An Introduction to Mathematical Tools in General Relativity

Apimook Watcharangkool

National Astronomical Institute of Thailand (NARIT)
SHEAP2018

Contents

1 Introduction

In Newtonian gravity, gravity is a mysterious force pulling objects together and spacetime is just a non-interactive background. In this picture, the trajectory of an object in the spacetime (space is a 3-dimensional Euclidean space) under the influence of gravitational field is not a shortest path. Unlike Newtonian gravity, general relativity (GR) is a theory of spacetime and how energy and matter affect the geometry of spacetime. In GR space, and time play a crucial role in the description of gravity, and any free falling object in GR always take the shortest path. Of course such a phenomenon does not occur in Euclidean space. The type of geometry that we use in GR is Riemannian geometry (or rather psuedo-Riemannian geometry), which is what we will discuss next.

2 Smooth manifold

A topological space is a set that we know how to define a continuous function on it. A smooth manifold is a topological space M with some extra structures which allow us to define the notion of smoothness (differentiable).

2.1 Definition of manifold

Definition 2.1 *(Topological manifold) A topological space M is a manifold if it satisfies the following properties*

- *M is a Hausdorff: any two points can be distinct by two disjoint open sets.*
- *second countable: there exists a countable basis for topology.*
- *Locally Euclidean: for any $p \in M$ there exists a neighbourhood U and a map $\varphi : U \to \mathbb{R}^n$ such that φ is homeomorphic onto its image. Note that the pair (U, φ) is called a coordinate chart.*

In general, we need to put more than one chart together to get a manifold. A very clear example is the World map (Fig. 1). Since the surface of the earth is not Euclidean, one needs to cut and stretch each chart in order to put everything on a sheet of paper. Why do we need locally Euclidean? Because it is well-known to us. We have vectors, tensors, and we know how to do calculus on it!

Once we have calculus on each chart, we need to be sure that the result of calculus translates properly when we move to a nearby chart.

Definition 2.2 *A manifold M is* **smooth** *if it admits an atlas: The atlas is the collection of charts $\{(U_\alpha, \varphi_\alpha)\}$ with a property that any overlapping charts U_α, U_β, the function $\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ is a smooth function.*

The function $\varphi_\alpha \circ \varphi_\beta^{-1}$ is called transition function, which is nothing but a Jacobian matrix. However, in the case of $\mathbb{R}^n$ we have the global chart, so we just need one Jacobian matrix, while on manifolds we need new Jacobian matrix on each pair of charts.

Example 2.3 *Smooth manifolds*

- $\mathbb{R}^n$ *n-dimensional Euclidean space*
- S^n *n-dimensional sphere*
- $SO(n)$ *special orthogonal group (Lie group)*

Figure 1: *World map [National Geographic]*

2.2 Smooth function

People went through all these troublesome definitions just to define the notion of smooth function on a manifold. However, start with smooth function, one can define many other mathematical creatures e.g. vector fields, tensors, which will be the subject of the next section.

Definition 2.4 *Let M, N be m, n-dimensional manifold respectively, and $f : M \to N$ a map. If $(U_\alpha, \varphi_\alpha)$ a coordinate chart for p and (V_i, ψ_i) a chart for $f(p)$, and if $\psi_i \circ f \circ \varphi_\alpha^{-1}$ is smooth on its domain, then f is smooth.*

Exercise Find the condition for $f : M \to \mathbb{R}$ to be a smooth function.

What the definition said is that one needs coordinate charts to link a function on a manifold with a function on a Euclidean space, and then borrow the power of calculus in the Euclidean space to decide whether a function is smooth or not.

3 Things that live on manifolds

To understand vectors on manifold, one needs to know the notion of vector bundle that is one can attach a vector space on each point of the manifold, so locally the tangent bundle should look like $U \times \mathbb{R}^n$.

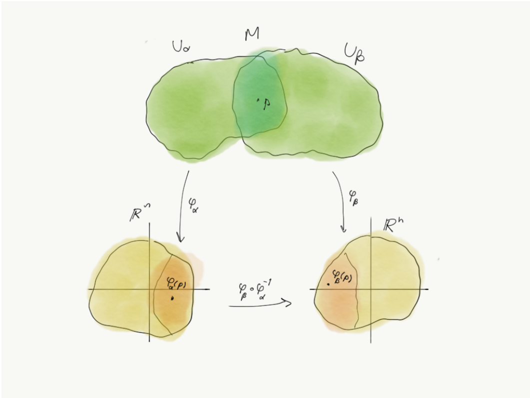

Figure 2: *smooth transition function*

3.1 Vector bundle and tangent space

Since we will not prove thing properly, it is sufficient to state an informal definition for vector bundle. A vector bundle (E, π, M), where $\pi : E \to M$ is a surjection, and E locally looks like $U \times V$, for some neighbourhood of point $p \in M$, V is a vector space. An important example of vector bundle is the tangent bundle, which can be constructed naturally on a smooth manifold.

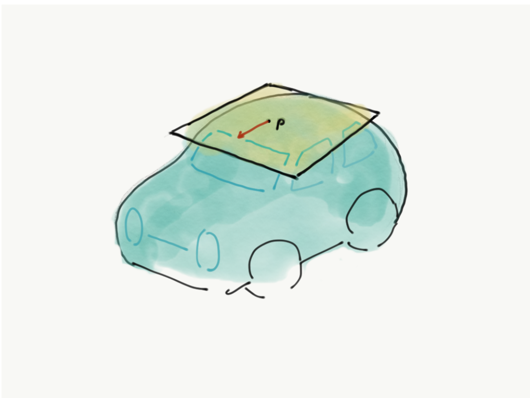

Figure 3: *tangent plane on a car like manifold*

Let $a \in \mathbb{R}^n$, $\mathbb{R}^n_a = \{(a, v); v \in \mathbb{R}^n\}$ and $T(\mathbb{R}^n)$ is the collection of maps $\tilde{v}_a : C^\infty(\mathbb{R}^n) \to$

$\mathbb{R}$. The two spaces can be identified by the following relation

$$\tilde{v}_a f = \left.\frac{d}{dt}\right|_{t=0} f(a+tv) \ . \tag{3.1}$$

From the definition we can deduce that

$$\begin{aligned}
\tilde{v}_a f &= \frac{d}{dt}\left(f(a) + tv^\mu \partial_\mu f(a) + \frac{t^2}{2!} v^\nu v^\mu \partial_\nu \partial_\mu f(a) + ...\right)_{t=0} \\
&= \left(v^\mu \partial_\mu f(a) + \frac{t}{2} v^\nu v^\mu \partial_\nu \partial_\mu f(a) + ...\right)_{t=0} \\
&= v^\mu \partial_\mu f(a) \ . \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.2)
\end{aligned}$$

It turns out that the map form $(a, v) \mapsto \tilde{v}_a$ is an isomorphism (bijective linear map), so one can conclude from Eq. (3.2) that $\{\partial_\mu\}$ is the basis for $T(\mathbb{R}^n)$.

On a manifold M, we call a linear map $X_p : C^\infty(M) \to \mathbb{R}$ derivative at $p \in M$ if it satisfies

$$X_p fg = f(p)X_p g + (X_p f)g(p) \ , \tag{3.3}$$

where $f, g \in C^\infty(M)$. The collection of this is denoted by T_pM, a tangent space (its elements are tangent vectors). One can express the basis of tangent space on the manifold with the basis of $T_a(\mathbb{R}^n)$ as follows

$$\left.\frac{\partial}{\partial x^\mu}\right|_p f = \left.\frac{\partial}{\partial x^\mu}\right|_{\varphi(p)} f \circ \varphi^{-1} \ . \tag{3.4}$$

The tangent bundle is the disjoint union of tangent space

$$TM := \coprod_{p \in M} T_pM \ , \tag{3.5}$$

such that TM is a smooth manifold. elements of TM are vector fields (or sections). A vector field X is smooth if Xf is a smooth function, and the set of smooth vector fields by is denoted by $\Gamma(TM)$.

Note that in order to get computable expression of vector, one always need to choose a local coordinate, therefore, the result of calculation is valid everywhere on the manifold if it does not depend on the choice of local chart.

3.2 Covector

The cotangent space T_p^*M is the collection of linear functional $\omega : T_pM \to \mathbb{R}$. Note that T_p^*M is also a vector space.

Example 3.1 *(Hamiltonian) Let $p \in M$ and $(q^1, ..., q^n)$ be a local coordinate. A point in the configuration space is written as $(\mathbf{q}, \dot{\mathbf{q}}) = (q^1, ..., q^n, \dot{q}^1, ..., \dot{q}^n)$. One immediately notices that this is an element of the tangent bundle, since an element is a cartesian product of coordinate and velocity vector. The Lagrangian is the function*

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} m \|\dot{\mathbf{q}}\|^2 - V(\mathbf{q}) \ , \tag{3.6}$$

so at a given point in the configuration space $L \in \mathbb{R}$. Define a canonical momentum $p_i = \partial L / \partial \dot{q}^i$, we then obtain a phase space, with an element $(\mathbf{q}, \mathbf{p}) = (q^1, ..., q^n, p_1, ..., p_n)$. From the the Legendre transformation

$$p_i \dot{q}^i = L + H \ , \tag{3.7}$$

where $H = \|\mathbf{p}\|^2/2m + V(\mathbf{q})$. So we know that $\mathbf{p}(\dot{\mathbf{q}}) \mapsto \mathbb{R}$. So $(\mathbf{q}, \mathbf{p}) \in T^*M$.

On the tangent space there is a natural inner product $g : T_pM \times T_pM \to \mathbb{R}$. This is the Riemannian metric (which we will discuss in details very soon.) Let $X_p, Y_p \in T_pM$

$$\begin{aligned} g(X_p, Y_p) &= g(X^\mu \partial_\mu|_p, Y^\nu \partial_\nu|_p) \\ &= X^\mu Y^\nu g(\partial_\mu, \partial_\nu) \\ &= X^\mu Y^\nu g_{\mu\nu} \ . \end{aligned}$$

Using the Riemannian metric, one can define covector

$$\omega_X := g(X_p, \ \cdot \ ) \ , \tag{3.8}$$

which is easy to check that it is a linear functional. From the **Riesz representation theorem**, all linear functional in an inner product space are in this form. We know that the basis of T_pM is $\{\partial_\mu\}$. The following proposition will give us a clue to find a basis for T_p^*M

Proposition 3.2 *Let $\{E_i\}$ be a basis of a vector space V. The set of linear functional $\{\varepsilon^i\}$ satisfying*

$$\varepsilon^i(E_j) = \delta^i_j \ , \tag{3.9}$$

is the basis for V^*, in particular $\dim V = \dim V^*$.

One can check that derivative of smooth function define by

$$df(X_p) = X_p f$$

is also a linear functional i.e. $df \in T_p^*M$. Since each x^μ is a smooth function, dx^μ is a covector, moreover

$$dx^\mu(\partial_\nu|_p) = \partial_\nu x^\mu(p) = \delta^\mu_\nu \ .$$

Hence from the previous proposition $\{dx^\mu\}$ is a basis for T_p^*M, and we also write

$$\omega = \omega_\mu dx^\mu \ . \tag{3.10}$$

From the Riesz's theorem, there exists a vector X such that $g(X, \ \cdot \ ) = \omega = \omega_\mu dx^\mu$, so the component of the covector is

$$\omega_\nu = g(X^\mu \partial_\mu, \partial_\nu) = X^\mu g_{\mu\nu} \ . \tag{3.11}$$

For convenient we use the same symbol but use a subscribe i.e. X_μ.

3.3 Tensors

Let V be a vector space. Tensor is a multilinear map (of rank n)

$$T : V \times V \times .... \times V \to \mathbb{R} \tag{3.12}$$

Suppose we have tensors T, and S of rank k and l respectively. Tensor product is a map $T \otimes S : \underbrace{V \times V \times ... \times V}_{k+l} \to \mathbb{R}$ defined by

$$T \otimes S(X_1, ..., X_{k+l}) = T(X_1, ..., X_k) S(X_{k+1}, ..., X_{k+l}) \tag{3.13}$$

Example 3.3 *If we take $V = T_pM$ or T_p^*M then*

- *vectors and covectors are tensors of rank $(0,1)$ and $(1,0)$ respectively.*
- *Riemannian metric g is a tensor $(2,0)$, in a local coordinate $g = g_{\mu\nu} dx^\mu dx^\nu$.*
- *The inverse metric $g^{-1} := g^{\mu\nu} \partial_\mu \otimes \partial_\nu$, s.t. $g^{\mu\nu} g_{\nu\sigma} = \delta^\mu_\sigma$, is a $(0,2)$ tensor.*

4 Stoke's Theorem

4.1 differential form

Now let us explore a special type of tensor called totally antisymmetric tensors. Let $\{X_\mu\}_{\mu=1,..,n}$ be a collection of linearly independent vector fields, consider a (2,0) tensor, which can be decomposed into symmetric and anti-symmetric parts i.e.

$$\begin{aligned} T(X_1, X_2) &= \frac{1}{2}\left(T(X_1, X_2) + T(X_2, X_1)\right) + \frac{1}{2}\left(T(X_1, X_2) - T(X_2, X_1)\right) \\ &:= \alpha(X_1, X_2) + \omega(X_1, X_2) \ , \end{aligned} \tag{4.1}$$

where α, ω are symmetric and anti-symmetric (2,0) tensor respectively. Likewise, for any (k,0) tensor, $T(X_1, ..., X_k)$ one can define a totally anti-symmetric tensor

$$\omega(X_1, ..., X_k) = \sum_\sigma \frac{\text{sign}}{k!}(\sigma) T(X_{\sigma(1)}, ..., T_{\sigma(k)}) \ . \tag{4.2}$$

The collection of such totally anti-symmetric tensor of rank k is called k-form, denoted by $\wedge^k T^* M$, or $\Omega^k(T^* M)$. Note that, a smooth function and its derivative f, df are defined as 0-form and 1-form respectively, and the maximum rank of forms is equal to the dimension of the manifold i.e. n-form.

Example 4.1 *(Wedge product) Let $\{x^\mu\}_{\nu=1,\dots,n}$ be a local coordinate, dx^μ is a $(1,0)$ tensor, so one defines*

$$dx^\mu \wedge dx^\nu = dx^\mu \otimes dx^\nu - dx^\nu \otimes dx^\mu \ . \tag{4.3}$$

Notice that, $dx^I := dx^\mu \wedge dx^\nu;\ \nu > \mu$ form a collection of linearly independent 2-froms, so there are $\begin{pmatrix} n \\ 2 \end{pmatrix}$ of them. The set $\{dx^I\}$ will be basis of 2-forms if it needs to span Ω^k. Since a 2-form is a tensor, it can be written as (not linearly independent)

$$\omega(\partial_\alpha, \partial_\beta) = \tilde{\omega}_{\mu\nu} dx^\mu \otimes dx^\nu (\partial_\alpha, \partial_\beta) \ ,$$

then using the fact that it is anti-symmetric, we have $\tilde{\omega}_{\alpha\beta} = -\tilde{\omega}_{\beta\alpha}$. The 2-form can be rewritten in the following way.

$$\begin{aligned} \tilde{\omega}_{\mu\nu} dx^\mu \otimes dx^\nu &= \frac{1}{2}(\tilde{\omega}_{\mu\nu} dx^\mu \otimes dx^\nu + \tilde{\omega}_{\nu\mu} dx^\nu \otimes dx^\mu) \\ &= \frac{\tilde{\omega}_{\mu\nu}}{2}(dx^\mu \otimes dx^\nu - dx^\nu \otimes dx^\mu) \\ &= \omega_{\mu\nu} dx^\mu \wedge dx^\nu \ . \end{aligned} \tag{4.4}$$

We can generalise this example to any k-form.

$$\omega = \omega_{\mu\nu\dots\sigma} dx^\mu \wedge dx^\nu \wedge \dots \wedge dx^\sigma \ , \tag{4.5}$$

The operation $\wedge : \Omega^k \times \Omega^l \to \Omega^{k+l}$, for $k + l \leq n$, is called **the wedge product** defined by

$$\omega \wedge \eta = \omega_{\mu_1\dots\mu_k} \eta_{\nu_1\dots\nu_l} (dx^{\mu_1} \wedge \dots \wedge dx^{\mu_k}) \wedge (dx^{\nu_1} \wedge \dots \wedge dx^{\nu_l}) \ . \tag{4.6}$$

The rank of forms is limited by dimension of manifold, the algebra of form defined by

$$\wedge^\bullet T^* M = \bigoplus_{k=0}^{n} \Omega^k(T^* M) \tag{4.7}$$

is called the exterior algebra, which is a finite dimensional algebra with $\dim \wedge^{\bullet} T^*M = 2^n$. The highest, or the top form is n-form, written as follow

$$\omega_{1...n} dx^1 \wedge ... \wedge dx^n \ . \tag{4.8}$$

Note that for any orthonormal basis $\{E_1, ..., E_n\}$ there is only one n-form, denoted by dV or Ω, such that $\Omega(E_1, ..., E_n) = 1$. This is the "volume form" on a manifold.

Exercise 4.2 *Show that, for a Riemannian manfold* $\Omega_{1...n} = \sqrt{\det g_{ij}}$.

If we have a volume form, one can define integration on a coordinate chart as

$$\int_U \omega = \int_{\varphi(U)} \omega_{12...n} dV \ , \tag{4.9}$$

where dV is a volume in $\mathbb{R}^n$. To obtain the integral over a manifold, one needs to sum over all charts modulo the overlap area (use partition of unity).

4.2 exterior derivative and interior product

Let $\omega \in \Omega^k$, we define a map $d : \Omega^k \to \Omega^{k+1}$ on a local coordinate as follow

$$d\omega = \left(\frac{\partial}{\partial x^\rho} \omega_{\mu\nu...\sigma} \right) dx^\rho \wedge dx^\mu \wedge dx^\nu \wedge ... \wedge dx^\sigma \tag{4.10}$$

Exercise 4.3 *Show that* d *is dempotemp operator ,i.e. for any* $\omega \in \Omega^k, d^2\omega = 0$.

One can also define a map that decrease the rank of p-form called the interior product. Let $V \in TM$, and $\omega \in \Omega^p(TM)$

$$i_V \omega(X_1, ..., X_p) = \omega(V, X_1, ..., X_p) \ . \tag{4.11}$$

This operator will play a part in Divergence theorem. Let consider a simple integral of a one-form df over a curve C connecting points p and q. We know that

$$\int_C df = f(q) - f(p) = \int_{\partial C} f \ .$$

We may think of this as an equality between pairing (similar to inner product) of one-form and a one-dimensionl submanifold with the zero-form and the zero-dimensional submanifold i.e. $\langle df, C \rangle = \langle f, \partial C \rangle$. Actually, this fact is true for all p-form, we call this the Stoke's theorm.

Theorem 4.4 *(Stoke's Theorem)*
If $\omega \in \Omega^p(T^*M)$ *and* N *is a* $p+1$*-dimensional submanifold, then*

$$\langle d\omega, N \rangle = \langle \omega, \partial N \rangle \tag{4.12}$$

5 Riemannian curvature

The special feature of Riemannian manifold is that one can always find the distance between any two points. The distance is defined in the similar manner as in the Euclidean space i.e. integrate over the norm of the velocity vector.

5.1 Infinitesimal and Riemannian metric

In Euclidean space, the distance function between $p, q \in \mathbb{R}^n$ along a curve $\gamma : [0,1] \to \mathbb{R}^n$ is

$$s(p,q) = \int_\gamma \sqrt{(dx^1(\gamma))^2 + ...(dx^n(\gamma))^2} = \int_0^1 dt \underbrace{\sqrt{(\dot{x}^1(t))^2 + ...(\dot{x}^n(t))^2}}_{\|\mathbf{v}\|} . \tag{5.1}$$

Consider the distance δs between p, q is sufficiently small such that $\|\mathbf{v}\|$ is constant, so that

$$\delta s^2 \sim \delta t^2 \left(\left(\frac{\delta x^1}{\delta t}\right)^2 + ... \left(\frac{\delta x^n}{\delta t}\right)^2 \right) = (\delta x^1)^2 + ... + (\delta x^n)^2 . \tag{5.2}$$

We call δs, or rather ds an infinitesimal. In a Riemannian manifold, the ds^2 is given by the local expression of a metric tensor

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu . \tag{5.3}$$

Thus, for a given coordinate, if one knows the metric tensor on that manifold, then in principle one can calculate the distance between any two points on that manifold.

Suppose $M = \{(x,y,z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = R^2\}$ is a subset of $\mathbb{R}^3$. Since M is in $\mathbb{R}^3$, it inherits the infinitesimal distance from $\mathbb{R}^3$ i.e.

$$ds^2 = dx^2 + dy^2 + dz^2 . \tag{5.4}$$

Note that the coordinates x, y and z are on all independent on the sphere, i.e. $xdx + ydy + zdz = 0$, and therefore,

$$dz = -\frac{xdx + ydy}{(R^2 - x^2 - y^2)^{1/2}} . \tag{5.5}$$

The infinitesimal distance on the sphere reads

$$ds^2 = \left(\frac{R^2 - y^2}{R^2 - x^2 - y^2}\right) dx^2 + \left(\frac{2xy}{R^2 - x^2 - y^2}\right) dxdy + \left(\frac{R^2 - x^2}{R^2 - x^2 - y^2}\right) dy^2 . \tag{5.6}$$

Hence in this example the metric (in the matrix form) reads

$$g = \frac{1}{R^2 - x^2 - y^2} \begin{pmatrix} R^2 - y^2 & xy \\ xy & R^2 - x^2 \end{pmatrix} \tag{5.7}$$

Exercise 5.1 *Show that if we let $x = R\sin\theta\cos\phi$, and $y = R\sin\theta\sin\phi$, then the metric will take a form*

$$g = R^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \tag{5.8}$$

5.2 Connection and curvature

The connection a way to define derivative for vector field. Why are we interested in finding derivative of vector field in the first place? Because we want to know acceleration of curves on manifold.

Definition 5.2 *An affine (or linear) connection is a map* $\nabla : TM \times \Gamma(TM) \to \Gamma(TM)$ *satisfying*

- $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$, *for* $f, g \in C^\infty(M)$
- $\nabla_X(aZ_1 + bZ_2) = a\nabla_X Z_1 + b\nabla_X Z_2$, *for* $a, b \in \mathbb{R}$
- $\nabla_X(fZ) = f\nabla_X Z + (Xf)Z$, *for* $f \in C^\infty(M)$

Let $\{x^\mu\}$ be a coordinate basis at point $p \in M$. Since the range of ∇ is in $\Gamma(TM)$

$$\nabla_{\partial_\mu}\partial_\nu = \Gamma^\sigma_{\mu\nu}\partial_\sigma \ , \tag{5.9}$$

where $\Gamma^\sigma_{\mu\nu}$ is a smooth function called **Christoffel symbol**, and for a smooth vector field $V \in \Gamma(TM)$

$$\begin{aligned}\nabla_{\partial_\mu} V^\nu \partial_\nu &= (\partial_\mu V^\nu)\partial_\nu + \Gamma^\sigma_{\mu\nu} V^\nu \partial_\sigma \\ &= (\partial_\mu V^\sigma + \Gamma^\sigma_{\mu\nu} V^\nu)\partial_\sigma \ .\end{aligned} \tag{5.10}$$

Proposition 5.3 *Every manifold admits an affine connection*

A given connection on TM can be extended on tensor bundle $T^k_l M$ such that

- $\nabla_{\partial_\mu}\omega_\nu = \partial_\mu \omega_\sigma - \Gamma^\sigma_{\mu\nu}\omega_\sigma$
- $\nabla_{\partial_\mu} T^{\nu_1 \dots \nu_k}{}_{\sigma_1 \dots \sigma_l} = \partial_\mu T^{\nu_1 \dots \nu_k}{}_{\sigma_1, \dots, \sigma_l} + \sum_{i=1}^k \Gamma^{\nu_i}_{\mu\alpha} T^{\nu_1 \dots \alpha \dots \nu_k}{}_{\sigma_1 \dots \sigma_l} - \sum_{i=1}^l \Gamma^\alpha_{\mu\sigma_i} T^{\nu_1 \dots \nu_k}{}_{\sigma_1 \dots \sigma_i \dots \sigma_l}$

Obviously the choice of connection is not unique on a Manifold; each choice of Christoffel symbol gives rise to different connection. However, In GR we are interested in a special type of connection called **Levi-Civita** connection.

Theorem 5.4 *Let* (M, g) *be a Riemannian (or pseudo-Riemannian) manifold. There exists unique affine connection that is* **metric compatible** *and* **torsion free***.*

We call this connection the **Levi-Civita** connection. We shall now describe the meaning of the words metric compatible and torsion free. Metric compatibility is the generalisation of the property of derivative on Euclidean space which is compatible with the inner product i.e. for $V, W \in \mathbb{R}^n$

$$\partial_i(V \cdot W) = (\partial_i V) \cdot W + V \cdot (\partial_i W) \ , \tag{5.11}$$

so we require that the Levi-Civita connection to be compatible with the metric

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \ , \tag{5.12}$$

for $X, Y, Z \in \Gamma(TM)$. A connection is torsion free if it satisfies a condition

$$\nabla_X Y - \nabla_Y X = [X, Y] \ . \tag{5.13}$$

The Christoffel symbol of the Levi-Civita connection can be written in terms of metric tensor as follow

$$\Gamma^{\sigma}_{\mu\nu} = \frac{g^{\sigma\alpha}}{2}(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}) \ . \tag{5.14}$$

Note that, for convenient, we will write $\nabla_{\partial_\mu} := \nabla_\mu$.

Exercise Show that (i) Eq. (5.12) and (ii) Eq. (5.13) lead to

(i). $\nabla_\mu g_{\alpha\beta} = 0$

(ii). $\Gamma^{\sigma}_{\mu\nu} = \Gamma^{\sigma}_{\nu\mu}$.

5.3 Curvature from acceleration

For mathematician, curvature is a local invariant that distinguishes one Riemannian manifold from another. However, physicists are more interested in the dynamics of objects moving in Riemannian manifold. To understand the notion of curvature, let us start with curves in 2-dimensional space. Suppose we have a circle of radius R and $\gamma : [0, 1] \to S^1$, is a curve with unit velocity i.e. $\|\dot{\gamma}(t)\| = 1$, the curvature at a point $p = \gamma(t_0)$ is defined by $\kappa(t_0) = \|\ddot{\gamma}(t_0)\|$. From classical mechanics we know that

$$\|\ddot{\gamma}(t)\| = \frac{\|\dot{\gamma}\|^2}{R} = \frac{1}{R} \ . \tag{5.15}$$

For more general curves, the curvature can be computed by attaching a circle with appropriate radius to the curve. Note that in this case, the curvature is quite easy to calculate since the manifold (curve) is embedded inside the higher dimension manifold ($\mathbb{R}^2$). However, one can obtain intrinsic (no embedding require) definition of curvature using the notion of **parallel transport**.

Parallel transport is the way to transport a vector along vector fields such that there is no acceleration. Let $Z \in \Gamma(TM)$ be a vector field, and $\gamma : [0, 1] \to M$ be a curve. A vector field is parallel transport along a curve γ if

$$\nabla_{\dot{\gamma}(t)} Z = 0 \ , \tag{5.16}$$

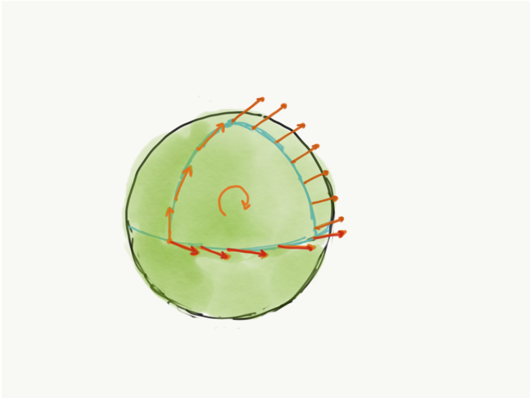

Figure 4: *parallel transport on a shpere*

In other words, the vector field is constant with respect to the velocity of the curve. In particular, we may consider the parallel transport of the velocity vector itself from p to $q \in M$. We may express the velocity in coordinate basis $\dot{\gamma} = (\dot{x}^1, ..., \dot{x}^n)$

$$\begin{aligned} 0 &= \nabla_{\dot{\gamma}}\dot{\gamma} = \nabla_{\dot{x}^\nu \partial_\nu} \dot{x}^\mu \partial_\mu \\ &= \dot{x}^\nu \nabla_\nu (\dot{x}^\mu \partial_\mu) = \dot{x}^\nu (\partial_\nu (\dot{x}^\mu)\partial_\mu + \dot{x}^\mu \Gamma^\sigma_{\nu\mu} \partial_\sigma) \\ &= (\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu} \dot{x}^\mu \dot{x}^\nu)\partial_\sigma \ . \end{aligned}$$

The equation

$$\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = 0 \ , \tag{5.17}$$

is called the geodesic equation. Think of parallel transport in Euclidean space, the trajectory of the parallel transport vector is the straight line which is the shortest path between any two points. Likewise, the solution of of this equation is the shortest curve connecting p and q in a curve space. In general, if a vector Z_p is parallel transported along a closed curve back to its starting point, then one may obtain a new vector $\tilde{Z}_p$ that is different from the original one. The infinitesimal difference between these vectors give rise to a linear map which we call curvature tensor. For $X, Y \in \Gamma(TM)$, the map $R(X,Y) : \Gamma(TM) \to \Gamma(TM)$ is a smooth linear map defined by

$$R(X,Y)Z := [\nabla_X, \nabla_Y]Z + \nabla_{[X,Y]}Z \ . \tag{5.18}$$

Suppose we choose X, Y to be basis vectors

$$
\begin{aligned}
R_{\mu\nu\rho}{}^{\sigma} Z^\rho \partial_\sigma := {} & R(\partial_\mu, \partial_\nu)^\sigma_\rho Z^\rho \partial_\sigma \\
= {} & \nabla_\mu \nabla_\nu (Z^\rho \partial_\rho) - \nabla_\nu \nabla_\mu (Z^\rho \partial_\rho) + \nabla_{[\partial_\mu, \partial_\nu]} Z^\rho \partial_\rho \\
= {} & \nabla_\mu (\partial_\nu Z^\rho \partial_\rho + Z^\rho \Gamma^\sigma_{\nu\rho} \partial_\sigma) - \nabla_\nu (\partial_\mu Z^\rho \partial_\rho + Z^\rho \Gamma^\sigma_{\mu\rho} \partial_\sigma) \\
= {} & (\partial_\mu \partial_\nu Z^\rho) \partial_\rho + \partial_\nu Z^\rho \Gamma^\sigma_{\mu\rho} \partial_\sigma + \partial_\mu Z^\rho \Gamma^\sigma_{\nu\rho} \partial_\sigma \\
& + Z^\rho (\partial_\mu \Gamma^\sigma_{\nu\rho}) \partial_\sigma + Z^\rho \Gamma^\sigma_{\nu\rho} \Gamma^\alpha_{\mu\sigma} \partial_\alpha \\
& - (\partial_\nu \partial_\mu Z^\rho) \partial_\rho - \partial_\mu Z^\rho \Gamma^\sigma_{\nu\rho} \partial_\sigma - \partial_\nu Z^\rho \Gamma^\sigma_{\mu\rho} \partial_\sigma \\
& - Z^\rho (\partial_\nu \Gamma^\sigma_{\mu\rho}) \partial_\sigma - Z^\rho \Gamma^\sigma_{\mu\rho} \Gamma^\alpha_{\nu\sigma} \partial_\alpha \\
= {} & (\partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\sigma_{\mu\alpha} \Gamma^\alpha_{\nu\rho} - \Gamma^\sigma_{\nu\alpha} \Gamma^\alpha_{\mu\rho}) Z^\rho \partial_\sigma \ ,
\end{aligned} \tag{5.19}
$$

so we have

$$
R_{\mu\nu\rho}{}^{\sigma} = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\sigma_{\mu\alpha} \Gamma^\alpha_{\nu\rho} - \Gamma^\sigma_{\nu\alpha} \Gamma^\alpha_{\mu\rho} \ . \tag{5.20}
$$

5.4 Properties of curvature tensor

Proposition 5.5 *The curvature tensor has the following properties*

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

The third property is called the algebraic Bianchi identity (or the first Bianchi identity). From curvature tensor, one can define Ricci tensor $R_{\mu\nu}$, and Ricci scalar (or scalar curvature) R as

$$
R_{\mu\nu} := g^{\rho\sigma} R_{\rho\mu\nu\sigma} \quad , \quad R := g^{\mu\nu} R_{\mu\nu} \ . \tag{5.21}
$$

We will se shortly that these two quantities play a crucial role in Einstein equation.

Exercise Show that $R_{\mu\nu}$ is symmetric tensor.

Proposition 5.6 *(Differential Bianchi identity) the total derivative of curvature tensor satisfies the following property*

$$
\nabla_\alpha R_{\mu\nu\rho\sigma} + \nabla_\rho R_{\mu\nu\sigma\alpha} + \nabla_\sigma R_{\mu\nu\alpha\rho} = 0 \ . \tag{5.22}
$$

Contract the differential Bianchi identity with the metric one obtains

$$
\begin{aligned}
0 &= g^{\mu\sigma} g^{\alpha\nu} (\nabla_\alpha R_{\mu\nu\rho\sigma} + \nabla_\rho R_{\mu\nu\sigma\alpha} + \nabla_\sigma R_{\mu\nu\alpha\rho}) \\
&= g^{\mu\sigma} (\nabla_\alpha R_{\rho\sigma\mu}{}^{\alpha} - \nabla_\rho R_{\mu\sigma} + \nabla_\sigma R_{\mu\rho}) \\
&= 2 \nabla_\alpha R_\rho{}^\alpha - \nabla_\rho R \\
&= 2 g_{\rho\beta} \nabla_\alpha (R^{\beta\alpha} - \frac{1}{2} g^{\beta\alpha} R) \ ,
\end{aligned} \tag{5.23}
$$

so we have that $G^{\mu\nu} := R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R$ has a vanishing divergence.

6 Einstein equation

Einstein equation is the equation that describes how mass and energy distort the curvature (which leads to acceleration) of spacetime.

6.1 Energy-Momentum tensor

An energy-momentum tensor $T^{\mu\nu}$ is a symmetric tensor defined by

- T^{00} energy density ρ
- T^{0i} energy flux through surface normal to x^i
- T^{ij} momentum flux in direction of x^i through surface normal to x^j

Let us look at an example of energy momentum tensor

Example 6.1 *Energy-momentum tensor*

- *Electromagnetic field*
$$T^{\mu\nu} = F^{\mu}_{\ \alpha}F^{\alpha\nu} - \frac{1}{4}g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta} \ , \tag{6.1}$$
where $F_{\mu\nu}$ is the field strength tensor.
- *Perfect fluid (which will be important when we start doing cosmology.)*
$$T^{\mu\nu} = (\rho + P)u^{\mu}u^{\nu} + Pg^{\mu\nu} \ , \tag{6.2}$$
where ρ is energy density, P is pressure and u is a normalised timelike 4-velocity i.e. $u^2 = -1$.

One can check that the two energy momentum tensors are divergence free (given that the matter fields obey their equation of motion.) Since Einstein tensor and energy momentum tensor are both divergence free, we can put

$$G^{\mu\nu} = \kappa T^{\mu\nu} \ , \tag{6.3}$$

where κ is some (dimensionful) constant. The value of κ can be determined when consider Newtonian limit i.e. $\kappa = 8\pi G$.

6.2 Einstein Hilbert action

There is an alternative way of deriving Einstein equation using an action functional (the map from the vector space of function to real number). The action is called **Einstein-Hilbert** action

$$S = \int_M dx^4 \sqrt{-g}(\frac{1}{\kappa}R + \mathcal{L}_m) \ . \tag{6.4}$$

Vary Eq. (6.4) with respect to the metric

$$0 = \delta S = \int_M dx^4 [\delta\sqrt{-g}(\frac{1}{\kappa}R + \mathcal{L}_m) + \sqrt{-g}(\frac{1}{\kappa}\delta R + \delta\mathcal{L}_m) \ . \tag{6.5}$$

To find the expression of $\delta\sqrt{-g}$ in terms of $\delta g^{\mu\nu}$, let us consider a symmetric matrix M (so that it is diagonalisable), and let $C^{-1}MC = D = \mathrm{diag}(D_{11}, D_{22}, ..., D_{nn})$. Suppose t is a very small parameter such that we can ignore $t^n, n \geq 2$

$$\begin{aligned}
\det(\mathbb{1} + tM) = & \det(C^{-1}C + tC^{-1}DC) \\
= & \det(\mathbb{1} + tD) \\
= & (1 + tD_{11})(1 + tD_{22})...(1 + tD_{nn}) \\
= & 1 + t\sum_{i=1}^{n} D_{ii} + t^2 \sum_{i \neq j}^{n} D_{ii}D_{jj} + ... \\
\approx & 1 + \mathrm{tr}(tD) = 1 + \mathrm{tr}(tM) \ .
\end{aligned} \tag{6.6}$$

Hence for determinant of the metric

$$\begin{aligned}
\delta g = & \det(g_{\mu\nu} + \delta g_{\mu\nu}) - g \\
= & g[\det(g^{\alpha\mu})\det(g_{\mu\nu} + \delta g_{\mu\nu}) - 1] \\
= & g[\det(\delta^\alpha_\nu + g^{\alpha\mu}\delta g_{\mu\nu}) - 1] \\
\approx & g \ \mathrm{tr}(g^{\alpha\mu}\delta g_{\mu\nu}) \\
= & g g^{\mu\nu}\delta g_{\mu\nu} = -g g_{\mu\nu}\delta g^{\mu\nu} \ .
\end{aligned} \tag{6.7}$$

Put this back in Eq. (6.5) we obtain

$$\begin{aligned}
0 = & \int_M dx^4 [-\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}(\frac{1}{\kappa}R + \mathcal{L}_m) + \sqrt{-g}(\frac{1}{\kappa}\delta g^{\mu\nu}R_{\mu\nu} + \frac{1}{\kappa}g^{\mu\nu}\delta R_{\mu\nu} + \delta\mathcal{L}_m) \\
= & \int_M d^4x\sqrt{-g}[\frac{1}{\kappa}(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R)\delta g^{\mu\nu} + (\frac{\delta\mathcal{L}_m}{\delta g^{\mu\nu}} - \frac{1}{2}g_{\mu\nu}\mathcal{L}_m)\delta g^{\mu\nu} + \frac{1}{\kappa}g^{\mu\nu}\delta R_{\mu\nu}] \ .
\end{aligned} \tag{6.8}$$

Notice that if we define

$$T^{\mu\nu} = -\frac{\delta\mathcal{L}_m}{\delta g_{\mu\nu}} + \frac{1}{2}g^{\mu\nu}\mathcal{L}_m \ , \tag{6.9}$$

and $g^{\mu\nu}\delta R_{\mu\nu}$ somehow vanishes then we will obtain Einstein equation. Let us consider $g^{\mu\nu}\delta R_{\mu\nu}$.

$$\begin{aligned}
g^{\mu\nu}\delta R_{\mu\nu} = \; & g^{\mu\nu}\Big(\partial_\alpha \delta\Gamma^\alpha_{\mu\nu} - \partial_\mu \delta\Gamma^\alpha_{\alpha\nu} + \delta\Gamma^\alpha_{\alpha\beta}\Gamma^\beta_{\mu\nu} + \Gamma^\alpha_{\alpha\beta}\delta\Gamma^\beta_{\mu\nu} \\
& - \delta\Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\alpha\nu} - \Gamma^\alpha_{\mu\beta}\delta\Gamma^\beta_{\alpha\nu}\Big) \\
= \; & g^{\mu\nu}\Big(\partial_\alpha \delta\Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\mu\beta}\delta\Gamma^\beta_{\alpha\nu} - \Gamma^\beta_{\alpha\nu}\delta\Gamma^\alpha_{\mu\beta} + \Gamma^\alpha_{\alpha\beta}\delta\Gamma^\beta_{\mu\nu} \\
& - \partial_\mu \delta\Gamma^\alpha_{\alpha\nu} + \Gamma^\beta_{\mu\alpha}\delta\Gamma^\alpha_{\beta\nu} + \Gamma^\beta_{\mu\nu}\delta\Gamma^\alpha_{\alpha\beta} - \Gamma^\beta_{\mu\alpha}\delta\Gamma^\alpha_{\beta\nu}\Big) \\
= & \left(\nabla_\alpha g^{\mu\nu}\delta\Gamma^\alpha_{\mu\nu} - \nabla_\mu g^{\mu\nu}\delta\Gamma^\alpha_{\alpha\nu}\right) \; ,
\end{aligned} \tag{6.10}$$

which is the boundary term, therefore, using divergence theorem, the integral of (6.10) vanishes. Hence, we obtain Einstein equation.