $\mathcal{M}$

$$\mathbb{S}_1^{d-1} := \{\mathbf{x} \in \mathcal{M} : |\mathbf{x} - \mathbf{r}_{\mathsf{A}}| := r_1\} \qquad \mathbb{S}_2^{d-1} = \{\mathbf{x} \in \mathcal{M} : |\mathbf{x} - \mathbf{r}_{\mathsf{B}}| = r_2\}$$

such that $W = \mathbb{S}_1^{d-1} \cap \mathbb{S}_2^{d-1} \neq \emptyset$. Let $\mathbf{X} \in W$ and consider $\mathbf{P} \in \mathbb{S}_1^{d-1}, \mathbf{Q} \in \mathbb{S}_2^{d-1}$. Since the space is isotropic about point A we can rotate the point P to X about A by an angle θ such that ρ remains same.

$$\rho(\mathbf{P}) = \rho(\mathbf{X}) \equiv \rho(R_\theta^A \mathbf{P})$$

Now $\mathbf{X} \in \mathbb{S}_2^{d-1}$ and since the space is isoptropic about point B, we can rotate it about point B by an angle ϕ to Q without changing ρ

$$\rho(\mathbf{X}) = \rho(\mathbf{Q}) \equiv \rho(R_\phi^B \mathbf{X})$$

Combining the above two expressions we get

$$\rho(\mathbf{Q}) = \rho(\mathbf{P})$$

Since P and Q were arbitrary two points, we can say that $\rho(\mathbf{x}) = \rho(\mathbf{y})$ for any $\mathbf{x} \in \mathbb{S}_1^{d-1}, \mathbf{y} \in \mathbb{S}_2^{d-1}$. By varying the radii we can construct different circles that follow this property and thus we get $\rho(\mathbf{x}) = \rho(\mathbf{y}) \; \forall \; \mathbf{x}, \mathbf{y} \in \mathcal{M}$

Thus, isotropy implies homogenity and at large scales, our Universe is both *homogenous* (independent of spatial position) and *isotropic*.

NOTE: Isotropy means that a space looks same in every direction while homogenity means that the space looks the same at every point. If the space was filled with some parallel strings, then the space would have been homogenous but not isoptropic, since depending on the orientations of the strings, the space would look different in different directions.

Now, another interesting question, is our Universe finite or infinite? And does this even matter?

This question is contained in the list of *fourteen unanswered questions* which Gautam Buddha refused to answer, stating that these kind of questions prevents us from liberation. And this question really doesn't matter that much.

25.3. Olber's Paradox

"The night sky should be bright!"

Consider a star at a distance r and of absolute luminosity L. Then the apparent luminosity of the star for an observer at the origin is

$$A(r) = \frac{L}{4\pi r^2}$$

Let the number density of the star at a distance r be $n(r)$. Then the apparent luminosity of the whole sky as viewed by the observer for the stars present between r and $r + \mathrm{d}r$ is

$$\mathrm{d}E = (4\pi r^2 \mathrm{d}r)n(r)A(r)$$

Then the total observed luminosity by the observer

$$E = \int \mathrm{d}E = \int_0^\infty n(r)L\ \mathrm{d}r \xrightarrow[\text{Universe}]{\text{homogenous}} n_0 L \int_0^\infty \mathrm{d}r \to \infty$$

Clearly this integral diverges and the observed luminosity is apparently infinite which tells us that the night sky should be bright if the observed luminosity (brightness) diverges which is clearly a contradiction. The divergence came because we carried out the integration till *infinity*. There might be some issue with this then and we infer that either

▷ The Universe is finite in age

OR

▷ Light can reach to us only from a finite, restricted part of the Universe

Modern models estimate that indeed Universe has a finite age and thus, we can in principle detect only objects which are situated as far light-years as the age of the Universe and anything beyond that, is out of our reach.

Lecture 26: Friedman Equations

We will discuss about the metric homogenous and isotropic spacetime. Isotropy implies that $\partial_i g_{\mu\nu} = 0$ and hence the metric is at best a function of time

$$g_{\mu\nu} \equiv g_{\mu\nu}(t)$$

Does spacetime really depend on time? During early 1900's, Edwin Hubble observed the spectra of different galaxies and found that all these were 'redshifted', that is, the spectra obtained from the galaxies deviatetd by some amount when compared with that of the normal spectra. We define that relative deviation to be the 'redshift'

$$z := \frac{\lambda_\text{o} - \lambda_\text{e}}{\lambda_\text{o}} \implies \lambda_\text{o} = (1+z)\lambda_\text{e}$$

where λ_o is the observed wavelength and λ_e is the emitted wavelength. Comparing this with the typical Doppler shift, we can say that the galaxies are moving away from us.

Hubble also found out that the recession speed of a distant galaxy follows a simple equation $v = H_0 d$ where H_0 is called the Hubble constant. Actually H_0 is just the current value of $H(t)$ which is called the Hubble parameter. So the Hubble constant tells us the expanding rate of the galaxies at the present moment.

Thus spacetime seems to depend on time, though it is independent of space. We now define the metric used to study such a spacetime.

26.1. Friedmann-Lemaître-Robertson-Walker (FLRW) metric

We define a general time dependent metric for the spacetime described above. Suppose that $\mathrm{d}x \to -\mathrm{d}x$, keeping all unchanged, then $\mathrm{d}x\mathrm{d}y \to -\mathrm{d}x\mathrm{d}y$, however, g_{xy} will not changed, since due to isotropy and homogenity, we expect the metric to depend on time only. Thus we cannot have any cross-terms like this since these will get cancelled.

Also due to isotropy, we expect all of $g_{\text{xx}}(t)$,$g_{\text{yy}}(t)$ and $g_{\text{zz}}(t)$ to be equal for a given t. Since the space-part of the metric should be positive, we define the metric to be of the form

$$\mathrm{d}s^2 = g_{\text{tt}}(t)\mathrm{d}t^2 + a^2(t)\big[\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\big]$$

We now make a choice for $g_{\text{tt}} = -1$ which is called the *proper time* choice. For this choice, the metric tensor looks like

$$[g_{\mu\nu}] = \begin{pmatrix} -1 & & & \\ & a^2 & & \\ & & a^2 & \\ & & & a^2 \end{pmatrix} \qquad [g^{\mu\nu}] = \begin{pmatrix} -1 & & & \\ & {}^1/_{a^2} & & \\ & & {}^1/_{a^2} & \\ & & & {}^1/_{a^2} \end{pmatrix}$$

▷ **Christoffel Symbols**

Since the metric is just dependent on time, finding the Christoffel symbols and other shitty quantities becomes a bit easier here. From the definition of the Christoffel symbol, $\Gamma^t{}_{\alpha\beta} = \frac{1}{2}g^{tt}[g_{\alpha t,\beta} + g_{t\beta,\alpha} - g_{\alpha\beta,t}]$. Now, $g_{tt} = -1$ is constant and hence $\Gamma^t{}_{tt} = 0$.

$$\Gamma^t{}_{xx} = -\frac{1}{2} \times (-2a\dot{a}) = a\dot{a} \qquad \Gamma^x{}_{xt} = \frac{1}{2}g^{xx}[g_{xx,t} + g_{xt,x} - g_{xt,x}] = \frac{1}{2a^2} \times 2a\dot{a} = {}^{\dot{a}}/_{a}$$

. From homogenity and isotropy of space, we can say that

$$\Gamma^x{}_{xt} = \Gamma^y{}_{yt} = \Gamma^z{}_{zt} = \frac{\dot{a}}{a} \qquad \Gamma^t{}_{xx} = \Gamma^t{}_{yy} = \Gamma^t{}_{zz} = a\dot{a}$$

The rest all Christoffel symbols are zero!

▷ **Riemann and Ricci Tensor**

$$\begin{aligned} R_{tt} = R^x{}_{txt} + R^y{}_{tyt} + R^z{}_{tzt} &= 3\Big(\cancelto{0}{\partial_x\Gamma^x{}_{tt}} - \partial_t\Gamma^x{}_{tx} + \Gamma^x{}_{x\lambda}\Gamma^\lambda{}_{tt} - \Gamma^x{}_{t\lambda}\Gamma^\lambda{}_{xt}\Big) \\ &= 3\Big[-\frac{\mathrm{d}({}^{\dot{a}}/_{a})}{\mathrm{d}t} - ({}^{\dot{a}}/_{a})^2\Big] \\ &= -3\Big[\frac{\mathrm{d}({}^{\dot{a}}/_{a})}{\mathrm{d}t} + ({}^{\dot{a}}/_{a})^2\Big] \end{aligned}$$

$$\begin{aligned} R_{xx} = R^\rho{}_{x\rho x} &= \Big(\partial_\rho\Gamma^\rho{}_{xx} - \partial_x\Gamma^\rho{}_{x\rho} + \Gamma^\rho{}_{\rho\lambda}\Gamma^\lambda{}_{xx} - \Gamma^\rho{}_{x\lambda}\Gamma^\lambda{}_{\rho x}\Big) \\ &= \Big(\partial_t\Gamma^t{}_{xx} - \cancelto{0}{\partial_x\Gamma^\rho{}_{x\rho}} + 3\Gamma^x{}_{xt}\Gamma^t{}_{xx} - \Gamma^t{}_{xx}\Gamma^x{}_{tx} - \Gamma^x{}_{xt}\Gamma^t{}_{xx}\Big) \\ &= \frac{\mathrm{d}(a\dot{a})}{\mathrm{d}t} + 3({}^{\dot{a}}/_{a})(a\dot{a}) - 2({}^{\dot{a}}/_{a})(a\dot{a}) \\ &= a\ddot{a} + \dot{a}^2 + \dot{a}^2 = a\ddot{a} + 2\dot{a}^2 \end{aligned}$$

By isotropy we can say that $R_{xx} = R_{yy} = R_{zz} = a\ddot{a} + 2\dot{a}^2$

▷ **Ricci Scalar**

$$\begin{aligned} R = g^{\mu\nu}R_{\mu\nu} = g^{tt}R_{tt} + 3g^{xx}R_{xx} &= 3\left[\frac{\mathrm{d}(\dot{a}/a)}{\mathrm{d}t} + (\dot{a}/a)^2\right] + \frac{3}{a^2}\left[a\ddot{a} + 2\dot{a}^2\right] \\ &= 3\left[\frac{\ddot{a}}{a} - \cancel{\frac{\dot{a}^2}{a}} + \cancel{\frac{\dot{a}^2}{a^2}}\right] + 3\frac{\ddot{a}}{a} + 6\frac{\dot{a}^2}{a^2} \\ &= 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right] \end{aligned}$$

26.2. Einstein's Equations

At large scale, the Universe is assumed to be a perfect fluid and thus we assume the same form of the stress-energy tensor

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

Now the macroscopic speed of the fluid cannot have a privileged direction, so it has only temporal component: $\bar{\mathbf{u}} = (u^t, 0, 0, 0)$. Now the norm of the velocity four vector is always -1

$$g_{\mu\nu}u^\mu u^\nu = g_{tt}(u^t)^2 = -(u^t)^2 = -1 \implies u^t = 1 \quad u_t = -1$$

▷ First Einstein Equation

We have the following non-zero component for the stress energy tensor

$$\mathcal{T}_{tt} = (\rho + P)u_{tt} + g_{tt}P = \rho + P - P = \rho \qquad \mathcal{T}_{xx} = g_{xx}P = Pa^2$$

$$\begin{aligned} R_{tt} - \frac{1}{2}g_{tt}R &= -3\left[\frac{\mathrm{d}(\dot{a}/a)}{\mathrm{d}t} + (\dot{a}/a)^2\right] + \frac{1}{2}\times 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right] \\ &= -\cancel{\left(\frac{3\ddot{a}}{a}\right)} + 3\cancel{\left(\frac{\dot{a}}{a}\right)^2} - 3\cancel{\left(\frac{\dot{a}}{a}\right)^2} + \cancel{\frac{3\ddot{a}}{a}} + 3\left(\frac{\dot{a}}{a}\right)^2 = 8\pi G\rho \end{aligned}$$

Then we obtain the first Einstein equation as

$$\boxed{\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3}} \tag{37}$$

The above equation is called the *Friedmann equation*.

Lecture 27: Scalar field in FLRW metric

In the previous section we found out the Friedmann equation and we will now derive the second equation.

▷ Second Einstein equation

$$R_{xx} - \frac{1}{2}g_{xx}R = a\ddot{a} + 2\dot{a}^2 - \frac{a^2}{2}\times 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right] = a\ddot{a} + 2\dot{a}^2 - 3a\ddot{a} - 3\dot{a}^2 = -2a\ddot{a} - \dot{a}^2 = 8\pi G a^2 P$$

Dividing by a^2 and using Eq. 37 to substitute the value of $\dot{a}/a$ above. We obtain a simplified expression

$$-\frac{2\ddot{a}}{a} - \frac{8\pi G\rho}{3} = 8\pi GP \implies \boxed{\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P)} \tag{38}$$

This is the second Friedmann equation that we obtain for the FLRW metric. We could obtain the time-derivative of the density from Eq. 37 and 38. From first equation we have

$$\frac{\dot{a}}{a} = c\sqrt{\rho} \qquad c = \sqrt{\frac{8\pi G}{3}} \to 3c^2 = 8\pi G$$

Differentiating the above expression we get

$$\begin{aligned} 3\left[\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2\right] = \frac{3c}{2\sqrt{\rho}}\dot{\rho} \quad \xrightarrow[\text{Eq. II}]{\text{Friedmann}} \quad & -\frac{3c^2}{2}(\rho + 3P) - 3\rho c^2 = \frac{3c^2}{2(\dot{a}/a)}\dot{\rho} \\ \implies & -\frac{1}{2}(\rho + 3P) - \rho = \frac{\dot{\rho}}{2(\dot{a}/a)} \\ \implies & -\frac{3}{2}(\rho + P) = \frac{\dot{\rho}}{2(\dot{a}/a)} \\ \implies & \boxed{\dot{\rho} = -3\left(\frac{\dot{a}}{a}\right)(\rho + P)} \end{aligned} \tag{39}$$

27.1. Scalar field in curved spacetime

Consider a scalar field $\phi(x)$ for which the action in flat spacetime takes the form

$$\mathcal{S}[\phi] = \int \mathrm{d}^4x \left(-\frac{1}{2}(\partial^\mu\phi)(\partial_\mu\phi) - V(\phi)\right)$$

where V is the potential. In arbitrary spacetime, we need to make the volume element invariant, and promote normal derivatives to covariant derivatives. Hence, in any arbitrary curved spacetime we have the action for the field

$$\mathcal{S}[\phi] = \int \mathrm{d}^4x\sqrt{-\mathfrak{g}} \left(-\frac{1}{2}g^{\mu\nu}(\nabla_\mu\phi)(\nabla_\nu\phi) - V(\phi)\right) \tag{40}$$

For the FLRW metric, the field is just a function of time due to homogenity of the space and $\sqrt{-\mathfrak{g}} = \sqrt{(a^2)^3} = a^3$. Then the action reduces to

$$\mathcal{S}[\phi] = \int \mathrm{d}^4x \; a^3\left(\frac{1}{2}(\partial_t\phi)^2 - V(\phi)\right)$$

We can now separate the integral into a space part and a time part, since field and potential all just depends on time

$$\mathcal{S}[\phi] = \int \mathrm{d}t \; a^3\left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right)\int \mathrm{d}^3x \quad \longrightarrow \quad \int \mathrm{d}t \; a^3\left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right)v_0$$

where v_0 is the free volume of the finite part of the Universe. The physical volume is of course given by $v = \int \mathrm{d}^3x \; \sqrt{-\mathfrak{g}} = a^3 v_0$. From this form of the action we define the Lagrangian as

$$\mathrm{L} = v\left[\frac{1}{2}\dot{\phi}^2 - V(\phi)\right]$$

The Lagrangian looks very much like the single particle Lagrangian with the kinetic and the potential terms. The conjugate momentum of the field is given as,

$$\Pi = \frac{\partial L}{\partial\dot{\phi}} = v\dot{\phi} \quad \xrightarrow[\text{Transformation}]{\text{Legendre}} \quad \mathrm{H} = \Pi\dot{\phi} - L = \frac{\Pi^2}{v} - \left[\frac{v}{2} \times \frac{\Pi^2}{v^2} - vV(\phi)\right]$$

From the above calculation we get the Hamiltonian as

$$\mathrm{H} = \frac{\Pi^2}{2v} + vV(\phi)$$

From the Hamiltonian we can find the energy density and the pressure as

$$\rho = \frac{\mathrm{H}}{v} = \frac{\Pi^2}{2v^2} + V(\phi) \qquad P = -\frac{\partial E}{\partial v} = -\Big(-\frac{\Pi^2}{2v^2} + V(\phi)\Big) = \Big(\frac{\Pi^2}{2v^2} - V(\phi)\Big)$$

If we now want to obtain the stress-energy tensor components using the above quantites, we have

$$\begin{aligned} \mathcal{T}_{tt} &= (\rho + P)u_t u_t + P g_{tt} = \frac{\Pi^2}{v^2} - P = \frac{\Pi^2}{2v^2} + V(\phi) = \frac{1}{2}\dot{\phi}^2 + V(\phi) \\ \mathcal{T}_{xx} &= 0 + P g_{xx} = a^2\Big(\frac{\Pi^2}{2v^2} - V(\phi)\Big) = a^2\Big(\frac{1}{2}\dot{\phi}^2 - V(\phi)\Big) \end{aligned}$$

The above components of the stress-energy tensor were derived purely from considering the thermodynamic relations and considering the perfect-fluid form. We now find the stress-energy tensor from the first principles, using the definition

$$\mathcal{T}^{\mu\nu} := \frac{-2}{\sqrt{-\mathfrak{g}}}\Big(\frac{\delta \mathcal{S}_{\mathsf{M}}}{\delta g_{\mu\nu}}\Big)$$

From Eq. 40 we write the action in terms of the Lagrangian density

$$\mathcal{S}[\phi] = \int \mathrm{d}^4x\ \sqrt{-\mathfrak{g}}\ \mathcal{L} \qquad \text{where, } \mathcal{L} = \Big(-\frac{1}{2}g^{\mu\nu}(\nabla_\mu\phi)(\nabla_\nu\phi) - V(\phi)\Big)$$

Recall that the variation of the metric determinant from **??** was given by

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

Using that above we obtain the variation of the action as

$$\begin{aligned} \delta\mathcal{S} &= \int \mathrm{d}^4x\ \big[\mathcal{L}\delta\sqrt{-\mathfrak{g}} + \sqrt{-\mathfrak{g}}\delta\mathcal{L}\big] \\ &= \int \mathrm{d}^4x\ \Big[-\frac{1}{2}\sqrt{-\mathfrak{g}}\ g_{\mu\nu}\delta g^{\mu\nu}\mathcal{L} - \frac{1}{2}\sqrt{-\mathfrak{g}}(\nabla_\mu\phi)(\nabla_\nu\phi)\delta g^{\mu\nu}\Big] \\ &= \int \mathrm{d}^4x\ \ -\frac{1}{2}\sqrt{-\mathfrak{g}}[\ g_{\mu\nu}\mathcal{L} + (\nabla_\mu\phi)(\nabla_\nu\phi)]\delta g^{\mu\nu} \equiv \int \mathrm{d}^4x\ \frac{\delta\mathcal{S}_{\mathsf{M}}}{\delta g^{\mu\nu}}\delta g^{\mu\nu} \end{aligned}$$

From this we can identify the stress-energy tensor to be

$$\mathcal{T}_{\mu\nu} = g_{\mu\nu}\mathcal{L} + (\nabla_\mu\phi)(\nabla_\nu\phi) = (\nabla_\mu\phi)(\nabla_\nu\phi) + g_{\mu\nu}\Big\{-\frac{1}{2}g^{\alpha\beta}(\nabla_\alpha\phi)(\nabla_\beta\phi) - V(\phi)\Big\}$$

Let us now calculate the components using this expression:

$$\begin{aligned} \mathcal{T}_{tt} &= \dot{\phi}^2 - \frac{1}{2}\dot{\phi}^2 + V(\phi) = \frac{1}{2}\dot{\phi}^2 + V(\phi) \\ \mathcal{T}_{xx} &= a^2\Big(\frac{1}{2}\dot{\phi}^2 - V(\phi)\Big) \end{aligned}$$

We see that the stress-energy tensor as calculated from the definition by varying the action with respect to the metric is the same as that calculated from thermodynamic relations, atleast for the case of the scalar field.

Lecture 28: Radiation dominated Universe

In the previous section we saw scalar fields in the FLRW metric. Here we will focus on a period in the Universe's past when it was (perhaps) dominated by radiation (photns). For that, note that in Minkowski space, the electromagnetic action is given by

$$\mathcal{S}_{\mathsf{EM}} = \int \mathrm{d}^4x \Big[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \Big]$$

where $F_{\mu\nu}$ is Maxwell's field tensor as defined in Eq.11. Note that this action is Lorentz invariant and also enjoys gauge symmetry. To generalise this, we use our previous prescription, that is, we replace the measure with the invariant volume, introduce the general metric and change the partial derivatives to covariant derivatives, doing which we obtain the action in arbitrary curved spacetime

$$\mathcal{S}_{\mathsf{EM}} = \int \mathrm{d}^4x \, \sqrt{-\mathfrak{g}} \, \Big[-\frac{1}{4} g^{\mu\alpha} g^{\nu\beta} F_{\mu\nu} F_{\alpha\beta} \Big] \tag{41}$$

where $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. However note that

$$\begin{aligned} F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu = \big(\partial\mu A_\nu - \Gamma^\sigma{}_{\mu\nu} A_\sigma\big) - \big(\partial\nu A_\mu - \Gamma^\sigma{}_{\nu\mu} A_\sigma\big) &= \big(\partial_\mu A_\nu - \cancel{\Gamma^\sigma{}_{\mu\nu} A_\sigma}\big) - \big(\partial_\nu A_\mu - \cancel{\Gamma^\sigma{}_{\mu\nu} A_\sigma}\big) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned}$$

Since Christoffel symbols are symmetric in lower indices, the covariant derivatives had no effect on the field tensor.

For the above action for free electromagnetic field, the stress-energy tensor is given by

$$\mathcal{T}_{\mu\nu} = g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} - \frac{1}{4} g_{\mu\nu} \big(g^{\rho\sigma} g^{\alpha\beta} F_{\rho\alpha} F_{\sigma\beta} \big)$$

Proof. Note that we define the stress-energy tensor using the variation of the action with respect to the metric.

$$\begin{aligned} \delta\mathcal{S} &= \int \mathrm{d}^4x \, \Big\{ (\delta\sqrt{-\mathfrak{g}}) \Big[-\frac{1}{4} g^{\mu\alpha} g^{\nu\beta} F_{\mu\nu} F_{\alpha\beta} \Big] - \frac{1}{4}\sqrt{-\mathfrak{g}} \Big[\delta g^{\mu\alpha} g^{\nu\beta} F_{\mu\nu} F_{\alpha\beta} + g^{\mu\alpha} \delta g^{\nu\beta} F_{\mu\nu} F_{\alpha\beta} \Big] \Big\} \\ &= \int \mathrm{d}^4x \, \Big\{ -\frac{1}{2}\sqrt{-\mathfrak{g}} \, g_{\mu\nu} \delta g^{\mu\nu} \Big(-\frac{1}{4} g^{\rho\alpha} g^{\sigma\beta} F_{\rho\sigma} F_{\alpha\beta} \Big) - \frac{1}{4}\sqrt{-\mathfrak{g}} \, \delta g^{\mu\nu} \Big(g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} + g^{\beta\alpha} F_{\beta\nu} F_{\alpha\mu} \Big) \Big\} \\ &= \int \mathrm{d}^4x \, \Big\{ -\frac{1}{2}\sqrt{-\mathfrak{g}} \, g_{\mu\nu} \delta g^{\mu\nu} \Big(-\frac{1}{4} g^{\rho\alpha} g^{\sigma\beta} F_{\rho\sigma} F_{\alpha\beta} \Big) - \frac{1}{4}\sqrt{-\mathfrak{g}} \, \delta g^{\mu\nu} \Big(g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} + g^{\alpha\beta} (-F_{\nu\beta})(-F_{\mu\alpha}) \Big) \Big\} \\ &= \int \mathrm{d}^4x \, -\frac{1}{2}\sqrt{-\mathfrak{g}} \Big\{ \Big(-\frac{1}{4} g_{\mu\nu} g^{\rho\alpha} g^{\sigma\beta} F_{\rho\sigma} F_{\alpha\beta} \Big) + g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} \Big\} \delta g^{\mu\nu} \equiv \int \mathrm{d}^4x \, \frac{\delta\mathcal{S}}{\delta g^{\mu\nu}} \delta^{\mu\nu} \end{aligned}$$

where we have used the expression for the variation of the metric determinant and the anti-symmetric property of the field tensor. Then from the definition of the stress-energy tensor, we obtain the form

$$\boxed{\mathcal{T}_{\mu\nu} = \Big(g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} - \frac{1}{4} g_{\mu\nu} g^{\rho\alpha} g^{\sigma\beta} F_{\rho\sigma} F_{\alpha\beta} \Big)}$$

Let us calculate the trace for this stress-energy tensor.

$$\mathcal{T}^\mu{}_\mu = g^{\mu\nu} \mathcal{T}_{\mu\nu} = \Big(g^{\mu\nu} g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} - \frac{1}{4} g^{\mu\nu} g_{\mu\nu} g^{\rho\alpha} g^{\sigma\beta} F_{\rho\sigma} F_{\alpha\beta} \Big) = \Big(F^{\nu\beta} F_{\nu\beta} - F^{\alpha\beta} F_{\alpha\beta} \Big) = 0$$

where we have used $g^{\mu\nu} g_{\mu\nu} = \delta^\mu{}_\mu = 4$. We see that the stress-energy tensor for the classical electromagnetic field is *traceless*. This is in general true for massless fields (like here, photons). However, when considering *quantum fields*, the traceless property no longer holds, giving rise to *trace anomaly*!

28.1. Equation of State

When the Universe was dominated by radiation, we expect the stress-energy tensor to be *traceless* and hence we can impose the traceless condition on $\mathcal{T}_{\mu\nu}$ for the perfect-fluid form, considering that this described, at one point in time, a radiation dominated Universe.

$$g^{\mu\nu}\mathcal{T}_{\mu\nu} = (\rho + P)g^{\mu\nu}u_\mu u_\nu + Pg^{\mu\nu}g_{\mu\nu} = -(\rho + P) + 4P = 3P - \rho$$

where we have used that the norm of the four velocity is always -1 (in natural units). Then imposing the traceless condition we get $\rho = 3P$. We call this as the equation of state (EoS) and define

$$w \equiv \frac{P}{\rho} = \frac{1}{3}$$

where w is called the EoS parameter. This equation of state describes a *barotropic fluid* since the pressure is entirely controlled by the pressure.

28.2. Friedmann Equations

We now apply the above implications in the Friedmann equations. Using Eq. 39 we have

$$\dot{\rho} = -3\left(\frac{\dot{a}}{a}\right)(\rho + P) = -\cancel{3}\left(\frac{\dot{a}}{a}\right) \times \frac{4\rho}{\cancel{3}} \implies \frac{\mathrm{d}\rho}{\rho} = -4\frac{\mathrm{d}a}{a} \implies \boxed{\rho = \rho_0\left(\frac{a_0}{a}\right)^4}$$

where a_0 is the value of a at some time τ. If we use the above form of the density in the Friendmann equation, we get

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho_0}{3}\left(\frac{a_0}{a}\right)^4$$

Now we take the positive square-root of the above expression.

$$\left(\frac{\dot{a}}{a}\right) = \underbrace{\sqrt{\frac{8\pi G\rho_0 a_0^4}{3}}}_{c_o^2/2}\frac{1}{a^2} \implies a\,\mathrm{d}a = \frac{c_0^2}{2}\mathrm{d}t \implies a^2 = a_0^2 + c_0^2(t - \tau) = c_0^2\left(t - \tau + \left(\frac{a_0}{c_0}\right)^2\right)$$

Define $t_0 := \left[\left(\frac{a_0}{c_0}\right)^2 - \tau\right]$ from which we obtain the expression for $a(t)$:

$$\boxed{a(t) = c_0\sqrt{t - t_0}}$$

Now, from the metric, we can interpret $a(t)$ as a *scale-factor* which connects physical length d to the coordinate length l. The coordinate length is the distance between two comoving points, as measured in some coordinate system while the physical length is the invariant distance between two points. We thus get the relation $d = a(t)l$.

Note that, as $t \to t_0$ we have $a(t) \to 0$ and hence $d \to 0$, meaning that, irrespective of the coordinate length l, the physical length is always zero. This means that all the matter in the Universe was concentrated at a single point at a *finite* time t_0.

Also note that t_0 is, in some way, the lower limit of *time*, since $t < t_0$ implies that a becomes imaginary. Hence we can say that the Universe was *born* as finite time ago! This phenomenon is called the *Big Bang*.

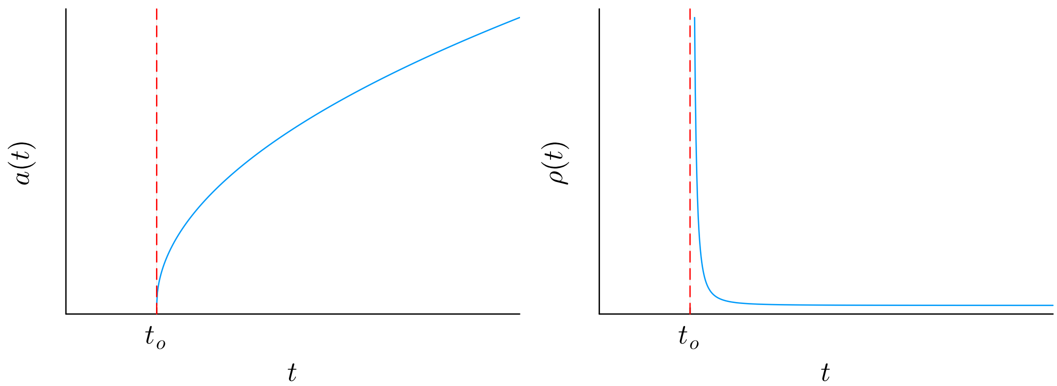


Figure 12: Plot of $a(t)$ and $\rho(t)$ with time in a radiation dominated Universe

Now, from the expression of $a(t)$ we can find the density profile,

$$\rho(t) = \rho_0 \left(\frac{a_0}{c_0\sqrt{t - t_0}} \right)^4 \sim \frac{1}{(t - t_0)^2}$$

This blows up as $t \to t_0$ and things kinda becomes *unphysical*! This implies that, during the Big Bang, the energy density of the Universe was infinite. We suspect that GR might break down at t_0 since this predicts something unphysical. Perhaps a theory of *quantisation of gravity*, if it occurs someday, may be able to rescue this apparent breakdown!

28.3. Pressureless matter

We will now look at a simpler case of pressureless *dust matter*, that is, $P = 0$. Using this in the Friedmann equation we obtain,

$$\dot{\rho} = -3\left(\frac{\dot{a}}{a}\right) \implies \frac{\mathrm{d}\rho}{\rho} = -3\frac{\mathrm{d}a}{a} \implies \boxed{\rho = \rho_0 \left(\frac{a_0}{a}\right)^3}$$

Now using this in the Friedmann equation we obtain

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho_0}{3}\left(\frac{a_0}{a}\right)^3 \implies \dot{a} = \sqrt{\frac{8\pi G\rho_0 a_0^3}{3}}\frac{1}{\sqrt{a}} \implies a(t) \sim (t - t_0)^{2/3}$$

And we also look at the energy density

$$\rho(t) = \rho_0\left(\frac{a_0}{a}\right)^3 \sim \frac{1}{(t - t_0)^2}$$

We see that the energy density profile remains same for the pressureless matter case also, that is, it has a singularity at $t = t_0$.

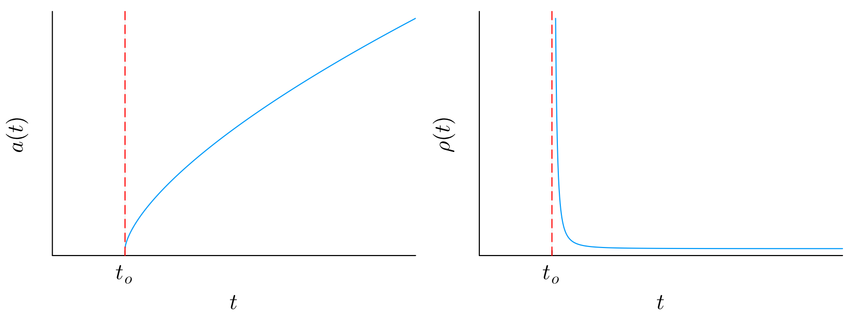


Figure 13: Plot of $a(t)$ and $\rho(t)$ with time for a pressureless dust matter

Lecture 29: Some issues...

Recall the definition of the Hubble parameter as the proportionality between the recession speed of a distant galaxy and its distance from us which can be written in terms of the scale factor $a(t)$

$$H = \frac{v}{d} = \frac{\dot{d}}{d} = \frac{l\dot{a}(t)}{la(t)} = \frac{\dot{a}}{a}$$

Then the Friedmann equation now becomes $3H^2 = 8\pi G\rho$. Let us define $\rho_{\mathrm{H}} = {}^{3H^2}/_{8\pi G}$ from which we have ${}^{\rho}/_{\rho_{\mathrm{H}}} = 1$. However, the observed baryonic energy density is $\rho_* = 0.05\rho_{\mathrm{H}}$. Thus there must be some other contributions to ρ.

It is kinda known that there is a $0.25\rho_{\mathrm{H}}$ contribution from *dark matter* (which is a form of almost non-interacting matter assumed to be present in the Universe). There are various evidences for the presence of DM however, concretely we still don't know what it is.

The rest $0.7\rho_{\mathrm{H}}$ contribution is still unknown and in a completely unimaginative way, we attribute it to *dark energy*[1].

29.1. Strong-Energy Condition

Recall that the stress-energy tensor $\mathcal{T}_{\mu\nu}$ was somewhat arbitrarily defined. In fact, for any given metric, we can always calculate $G_{\mu\nu}$ and then demand that this is the stress-energy tensor, however, this will not always be physical. For 'realistic' sources of energy, we thus impose some coordinate-invariant conditions on the stress-energy tensor which are called *energy conditions*. There are different energy conditions, of which we focus on:

- **Weak Energy Condition** which states that $\mathcal{T}_{\mu\nu}t^\mu t^\nu \geq 0$ for all time-like vectors t^μ. For the perfect fluid form, this condition is equivalent to $\rho + P \geq 0$ and $\rho \geq 0$
- **Strong Energy Condition** which states that $\mathcal{T}_{\mu\nu}t^\mu t^\nu \geq \frac{1}{2}\mathcal{T}^\lambda_{\ \lambda}\, t^\sigma t_\sigma$ for any time-like vector t^μ which, for the perfect fluid form, is equivalent to $\rho + P \geq 0$ and $\rho + 3P \geq 0$

Let us define a quantity

$$q = -\frac{{}^{\ddot{a}}/_{a}}{({}^{\dot{a}}/_{a})^2} = -\frac{\frac{4\pi G}{3}(\rho + 3P)}{\frac{8\pi G}{3}\rho} = \frac{1}{2}\left(1 + \frac{3P}{\rho}\right) = \frac{1}{2}(1 + 3w)$$

where w is the EoS parameter. The quantity q is the measure of the cosmic acceleration. If $q > 0$ then Universe decelerates. Note that, if the strong energy condition holds, then this directly implies that Universe is decelerating since $w = \frac{P}{\rho}$ and for physical matter, in general $\rho > 0$!

Recent observations tells the contrary fact that Universe is accelerating ($q < 0$ and $w < -{}^1/_3$), that is, somehow gravity is behaving as a *repulsive force*!

29.2. Cosmological Constant

Recall from Eq.16 that the complete Einstein-Hilbert action had the term Λ which was termed as the cosmological constant.

$$\mathcal{S}_\Lambda = -\frac{1}{8\pi G}\int \mathrm{d}^4x\, \sqrt{-g}\,\Lambda \implies \mathcal{T}^\Lambda_{\mu\nu} = \frac{-2}{\sqrt{g}}\frac{\delta\mathcal{S}_\Lambda}{\delta g^{\mu\nu}} = -\frac{\Lambda}{8\pi G}g_{\mu\nu}$$

[1] The mere fact that so much unknown energy is attributed to *darkness* is a vivid example of how progressive our society has become!

Also if we assume the perfect fluid form, we have $\mathcal{T}_{\mu\nu} = (\rho + P)u_\mu u_\nu + Pg_{\mu\nu}$. If we want to satisfy both of these, then we get

$$\boxed{P = -\frac{\Lambda}{8\pi G}} \qquad \boxed{P + \rho = 0} \implies w = -1 < -\frac{1}{3} \quad \rho = \frac{\Lambda}{8\pi G} > 0$$

Thus we see that adding the cosmological constant naturally gave us a way to explain the accelerating Universe and thus we can kinda attribute the rest of the energy density contribution to Λ.

$$3H^2 = 8\pi G(\rho_{\text{baryon}} + \rho_{\text{DM}} + \rho_\Lambda)$$

Λ solves the *missing energy* and *missing acceleration* issue, however, other subtleties do remain!

Lecture 30: Lie Derivatives

We will shift our focus from the physical aspects to a more mathematical aspect, of finding the symmetries of a given arbitrary metric. Coordinate transformations which keep the metric invariant are called *isometries*[1]. Recall the transformation of different tensorial quanttities, where now, the dependence on the coordinates is explicitly shown:

- **Scalar**: $\varphi'(x') = \varphi(x)$
- **Vector**: $v^{\mu'}(x') = \Lambda^\mu{}_\nu v^\nu(x)$
- **(0,2) Tensor**: $L'_{\mu\nu}(x') = \left(\Lambda^{-1}\right)^\beta{}_\nu \left(\Lambda^{-1}\right)^\alpha{}_\mu L_{\alpha\beta}(x)$

We need a general procedure to find the isometries of a given metric. For that, let us now consider an infinitesimal coordinate transformation by some vector ξ^μ

$$x'^\mu = x^\mu + \varepsilon\xi^\mu(\bar{\mathbf{x}}) \qquad 0 < \varepsilon \ll 1$$

The transformation matrix, upto order $\mathcal{O}(\varepsilon^2)$, for the above transformation is given as:

$$\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}(x^\mu + \varepsilon\xi^\mu) = \delta^\mu{}_\nu + \varepsilon\frac{\partial \xi^\mu}{\partial x^\nu}$$

The inverse transformation is similarly given as:

$$\left(\Lambda^{-1}\right)^\mu{}_\nu = \delta^\mu{}_\nu - \varepsilon\frac{\partial \xi^\mu}{\partial x^\nu}$$

We can check that this is indeed the inverse upto $\mathcal{O}(\varepsilon^2)$, since

$$\begin{aligned}
\Lambda^\alpha{}_\nu \left(\Lambda^{-1}\right)^\nu{}_\beta &= \left(\delta^\alpha{}_\nu + \varepsilon\frac{\partial \xi^\alpha}{\partial x^\nu}\right)\left(\delta^\nu{}_\beta - \varepsilon\frac{\partial \xi^\nu}{\partial x^\beta}\right) \\
&= \delta^\alpha{}_\beta + \varepsilon\left[\frac{\partial \xi^\alpha}{\partial x^\nu}\delta^\nu{}_\beta - \frac{\partial \xi^\nu}{\partial x^\beta}\delta^\alpha{}_\nu\right] \\
&= \delta^\alpha{}_\beta + \varepsilon\left[\cancel{\frac{\partial \xi^\alpha}{\partial x^\beta}} - \cancel{\frac{\partial \xi^\alpha}{\partial x^\beta}}\right] \\
&= \delta^\alpha{}_\beta
\end{aligned}$$

NOTE: In the following part, we will define a new kind of derivative called the *Lie derivative*. We will use the notion of coordinate transformations to define this, however, the precise definition is by considering the 'flow'

[1]Mathematically, *isometries* are *diffeomorphisms* such that the *pullback* of the metric is equal to the metric, which basically means that isometries preserve the metric, but this is justsome sophisticated differential geometric jargon, made up to sound cool (and also to precisely define things without being philosophical and hand-wavy)

generated by some vector field.

There is a nice reason for such a definition of the Lie derivative. Suppose $\boldsymbol{\xi}$ is a vector field, which means that, if we specify a point in spacetime, $\boldsymbol{\xi}$ will assign a vector to that point. Now we can define some curves (called *flow lines* or *integral curves*) such that at each point p on the curve, if we define a tangent vector, this tangent vector will be equal to the vector that is assigned by $\boldsymbol{\xi}$ at p.

Consider a flow-line curve parameterised by some ε, $\sigma(\varepsilon)$ generated by the vector field $\boldsymbol{\xi}$ and a point p with coordinates x^μ on that flow-line. We want to compare how some given vector field, say T, changes when we move along the flow line to point q.

At each point on the flow line, T will have different 'values'. However, we cannot compare directly between two different points since each different point is associated with different vector spaces (as the spacetime can have different geometry at different points, vectors spaces for each point is defined 'locally') and two different vector spaces cannot be directly compared.

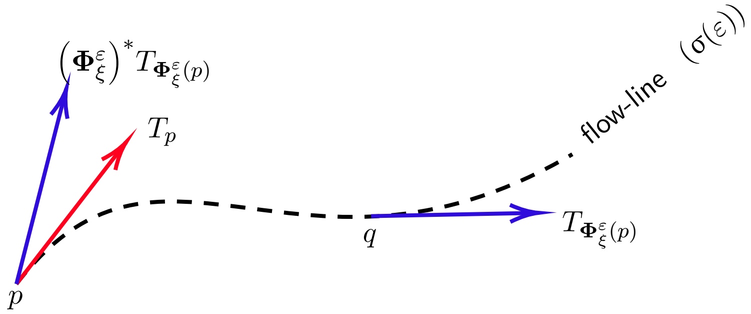


Figure 14: Lie transport of a vector along the flow-line generated by some vector field $\boldsymbol{\xi}$

So for q with coordinates x'^μ, we first need to bring the vector there back to the vector at p (so that these are both in the same vector space) and then compare. This 'bringing back' is what is done by the *pullback*.

As a final scare I am going to specify the definition of Lie derivative according to the above intuition. Consider $\mathbf{\Phi}_\xi^\varepsilon$ to be the map which takes points along a flow-line and consider the vector field T, which will take values T_p at point p and T_q at point q. Thus T_p and T_q are tangent vectors at those points. Now, q can be written as $\mathbf{\Phi}_\xi^\varepsilon(p)$ and thus the tangent vector at q is then $T_{\mathbf{\Phi}_\xi^\varepsilon(p)}$.

Now let us denote the pull-back map by $\left(\mathbf{\Phi}_\xi^\varepsilon\right)^*$ which will act on the vector field T_q. Note that since $\mathbf{\Phi}_\xi^\varepsilon(p)$ generates the flow, we need to define the pullback with respect to this map only. And then the Lie derivative compares between the dragged vector and the original vector,

$$\mathcal{L}_{\boldsymbol{\xi}} T(p) = \lim_{\varepsilon \to 0} \frac{\left(\mathbf{\Phi}_\xi^\varepsilon\right)^* T_{\mathbf{\Phi}_\xi^\varepsilon(p)} - T_p}{\varepsilon}$$

This is as fas as I can go due to my own inability to understand these complex math stuff. Hence for now, we will shove this into a box and move on with a much simpler definition of Lie derivative using coordinate transformations.

30.1. Lie Derivative of a scalar field

Under a coordinate transformation, we define the *Lie derivative* of a scalar as follows:

$$
\begin{aligned}
\mathcal{L}_{\boldsymbol{\xi}}\varphi &:= \lim_{\varepsilon\to 0}\frac{\varphi(x')-\varphi'(x')}{\varepsilon} \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[\varphi(x^\alpha+\varepsilon\xi^\alpha)-\varphi(x^\alpha)] \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big[\varphi(x^\alpha)+\varepsilon\frac{\partial\varphi}{\partial x^\alpha}\xi^\alpha-\varphi(x^\alpha)\Big] \\
&= \xi^\alpha\partial_\alpha\varphi
\end{aligned}
$$

Since for scalars, covariant and partial derivatives coincide, we can write

$$
\boxed{\mathcal{L}_{\boldsymbol{\xi}}\varphi = \xi^\alpha\nabla_\alpha\varphi}
$$

Hence the Lie derivative of a scalar field under some coordinate transformation generated by $\boldsymbol{\xi}$, is just the 'directional derivative' along $\boldsymbol{\xi}$.

30.2. Lie Derivative of a contravector field

In similar lines to the above, we can define the Lie derivative of a contra-vector as

$$
\begin{aligned}
\mathcal{L}_{\boldsymbol{\xi}}V^\mu &:= \lim_{\varepsilon\to 0}\frac{V^\mu(x')-V^{\mu'}(x')}{\varepsilon} \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big[V^\mu(x^\alpha+\varepsilon\xi^\alpha)-\Lambda^\mu{}_\beta V^\beta(x^\alpha)\Big] \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big[V^\mu(x^\alpha)+\varepsilon\frac{\partial V^\mu}{\partial x^\alpha}\xi^\alpha-\Big(\delta^\mu{}_\beta+\varepsilon\frac{\partial\xi^\mu}{\partial x^\beta}\Big)V^\beta(x^\alpha)\Big] \\
&= \xi^\alpha\partial_\alpha V^\mu - V^\alpha\partial_\alpha\xi^\mu
\end{aligned}
$$

Now, the above expression can be written in terms of covariant derivatives.

$$
\mathcal{L}_{\boldsymbol{\xi}}V^\mu = \xi^\alpha(\nabla_\alpha V^\mu-\Gamma^\mu{}_{\alpha\sigma}V^\sigma)-V^\alpha(\nabla_\alpha\xi^\mu-\Gamma^\mu{}_{\alpha\sigma}\xi^\sigma) = (\xi^\alpha\nabla_\alpha V^\mu - V^\alpha\nabla_\alpha\xi^\mu)
$$

The terms with the Christoffel symbols cancel after renaming the dummy indices $\alpha\leftrightarrow\sigma$ in one of the terms. Hence for a contravariant vector we find that

$$
\boxed{\mathcal{L}_{\boldsymbol{\xi}}V^\mu = \xi^\alpha\nabla_\alpha V^\mu - V^\alpha\nabla_\alpha\xi^\mu}
$$

30.3. Lie Derivative of a tensor field

For our purpose, let us consider a $(0,2)$ tensor field T and consider the change under coordinate transformation.

$$
\begin{aligned}
T'_{\mu\nu}(\bar{\mathbf{x}}') &= \big(\delta^\alpha{}_\mu-\varepsilon\partial_\mu\xi^\alpha\big)\Big(\delta^\beta{}_\nu-\varepsilon\partial_\nu\xi^\beta\Big)T_{\alpha\beta}(\bar{\mathbf{x}}) \\
&= T_{\mu\nu}(\bar{\mathbf{x}})-\varepsilon\Big(\partial_\nu\xi^\beta T_{\mu\beta}+\partial_\mu\xi^\alpha T_{\alpha\nu}\Big)+\mathcal{O}(\varepsilon^2)
\end{aligned}
$$

Then let us calculate the difference:

$$
\begin{aligned}
T_{\mu\nu}(\bar{\mathbf{x}}')-T'_{\mu\nu}(\bar{\mathbf{x}}') &= T_{\mu\nu}(\bar{\mathbf{x}}+\varepsilon\boldsymbol{\xi})-T_{\mu\nu}(\bar{\mathbf{x}})+\varepsilon\Big(\partial_\nu\xi^\beta T_{\mu\beta}+\partial_\mu\xi^\alpha T_{\alpha\nu}\Big) \\
&= \cancel{T_{\mu\nu}(\bar{\mathbf{x}})}+\varepsilon\xi^\rho\partial_\rho T_{\mu\nu}(\bar{\mathbf{x}})-\cancel{T_{\mu\nu}(\bar{\mathbf{x}})}+\varepsilon\Big(\partial_\nu\xi^\beta T_{\mu\beta}+\partial_\mu\xi^\alpha T_{\alpha\nu}\Big)
\end{aligned}
$$

Then from the previous definition the Lie derivative is found out to be:

$$
\boxed{\mathcal{L}_{\boldsymbol{\xi}}T_{\mu\nu} = \xi^\rho\nabla_\rho T_{\mu\nu}+\nabla_\nu\xi^\beta T_{\mu\beta}+\nabla_\mu\xi^\alpha T_{\alpha\nu}}
$$

30.4. Killing vectors

We can now characterise the isometries of the metric by defining them to be transformations under which the Lie derivative of the metric vanishes, that is,

$$\mathcal{L}_{\boldsymbol{\xi}} g_{\mu\nu} = 0$$

Such a vector field $\boldsymbol{\xi}$ is called a *Killing vector field*. From the expression above of the Lie derivative of a $(0, 2)$ tensor field, the above condition becomes,

$$\mathcal{L}_{\boldsymbol{\xi}} g_{\mu\nu} = \xi^\rho \nabla_\rho g_{\mu\nu} + \nabla_\nu \xi^\beta g_{\mu\beta} + \nabla_\mu \xi^\alpha g_{\alpha\nu}$$

From metric compatibility, we know that $\nabla_\rho g_{\mu\nu} = 0$ and hence the first term vanishes. And we can put the metric inside the covariant derivatives in the remaining terms, which will lower the indices of ξ. We finally obtain:

$$\mathcal{L}_{\boldsymbol{\xi}} g_{\mu\nu} \equiv \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0$$

This is a differential equation, referred to as the *Killing equation*. After solving this, we will get the Killing vector fields which will provide us with the isometries of the metric.

FUN FACT: The collection of all isometries form a group, called the *isometry group* and it is typically a Lie group. The Killing vector fields are the *generators* of this Lie group.

Reflection might also be an *isometry* of the metric, however, these are discrete transformations and hence cannot be obtained from the Killing equation, since these only generate the continuous transformations.

Lecture 31: Killing vectors for 2D metric

In the previous section, we discussed about the Killing equation and how we can derive the *isometries* of a metric from the equation. We will now see an explicit example of that, using the metric for 2D flat space,

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$$

From this we can find the metric and the inverse metric,

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Intuitively, we can 'see' from the metric itself, that the isometries are going to be *rotations* and *translations*. Let us now see whether the Killing equation says the same or not.

Note that, since the metric is constant, the Christoffel symbols are all zero and hence covariant and partial derivatives coincide. We then have three equations:

$$\begin{aligned} \partial_x \xi_x + \partial_x \xi_x &= 0 \\ \partial_y \xi_y + \partial_y \xi_y &= 0 \\ \partial_x \xi_y + \partial_y \xi_x &= 0 \end{aligned}$$

The first (second) equation tells us that ξ_x (ξ_y) is independent of x (y) and hence we can write,

$$\xi_x \equiv \xi_x(y) \qquad \xi_y \equiv \xi_y(x)$$

Now let us take the third equation and take the partial derivative with respect to x,

$$0 = \partial_x^2 \xi_y + \partial_x \partial_y \xi_x = \partial_x^2 \xi_y + \partial_y \cancelto{0}{\partial_x \xi_x} = \partial_x^2 \xi_y \implies \boxed{\xi_y = A_2 x + B_2}$$

where A_2 and B_2 are constants to be determined. Similarly, taking partial derivative with respect to y we obtain,

$$0 = \partial_y^2 \xi_x + \partial_y \partial_x \xi_y = \partial_y^2 \xi_x + \partial_x \cancelto{0}{\partial_y \xi_y} = \partial_y^2 \xi_x \implies \boxed{\xi_x = A_1 y + B_1}$$

Substituting these in the third Killing equation again gives us $A_1 + A_2 = 0 \implies A_1 = -A_2$. Now, we want to characterise the vector field $\boldsymbol{\xi}$ and for that, we can write it in terms of some basis as

$$\boldsymbol{\xi} = \xi^1 \mathbf{e}_1 + \xi^2 \mathbf{e}_2$$

As we are using the Cartesian coordinates, a natural choice of basis is $\mathbf{e}_i = \partial_i$ and using this, we obtain the Killing vector field as:

$$\boldsymbol{\xi} = (A_1 y + B_1)\partial_x + (-A_1 x + B_2)\partial_y = B_1 \partial_x + B_2 \partial_y - A_1(x\partial_y - y\partial_x)$$

It is evident that the Killing vector fields are composed of three different quantities and we can now choose to focus on them one by one.

▷ $\boxed{B_1 = 1, B_2 = 0, A_1 = 0}$

$\boldsymbol{\xi}^{(1)} = \partial_x$ is one of the infinitesimal generators of the isometry group. To see the finite transformation generated by this, we will exponentiate this.

$$\exp\left(a\boldsymbol{\xi}^{(1)}\right) = \exp\left(a\partial_x\right) = \sum_{n=0}^{\infty} \frac{a^n}{n!}\partial_x^n \equiv \mathsf{T}^1_{\mathsf{a}}$$

Now, let us act this on the coordinates:

$$\begin{aligned} \mathsf{T}^1_{\mathsf{a}}(x) &= \sum_{n=0}^{\infty} \frac{a^n}{n!}\frac{\partial x}{\partial x^n} = x + a \\ \mathsf{T}^1_{\mathsf{a}}(y) &= \sum_{n=0}^{\infty} \frac{a^n}{n!}\frac{\partial y}{\partial x^n} = y + 0 = y \end{aligned}$$

Thus, this generates a *translation in x direction*.

▷ $\boxed{B_1 = 0, B_2 = 1, A_1 = 0}$

$\boldsymbol{\xi}^{(2)} = \partial_y$ is another infinitesimal generator of the isometry group. Similarly as before, we will exponentiate this.

$$\exp\left(a\boldsymbol{\xi}^{(2)}\right) = \exp\left(a\partial_y\right) = \sum_{n=0}^{\infty} \frac{a^n}{n!}\partial_y^n \equiv \mathsf{T}^2_{\mathsf{a}}$$

Now, let us act this on the coordinates:

$$\begin{aligned} \mathsf{T}^2_{\mathsf{a}}(x) &= \sum_{n=0}^{\infty} \frac{a^n}{n!}\frac{\partial x}{\partial y^n} = x + 0 \\ \mathsf{T}^2_{\mathsf{a}}(y) &= \sum_{n=0}^{\infty} \frac{a^n}{n!}\frac{\partial y}{\partial y^n} = y + a = y + a \end{aligned}$$

Thus, this generates a *translation in y direction*.

▷ $\boxed{B_1 = 0, B_2 = 0, A_1 = -1}$

$\boldsymbol{\xi}^{(3)} = x\partial_y - y\partial_x$, which is a bit non-trivial but manageable. For that, first observe that,

$$\boldsymbol{\xi}^{(3)}(x) = -y \quad [\boldsymbol{\xi}^{(3)}]^2(x) = -x \qquad \boldsymbol{\xi}^{(3)}(y) = x \quad [\boldsymbol{\xi}^{(3)}]^2(y) = -y$$

Now, let us exponentiate the infinitesimal generator again to have some finite transformation,

$$\exp\left(\theta\boldsymbol{\xi}^{(3)}\right) = \sum_{n=0}^{\infty} \frac{\theta^n}{n!}[\boldsymbol{\xi}^{(3)}]^n \equiv \mathsf{R}_\theta$$

We can separate this into odd and even parts, since we know what happens when the generator acts odd and even number of times on a particular coordinate.

$$\begin{aligned}\mathsf{R}_\theta(x) &= \sum_{n\in\text{odd}}^{\infty} \frac{\theta^n}{n!}[\boldsymbol{\xi}^{(3)}]^n(x) + \sum_{n\in\text{even}}^{\infty} \frac{\theta^n}{n!}[\boldsymbol{\xi}^{(3)}]^n(x) \\ &= \sum_{n=0}^{\infty} \frac{\theta^{2n+1}(-1)^{n+1}}{(2n+1)!}\, y + \sum_{n=0}^{\infty} \frac{\theta^{2n}(-1)^n}{(2n)!}\, x \\ &= x\cos\theta - y\sin\theta\end{aligned}$$

Similarly, acting on the coordinate y, we have:

$$\begin{aligned}\mathsf{R}_\theta(y) &= \sum_{n\in\text{odd}}^{\infty} \frac{\theta^n}{n!}[\boldsymbol{\xi}^{(3)}]^n(y) + \sum_{n\in\text{even}}^{\infty} \frac{\theta^n}{n!}[\boldsymbol{\xi}^{(3)}]^n(y) \\ &= \sum_{n=0}^{\infty} \frac{\theta^{2n+1}(-1)^{n}}{(2n+1)!}\, x + \sum_{n=0}^{\infty} \frac{\theta^{2n}(-1)^n}{(2n)!}\, y \\ &= x\sin\theta + y\cos\theta\end{aligned}$$

Thus, the finite transformation generated is rotation by some angle θ.

We see that the generators of the isometries of the metric are two translations and one rotation as was expected. The *isometry group* is called the proper *Euclidean group*, SE(2). This does not include reflections, though. If we included reflections too, the isometry group would be generalised to E(n), for an Euclidean space $\mathbb{E}^n$.

Lecture 32: Killing vectors for Minkowski space

In the previous section, we saw the isometry group of the 2D flat metric to be the proper Euclidean group. We will now focus on the Minkowski metric and characterise its isometries.

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \qquad g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

In this case also, the Christoffel symbols are zero, since the metric is constant and covariant and partial derivatives are interchangeable. The Killing equation thus becomes

$$\partial_\mu\xi_\nu + \partial_\nu\xi_\mu = 0 \tag{42}$$

Let us differentiate this expression,

$$\partial_\sigma\partial_\mu\xi_\nu + \partial_\sigma\partial_\nu\xi_\mu = 0 \tag{43}$$

Replacing $\sigma \to \mu$, $\mu \to \nu$ and $\nu \to \sigma$ in Eq. 43 we get:

$$\partial_\mu\partial_\nu\xi_\sigma + \partial_\mu\partial_\sigma\xi_\nu = 0 \tag{44}$$

Replacing $\sigma \to \mu$, $\mu \to \nu$ and $\nu \to \sigma$ in Eq. 44 we get:

$$\partial_\nu \partial_\sigma \xi_\mu + \partial_\nu \partial_\mu \xi_\sigma = 0 \tag{45}$$

Adding Eq. 43 and Eq. 44 and subtracting Eq. 45 we get:

$$2\partial_\sigma(\partial_\mu \xi_\nu) = 0 \implies \partial_\mu \xi_\nu = \mathcal{B}_{\mu\nu} \longrightarrow \text{ constant tensor}$$

Integrating the expression we obtain

$$\xi_\nu = \mathcal{A}_\nu + \mathcal{B}_{\rho\nu} x^\rho$$

This is the general solution of the Killing equation and hence, must satisfy Eq. 42. Substituting this in Eq. 42 we get:

$$\mathcal{B}_{\mu\nu} + \mathcal{B}_{\nu\mu} = 0$$

Hence $\mathcal{B}_{\mu\nu}$ is an anti-symmetric tensor and hence it can have atmost sixe independent entries (diagonals are zero and upper triangle elements are negative of lower triangle elements. So total $(16-4)/2 = 6$ independent entries).

$\mathcal{A}_\mu$ has no such restrictions so it has *four* independent components and $\mathcal{B}_{\mu\nu}$ has *six* independent components, so in total, 10 independent constants are there which corresponds to 10 independent Killing vector fields possible for the Minkowski metric.

Let us now write the Killing vector field in terms of the basis

$$\boldsymbol{\xi} = \xi^\mu \partial_\mu = g^{\mu\nu} \xi_\nu \partial_\mu = g^{\mu\nu}(\mathcal{A}_\nu + \mathcal{B}_{\rho\nu} x^\rho)\partial_\mu$$

Let us again consider different cases to explicitly see the different generators of the isometry group.

- $\boxed{\mathcal{A}_0 = -1, \mathcal{A}_i = 0, \mathcal{B}_{\mu\nu} = 0}$

  $\boldsymbol{\xi}^{(1)} = g^{00}\partial_0 = \partial_t$ is the infinitesimal generator. To obtain the finite transformation, we use exponentiation as before:

  $$\exp\left(\varepsilon \boldsymbol{\xi}^{(1)}\right) \equiv \exp(\varepsilon \partial_t) = \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} \partial_t^n \equiv \mathsf{T}_\varepsilon$$

  Then note that $\mathsf{T}_\varepsilon(t) = t + \varepsilon$ and $\mathsf{T}_\varepsilon(x^i) = 0$. Thus this generator corresponds to *time translation* or *time evolution*.

- $\boxed{\mathcal{A}_i = 1, \mathcal{A}_0 = 0, \mathcal{B}_{\mu\nu} = 0}$

  $\boldsymbol{\xi}^{(i+1)} = g^{ii}\partial_i = \partial_i$ is the infinitesimal generator. We had already seen this before, that the generator corresponds to *spatial translation* in the x, y or z direction. We already found four generators corresponding to $\mathcal{A}_\mu$ which are temporal and spatial translations.

- $\boxed{\mathcal{A}_\mu = 0, \mathcal{B}_{0i} = 1 = -\mathcal{B}_{i0}}$

  We then have the generator as, $\boldsymbol{\xi}^{(i+4)} = g^{ii}\mathcal{B}_{0i} x^0 \partial_i + g^{00}\mathcal{B}_{i0} x^i \partial_0 = t\partial_i + x^i \partial_t$. This is a somewhat non-trivial thing, though this looks kinda like the rotation generator as seen in the 2D flat metric. Let us exponentiate to investigate this further.

  $$\exp\left(w \boldsymbol{\xi}^{(i+4)}\right) \equiv \exp\left(w(t\partial_i + x^i \partial_t)\right) = \sum_{n=0}^{\infty} \frac{w^n}{n!}(t\partial_i + x^i\partial_t)^n \equiv \Lambda_w^i$$

  Note that $(t\partial_i + x^i\partial_t)t = x^i$ and $(t\partial_i + x^i\partial_t)x^i = t$ and hence acting this on t odd number of times, generates x^i and even number of times, generates t (and vice versa).

  $$\Lambda_w^i(t) = \sum_{n=0}^{\infty} \frac{w^{2n}}{(2n)!}\, t + \sum_{n=0}^{\infty} \frac{w^{2n+1}}{(2n+1)!}\, x = t\cosh w + x^i \sinh w$$

  $$\Lambda_w^i(x) = \sum_{n=0}^{\infty} \frac{w^{2n}}{(2n)!}\, x^i + \sum_{n=0}^{\infty} \frac{w^{2n+1}}{(2n+1)!}\, t = t\sinh w + x^i \cosh w$$

This might be a little bit difficult to identify at the first glance. Consider the usual Lorentz transformations,

$$t' = \gamma(t - vx) \equiv \gamma(t - \beta x) \qquad x'^i = \gamma(x^i - vt) \equiv \gamma(x^i - \beta t)$$

where $\beta = v$. Define a quantitiy *rapidity*, w, such that $\tanh w := -\beta$ which gives us the parameters,

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\tanh^2 w}} = \frac{\cosh w}{\sqrt{\cosh^2 w - \sinh^2 w}} = \cosh w$$
$$\beta\gamma = \tanh w \cosh w = -\sinh w$$

Then the Lorentz transformation in terms of rapidity becomes:

$$t' = t\cosh w + x \sinh w \qquad x'^i = x^i \cosh w + t \sinh w$$

This is exactly the form that we had obtained from the Killing equation and hence we can safely say that this Killing vector field generates *Lorentz boosts* along x,y or z direction.

▷ $\boxed{\mathcal{A}_\mu = 0, \mathcal{B}_{ij} = 1 = -\mathcal{B}_{ji}}$

This is again familiar to us as $\boldsymbol{\xi}^{(i+7)} = g^{jj}\mathcal{B}_{ij}x^i\partial_j + g^{ii}\mathcal{B}_{ji}x^j\partial_i = x^i\partial_j - x^j\partial_i$ was seen to be the generator of rotation about x,y or z axis.

This completes the identification of all the 10 Killing vector fields for the Minkowski metric. These include three *spatial translations*, one *time translation*, three *rotations* and three *Lorentz boosts*. This is exactly the Poincaré group and hence, the *isometry group* of the Minkowski metric is the *Poincaré group*.

Lecture 33: Parallel Transport

HI

Appendices

Appendix A Code for Christoffel Symbols

In[78]:=
```
(*Defining our coordinates*)coords = {t, r, θ, ϕ};

(*Defining the metric*)
g = {{f[r], 0, 0, 0}, {0, h[r], 0, 0}, {0, 0, r^2, 0}, {0, 0, 0, r^2 Sin[θ]^2}};

(*Defining the inverse metric*)
ginv = Simplify[Inverse[g]];

n = Length[coords];

(*Defining the Christoffel symbols  Γ^ρ_μν*)
Gamma1 =
  Table[1/2 Sum[ginv[[ρ, σ]] (D[g[[σ, ν]], coords[[μ]]] + D[g[[σ, μ]], coords[[ν]]] - D[g[[μ, ν]], coords[[σ]]]),
     {σ, 1, n}], {ρ, 1, n}, {μ, 1, n}, {ν, 1, n}];

(*Simplifying the results*)
Gamma2 = Simplify[Gamma1];

(*Printing the Christoffel symbols in table format*)
TableForm[Gamma2, TableHeadings → {coords, coords}]
```

Out[84]//TableForm=

	t	r	θ	ϕ
t	0 $\frac{f'[r]}{2\,f[r]}$ 0 0	$\frac{f'[r]}{2\,f[r]}$ 0 0 0	0 0 0 0	0 0 0 0
r	$-\frac{f'[r]}{2\,h[r]}$ 0 0 0	0 $\frac{h'[r]}{2\,h[r]}$ 0 0	0 0 $-\frac{r}{h[r]}$ 0	0 0 0 $-\frac{r\,\mathrm{Sin}[\theta]^2}{h[r]}$
θ	0 0 0 0	0 0 $\frac{1}{r}$ 0	0 $\frac{1}{r}$ 0 0	0 0 0 $-\mathrm{Cos}[\theta]\,\mathrm{Sin}[\theta]$
ϕ	0 0 0 0	0 0 0 $\frac{1}{r}$	0 0 0 $\mathrm{Cot}[\theta]$	0 $\frac{1}{r}$ $\mathrm{Cot}[\theta]$ 0

Appendix B Code for Riemann and Ricci Tensor

In[103]:=

```
coords = {t, r, θ, ϕ};

(*Defining the metric*)
g = {{f[r], 0, 0, 0}, {0, h[r], 0, 0}, {0, 0, r^2, 0}, {0, 0, 0, r^2 Sin[θ]^2}};

(*Defining the inverse metric*)
ginv = Simplify[Inverse[g]];

n = Length[coords];
Gammachris[ρ_, μ_, ν_] :=
  1/2 Sum[ginv[[ρ, σ]] (D[g[[σ, ν]], coords[[μ]]] + D[g[[σ, μ]], coords[[ν]]] - D[g[[μ, ν]], coords[[σ]]]), {σ, 1, n}]

RieTensor[ρ_, σ_, μ_, ν_] :=
  D[ Gammachris[ρ, ν, σ], coords[[μ]]] - D[ Gammachris[ρ, μ, σ], coords[[ν]]] +
   Sum[Gammachris[ρ, μ, λ] × Gammachris[λ, ν, σ], {λ, 1, n}] -
   Sum[Gammachris[ρ, ν, λ] × Gammachris[λ, μ, σ], {λ, 1, n}]
RiccTensor[μ_, ν_] := Sum[RieTensor[ρ, μ, ρ, ν], {ρ, 1, n}]

Print["Non-zero components of Ricci tensor R_{μν}:"];

Do[Module[{val = Simplify[RiccTensor[μ, ν]]},
   If[val =!= 0, Print["R_", coords[[μ]], coords[[ν]], " = ", val]]], {μ, 1, n}, {ν, 1, n}];
```

Non-zero components of Ricci tensor R_{μν}:

$$\text{R\_tt} = \frac{r\, h[r]\, f'[r]^2 + f[r]\left(r\, f'[r]\, h'[r] - 2\, h[r]\left(2\, f'[r] + r\, f''[r]\right)\right)}{4\, r\, f[r]\, h[r]^2}$$

$$\text{R\_rr} = \frac{f[r]\left(4\, f[r] + r\, f'[r]\right) h'[r] + r\, h[r]\left(f'[r]^2 - 2\, f[r]\, f''[r]\right)}{4\, r\, f[r]^2\, h[r]}$$

$$\text{R\_}\theta\theta = \frac{1}{2}\left(2 - \frac{2 + \frac{r\, f'[r]}{f[r]}}{h[r]} + \frac{r\, h'[r]}{h[r]^2}\right)$$

$$\text{R\_}\phi\phi = \frac{\mathrm{Sin}[\theta]^2\left(-r\, h[r]\, f'[r] + f[r]\left(-2\, h[r] + 2\, h[r]^2 + r\, h'[r]\right)\right)}{2\, f[r]\, h[r]^2}$$

Appendix C Code for Interior solution

```
In[27]:= (*Defining the Coordinates*)
coords = {t, r, θ, ϕ};

(*Defining the Metric*)
g = {{-Exp[2 Φ[r]], 0, 0, 0}, {0, Exp[2 ν[r]], 0, 0}, {0, 0, r^2, 0}, {0, 0, 0, r^2 Sin[θ]^2}};

ginv = Simplify[Inverse[g]];

(*Definining the Christoffel symbols*)
Γ[ρ_, μ_, ω_] := 1/2 Sum[ginv[[ρ, sigma]] (D[g[[sigma, ω]], coords[[μ]]] +
        D[g[[sigma, μ]], coords[[ω]]] - D[g[[μ, ω]], coords[[sigma]]]), {sigma, 1, 4}];

(*Defining the Ricci tensor*)
R[μ_, ω_] := Sum[D[Γ[ρ, ω, μ], coords[[ρ]]] - D[Γ[ρ, ρ, μ], coords[[ω]]] + Sum[
      Γ[ρ, ρ, lambda] Γ[lambda, ω, μ] - Γ[ρ, ω, lambda] Γ[lambda, ρ, μ], {lambda, 1, 4}], {ρ, 1, 4}];

(*Defining the Ricci scalar*)
Rscal = Sum[ginv[[μ, ω]] R[μ, ω], {μ, 1, 4}, {ω, 1, 4}];

(*Calculating Einstein tensor*)
G[μ_, ω_] := Simplify[R[μ, ω] - 1/2 g[[μ, ω]] Rscal];

(*Printing non-zero elements of Einstein tensor*)
Print["Non-zero components of Einstein tensor G_{μν}:"];
Do[Module[{val = Simplify[G[μ, ω]]},
   If[val =!= 0, Print["G_", coords[[μ]], coords[[ω]], " = ", val]]], {μ, 1, 4}, {ω, 1, 4}];
```

Non-zero components of Einstein tensor G_{μν}:

G_tt = $\dfrac{e^{-2\nu[r]+2\Phi[r]}\left(-1+e^{2\nu[r]}+2\, r\, \nu'[r]\right)}{r^2}$

G_rr = $\dfrac{1-e^{2\nu[r]}+2\, r\, \Phi'[r]}{r^2}$

G_θθ = $e^{-2\nu[r]}\, r\left(\Phi'[r]+r\,\Phi'[r]^2-\nu'[r]\left(1+r\,\Phi'[r]\right)+r\,\Phi''[r]\right)$

G_ϕϕ = $e^{-2\nu[r]}\, r\, \mathrm{Sin}[\theta]^2\left(\Phi'[r]+r\,\Phi'[r]^2-\nu'[r]\left(1+r\,\Phi'[r]\right)+r\,\Phi''[r]\right)$