General Relativity and Black Holes I

2026 IF Summer School for Young Physicists
on Cosmology, Quantum Information, and
Quantum Field Theory

The Institute for Fundamental study,
Naresuan University

May 27, 2026

Contents

1 From Newtonian Theory to Special Relativity

Black holes is one of the interesting astronomical objects in which even light cannot escape from a particular radius called the Schwarzschild radius. The existence of black holes is formally based on a solution of the Einstein field equation, which plays the main role in the gravitational theory called General Relativity (GR). So, one of the main topics in this lecture, I will give you the main concept of GR before providing you with the formal description of black holes.

In order to formulate GR properly, it is worthwhile to introduce how it significantly differs from the well-known gravitational theory, Newtonian theory. In this lecture, I will review the important concept of Newtonian theory in order to further connect to GR. Newtonian theory is a well-known theory on the motion of objects influenced by something we call force. Even though this theory is suitable to deal with particle motion, it cannot be valid for the high velocity motion of a particle and high gravitational force. For the high velocity (The velocity approaches the speed of light), the motion can be well described by Special Relativity. I will give you an introduction to SR in this lecture after providing the important concept of Newtonian theory.

1.1 Newton's laws

Newtonian theory is based on the Newton's laws. In this subsection, I will give you the description of the Newton's laws in both intuitive and mathematical viewpoints.

Conceptual description

1. Every body continues in its state of the rest, or uniform motion in a straight line, unless it is compelled to change that state by forces impressed upon it.
   - This law describes a common property of matter, namely ”inertia”. Inertia is the resistance of all matter to having its motion changed.
   - This law is used to define the reference frame or initial frame. There exists a reference frame in which the particle motion can be mathematically described.
   - The Galilean transformation will connect to this law and then can be generalized to the Lorentz transformation, which is the key issue for

SR. (We will discuss in detail later.)

2. The change of motion is proportional to the motive force impressed and is made in the direction of the line in which that force is impressed.
   - This law is used to define the motion of particles influenced by the force acting on them.
   - For the gravitational force, this law will be generalized to GR by using the concept of spacetime curvature. (We will discuss in detail later.)

3. To every action there is always imposed an equal reaction; or the mutual actions of two bodies upon each other are always equal and directed to the contrary part.
   - It is a consequence of the second law in case the particle does not move following the force. The particle has to be impressed with the force on the opposite side.
   - This is a kind to define the conservation law initiated from our intuition.

Actually, the laws are based on the common sense of our intuition. Everyone seems to know these in everyday life. Newton just states them precisely in order to collaborate with mathematical language.

Mathematical description

1. One of the most important things to describe the particle motion is to identify the position of the particle at time t. For a given reference frame system (discuss later), the position of a particle can be specified by a vector $\vec{r}(t)$. For the Cartesian coordinates the vector can be written as
$$\vec{r}(t) = x\hat{i} + y\hat{j} + z\hat{k}.$$
   - Since the position of the particle changes with time, the velocity of the particle can be given by
$$\vec{v}(t) = \frac{d\vec{r}}{dt} = v_x\hat{i} + v_y\hat{j} + v_z\hat{k}.$$
   - The rate of change of the velocity of the particle can be described by

a acceleration vector

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2} = a_x\hat{i} + a_y\hat{j} + a_z\hat{k}.$$

2. The motion of particle of mass m due to the force $\vec{F}_i$ can be described by
$$\sum_i \vec{F}_i = m\vec{a}.$$

3. In the case of particle case the particle does not move following the force, there exist the reaction such that
$$\sum_i \vec{F}_i = 0, \quad \rightarrow \quad \vec{F}_{ac} = -\vec{F}_{re}.$$

1.2 Conservation law

Linear momentum

It is worthwhile to write the second law in the other form by using the notion of linear momentum $\vec{p} = m\vec{v}$ (The ability of the particle to impact the other objects),

$$\vec{F} = m\frac{d\vec{v}}{dt} = \frac{d(m\vec{v})}{dt} = \frac{d(\vec{p})}{dt}.$$

From this equation, one can obtain the conservation law for the linear momentum of a particle such that "If the total force is zero, then the linear momentum is conserved $\dot{\vec{p}} = 0$".

Angular momentum

For the particle moving with a direction change, one can find the pivot point and then can define the angular momentum of the particle as follows

$$\vec{L} = \vec{r} \times \vec{p}.$$

The moment of the force or torque about the pivot point can be defined as

$$\vec{N} = \vec{r} \times \vec{F}.$$

By playing the same fashion of the linear momentum case, the torque can be written in terms of the angular momentum as

$$\begin{aligned}\vec{N} &= \vec{r}\times\frac{d}{dt}\left(m\vec{v}\right),\\ &= \frac{d}{dt}\left(\vec{r}\times m\vec{v}\right)-\frac{d\vec{r}}{dt}\times m\vec{v},\\ &= \frac{d}{dt}\left(\vec{r}\times\vec{p}\right)-\cancelto{0}{\vec{v}\times m\vec{v}},\\ &= \frac{d}{dt}\left(\vec{L}\right).\end{aligned}$$

As a result, the conservation of the angular momentum can be obtained such that if the total torque is zero, the angular momentum is conserved, $\dot{\vec{L}}=0$.

Work done
The work done can be defined as the force $\vec{F}$ upon the particle in going from point 1 to point 2. Mathematically, one can write the work done as

$$\begin{aligned}W_{12} &= \int_1^2 \vec{F}\cdot d\vec{s} = \int_1^2 m\frac{d\vec{v}}{dt}\cdot\frac{d\vec{s}}{dt}dt,\\ &= m\int_1^2\frac{d\vec{v}}{dt}\cdot\vec{v}dt = \frac{m}{2}\int_1^2\frac{d}{dt}\left(\vec{v}\cdot\vec{v}\right)dt,\\ &= \frac{m}{2}v^2|_1^2 = \frac{m}{2}\left(v_2^2-v_1^2\right) = T_2-T_1,\end{aligned}$$

where $T=mv^2/2$ is the kinetic energy.

Conservative force
"If the force is such that the work done W_{12} is the same for any physically possible path between point 1 and point 2, then the force is said to be conservative"

$$\oint\vec{F}\cdot d\vec{s} = \int_1^2\vec{F}\cdot d\vec{s}+\int_2^1\vec{F}\cdot d\vec{s}=0.$$

From Stokes' theorem, we have

$$\oint\vec{F}\cdot d\vec{s} = \int\left(\vec{\nabla}\times\vec{F}\right)\cdot d\vec{A}=0,\rightarrow\vec{\nabla}\times\vec{F}=0.$$

From mathematical identity, if $\vec{\nabla}\times\vec{X}=0$, then the vector $\vec{X}$ can be written in terms of the gradient of a scalar function Y as $\vec{X}=\vec{\nabla}Y$. Therefore, if the force is conservative, one can define the potential energy V via such a scalar function as

$$\vec{F} = -\vec{\nabla} V.$$

Note that the existence of the minus sign is just a convention, such that one can find the reference of the potential energy at infinity. Actually, the potential energy can be defined as the work done from infinity to the considered point.

$$\oint \vec{F} \cdot d\vec{s} \;\; = \;\; -\int_1^2 dV = V_1 - V_2 = W_{12} = T_2 - T_1.$$

As a result, one can define the total energy as the sum of potential energy and kinetic energy $E = V + T$, then the conservation of the total energy can be obtained $E_1 = E_2$. "If the force acting on the particle is conservative, then the total energy of the particle is conserved."

1.3 Central force

One can see that if the force is conservative, one can formulate a proper mathematical description of the motion. In fact, for the central force $\vec{F} = \vec{F}(r)$, the force is conservative, so the mathematical description of the motion can be obtained properly. This allows us to define the property of mass as the field around it called the gravitational field.

Field and potential

One of the important properties of the object of mass M is the gravitational field. The gravitational field $\vec{E}_g$ can be defined in the same fashion as the electric field. The gravitational force per unit mass acting on the test mass m_0 at every point in space.

$$\vec{E}_g = \frac{\vec{F}}{m_0} = -\frac{GM}{r^2}\hat{r}.$$

For the gravitation field associated with gravitational force which is conservative, one can define the gravitational potential as an analog quantities of the potential energy as

$$\vec{E}_g = -\vec{\nabla}\Phi \rightarrow \Phi = -\int_\infty^p \vec{E}_g \cdot d\vec{s} = \frac{GM}{r}.$$

One can apply the Gauss's law to the gravitational case as found in the electric case.

$$\int \vec{E}_g \cdot d\vec{A} = -4\pi GM \to E_g = -\frac{GM}{r^2}.$$

By using the divergence theorem, one can obtain the Poisson equation for the gravitational field as

$$\begin{aligned}
\int \vec{E}_g \cdot d\vec{A} &= -4\pi GM, \\
\int \vec{\nabla} \cdot \vec{E}_g \, dV &= -4\pi G \int \rho \, dV, \\
\int \nabla^2(-\Phi) dV &= -4\pi G \int \rho \, dV, \\
\nabla^2 \Phi &= 4\pi G\rho.
\end{aligned}$$

Note that this equation will be connected to the Einstein equation in GR (We will see later).

Effective potential

Now we will apply the mathematical tool we have to analyze the motion of a particle influenced by the gravitational force. It is worthwhile to analyze the motion by using the effective potential. For motion due to the central force, the system has spherical symmetry. For example, the force does not depend on the angular parts θ, ϕ. In this case, there exists the conserved quantity as the angular momentum, $\vec{L} = \vec{r} \times m\dot{\vec{r}} = const.$ It will be clearer if we consider the system with a Lagrangian formulation. However, I do not think we have enough time to study. Then, we can see that the motion is always perpendicular to the vector $\vec{L}$ or the motion is always in the plane perpendicular to $\vec{L}$. In this case, one can choose the proper direction of $\vec{L}$ (mostly pointing to the z direction), and then the magnitude can be written as $L = mr^2\dot{\theta}$. Now, let us consider the total energy which can be given by

$$\begin{aligned}
E &= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}m\, r^2\dot{\theta}^2 - \frac{GMm}{r}, \\
\frac{E}{m} &= \frac{1}{2}\dot{r}^2 + \frac{1}{2}\, r^2\dot{\theta}^2 - \frac{GM}{r}, \\
\bar{E} &= \frac{1}{2}\dot{r}^2 + \frac{l^2}{2r^2} - \frac{GM}{r}, \text{ where } l = \frac{L}{m}, \\
\bar{E} &= \frac{1}{2}\dot{r}^2 + V_{eff}(r), \text{ where } V_{eff}(r) = \frac{l^2}{2r^2} - \frac{GM}{r}.
\end{aligned}$$

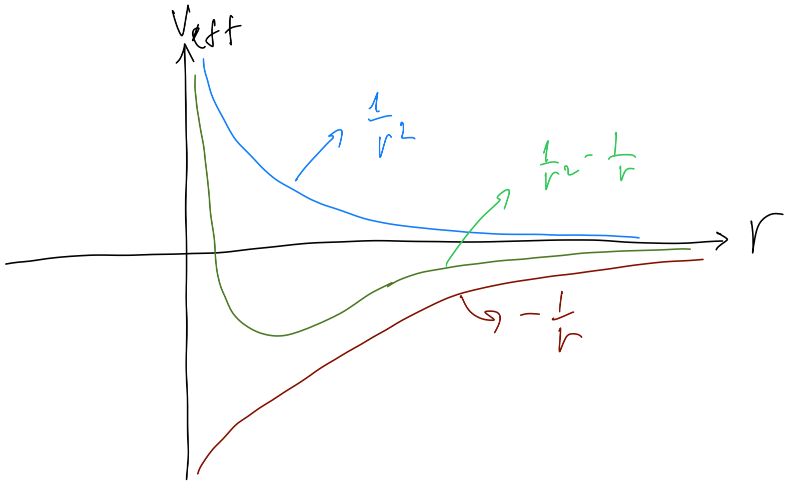


Figure 1: The effective potential

In order to analyze the motion, let us sketch the effective potential as shown in Fig. 1.

From this figure, it is a combination of $-1/r$ and $1/r^2$ terms. Moreover, one can see that the motion can be classified into two main cases depending on the value of $\bar{E}$. For $\bar{E} \leq 0$, the orbit is open, corresponding to parabolic $\bar{E} = 0$ and hyperbolic $\bar{E} > 0$ as shown in Fig. 2. For the closed orbit, it

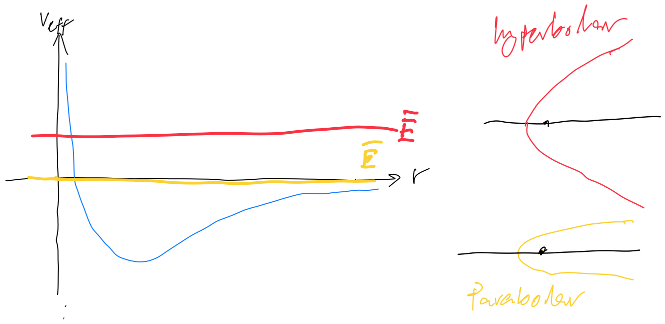


Figure 2: The effective potential and the motion with $\bar{E} \leq 0$.

can be classified into two cases. For $V_{min} < \bar{E} < 0$, the motion is ellipsoidal, and for $\bar{E} = V_{min}$, the motion is circular orbit as shown in Fig. 3. We can

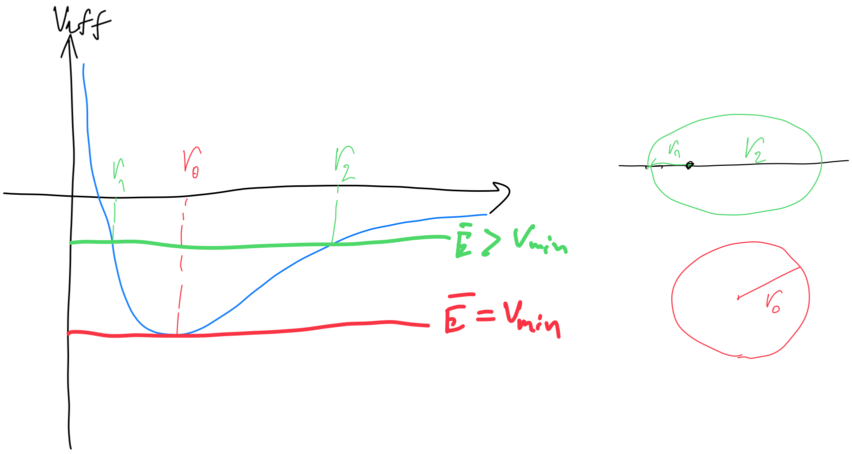


Figure 3: The effective potential and the motion with $V_{min} \leq \bar{E} < 0$.

see that all possible motions can be analyzed by using the effective potential. From the Newtonian theory, the light cannot be influenced by gravitation. How can the motion of light be inferred from this case? (The light motion is a straight line, How?). This is one of the significant concepts different from GR where the light can be influenced by the gravity so that it is possible to have a light orbit. We will address this later.

1.4 Special Relativity

Inertial frame

Before moving our consideration to the SR, let us go back to the 1st law to discuss in detail what the inertial frame of reference is. As I have mentioned before, the existence of this frame allows us to define the coordinates to identify the position of the particle. In this consideration, the frame can be defined by demanding 3 main properties:

1. Rigidity: One can imagine the rigid frame as a rigid body as shown in Fig. 4. In GR, it is allowed to be flexible

2. Synchronized time: all positions in the rigid body have a clock that is synchronized with the others. In GR, it is possible to have clocks ticking at different rates.

3. Euclidean space: The interval can be written as $\Delta s^2 = \Delta x^2 + \Delta y^2 +$

Δz^2. The interval is invariant under frame transformation, which is known as the Galilean transformation. In SR, it is possible to include time, and then the interval is invariant under Lorentz transformation.

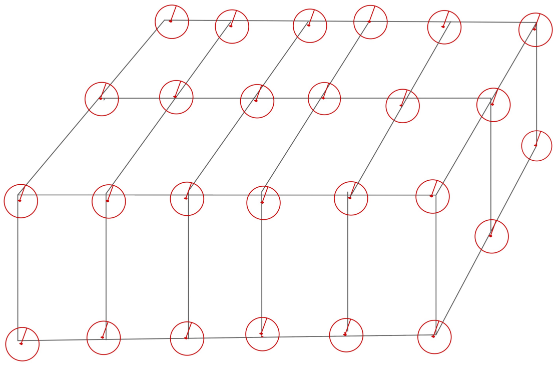


Figure 4: The image of the inertial frame.

Galilean transformation

The inertial frame is constructed in order to define the dynamics of a particle properly. This kind of frame is ideal but still sufficiently adopted, e.g., the Earth frame. This satisfies our intuition. The key importance of this frame is that there exists the transformation that leaves the interval invariant. This also satisfies our intuition. For example, if we consider the moving frame with constant velocity u in x direction as shown in Fig 5, the position of the two frames can be related by

$$x' = x - ut, \quad y' = y, \quad z' = z, \quad t' = t.$$

Moreover, for the rotating frame, it can be seen explicitly that the Δs^2 is invariant as shown in Fig. 6. In this case, one can mathematically write down as

$$\Delta s^2 = \Delta x^2 + \Delta y^2 = \Delta x'^2 + \Delta y'^2.$$

The important key here is that the intervals Δs and Δt are invariant separately. Note that we will see this transformation formally in the next section.

SR postulates

The Newtonian mechanics seems to be useful to describe the particle motion. However, it has been demonstrated since 1888 that the speed of light

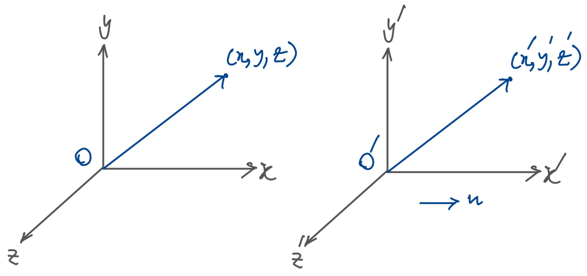


Figure 5: The image of a moving frame with respect to the inertial frame.

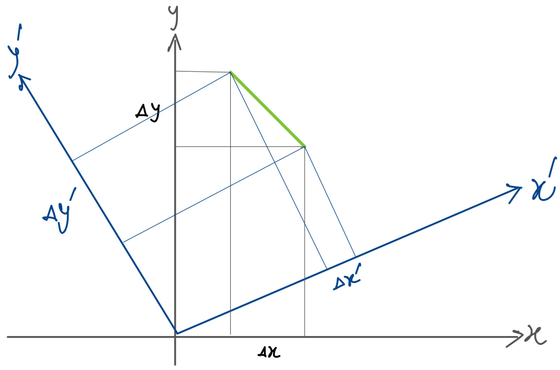


Figure 6: The image of rotating frame with respect to the inertial frame.

is the same in all reference frames, $c = 3 \times 10^8 m/s$. (1888 was the year of the Michaelson-Morley experiment). This does not satisfy our common sense. Eventually, the resolution came from Einstein (principle of relativity in 1905). The principle of relativity is based on two postulates:

1. All inertial frames are equivalent for the performance of all physical experiments. This postulate implies that there is no absolute space, there is no ether, there is no absolute velocity of an inertial frame, and all inertial frames are totally equivalent.

2. Light travels rectilinearly at speed c in every direction and in every inertial frame. This postulate implies that we need the space and time transform together in order to keep the speed of light the same. Then the Galilean transformation needed to be replaced by another proper

one.

It was found that Maxwell's equations in electrodynamic theory are invariant under the Lorentz transformation;

$$\begin{aligned} x' &= \gamma\left(x-ut\right), \quad y'=y, \quad z'=z, \quad t'=\gamma\left(t-\frac{ux}{c^2}\right), \text{ or} \\ x &= \gamma\left(x+ut'\right), \quad y=y', \quad z=z', \quad t=\gamma\left(t'-\frac{ux'}{c^2}\right), \\ \text{where } \gamma &= \left(1-\frac{u^2}{c^2}\right)^{-1/2}. \end{aligned}$$

Einstein introduced this transformation to mechanics, while keeping the speed of light constant, and proposed the idea that "time is relative and is not absolute. As a result, the interval is generalized to

$$\Delta s^2 = -c^2\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2.$$

Then the Euclidean space is now promoted to the Minkowski spacetime.

Time dilation and Length contraction

Let us find the consequences of the Lorentz transformation in mechanics. Considering two trains labeled by A and B with length L, the train passes another with the relative velocity v in the x direction. Two events E_1 and E_2 are assumed as follows

E_1 : Front of B passes front of A $(t'_1=0, x'_1=0)$, $(t_1=0, x_1=0)$

E_2 : Rear of B passes front of A $(t'_2=T', x'_2=-L)$, $(t_2=T, x_2=0)$

These events can be illustrated in Fig. 7. Now, let us apply these events to the Lorentz transformation.

$$\begin{aligned} \Delta x' = (x'_2 - x'_1) = -L &= \gamma\left(0 - v(t_2-t_1)\right), \\ &= -\gamma v T, \\ \Rightarrow & \quad \boxed{L = \gamma v T.}, \end{aligned}$$

$$\Delta t' = t'_2 - t'_1 = T' = \gamma T$$

Since $\gamma > 1$, the time interval in the moving frame is greater than one in the stationary frame. In other words, the A's clock goes slower than B's clock.

In other difference situation, when there are no other references, one can think that the train A is moving to the left while the train B stays at rest.

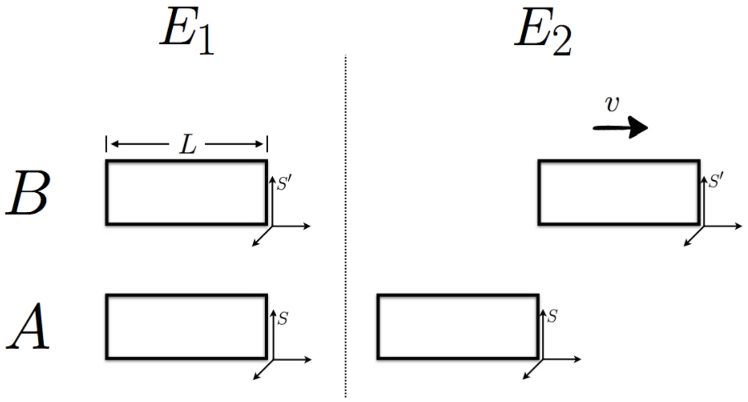


Figure 7: Einstein Trains

In this case, the people in train B will see the clock slower than that in the train A.
Whose clock is really going slower? (Twin Paradox)
This is one of the key ideas of SR, "only relative motion has physical significance". There is no meaning to say that A is at rest, B is moving, or otherwise. The situation is known as time dilation. Now, let us move to the other consequence. In this case, one can figure out what the length of the train B observed by observer in train A, labeled by L_A, By applying the Lorentz transformation, one obtains

$$L_A = x_2 - x_1 = \gamma^{-1}\left(x_2' - x_1'\right) = \gamma^{-1} L_B = \frac{L}{\gamma}.$$

From this expression, one can see that the observer in the train A measures the length of the train B as shorter than that in the train B. Also, in a similar manner, if an observer in the train B tries to measure the length of the train A, he would measure the length of the train A shorter than that of the train A.

One can see that the time interval and length are not invariant under the Lorentz transformation. In other words, $dl^2 = dx^2 + dy^2 + dz^2$ is not invariant. As we have mentioned, the quantity which in invariant under Lorentz transformation is the spacetime interval $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$. In our case, $dy = dz = 0$, then we have

$$ds^2 = -c^2dt^2 + dx^2 = -c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 = -c^2T^2.$$

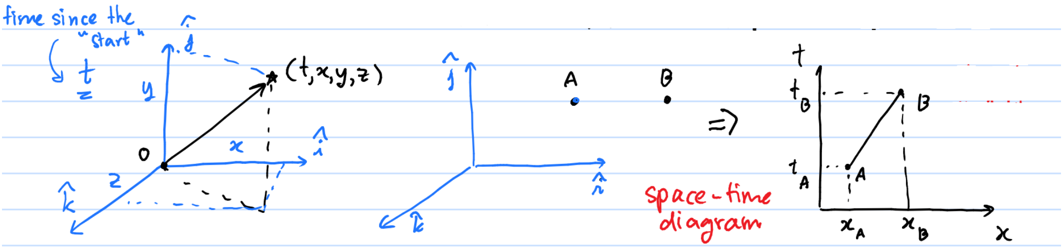
Let us check that this is invariant for one in the moving frame.

$$
\begin{aligned}
ds^2 &= -c^2(t'_2-t'_1)^2+(x'_2-x'_1)^2+0+0,\\
&= -c^2\gamma^2T^2+\gamma^2v^2T^2=-c^2T^2\cancelto{1}{\left[\gamma^2\left(1-\frac{v^2}{c^2}\right)\right]},\\
&= -c^2T^2.
\end{aligned}
$$

2 Special Relativity: Formal approach

2.1 Galilean transformation

A reference frame is a 1-1 correspondence between a physical event and $\mathbb{R}^4$ space (t,x,y,z) where each value represents how far the event is from the reference point in space & time



For example, a train leaving station A at time $t_A \rightarrow (t_A, x_A, 0, 0)$, after that the train arrives at the station B at time $t_B \rightarrow (t_B, x_B, 0, 0)$, the train speed in this reference frame $= (x_B - x_A)/(t_B - t_A)$

Inertial frame = a reference frame in which the motions of free particles are rectilinear

$$\vec{r}(t) = \vec{u}t + \vec{r}_0$$

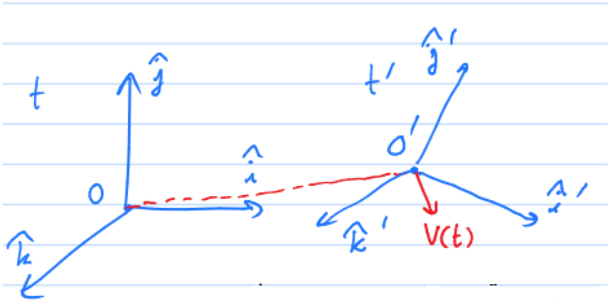
where $\vec{u}$ and $\vec{r}_0$ are constants.

Remarks:

1. $\vec{r}$ is a 3D vector (x, y, z)
2. Free (inertial) particles = particles with no force (influence) acting on them
3. This is actually the first law of motion in Newtonian (writing is a fancy but simple way)

What the Galilean transformations are

A transformation of spacetime is a change in reference frame or a change in the reference point (space & time).



For example, the reference frame O' (point O') could be moving with respect to O, or the frame of axes is rotating with respect to O.

- A change in reference frame → a change in the value of (t, x, y, z)

- Galilean transformation = a spacetime transformation which leaves the following structures invariant

1. Time intervals of any two events $\Delta t = t_2 - t_1$ (Synchronized time)
2. Spatial distances of any two events which happen at the same time (a set of simultaneous events) $\Delta S = |\vec{r}_2 - \vec{r}_1|$ (Rigid)
3. The rectilinear motion of free particles $\vec{r}(t) = \vec{u}t + \vec{r}_0$ where $\vec{u}, \vec{r}_0$ are arbitrary constant vectors (Here, invariance means $\vec{u}, \vec{r}_0$ are different after transformation, but the motion takes the same form)

It turns out that all Galilean transformations have the form (in coordinate transformation form)

$$\begin{aligned} t' &= t + a \quad (a = \text{constant}) && (1) \\ \vec{r}' &= \tilde{R}\vec{r} - \vec{v}t + \vec{b} && (2) \end{aligned}$$

where $\tilde{R}$ is a rotation matrix $\tilde{R}^T\tilde{R} = \mathbb{1}$ and $\vec{v}, \vec{b} =$ constants

<u>Proof</u>:
property 1: (time interval invariant) is easy to see

$$\Delta t' = t'_2 - t'_1 = t_2 + a - (t_1 + a) = t_2 - t_1 = \Delta t \tag{3}$$

property 2: (spatial distance invariant) needs a little calculation. Let's us consider the first part only $\vec{r}' = \tilde{R}\vec{r}$

$$\begin{aligned}\Delta S' &= |\vec{r}_2{}' - \vec{r}_1{}'| = \sqrt{(\vec{r}_2{}' - \vec{r}_1{}') \cdot (\vec{r}_2{}' - \vec{r}_1{}')} \qquad (4)\\ \Delta S'^2 &= (\vec{r}_2{}' - \vec{r}_1{}') \cdot (\vec{r}_2{}' - \vec{r}_1{}') = \Delta\vec{r}' \cdot \Delta\vec{r}' \qquad (5)\end{aligned}$$

One useful thing about the inner product of vectors is the matrix representation of vectors: in Cartesian coordinates

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z = \begin{pmatrix} A_x & A_y & A_z \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}$$

This means that we can write a vector in column form.

(choosing & agreeing or using Cartesian coordinates)

$$\vec{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \rightarrow |\vec{A}|^2 = \vec{A} \cdot \vec{A} = A^T A = \begin{pmatrix} A_x & A_y & A_z \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}$$

This will become a very useful representation later.

Now it is easy to see that the spatial distance is invariant under the rotation part

$$\vec{r}' = R\vec{r} \rightarrow \Delta\vec{r}' = \vec{r}_2{}' - \vec{r}_1{}' = \tilde{R}(\vec{r}_2 - \vec{r}_1) = \tilde{R}\Delta\vec{r} \qquad (6)$$

$$\Delta S'^2 = \Delta\vec{r}' \cdot \Delta\vec{r}' = \Delta r'^T \Delta r' = \Delta r^T R^T R \Delta r = \Delta r^T \Delta r = \Delta S^2 \qquad (7)$$

For the translation part of the Galilean transformation, the simultaneity is crucial for the invariance since $\vec{r}' = \tilde{R}\vec{r} - \vec{v}t - \vec{b}$

$$\Delta\vec{r}' = \vec{r}_2{}'(t) - \vec{r}_1{}'(t) = (\tilde{R}\vec{r}_2 - \cancel{\vec{v}t} + \cancel{\vec{b}}) - (\tilde{R}\vec{r}_1 - \cancel{\vec{v}t} + \cancel{\vec{b}}) = \tilde{R}(\vec{r}_2 - \vec{r}_1) = \tilde{R}\Delta\vec{r} \qquad (8)$$

Hence,

$$\Delta S'^2 = \Delta r'^T \Delta r' = \Delta r^T R^T R \Delta r = \Delta r^T \Delta r = \Delta S^2 \qquad (9)$$

property 3: If $\vec{r}(t) = \vec{u}t + \vec{r}_0$ then the transformation leads to

$$\vec{r}'(t) = \tilde{R}\vec{r}(t) - \vec{v}t + \vec{b} = \tilde{R}(\vec{u}t + \vec{r}_0) - \vec{v}t + \vec{b} = (\tilde{R}\vec{u} - \vec{v})t + \tilde{R}\vec{r}_0 + \vec{b} \qquad (10)$$

which is a rectilinear motion (with velocity $\tilde{R}\vec{u} - \vec{v}$)

2.2 Lorentz transformation

From the previous section, the velocity after the Galilean transformation ($\tilde{R} = \mathbb{1}$) is simply additive. i.e. $\vec{u}' = \vec{u} - \vec{v}$

However, it has been demonstrated since 1888 that this velocity transformation does not quite work with the speed of light. (1888 was the year of the Michaelson-Morley experiment). Eventually, the resolution came from Einstein (principle of relativity in 1905)

Poincaré and Lorentz transformations

In classical mechanics, we assume that the space (set) of events (t, x, y, z) forms a Galilean space (a space equipped with Galilean transformation). In Special Relativity (SR), the structure for invariance is different. Instead of a separate time interval & spatial distance, there is a single interval defined between pairs of events:

$$\Delta S^2 = -(c\Delta t)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{11}$$

• First shortcut! There will be too many c's to write in the future, and this is just a constant. → We will choose a system of units in such a way that $c = 1$. This is not so surprising as it seems since a constant can be different in different units, for example $c = 2.99 \times 10^8 m/s = 1.08 \times 10^9 km/hr$. However, a peculiar thing about this is that we don't write unit at all for c! → In this new system of units, the length is space & the length in time shares the same unit!

Transformations which leave $\Delta S^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$ invariant are called Poincare' transformation

In order to discuss the general form of Poincaré's transformation. We start by writing the interval ΔS in matrix representation:

$$\Delta S^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 = \Delta r^T \tilde{G} \Delta r \tag{12}$$

where

$$r = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \quad \Delta r = \begin{pmatrix} \Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \quad \text{and} \quad \tilde{G} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{13}$$

It is easy to see that

$$\begin{pmatrix} \Delta t & \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \Delta S^2 \tag{14}$$

Note that the spatial distance in Galilean space can be written in a similar structure $\Delta S^2 = \Delta r^T \Delta r = \Delta r^T \mathbb{1} \Delta r$

The Poincaré transformation takes a general form as

$$r' = \tilde{\Lambda} r + a \tag{15}$$

where $\tilde{\Lambda}$ is 4×4 matrix, a is a constant column vector

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \tag{16}$$

The Subset of Poincaré's transformation with no translation ($a = 0$) is called Lorentz transformation. Then we have

$$\begin{aligned} \Delta r' &= r_2' - r_1' = \tilde{\Lambda} r_2 + a - (\tilde{\Lambda} r_1 + a) \\ &= \tilde{\Lambda}(r_2 - r_1) = \tilde{\Lambda} \Delta r \end{aligned} \tag{17}$$

Therefore, the invariance of the interval ΔS requires $\tilde{\Lambda}$ to have the following porperty:

$$\begin{aligned} &\Delta S'^2 = \Delta r'^T \tilde{G} \Delta r' = \Delta r^T \tilde{\Lambda}^T \tilde{G} \tilde{\Lambda} \Delta r \\ &\Delta S'^2 = \Delta r^T \tilde{G} \Delta r \end{aligned} \tag{18}$$

$$\boxed{\tilde{\Lambda}^T \tilde{G} \tilde{\Lambda} = \tilde{G}} \tag{19}$$

This is the property of the Lorentz transformation.

• The second shortcut! (The Einstein summation convention)
In the component form of vector/matrix operation, we have

$$x' = \tilde{\Lambda} x \to \begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

where $x^0 \equiv t, x^1 \equiv x, x^2 \equiv y, x^3 \equiv z$

Note: The motivation for choosing the position of indices are 1) $x'\&x$ are objects of the same type → both are up 2) The summation must happen with up-down $\Lambda^\nu{}_\mu x^\mu$

Writing 4 equations explicitly as

$$x'^0 = \sum_{\mu=0}^{3} \Lambda^0{}_\mu x^\mu, x'^1 = \sum_{\mu=0}^{3} \Lambda^1{}_\mu x^\mu, x'^2 = \sum_{\mu=0}^{3} \Lambda^2{}_\mu x^\mu, x'^3 = \sum_{\mu=0}^{3} \Lambda^3{}_\mu x^\mu$$

We can do better by writing generically as

$$x' = \tilde{\Lambda} x \to x'^\nu = \sum_{\mu=0}^{3} \Lambda^\nu{}_\mu x^\mu$$

The Einstein's convention is the realisation that every time the $\sum$ appears, there are 2 repeating indices on the same side of the equation
There is no need to write $\sum$ at all!!

$$x' = \tilde{\Lambda} x \to x'^\nu = \Lambda^\nu{}_\mu x^\mu$$

The rule is repeating indices = summation

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

index ν is free index that must appear on both sides of the equation, and index μ is a dummy index that can be changed to anything.

The inverse Lorentz transformation can be written as

$$x' = \Lambda x \to x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{20}$$

$$\Lambda^{-1} x' = \Lambda^{-1} \Lambda x \to (\Lambda^{-1})^\rho{}_\mu x'^\mu = (\Lambda^{-1})^\rho{}_\mu \Lambda^\mu{}_\nu x^\nu \tag{21}$$

using $(\Lambda^{-1})^\rho{}_\mu \Lambda^\mu{}_\nu = (\Lambda^{-1}\Lambda)^\rho{}_\nu = (\mathbb{1})^\rho{}_\nu$

$$\Lambda^{-1}x' = \mathbb{1}x = x \rightarrow (\Lambda^{-1})^\rho{}_\mu x'^\mu = (\mathbb{1})^\rho{}_\nu x^\nu = \delta^\rho{}_\nu x^\nu \tag{22}$$

Since $\mathbb{1}$ is the identity matrix, we define the Kronecker delta:

$$\delta^\mu{}_\nu = \begin{cases} 1, & \text{if } \mu = \nu, \\ 0, & \text{if } \mu \neq \nu. \end{cases} \tag{23}$$

We get

$$(\Lambda^{-1})^\rho{}_\mu x'^\mu = x^\rho \tag{24}$$

Back to our space-time interval. We can also write it in component form with Einstein's convention

$$\Delta S^2 = \Delta x^T \tilde{G} \Delta x \rightarrow \Delta S^2 = \Delta x^\mu (\tilde{G})_{\mu\nu} \Delta x^\nu \tag{25}$$

Now we define $(\tilde{G})_{\mu\nu} \equiv \eta_{\mu\nu}$ as a component form of the metric tensor where

$$\eta_{\mu\nu} = \begin{cases} 0, & \text{if } \mu \neq \nu, \\ -1, & \text{if } \mu = \nu = 0, \\ 1, & \text{if } \mu = \nu > 0 \end{cases} \tag{26}$$

Exercise1: Derive the definition of Lorentz transformation in component form (Ans: $\tilde{G} = \Lambda^T \tilde{G} \Lambda \rightarrow \eta_{\mu\nu} = \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu$)

Exercise2: Check that $\det(\tilde{\Lambda}) = \pm 1$

Exercise3: Check that for any Lorentz transformation, we have either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$ [Hint: work out $\begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix} = \begin{pmatrix} \Lambda^0{}_0 & A^T \\ B & R^T \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \Lambda^0{}_0 & B^T \\ A & R \end{pmatrix}$]

Exercise4 What is the physical meaning of Lorentz transformation with $\Lambda^0{}_0 \geq 1$ and Lorentz transformation with $\Lambda^0{}_0 \leq -1$

[Hint: look at $\Delta t = \begin{pmatrix} \Delta t \\ 0 \\ 0 \\ 0 \end{pmatrix}$]

The set $\mathbb{R}^4$ of spacetime coordinates with this Poincaré invariant interval structure is called Minkowski space(= the set of events)

However, Minkowski space $\neq$ vector space in the linear algebra sense. Since the linear combination of 2 events does not necessarily have the Poincaré invariant property. Let x^μ & y^μ be 2 events, $x^\mu + by^\mu$ does not transform as Poincare' transformation

$$\begin{aligned} x'^\mu + by'^\mu &= \Lambda^\mu{}_\nu x^\nu + a^\mu + b(\Lambda^\mu{}_\nu y^\nu + a^\mu) \\ &= \Lambda^\mu{}_\nu(x^\nu + by^\nu) + \underbrace{(1+b)a^\mu}_{\text{not invariant}} \end{aligned} \tag{27}$$

But the difference between any two events does form a vector space since $x'^\mu - y'^\mu = \Lambda^\mu{}_\nu(x^\nu - y^\nu)$

● Minkowski space is an <u>affine space</u>

2.3 4-vectors

- Objects $x = x^\mu \hat{e}_\mu = x^0\hat{e}_0 + x^1\hat{e}_1 + x^2\hat{e}_2 + x^3\hat{e}_3$ are called 4-vectors
- The metric tensor has the inverse, $\eta^{\mu\nu}$ defined by

$$\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu_\rho \tag{28}$$

- One can immediately check that the inverse has identical elements

$$\eta^{\mu\nu} = \begin{cases} 0, & \text{if } \mu \neq \nu, \\ -1, & \text{if } \mu = \nu = 0, \\ 1, & \text{if } \mu = \nu > 0 \end{cases} \tag{29}$$

- $\eta_{\mu\nu}$ & $\eta^{\mu\nu}$ will be used to " switch" between lower & upper index objects For example, $u_\mu = \eta_{\mu\nu}u^\nu, u^\mu = \eta^{\mu\nu}u_\nu$
- For two 4-vectors $A = A^\mu\hat{e}_\mu$ and $B = B^\mu\hat{e}_\mu$, we can define their inner product to be the scalar

$$\eta(A, B) \equiv A^\mu B_\mu = A^\mu\eta_{\mu\nu}B^\nu = A_\nu B^\nu = A_\nu\eta^{\nu\mu}B_\mu \tag{30}$$

- We say that 2 vectors are orthogonal if $A^\mu B_\mu = 0$
- The magnitude of a 4-vector is defined as $A^\mu A_\mu$
- A non-zero 4-vector A^μ is called
  - null if $A^\mu A_\mu = 0$
  - timelike if $A^\mu A_\mu < 0$
  - spacelike if $A^\mu A_\mu > 0$

- The set of all null 4 vectors is called the null cone
- The light cone at point p in Minkowski is the set of points in M that are connected to p by a null vector

Exercise5: Check that the "boost" with velocity v in the x-direction satisfies $\tilde{G} = \Lambda^T \tilde{G} \Lambda$

$$\Lambda = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{31}$$

where $\gamma = 1/\sqrt{1-v^2}$

This transformation is equivalent to changing the frame of reference to an inertial frame with speed v in the x-direction

Exercise6: Verify that performing two Lorentz transformations with velocities v_1 and v_2 in x-direction successively is equivalent to a single Lorentz transformation with velocity

$$v = \frac{v_1 + v_2}{1 + v_1 v_2} \tag{32}$$

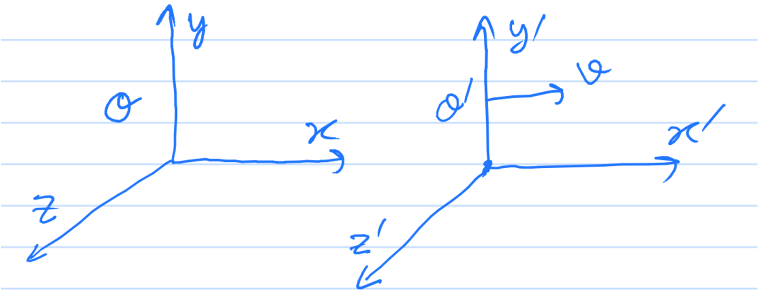
2.4 Relativity

- Consider 2 events $p = (t_1, x_1, y, z)$ and $q = (t_2, x_2, y, z)$ which happens at the same time $\Delta t = t_2 - t_1 = 0$. They are called simultaneous in an inertial frame O.
- However the concept of simultaneity does not make sense in special relativity since a boost in x-direction is enough to break the simultaneity of two events

  $$\Delta x' = \gamma(\Delta x - v\Delta t), \quad \Delta t' = \gamma(\Delta t - v\Delta x) \tag{33}$$

  For $\Delta t = 0$ in O frame, $\Delta t' = -\gamma v \Delta x \neq 0$ if $x_2 \neq x_1$

  → relativity of simultaneity



- Now consider a clock at rest in O' frame make a successive ticks at (t_1', x', y', z') & (t_2', x', y', z'), time difference in O frame (let O' be a moving frame with v with respect to O) is given by the inverse Lorentz transformation

$$\begin{aligned} \Delta t &= \gamma(\Delta t' + v\Delta x') = \gamma \Delta t' \quad (\Delta x' = 0 \quad \text{in} \quad O') \\ \Delta t &= \frac{\Delta t'}{\sqrt{1-v^2}} \geq \Delta t' \quad \text{since} \quad v < 1 (v < c) \end{aligned} \tag{34}$$

  This is an effect known as time dilation ("a moving clock appears to slow down")

- Now consider a rod of length $l = \Delta x$ at rest in O

- How do we measure a " moving " rod in O' frame?

  → We measure 2 simultaneous events at the endpoints (in O'). Consider the inverse Lorentz transformation

$$\begin{aligned} l = \Delta x &= \gamma(\Delta x' + v\Delta t') = \gamma \Delta x' \\ l' = \Delta x' &= \frac{1}{\gamma}\Delta x = \sqrt{1-v^2} l \leq l \end{aligned} \tag{35}$$

  This is the effect known as length contraction (" a rod is contracted when viewed by a moving observer ")

- Let a particle have velocity $\vec{u} = (u_x, u_y, u_z)$ in O frame and $\vec{u}' = (u_x', u_y', u_z')$ in O' frame

- Let set $u_x = \frac{dx}{dt}, u_x' = \frac{dx'}{dt'}, u_y = \frac{dy}{dt}$, etc...

- From the Lorentz transformation, it is easy to derive the relativistic transformation of velocities

$$u_x' = \frac{dx'}{dt'} = \frac{d}{dt'}(\gamma(x - vt)) = \gamma \frac{d}{dt}(x - vt)\frac{dt}{dt'} = \gamma(u_x - v)\frac{dt}{dt'} \tag{36}$$

Consider

$$\frac{dt'}{dt} = \frac{d}{dt}\gamma(t - vx) = \gamma(1 - vu_x) \tag{37}$$

Combining the two equations, we get

$$u'_x = \frac{u_x - v}{1 - vu_x} \tag{38}$$

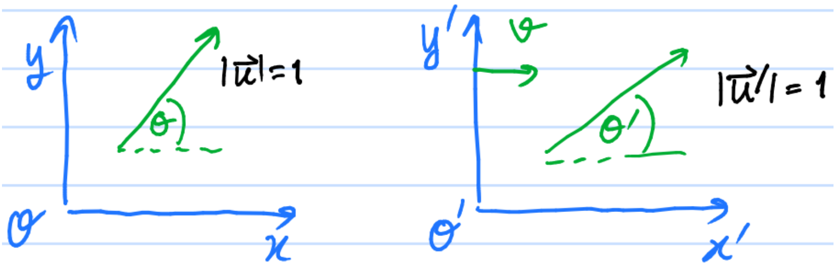
Exercise7: show that

$$u'_y = \frac{u_y}{\gamma(1 - vu_x)}, \quad u'_z = \frac{u_z}{\gamma(1 - vu_x)} \tag{39}$$

Exercise8: Using the inverse Lorentz transformation to show that

$$u_x = \frac{u'_x + v}{1 + vu'_x}, \quad u_y = \frac{u'_y}{\gamma(1 + vu'_x)}, \quad u_z = \frac{u'_z}{\gamma(1 + vu'_x)} \tag{40}$$

Exercise9: Show that if the speed of particle is $|\vec{u}| = 1$ the velocity in any inertial frame is also $|\vec{u}'| = 1$

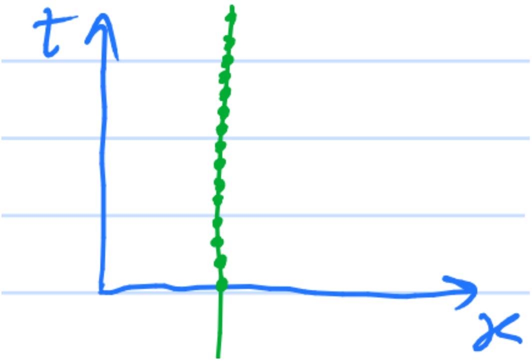


Exercise10: Show that the relativistic formula of light aberration

$$\sin\theta = \frac{\sqrt{1 - v^2}\sin\theta'}{1 + v\cos\theta'} \tag{41}$$

2.5 Particle Dynamics

- In Minkowski space, a series of continuous events can form a "World-line"
- We can parametrise the World-line using a parameter λ
- For example, a particle at rest has a straight line as the World line in space time diagrame



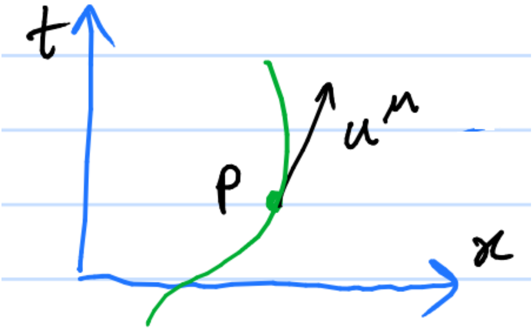
The “natural” parametrisation of this world line is just “t”

- Let’s define the tangent 4-vector to the curve at point p to be the 4-vector u given by

$$u = u^{\mu} e_{\mu} \quad \text{where} \quad u^{\mu} = \frac{dx^{\mu}}{d\lambda}\Big|_{\lambda=\lambda_0} \tag{42}$$

where p is at $x^{\mu}(\lambda_0)$

- Note that this 4-vector is independent of the basis choice $(\hat{e}_{\mu})$
- The direction in the space-time diagram = direction of the world-line



- The world-line of a physical object is assumed to stay inside the light-cone (null cone) of every point in the World-line
- This translates to the requirement on the tangent 4-vector to be time-like:

$$\eta(u(\lambda), u(\lambda)) = \eta_{\mu\nu} u^{\mu}(\lambda) u^{\nu}(\lambda) = \eta_{\mu\nu} \frac{dx^{\mu}}{d\lambda}\frac{dx^{\nu}}{d\lambda} < 0 \tag{43}$$

we have

$$\begin{aligned} 0 &> \eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = -\left(\frac{dt}{d\lambda}\right)^2 + \left(\frac{dx}{d\lambda}\right)^2 + \left(\frac{dy}{d\lambda}\right)^2 + \left(\frac{dz}{d\lambda}\right)^2 \\ 0 &> \left(\frac{dt}{d\lambda}\right)^2 \left[-1 + \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right] \\ 0 &> \left(\frac{dt}{d\lambda}\right)^2 \left(-1+v^2\right) \quad \rightarrow v < 1 \end{aligned} \tag{44}$$

→ the velocity of physical objects is always less than 1

- Now let's talk about the choice of parametrisation λ First consider two neighbouring events on the Wolrd-line $x(\lambda)$ and $x(\lambda + \Delta\lambda)$ where the difference is timelike vector ($\Delta x^\mu = x^\mu(\lambda + \Delta\lambda) - x^\mu$):

$$\Delta S^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu < 0 \tag{45}$$

- There always exists an inertial frame of reference where $x^\mu(\lambda + \Delta\lambda)$ & x^μ are seperated only in time coordinate

  → The object stays at rest in this period $\Delta\lambda$

  → this is called “instantaneous rest frame ”(i.r.f) which can be varied from point to point

  In this frame we can define ΔS as the difference in time in i.r.f (We can do this since ΔS is invariant & $\Delta x = \Delta y = \Delta z = 0$ in i.r.f.) $\rightarrow -\Delta\tau^2 \equiv \Delta S^2 = \eta_{\mu\nu}\Delta x^\mu \Delta x^\nu < 0$

  Now, let's take the limit $\Delta\lambda \rightarrow 0$

$$\begin{gathered} -\Delta\tau^2 = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}(\Delta\lambda)^2 = \left[-\left(\frac{dt}{d\lambda}\right)^2 + \left(\frac{dx}{d\lambda}\right)^2 + \left(\frac{dy}{d\lambda}\right)^2 + \left(\frac{dz}{d\lambda}\right)^2\right]\Delta\lambda^2 \\ \Delta\tau^2 = -(-1+v^2)\left(\frac{dt}{d\lambda}\right)^2\Delta\lambda^2 = (1-v^2)\left(\frac{dt}{d\lambda}\right)^2\Delta\lambda^2 = \frac{1}{\gamma^2}\Delta t^2 \end{gathered} \tag{46}$$

- Therefore, we get the usual time dilation formula $\Delta\tau = \frac{1}{\gamma}\Delta t$ where $\Delta\tau$ is interpreted as the clock carried by the particle
- We can take this infinitesimal piece and integrate along the world-line

$$\tau_{pq} = \int_p^q d\tau = \int_{t_p}^{t_q} \frac{dt}{\gamma} \tag{47}$$

- This is called the proper time from $p \to q$
- If we fix p and let q vary we can use this as a parameter λ

$$\lambda = \tau = \int_{t_p}^{t} \frac{dt}{\gamma} \tag{48}$$

- In this propertime parametrisation, the tangent 4-vector becomes the 4-velocity defined as

$$V^\mu = \frac{dx^\mu}{d\lambda} = \frac{dx^\mu}{d\tau} = \gamma \frac{dx^\mu}{dt} \to V = \gamma \begin{pmatrix} \frac{dt}{dt} \\ \frac{dx}{dt} \\ \frac{dy}{dt} \\ \frac{dz}{dt} \end{pmatrix} = \gamma \begin{pmatrix} 1 \\ \vec{v} \end{pmatrix} \tag{49}$$

- It is quicker to write the vector component in 1 row as follows

$$V^\mu = \gamma(1, \vec{v}) \tag{50}$$

Exercise11: Show that 4-velocity always has magnitude $= -1$

Exercise12: Show that the 4-acceleration can be written as

$$A^\mu = \frac{dV^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2} = \gamma \left(\frac{d\gamma}{dt}, \vec{v}\frac{d\gamma}{dt} + \gamma \frac{d\vec{v}}{dt} \right) \tag{51}$$

Exercise13: Show that 4-velocity is always $\perp$ to 4-acceleration
[Hint: consider $\frac{d}{dt}(V_\mu V^\mu) = \frac{d}{dt}(1) = 0$]

2.6 Relativistic Particle Dynamics

We assume that each particle has a constant, m, attached to it called "rest mass". This can be considered as the inertia mass in an i.r.f. of the particle satisfying $\vec{F} = m\vec{a}$ in that frame

- The 4-momentum of the particle is defined as

$$\begin{aligned} P^\mu &\equiv mV^\mu = m\gamma(1, \vec{v}) \\ &= (m\gamma, \gamma m \vec{v}) \end{aligned} \tag{52}$$

- We will defined $P^\mu = (E, \vec{p})$, where E is energy, $\vec{p} =$ momentum

$$E = m\gamma = \frac{m}{\sqrt{1 - v^2}}, \quad \vec{p} = \gamma m \vec{v} = \frac{m\vec{v}}{\sqrt{1 - v^2}} \tag{53}$$

- We call them energy & momentum because in $v << 1$ limit

$$\begin{aligned}\vec{p} &= \frac{m\vec{v}}{(1-v^2)^{1/2}} = m\vec{v}(1-v^2)^{-1/2} \cong m\vec{v}(1+\frac{1}{2}v^2+O(v^4)) \cong m\vec{v} \\ E &= \frac{m}{(1-v^2)^{1/2}} = m(1-v^2)^{-1/2} \cong m(1+\frac{1}{2}v^2+O(v^4)) \cong m+\frac{1}{2}m\vec{v}^2\end{aligned} \tag{54}$$

  (in the usual unit the last equation reads $E = mc^2 + \frac{1}{2}mv^2$)

  The energy contribution $E = m$, which is the energy of a particle at rest

- The 4-momentum has magnitude m^2 since $P^\mu P_\mu = m^2 V^\mu V_\mu = -m^2$

- Therefore, we have the dispersion relation of a relativistic particle

$$-m^2 = P^\mu P_\mu = \begin{pmatrix} E & \vec{p} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} E \\ \vec{p} \end{pmatrix} = -E^2 + \vec{p}^2 = -m^2 \tag{55}$$

$$\boxed{E^2 = \vec{p}^2 + m^2} \tag{56}$$

- Also

$$\frac{\vec{p}}{E} = \frac{\gamma m\vec{v}}{\gamma m} = \vec{v} \tag{57}$$

- The above relation still holds even in the case $v \to 1$ as long as $m = 0$. This is a photon which satisfy

$$E = |\vec{p}| \quad \& \quad P^\mu = (E, E\hat{n}) \quad \& \quad P^\mu P_\mu = 0 \tag{58}$$

  where $\hat{n}$ is the direction of propagation.

- We can define a 4-force as follows

$$F^\mu = \frac{dP^\mu}{d\tau} = mA^\mu \tag{59}$$

  which is always orthogonal to V^μ

- Define the 3-force in the usual way as $\vec{f} = \frac{d\vec{p}}{dt}$

- We have

$$\begin{aligned}F^\mu &= \frac{dP^\mu}{d\tau} = \frac{d}{dt}(E, \vec{p}) \cdot \frac{dt}{d\tau} \\ F^\mu &= \gamma\left(\frac{dE}{dt}, \vec{f}\right)\end{aligned} \tag{60}$$

Exercise14: Show that the definition of power $\vec{f} \cdot \vec{v} = \frac{dE}{dt}$ can be derived from $F^\mu V_\mu = 0$.

3 Equivalence Principle

The equivalence principle (EP) is one of the crucial ideas leading to the argument why we think the spacetime is curved, as well as how to generalize the idea from SR to GR. In this lecture note, we will classify EP into two parts: Weak Equivalence Principle (WEP) and Strong Equivalence Principle (SEP).

What is the WEP?
For the WEP, it states that "No experiment in mechanics can distinguish between a gravitational field and an accelerating frame of reference." This idea has a root from Galileo's demonstration shown in Fig. 8. This implies that "All bodies fall at the same rate in a gravitational field." Now let us do the Galileo in a rest box influenced by a gravitational field as illustrated in the left panel of Fig. 9. Then changing the situation to a box with an acceleration in the same magnitude of the gravitational field, as shown in the right panel of Fig. 9. By means of WEP, the results of the experiment are the same. We cannot distinguish between the two situations by using mechanical experiments.

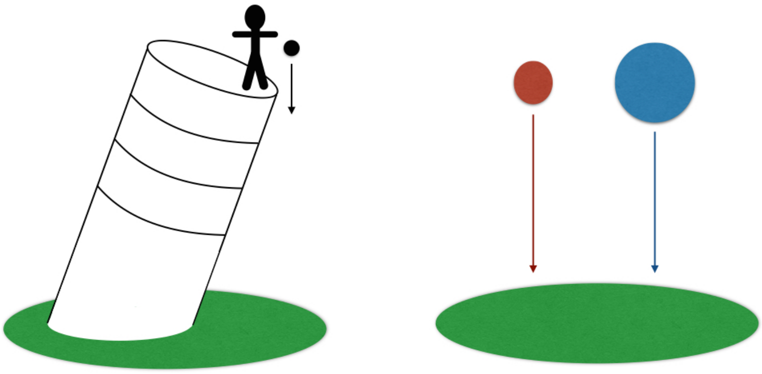


Figure 8: Objects with different masses are dropped to the ground at the same time.

What is the consequence of WEP?
According to Newton's second law, the particle of mass m_i can be accelerated by the applied force as

$$\vec{F} = m_i\vec{a}.$$

Note that the mass m_i denotes the inertial mass of a particle. This kind of mass plays the role to resist of the moment of the particle. The greater the

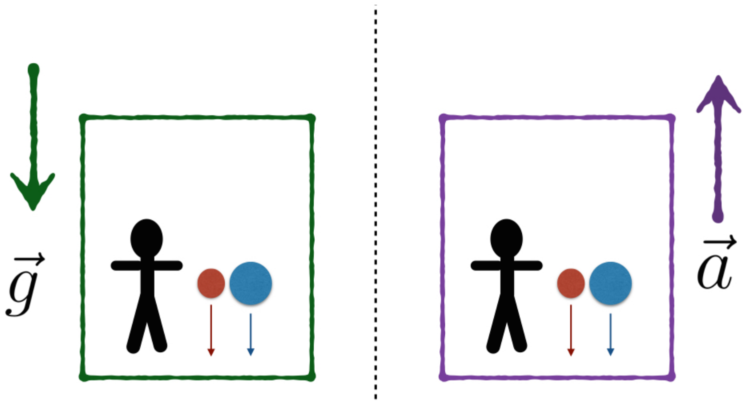


Figure 9: Observers do the experiment in the boxes, then he obtains the same result.

value of m_i, the more force to make its movement.

There is another kind of mass called "gravitational mass". It measures the response of the particle to the gravitational field,

$$\vec{F} = m_g\vec{E}_g = -m_g\vec{\nabla}\Phi.$$

It plays the similar role to electric charge in terms of electric force, $\vec{F} = q\vec{E}$. Therefore, m_g is essentially gravitational charge. These two kinds of mass are completely different in this sense. By using the argument from the WEP, we then have

$$\vec{F} = m_g\vec{E}_g = m_i\vec{a}, \quad \Rightarrow \quad m_i = m_g.$$

Sometimes, the WEP is known such that "The inertial mass and gravitational mass of any object are equal".

It is important to note that EP is a local principle. To see this property, let us consider the Einstein box, but now it is much bigger as seen in Fig. 10. From this figure, the masses are falling to the center of the Earth, then (for the box large enough) the masses become closer as they fall. In this case, we can distinguish between $\vec{E}_g$ and $\vec{a}$ so that the principle is not valid. Actually, the principle is valid for a small enough box. Note that most gravitational fields are not uniform, so the EP treats a gravitational field at a single point, which is equivalent to the uniform acceleration.

Strong Equivalence Principle

The WEP, it is only for the experiments associated with mechanics. It may

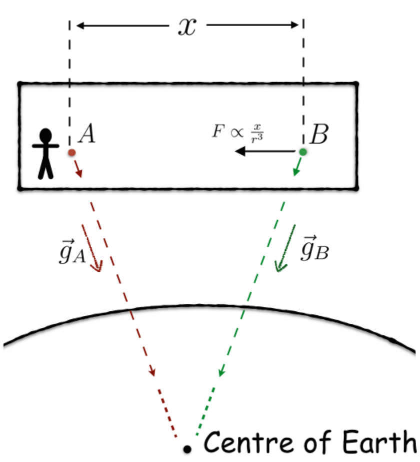


Figure 10: Two masses become closer while they are falling in the large enough box in a gravitational field.

be asked if it is possible to use this principle in other areas of physics? In this consideration, we will discuss the experiments in optics. To see this, let us consider the Einstein box with the laser attached on the top and the receiver at the bottom as illustrated in Fig. 11. By considering two situations, with and without acceleration. From this experiment, one can measure the travel time for both cases. By comparing these, we can find the difference in the

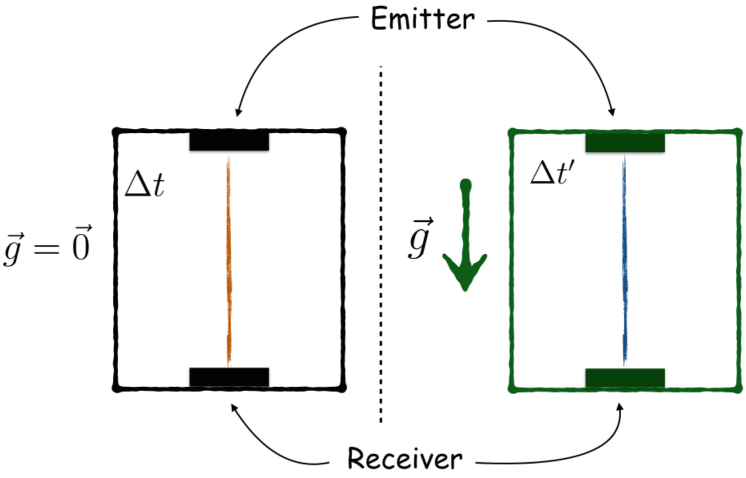


Figure 11: The laser and receiver in the Einstein box with/without a gravitational field.

light wavelength since the speed of light is constant, as

$$\frac{\Delta\lambda}{\lambda} = \frac{al}{c^2},$$

where l is the height of the box. This actually is due to the Doppler shift. One can perform the experiment in the same way, but now consider the box at rest in a gravitational field. It is found that the light wavelength is also shifted in the same way, but now it is not the Doppler effect since the box is at rest. We called this effect the gravitational redshift. From this, one can summarize that ”No experiment in optics can distinguish between a gravitational field and an accelerating frame of reference”. Moreover, we can generalize the idea to all areas of physics, and this is the statement of the SEP: ”All the laws of nature are affected in the same way by a gravitational field and a constant acceleration.”

More thought-experiment in optics

Set up the experiment in a similar way, but now let us attach the emitter and receiver in the horizontal part as shown in Fig. 12. Using EP and this thought experiment, one can see that the light travels in a curved path in a gravitational field. Note that this conclusion is also obtained from the variational principle: the light follows the path with using minimum time. In any Euclidean space, the shortest path is a straight line, but the curved path in a gravitational field. This leads Einstein to obtain the idea that the gravity makes a curved spacetime, and then the light moving in such a curved spacetime making the curved path.

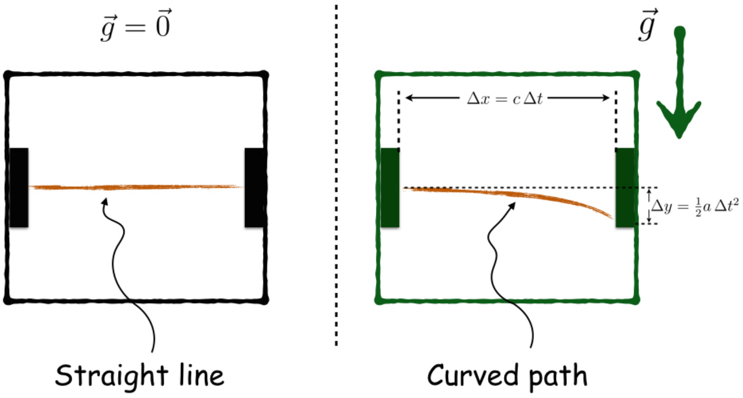


Figure 12: The laser and receiver attaching in the horizontal part of the Einstein box with/without a gravitational field.

In order to study the gravity, one has to study the curved spacetime. This leads to learning a vector of tensors in curved spacetime and also how to differentiate them. This subject is formally known as ”differential geometry”.

4 Spacetime curvature

4.1 Intrinsic and Extrinsic curvature

Physically, the curvature can be classified into two types: Intrinsic curvature and extrinsic curvature.

☐ Intrinsic curvature: a curvature measured by one in the surface itself does not need the information from the higher dimension.

☐ Extrinsic curvature: a curvature measured by one who needs the information from the higher dimension.

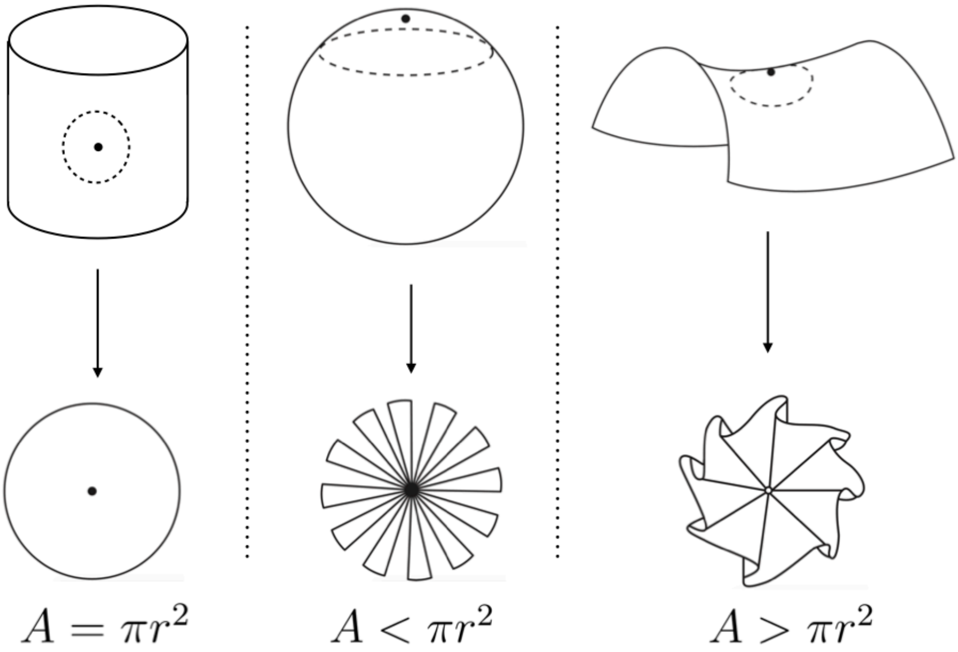


Figure 13: Show how cylinder, sphere, and saddle differ in terms of curvature.

Example

1. Cylinder (see the left figure of Fig. 13): In 2 dimensions, it is flat. However, its shape is curved in 3 dimensions. We need the information from 3 dimensions in order to identify the curve of the cylinder ⇒ Extrinsic curvature. If we unfold the cylinder into 2D, it will look like flat paper. In

other words, if we write down the circle on the cylinder, it can be put in the flat completely, or the area is still the same $A = \pi r^2$ as illustrated in the left panel of Fig. 13. In this case, the cylinder is intrinsically flat.

2. Sphere (see the middle figure of Fig13): $\Rightarrow$ From the middle panel, one can see that the circle on the sphere cannot be put in the plane completely, or the area is less than the usual one $A < \pi r^2$. This object is intrinsically curved.

3. Saddle (see the right figure of Fig13): $\Rightarrow$ From the right panel, one can see that the circle on the saddle cannot be put in the flat completely, or the area is greater than the usual one $A > \pi r^2$. This object is also intrinsically curved.

Let us consider how to define curvature from a mathematical point of view. In order to define the curve properly, one can take a vector in close path. If the vector becomes the same at the same point, the surface we take such the vector is called flat. If the vector is not the same, the surface is curved.

☐ Flat: From Fig. 14, one can see that after we take the vector to the closed path, the vector will be the same. The surface is called flat in this case.

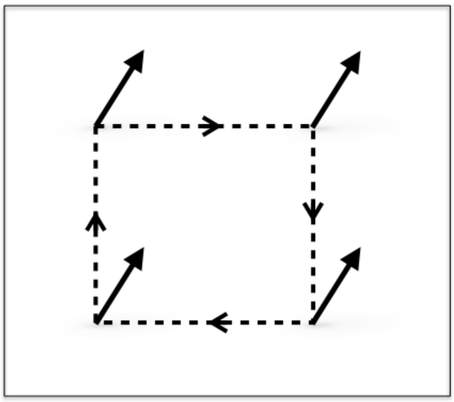


Figure 14: Taking a round trip of the vector on a flat surface, and then we will find that the vector becomes the same at the same point.

☐ Sphere (example of curve): From Fig. 15, We move the vector along $A \to B \to C \to D \to A$. It is found that we do not obtain the same vector $\vec{V}_i = \vec{V}_f \neq \vec{V}_i$. $\to$. It is the curved surface.

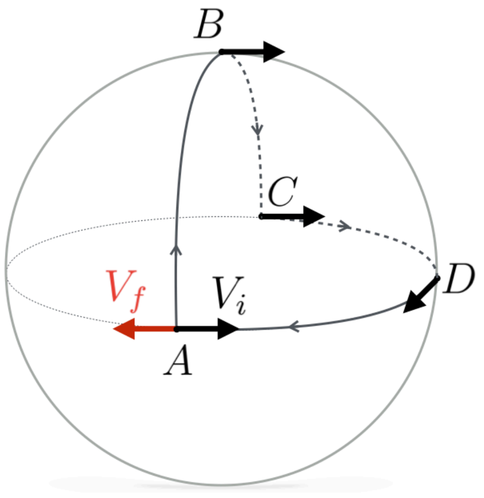


Figure 15: Taking a round trip of the vector on a non-flat surface, and then we will find that the vector is not the same.

4.2 Vector in curved space

From the previous section, in order to properly deal with spacetime curvature, we have to define the vector on the surface. However, we need to first define the mathematical object referred to the curved surface. Mathematically, such an object is called a manifold. Conceptually, the manifold is a mathematical object that is smooth and locally flat. The locally flat property is compatible with the notion obtained in EP. This allows us to connect the vector defined in flat space to a curved manifold. For rigorous speaking, the definition of manifold is any set that can be continuously parameterized (We would not consider more details about this definition).

Now, let us move to consider the vector. Recall the 3-vector, it can be written in terms of basis and components as follows

$$\begin{aligned} \vec{V} &= V_x\hat{i} + V_y\hat{j} + V_k\hat{k}, & (61)\\ &= V_r\vec{e}_r + V_r\theta\vec{e}_\theta + V_\phi\vec{e}_\phi, & (62)\\ &= V^i\vec{e}_i. & (63) \end{aligned}$$

For four-dimensional spacetime (in SR), it can be promoted to a vector in an n-dimensional manifold as

$$V = V^\mu\vec{e}_\mu, \tag{64}$$

where μ runs over $0, 1, 2, 3$.

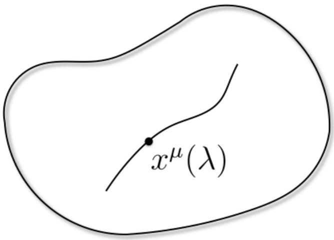


Figure 16: A curve on the manifold $\mathbb{M}$ parametrized by λ with assignment of coordinates $x^\mu(\lambda)$ at a particular point

☐ For a manifold $\mathbb{M}$, the curve on the manifold can be parametrized by a parameter λ.
☐ A point on the curve can be assigned by the local coordinates $x^\mu(\lambda)$ as shown in Fig. 16.
☐ Supposed $f(x^\mu)$ is a function on manifold (Mapping from $\mathbb{M}$ to $\mathbb{R}$), then we can write

$$f\left(x^\mu(\lambda)\right) = g(\lambda) \in \mathbb{R}. \tag{65}$$

The differential of the function along the curve can be written as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} g(\lambda) = \frac{\partial f}{\partial x^\mu}\frac{\partial x^\mu}{\partial \lambda}. \tag{66}$$

As a result, one can write the object in a similar form of a vector in flat space as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{\partial x^\mu}{\partial \lambda}\frac{\partial}{\partial x^\mu} = \frac{\partial x^\mu}{\partial \lambda}\partial_\mu, \tag{67}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \text{ is a vector,} \qquad \frac{\partial x^\mu}{\partial \lambda} \text{ is a component of vector,} \qquad \partial_\mu \text{ is a basis of vector.} \tag{68}$$

Note that since $\frac{\mathrm{d}}{\mathrm{d}\lambda}$ varies with point in manifold, it plays the role of a vector field. Note also that there are other non-coordinate basis but we do not consider them in this lecture. In this lecture, we restrict our attention to the coordinate basis.

☐ **Tangent space**

A vector defined at point p lies in the tangent space, denoted by T_p. At point p, there are actually many curves, and then the plane tangent to the surface at this point is visualized as T_p as shown in Fig 17. In other words,

T_p is obtained by taking all possible curves passing through point p. Note that vectors defined at two different points (different T_p) have no relation to each other. This is a crucial property of the vector on a manifold which significantly differs from that in flat space.

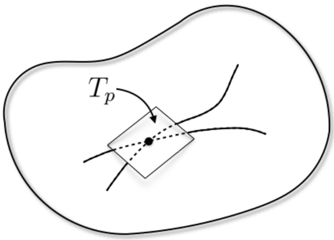


Figure 17: Tangent space

☐ **General coordinate transformation**

In SR, the vector is constructed under the Lorentz transformation. It is valid only in flat spacetime (Minkowski spacetime). This is just a local frame of the curved manifold. For GR, we need to generalize the transformation to cover the whole manifold. In this sense, we introduce the general coordinate transformation (GCT) as

$$x^\mu \to x'^\mu(x). \tag{69}$$

Now let us consider how the component of the vector changes under GCT

$$\begin{aligned}\partial_\mu f(x) &= \partial_\mu f(x'(x)),\\ &= \frac{\partial f}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\mu},\\ &= \frac{\partial x'^\nu}{\partial x^\mu}\partial'_\nu f.\end{aligned} \tag{70}$$

Then, we obtain the transformation of the basis as follows

$$\partial_\mu = \frac{\partial x'^\nu}{\partial x^\mu}\partial'_\nu, \qquad \text{or} \qquad \vec{e}_\mu = \frac{\partial x'^\nu}{\partial x^\mu}\vec{e}\,'_\nu, \tag{71}$$

where $\frac{\partial x'^\nu}{\partial x^\mu}$ is GCT matrix. For the case of the matrix is invertible, we can write

$$\partial'_\nu = \frac{\partial x^\mu}{\partial x'^\nu}\partial_\mu, \qquad \text{or} \qquad \vec{e}\,'_\nu = \frac{\partial x^\mu}{\partial x'^\nu}\vec{e}_\mu. \tag{72}$$

with the property

$$\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial x'^\rho}{\partial x^\nu} = \delta^\mu_\nu \tag{73}$$

As a result, the vector, V under GCT can be rewritten as

$$\begin{aligned}\vec{V} &= V^{\mu}\vec{e}_{\mu}, \\ &= V^{\mu}\left(\frac{\partial x'^{\nu}}{\partial x^{\mu}}\vec{e}'_{\nu}\right),\end{aligned}$$

In order to obtain that the vector is unchanged ($\vec{V} = V^{\mu}\vec{e}_{\mu} = V'^{\nu}\vec{e}'_{\nu} = \vec{V}'$), the component must be transformed as

$$V'^{\nu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}}V^{\mu}. \tag{74}$$

Note that the Lorentz transformation is a special case of GCT.

$$\Lambda^{\mu}{}_{\nu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}}. \tag{75}$$

□ Dual vector

There exist other objects on a manifold called "dual vector". To visualize this kind of vector, let us consider the infinitesimal transformation of coordinates, $x \to x + \mathrm{d}x$. For any function $f = f(x)$ on manifold, it is changed as

$$\mathrm{d}f = \frac{\partial f}{\partial x^{\mu}}\mathrm{d}x^{\mu} \tag{76}$$

By comparing to the usual form of vector, we have

$$\begin{aligned} df &\quad \text{is a dual vector,} \\ \frac{\partial f}{\partial x^{\mu}} &\quad \text{is a component of dual vector,} \\ dx^{\mu} &\quad \text{is a basis of dual vector.}\end{aligned}$$

Mathematically, the dual vector is a map of the vector to $\mathbb{R}$. In the same fashion as vector, $\vec{V}$ defined in T_p, the dual vector $\vec{w}$ is defined in T_p^* (called dual tangent space or cotangent space). By choosing f to be the new coordinate, x'^{μ}, we obtain the transformation rule of the coordinate of the dual vector as

$$\mathrm{d}x'^{\nu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}}\mathrm{d}x^{\mu}. \tag{77}$$

Similar to a vector, the component of dual vector should be transform

$$w'_{\nu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}}w_{\mu}, \tag{78}$$

in order to obtain the invariant dual vector under GCT, $w = w_\mu \mathrm{d}x^\mu = w'_\mu \mathrm{d}x'^\mu = w'$.

□ **Tensor**

A tensor is an object that combines the description of both a dual vector and a vector. More precisely, an object that has components such that the Cartesian product of "r" basis vector and "s" basis of dual vector. Therefore, rank (r, s) tensor can be defined

$$T = T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\,\partial_{\mu_1}\otimes\cdots\otimes\partial_{\mu_r}\otimes dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_s}, \tag{79}$$

The transformation rule can be written as

$$T'^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} = \frac{\partial x'^\mu_1}{\partial x^\rho_1}\cdots\frac{\partial x'^\mu_r}{\partial x^\rho_r}\frac{\partial x^\sigma_1}{\partial x'^\nu_1}\cdots\frac{\partial x^\sigma_s}{\partial x'^\nu_s}T^{\rho_1\cdots\rho_r}{}_{\sigma_1\cdots\sigma_s} \tag{80}$$

<u>Notation Convention</u>

$$T^{\mu_1\cdots\mu_{r-1}\rho}{}_{\nu_1\cdots\nu_{s-1}\rho} \quad : \quad (r-1, s-1)\text{ tensor}, \tag{81}$$

$$T^{(\mu_1\cdots\mu_r)}{}_{\nu_1\cdots\nu_s} = \frac{1}{r!}(\text{Sum over all permutation}), \tag{82}$$

$$T^{\mu_1\cdots\mu_r}{}_{[\nu_1\cdots\nu_s]} = \frac{1}{s!}(\text{Alternative sum over all permutation}). \tag{83}$$

e.g.

$$T^{(\rho\sigma)}{}_\mu = \frac{1}{2}\left(T^{\rho\sigma}{}_\mu + T^{\sigma\rho}{}_\mu\right), \tag{84}$$

$$T^\rho{}_{[\mu\nu]} = \frac{1}{2}\left(T^\rho{}_{\mu\nu} - T^\rho{}_{\nu\mu}\right). \tag{85}$$

It is important to note that, if there exists the metric tensor $g_{\mu\nu}$ on the manifold, it is called a Riemannian manifold. From SR, the spacetime is a kind of rigid body, and the metric tensor $\eta_{\mu\nu}$ is somehow fixed. However, as we have mentioned, the spacetime is a kind of flexible object in GR viewpoint. Physically, the metric tensor is the object that characterizes the flexibility of spacetime. Therefore, the curvature of the spacetime will be related to the property of the metric tensor. We will see later on this issue.

Note that we will use $g_{\mu\nu}$ for curved spacetime, while the flat Minkowski spacetime is dented by $\eta_{\mu\nu}$. As a result, the transformation for the metric tensor can be generalized as

$$\eta'_{\mu\nu} = \Lambda^\rho_\mu \Lambda^\sigma_\nu \eta_{\rho\sigma} \Rightarrow g'_{\mu\nu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}g_{\rho\sigma}. \tag{86}$$

The other properties such as dot product of vectors, raise or lower indecies still be the same.

4.3 Covariant derivatives

In order to find the proper derivatives on curved spacetime, let us consider the transformation of the derivative of V^μ,

$$\begin{aligned}
\partial'_\mu V'^\nu &= \frac{\partial}{\partial x'^\mu} V'^\nu, \\
&= \left(\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial}{\partial x^\rho}\right)\left(\frac{\partial x'^\nu}{\partial x^\sigma} V^\sigma\right), \\
&= \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\sigma}\frac{\partial}{\partial x^\rho} V^\sigma + \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\rho \partial x^\sigma} V^\sigma, \\
&= \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\sigma}\partial_\rho V^\sigma + \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\rho \partial x^\sigma} V^\sigma. \qquad (87)
\end{aligned}$$

We see that $\partial_\mu V^\nu$ does not transform like a (1,1) tensor (because there exists the exceed term, $\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial^2 x'^\nu}{\partial x^\rho \partial x^\sigma} V^\sigma$ in the above equation). This is due to the differentiation of a vector in this manner, defined by comparing two vectors in different tangent spaces. In this sense, the normal derivative can be performed only in the flat spacetime of comparing $V^i(x^j)$ and $V^i(x^j + dx^j) = V^i + dV^i$ (see Fig. 18) as follows

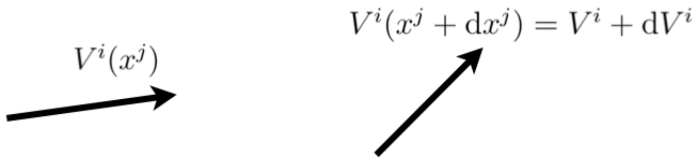


Figure 18: In 3D, two vectors in different points can be compared only in flat space

$$\frac{\partial V^i}{\partial x^j} = \lim_{\mathrm{d}x^j \to 0} \frac{V^i(x^j + \mathrm{d}x^j) - V^i(x^j)}{\mathrm{d}x^j}. \qquad (88)$$

According to the curved spacetime, one has to compare two vectors at the same point (same tangent space). Therefore, we have to move the vector to the same point. Let denote the vector resulting from this transport as $V^\mu(x^\nu) + \delta V^\mu$ as shown in Fig. 19.

Thus, the genuine derivative should be constructed from the difference between $V^\mu(x^\nu) + \delta V^\mu$ and $V^\mu(x^\nu) + \mathrm{d}V^\mu$ as

$$\mathrm{D}V^\mu = (V^\mu(x^\nu) + \mathrm{d}V^\mu) - (V^\mu(x^\nu) + \delta V^\mu) = \mathrm{d}V^\mu - \delta V^\mu. \qquad (89)$$

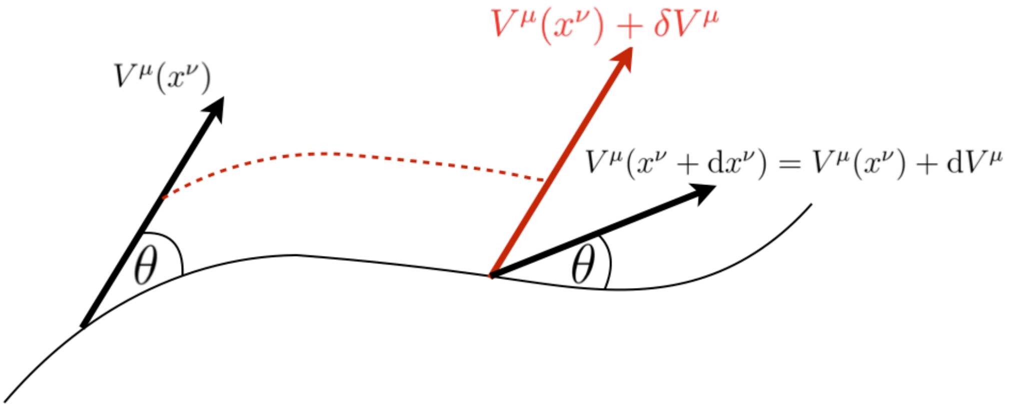


Figure 19: Two vectors in different points can be compared in curved spacetime, if we move one of them to the same point.

Since δV^μ is obtained by parallel transport of vector V^ρ, it should depend on the vector V^ρ itself. Moreover, it is in the same tangent vector of $\mathrm{d}V^\sigma$ so that it has to be proportional to $\mathrm{d}x^\sigma$. As a result, we have,

$$\delta V^\mu \propto V^\rho \mathrm{d}x^\sigma \qquad \rightarrow \qquad \delta V^\mu = -\Gamma^\mu_{\rho\sigma} V^\rho \mathrm{d}x^\sigma, \tag{90}$$

where, $\Gamma^\mu_{\rho\sigma}$ is the connection coefficient. Considering derivative (89) along the curve parametrized by a parameter λ, it reads

$$\begin{aligned} \frac{\mathrm{D}V^\mu}{\mathrm{d}\lambda} &= \frac{\mathrm{d}V^\mu}{\mathrm{d}\lambda} - \frac{\delta V^\mu}{\mathrm{d}\lambda}, \\ &= \frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}\frac{\partial V^\mu}{\partial x^\sigma} + \Gamma^\mu_{\rho\sigma} V^\rho \frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}, \\ \frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}\nabla_\sigma V^\mu &= \left(\frac{\partial V^\mu}{\partial x^\sigma} + \Gamma^\mu_{\rho\sigma} V^\rho\right)\frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}. \end{aligned} \tag{91}$$

Then the component of the covariant derivative of vector V^μ can be written as

$$\nabla_\sigma V^\mu = \partial_\sigma V^\mu + \Gamma^\mu_{\rho\sigma} V^\rho. \tag{92}$$

For the dual vector, one can find the covariant derivative from the fact that a scalar quantity is covariant and is not affected by the transport,

$$\nabla_\sigma (V^\mu W_\mu) = \partial_\sigma (V^\mu W_\mu). \tag{93}$$

<u>Exercise</u> show that the component of the covariant derivative of the dual

vector can be written as

$$\nabla_\sigma W_\mu = \partial_\sigma W_\mu - \Gamma^\rho_{\sigma\mu} W_\rho. \tag{94}$$

Generally, one can find the covariant derivative of the tensor as

$$\begin{aligned}\nabla_\rho T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} &= \partial_\rho T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} + \sum_{i=1}^{p} \Gamma^{\mu_i}_{\rho\sigma} T^{\mu_1\cdots\mu_{i-1}\sigma\mu_{i+1}\cdots\mu_p}{}_{\nu_1\cdots\nu_q} \\ &\quad - \sum_{i=1}^{q} \Gamma^{\sigma}_{\rho\nu_i} T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_{i-1}\sigma\nu_{i+1}\cdots\nu_q}. \end{aligned} \tag{95}$$

<u>Exercise</u> Show that the connection, $\Gamma^\rho_{\mu\nu}$ transforms as the transformation rule

$$\Gamma'^\rho_{\mu\nu} = \frac{\partial x'^\rho}{\partial x^\sigma}\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\Gamma^\sigma_{\alpha\beta} + \frac{\partial x'^\rho}{\partial x^\sigma}\frac{\partial^2 x^\sigma}{\partial x'^\mu \partial x'^\nu}. \tag{96}$$

One found that $\Gamma^\rho_{\mu\nu}$ does not transform as a tensor. However, $\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu}$ transforms as tensor. Therefore, it is possible the define the torsion tensor as

$$T^\rho_{\mu\nu} = \Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\nu\mu} \tag{97}$$

Note that the torsion is a result of the round-trip transport of the scalar function $T^\rho_{\mu\nu} \propto [\nabla_\mu, \nabla_\nu] f$. In GR, we consider only the torsion-free spacetime $T^\rho_{\mu\nu} = 0$. In principle, $\Gamma^\rho_{\mu\nu}$ does not depend on the metric tensor $g_{\mu\nu}$. It is a structure we introduce to the manifold, like $g_{\mu\nu}$. However, if we impose the metric compatibility $\nabla_\rho g_{\mu\nu} = 0$, it will depend on the metric tensor can called the Christoffel symbol.

<u>Exercise</u> By using the metric compatibility $\nabla_\rho g_{\mu\nu} = 0$, show that

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right). \tag{98}$$

<u>Exercise</u> Show that

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} V^\mu\right). \tag{99}$$

4.4 Parallel transport and geodesic equation

Considering directional derivative along a vector $\vec{U}$, it reads

$$\left.\frac{\mathrm{D}V^\nu}{d\lambda}\right|_U = U^\mu \nabla_\mu V^\nu. \tag{100}$$

Now let us specify the direction of the derivatives along the tangent vector ($U \to t$),

$$\left.\frac{\mathrm{D}\ V^\nu}{d\lambda}\right|_t = t^\mu \nabla_\mu V^\nu, \tag{101}$$

$$= \frac{dx^\mu}{d\lambda}\nabla_\mu V^\nu. \tag{102}$$

From this equation, one can see that it implies the transport of the vector along a curve parametrized by λ. If we want to keep the vector constant along the path, the covariant derivative of the vector should not be changed. In this sense, one can define the parallel transport as

$$\left.\frac{\mathrm{D}V^\nu}{d\lambda}\right|_t = t^\mu \nabla_\mu V^\nu = 0. \tag{103}$$

The properties of parallel transport can be listed as

$$\to \qquad \frac{\mathrm{D}}{\mathrm{d}\lambda}T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} = t^\mu \nabla_\mu T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} = 0. \tag{104}$$

$$\to \qquad \frac{\mathrm{D}}{\mathrm{d}\lambda}(g_{\mu\nu}V^\mu W^\nu) = 0. \tag{105}$$

$$\left(\text{Check}: \cancelto{0}{\frac{\mathrm{D}}{\mathrm{d}\lambda}g_{\mu\nu}V^\mu W^\nu} + g_{\mu\nu}\cancelto{0}{\frac{\mathrm{D}V^\mu}{\mathrm{d}\lambda}}W^\nu + g_{\mu\nu}V^\mu\cancelto{0}{\frac{\mathrm{D}W^\nu}{\mathrm{d}\lambda}} = 0\right).$$

→ If $C(\lambda)$ is an arbitrary curve, the tangent vectors are not parallel-transported into the tangent vectors. (106)

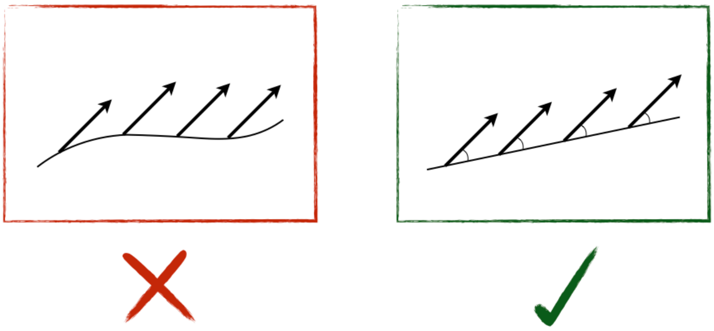


Figure 20: There exists a special subset of all arbitrary curves to satisfy the tangent vectors are parallel-transported into the tangent vectors

There exists a special subset of all arbitrary curves that satisfy the tangent vectors are parallel-transported into the tangent vectors. This curve is called

”geodesic path” or auto-parallel curve as shown in Fig 20. From the condition for parallel transport,

$$\begin{aligned} t^\mu \nabla_\mu t^\rho &= 0, \\ \left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\right)\left(\partial_\mu t^\rho + \Gamma^\rho_{\mu\nu} t_\nu\right) &= 0, \\ \frac{\mathrm{d}t^\rho}{\mathrm{d}\lambda} + \Gamma^\rho_{\mu\nu} t_\mu t_\nu &= 0, \\ \frac{\mathrm{d}^2 x^\rho}{\mathrm{d}\lambda^2} + \Gamma^\rho_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} &= 0. \end{aligned} \tag{107}$$

This is called the geodesic equation. The parameter, λ, satisfying the geodesic equation is called the affine parameter. In the other word,

Parallel transport of V^μ along a geodesic path.
$\Downarrow$
Moving vector by fixing direction and magnitude.

It is important to note that the geodesic equation can be derived from the variational principle. Generally, it can be written as

$$\ddot{x}^\rho + \Gamma^\rho_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = f(\lambda)\dot{x}^\rho, \tag{108}$$

From this equation, it follows that the geodesic equation can be obtained from the notion of the parallel transport by fixing only the direction of the vector V^μ, (not fixing the magnitude). However, this is still equivalent since we can find another proper affine parameter λ.
Exercise Show that it is possible to find the other affine parameter to obtain the usual geodesic equation.

4.5 Curvature tensor

As we have mentioned before, in order to find the intrinsic curvature, we have to move the vector in a closed path. In this sense, it is natural to use the parallel transport to move the vector. Therefore, in this subsection, we will find the spacetime curvature by performing the parallel transport of a vector along the closed path as shown in Fig. 21.

It is found that, in general, the change of vector is

$$\delta V^\rho = -\Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu. \tag{109}$$

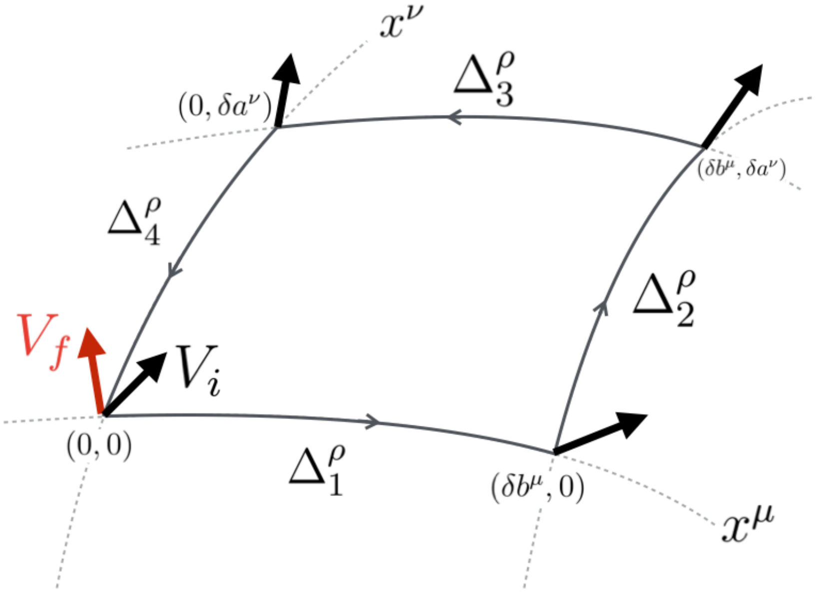


Figure 21: Round trip parallel transport of a vector

According to the above figure, we have

$$\Delta_1^\rho = V_i^\rho(B) - V_i^\rho(A) = -\int_{x^\nu=0} \Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu, \tag{110}$$

$$\Delta_2^\rho = V_i^\rho(C) - V_i^\rho(B) = -\int_{x^\nu=\delta b^\nu} \Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu, \tag{111}$$

$$\Delta_3^\rho = V_i^\rho(D) - V_i^\rho(C) = \int_{x^\mu=\delta a^\mu} \Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu, \tag{112}$$

$$\Delta_4^\rho = V_i^\rho(A) - V_i^\rho(D) = \int_{x^\mu=0} \Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu. \tag{113}$$

Then

$$\begin{aligned}
\Delta_1^\rho + \Delta_3^\rho &= \left(\int_{x^\nu=\delta a^\nu} - \int_{x^\nu=0}\right) \Gamma^\rho_{\mu\sigma} V^\sigma \mathrm{d}x^\mu, \\
&\approx \delta a^\nu \int \partial_\nu \left(\Gamma^\rho_{\mu\sigma} V^\sigma\right) \mathrm{d}x^\mu, \\
&\approx \delta a^\nu \delta b^\mu \partial_\nu \left(\Gamma^\rho_{\mu\sigma} V^\sigma\right). \qquad (114)
\end{aligned}$$

and

$$\begin{aligned}\Delta_2^\rho + \Delta_4^\rho &= \left(-\int_{x^\mu=\delta b^\mu} + \int_{x^\mu=0}\right)\Gamma^\rho_{\mu\sigma}V^\sigma \mathrm{d}x^\mu,\\ &\approx -\delta b^\mu \int \partial_\mu\left(\Gamma^\rho_{\nu\sigma}V^\sigma\right)\mathrm{d}x^\nu,\\ &\approx -\delta b^\mu \delta a^\nu \partial_\mu\left(\Gamma^\rho_{\nu\sigma}V^\sigma\right). \end{aligned} \tag{115}$$

The whole parallel transportation can be written as

$$\begin{aligned}\Delta^\rho &= \sum_i \delta_i^\rho,\\ &\approx \delta a^\nu \delta b^\mu \left[\left(\partial_\nu \Gamma^\rho_{\mu\sigma} - \partial_\mu \Gamma^\rho_{\nu\sigma}\right)V^\sigma + \Gamma^\rho_{\nu\sigma}\partial_\mu V^\sigma - \Gamma^\rho_{\mu\sigma}\partial_\nu V^\sigma\right],\\ &= \delta a^\nu \delta b^\mu \left[\left(\partial_\nu \Gamma^\rho_{\mu\sigma} - \partial_\mu \Gamma^\rho_{\nu\sigma}\right)V^\sigma + \Gamma^\rho_{\nu\sigma}\Gamma^\sigma_{\mu\lambda}V^\lambda - \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\lambda}V^\lambda\right],\\ &= \delta a^\nu \delta b^\mu \left(\partial_\nu \Gamma^\rho_{\mu\sigma} - \partial_\mu \Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma}\right)V^\sigma,\\ &= \delta a^\nu \delta b^\mu R^\rho{}_{\sigma\nu\mu}V^\sigma. \end{aligned} \tag{116}$$

where the tensor $R^\rho{}_{\sigma\nu\mu} \equiv \partial_\nu \Gamma^\rho_{\mu\sigma} - \partial_\mu \Gamma^\rho_{\nu\sigma} + \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma}$ called Riemannian tensor which describes how the manifold curves.

Next, we will discuss the alternative approach to defining the Riemannian tensor. Let us consider the parallel transportation of a vector along different 2 paths in which the starting and ending points are the same, as seen in Fig. 22.

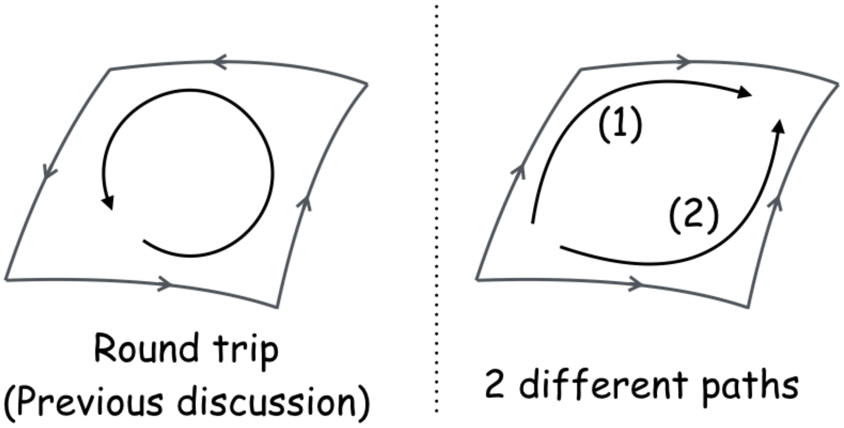


Figure 22: Other way to obtain the curvature tensor

The difference of parallel transported vectors along path 1 and path 2 can be written as

$$\begin{aligned}
\left(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu\right) V^\rho &= \partial_\mu \left(\nabla_\nu V^\rho\right) - \Gamma^\sigma_{\mu\nu} \nabla_\sigma V^\rho + \Gamma^\rho_{\mu\sigma} \nabla_\nu V^\sigma - \partial_\nu \left(\nabla_\mu V^\rho\right) + \Gamma^\sigma_{\nu\mu} \nabla_\sigma V^\rho - \Gamma^\rho_{\nu\sigma} \nabla_\mu V \\
&= \partial_\mu \left(\cancel{\partial_\nu V^\rho} + \Gamma^\rho_{\nu\gamma} V^\gamma\right) - \partial_\nu \left(\cancel{\partial_\mu V^\rho} + \Gamma^\rho_{\mu\gamma} V^\gamma\right) \\
&\quad + \Gamma^\rho_{\mu\sigma} \left(\partial_\nu V^\sigma + \Gamma^\sigma_{\nu\gamma} V^\gamma\right) - \Gamma^\rho_{\nu\sigma} \left(\partial_\mu V^\sigma + \Gamma^\sigma_{\mu\gamma} V^\gamma\right) - \left(\Gamma^\sigma_{\mu\nu} - \Gamma^\sigma_{\nu\mu}\right) \nabla_\sigma V^\rho, \\
&= \partial_\mu \Gamma^\rho_{\nu\gamma} V^\gamma + \cancel{\Gamma^\rho_{\nu\gamma} \partial_\mu V^\gamma} - \partial_\nu \Gamma^\rho_{\mu\gamma} V^\gamma - \cancel{\Gamma^\rho_{\mu\gamma} \partial_\nu V^\gamma} \\
&\quad + \Gamma^\rho_{\mu\sigma} \left(\cancel{\partial_\nu V^\sigma} + \Gamma^\sigma_{\nu\gamma} V^\gamma\right) - \Gamma^\rho_{\nu\sigma} \left(\cancel{\partial_\mu V^\sigma} + \Gamma^\sigma_{\mu\gamma} V^\gamma\right) - \left(\Gamma^\sigma_{\mu\nu} - \Gamma^\sigma_{\nu\mu}\right) \nabla_\sigma V^\rho, \\
&= \left(\partial_\mu \Gamma^\rho_{\nu\gamma} - \partial_\nu \Gamma^\rho_{\mu\gamma} + \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\nu\gamma} - \Gamma^\rho_{\nu\sigma} \Gamma^\sigma_{\mu\gamma}\right) V^\gamma - 2\Gamma^\sigma_{[\mu\nu]} \nabla_\sigma V^\rho, \\
&= R^\rho{}_{\gamma\mu\nu} V^\gamma - 2\Gamma^\sigma_{[\mu\nu]} \nabla_\sigma V^\rho. \qquad (11
\end{aligned}$$

Notice that this difference between 2 vectors is considered in the spacetime with torsion because the term $\Gamma^\sigma_{[\mu\nu]}$ does not vanish. Hence, for torsionless spacetime, we also have

$$\left(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu\right) \left(V^\rho W_\rho\right) \begin{cases} = 0, & \text{for torsionless spacetime,} \\ \neq 0, & \text{for spacetime with torsion.} \end{cases} \tag{118}$$

Next, we will consider the torsionless spacetime. Let us first define

$$O_{\mu\nu} \equiv \nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu. \tag{119}$$

From (4.5), we can write

$$\begin{aligned}
O_{\mu\nu} V^\rho W_\rho + V^\rho O_{\mu\nu} W_\rho &= 0, \\
V^\rho O_{\mu\nu} W_\rho &= -R^\rho{}_{\gamma\mu\nu} V^\gamma W_\rho.
\end{aligned} \tag{120}$$

The difference of the dual vector parallel transported along paths gives us

$$O_{\mu\nu} W_\rho = -R^\gamma{}_{\rho\mu\nu} W_\gamma. \tag{121}$$

In general, for any (p,q) tensor, we have

$$\begin{aligned}
O_{\mu\nu} T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} &= R^{\mu_1}{}_{\gamma\mu\nu} T^{\gamma\mu_2 \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} + \cdots + R^{\mu_i}{}_{\gamma\mu\nu} T^{\mu_1 \cdots \mu_{i-1} \gamma \mu_{i+1} \cdots \mu_p}{}_{\nu_1 \cdots \nu_q} + \cdots + R^{\mu_p}{}_{\gamma\mu\nu} T^{\mu_1 \cdots \mu_{p-1} \gamma}{}_{\nu_1 \cdots \nu_q} \\
&\quad - R^\gamma{}_{\nu_1 \mu\nu} T^{\mu_1 \cdots \mu_p}{}_{\gamma \nu_2 \cdots \nu_q} - \cdots - R^\gamma{}_{\nu_i \mu\nu} T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \cdots \nu_{i-1} \gamma \nu_{i+1} \nu_q} - \cdots - R^\gamma{}_{\nu_q \mu\nu} T^{\mu_1 \cdots \mu_p}{}_{\nu_1 \cdots \nu_{q-1} \gamma}.
\end{aligned} \tag{122}$$

4.6 Properties of Riemann tensor

From the definition of Riemanian tensor,

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma}. \tag{123}$$

Contracting with $g_{\gamma\rho}$,

$$\begin{aligned} g_{\gamma\rho}R^{\rho}{}_{\sigma\mu\nu} &= R_{\gamma\sigma\mu\nu} = g_{\gamma\rho}\partial_\mu\Gamma^\rho_{\nu\sigma} - g_{\gamma\rho}\partial_\nu\Gamma^\rho_{\mu\sigma} + g_{\gamma\rho}\Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - g_{\gamma\rho}\Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}, \\ &= \partial_\mu(g_{\gamma\rho}\Gamma^\rho_{\nu\sigma}) - \partial_\mu g_{\gamma\rho}\Gamma^\rho_{\nu\sigma} - \partial_\nu(g_{\gamma\rho}\Gamma^\rho_{\mu\sigma}) + \partial_\nu g_{\gamma\rho}\Gamma^\rho_{\mu\sigma} + g_{\gamma\rho}\Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - g_{\gamma\rho}\Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}, \end{aligned} \tag{124}$$

Choosing to consider in geodesic coordinate ($\Gamma^\rho_{\mu\nu} = 0$, but $\partial_\sigma\Gamma^\rho_{\mu\nu} \neq 0$), we can write

$$\begin{aligned} R_{\gamma\sigma\mu\nu} &= \partial_\mu(g_{\gamma\rho}\Gamma^\rho_{\nu\sigma}) - \partial_\nu(g_{\gamma\rho}\Gamma^\rho_{\mu\sigma}), \\ &= \frac{1}{2}\partial_\mu\left[g_{\gamma\rho}g^{\rho\lambda}\left(\cancel{\partial_\nu g_{\sigma\lambda}} + \partial_\sigma g_{\nu\lambda} - \partial_\lambda g_{\nu\sigma}\right)\right] - \frac{1}{2}\partial_\nu\left[g_{\gamma\rho}g^{\rho\lambda}\big(\cancel{\partial_\mu g_{\sigma\lambda}} + \partial_\sigma g_{\mu\lambda} - \partial_\lambda g_{\mu\sigma}\big)\right] \\ &= \frac{1}{2}\left(\partial_\mu\partial_\sigma g_{\gamma\nu} - \partial_\mu\partial_\gamma g_{\sigma\nu} + \partial_\nu\partial_\gamma g_{\sigma\mu} - \partial_\nu\partial_\sigma g_{\gamma\mu}\right). \end{aligned} \tag{125}$$

It is easy to see that

$$R_{\gamma\sigma\mu\nu} = -R_{\gamma\sigma\nu\mu}, \tag{126}$$

$$R_{\gamma\sigma\mu\nu} = -R_{\sigma\gamma\mu\nu}, \tag{127}$$

$$R_{\gamma\sigma\mu\nu} = R_{\mu\nu\gamma\sigma}. \tag{128}$$

These properties not only exist in the geodesic coordinates, but also in all coordinates.

Let consider

$$\vec{d}\vec{w}_1 = \nabla_{[\nu}a_{\rho]}\,\vec{\theta^\nu}\wedge\vec{\theta^\rho}, \tag{129}$$

$$\vec{d}^2\vec{w}_1 = \nabla_{[\mu}\nabla_\nu a_{\rho]}\,\vec{\theta^\mu}\wedge\vec{\theta^\nu}\wedge\vec{\theta^\rho} = 0. \tag{130}$$

Since the basis does not vanish, the components are

$$\begin{aligned} 0 &= \nabla_{[\mu}\nabla_\nu a_{\rho]}, \\ &= \frac{1}{2}\left(\nabla_{[\mu}\nabla_\nu a_{\rho]} - \nabla_{[\nu}\nabla_\mu a_{\rho]}\right), \\ &= \frac{1}{2}R^\gamma{}_{[\rho\mu\nu]}a_\gamma. \end{aligned} \tag{131}$$

We obtain another property of the Riemannian tensor

$$R^\gamma{}_{[\sigma\mu\nu]} = 0. \tag{132}$$

It can be alternatively written as

$$R_{[\gamma\sigma\mu\nu]} = 0. \tag{133}$$

Eventually, we have 4 independent properties which are (126)-(128) and (132)(or (133)).

Counting degrees of freedom

In n-dimensional spacetime, from the property in (128), we can consider the Riemannian tensor as the 2-rank tensor,

$$R_{\gamma\sigma\mu\nu} = T_{ab}, \tag{134}$$

where the new indices are defined as the pair of the old ones, $a \equiv \gamma\sigma$ and $b \equiv \mu\nu$. Suppose that each a and b contains m degrees of freedom. Since T_{ab} is symmetric, the number of degrees of freedom is $m(m+1)/2$.

From the properties in (126) and (127), it is anti-symmetric for each pair $\gamma\sigma$ and $\mu\nu$. Thus a and b contain $n(n-1)/2$ degrees of freedom. Then the properties (126)-(128) give us that

$$\#\text{ of d.o.f of } R_{\gamma\sigma\mu\nu}\Big|_{\text{using }(126)-(128)} = \frac{\left(\frac{n(n-1)}{2}\right)\left(\frac{n(n-1)}{2}+1\right)}{2}. \tag{135}$$

However, there is 1 property left which is (132) or (133). We choose to use (133). It is found that (133) will eliminate $n(n-1)(n-2)(n-3)/4!$ degrees of freedom. Thus we have

$$\begin{aligned}
\#\text{ of d.o.f of } R_{\gamma\sigma\mu\nu} &= \frac{\left(\frac{n(n-1)}{2}\right)\left(\frac{n(n-1)}{2}+1\right)}{2} - \frac{n(n-1)(n-2)(n-3)}{4!}, \\
&= \frac{n(n-1)}{24}\left[3n(n-1)+6-(n-2)(n-3)\right], \\
&= \frac{n(n-1)}{24}(2n^2+2n), \\
&= \frac{1}{12}n^2(n-1)(n+1). \qquad (136)
\end{aligned}$$

We can see that there are 20 degrees of freedom for 4-dimensional spacetime. Notice that the number of degrees of freedom in 1 dimension is zero. This means that there is no curvature in 1 dimension. We need information in higher dimensions in order to identify whether the line in 1 dimension is straight or curved.

Bianchi Identity

Considering

$$O_{[\mu\nu}\nabla_{\rho]}W_\sigma = -\overset{=\,0\text{ by }(132)}{\cancel{R^{\lambda}}}{}_{[\rho\mu\nu]}\nabla_\lambda W_\sigma - R^{\lambda}{}_{\sigma[\mu\nu}\nabla_{\rho]}W_\lambda \tag{137}$$

and

$$\begin{aligned}\nabla_{[\rho}O_{\mu\nu]}W_\sigma &= -\nabla_{[\rho}\left(R^\lambda{}_{|\sigma|\mu\nu]}W_\lambda\right),\\ &= -\nabla_{[\rho}R^\lambda{}_{|\sigma|\mu\nu]}W_\lambda - R^\lambda{}_{\sigma[\mu\nu}\nabla_{\rho]}W_\lambda. \qquad (138)\end{aligned}$$

In order to obtain (137) = (138), the term $\nabla_{[\rho}R^\lambda{}_{|\sigma|\mu\nu]}W_\lambda$ must be zero. Then

$$\nabla_{[\rho}R_{|\lambda\sigma|\mu\nu]}W^\lambda = 0, \qquad (139)$$

or

$$\begin{aligned}\nabla_{[\rho}R_{|\lambda\sigma|\mu\nu]} &= 0,\\ \nabla_{[\rho}R_{\mu\nu]\lambda\sigma} &= 0, \qquad (140)\end{aligned}$$

which is called the Bianchi identity.

Ricci tensor

In order to $R^\gamma{}_{\gamma\mu\nu} = 0$

Ricci tensor

$$R^\gamma{}_{\mu\gamma\nu} = R_{\mu\nu} \qquad (141)$$

Ricci scalar

Ricci scalar

$$R = g^{\mu\nu}R_{\mu\nu}. \qquad (142)$$

Einstein tensor

Contracting (140) with $g^{\rho\lambda}g^{\mu\sigma}$,

$$\begin{aligned}g^{\rho\lambda}g^{\mu\sigma}\nabla_{[\rho}R_{\mu\nu]\lambda\sigma} &= g^{\rho\lambda}g^{\mu\sigma}\frac{1}{3}\left(\nabla_\rho R_{\mu\nu\lambda\sigma} + \nabla_\mu R_{\nu\rho\lambda\sigma} + \nabla_\nu R_{\rho\mu\lambda\sigma}\right),\\ 0 &= \frac{1}{3}\left[\nabla^\lambda(-R_{\nu\lambda}) + \nabla^\sigma(-R_{\nu\sigma}) + \nabla_\nu R\right],\\ &= -\frac{2}{3}\nabla^\mu\left(R_{\nu\mu} - \frac{1}{2}g_{\nu\mu}R\right),\\ &= -\frac{2}{3}\nabla^\mu G_{\nu\mu}. \qquad (143)\end{aligned}$$

The tensor $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is called the Einstein tensor. Note that this tensor is divergence-less, $\nabla^\mu G_{\mu\nu} = 0$, corresponding to the behaviour of matter which obeys the law of conservation of energy. Hence, the Einstein tensor is useful to explain the curvature of spacetime due to the existence of matter, as we will discuss later.