Introduction to General Relativity

Lectured by Jeong-Hyuck Park

Department of Physics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 04107, Korea

Scribed by Hyunwoo Kwon

Department of Mathematics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 04107, Korea

Abstract

These are lecture notes that I typed up for Professor Jeong-Hyuck Park's course (PHY4010) on General Relativity in Spring 2017. I should note that these notes are not polished and hence might be riddled with errors. If you notice any typos or errors, please do contact me at willkwon@sogang.ac.kr[1]

[1]Revised: February 7, 2019

Contents

1

Lorentz group and Lorentz transformation

1.1 Lorentz transformation

For temporary, we introduce some notation. Let

$$\mathbb{R}^{1+3} = \{(t, x_1, \ldots, x_3) : t, x_i \in \mathbb{R}\}.$$

We regard t as a *time variable* and $(x_1, \ldots, x_3)$ as a *spatial variable*. In the case of vector field, we use boldface notation. For $\boldsymbol{E} : \mathbb{R}^{1+3} \to \mathbb{R}^3$, we define

$$(\nabla \times \boldsymbol{E})^i = \sum_{j,k} \varepsilon^{ijk} \frac{\partial}{\partial x_j} E^k,$$

$$\nabla \cdot \boldsymbol{E} = \sum_i \frac{\partial E^i}{\partial x_i},$$

the *curl* of $\boldsymbol{E}$ and the *divergence* of $\boldsymbol{E}$. Here ε^{ijk} is the Levi-Civita symbol defined by

$$\varepsilon^{ijk} = \begin{cases} 1 & (i,j,k) \text{ is a even permutation of } (1,2,3) \\ -1 & (i,j,k) \text{ is a odd permutation of } (1,2,3) \\ 0 & \text{otherwise.} \end{cases}$$

We take curl and divergence in spatial variable only. We denote ∇ the gradient defined by

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right)$$

and $\nabla^2 f$ the *Laplacian* defined by

$$\nabla^2 f = \sum_{i=1}^{3} \frac{\partial^2 f}{\partial x_i^2}.$$

Consider the following Maxwell equation in the differential form:

$$\begin{cases} \nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} \\ \nabla \cdot \boldsymbol{B} = 0 \\ \nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \\ \nabla \times \boldsymbol{B} = \mu_0 \left(\boldsymbol{J} + \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \right) \end{cases}$$

Here $\boldsymbol{E} : \mathbb{R}^{1+3} \to \mathbb{R}^3$, $\boldsymbol{B} : \mathbb{R}^{1+3} \to \mathbb{R}^3$ denote the electric field, magnetic field, respectively. ρ denotes the electric charge density and $\boldsymbol{J}$ denotes the current density. We denote ε_0 the permittivity of free space and μ_0 the permeability of free space.

Exercise 1.1. If $\rho = 0$ and $\boldsymbol{J} = \boldsymbol{0}$, prove that $\boldsymbol{E}$ and $\boldsymbol{B}$ satisfies the equation $\square \boldsymbol{F} = \boldsymbol{0}$, where

$$\square = \left(\frac{\partial^2}{\partial t^2} - \frac{1}{c^2} \nabla^2 \right)$$

and $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$.

We call $\square$ the *wave operator* or the *d'Alembertian* on $\mathbb{R}^{1+3}$.

Remark. Let us recall how to solve the wave equation in $\mathbb{R}^{1+1}$:

$$\frac{\partial^2 u}{\partial t^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial x^2} = 0.$$

Now we consider the following new set of variables $\xi = x+ct, \eta = x-ct$. Define $v\left(\xi,\eta\right) = u\left(x,t\right)$. Then the change of variables formula shows that

$$\frac{\partial^2 v}{\partial\xi\partial\eta} = 0.$$

Integrating this twice gives

$$v\left(\xi,\eta\right) = F\left(\xi\right) + G\left(\eta\right),$$

i.e.,

$$u\left(x,t\right) = F\left(x+ct\right) + G\left(x-ct\right).$$

For the case of higher dimension, see [1, 2].

In what follows, c denote the speed of light. Unfortunately, the Maxwell equation is not Galieli invariance. We need some symmetry to study the Maxwell equation well. The Lorentz transform gives a positive answer for this question. First, let us explain the Lorentz transform in physical way.

Consider two parallel mirrors and let L to be the distance between mirrors. Then light travels once in $\Delta t = \frac{2L}{c}$.

Now we consider a train moving with a constant speed v in the x-direction. Then if we look the light outside of the train in the coordinate (t,x,y,z) with $(t,x,0,0)$ we have

$$\frac{2\sqrt{L^2 + v^2\left(\frac{\Delta t}{2}\right)^2}}{\Delta t} = c,$$

Here Δt denotes the time which light travels between mirrors. Solving this equation, we obtain

$$\Delta t = \frac{2L}{\sqrt{c^2 - v^2}}.$$

If we see the light in the train in the moving coordinate (t',x',y',z') with $(t',0,0,0)$, we have

$$\Delta t' = \frac{2L}{c}.$$

Since L is invariant, we obtain

$$c\Delta t' = \sqrt{c^2 - v^2}\left(\Delta t\right),$$

i.e.,

$$\Delta t = \frac{\Delta t'}{\sqrt{1-\eta^2}},$$

where $\eta = \frac{v}{c}$. We obtain the time-dillation. Write $\gamma = \frac{1}{\sqrt{1-\eta^2}}$. Then

$$(t',0,0,0) \iff (\gamma t', v\gamma t', 0, 0) = \left(t,x,y,z\right).$$

This is the special case of Lorentz transformation.

Now let's move to more general case. Consider a fixed frame S and moving frame S' with velocity v in the x-direction.

Consider an event A in (t'_A, x'_A, y'_A, z'_A) if we write it in S' frame and in (t_A, x_A, y_A, z_A) if we write it in S frame. Also consider an event B in $(t'_B, 0, 0, 0)$ if we write it in S' frame and $(t_B, x_B, 0, 0)$ if we write it in S frame.

Now consider a light starting from A to the B. Then the light travels

$$t'_B = t'_A + \frac{\sqrt{(x'_A)^2 + (y'_A)^2 + (z'_A)^2}}{c}$$

in S' and

$$\begin{aligned} t_B &= t_A + \frac{\sqrt{(x_A - vt_B)^2 + y_A^2 + z_A^2}}{c} \\ x_B &= vt_B. \end{aligned} \tag{1.1}$$

Then solve the equation (1.1) to get

$$\begin{aligned} t_B &= \frac{c^2 t_A - v x_A + \sqrt{(c^2 t_A - v x_A)^2 - (c^2 - v^2)(c^2 t_A^2 - x_A^2 - y_A^2 - z_A^2)}}{c^2 - v^2} \\ &= \gamma^2 \left\{ t_A - \frac{\eta}{c} x_A + \frac{1}{c}\sqrt{(x_A - vt_A)^2 + \frac{1}{\gamma^2}(y_A^2 + z_A^2)} \right\}. \end{aligned}$$

Since $t_B = \gamma t'_B$,

$$\gamma \left\{ t_A - \frac{\eta}{c} x_A + \frac{1}{c}\sqrt{(x_A - vt_A)^2 + \frac{1}{\gamma^2}(y_A^2 + z_A^2)} \right\} = t'_A + \frac{\sqrt{(x'_A)^2 + (y'_A)^2 + (z'_A)^2}}{c}.$$

Since $y_A = y'_A$ and $z_A = z'_A$

$$t'_A + \frac{\sqrt{(x'_A)^2 + (y_A)^2 + (z_A)^2}}{c} = \gamma \left[t_A - \frac{\eta}{c} x_A \right] + \frac{1}{c}\sqrt{(x_A - vt_A)^2 + y_A^2 + z_A^2}.$$

The equation holds for all y, z. Hence we obtain

$$\begin{cases} t' &= \gamma\left(t - \frac{\eta}{c}x\right) \\ x' &= \gamma\left(x - vt\right) \\ y' &= y \\ z' &= z. \end{cases}$$

In the matrix form, we write

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\frac{\eta}{c} & 0 & 0 \\ -\frac{\eta}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}.$$

We call this transform as the *Lorentz transform.*

Observe that $\gamma^2 - (\gamma\eta)^2 = 1$ since $\gamma = \frac{1}{\sqrt{1-\eta^2}}$. Then there is ϕ such that $\cosh\phi = \gamma$ and $\sinh\phi = \gamma\eta$. So we can write

$$\begin{bmatrix} ct \\ x \end{bmatrix} = \begin{bmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{bmatrix} \begin{bmatrix} ct' \\ x' \end{bmatrix}.$$

It seems similar to the rotation matrix in $\mathbb{R}^2$. In the next section, we analyze this kind of matrices in detail.

1.2 The Lorentz group

Let us compare two matrices

$$A = \begin{bmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{bmatrix},$$
$$B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

In the case of B, $B^T B = I$ and

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = B \begin{bmatrix} x \\ y \end{bmatrix}$$

implies

$$(x')^2 + (y')^2 = x^2 + y^2.$$

Indeed,

$$\begin{aligned} (x')^2 + (y')^2 &= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} B^T B \begin{bmatrix} x \\ y \end{bmatrix} \\ &= x^2 + y^2. \end{aligned}$$

We can regard $B^T B = I$ as $B^T I B = I$, i.e., B stablizes I.

By considering $\cosh^2\phi - \sinh^2\phi = 1$, we obtain the following relation holds:

$$A^T \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Also by the similar reason as before, we obtain

$$-(ct)^2 + x^2 = -(c't')^2 + (x')^2$$

if

$$\begin{bmatrix} ct \\ x \end{bmatrix} = A \begin{bmatrix} ct' \\ x' \end{bmatrix}.$$

Write

$$B(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Then entries by entries differentiation with respect to θ gives

$$B'(\theta) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} B(\theta).$$

Inductively, we see

$$\frac{d^n}{d\theta^n} B(\theta) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^n B(\theta).$$

Evaluating at $\theta = 0$, we get

$$\left.\frac{d^n}{d\theta^n} B\right|_{\theta=0} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^n.$$

So we obtain the series expression of $B(\theta)$:

$$\begin{aligned} B(\theta) &= \sum_{n=0}^{\infty} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^n \frac{1}{n!}\theta^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix}^n \\ &=: \exp\left(\begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix} \right). \end{aligned}$$

Note that $\frac{dB}{d\theta} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} B(\theta)$ and $\frac{dB^T}{d\theta} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} B(\theta)$. Hence

$$\frac{d}{d\theta}\left(B^T I B\right) = \frac{d}{d\theta}\left(B^T\right) B + B^T \frac{dB}{d\theta} = O.$$

Similarly,

$$\frac{dA}{d\phi} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A$$

and inductively we have

$$\frac{d^n A}{d\phi^n} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^n A.$$

Also $\frac{dA}{d\phi} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A$ and $\frac{dA^T}{d\phi} = A^T \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Hence

$$\begin{aligned} \frac{d}{d\phi}\left(A^T \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} A\right) &= A^T \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) A \\ &= A^T \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right) A \\ &= O. \end{aligned}$$

Note that we have $A(\phi_1) A(\phi_2) = A(\phi_1 + \phi_2)$ and $B(\theta_1) B(\theta_2) = B(\theta_1 + \theta_2)$ by the addition law of hyperbolic cosine, sine and cosine and sine, respectively.

Let us recall the definition of group.

Definition 1.2. A group $(G, \cdot)$ is a set G with binary operation $\cdot : G \times G \to G$ satisfying

1. $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ for all $g_1, g_2, g_3 \in G$;
2. there is $e \in G$ such that $g \cdot e = e \cdot g = g$ for all $g \in G$;
3. for any $g \in G$, there exists $g' \in G$ such that $g' \cdot g = e$.

There are many examples of groups. $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}\setminus\{0\}, \times)$ are groups.

We define

$$O(n) = \left\{ A \in \mathcal{M}_{n\times n}(\mathbb{R}) : A^T A = I_n \right\}.$$

With matrix multiplication, it is a group and we call $O(n)$ as the *orthogonal group* of order n. There is a subgroup

$$SO(n) = \{A \in O(n) : \det(A) = 1\}$$

of $O(n)$ and we call this the *special orthogonal group* of order n.

We define $O(p,q)$ the set of all $(p+q)\times(p+q)$ invertible matrices C satisfying

$$C^T\eta C=\eta$$

where $\eta=\operatorname{diag}\left(\underbrace{-1,\ldots,-1}_{p-\text{times}},\underbrace{1,\ldots,1}_{q-\text{times}}\right)$. It is easy to check that $O(p,q)$ is a group with matrix multiplication. We call this group as *indefinite-orthogonal* group. Similarly, there is a subgroup

$$SO(p,q)=\{A\in O(p,q):\det(A)=1\}$$

of $O(p,q)$ and we call this the *special indefinite-orthogonal group* of order p and q.

In the special case $SO(1,3)$, we call this group as the *Lorentz group*. Usually, we write L as an element of $SO(1,3)$. Under the relation

$$\begin{bmatrix}ct'\\x'\\y'\\z'\end{bmatrix}=L\begin{bmatrix}ct\\x\\y\\z\end{bmatrix},$$

if we denote $\eta=\operatorname{diag}(-1,1,1,1)$, then we have

$$\begin{aligned}&=-\left(cdt'\right)^2+\left(dx'\right)^2+\left(dy'\right)^2+\left(dz'\right)^2\\&=-\left(cdt\right)^2+dx^2+dy^2+dz^2\end{aligned}$$

In this sense, we denote $ds^2=-\left(cdt\right)^2+dx^2+dy^2+dz^2$ and we regard this as the proper length in our setting.

1.3 The Einstein convention

In order to describe the theory of general relativity, we need to introduce some notations and rules. It might confuse in the first time.

From now on we write

$$x^\mu=\left(x^0,x^1,x^2,x^3\right),\quad \mu=0,1,2,3$$

and we call μ as the space-time indices. Here we regard x^0 as ct. We call this coordinate as *Minkowskian* 4*-dimensional spacetime coordinate*. We call

$$\eta=\begin{bmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

as the Minkowski metric.

Now we study the Einstein conventions. There are four types of expressing matrices: $M_{\mu\nu}$, $K^{\mu\nu}$, $L^{\mu}{}_{\nu}$, $N_{\mu}{}^{\nu}$. Multiplication is allowed if there is a contraction between indices. As an example, MK is allowed since

$$M_{\mu\nu}K^{\nu\lambda}=(ML)_{\mu}^{\lambda}$$

since ν is a contraction. Also ML is allowed since

$$M_{\mu\nu}L^{\nu}_{\lambda}=(ML)_{\mu\lambda}.$$

However MM and MN are not allowed since between M and M, there is no contraction. Also $M_{\mu\nu}N_{\mu}^{\ \nu}$ has no contraction.

If $M_{\mu\nu}$, $K^{\mu\nu}$, $L^{\mu}_{\ \nu}$, $N_{\mu}^{\ \nu}$, then their inverses are

$$M^{-1} = M^{\mu\nu}, \quad K^{-1} = K_{\mu\nu}, \quad L^{-1} = L_{\nu}^{\ \mu}, \quad N^{-1} = N^{\nu}_{\ \mu}$$

and their transpose are

$$\left(M^T\right)_{\mu\nu} \quad \left(K^T\right)^{\mu\nu} \quad \left(L^T\right)_{\mu}^{\ \nu} \quad \left(N^T\right)^{\mu}_{\ \nu}.$$

From this

$$L^T \eta L = \eta \iff \eta_{\mu\nu} = L_{\mu}^{\ \lambda}\eta_{\lambda\rho}L^{\rho}_{\ \nu}$$

We write

$$\frac{\partial}{\partial x^\mu} = \partial_\mu.$$

Then note that

$$\frac{\partial \left(x'\right)^\mu}{\partial x^\nu} = L^{\mu}_{\ \nu}, \quad \frac{\partial x^\mu}{\partial \left(x'\right)^\nu} = \left(L^{-1}\right)_{\mu}^{\ \nu}.$$

Then by chain rule,

$$\partial'_\mu = \frac{\partial}{\partial \left(x'\right)^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\partial_\mu = \left(L^{-1}\right)_{\mu}^{\ \nu}\partial_\nu.$$

For the case $\eta_{\mu\nu}$, $\left(\eta^{-1}\right)^{\mu\nu} \equiv \eta^{\mu\nu}$ by definition. Metric interchanges the placement of indicies. As an example

$$v^\mu \eta_{\mu\nu} = v_\nu, \quad \eta^{\mu\nu}v_\nu = v^\mu.$$

So $dx_\nu = \eta_{\mu\nu}dx^\nu$, $ds^2 = dx^\mu dx^\nu \eta_{\mu\nu} \equiv dx^\mu dx_\mu$.

If we can write some equality which can be written by Einstein convention, we write

$$\eta_{\mu\nu} = \eta_{\lambda\rho}L_{\mu}^{\ \lambda}L^{\rho}_{\ \nu}$$

instead of

$$\eta_{\mu\nu} = L_{\mu}^{\ \lambda}\eta_{\lambda\rho}L^{\rho}_{\ \nu}.$$

Note that

$$\begin{aligned}
\left(M_1 M_2\right)^{\lambda}_{\ \nu} &= \left(M_1\right)^{\lambda}_{\ \rho}\left(M_2\right)^{\rho}_{\ \nu} = \left(M_2\right)^{\rho}_{\ \nu}\left(M_1\right)^{\lambda}_{\ \rho} \\
\left(M_2 M_1\right)^{\lambda}_{\ \nu} &= \left(M_2\right)^{\lambda}_{\ \rho}\left(M_1\right)^{\rho}_{\ \nu} = \left(M_1\right)^{\rho}_{\ \nu}\left(M_2\right)^{\lambda}_{\ \rho}.
\end{aligned}$$

So in general $(M_1 M_2) \neq (M_2 M_1)$.

Note that

$$\begin{aligned}
\delta^\mu_\nu &= \eta^{\mu\kappa}L^{\lambda}_{\ \nu}L^{\rho}_{\kappa}\eta_{\lambda\rho} \\
&= \left(\eta_{\lambda\rho}L^{\rho}_{\kappa}\eta^{\kappa\mu}\right)L^{\lambda}_{\ \nu} \\
&= L^{\mu}_{\ \lambda}L^{\lambda}_{\ \nu}.
\end{aligned}$$

Hence $L_{\lambda}^{\ \mu} = \left(L^{-1}\right)^{\mu}_{\ \lambda}$.

In this convention,

$$\begin{aligned}
\partial'_\mu &= \left(L^{-1}\right)^{\nu}_{\ \mu}\partial_\nu = L_{\mu}^{\ \nu}\partial_\nu \\
d\left(x'\right)^\mu &= L^{\mu}_{\ \nu}dx^\nu.
\end{aligned}$$

We define

$$d\tau^2 = dt^2 - \vec{dx} \cdot \vec{dx}$$

and is called the *proper time*. In nature, $d\tau^2 \geq 0$. So $c^2 \geq \left|\frac{dx}{dt}\right|^2$.

In some sense $\frac{dx^\mu}{dt} = \frac{dx^\mu}{\frac{dx^0}{c}}$ is not inappropriate. From now on, we assume $c \equiv 1$. Note that

$$\frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt\sqrt{1 - \vec{v} \cdot \vec{v}}}.$$

So

$$\frac{d\vec{x}}{d\tau} = \frac{\vec{v}}{\sqrt{1 - \vec{v}^2}}$$

Note that

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - v^2}} \approx 1 + \frac{1}{2}v^2.$$

Hence the momentum of

$$P^\mu = m\frac{dx^\mu}{d\tau},$$

in particular,

$$P^0 = m + \frac{1}{2}mv^2 + \cdots.$$

We write the Maxwell equation in simple way. Let $F = F_{\mu\nu}$ be a 2-form defined by

$$E_i = F_{0i}, \quad B_i = \frac{1}{2}\varepsilon_{ijk}F^{jk},$$

where i, j, k runs $\{1, 2, 3\}$.

Then

$$\begin{aligned}
&\partial_t F_{0i} - \varepsilon^{ijk}\partial^j B^k \\
&= \partial_0 F_{0i} + \frac{1}{2}\varepsilon_{ijk}\varepsilon^{iab}\partial^j F_{ab} \\
&= \partial_0 F_{0i} + \frac{1}{2}(\delta_i^a\delta_j^b - \delta_i^b\delta_j^a)\partial^j F_{ab} \\
&= \partial_0 F_{0i} + \frac{1}{2}(\partial^j F_{ij} + \partial^j F_{ij}) \\
&= \partial_0 F_{0i} + \partial^j F_{ij} \\
&= -\partial_0 F_{i0} + \partial^j F_{ij} \\
&= \eta^{\mu\nu}\partial_\mu F_{i\nu}.
\end{aligned}$$

So $\eta^{\mu\nu}\partial_\lambda\mu F_{i\nu} = j_i$. Also $\nabla \cdot E = \rho$ implies $\partial_j F_0^i = \rho$. So $\eta^{\mu\nu}\partial_\mu F_\nu = \rho$. We write $J = (\rho, j_1, j_2, j_3)$, then we get

$$\eta^{\mu\nu}\partial_\mu F_{\lambda\nu} = J_\lambda.$$

From $\nabla \cdot B = 0$ and $\partial_t B + \nabla \cdot \times E = 0$, we obtain

$$\partial_\lambda F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0.$$

Hence, we rewrite the Maxwell equation in the following way:

$$\begin{cases} \eta^{\mu\nu}\partial_\mu F_{\lambda\nu} = J_\lambda. \\ \partial_\lambda F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0. \end{cases}$$

Recall

$$x^\mu \mapsto (x')^\mu = L^\mu{}_\nu x^\nu$$
$$\partial_\mu \mapsto \partial'_\mu = \left(L^{-1}\right)^\nu{}_\mu \partial_\nu = L_\mu{}^\nu \partial_\nu$$

under the Lorentz transformation. Hence

$$F_{\mu\nu} \mapsto L_\mu{}^\lambda L_\nu{}^\rho F_{\lambda\rho}$$

gives the Maxwell equation invariant under the Lorentz transformation.

Note that $J^\mu \mapsto J'^\mu = L^\mu{}_\nu J$ gives $\partial_\lambda F^{\lambda\mu} = J^\mu$ to

$$\partial'_\lambda F'^{\lambda\mu} = J^\mu$$

to

$$L_\lambda{}^\rho \partial_\rho \left(L^\lambda{}_\kappa L^\mu{}_\rho F^{\kappa\sigma}(x)\right) = L^\mu{}_\nu J^\nu.$$

On the LHS, this is equal to

$$\delta^\rho_\kappa \partial_\rho F^{\kappa\sigma} L^\mu{}_\sigma = L^\mu{}_\sigma \partial_\rho F^{\rho\sigma}.$$

So

$$L^\mu{}_\sigma \left(\partial_\rho F^{\rho\sigma} - J^\sigma\right) = 0,$$

which reduces to the original Maxwell equation.

1.4 Baker-Campbell-Hausdorff formula

Recall

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!},$$
$$\ln(1+x) = \sum_{n=0}^{\infty} -\frac{(-x)^n}{n!}$$
$$\ln\left(e^A\right) = A.$$

We write

$$e^A e^B = e^C.$$

If A and B commutes, then $C = A + B$. But this is not generally holds. However, we have the following formula:

Theorem 1.3 (Baker-Campbell-Hausdorff formula)**.**

$$\begin{aligned}\ln\left(e^A e^B\right) &= B + \int_0^1 \mathrm{d}t \sum_{n=1}^{\infty} \frac{1}{n}\left(1 - e^{t\mathrm{ad}A} e^{\mathrm{adB}}\right)^{n-1} A \\ &= A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{12}[A[A,B]] + \tfrac{1}{12}[B,[B,A]] + +\tfrac{1}{24}[A,[[A,B],B]] + \cdots.\end{aligned} \tag{1.2}$$

which holds for an arbitrary pair of operators, A and B.

Proof. Introduce a real parameter, $t \in \mathbb{R}$ and define

$$C(t) := \ln\left(e^{tA} e^B\right), \qquad e^{C(t)} = e^{tA} e^B. \tag{1.3}$$

We have

$$C(0)=B\,,\qquad \tfrac{\mathrm{d}}{\mathrm{d}t}e^{C(t)}=Ae^{C(t)}\,,\qquad A=\left[\tfrac{\mathrm{d}}{\mathrm{d}t}e^{C(t)}\right]e^{-C(t)}\,. \tag{1.4}$$

Further for an arbitrary, M,

$$e^{\mathrm{ad}C(t)}M=e^{C(t)}Me^{-C(t)}=e^{tA}e^{B}Me^{-B}e^{-tA}=e^{t\mathrm{ad}A}e^{\mathrm{ad}B}M\,, \tag{1.5}$$

and hence, like (1.3),

$$e^{\mathrm{ad}C(t)}=e^{t\mathrm{ad}A}e^{\mathrm{ad}B}\,. \tag{1.6}$$

Further, we set with one more real parameter, $s\in\mathbb{R}$,

$$F(s,t):=\left[\frac{\partial}{\partial t}e^{sC(t)}\right]e^{-sC(t)}\,. \tag{1.7}$$

It is straightforward to see

$$F(0,t)=0\,,\qquad F(1,t)=A\,, \tag{1.8}$$

and

$$\frac{\partial}{\partial s}F(s,t)=e^{sC(t)}\left[\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right]e^{-sC(t)}=e^{s\mathrm{ad}C(t)}\left[\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right]\,. \tag{1.9}$$

Hence,

$$\frac{\mathrm{d}C(t)}{\mathrm{d}t}=e^{-s\mathrm{ad}C(t)}\frac{\partial}{\partial s}F(s,t)\,, \tag{1.10}$$

and

$$\begin{aligned} A &= \int_0^1 \mathrm{d}s\,\frac{\partial}{\partial s}F(s,t)\\ &=\int_0^1\mathrm{d}s\; e^{s\mathrm{ad}C(t)}\left[\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right]\\ &=\int_0^1\mathrm{d}s\sum_{n=0}^{\infty}\frac{s^n}{n!}\,[\mathrm{ad}C(t)]^n\left[\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right]\\ &=\sum_{n=0}^{\infty}\frac{[\mathrm{ad}C(t)]^n}{(n+1)!}\left[\frac{\mathrm{d}C(t)}{\mathrm{d}t}\right]\\ &=G(\mathrm{ad}C(t))\left[\tfrac{\mathrm{d}C(t)}{\mathrm{d}t}\right]\,,\end{aligned} \tag{1.11}$$

where we put a function,

$$\begin{aligned} G(x)&=\frac{e^x-1}{x}=\sum_{n=0}^{\infty}\frac{x^n}{(n+1)!}\,,\\ G(x)^{-1}&=\frac{x}{e^x-1}=-\sum_{k=0}^{\infty}xe^{kx}=\frac{\ln(e^x)}{e^x-1}=-\frac{\ln[1-(1-e^x)]}{1-e^x}=\sum_{n=1}^{\infty}\frac{(1-e^x)^{n-1}}{n}\,.\end{aligned} \tag{1.12}$$

Therefore, we see, with (1.6),

$$\frac{\mathrm{d}C(t)}{\mathrm{d}t}=\sum_{n=1}^{\infty}\frac{1}{n}\left(1-e^{\mathrm{ad}C(t)}\right)^{n-1}A=\sum_{n=1}^{\infty}\frac{1}{n}\left(1-e^{t\mathrm{ad}A}e^{\mathrm{adB}}\right)^{n-1}A\,, \tag{1.13}$$

and finally,

$$\ln\left(e^Ae^B\right)=B+\int_0^1\mathrm{d}t\;\sum_{n=1}^{\infty}\frac{1}{n}\left(1-e^{t\mathrm{ad}A}e^{\mathrm{adB}}\right)^{n-1}A\,. \tag{1.14}$$

This completes the proof. □

In the quantum mechanical setting, since $[\hat{x},\hat{p}] = i\hbar$,

$$e^{a\hat{x}}e^{b\hat{p}} = e^{a\hat{x}+b\hat{p}+\frac{i}{2}\hbar ab}$$

If L is a Lorentz transformation, then $L = e^M$, where $M^T\eta + \eta M = 0$. We will show it. Note that

$$(\eta M_1)^T = -\eta M_1, \quad (\eta M_2)^T = -\eta M_2.$$

So

$$\begin{aligned}(\eta M_1 M_2)^T &= M_2^T (\eta M_1)^T \\ &= -M_2^T \eta M_1 \\ &= \eta M_2 M_1 \\ &\neq -\eta M_1 M_2.\end{aligned}$$

However, we have

$$\begin{aligned}[\eta (M_1 M_2 - M_2 M_1)]^T &= -\eta (M_1 M_2 - M_2 M_1) \\ &= -\eta [M_1, M_2].\end{aligned}$$

In the case $SO(2)$, we write $\exp\left(\theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right)$ and $SO(1,1)$, we write $\exp\left(\phi \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)$. We call $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is a generator[1] of $SO(2)$. Similarly, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a generator of $SO(1,1)$.

Let M_a, $a = 1,\ldots,n$. We define

$$\mathfrak{g} = \{\theta^a M_a : \theta^a \in \mathbb{R}, a = 1,\ldots,n\}$$

and $[M_a, M_b] = f_{ab}^c M_c$ for some f_{ab}^c. We call $\mathfrak{g}$ as a *Lie algebra* and $M_1,\ldots,M_n$ are said to be generator of $\mathfrak{g}$.

Roughly speaking, we define the Lie group as the exponential of Lie algebra.

For $L \in SO(p,q)$, we write $L = \exp(\theta M)$. Note that $L^T \eta L = \eta$ and $\frac{dL}{d\theta} |_{\theta=0} = M$ and $L |_{\theta=0} = I$. Taking derivative, we get

$$M^T\eta + \eta M = 0,$$

i.e., $(\eta M)^T = -\eta M$.

If $e^{\theta M} \equiv L$ and $M^T\eta + \eta M = 0$, then $\frac{d}{d\theta}L = ML = LM$. So

$$\begin{aligned}\frac{d}{d\theta}\left(L^T \eta L\right) &= L^T M^T \eta L + L^T \eta M L \\ &= L^T \left(M^T\eta + \eta M\right) L \\ &= 0.\end{aligned}$$

Hence $L^T\eta L = \eta$.

[1] In the mathematical terminology, this is a basis

Note that there are $\frac{d(d-1)}{2}$ matrices ηM which is anti-symmetric. In the case $d = 4$, $6 = 3 + 3$. Note that

$$\begin{aligned} M = \eta^{-1} \begin{bmatrix} 0 & -\phi_1 & -\phi_2 & -\phi_3 \\ \phi_1 & 0 & \theta_3 & -\theta_2 \\ \phi_2 & -\theta_3 & 0 & \theta_1 \\ \phi_3 & \theta_2 & -\theta_1 & 0 \end{bmatrix} \\ = \begin{bmatrix} 0 & \phi_1 & \phi_2 & \phi_3 \\ \phi_1 & 0 & \theta_3 & -\theta_2 \\ \phi_2 & -\theta_3 & 0 & \theta_1 \\ \phi_3 & \theta_2 & -\theta_1 & 0 \end{bmatrix}. \end{aligned}$$

Hence $(\theta_1, \theta_2, \theta_3)$ represents a rotation and (ϕ_1, ϕ_2, ϕ_3) represents boosts.

Remark. $\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $\sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

$$i\theta_a \sigma^a$$

Recall SU (2), the set of all 2×2 matrices which entries are complex and $U^\dagger U = 1$ and $|\det U| = 1$. Write $u = e^{\theta M}$. Taking derivatives, we have $M^\dagger + M = 0$. So

$$\left| e^{\theta M} \right| = e^{\theta \mathrm{Im} M}.$$

Roughly speaking, SU (2) and SO (3) are same.

1.5 Covariant transformation

Recall the definition of wave function. We say a complex-valued function ψ as a wave function if

$$\int |\psi|^2 \, dx = 1.$$

Write $\psi = |\psi| \, e^{i\varphi}$. One may ask whether phase φ has physical meaning. Unfortunately, φ is non-physical. To deal the theory of Quantum mechanics, complex-valued function is essential, not a real-valued function.

For arbitrary function θ on spacetime ,consider a transformation $\psi \to \psi' = \psi e^{i\theta}$. Consider the Schrödinger equation

$$i\hbar \partial_t \psi = \frac{1}{2m} \left(-i\hbar \nabla \right)^2 \psi + V\psi.$$

Write

$$E\psi = \frac{1}{2m} \vec{p}^2 \psi + V\psi$$

and consider

$$\begin{aligned} E &\to i\hbar \frac{\partial}{\partial t} \\ \vec{p} &\to -i\hbar \nabla. \end{aligned} \tag{1.15}$$

Note that

$$\begin{aligned} P_\mu &= -i\hbar \partial_\mu = (-E, \vec{p}) \\ P^\mu &= m \frac{dx^\mu}{d\tau} = (E, \vec{p}). \end{aligned}$$

So the above transformation (1.15) makes sense.

We want to claim that the transformation $\psi \mapsto \psi'$ does not change the law of physics. Note that

$$\partial_\mu \psi \to \partial_\mu \psi' = \partial_\mu \left(\psi e^{i\theta}\right) = \left(\partial_\mu \psi + i \partial_\mu \theta\right) e^{i\theta}.$$

Then from this transformation, we can check that ψ' does not satisfies the Schrödinger's equation since ∂_μ is not covariant. To overcome this difficulty, we introduce the "covariant derivative"

$$\nabla_\mu = \partial_\mu - iqA_\mu, \quad q \equiv 1.$$

We choose A so that

$$\nabla_\mu \psi \to \nabla'_\mu \psi' = \left(\nabla_\mu \psi\right) e^{i\theta}.$$

Indeed,

$$\begin{aligned}
\nabla'_\mu \psi' &= \left(\partial_\mu - iA'_\mu\right) \psi' \\
&= \left(\partial_\mu \psi' - iA'_\mu \psi'\right) \\
&= \partial_\mu \left(e^{i\theta} \psi\right) - iA'_\mu e^{i\theta} \psi \\
&= \left(\partial_\mu \psi + i \partial_\mu \theta - iA'_\mu\right) e^{i\theta} \psi.
\end{aligned}$$

So if we write

$$A'_\mu = A_\mu - i \partial_\mu \theta,$$

∇_μ behaves 'covariant' under the transformation

$$\begin{aligned}
\psi &\mapsto \psi e^{i\theta} \\
A_\mu &\mapsto A_\mu - \partial_\mu \theta.
\end{aligned}$$

We call this kind of transform as $U(1)$ gauge transform. Note that the transform does not change the physics since $|\psi|$ is invariant.

There are two kind of 'symmetry'. One is local symmetry, that is, there exists a parameter θ depending on spacetime point. This is however is non-physical symmetry. The other is global symmetry, i.e., this is a parameter is constant. We call this symmetry as physical symmetry.

Although the gauge symmetry is not physical symmetry, the concept gauge symmetry is a principle concept in 20th-21st physics.

So far we didn't defined the terminology 'symmetry' in rigorous way. From now on, we define 'symmetry' and give the fundamental theorem due to Noether.

We say a transform is *symmetry* of the Lagrangian L if its action is invariant under the transformation. If the transform is infinitesimal, we say this symmetry as Noether symmetry.

Note that

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} = \frac{d}{dt} K.$$

We state the Noether theorem informally.

Theorem 1.4. *Let L be a Lagrangian. If there is a Noether symmetry, then there exists a conserved charge*

$$Q = \frac{\partial L}{\partial \dot{q}^a} \delta q^a - K.$$

Note that

$$\begin{aligned}\frac{dQ}{dt} &= \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}^a}\delta q^a - K\right] \\ &= \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}^a}\right]\delta q^a + \frac{\partial L}{\partial \dot{q}^a}\left[\delta \dot{q}^a\right] - \frac{dK}{dt} \\ &= \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial \dot{q}^a}\right]\delta q^a + \left[\delta L - \frac{d}{dt}K\right].\end{aligned}$$

Example 1.5. Consider $q(t) \mapsto q(t+a)$, time translation. Then

$$q(t+a) \simeq q(t) + a\dot{q}(t).$$

So

$$\delta L = \dot{q}\frac{\partial L}{\partial q} + \ddot{q}\frac{\partial L}{\partial \dot{q}}, \quad \dot{q} = \frac{\partial L}{\partial t}.$$

Hence $K = L$ and so

$$Q = \frac{\partial L}{\partial q^a}\dot{q}^a - L = H,$$

the Hamiltonian. So by the Noether's theorem, the Hamiltonian is conserved. So we call this kind of symmetry as *time symmetry*.

Example 1.6. Let $\mathcal{L} = \frac{m}{2}\dot{x}^2$, $x \mapsto x + c$, spartial symmetry.
$Q = m\dot{x}$. Then the momentum is conserved by the Noether's theorem.

Example 1.7. $\mathcal{L} = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right)$, consider rotation.
Note that

$$\begin{aligned}\delta\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} y \\ -x \end{bmatrix} \\ \delta\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} &= \begin{bmatrix} \dot{y} \\ -\dot{x} \end{bmatrix}\end{aligned}$$

Then

$$\delta L = m\left(\dot{x}\delta\dot{x} + \dot{y}\delta\dot{y}\right) = 0.$$

So

$$\begin{aligned}Q &= \frac{\partial L}{\partial \dot{x}}\delta\dot{x} + \frac{\partial L}{\partial \dot{y}}\delta\dot{y} \\ &= m\dot{x}y - m\dot{y}x = p_x y - p_y x.\end{aligned}$$

So by the Noether's theorem, the angular momentum is conserved.

Observe the symmetry is global. There is no local dependence. Also, Noether charge satsifies

$$\{Q, H\} \equiv 0,$$

and

$$\{Q, q\} = \delta q, \quad \{Q, P\} = \delta p, \quad \{Q, H\} = 0.$$

Back to the original problem, to recover the covariance, we need

$$i\hbar D_t\psi = \frac{1}{2m}\left(-i\hbar\vec{D}\right)^2\psi,$$

i.e.,

$$i\hbar\left(\partial_t\psi - iA_0\psi\right) = -\frac{\hbar^2}{2m}\left(\vec{\nabla} - i\vec{A}\right)^2\psi.$$

So

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\left(\vec{\nabla} - i\vec{A}\right)^2\psi - \hbar A_0\psi,$$

where

$$A_\mu = \left(\phi, \vec{A}\right)$$

and ϕ is a columb potential and $\vec{A}$ is magentic potential. We call this potential as *electromagnetic potential.*

In Schrödinger's equation, local symmetry, i.e., Gauge symmetry required. The solution $\psi\left(t, x, y, z\right)$ itself does not have physical meaning. However, $|\psi|^2$ has a physical meaning, the probability density. So we can transform ψ to $e^{i\theta(t,x,y,z)}\psi\left(t, x, y, z\right) \equiv \psi'\left(t, x, y, z\right)$, where $\theta\left(t, x, y, z\right)$ is arbitrary function on spacetime. The θ depends on a point. So we can regard it as it has a local symmetry. However, the classical Schrödinger equation does not transform in covariant sense. So we replace ∂_μ to $D_\mu = \partial_\mu - iA_\mu$, where A_μ is vector potential (in the mathematical way, gauge connection).satisfying

$$\begin{aligned} \psi &\mapsto \psi' = e^{i\theta}\psi \\ A_\mu &\mapsto A'_\mu = A_\mu + \partial_\mu\theta \end{aligned}$$

(gauge) transformation rule.

Is their any physical interpretation of A_μ? No since the gauge transformation is non-physical symmetry, i.e.,

$$\eta^{\mu\nu}A_\mu A_\nu \neq \eta^{\mu\nu}A'_\mu A'_\nu.$$

However, we can make a physical quantity from A,

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

field strength of A. It is gauge invariance since

$$\begin{aligned} F'_{\mu\nu} &= \partial_\mu A'_\nu - \partial_\nu A'_\mu \\ &= \partial_\mu\left(A_\nu + \partial_\nu\theta\right) - \partial_\nu\left(A_\mu + \partial_\mu\theta\right) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu\partial_\nu\theta - \partial_\nu\partial_\mu\theta \\ &= F_{\mu\nu}. \end{aligned}$$

It is physical quantity. Note that $F_{\mu\nu} = -F_{\nu\mu}$. Actually, it is a electromagnetic field. Why we have a light? Some people says it is due to gauge symmetry.

Note that $e^{i\theta} \in U\left(1\right)$. We call this kind of symetry as $U\left(1\right)$-gauge symmetry.

Remark. We can generalize this concept to various groups. In the case of $SU\left(3\right)$, it corresponds to the strong force. In the case of $SU\left(2\right)$, it corresponds to weak force. In the case of diffeomorphism, it corresponds to general relativity.

$SU\left(3\right) \times SU\left(2\right) \times U\left(1\right)$: Gauge symmetry of the standard model in particle physics, which is quite accurate model in physics in theoratical way and experimental way. It is accepted as a theory in nowadays, not just a model.

Remark. They are two kind of feature in quantum mechanics.

Schrödinger picture: $|E\left(t\right)>$, $\hat{p}$, $\hat{x}$, i.e., time variance on state, but operator does not change in time.

Heisenberg picture: $\mid E >$, $\hat{p}(t), \hat{x}(t)$, i.e., operator change in time but the state does not depend on time.

$$\begin{aligned}\left\langle \Psi \mid \hat{A} \mid E \right\rangle &= \left\langle E(t) \mid \hat{A} \mid E(t) \right\rangle \\ &= \left\langle \Psi \mid e^{i\hbar \hat{H}(t)} \hat{A} e^{-i\hbar \hat{H}(t)} \mid \Psi \right\rangle \\ &= \left\langle \Psi \mid \hat{A}(t) \mid \Psi \right\rangle\end{aligned}$$

Hence in the sense of Heisenberg picture, we have gauge symmetry also.

The gauge symmetry is a principle in physics in nowadays although it is not physical.

Note that $U(1)$ is abelian, however, $SU(2)$, $SU(3)$, diffeomorphism are not abelian. Actually,

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

in general.

We will figure out why this is natural. Consider a Gauge group $\mathbb{G}$ (ex. $SU(3)$, $SU(2)$, $U(1)$, etc)[2] and transform

$$\psi(t, x, y, z) \mapsto \psi' = g\psi, \quad g \in \mathbb{G}$$

with $D_\mu \psi := \partial_\mu \psi + A_\mu \psi$. Note that A_μ is in general matrix. We require

$$D'_\mu \psi' = \partial_\mu \psi' + A'_\mu \psi' = g D_\mu \psi.$$

Then by computation

$$\begin{aligned}D'_\mu \psi' &= \partial_\mu \psi' + A'_\mu \psi' \\ &= \partial_\mu g \psi + g \partial_\mu \psi + A'_\mu g \psi \\ &= g \left[\partial_\mu \psi + \left(g^{-1} A'_\mu g + g^{-1} \partial_\mu g\right) \psi\right].\end{aligned}$$

So we require

$$A_\mu = g^{-1} A'_\mu g + \partial^{-1} \partial_\mu g.$$

So

$$A'_\mu = g\left(A_\mu - g^{-1} \partial_\mu g\right) g^{-1}.$$

Hence the transform must satisfy

$$A_\mu \mapsto g\left(A_\mu - g^{-1} \partial_\mu g\right) g^{-1}.$$

Under this transformation, it is easy to check by computation

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \mapsto F'_{\mu\nu} = g F_{\mu\nu} g^{-1}.$$

So

$$F_{\mu\nu} \psi \mapsto g F_{\mu\nu} g^{-1} g \psi = g F_{\mu\nu} \psi.$$

[2]In mathematical language, it is Lie group,

$$g(\theta^a) = \exp\left[\sum_a \theta^a T_a\right], \quad \{T_a\} : \text{Lie algebra},$$

We say it has a Gauge symmetry if $\theta^a(t, x, y, z)$: arbitrary / local parameter.

Note that $D_\mu\psi \mapsto gD_\mu\psi$. So $D_\mu D_\nu\psi \mapsto D'_\mu D'_\nu = gD_\mu D_\nu\psi$. Note that

$$\begin{aligned} D_\mu D_\nu\psi &= \partial_\mu (D_\nu\psi) + A_\mu D_\nu\psi \\ &= \partial_\mu (\partial_\nu\psi + A_\nu\psi) + A_\mu (\partial_\nu\psi + A_\nu\psi) \\ &= \partial_\mu\partial_\nu\psi + \partial_\mu A_\nu\psi + A_\nu\partial_\mu\psi + A_\mu\partial_\nu\psi + A_\mu A_\nu\psi. \end{aligned}$$

So

$$[D_\mu, D_\nu]\psi = (\partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu)\psi = F_{\mu\nu}\psi.$$

Proposition 1.8. If A, B are in Lie algebra, then $e^A B e^{-A}$ is in Lie algebra.

Proof. We claim that

$$e^{\mathrm{ad}A} B = e^A B e^{-A},$$

where

$$e^{\mathrm{ad}A} B = e^{[A,\cdot]} B = B + [A,B] + \frac{1}{2}[A,[A,B]] + \cdots + \frac{1}{n!}[A[A,\cdots,[A,B]]] + \cdots.$$

Here

$$\frac{1}{n!}(\mathrm{ad}A)^n B = \frac{1}{n!}[A[A,\cdots,[A,B]]].$$

Since the Lie algebra is closed under commutator, we are done. Hence it suffices to show the claim.

Note that

$$\frac{d}{dt}\left(e^{tA} B e^{-tA}\right) = \left[A, e^{tA} B e^{-tA}\right]$$

and

$$\frac{d}{dt}\left(e^{t\mathrm{ad}A} B\right) = \left[A, e^{t\mathrm{ad}A} B\right].$$

Hence by the uniqueness of initial value problem, we are done. □

For $g \in \mathbb{G}$, write

$$g(\theta) = e^{\theta^a T_a}.$$

Note that

$$\begin{aligned} g(\theta_1) g(\theta_2) &= e^{\theta_1^a T_a} e^{\theta_2^b T_b} = e^{f^a(\theta_1,\theta_2) T_a} \\ &= g(f(\theta_1,\theta_2)). \end{aligned}$$

Here f^a satisfies $f^a(\theta,0) = f^a(0,\theta) = \theta^a$. Since $g(-\theta) = e^{-\theta^a T_a} = g(\theta)^{-1}$,

$$f^a(\theta,-\theta) = f^a(-\theta,\theta) = 0.$$

Then

$$\begin{aligned} \partial_\mu g g^{-1} &= \partial_\mu g(\theta) g(-\theta) \\ &= \partial_\mu g(\theta) g(-\phi)\big|_{\phi=\theta} \\ &= \partial_\mu [g(\theta) g(-\phi)]\big|_{\phi=\theta} \\ &= \partial_\mu g(f(\theta,-\phi))\big|_{\phi=\theta} \\ &= \partial_\mu f^a(\theta,-\phi) \frac{\partial}{\partial f^a} g(f)\bigg|_{\phi=\theta} \\ &= \partial_\mu f^a(\theta,-\phi)\big|_{\phi=\theta} \frac{\partial}{\partial f^a} g(f)\bigg|_{f=0} \\ &= \partial_\mu f^a(\theta,-\phi)\big|_{\phi=\theta} T_a \in \text{Lie algebra}. \end{aligned}$$

So

$$\frac{d}{d\theta^a} g(\theta) \mid_{\theta=0} = T_a.$$

From this, $A_\mu \mapsto A'_\mu \in \mathbb{G}$. So by previous proposition, $F_{\mu\nu} \in \mathbb{G}$. Hence we can write

$$A_\mu = \sum_a A_\mu^a T_a.$$

Remark. Note that

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

is Gauge invariant. However,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

is Gauge covariant. Although it is in some sense is good, it is not observable by experiment. However,

$$\text{Tr}\left(F_{\mu\nu}\right), \quad \text{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right)$$

are Gauge invariant.

Yang-Mills first gives Gauge symmetry argument in the theory of standard model of particle physics.

In the case of free particle, we write a Schrödinger equation

$$i\hbar\left(\partial_t - iA_0\right)\psi = \frac{1}{2m}\left(-i\hbar\left(\nabla - i\vec{A}\right)\right)^2\psi,$$

under Gauge transform consideration. We rewrite

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\left(\nabla - i\vec{A}\right)^2\psi - A_0\psi.$$

Here A_0 is a coulomb potential. This is so-called the magnetic Schrödinger equation.

General Relativity 2

2.1 The geodesic equation

In this chapter, we introduce the general relativity which is the Einstein's theory of gravity. In the previous chapter,

$$d\theta^2 = dx^\mu dx^\nu \eta_{\mu\nu}$$
$$\eta_{\mu\nu} = \mathrm{diag}\,(-+++)\,.$$

Recall that freely falling frame is locally inertial frame since it has a tidal force effect. In this setting, gravity disappears. Hence the special relativity works in this setting.

Let x^μ denote a general coordinate system and consider y^μ a coordinates of the locally inertial frame at $x^\mu = X^\mu$. Then there is a local coordinate transform $x^\mu(y)$, $y^\mu(x)$, e.g., $x^\mu(y \equiv 0) = X^\mu$. Then in the near $y = 0$,

$$ds^2 = \eta_{\mu\nu} dy^\mu dy^\nu.$$

By considering the transformation, we have

$$\begin{aligned} ds^2 &= \eta_{\mu\nu} \left.\frac{\partial y^\mu}{\partial x^\lambda}\right|_{x=X} dx^\lambda \left.\frac{\partial y^\nu}{\partial x^\rho}\right|_{x=X} dx^\rho \\ &= \left(\eta_{\mu\nu} \left.\frac{\partial y^\mu}{\partial x^\lambda}\right|_{x=X} \left.\frac{\partial y^\nu}{\partial x^\rho}\right|_{x=X}\right) dx^\lambda dx^\rho \\ &= g_{\mu\nu}(X)\, dx^\mu dx^\nu. \end{aligned}$$

Here

$$g_{\mu\nu}(x) = \eta_{ab} \left.\frac{\partial y^a}{\partial x^\lambda} \frac{\partial y^b}{\partial x^\rho}\right|_{y=0}.$$

Hence to describe the physics, we need to consider the above metric. Hence

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu dx^\nu.$$

Note that $g_{\mu\nu}$ is metric and this is the only geometric quantity. This describes the gravitational effect in physics. Our aim is to study the PDE which gives a solution $g_{\mu\nu}$, so called the Einstein field equation.

If we consider a $\mathbb{R}^2$ which is in the standard coordinate system and polar coordinate system, then one metric is constant but the other has variable. If one can transform the metric into Lorentz metric, then the surface is flat. However, it is not. We will figure out in detail.

Under coordinate transformations, $y^\mu \to x^\mu(y)$, the velocity of a particle, $\dot{y}^\mu = \frac{\mathrm{d}y^\mu}{\mathrm{d}\tau}$, transforms as

$$\dot{y}^\mu \quad \Longrightarrow \quad \dot{x}^\mu = \frac{\partial x^\mu}{\partial y^\nu}\,\dot{y}^\nu\,,$$

where τ denotes the proper time.

A chain rule gives

$$\ddot{x}^\mu = \frac{\partial x^\mu}{\partial y^\nu}\ddot{y}^\nu + \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial x^\mu}{\partial y^\nu}\right)\dot{y}^\nu.$$

Changing the role of x and y, we have

$$\ddot{y}^\mu = \frac{\partial y^\mu}{\partial x^\sigma}\ddot{x}^\sigma + \frac{\partial^2 y^\mu}{\partial x^\sigma \partial x^\nu}\dot{x}^\nu \dot{x}^\sigma.$$

Now multiplying $\frac{\partial x^\lambda}{\partial y^\mu}$ to the both sides, we get

$$\begin{aligned}
&\frac{\partial x^\lambda}{\partial y^\mu}\ddot{y}^\mu \\
&= \frac{\partial x^\lambda}{\partial y^\mu}\frac{\partial y^\mu}{\partial x^\sigma}\ddot{x}^\sigma + \frac{\partial x^\lambda}{\partial y^\mu}\frac{\partial^2 y^\mu}{\partial x^\sigma \partial x^\nu}\dot{x}^\nu \dot{x}^\sigma . \\
&= \delta^\lambda_\sigma \ddot{x}^\sigma + \frac{\partial x^\lambda}{\partial y^\mu}\frac{\partial^2 y^\mu}{\partial x^\sigma \partial x^\nu}\dot{x}^\nu \dot{x}^\sigma \\
&= \ddot{x}^\lambda + \frac{\partial x^\lambda}{\partial y^\mu}\frac{\partial^2 y^\mu}{\partial x^\sigma \partial x^\nu}\dot{x}^\nu \dot{x}^\sigma .
\end{aligned}$$

Hence we see that under this coordinate transformation, the acceleration becomes

$$\frac{\partial x^\mu}{\partial y^\nu}\,\ddot{y}^\nu = \ddot{x}^\mu + \frac{\partial x^\mu}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\alpha \partial x^\beta}\,\dot{x}^\alpha \dot{x}^\beta .$$

From the invariance of the proper length,

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu = \tilde{g}_{\mu\nu}(y)dy^\mu dy^\nu ,$$

the metric transforms as

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

We may derive straightforwardly

$$\partial_\lambda g_{\mu\nu} = \tilde{\partial}_\alpha \tilde{g}_{\rho\sigma}\partial_\lambda y^\alpha \partial_\mu y^\rho \partial_\nu y^\sigma + \tilde{g}_{\rho\sigma}\partial_\lambda\partial_\mu y^\rho \partial_\nu y^\sigma + \tilde{g}_{\rho\sigma}\partial_\mu y^\rho \partial_\lambda\partial_\nu y^\sigma ,$$

where we put $\tilde{\partial}_\alpha = \frac{\partial}{\partial y^\alpha}$.

Exercise 2.1. Using the above result, derive

$$\begin{aligned}
&\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu} \\
&= \left(\tilde{\partial}_\alpha \tilde{g}_{\gamma\beta} + \tilde{\partial}_\beta \tilde{g}_{\alpha\gamma} - \tilde{\partial}_\gamma \tilde{g}_{\alpha\beta}\right)\partial_\lambda y^\gamma \partial_\mu y^\alpha \partial_\nu y^\beta + 2\tilde{g}_{\rho\sigma}\partial_\mu\partial_\nu y^\rho\,\partial_\lambda y^\sigma .
\end{aligned}$$

For convenience, we define the *Christoffel symbol*

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

Then we get the following transformation rule for Christoffel symbol.

$$\Gamma^\lambda_{\mu\nu} = \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial y^\beta}{\partial x^\nu}\frac{\partial x^\lambda}{\partial y^\gamma}\tilde{\Gamma}^\gamma_{\alpha\beta} + \frac{\partial x^\lambda}{\partial y^\sigma}\frac{\partial^2 y^\sigma}{\partial x^\mu \partial x^\nu} . \tag{2.1}$$

Exercise 2.2. Prove the identity (2.1).

Combining the above results, we can show the ‘covariance’:

$$\frac{\partial x^\mu}{\partial y^\nu}\left(\ddot{y}^\nu + \tilde{\Gamma}^\nu_{\rho\sigma}\dot{y}^\rho \dot{y}^\sigma\right) = \ddot{x}^\mu + \Gamma^\mu_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma .$$

In local frame, when it is in free fall, the object moves in linear motion. So $\ddot{y}^a = 0$. From the above equation, we have

$$\ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0. \tag{2.2}$$

We call this equation as *Geodesic equation* . We derive the geodesic equation in geometrically. Recall

$$d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu.$$

We consider the relativisitic point particle action:

$$\mathcal{S} = \int_{x_i}^{x_f} d\tau = \int d\sigma \sqrt{-g_{\mu\nu}\left(x\right)\dot{x}^\mu \dot{x}^\nu}.$$

Here

$$\dot{x}^\mu = \frac{d}{d\sigma} x^\mu.$$

Then

$$\mathcal{S} = \int d\sigma \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}}.$$

Note that the above action is invariant $x^\mu\left(\sigma\right) \mapsto x^\mu\left(\sigma'\right)$. So it has a local symmetry. Hence we can choose $\sigma \equiv \tau$ to $\sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} = 1$.

Therefore, by the least action principle, we deduce the Euler-Lagrange equation. From this, we will derive the geodesic equation.

Taking variation, we have

$$\begin{aligned}
0 = \delta\mathcal{S} &= \int d\tau \delta \sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} \\
&= \int d\tau \frac{1}{2}\left(\frac{\delta\left(-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right)}{\sqrt{-g_{\lambda\rho}\dot{x}^\lambda\dot{x}^\rho}}\right) \\
&= -\int d\tau \frac{\delta x^\lambda \partial_\lambda g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu + g_{\mu\nu}\delta\dot{x}^\mu\dot{x}^\nu + g_{\mu\nu}\dot{x}^\mu\delta\dot{x}^\nu}{2\sqrt{-g_{\lambda\rho}\dot{x}^\lambda\dot{x}^\rho}} \\
&= \int d\tau \frac{\delta x^\lambda\left(-\partial_\lambda g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right)}{2\sqrt{-g_{\lambda\rho}\dot{x}^\lambda\dot{x}^\rho}} + \delta x^\mu \frac{d}{d\tau}\left[\frac{g_{\mu\nu}\dot{x}^\nu}{\sqrt{-g_{\lambda\rho}\dot{x}^\lambda\dot{x}^\rho}}\right] \\
&= \int d\tau - \frac{1}{2}\delta x^\lambda\left(\partial_\lambda g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right) + \delta x^\mu\left[\dot{x}^\lambda\partial_\lambda g_{\mu\nu}\dot{x}^\nu + g_{\mu\nu}\ddot{x}^\nu\right] \\
&= \int d\tau \delta x^\sigma\left[g_{\sigma\lambda}\ddot{x}^\lambda + \left(\partial_\mu g_{\nu\sigma} - \frac{1}{2}\partial_\sigma g_{\mu\nu}\right)\dot{x}^\mu\dot{x}^\nu\right] \\
&= \int d\tau dx^\rho g_{\rho\lambda}\left[\ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right].
\end{aligned}$$

Hence we obtain the geodesic equation

$$\ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0,$$

which we already derived in (2.2).

On the other hand,

$$\mathcal{S} = \int d\tau \sqrt{-\eta_{ab}\dot{y}^a\dot{y}^b}.$$

The equation describes linear motion on a manifold. Since g is metric, it has geometrical meaning. Hence Christoffel symbol. Here note that the equation has no mass part.

Write $x^\mu = \left(x^0 = ct, x^i\right)$. Then $\frac{\partial}{\partial x^0} = \frac{\partial}{c\partial t}$. For large c and slowing moving paricle, we have $\frac{\partial}{\partial x^0} \approx 0$

$$\dot{x}^\mu = \left(\frac{cdt}{d\tau}, \frac{dx^i}{d\tau}\right) \approx \left(c, 0\right).$$

From this assumption, we get

$$\partial_\mu = (0, \partial_i)$$

and so

$$\begin{aligned}\Gamma^\lambda_{\mu\nu}\dot{x}^\mu\dot{x}^\nu &\approx \Gamma^\lambda_{00}c^2 \\ &= \frac{1}{2}c^2 g^{\lambda\rho}\left(\partial_0 g_{\rho 0} + \partial_0 g_{0\rho} - \partial_\rho g_{00}\right) \\ &= -\frac{1}{2}c^2 g^{\lambda i}\partial_i g_{00}.\end{aligned}$$

Hence

$$\ddot{x}^\lambda \approx \frac{1}{2}c^2 g^{\lambda i}\partial_i g_{00}.$$

If we have a weak gravity, it goes almostly flat. Hence in the weak gravity limit, we have

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}.$$

So

$$g^{\mu\nu} \approx \eta^{\mu\nu} - \eta^{\mu\lambda}h_{\lambda\rho}\eta^{\rho\nu}$$

up to second-order approximation.

From this, we see

$$\begin{aligned}\ddot{x}^\lambda &\approx \frac{1}{2}c^2\left(\eta^{\mu i} - \eta^{\mu\rho}h_{\rho\sigma}\eta^{\sigma i}\right)\partial_i\left(\eta_{00} + h_{00}\right) \\ &\approx \frac{1}{2}c^2\eta^{\lambda i}\partial_i h_{00}.\end{aligned}$$

In spartial part,

$$\ddot{x}^i \approx \frac{1}{2}c^2\partial_i h_{00}.$$

Recall the theory of Newton:

$$m\ddot{x}^i \equiv -\partial_i U_{\text{newton}},$$

So

$$h_{00} \approx -\frac{2}{c^2}\frac{U_{\text{newton}}}{m}.$$

In particular, when we have a spherical symmetry, the potential is given by

$$U_{\text{newton}}(r) = -\frac{mMG}{r}.$$

So we get

$$h_{00} \approx \frac{2MG}{c^2 r}.$$

Hence

$$g_{00} \approx -1 + \frac{2MG/c^2}{r^2}.$$

Therefore the theory of gravity by newton is a special case of Einstein's. Actually, the approximation is equal. This was proved by Schwarzschild.

All physical object must obey the geodesic equation. Light also obey the equation. From this, Einstein predicts the light should band. Eddington examine the Einstein's prediction by observing total solar eclipse.

If x is a solution of geodesic equation, then

$$\begin{aligned}
&\frac{d}{d\tau}\left[\dot{x}^{\mu}\dot{x}^{\nu}g_{\mu\nu}\left(x\right)\right] \\
&= \left(\ddot{x}^{\mu}\dot{x}^{\nu}+\dot{x}^{\mu}\ddot{x}^{\nu}\right)g_{\mu\nu}+\dot{x}^{\mu}\dot{x}^{\nu}\dot{x}^{\lambda}\partial_{\lambda}g_{\mu\nu} \\
&= 2\ddot{x}^{\mu}\dot{x}^{\nu}g_{\mu\nu}+\dot{x}^{\mu}\dot{x}^{\nu}\dot{x}^{\lambda}\partial_{\lambda}g_{\mu\nu} \\
&= -2g^{\mu\rho}\left(\partial_{\kappa}g_{\rho\sigma}+\partial_{\sigma}g_{\kappa\rho}-\partial_{\rho}g_{\kappa\rho}\right)\dot{x}^{\kappa}\dot{x}^{\sigma}\dot{x}^{\nu}g_{\mu\nu}+\dot{x}^{\mu}\dot{x}^{\nu}\dot{x}^{\lambda}\partial_{\lambda}g_{\mu\nu} \\
&= -\dot{x}^{\kappa}\dot{x}^{\sigma}\dot{x}^{\rho}\left(\partial_{\kappa}g_{\rho\sigma}+\partial_{\sigma}g_{\kappa\rho}-\partial_{\rho}g_{\kappa\rho}-\partial_{\kappa}g_{\sigma\rho}\right) \\
&= 0.
\end{aligned}$$

From this, $\dot{x}^{\mu}\dot{x}^{\nu}g_{\mu\nu}$ is constant with respect to τ. There are three cases. $>$, $=$, $<$. We say $\dot{x}^{\mu}$ is *space-like* if the quantity is strictly positive. Similarly, $\dot{x}^{\mu}$ is *null, time-like* if the quantity is $=0$, strictly negative, respectively.

We can normalize so that

$$\begin{aligned}
\dot{x}^{\mu}\dot{x}^{\nu}g_{\mu\nu} &= +1 \quad \text{if it is space-like} \\
\dot{x}^{\mu}\dot{x}^{\nu}g_{\mu\nu} &= -1 \quad \text{if it is time-like.}
\end{aligned}$$

In space-like, $ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$. In time-like, $-c^2 d\tau^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$. In the case of null, we do not call τ as a proper-time rather than affine parameter.

For ordinary massive particle, it is time-like. For massless particle, or photon, it is null. In the space-like, there is no such particle since the particle must exceed the speed of light. But we call it just Tachyon.

Write $m\dot{x}^{\mu} = p^{\mu}$. In the time-like, we have

$$p^{\mu}p^{\nu}g_{\mu\nu} = -m^2c^2.$$

From now on, we will derive an equation for $g_{\rho\nu}$, the Einstein field equation.

2.2 Tensor and covariant derivatives

Consider the diffeomorphism

$$\begin{aligned}
x^{\mu} &\mapsto x'^{\mu}\left(x\right) \\
x'^{\mu} &\to x^{\mu}\left(x'\right).
\end{aligned}$$

Then

$$\begin{aligned}
dx^{\mu} \mapsto & dx'^{\mu} = dx^{\nu}\frac{\partial x'^{\mu}}{\partial x^{\nu}} \\
\partial_{\mu} \mapsto & \frac{\partial x^{\nu}}{\partial x'^{\mu}}\partial_{\nu}.
\end{aligned}$$

Note that ds^2 must be invariant under diffeomorphism. From this assumption, we have

$$g_{\mu\nu}dx^{\mu}dx^{\nu} = g'_{\mu\nu}\frac{\partial x'^{\mu}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}dx^{\rho}dx^{\sigma}.$$

So

$$g_{\mu\nu} \mapsto g'_{\mu\nu}\left(x'\right) = \frac{\partial x^{\rho}}{\partial x'^{\mu}}\frac{\partial x^{\rho}}{\partial x'^{\nu}}g_{\sigma\rho}\left(x\right).$$

Definition 2.3. We define (p,q)-Tensor$(p,q=0,1,2,\dots)$ T if it satisfies

$$T^{\mu_1\mu_2\cdots\mu_p}{}_{\nu_1\nu_2\cdots\nu_q}(x)\mapsto\frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}}\cdots\frac{\partial x'^{\mu_p}}{\partial x^{\rho_p}}\frac{\partial x'^{\sigma_1}}{\partial x^{\nu_1}}\cdots\frac{\partial x'^{\sigma_q}}{\partial x^{\nu_q}}T^{\rho_1\cdots\rho_p}{}_{\sigma_1\cdots\sigma_q}(x)$$

under diffeomorphism. We call this invariance as the *covariant transformation rule*.

Remark. All the physics laws must be expressed / expressible by tensors. This ensures that physics law is invariant under coordinate transformations.

Example 2.4. dx^μ, ∂_μ , $g_{\mu\nu}$ are tensor. But x^μ is not a tensor in general.

Example 2.5. Scalar is $(0,0)$-tensor: $\phi(x)\mapsto\phi'(x')=\phi(x)$.
Vector is $(1,0)$-tensor: $v^\mu(x)\mapsto v'^\mu(x')=\frac{\partial x'^\mu}{\partial x^\nu}v^\nu(x)$
Contravariant vector is$(0,1)$-tensor: $w_\mu(x)\mapsto w'_\mu(x')=\frac{\partial x^\nu}{\partial x'^\mu}w_\nu(x)$
Metric is $(0,2)$-tensor: $g_{\mu\nu}(x)\mapsto g'_{\mu\nu}(x')=\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu}g_{\lambda\rho}(x)$.

Physics law requires differentiation. Let us observe something.
If ϕ is scalar, i.e., $(0,0)$-tensor, then

$$\partial_\mu\phi(x)\mapsto\partial'_\mu\phi'(x')=\frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu\phi(x)\,.$$

So the dervative of scalar is $(0,1)$-tensor.
If v is vector, i.e., $(1,0)$-tensor, then

$$\begin{aligned}\partial_\mu v^\nu(x)\mapsto\partial'_\mu v'^\nu(x')&=\frac{\partial x^\lambda}{\partial x'^\mu}\partial_\lambda\left[\frac{\partial x'^\nu}{\partial x^\rho}v^\rho(x)\right]\\&=\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\rho}\partial_\lambda v^\rho(x)+\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial^2x'^\nu}{\partial x^\lambda\partial x^\rho}v^\rho(x)\,.\end{aligned}$$

Note that

$$\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\rho}\partial_\lambda v^\rho(x)$$

is $(1,1)$-tensor but

$$\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial^2x'^\nu}{\partial x^\lambda\partial x^\rho}v^\rho(x)$$

is not.
Similarly, if A_ν is a contravariant vector, then

$$\begin{aligned}\partial_\mu A_\nu(x)\mapsto\partial'_\mu A'_\nu(x')&=\frac{\partial x^\lambda}{\partial x'^\mu}\partial_\lambda\left[\frac{\partial x^\rho}{\partial x'^\nu}A_\rho(x)\right]\\&=\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\lambda}{\partial x'^\nu}\partial_\lambda A_\rho+\frac{\partial^2x^\rho}{\partial x'^\mu\partial x'^\nu}A_\rho(x)\,.\end{aligned}$$

So first part is tensor but the second one is not. In all, the usual derivative of (p,q)-tensor is not tensor except $(p,q)=(0,0)$.
To eliminate anomalous term, we define ∇_σ by

$$\begin{aligned}\nabla_\sigma T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}(x)&=\partial_\sigma T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}(x)\\&\quad+\sum_{i=1}^p\Gamma^{\mu_i}_{\sigma\,\rho}T^{\mu_1\cdots\mu_{i-1}\rho\mu_{i+1}\cdots\mu_p}-\sum_{j=1}^qT^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_{j-1}\rho\nu_{j+1}\cdots\nu_q}\Gamma^\rho_{\sigma\,\nu_j}\,.\end{aligned}$$

Note that this derivative ∇_σ satisfies the Leibniz rule:

$$\nabla_\sigma \left(V^\mu W_\mu\right) = \left(\nabla_\sigma V^\mu\right) W_\mu + V^\mu \left(\nabla_\sigma W_\mu\right).$$

Also

$$\nabla_\sigma \delta^\nu_\mu = 0.$$

Indeed,

$$\begin{aligned} \nabla_\sigma V^\mu &= \partial_\sigma V^\mu + \Gamma^{\ \mu}_{\sigma\,\rho} V^\rho \\ \nabla_\sigma W_\mu &= \partial_\sigma W_\mu - W_\rho \Gamma^{\ \rho}_{\sigma\,\mu} \end{aligned}$$

implies

$$\begin{aligned} &\left(\nabla_\sigma V^\mu\right) W_\mu + V^\mu \left(\nabla_\sigma W_\mu\right) \\ &= \partial_\sigma \left(V^\mu W_\mu\right) \\ &= \nabla_\sigma \left(V^\mu W_\mu\right). \end{aligned}$$

Also

$$\begin{aligned} \nabla_\sigma \delta^\mu_\nu &= \partial_\sigma \delta^\mu_\nu + \Gamma^{\ \mu}_{\sigma\,\rho} \delta^\rho_\nu - \delta^\mu_\rho \Gamma^{\ \rho}_{\sigma\,\nu} \\ &= 0 + \Gamma^{\ \mu}_{\sigma\,\nu} - \Gamma^{\ \mu}_{\sigma\,\nu} = 0. \end{aligned}$$

∇_σ should send tensor to tensor under diffeomorphism. To find the condition, we calculate

$$\begin{aligned} \nabla_\sigma T^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}(x) &= \nabla'_\sigma \left(T'\right)^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} \\ &= \frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\mu_p}}{\partial x^{\lambda_\rho}} \frac{\partial x^\tau}{\partial x'^\sigma} \frac{\partial x^{\sigma_1}}{\partial x'^{\upsilon_q}} \\ &\quad \times \left[\partial_\tau T^{\lambda_1\cdots\lambda_\rho}{}_{\kappa_1\cdots\kappa_q} + \sum_i \Gamma^{\lambda_i}_{\tau\rho} T^{\cdots\rho\cdots}{}_{\kappa_1\cdots\kappa_q} - \sum_j T^{\lambda_1\cdots\lambda_p}{}_{\cdots\rho\cdots} \Gamma^{\ \rho}_{\tau\,\kappa_j}\right] \\ &= \frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\mu_p}}{\partial x^{\lambda_\rho}} \frac{\partial x^\tau}{\partial x'^\sigma} \frac{\partial x^{\sigma_1}}{\partial x'^{\upsilon_q}} \left[\partial_\tau T^{\lambda_1\cdots\lambda_\rho}{}_{\kappa_1\cdots\kappa_q}\right] \\ &\quad + \frac{\partial}{\partial x'^\sigma} \left[\frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\mu_p}}{\partial x^{\lambda_\rho}} \frac{\partial x^\tau}{\partial x'^\sigma} \frac{\partial x^{\sigma_1}}{\partial x'^{\upsilon_q}}\right] T^{\lambda_1\cdots\lambda_p}{}_{\kappa_1\cdots\kappa_q} \\ &\quad + \sum_i \left(\Gamma'\right)^{\mu_i}_{\sigma\,\rho} \left(T'\right)^{\cdots\rho\cdots}{}_{\nu_1\cdots\nu_q} - \sum_j \left(T'\right)^{\mu_1\cdots\mu_p}{}_{\cdots\rho\cdots} \left(\Gamma'\right)^{\rho}_{\sigma\,\nu_j}. \end{aligned}$$

Write

$$\left(\Gamma'\right)^\mu_{\lambda\nu}\left(x'\right) = \frac{\partial x'^\mu}{\partial x^\kappa} \frac{\partial x^\rho}{\partial x'^\nu} \Gamma^\kappa_{\rho\sigma}(x) + \Delta^\mu_{\lambda\nu}.$$

Then

$$\begin{aligned} 0 &= \nabla'_\lambda T'^{\mu_1\cdots}{}_{\nu_1\cdots} - \frac{\partial x^\alpha}{\partial x'^\lambda} \frac{\partial x'^{\mu_1}}{\partial x^{\beta_1}} \cdots \frac{\partial x^{\gamma_1}}{\partial x'^{\nu_1}} \cdots \nabla_\alpha T^{\beta_1\cdots}{}_{\gamma_1\cdots} \\ &= \sum_i \left[\left[\frac{\partial}{\partial x'^\lambda}\left(\frac{\partial x'^{\mu_i}}{\partial x^{\lambda_i}}\right)\right] \frac{\partial x^{\beta_i}}{\partial x'^\rho} \cdots + \Delta^{\mu_i}_{\lambda\,\rho}\right] T'^{\cdots\rho}{}_{\nu_1} + \sum_j \left[\frac{\partial^2 x^{\gamma_j}}{\partial x'^\lambda \partial x'^{\nu_j}} \frac{\partial x'^\rho}{\partial x^{\gamma_j}} - \Delta^{\rho}_{\lambda\,\nu_j}\right] T'^{\mu_1\cdots}{}_{\cdots\rho\cdots}\cdots \end{aligned}$$

Now

$$\begin{aligned}
\Delta^{\rho}_{\lambda\nu} &\equiv \frac{\partial^2 x^\gamma}{\partial x'^\lambda \partial x'^\mu}\frac{\partial x'^\rho}{\partial x^\gamma} \\
\Delta^{\mu}_{\lambda\rho} &\equiv -\frac{\partial}{\partial x'^\lambda}\left(\frac{\partial x'^\mu}{\partial x^\beta}\right)\frac{\partial x^\beta}{\partial x'^\rho} \\
&\equiv -\frac{\partial}{\partial x'}\left(\frac{\partial x'^\mu}{\partial x^\beta}\frac{\partial x^\beta}{\partial x'^\rho}\right) + \frac{\partial x'^\mu}{\partial x^\beta}\frac{\partial^2 x^\beta}{\partial x'^\lambda \partial x'^\rho} \\
&= \frac{\partial^2 x^\beta}{\partial x'^\lambda \partial x'^\rho}\frac{\partial x'^\mu}{\partial x^\beta}
\end{aligned}$$

since $\frac{\partial x'^\mu}{\partial x^\beta}\frac{\partial x^\beta}{\partial x'^\rho} = \delta^\mu_\rho$ and $\delta M^{-1} = -M^{-1}\delta M M^{-1}$(from $MM^{-1} = I$, $\delta\left(MM^{-1}\right) = 0$.)

Hence we obtain the desired covariant derivative $\nabla_\mu = \partial_\mu + \Gamma_\mu$, where

$$\begin{aligned}
T^{\mu_1\cdots}{}_{\nu_1\cdots}(x) \mapsto & T'^{\mu_1\cdots}{}_{\nu_1\cdots}(x') = \frac{\partial x'^{\mu_1}}{\partial x^{\alpha_1}}\cdots\frac{\partial x^{\beta_1}}{\partial x'^{\nu_1}}\cdots T^{\alpha_1\cdots}{}_{\beta_1\cdots}(x) \\
\Gamma^\mu_{\lambda\,\nu}(x) \mapsto & \Gamma'^\mu_{\lambda\nu}(x') = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\nu}\Gamma^\alpha_{\beta\gamma}(x) + \frac{\partial^2 x^\rho}{\partial x'^\lambda \partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho} \\
\nabla_\lambda T^{\mu_1\cdots}{}_{\nu_1\cdots}(x) \mapsto & \nabla'_\lambda T'^{\mu_1\cdots}{}_{\nu_1\cdots}(x') = \frac{\partial x^\alpha}{\partial x'^\lambda}\frac{\partial x'^{\mu_1}}{\partial x^{\beta_1}}\cdots\frac{\partial x^{\gamma_1}}{\partial x'^{\nu_1}}\nabla_\alpha T^{\beta_1\cdots}{}_{\gamma_1\cdots}(x)\,.
\end{aligned}$$

So

$$\nabla_\lambda T^{\mu_1\cdots}{}_{\nu_1} := \partial_\lambda T^{\mu_1\cdots}{}_{\nu_1\cdots} + \sum_i \Gamma^{\mu_i}_{\lambda\,\rho} T^{\cdots\rho\cdots}{}_{\nu_1\cdots} - \sum_j T^{\mu_1\cdots}{}_{\cdots\rho\cdots}\Gamma^\rho_{\lambda\nu_j}.$$

One can check the derivative behaves covariantly.

Example 2.6. Recall the Maxwell equation. We write

$$\begin{aligned}
\partial_\lambda F^{\lambda\mu} &= J^\mu \\
\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} &= 0. \qquad (2.3)
\end{aligned}$$

In fact, we should write

$$\begin{aligned}
\nabla_\lambda F^{\lambda\mu} &= J^\mu \qquad (2.4) \\
\nabla_\lambda F_{\mu\nu} + \nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} &= 0.
\end{aligned}$$

to be invariant under coordinate transform. Note that (2.3) is a special case of (2.4). In fact, if $\Gamma^\nu_{\lambda\mu} \equiv 0$(flat spacetime), then (2.4) becomes (2.3).

So we want to find the condition on flatness. In free falling, the metric $g_{\mu\nu}$ acting on a particle satisfies

$$\begin{aligned}
g_{\mu\nu}\,|_{\text{origin}} &\to \eta_{\mu\nu} \\
\partial_\lambda g_{\mu\nu}\,|_{\text{origin}} &= 0.
\end{aligned}$$

So by considering the Taylor expansion, we have

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \frac{1}{2}x^\lambda x^\rho \partial_\lambda \partial_\rho g_{\mu\nu}(x) + \cdots$$

If $\partial_\lambda g_{\mu\nu}\,|_{\text{origin}} = 0$ is of physics law, this should be $\nabla_\lambda g_{\mu\nu}(x) = 0$. We call this as the metric is *covariantly constant* (metric compatiblity condition). So the one of the axiom of the Einstein theory is

- Metric is covariantly constant, i.e., $\nabla_\lambda g_{\mu\nu}(x) = 0$.

From this,

$$0 = \nabla_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - g_{\rho\nu}\Gamma^\rho_{\lambda\mu} - g_{\mu\rho}\Gamma^\rho_{\lambda\nu}.$$

Note that

$$\begin{aligned}
\Gamma^\mu_{\lambda\nu} - \Gamma^\mu_{\nu\lambda} \mapsto & \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\nu}\Gamma^\alpha_{\beta\gamma}(x) + \frac{\partial^2 x^\rho}{\partial x'^\lambda \partial x'^\nu}\frac{\partial x'^\mu}{\partial x^\rho} - \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial x^\gamma}{\partial x'^\lambda}\Gamma^\alpha_{\beta\gamma}(x) - \frac{\partial^2 x^\rho}{\partial x'^\nu \partial x'^\lambda}\frac{\partial x'^\mu}{\partial x^\rho} \\
&= \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\nu}\Gamma^\alpha_{\beta\gamma}(x) - \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial x^\gamma}{\partial x'^\lambda}\Gamma^\alpha_{\beta\gamma}(x) \\
&= \frac{\partial x'^\mu}{\partial x^\alpha}\left(\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\nu} - \frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial x^\gamma}{\partial x'^\lambda}\right)\Gamma^\alpha_{\beta\gamma}(x) \\
&= \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\nu}\left(\Gamma^\alpha_{\beta\gamma}(x) - \Gamma^\alpha_{\gamma\beta}(x)\right).
\end{aligned}$$

Hence $\Gamma^\mu_{\lambda\nu} - \Gamma^\mu_{\nu\lambda}$ is a tensor. We call this tensor as a *Torsion tensor*.

In the theory of Einstein's gravitiy, we assume $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu}$, i.e., Torsionless connection. From this assumption,

$$\begin{aligned}
0 &= \nabla_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - g_{\rho\nu}\Gamma^\rho_{\lambda\mu} - g_{\mu\rho}\Gamma^\rho_{\lambda\nu} \\
0 &= \nabla_\mu g_{\lambda\nu} = \partial_\mu g_{\lambda\nu} - g_{\rho\nu}\Gamma^\rho_{\mu\lambda} - g_{\lambda\rho}\Gamma^\rho_{\mu\nu} \\
0 &= \nabla_\nu g_{\lambda\mu} = \partial_\nu g_{\lambda\mu} - g_{\rho\mu}\Gamma^\rho_{\nu\lambda} - g_{\lambda\rho}\Gamma^\rho_{\nu\mu}.
\end{aligned}$$

From the above, we obtain

$$0 = \partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\mu g_{\lambda\mu} - 2g_{\rho\nu}\Gamma^\rho_{\lambda\mu}.$$

Hence

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

Taking inverse $g^{\rho\nu}$, we have

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

It is easy to see that $\Gamma^\rho_{\lambda\mu} = \Gamma^\rho_{\mu\lambda}$.

In summary, from $\nabla_\lambda g_{\mu\nu} = 0$ and $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu}$, we obtain

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

.and we call this the Christoffel symbol. Here note that we obtained the above formula from the geodesic equation (2.2). We call $\nabla_\lambda g_{\mu\nu} = 0$ as an *equivalence principle*, which is a generalization of free falling or no gravitiy.

Remark. Note that in (x, y, z) coordinate,

$$ds^2 = \delta_{ij}dx^i dx^j.$$

In (r, θ, φ) coordinate, we have

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2.$$

Note that

$$\begin{aligned}\nabla^\mu\nabla_\mu\phi &= g^{\mu\nu}\nabla_\mu\nabla_\nu\phi\\ &= g^{\mu\nu}\nabla_\mu\left(\partial_\nu\phi\right)\\ &= g^{\mu\nu}\left(\partial_\mu\partial_\nu\phi-\Gamma^\lambda_{\mu\nu}\partial_\lambda\phi\right)\\ &= g^{\mu\nu}\partial_\mu\partial_\nu\phi-g^{\mu\nu}\Gamma^\lambda_{\mu\nu}\partial_\lambda\phi\\ &= \partial^\mu\partial_\mu\phi-g^{\mu\nu}\Gamma^\lambda_{\mu\nu}\partial_\lambda\phi..\end{aligned}$$

Note

$$\begin{aligned}g^{\mu\nu}\Gamma^\lambda_{\mu\nu} &= g^{\mu\nu}\frac{1}{2}g^{\lambda\rho}\left(\partial_\mu g_{\rho\nu}+\partial_\nu g_{\beta\mu}-\partial_\rho g_{\mu\nu}\right)\\ &= \frac{1}{2}g^{\lambda\rho}\left(g^{\mu\nu}\partial_\mu g_{\rho\nu}+g^{\mu\nu}\partial_\nu g_{\beta\mu}-g^{\mu\nu}\partial_\rho g_{\mu\nu}\right)\\ &= g^{\lambda\rho}g^{\mu\nu}\partial_\mu g_{\rho\nu}-\frac{1}{2}g^{\lambda\rho}g^{\mu\nu}\partial_\rho g_{\mu\nu}\\ &= -\partial_\mu g^{\mu\lambda}-\frac{1}{2}\left(\partial^\lambda g_{\mu\nu}\right)g^{\mu\nu}.\end{aligned}$$

In (r,θ,φ) coordinate, $g_{rr}=1$, $g_{\theta\theta}=r^2$, $g_{\varphi\varphi}=r^2\sin^2\theta$. So $g^{rr}=1$, $g^{\theta\theta}=\frac{1}{r^2}$, $g^{\varphi\varphi}=\frac{1}{r^2\sin^2\theta}$. So

$$\begin{aligned}\partial_\mu g^{\mu r} &= \partial_r g^{rr}+\partial_\theta g^{\theta r}+\partial_\varphi g^{\varphi r}=0,\\ \partial_\mu g^{\mu\theta} &= \partial_\theta g^{\theta\theta}=0\\ \partial_\mu g^{\mu\varphi} &= \partial_\varphi g^{\varphi\varphi}=0.\end{aligned}$$

Now

$$\begin{aligned}g^{\mu\nu}\partial_\lambda g_{\mu\nu} &= g^{rr}\partial_\lambda g_{rr}+g^{\theta\theta}\partial_\lambda g_{\theta\theta}+g^{\varphi\varphi}\partial_\lambda g_{\varphi\varphi}\\ &= \frac{1}{r^2}\partial_\lambda\left(r^2\right)+\frac{1}{r^2\sin^2\theta}\partial_\lambda\left(r^2\sin^2\theta\right).\end{aligned}$$

So

$$\begin{aligned}&g^{\mu\nu}\partial_r g_{\mu\nu}=\frac{4}{r}\Rightarrow g^{\mu\nu}\partial^r g_{\mu\nu}=\frac{4}{r}\\ &g^{\mu\nu}\partial_\theta g_{\mu\nu}=2\cot\theta\Rightarrow g^{\mu\nu}\partial^\theta g_{\mu\nu}=\frac{2}{r^2}\cot^2\theta.\\ &g^{\mu\nu}\partial_\varphi g_{\mu\nu}=0.\end{aligned}$$

Hence

$$\begin{aligned}\nabla^\mu\nabla_\mu\phi &= \left(\frac{\partial^2}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right)\phi+\frac{2}{r}\frac{\partial}{\partial r}\phi+\frac{1}{r^2}\cot\theta\frac{\partial}{\partial\theta}\phi\\ &= \frac{1}{r^2}\frac{\partial^2}{\partial r^2}\left(r^2\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\phi\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\varphi^2}.\end{aligned}$$

For the case of vector,

$$\nabla^\mu\nabla_\mu V^\nu\neq\left(\partial^\mu\partial_\mu-g^{\mu\rho}\Gamma^\lambda_{\mu\rho}\partial_\lambda\right)V^\nu.$$

In fact,

$$\begin{aligned}\nabla^\mu\nabla_\mu V^\nu &= g^{\mu\lambda}\nabla_\lambda\nabla_\mu V^\nu\\ &= g^{\mu\nu}\left(\partial_\lambda\nabla_\mu V^\nu+\Gamma^\nu_{\lambda\rho}\nabla_\mu V^\rho-\Gamma^\rho_{\lambda\mu}\nabla_\rho V^\nu\right)\\ &= g^{\mu\lambda}\left[\partial_\lambda\left(\partial_\mu V^\nu+\Gamma^\nu_{\mu\rho}V^\rho\right)+\Gamma^\nu_{\lambda\rho}\left(\partial_\mu V^\rho+\Gamma^\rho_{\mu\kappa}V^\kappa\right)-\Gamma^\rho_{\lambda\mu}\left(\partial_\rho V^\nu+\Gamma^\nu_{\rho\kappa}V^\kappa\right)\right].\end{aligned}$$

2.3 Existence of Metric compatibility

In this section, we show the existence of metric $g_{\mu\nu}$ satisfying

$$\nabla_\lambda g_{\mu\nu} = 0.$$

Recall the geodesic equation of motion

$$\ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0. \tag{2.5}$$

If $x^\mu(\tau)$ is a solution to the geodesic motion, so is $x^\mu(k\tau)$ for $k \in \mathbb{R}$, i.e., if $x'^\mu(\tau) \equiv x^\mu(k\tau)$, then

$$\ddot{x}'^\lambda + \Gamma^\lambda_{\mu\nu}(x')\,\dot{x}'^\mu\dot{x}'^\nu = 0.$$

Indeed,

$$\frac{d}{d\tau}x'^\mu(\tau) = k\dot{x}^\mu(k\tau)$$
$$\frac{d^2}{d\tau^2}x'^\mu(\tau) = k^2\ddot{x}^\mu(k\tau)$$

gives

$$\ddot{x}'^\lambda + \Gamma^\lambda_{\mu\nu}(x')\,\dot{x}'^\mu\dot{x}'^\nu = k^2\left[\ddot{x}^\lambda + \Gamma^\lambda_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\right] = 0.$$

Now we consider initial-value problem. If x is an analytic solution of (2.5), then

$$x^\mu(\tau) = x^\mu_0 + \tau v^\mu - \frac{\tau^2}{2}\Gamma^\mu_{\nu\rho}(x_0)\,v^\mu v^\rho + \cdots,$$

where x_0 denotes the initial position and v denotes the initial velocity.

Note that for $k \in \mathbb{R}$,

$$x^\mu\left(k\cdot 0, v^\lambda\right) = 0,$$
$$\frac{d}{d\tau}x^\mu\left(k\tau, v^\lambda\right)\Big|_{\tau=0} = kv^\lambda.$$

From this observation, we get

$$x^\mu\left(\tau, v^\lambda\right) = x^\mu\left(\hat{\tau}, \frac{\tau}{\hat{\tau}}v^\lambda\right),$$

where $\hat{\tau}$ is an arbitrary fixed time.

If f is an analytic solution of (2.5), then we can choose a coordinate transform $x^\mu \mapsto y^\mu(x)$ with $x^\mu = x^\mu_0$ and $y^\mu = 0$ satisfying

$$x^\mu = f^\mu\left(\hat{\tau}, \frac{y^\lambda}{\hat{\tau}}\right).$$

The idea is to choose a coordinate in terms of velocity vector in the unit time. This identity defines a coordinate transformation from x^μ to y^μ.

Note that the geodesic equation transforms like a tensor under coordinate transforms. Indeed,

$$\ddot{x}'^\mu = \ddot{x}^\nu\frac{\partial x'^\mu}{\partial x^\nu} + \dot{x}^\nu\dot{x}^\rho\frac{\partial^2 x'^\mu}{\partial x^\nu\partial x^\rho}$$

and

$$\begin{aligned}\Gamma'^\mu_{\nu\rho}(x')\,\dot{x}'^\nu\dot{x}'^\rho &= \frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma + \frac{\partial^2 x^\kappa}{\partial x'^\rho\partial x'^\sigma}\frac{\partial x'^\rho}{\partial x^\alpha}\frac{\partial x^\sigma}{\partial x^\beta}\dot{x}^\alpha\dot{x}^\beta \\ &= \frac{\partial x'^\mu}{\partial x^\kappa}\Gamma^\kappa_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma - \frac{\partial^2 x'^\mu}{\partial x^\alpha\partial x^\beta}\dot{x}^\alpha\dot{x}^\beta.\end{aligned}$$

So

$$\begin{aligned}&\ddot{x}'^{\lambda} + \Gamma'^{\lambda}_{\mu\nu}\left(x'\right)\dot{x}'^{\mu}\dot{x}'^{\nu}\\&= \frac{\partial x'^{\lambda}}{\partial x^{\kappa}}\left(\ddot{x}^{\kappa} + \Gamma^{\kappa}_{\rho\sigma}\left(x\right)\dot{x}^{\rho}\dot{x}^{\sigma}\right).\end{aligned}$$

This shows x^{μ} is a solution of (2.5). Note that

$$\begin{aligned}x^{\mu}\left(\tau, v^{\lambda}\right) &= f^{\mu}\left(\tau, v^{\lambda}\right)\\&= f^{\mu}\left(\hat{\tau}, \frac{\tau v^{\lambda}}{\hat{\tau}}\right).\end{aligned}$$

From this, $y^{\mu}\left(\tau, v^{\lambda}\right) = \tau v^{\mu}$. This describes a geodesic motion in y-coordinate system, i.e., this satisfies

$$\ddot{y}^{\mu} + \Gamma^{\mu}_{\nu\rho}\left(y\right)\dot{y}^{\nu}\dot{y}^{\rho} = 0.$$

and this implies

$$\Gamma^{\mu}_{\nu\rho}\left(\tau v\right) = v^{v}v^{\rho} = 0.$$

Especially at $\tau = 0$, for the y-coordinate system

$$\Gamma^{\lambda}_{\mu\nu}\left(y = 0\right)v^{\mu}v^{\nu} = 0$$

for arbitrary v^{μ}. So $\Gamma^{\lambda}_{\mu\nu}\left(y = 0\right) = 0$. Hence $\frac{\partial}{\partial y^{\lambda}} g_{\mu\nu}\left(y\right)\Big|_{y=0} = 0$. Thus, y^{μ} is identified as the locally inertial frame. We call this coordinate as a *Riemann normal coordinate*.

Usually, we write

$$\frac{D^2 x^{\mu}}{D\tau^2} := \ddot{x}^{\mu} + \Gamma^{\mu}_{\nu\rho}\dot{x}^{\nu}\dot{x}^{\rho}$$

and we call this as a *covariant acceleration*

2.4 The Einstein field equation, Tensor density

In this section, we study the Einstein field equation:

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

Here G is a Newton constant, $R_{\mu\nu}$ and R will be defined later. $T_{\mu\nu}$ is an energy-momentum tensor. On the left hand side, this concerns geometry. On the right hand side, this concerns matter, source of gravity or curved spacetime.

As an example, if we consider a point particle, one of example of energy-momentum tensor is

$$T^{\mu\nu} \sim m\frac{dx^{\mu}}{d\tau}\left(\tau\right)\frac{dx^{\nu}}{d\tau}\left(\tau\right).$$

There is an energy-momentum tensor conservation

$$\begin{aligned}\partial_{\mu}T^{\mu\nu} &= 0,\\\nabla_{\mu}T^{\mu\nu} &= 0.\end{aligned}$$

There is some ambiguity on the equation since

$$\nabla_{\mu}\left(\Lambda g^{\mu\nu}\right) = 0,$$

where Λ is a constant. So later Einstein add some terms on Einstein equation

$$\Lambda g_{\mu\nu} + R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}.$$

Here Λ is a cosmology constant. Einstein thought that universe is static. However, due to Hubble, it is not. Einstein later mentioned that "adding that term is the one of my big mistakes!"

One might consider

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu}.$$

Later we will explain the meaning of Λterm.

From now on, we consider $\Lambda = 0$.

Consider a transformation

$$g_{\mu\nu}(x) \mapsto g'_{\mu\nu}(x) = \frac{\partial x^\lambda}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} g_{\lambda\rho}(x).$$

$$g = \det g_{\mu\nu} = \|g\|_{\mu\nu} \to g' = \det\left(g'_{\mu\nu}\right) = \left|\frac{\partial x}{\partial x'}\right|^2 g = \left|\frac{\partial x'}{\partial x}\right|^{-2} g.$$

So

$$g \mapsto \left|\frac{\partial x}{\partial x'}\right|^2 g.$$

We call this as scalar density with weight $\omega = 2$. Based on this observation, one may define the tensor density with weight ω

$$T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}(x) \mapsto T'^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}(x') = \left|\frac{dx'}{dx}\right|^{-\omega} \frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\mu_p}}{\partial x^{\lambda_q}} \frac{\partial x^{\rho_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\rho_q}}{\partial x'^{\nu_q}} T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}.$$

We require

$$\nabla_\lambda \left(\frac{T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}}{\left(\sqrt{-g}\right)^\omega}\right) = \frac{\nabla_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}}{\left(\sqrt{-g}\right)^\omega} + T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} \nabla_\lambda \left(\sqrt{-g}\right)^{-\omega}.$$

By computation,

$$\begin{aligned}
\nabla_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} &= \left(\sqrt{-g}\right)^\omega \nabla_\lambda \left(\frac{T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}}{\left(\sqrt{-g}\right)^\omega}\right) - \left(\sqrt{-g}\right)^\omega \nabla_\lambda \left(\sqrt{-g}\right)^{-\omega} T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} \\
&= \partial_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} + \sum_i \Gamma^{\mu_i}_{\lambda\rho} T_\omega{}^{\cdots\rho\cdots}{}_{\nu_1\cdots\nu_q} - \sum_j \Gamma^\rho_{\lambda\nu_j} T_\omega{}^{\mu_1\cdots\mu_p}{}_{\cdots\rho\cdots} \\
&+ \left(\sqrt{-g}\right)^\omega \left[\partial_\lambda \left(\sqrt{-g}\right)^{-\omega} - \nabla_\lambda \left(\sqrt{-g}\right)^{-\omega}\right] T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}.
\end{aligned}$$

Also by computation,

$$\begin{aligned}
\nabla_\lambda \left(\sqrt{-g}\right)^\omega &= \omega \left(\sqrt{-g}\right)^{\omega-1} \nabla_\lambda \sqrt{-g} \\
\nabla_\lambda \left(\sqrt{-g}\right)^\omega &= \frac{\omega}{2} (-g)^{\frac{\omega}{2}-1} \nabla_\lambda (-g).
\end{aligned}$$

Later we show

$$\delta \ln(\det M) = \text{Tr}\left(M^{-1} \delta M\right).$$

To show this, recall

$$|M| = \det M = \varepsilon^{a_1 a_2 \cdots a_n} M_{1a_1} \cdots M_{na_n},$$

where $\varepsilon^{a_1\cdots a_n}$ denotes the Levi-Civita symbol. Now taking variation to get

$$\begin{aligned}
\delta\left|M\right| &= \varepsilon^{a_1 a_2\cdots a_n}\sum_{i=1}^{n}\delta M_{ia_i}M_{1a_1}\cdots M_{i-1a_{i-1}}M_{i+1a_{i+1}}\cdots M_{na_n}\\
&= \varepsilon^{a_1\cdots a_n}\sum_{i=1}^{n}\delta M_{ik}\delta^k_{a_i}M_{1a_1}\cdots M_{i-1a_{i-1}}M_{i+1a_{i+1}}\cdots M_{na_n}\\
&= \varepsilon^{a_1\cdots a_n}\sum_{i=1}^{n}\delta M_{ik}\cdots\left(M^{-1}\right)^{kl}M_{la_i}M_{a_1}\cdots M_{i-1a_{i-1}}M_{i+1a_{i+1}}\cdots M_{na_n}\\
&= \sum_{i=1}^{n}\delta M_{ik}\left(M^{-1}\right)^{kl}\varepsilon^{a_1\cdots a_n}M_{1a_1}\cdots M_{i-1a_{i-1}}M_{la_i}M_{i+1a_{i+1}}\cdots M_{Na_N}\\
&= |M|\sum_{i=1}^{n}\delta M_{ik}\left(M^{-1}\right)^{kl}\delta^i_l\\
&= |M|\,\mathrm{Tr}\left(M^{-1}\delta M\right).
\end{aligned}$$

So from this we get

$$\partial_\lambda g = g\, g^{\mu\nu}\partial_\lambda g_{\nu\mu}.$$

Now define

$$\begin{aligned}
\nabla_\lambda g :&= \partial_\lambda g - \left(g^{\mu\nu}\partial_\lambda g_{\nu\mu}\right)g\\
&= \partial_\lambda g - 2\Gamma^\mu_{\lambda\mu}g.
\end{aligned}$$

Hence

$$\nabla_\lambda\sqrt{-g} := \partial_\lambda\sqrt{-g} - \Gamma^\mu_{\lambda\mu}\sqrt{-g}\equiv 0.$$

Similarly,

$$\nabla_\lambda\left(\sqrt{-g}\right)^\omega\equiv 0.$$

Observe that

$$\begin{aligned}
\partial_\lambda\left(\sqrt{-g}\right) &= \frac{\partial_\lambda\left(-g\right)}{2\sqrt{-g}}\\
&= \frac{-2\Gamma^\mu_{\lambda\mu}g}{2\sqrt{-g}} = \Gamma^\mu_{\lambda\mu}\sqrt{-g}.
\end{aligned}$$

From this observation, we define

$$\nabla_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} = \partial_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q} + \sum_i\Gamma^{\mu_i}_{\lambda\rho}T_\omega{}^{\cdots\rho\cdots}{}_{\nu_1\cdots\nu_q} - \sum_j\Gamma^{\rho}_{\lambda\nu_j}T_\omega{}^{\mu_1\cdots\mu_p}{}_{\cdots\rho\cdots} - \omega\Gamma^\rho_{\lambda\rho}T_\omega{}^{\mu_1\cdots\mu_\rho}{}_{\nu_1\cdots\nu_q}.$$

Now if

$$T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}\left(x\right)\mapsto T'^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}\left(x'\right) = \left|\frac{dx'}{dx}\right|^{-\omega}\frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}}\cdots\frac{\partial x'^{\mu_p}}{\partial x^{\lambda_q}}\frac{\partial x^{\rho_1}}{\partial x'^{\nu_1}}\cdots\frac{\partial x^{\rho_q}}{\partial x'^{\nu_q}}T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q},$$

then

$$\nabla_\lambda T_\omega{}^{\mu_1\cdots\mu_p}{}_{\nu_1\cdots\nu_q}\to\left|\frac{\partial x'}{\partial x}\right|^{-\omega}\frac{\partial x^\kappa}{\partial x'^\lambda}\frac{\partial x'^\mu}{\partial x^\rho}\cdots\frac{\partial x^\sigma}{\partial x'^\nu}\nabla_\kappa T_\omega{}^{\rho\cdots}{}_{\sigma\cdots}.$$

Thus the covariant derivative behaves like tensor under the coordinate transform.

Example 2.7. The Levi-Civita symbol is Tensor density weight 1 since

$$\varepsilon^{\mu_1\cdots\mu_4} = \left|\frac{\partial x'}{\partial x}\right|^{-1} \frac{\partial x'^{\mu_1}}{\partial x^{\nu_1}} \cdots \frac{\partial x'^{\mu_4}}{\partial x^{\nu_4}}.$$

From this, we have

$$\begin{aligned}\nabla_\lambda \varepsilon^{\mu_1\cdots\mu_4} &= \partial_\lambda \varepsilon^{\mu_1\cdots\mu_4} + \sum_i \Gamma^{\mu_i}_{\lambda\rho} \varepsilon^{\mu_1\cdots\rho\cdots\mu_4} - \Gamma^{\rho}_{\lambda\rho}\varepsilon^{\mu_1\cdots\mu_4} \\ &= \partial_\lambda \varepsilon^{\mu_1\cdots\mu_4}.\end{aligned}$$

Example 2.8. Vector density with weight one, $\omega = 1$ J^μ

$$J^\mu(x) \mapsto J'^\mu(x') = \left|\frac{\partial x}{\partial x'}\right| \frac{\partial x'^\mu}{\partial x^\nu} J^\nu(x).$$

Then

$$\begin{aligned}\nabla_\mu J^\mu &= \partial_\mu J^\mu + \Gamma^\mu_{\mu\nu} J^\nu - \Gamma^\rho_{\mu\rho} J^\mu \\ &= \partial_\mu J^\mu.\end{aligned}$$

So the covariant divergence is the classical divergence.

For $x^\mu_a(\tau)$, q_a a charge,

$$J^\mu(x) = \sum_a \int d\tau q_a \dot{x}^\mu_a(\tau)\,\delta(x - x_a(\tau))$$

We call this as a current vector density with weight $\omega = 1$. Then

$$\begin{aligned}\partial_\mu J^\mu &= \frac{\partial}{\partial x^\mu}\left[\sum_a \int d\tau q_a \dot{x}^\mu(\tau)\,\delta(x - x_a(\tau))\right] \\ &= \sum_a \int d\tau q_a \dot{x}^\mu(\tau) \frac{\partial}{\partial x^\mu}\delta(x - x_a(\tau)) \\ &= \sum_a \int d\tau q_a \dot{x}^\mu(\tau)\left[-\frac{\partial}{\partial x^\mu_a}\delta(x - x_a(\tau))\right] \\ &= \sum_a \int d\tau q_a \frac{d}{d\tau}\delta(x - x_a(\tau)) \\ &= \sum_a q_a\ \delta(x - x_a(\tau))\big|^{\infty}_{\tau=-\infty} = 0.\end{aligned}$$

Since $t(\tau) \sim \tau$.

Let f be a scalar function, i.e., $f(x) \mapsto f(x') = f(x)$. Then

$$\begin{aligned}\int d^4x f(x)\,\delta(x-y)\,dy = f(y) \mapsto \int d^4x' f'(x')\,\delta'(x'-y') &= f(x) \\ &= \int d^4x \left|\frac{dx'}{dx}\right| f(x)\,\delta'(x'-y') \\ &= \int d^4x f(x)\,\delta(x-y) \\ &= f(y).\end{aligned}$$

So J^μ is a vector density weight $\omega = 1$. So for the unweighted version, we have

$$J^\mu_{\omega=0} = \sum_a \int d\tau q_a \dot{x}^\mu_a \frac{\delta\left(x - x_a\left(q\right)\right)}{\sqrt{-g}}.$$

Now we define

$$T^{\mu\nu}\left(x\right) = \sum_a \int d\tau m_a \dot{x}^\mu \dot{x}^\nu \frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g\left(x\right)}}.$$

Then

$$\nabla_\mu T^{\mu\nu}\left(x\right) = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\rho} T^{\rho\nu} + \Gamma^\nu_{\mu\rho} T^{\mu\rho}.$$

Note that

$$\partial_\mu T^{\mu\nu} = \sum_a \int d\tau m_a \dot{x}^\mu_a\left(\tau\right)\dot{x}^\nu_a\left(\tau\right)\left[\frac{\partial_\mu \delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g}} + \delta\left(x - x_a\left(\tau\right)\right)\partial_\mu\left(-g\right)^{-\frac{1}{2}}\right].$$

Observe that

$$\begin{aligned}\int d\tau \dot{x}^\mu_a\left(\tau\right)\dot{x}^\nu_a\left(\tau\right)\frac{\partial_\mu \delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g\left(x\right)}} &= \int d\tau\left(-\dot{x}^\nu_a\left(\tau\right)\right)\frac{\frac{d}{d\tau}\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g\left(x\right)}} \\ &= \int d\tau \ddot{x}^\nu_a\left(\tau\right)\frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g\left(x\right)}}.\end{aligned}$$

From this observation, we continue the calculation. Then

$$\partial_\mu T^{\mu\nu} = \sum_a \int d\tau m_a\left[\ddot{x}^\nu_a\left(\tau\right)\frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g}} + \dot{x}^\mu_a\left(\tau\right)\dot{x}^\nu_a\left(\tau\right)\delta\left(x - x_a\left(\tau\right)\right)\times\frac{\left(-\Gamma^\rho_{\mu\rho}\right)}{\sqrt{-g}}\right].$$

From

$$0 = \nabla_\mu\left(-g\right)^{-\frac{1}{2}} = \partial_\mu\left(-g\right)^{-\frac{1}{2}} + \Gamma^\rho_{\mu\rho}\left(-g\right)^{-\frac{1}{2}},$$

we get

$$\partial_\mu T^{\mu\nu} = \sum_a \int d\tau m_a \frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g}}\left[\ddot{x}^\nu_a\left(\tau\right) - \dot{x}^\nu_a\left(\tau\right)\dot{x}^\mu_a\left(\tau\right)\Gamma^\rho_{\mu\rho}\right].$$

Hence

$$\begin{aligned}\nabla_\mu T^{\mu\nu}\left(x\right) &= \sum_a \int d\tau m_a \frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g}}\left[\ddot{x}^\nu_a\left(\tau\right) - \dot{x}^\nu_a\left(y\right)\dot{x}^\mu_a\left(\tau\right)\Gamma^\rho_{\mu\rho} + \Gamma^\mu_{\mu\rho}\dot{x}^\rho_a\dot{x}^\nu_a + \Gamma^\nu_{\mu\rho}\dot{x}^\mu_a\dot{x}^\nu_a\right] \\ &= \sum_a \int d\tau m_a \frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g}}\left[\ddot{x}^\nu_a\left(\tau\right) + \Gamma^\nu_{\mu\rho}\left(x_a\left(\tau\right)\right)\dot{x}^\mu_a\dot{x}^\nu_a\right].\end{aligned}$$

In contrast to the current vector case,

$$\nabla_\mu T^{\mu\nu} \equiv 0$$

up to geodesic motion. This shows energy-momentum is conserved if there is no exterior force.

In the case of static case, $\dot{x}^\mu\left(\tau\right) = \left(\frac{dt}{d\tau}, 0, 0, 0\right) = \left(1, 0, 0, 0\right)$. So $T^{00} \neq 0$ otherwise $T^{\mu\nu} = 0$. In this case,

$$T^{00} = \sum_a \int d\tau m_a \frac{\delta\left(x - x_a\left(\tau\right)\right)}{\sqrt{-g\left(x\right)}}.$$

Note that the source of gravity is mass. From this, the right-hand side of Einstein equation can be regareded as a generalization of Newton's mechanics since

$$\nabla^2 \Phi = 4\pi G \rho$$

Let us explain the left hand side of the Einstein field equation. If we have a connection, we consider its curvature. To find a curvature, it suffices to calculate $[\nabla_\mu, \nabla_\nu]$. For scalar function,

$$\begin{aligned}\nabla_\mu \nabla_\nu \phi &= \nabla_\mu \partial_\nu \phi \\ &= \partial_\mu \partial_\nu \phi - \Gamma^\rho_{\mu\nu} \partial_\rho \phi.\end{aligned}$$

So $[\nabla_\mu, \nabla_\nu] \phi = 0$.

Let us consider

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

Similarly,

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

Hence

$$\begin{aligned}\left[\nabla_\mu, \nabla_\nu \right] V^\lambda &= \left[\partial_\mu \Gamma^\lambda_{\nu\,\sigma} - \partial_\nu \Gamma^\lambda_{\mu\,\sigma} + \Gamma^\lambda_{\mu\,\rho} \Gamma^\rho_{\nu\,\sigma} - \Gamma^\lambda_{\nu\,\rho} \Gamma^\rho_{\nu\,\sigma} \right] V^\sigma, \\ \left[\nabla_\mu, \nabla_\nu \right] U_\lambda &= -U_\sigma \left[\partial_\mu \Gamma^\sigma_{\nu\,\lambda} - \partial_\nu \Gamma^\sigma_{\mu\,\rho} + \Gamma^\sigma_{\mu\,\rho} \Gamma^\rho_{\nu\,\lambda} - \Gamma^\lambda_{\nu\,\rho} \Gamma^\rho_{\mu\,\lambda} \right],\end{aligned}$$

From this we define

$$R^\kappa_{\lambda\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + \Gamma^\kappa_{\mu\,\rho} \Gamma^\rho_{\nu\,\lambda} - \Gamma^\kappa_{\nu\,\rho} \Gamma^\rho_{\mu\,\lambda} = -R^\kappa_{\lambda\nu\mu}$$

and we call this as a *Riemann curvature*.

It is better to memorize

$$\begin{cases} R^\times_{\times\mu\nu} = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + \left[\Gamma_\mu, \Gamma_\nu \right] \\ \left[\nabla_\mu, \nabla_\nu \right] = \left[\partial_\mu + \Gamma_\mu, \partial_\nu + \Gamma_\nu \right] = \partial_\mu \Gamma_\nu - \partial_\nu \Gamma_\mu + \left[\Gamma_\mu, \Gamma_\nu \right]. \end{cases}$$

From the above calculation law,

$$\left[\nabla_\mu, \nabla_\nu \right] T^{\kappa_1 \cdot \kappa_p}{}_{\lambda_1 \cdots \lambda_q} = \sum_i R^{\kappa_i}_{\rho\mu\nu} T^{\kappa_1 \cdots \rho \cdots \kappa_p} \qquad - \sum_j T^{\kappa_1 \cdots \kappa_p}{}_{\lambda_1 \cdots \rho \cdots \lambda_q} R^\rho_{\lambda_j \mu\nu}.$$

Note

$$\begin{aligned}R_{\kappa\lambda\mu\nu} &= g_{\kappa\rho} R^\rho_{\lambda\mu\nu} \\ &= g_{\kappa\rho} \left[\partial_\mu \Gamma^\rho_{\nu\,\lambda} - \partial_\nu \Gamma^\rho_{\mu\,\lambda} + \Gamma^\rho_{\mu\,\sigma} \Gamma^\sigma_{\nu\,\lambda} - \Gamma^\rho_{\nu\,\sigma} \Gamma^\sigma_{\mu\,\lambda} \right] \\ &= \partial_\mu \left(g_{\kappa\rho} \Gamma^\rho_{\nu\lambda} \right) - \partial_\nu \left(g_{\kappa\rho} \Gamma^\rho_{\mu\nu} \right) - \partial_\mu g_{\kappa\rho} \Gamma^\rho_{\nu\lambda} + \partial_\nu g_{\kappa\rho} \Gamma^\rho_{\mu\lambda} + g_{\kappa\rho} \Gamma^\rho_{\mu\sigma} T^\sigma_{\nu\lambda} - g_{\kappa\rho} \Gamma^\rho_{\nu\sigma} \Gamma^\sigma_{\mu\lambda} \\ &= \partial_\mu \left(g_{\kappa\rho} \Gamma^\rho_{\nu\lambda} \right) - \partial_\nu \left(g_{\kappa\rho} \Gamma^\rho_{\mu\nu} \right) - \left(g_{\kappa\sigma} \Gamma^\sigma_{\mu\rho} + g_{\rho\sigma} \Gamma^\sigma_{\mu\kappa} \right) \Gamma^\rho_{\nu\lambda} \\ &\quad + \left(g_{\kappa\sigma} \Gamma^\sigma_{\nu\rho} + g_{\rho\sigma} \Gamma^\sigma_{\nu\kappa} \right) \Gamma^\rho_{\mu\lambda} + g_{\kappa\rho} \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\nu\lambda} - g_{\kappa\rho} T^\rho_{\nu\sigma} \Gamma^\sigma_{\mu\lambda} \\ &= \partial_\mu \left(g_{\kappa\rho} \Gamma^\rho_{\nu\lambda} \right) - \partial_\nu \left(g_{\kappa\rho} \Gamma^\rho_{\mu\nu} \right) + g_{\mu\nu} \left(\Gamma^\rho_{\mu\lambda} \Gamma^\sigma_{\nu\kappa} - \Gamma^\rho_{\nu\lambda} \Gamma^\sigma_{\mu\kappa} \right).\end{aligned}$$

Here we used

$$0 = \nabla_\mu g_{\kappa\rho} = \partial_\mu g_{\kappa\rho} - \Gamma^\sigma_{\mu\kappa} g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho} g_{\kappa\sigma}$$
$$\partial_\mu g_{\kappa\rho} = g_{\rho\sigma}\Gamma^\sigma_{\mu\kappa} + g_{\kappa\sigma}\Gamma^\sigma_{\mu\rho}.$$

Now note that

$$g_{\kappa\rho}\Gamma^\rho_{\nu\lambda} = \frac{1}{2}\left(\partial_\nu g_{\kappa\lambda} + \partial_\lambda g_{\kappa\nu} - \partial_\kappa g_{\nu\lambda}\right).$$

So

$$\begin{aligned} R_{\kappa\lambda\mu\nu} &= \frac{1}{2}\left[\partial_\mu\partial_\nu g_{\kappa\lambda} + \partial_\mu\partial_\lambda g_{\kappa\nu} - \partial_\mu\partial_\kappa g_{\nu\lambda} - \partial_\nu\partial_\mu g_{\kappa\lambda} - \partial_\nu\partial_\lambda g_{\kappa\mu} + \partial_\nu\partial_\kappa g_{\mu\lambda}\right] \\ &\quad + g_{\rho\sigma}\left(\Gamma^\rho_{\mu\lambda}\Gamma^\sigma_{\nu\kappa} - \Gamma^\rho_{\nu\lambda}\Gamma^\sigma_{\mu\kappa}\right) \\ &= \frac{1}{2}\left(\partial_\mu\partial_\lambda g_{\kappa\nu} + \partial_\nu\partial_\kappa g_{\mu\lambda} - \partial_\mu\partial_\kappa g_{\nu\lambda} - \partial_\nu\partial_\lambda g_{\kappa\mu}\right) + g_{\rho\sigma}\left(\Gamma^\rho_{\mu\lambda}\Gamma^\sigma_{\nu\kappa} - \Gamma^\rho_{\nu\lambda}\Gamma^\sigma_{\mu\kappa}\right). \end{aligned}$$

From the above identities, we can derive some properties of Riemann tensor.

- $R^\lambda_{\ \rho\mu\nu} = -R^\lambda_{\ \rho\nu\mu}$.
- $R_{\kappa\lambda\mu\nu} = -R_{\kappa\lambda\nu\mu}$.
- $R_{\kappa\lambda\mu\nu} = -R_{\lambda\kappa\mu\nu}$.
- $R_{\kappa\lambda\mu\nu} = R_{\mu\nu\kappa\lambda}$.
- $R_{\kappa\lambda\mu\nu} + R_{\kappa\mu\nu\lambda} + R_{\kappa\nu\lambda\mu} = 0$.

Actually, by considering equilibrium principle, we can prove this theorem more easily.

Consider

$$R_{\kappa\lambda\mu\nu}(x) \mapsto R'_{\kappa\lambda\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\kappa}\frac{\partial x^\beta}{\partial x'^\lambda}\frac{\partial x^\gamma}{\partial x'^\mu}\frac{\partial x^\delta}{\partial x'^\nu} R_{\alpha\beta\gamma\delta}(x).$$

Recall the Jacobi's identity

$$[\nabla_\lambda[\nabla_\mu, \nabla_\nu]] + [\nabla_\mu[\nabla_\nu, \nabla_\lambda]] + [\nabla_\nu[\nabla_\lambda, \nabla_\mu]] = 0.$$

Note that

$$\begin{aligned} [\nabla_\lambda[\nabla_\mu, \nabla_\nu]]V^\kappa &= \nabla_\lambda[\nabla_\mu, \nabla_\nu]V^\kappa - [\nabla_\mu, \nabla_\nu]\nabla_\lambda V^\kappa \\ &= \nabla_\lambda\left(R^\lambda_{\ \sigma\mu\nu}V^\sigma\right) - R^\kappa_{\ \sigma\mu\nu}\nabla_\lambda V^\sigma + \nabla_\sigma V^\kappa R^\sigma_{\ \lambda\mu\nu} \\ &= \nabla_\lambda R^\kappa_{\ \sigma\mu\nu}V^\sigma + R^\kappa_{\ \sigma\mu\nu}\nabla_\lambda V^\sigma - R^\kappa_{\ \sigma\mu\nu}\nabla_\lambda V^\sigma + \nabla_\sigma V^\kappa R^\sigma_{\ \lambda\mu\nu} \\ &= \nabla_\lambda R^\kappa_{\ \sigma\mu\nu}V^\sigma + \nabla_\sigma V^\kappa R^\sigma_{\ \lambda\mu\nu}. \end{aligned}$$

By considering metric, we have

$$[\nabla_\lambda[\nabla_\mu, \nabla_\nu]]V_\kappa = -V_\sigma\nabla_\lambda R^\sigma_{\ \kappa\mu\nu} + \nabla_\sigma V_\kappa R^\sigma_{\ \lambda\mu\nu}.$$

Hence by considering cyclicity on indices and Jacobi identity, we have

$$\begin{aligned} 0 = &\left(\nabla_\lambda R^\kappa_{\ \sigma\mu\nu} + \nabla_\mu R^\kappa_{\ \sigma\nu\lambda} + \nabla_\nu R^\kappa_{\ \sigma\lambda\mu}\right)V^\sigma \\ &+ \nabla_\sigma V^\kappa\left(R^\sigma_{\ \lambda\mu\nu} + R^\sigma_{\ \mu\nu\lambda} + R^\sigma_{\ \nu\lambda\mu}\right). \end{aligned}$$

By the last property of Riemann tensor and arbitrariness of V, we finally get

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

2.5 Riemann curvature and spacetime

Theorem 2.9 (Poincare's Lemma)**.** *Let A be a 1-form on a simply connected domain. Then there exists a function ϕ such that $A_\mu = \partial_\mu \phi$ if and only if $F_{\mu\nu} = 0$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.*

Proof. One direction is obvious. Define

$$x^\mu (s) = s\left(x^\mu - x_0^\mu\right) + x_0^\mu =: x_s^\mu$$

where $0 \leq s \leq 1$. with $x^\mu (0) = x_0^\mu$, $x^\mu (1) = x^\mu$.Then

$$\frac{d}{ds} x^\mu (s) = x^\mu - x_0^\mu .$$

Define

$$\Phi(x) = \int_0^1 dx \frac{dx^\mu}{ds} A_\mu (x(s)) .$$

Then

$$\begin{aligned}
\frac{\partial}{\partial x^\mu}\Phi &= \frac{\partial}{\partial x^\mu}\int_0^1 ds \left(x^\lambda - x_0^\lambda\right) A_\lambda (x(s)) \\
&= \int_0^1 ds A_\mu (x_s) + \left(x^\lambda - x_0^\lambda\right) \frac{\partial x^\nu}{\partial x^\mu}(s) \frac{\partial A_\lambda (x_s)}{\partial x_s^\nu} \\
&= \int_0^1 ds A_\mu (x_s) + \left(x^\lambda - x_0^\lambda\right) \frac{\partial A_\lambda (x_s)}{\partial x_s^\mu} \\
&= \int_0^1 ds A_\mu (x_s) + s\left(x^\lambda - x_0^\lambda\right)\left[F_{\mu\nu}(x_s) + \frac{\partial A_\mu}{\partial x_s^\lambda}(x_s)\right] \\
&= \int_0^1 ds A_\mu (x_s) + s\frac{dx_s^\lambda}{ds}\frac{\partial A_\mu}{\partial x_s^\lambda}(x_s) \\
&= \int_0^1 ds A_\mu (x_s) + s\frac{d}{ds} A_\mu (x_s) \\
&= \int_0^1 ds \frac{d}{ds}\left(s A_\mu (x_s)\right) \\
&= A_\mu (x) .
\end{aligned}$$

This completes the proof of Poincaré's lemma. □

Theorem 2.10. *$R^\kappa_{\lambda\mu\nu} = 0$ if and only if spacetime is flat, i.e, there exists a coordinate system $\partial_\mu g_{\nu\rho} = 0$.*

Hence the Riemann curvature determines whether the spacetime is flat.

Proof. ($\Leftarrow$): Obvious.

($\Rightarrow$): Define $\Gamma(s)$ a 4×4 matrix amd $\Gamma(s)^\mu_{\ \nu} = -\left(x^\lambda - x_0^\lambda\right)\Gamma^\mu_{\lambda\,\nu}(x_s)$, where $x_s^\mu = s\left(x^\mu - x_0^\mu\right) + x_0^\mu$. Now define $w(s)$ a matrix

$$\frac{d}{ds} w(s) = \Gamma(s) w(s), \quad w(s=0) = I.$$

Note that such $w(s)$ exists since

$$w(s) = I + \int_0^s ds_1 \Gamma(s_1) w(s_1)$$

Taking iteration, then we get

$$w(s) = I + \sum_{n=1}^{\infty} \int_0^s ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \Gamma(s_1) \cdots \Gamma(s_n).$$

We denote this as $P \exp\left[\int ds \Gamma(s)\right]$, where $0 \le s_n \le s_{n-1} \le \cdots \le s_1 \le s$.

The above w satisfies

$$w(s_f, s_m) w(s_m, s_i) = w(s_f, s_i)$$

for $0 \le s_i \le s_m \le s_f$.

Now

$$\partial_\mu w = \sum_{n=1}^{\infty} \sum_{j=1}^{n} \int_0^s ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \Gamma(s_1) \cdots \partial_\mu \Gamma(s_j) \cdots \Gamma(s_n). \tag{2.6}$$

Note that

$$\begin{aligned} \partial_\mu \Gamma(s)^\lambda{}_\nu &= \partial_\mu \left[-(x^\rho - x_0^\rho) \Gamma^\lambda_{\rho\,\nu}(x_s) + x_0^\lambda \right] \\ &= -\Gamma^\lambda_{\mu\,\nu}(x_s) - s(x^\rho - x_0^\rho) \frac{\partial}{\partial x_s^\mu} \Gamma^\lambda_{\rho\nu}(x_s). \end{aligned}$$

Since $R^\kappa_{\lambda\mu\nu} = 0$, we have

$$\begin{aligned} \partial_\mu \Gamma(s) &= -\Gamma_\mu(x_s) - s(x^\rho - x_0^\rho) \left[\frac{\partial}{\partial x_s^\rho} \Gamma_\mu(x_s) - [\Gamma_\mu, \Gamma_\rho] \right] \\ &= -\frac{d}{ds}(s \Gamma_\mu(x_s)) - [\Gamma_\mu(x_s), s\Gamma(x_s)]. \end{aligned}$$

Plugin this relation to (2.6). Then we have

$$\begin{aligned} \partial_\mu w = & \sum_{n=1}^{\infty} \sum_{j=1}^{n} \int_0^1 ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n \Gamma(s_1) \cdots \Gamma(s_{j-1}) \\ & \times \left[-\frac{d}{ds} \left(\Gamma_\mu \left(x_{s_j} \right) \right) - s \Gamma_\mu \left(x_{s_j} \right) \Gamma \left(x_{s_j} \right) + \Gamma \left(x_{s_j} \right) s \Gamma_\mu \left(x_{s_j} \right) \right] \times \Gamma(x_{j+1}) \cdots \Gamma(x_{s_n}) \\ = & -\int_0^1 ds' \frac{d}{ds'} \left[w(s, s') s' \Gamma_\mu(x'_s) w(s', 0) \right] \\ = & \, \Gamma_\mu(x) w. \end{aligned}$$

So

$$\begin{aligned} \Gamma^\lambda_{\mu\,\rho} &= \partial_\mu w^\lambda_\sigma w^{-1\sigma}_\rho \\ &= -w^\lambda_\sigma \partial_\mu w^{-1\sigma}_\rho \\ \Gamma^\lambda_{\mu\rho} &= \Gamma^\lambda_{\rho\mu} \end{aligned}$$

yields $\partial_\mu w^{-1\sigma}_\rho = \partial_\rho w^{-1\sigma}_\mu$. Hence by Poincaré's lemma, there exists f such that $w^{-1\sigma}_\mu = \partial_\mu f^\sigma$. So

$$\Gamma^\lambda_{\mu\rho} = -\frac{\partial x^\lambda}{\partial f^\sigma} \frac{\partial^2 f^\sigma}{\partial x^\mu \partial x^\rho}.$$

Finally, consider $x^\mu \mapsto f^\mu(x)$ coordinate transform. Then we have $\Gamma^\lambda_{\mu\nu} = 0$, i.e., $\partial_\lambda g_{\mu\nu} = 0$. □

Remark. (i) We call w as a Wilson line.

$$\frac{\partial}{\partial s} w = \Gamma(s) w$$
$$\frac{\partial}{\partial s_i} w = -w\Gamma(s_i).$$

(ii) For $0 \leq s_n \leq s_{n-1} \leq \cdots \leq s_1 \leq 1$, by Fubini's theorem, we have

$$\int_0^1 ds_1 \int_0^{s_1} ds_2 \cdots \int_0^{s_{n-1}} ds_n F(s_1, \ldots, s_n)$$
$$= \int_0^1 ds_1 \int_{s_n}^1 ds_{n-1} \int_{s_{n-1}}^1 ds_{n-2} \cdots \int_{s_2}^1 ds_1 F(s_1, \ldots, s_n).$$

Substitute $0 \mapsto s_i$ and $1 \mapsto s$, we can prove identities in (i).

Now we study some contraction on Riemann curvature. By antisymmetricity of Riemann curvature, we have

$$g^{\kappa\lambda} R_{\kappa\lambda\mu\nu} = 0$$
$$g^{\mu\nu} R_{\kappa\lambda\mu\nu} = 0.$$

Also

$$g^{\kappa\mu} R_{\kappa\lambda\mu\nu} = -g^{\kappa\mu} R_{\kappa\lambda\nu\mu} = -g^{\kappa\mu} R_{\lambda\kappa\mu\nu} = g^{\kappa\mu} R_{\lambda\kappa\nu\mu}.$$

We define

$$R_{\lambda\nu} := g^{\kappa\mu} R_{\kappa\lambda\mu\nu}.$$

Note that

$$R_{\lambda\nu} = g^{\kappa\mu} R_{\lambda\kappa\nu\mu} = g^{\kappa\mu} R_{\kappa\nu\mu\lambda} = R_{\nu\lambda}.$$

So the Ricci curvature is symmetric.

By considering some contraction, we have

$$R = g^{\mu\nu} R_{\mu\lambda} = g^{\mu\nu} g^{\kappa\lambda} R_{\kappa\mu\lambda\nu},$$

and we call this as *Scalar curvature*.

In summary,

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

From the Bianci's identity

$$\nabla_\rho R_{\kappa\lambda\mu\nu} + \nabla_\mu R_{\kappa\lambda\nu\rho} + \nabla_\nu R_{\kappa\lambda\rho\mu} = 0,$$

we have

$$0 = g^{\mu\lambda} \left(\nabla_\rho R_{\kappa\lambda\mu\nu} + \nabla_\mu R_{\kappa\lambda\nu\rho} + \nabla_\nu R_{\kappa\lambda\rho\mu}\right)$$
$$= -\nabla_\rho \left(g^{\mu\lambda} R_{\kappa\lambda\mu\nu}\right) + \nabla^\lambda R_{\kappa\lambda\nu\rho} + \nabla_\nu \left(g^{\mu\lambda} R_{\kappa\lambda\rho\mu}\right)$$
$$= -\nabla_\rho R_{\kappa\nu} - \nabla^\lambda R_{\lambda\kappa\nu\rho} + \nabla_\nu R_{\kappa\rho}.$$

Hence we obtain

$$\nabla^\kappa R_{\kappa\lambda\mu\nu} = \nabla_\mu R_{\nu\lambda} - \nabla_\nu R_{\mu\lambda}.$$

Contract $\lambda\mu$. Then

$$-\nabla^\kappa R_{\kappa\nu} = \nabla^\lambda R_{\lambda\nu} - \nabla_\nu R.$$

Hence

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R$$

So

$$0 = \nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right).$$

From this, we may define

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

and call the Einstein curvature or tensor. Note that $G_{\mu\nu}$ has conservation law $\nabla^\mu G_{\mu\nu} = 0$.

Now we have to connect $G_{\mu\nu}$ and $T_{\mu\nu}$. We will show that

$$G_{\mu\nu} = -8\pi G_N T_{\mu\nu},$$

where G_N is the Newton constant. Note that

$$G^\mu_\mu = R - \frac{1}{2} DR = \left(1 - \frac{1}{2} D\right) R.$$

So

$$\left(1 - \frac{1}{2} D\right) R = -8\pi G T^\mu_\mu.$$

Hence

$$R = \frac{16\pi G T^\mu_\mu}{D-2}.$$

So

$$\begin{aligned} R_{\mu\nu} &= -8\pi G T_{\mu\nu} + \frac{1}{2} g_{\mu\nu} R \\ &= -8\pi G \left[T_{\mu\nu} + \frac{1}{D-2} g_{\mu\nu} T^\lambda_\lambda \right]. \end{aligned}$$

Assume gravity is weak and the situation is static(large c limit). Write

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

Its inverse is

$$g^{\mu\nu} \approx \eta^{\mu\nu} - \eta^{\mu\nu}\eta^{\nu\rho} h_{\lambda\rho}.$$

So

$$\begin{aligned} \partial_\lambda g_{\mu\nu} &= \partial_\lambda h_{\mu\nu} \\ g_{tt} = g_{00} &\approx -1 + \frac{2MG}{r} \end{aligned}$$

Then

$$\begin{aligned} R_{00} &= R^\lambda_{0\lambda 0} \\ &= \partial_\lambda \Gamma^\lambda_{0\,0} - \partial_0 \Gamma^0_{\lambda 0} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{00} - \Gamma^\lambda_{0\rho}\Gamma^\rho_{\lambda 0} \\ &= \partial_\lambda \Gamma^\lambda_{0\,0} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{00} - \Gamma^\lambda_{0\rho}\Gamma^\rho_{\lambda 0} \end{aligned}$$

since $\partial_0 \Gamma^0_{\lambda 0} = 0$. Note that

$$\Gamma^\lambda_{\lambda\nu} = \frac{1}{2} g^{\lambda\rho} \partial_\nu g_{\lambda\rho} \approx \frac{1}{2} \eta^{\lambda\rho} \partial_\nu h_{\lambda\rho}$$
$$\Gamma^\lambda_{0\nu} \approx \frac{1}{2} g^{\lambda\rho} \left(\partial_\nu h_{0\rho} - \partial_\rho h_{0\nu}\right)$$
$$\Gamma^\lambda_{00} \approx -\frac{1}{2} g^{\lambda\rho} \partial_\rho h_{00} \approx -\frac{1}{2} \eta^{\lambda\rho} \partial_\rho h_{00}.$$

From these rule, we get $R_{00} \approx -\nabla^2 h_0$. From $g_{00} \approx -1 + \frac{2MG}{r}$. we have

$$R_{00} \approx 2\nabla^2 \Phi_{\text{newton}} = -2\left(\rho 4\pi G\right)$$

where ρ is the mass density. Since $T \approx \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, we have

$$T_{00} - \frac{1}{D-2} g_{00} T^\lambda_\lambda = \frac{D-3}{D-2}\rho$$

2.6 Some solutions of Einstein's equation

Recall the Einstein equation

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G_N T_{\mu\nu}.$$

The left hand side describes the spacetime geometry and the right-hand side describes matters.

2.6.1 Weak curvature limit

If $g_{\mu\nu} \sim \eta_{\mu\nu} + h_{\mu\nu}$, then $g^{\mu\nu} \sim \eta^{\mu\nu} - \eta^{\mu\lambda}\eta^{\nu\rho} h_{\lambda\rho}$. From this, Einstein solved the problem on mercury.

2.6.2 Exact solution: Schwarzschild solution

Here we derive the most general, asymptotically flat, spherically symmetric, static, regular solution to the $D = 4$ Einstein Field Equations. We require the metric and Energy-Momentum tensor to be spherically symmetric,

$$\mathcal{L}_{\xi_a} g_{\mu\nu} = 0\,, \qquad \mathcal{L}_{\xi_a} T_{\mu\nu} = 0\,, \tag{2.7}$$

with three Killing vectors, ξ^μ_a, $a = 1, 2, 3$, corresponding to the usual angular momentum differential operators,

$$\xi_1 = \sin\varphi \partial_\vartheta + \cot\vartheta \cos\varphi \partial_\varphi\,, \qquad \xi_2 = -\cos\varphi \partial_\vartheta + \cot\vartheta \sin\varphi \partial_\varphi\,, \qquad \xi_3 = -\partial_\varphi\,. \tag{2.8}$$

They satisfy the $\mathbf{so}(3)$ commutation relation,

$$\left[\xi_a, \xi_b\right] = \sum_c \epsilon_{abc}\, \xi_c\,. \tag{2.9}$$

It follows from (2.7) that

$$g_{\vartheta\varphi} = g_{\varphi\vartheta} = 0\,, \qquad g_{\varphi\varphi} = \sin^2\vartheta g_{\vartheta\vartheta}\,, \qquad T_{\vartheta\varphi} = T_{\varphi\vartheta} = 0\,, \qquad T_{\varphi\varphi} = \sin^2\vartheta T_{\vartheta\vartheta}\,. \tag{2.10}$$

Furthermore, without loss of generality, we can put $g_{tr} = 0$, $g_{\vartheta\vartheta} = r^2$, and set the metric to be diagonal, utilizing diffeomorphisms (see *e.g.* [9]),

$$\mathrm{d}s^2 = -B(r)\mathrm{d}t^2 + A(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{2.11}$$

where we put as shorthand notation,

$$\mathrm{d}\Omega^2 := \mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\,. \tag{2.12}$$

The nonvanishing Christoffel symbols are then exhaustively,

$$\begin{aligned}
\gamma^t_{tr} = \gamma^t_{rt} &= \frac{B'}{2B}\,, & \gamma^r_{tt} &= \frac{B'}{2A}\,, & \gamma^r_{rr} &= \frac{A'}{2A}\,,\\
\gamma^r_{\vartheta\vartheta} &= -\frac{r}{A}\,, & \gamma^r_{\varphi\varphi} &= -\sin^2\vartheta\,\frac{r}{A}\,, & \gamma^\vartheta_{r\vartheta} = \gamma^\vartheta_{\vartheta r} &= \frac{1}{r}\,,\\
\gamma^\vartheta_{\varphi\varphi} &= -\sin\vartheta\cos\vartheta\,, & \gamma^\varphi_{r\varphi} = \gamma^\varphi_{\varphi r} &= \frac{1}{r}\,, & \gamma^\varphi_{\vartheta\varphi} = \gamma^\varphi_{\varphi\vartheta} &= \cot\vartheta\,,
\end{aligned}$$

and subsequently the Ricci curvature, $R_{\mu\nu}$, becomes diagonal, with components

$$\begin{aligned}
R_{tt} &= \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\frac{B'}{A}\\
R_{rr} &= -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\frac{A'}{A}\\
R_{\vartheta\vartheta} &= 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A}\\
R_{\varphi\varphi} &= \sin^2\vartheta\,R_{\vartheta\vartheta}\,.
\end{aligned} \tag{2.13}$$

Since both the Ricci curvature and the metric are diagonal, the Einstein Field Equations imply that the Energy-Momentum tensor must also be diagonal, thus fixing $T_{tr} = 0$,

$$T_{\mu\nu} = \mathrm{diag}\big(T_{tt}\,,\quad T_{rr}\,,\quad T_{\vartheta\vartheta}\,,\quad T_{\varphi\varphi} = \sin^2\vartheta\,T_{\vartheta\vartheta}\big)\,. \tag{2.14}$$

Now the conservation of the Energy-Momentum tensor, $\bigtriangledown_\mu T^\mu{}_\nu$, boils down to a single equation:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(T^r{}_r\right) + \frac{2}{r}\left(T^r{}_r - T^\vartheta{}_\vartheta\right) + \frac{B'}{2B}\left(T^r{}_r - T^t{}_t\right) = 0\,. \tag{2.15}$$

The Einstein Field Equations reduce to

$$\begin{aligned}
R_{tt} &= \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\frac{B'}{A} = -4\pi GB\left(T^t{}_t - T^r{}_r - 2T^\vartheta{}_\vartheta\right)\,,\\
R_{rr} &= -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{1}{r}\frac{A'}{A} = -4\pi GA\left(T^t{}_t - T^r{}_r + 2T^\vartheta{}_\vartheta\right)\,,\\
R_{\vartheta\vartheta} &= 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A} = -4\pi Gr^2\left(T^t{}_t + T^r{}_r\right)\,,
\end{aligned} \tag{2.16}$$

which are linearly equivalent to

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}r}\left[r\left(1 - \frac{1}{A}\right)\right] &= \frac{rA'}{A^2} + 1 - \frac{1}{A} = -8\pi Gr^2T^t{}_t\,,\\
\frac{\mathrm{d}}{\mathrm{d}r}\ln(AB) &= \frac{A'}{A} + \frac{B'}{B} = -8\pi GAr(T^t{}_t - T^r{}_r)\,,\\
\frac{B''}{B} - \frac{B'}{2B}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{1}{r}\left(\frac{A'}{A} - \frac{B'}{B}\right) &= 16\pi GAT^\vartheta{}_\vartheta\,.
\end{aligned} \tag{2.17}$$

The first equation can be integrated to give

$$\tfrac{1}{2}r\left(1-\frac{1}{A}\right)=-G\int_0^r \mathrm{d}r' 4\pi r'^2 T^t{}_t(r')\,, \tag{2.18}$$

where we have assumed 'regularity' at the origin,

$$\lim_{r\to 0} r\left(1-\frac{1}{A}\right)=0\,. \tag{2.19}$$

This fixes the function $A(r)$,

$$A(r)=\frac{1}{1-\frac{2GM(r)}{r}}\,, \tag{2.20}$$

for which we have defined

$$M(r):=-\int_0^r \mathrm{d}r' 4\pi r'^2\, T^t{}_t(r')\,. \tag{2.21}$$

The regularity condition (2.19) is then equivalent to

$$\lim_{r\to 0} M(r)=0\,. \tag{2.22}$$

Furthermore, the positive energy (density) condition implies $T^{tt}\geq 0$, such that, owing to the convention of the mostly plus signature of the metric (2.11), $M(r)$ is generically positive. Similarly, assuming the 'flat' boundary condition at infinity,

$$\lim_{r\to\infty} A(r)B(r)=1\,, \tag{2.23}$$

the second equation in (2.17) can be integrated to fix $B(r)$,

$$B(r)=\left[1-\frac{2GM(r)}{r}\right]\exp\left[8\pi G\int_r^\infty \mathrm{d}r' r' A(r')\left\{T^t{}_t(r')-T^r{}_r(r')\right\}\right]\,, \tag{2.24}$$

such that the metric takes the final form:

$$\mathrm{d}s^2=-e^{-2\Delta(r)}\left(1-\frac{2GM(r)}{r}\right)\mathrm{d}t^2+\frac{\mathrm{d}r^2}{1-\frac{2GM(r)}{r}}+r^2\mathrm{d}\Omega^2\,. \tag{2.25}$$

Here $M(r)$ is given by (2.21) and $\Delta(r)$ is defined by

$$\Delta(r):=4\pi G\int_r^\infty \mathrm{d}r'\,\frac{\left\{T^r{}_r(r')-T^t{}_t(r')\right\}r'}{1-\frac{2GM(r')}{r'}}\,. \tag{2.26}$$

Finally, we show that, up to the first and the second relations in (2.17) and their solutions, (2.20), (2.24), the third relation is equivalent to the conservation of the Energy-Momentum tensor (2.15). For this, we solve for $T^r{}_r$ and $T^r{}_r-T^t{}_t$ from the first and the second relations in (2.17),

$$\begin{aligned}T^r{}_r&=\frac{1}{8\pi G}\left(\frac{B'}{ABr}+\frac{1}{Ar^2}-\frac{1}{r^2}\right)\\ T^r{}_r-T^t{}_t&=\frac{1}{8\pi GAr}\left(\frac{A'}{A}+\frac{B'}{B}\right)\end{aligned} \tag{2.27}$$

and substitute these two expressions into the left-hand side of the conservation relation (2.15), to obtain

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}r}\left(T^r{}_r\right)+\frac{2}{r}\left(T^r{}_r-T^\vartheta{}_\vartheta\right)+\frac{B'}{2B}\left(T^r{}_r-T^t{}_t\right)\\ &=\frac{1}{8\pi GAr}\left[\frac{B''}{B}-\frac{B'}{2B}\left(\frac{A'}{A}+\frac{B'}{B}\right)-\frac{1}{r}\left(\frac{A'}{A}-\frac{B'}{B}\right)-16\pi GAT^\vartheta{}_\vartheta\right]\,.\end{aligned} \tag{2.28}$$

This result clearly establishes the equivalence between the Energy-Momentum conservation equation (2.15) and the third relation in (2.17).

Remark. Some comments are in order.

– When the matter is localized up to a finite radius r_c, such that outside this radius, $r > r_c$, we have $T_{\mu\nu}(r) = 0$ and $\Delta(r) = 0$, we recover the Schwarzschild solution, in which the mass agrees with the ADM mass [10] and from (2.21) is further given by the volume integral,

$$M = -\int_0^{r_c} \mathrm{d}r\, 4\pi r^2\, T^t{}_t(r) = -\int_0^{\infty} \mathrm{d}r\, 4\pi r^2\, T^t{}_t(r)\,. \tag{2.29}$$

However, this differs from the Noether charge (**??**) of the time translational Killing vector,

$$M \neq -Q[\partial_t] = -\int_0^{\infty} \mathrm{d}r\, 4\pi r^2\, e^{-\Delta(r)}\, T^t{}_t(r)\,, \tag{2.30}$$

since from (2.19) the integral measure is nontrivial,

$$\sqrt{-g} = e^{-\Delta(r)}\, r^2 \sin\vartheta \neq r^2 \sin\vartheta\,. \tag{2.31}$$

This discrepancy and its remedy by an extra surface integral are rather well known.

– If there is a spherical void in which $T_{\mu\nu} = 0$ for $r_1 < r < r_2$, both $M(r)$ and $\Delta(r)$ become constant inside the void as $M(r) = M(r_1)$ and $\Delta(r) = \Delta(r_2)$. After a constant rescaling of the time, $t_{\text{new}} = e^{-\Delta(r_2)}t$, the local geometry inside the void coincides precisely with the Schwarzschild solution. We note that the mass is determined through the integral over $0 < r < r_1$ only and is independent of the matter distribution outside, $r > r_2$. While this is certainly true in Newtonian gravity (namely the iron sphere theorem), if we solved the vacuum Einstein Field Equations with vanishing Energy-Momentum tensor inside the void, we would merely recover the Schwarzschild geometry in accordance with Birkhoff's theorem. Nevertheless it would be hard to conclude that the constant mass parameter is unaffected by the outer region.

– The radial derivative of $B(r)$ amounts to the gravitational acceleration for a circular geodesic [11],

$$r\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}t}\right)^2 = -\frac{1}{2}\frac{\mathrm{d}g_{tt}(r)}{\mathrm{d}r} = \tfrac{1}{2}B' = \left[\frac{GM(r)}{r^2} + 4\pi G r T^r{}_r(r)\right] e^{-2\Delta(r)}\,. \tag{2.32}$$

Again, inside a void or the outer vacuum region, we may absorb the constant factor of $e^{-2\Delta(r)}$ into the rescaled time, and thus recover the Keplerian acceleration,

$$r\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}t}\right)^2 = \frac{GM}{r^2}\,. \tag{2.33}$$

2.6.3 Spherical symmetric with Circular motion

If r is constant, and $\theta = \frac{\pi}{2}$, $\phi\,(t)$, then the orbital velocity is

$$v = r\dot{\phi} = \sqrt{\frac{MG}{r}}.$$

2.6.4 Gravitational time dilation

If $dr = d\theta = d\phi = 0$, then

$$\Delta\tau = \sqrt{1 - \frac{2MG}{r}}\Delta t < \Delta t.$$

If $r_1 < r_2$, then

$$\frac{\Delta t_2}{\Delta t_1} = \sqrt{\frac{1 - \frac{2MG}{r_1}}{1 - \frac{2MG}{r_2}}} \approx \frac{1 - \frac{MG}{r_1}}{1 - \frac{MG}{r_2}} \approx 1 + MG\left(\frac{1}{r_2} - \frac{1}{r_1}\right) < 1.$$

Hence we obtain the gravitational time dilation. The stronger grativty we have, the slower time flows. This effect is observed in experiment.

2.6.5 Kerr solution

In rotating, Kerr obtained another black-hole solution.

2.6.6 Cosmology

The theory of cosmology is to apply the Einstein equation to universe.

Fact (Cosmological principle)**.** Space is isotropy and homogeneity.

Then the metric of universe is approximated to

$$ds^2 = -dt^2 + a^2(t)\left(dx^2 + dy^2 + dz^2\right)$$

and

$$T = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{bmatrix}.$$

Put it into the Einstein's equation to get some equation. We say $a(t)$ is a scaler factor and $\frac{\dot{a}}{a}$ the Hubble 'constant'

It was observed $\dot{a} \neq 0$ and $\ddot{a} > 0$. $\dot{a} \neq 0$ implies universe expansion. It is conjecture that $\ddot{a} > 0$ in more theoratical way. It seems there is no bouncing force. Some people introduced some cosmological constant

$$R_{\mu\nu} - \frac{1}{2}\rho_{\mu\nu} = T_{\mu\nu} + \Lambda g_{\mu\nu}.$$

But this model is unsatisfactory since we must assume Λ is very small. This is called the Dark Energy Problem.

There is another problem: Dark Matter Problem. The result is not satisfactory even in nowadays.

Bibliography

[1] L. C. Evans, *Partial Differential Equations*, AMS

[2] E. M. Stein and R. Sharkarchi, *Fourier Analysis: An introduction*

[3] K. Kim, *Riemannian geometry*

[4] B. E. Hall, *Lie Groups, Lie Algebras and Representations*

[5] http://www.damtp.cam.ac.uk/user/tong/qft/qft.pdf

[6] Goldstein

[7] http://www.maths.ed.ac.uk/~jmf/Teaching/Lectures/CoV.pdf

[8] S. Angus, K. Cho and J. H. Park, "Einstein Double Field Equations," arXiv:1804.00964 [hep-th].

[9] S. Weinberg, "Gravitation and Cosmology : Principles and Applications of the General Theory of Relativity," ISBN-10: 0471925675, WILEY USA (1972).

[10] R. L. Arnowitt, S. Deser and C. W. Misner, "The Dynamics of general relativity," Gen. Rel. Grav. **40** (2008) 1997 doi:10.1007/s10714-008-0661-1 [gr-qc/0405109].

[11] S. M. Ko, J. H. Park and M. Suh, "The rotation curve of a point particle in stringy gravity," JCAP **1706** (2017) no.06, 002 doi:10.1088/1475-7516/2017/06/002 [arXiv:1606.09307 [hep-th]].