A short note to Riemann Manifold

Yuandong Tian

December 20, 2009

1 Short Definition

A manifold M is like a curve in 2D and a surface in 3D. It is a set of points with topology (or neighbor structures), and locally looks like $\mathbb{R}^n$.

Example: our earth. Locally it is planar.

Rigidly, for a patch $U \subset M$, we have a local coordinate system $x_U^i : U \mapsto \mathbb{R}$ as $i = 1 \cdot n$ local coordinates. Typically M can be covered by several such patches. (eg, A sphere can be covered by 2 patches but *not one*). In general, we omit the subscript U for clarity.

Tangent space Given any point $p \in M$, it has a *tangent space* T_pM isometric to $\mathbb{R}^n$. If we have a metric (inner-product) in this space $< \cdot, \cdot >_p$: $T_pM \times T_pM \mapsto \mathbb{R}$ defined on every point $p \in M$, we thus call M *Riemann Manifold*.

Usually (except for general relativity) the manifold M is embedded in an *ambient space*. For example, a 2D curve is embedded in $\mathbb{R}^2$; our earth is embedded in $\mathbb{R}^3$. In such cases, the tangent spaces are subspaces of the ambient space and we can have an induced inner product from ambient space. Riemann geometry aims to study the property of M without the help of ambient space (which is the typical idea of mathematicians); however, a better way to understand many important concepts is to start with a manifold with its ambient space.

The union of tangent space over all the points in M is called the tangent bundle TM.

A vector $\mathbf{v} \in T_pM$ can be written as

$$\mathbf{v} = \sum_i v^i \frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x^i} = \sum_i v^i \boldsymbol{\partial}_i \tag{1}$$

where v^i is its coefficients under the coordinate x^i. The symbol $\boldsymbol{\partial}_i \equiv \boldsymbol{\partial}/\boldsymbol{\partial} x_i$ serves as the basis vectors of the tangent space. Intuitively, $\boldsymbol{\partial}_i$ points a direction that increases the corresponding coordinate x^i. Note ∂_i is also the partial derivative operator if we write $\partial_i f$, which is $\partial_i f = \partial f/\partial x^i$. In the following, we also use $f_{,i}$ as $\partial_i f$.

Note: as wee shall see, **A vector is not the same things as its coefficients.**

Similarly, we can define the vector field $\mathbf{v}$ that maps each point $p \in M$ on its T_pM.

The inner product of two vectors $\mathbf{v}, \mathbf{w} \in T_pM$ can be written as

$$\begin{aligned} < \mathbf{v}, \mathbf{w} >_p &= < \sum_i v^i \boldsymbol{\partial}_i, \sum_j w^j \boldsymbol{\partial}_j > & (2) \\ &= \sum_i \sum_j v^i w^j < \boldsymbol{\partial}_i, \boldsymbol{\partial}_j >_p & (3) \\ &= \sum_i \sum_j v^i w^j g_{ij} & (4) \end{aligned}$$

where $g_{ij}(p) =< \boldsymbol{\partial}_i, \boldsymbol{\partial}_j >_p$ is the coefficients of Riemann metric under the coordinate x^i. Using Einstein summation notation, we can just write it as $v^i w^j g_{ij}$.

In differential geometry, $v^i w^j g_{ij}$ is also called First Fundamental Form. It measures the local length/angle of the tangent vectors.

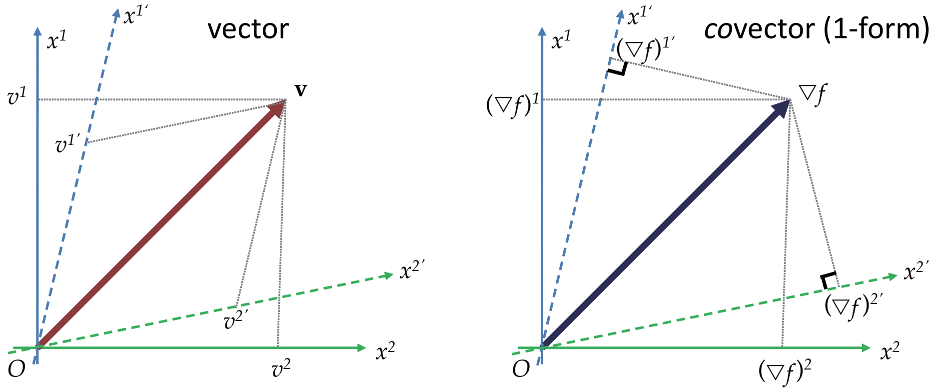

Figure 1: As a covector, the function gradient grad$f$ transforms differently from a vector. This is because in a certain coordinate frame, the representation of a vector is determined by the quadrilateral rule, while the representation of a covector is determined by its "projection" on frame bases.

2 Tensors and Exterior Forms

2.1 Tensors

We often think

- a_i as a vector because it has one subindex.
- a_{ij} as a matrix because it has two subindices.
- a_{ijk} as a tensor because it has three or more subindices.

That is incorrect in manifold.

Object and its representation Consider a vector pointing to some direction, if we change the coordinates, its representation (each element) will change, but its pointing direction does not change. Similarly, for the operator of an inner product, its elements could change when the coordinate changes, but given two vectors, their inner product is invariant under coordinate changes.

The choice of coordinate should not affect the intrinsic structure of the problem.

A tensor is such an object.

A vector $\mathbf{v}$ is a tensor with its representation v^i; the inner product $< \cdot, \cdot >_p$ is a tensor with its representation g_{ij}. As expected, the representation will change when the coordinate changes:

$$v^{j'} = v^i \frac{\partial x'_j}{\partial x_i} \quad (5)$$

$$g'_{kl} = g_{ij} \frac{\partial x_i}{\partial x'_k} \frac{\partial x_j}{\partial x'_l} \quad (6)$$

Did you notice how clever the notation is — That's amazing.

According to the transformation, the vector $\mathbf{v}$ (or its component v^i) is called first-order *contravariant tensor*. The metric (or its component g_{ij}) is called second-order *covariant tensor*.

Note according to this criterion, the function gradient grad$f = [\frac{\partial f}{\partial x_i}]$ is *not a vector* because it follows the covariant rules:

$$\frac{\partial f}{\partial x'_i} = \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial x'_i} \quad (7)$$

Fig. 1 shows how this could happen. Actually gradient is a linear *operator* acting on a direction $\mathbf{v}$ to obtain the directional derivative of f, written as $\nabla_{\mathbf{v}} f$, or $\mathbf{v}(f)$. Since $\mathbf{v}(f)$ should not change when the coordinate changes, the transformation rule of function gradient is the opposite to the rule for vectors.

The bases $\boldsymbol{\partial\partial} x_i$ is itself a covariant tensor. By writting Eqn. 1 The symbolic representation of $\mathbf{v}$ will indeed give the symbolic invariance of $\mathbf{v}$ under transformation. Similarly, we define $\mathrm{d}x^i$ to be the bases of first-order covariant tensors, or 1-forms. The function gradient $\mathrm{d}f$ thus can be written as a linear combination of $\mathrm{d}x^i$:

$$\mathrm{d}f = \frac{\partial f}{\partial x_i}\mathrm{d}x^i = f_{,i}\mathrm{d}x^i \tag{8}$$

with $f_{,i} \equiv \frac{\partial f}{\partial x_i}$. The concides with the "function differential" as defined in the elementary calculus; however, all the symbols have new (and) deeper meanings.

Note $\mathrm{d}x^i$ itself is a contravariant tensor. since as a function differential, by the chain rule, $\mathrm{d}x^{j'}$ in the new coordinate x' can be written as

$$\mathrm{d}x^{j'} = \frac{\partial x^{j'}}{\partial x^i}\mathrm{d}x^i \tag{9}$$

when doing coordinate transformation $x' = x'(x)$. So $\mathrm{d}f$ is also a symbolic representation invariant to coordinate transform.

The metric tensor now can be written similarly as

$$g = g_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{10}$$

so that it takes two vectors $\mathbf{v}$ and $\mathbf{w}$, and gives a scalar $g(\mathbf{v},\mathbf{w})$. Sometimes we write $\mathrm{d}x^i \otimes \mathrm{d}x^j$ instead of $\mathrm{d}x^i\mathrm{d}x^j$. Conceptually, one can regard $\mathrm{d}x^i$ as an operator to "extract" i-th component out of a vector.

2.2 Forms as functionals in T_pM

You may have noticed that there are operations defined on the tensors. Specifically, *if we multiply a first-order contravariant tensor with a 1-form, we get a scalar, which is zero-order tensor.* Symbolically, if we have tensor v^i and w_i, we get the scalar $v^i w_i$.

This is called tensor contraction (Matlab doesn't have such operations, what a pity!). Under this operation, the 1-form can be regard as the mapping $T_pM \mapsto \mathbb{R}$ (Remember that T_pM contains all the first-order contravariant tensors, or vectors), or $TM \mapsto C(M)$, where $C(M)$ contains all the function defined in M (Don't ask me about the continuity...). Similarly we can think n-forms this way.

So for a 1-form α, we can write $\alpha(\mathbf{v})$ which gives a scalar. And we define

$$\mathrm{d}x^i\left(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial} x_j}\right) = \mathrm{d}x^i(\boldsymbol{\partial}_j) \equiv \delta^i_j \tag{11}$$

so that $\mathrm{d}f(\mathbf{v})$ gives a scalar which is the directional gradient of f along the direction of $\mathbf{v}$. Using this definition, 2-form g can take two vectors and compute their inner products invariant to coordinate x^i.

The set of all 1-forms at all points of M are called the cotangent bundle T^*M. (Very important in Hamitonian mechanics.)

2.3 Exterior Form

Given the bases $\mathrm{d}x^i$, we thus build high-order forms using the following *wedge operator* as the building blocks:

$$(\mathrm{d}x^i \wedge \mathrm{d}x^j)(\mathbf{v},\mathbf{w}) \equiv \mathrm{d}x^i(\mathbf{v})\mathrm{d}x^j(\mathbf{w}) - \mathrm{d}x^i(\mathbf{w})\mathrm{d}x^j(\mathbf{v}) \tag{12}$$

The exterior forms are linear combination of such blocks. All exterior forms are anti-symmetric, i.e., if we swap two of its arguments, it negates; if two of its arguments are identical, it gives zero.

Then we can define a linear exterior differentiation operator d on exterior p-forms, to make it become a $p+1$-form:

$$\begin{aligned} \mathrm{d}f &\equiv f_i\mathrm{d}x_i && (13)\\ \mathrm{d}(\alpha\wedge\beta) &\equiv \mathrm{d}\alpha\wedge\beta + (-1)^p\alpha\wedge\mathrm{d}\beta && (14)\\ \mathrm{d}(\mathrm{d}\alpha) &\equiv 0 && (15) \end{aligned}$$

for 0-form f and p-form α and β.

So what's the goal of defining all such complicated stuff? Ok, all the suffering ends once we see Stokes' theorem. For a p-dimensional manifold V with its boundary ∂V, given a $p-1$-form on V, we have:

$$\int_V \mathrm{d}\omega = \int_{\partial V}\omega \tag{16}$$

This includes the famous Green formula ($p = 2$) and Stokes formula ($p = 3$) as we have learned in physics. In Green case, we have

$$\omega = a(x, y)\mathrm{d}x + b(x, y)\mathrm{d}y \tag{17}$$

so we have

$$\begin{aligned} \mathrm{d}\omega &= \mathrm{d}a \wedge \mathrm{d}x + \mathrm{d}b \wedge \mathrm{d}y & (18) \\ &= (a_x \mathrm{d}x + a_y \mathrm{d}y) \wedge \mathrm{d}x + (b_x \mathrm{d}x + b_y \mathrm{d}y) \wedge \mathrm{d}y & (19) \\ &= (b_x - a_y)\mathrm{d}x \wedge \mathrm{d}y & (20) \end{aligned}$$

Note $\mathrm{d}x^i \wedge \mathrm{d}x^i = 0$. For 2D region V and its positively oriented boundary ∂V, we have

$$\begin{aligned} \int_V \mathrm{d}\omega &= \iint_V \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) \mathrm{d}x\mathrm{d}y & (21) \\ \int_{\partial V} \omega &= \oint_{\partial V} a(x, y)\mathrm{d}x + b(x, y)\mathrm{d}y & (22) \end{aligned}$$

According to Stokes' theorem, we thus have

$$\iint_V \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \oint_{\partial V} a(x, y)\mathrm{d}x + b(x, y)\mathrm{d}y \tag{23}$$

which is Green formula! Note I haven't define what does it means by intergrating a p-form, but if you are interested you can check online; if you are tired, there is no bother to introduce it here.

Hopefully you enjoy the beauty of mathematics.

3 Connection, Parallel Transportation and Curvature

3.1 Connection

From the abstract definition of manifold M, there is no way to compare two vectors in the different tangent space corresponding to different point on M, or *build connection* between them.

In the situation that we have an ambient space and the inner product $< \cdot, \cdot >$ is the same as in ambient space, that's easy and natural. Suppose we have $\mathbf{v} \in T_p M$ and $\mathbf{w} \in T_{p'} M$. We want to *find their difference*. We do the following:

- Parallelly move $\mathbf{w}$ to position p. Here "parallel" is in the sense of ambient space.
- Orthogonally project $\mathbf{w}$ on $T_p M$ to get $\mathbf{w}' \in T_p M$. Here "orthogonal" is in the sense of ambient space.
- $\mathbf{w}'$ is now in the same space as $\mathbf{v}$ and can be compared with $\mathbf{v}$.
- $\mathbf{v} - \mathbf{w}'$ is the result.

In other words, The difference of $\mathbf{v}$ and $\mathbf{w}$ is not $\mathbf{v} - \mathbf{w}$ (which is not in the tangent space of p), but its projection on the tangent space (See Fig. 2). To do that, we need *parallel transport* $\mathbf{w} \in T_{p'} M$ to $\mathbf{w}' \in T_p M$. Although they are not parallel in the ambient space, they are parallel in the sense of manifold M.

3.1.1 An example in $\mathbb{R}^3$

For example, consider a 2D surface $\mathbf{x} = \mathbf{x}(u^1, u^2)$ embedded in $\mathbb{R}^3$ with

$$\mathbf{x}_\alpha \equiv \frac{\partial \mathbf{x}}{\partial u^\alpha} \quad \alpha = 1, 2 \tag{24}$$

as the bases of tangent spaces in each position $p = (u^1, u^2)$. Define

$$\mathbf{N} = \frac{\mathbf{x}_1 \times \mathbf{x}_2}{\|\mathbf{x}_1 \times \mathbf{x}_2\|} \tag{25}$$

as the normal vector in each position. Eqn. 25 is also called *Gauss Normal map*.

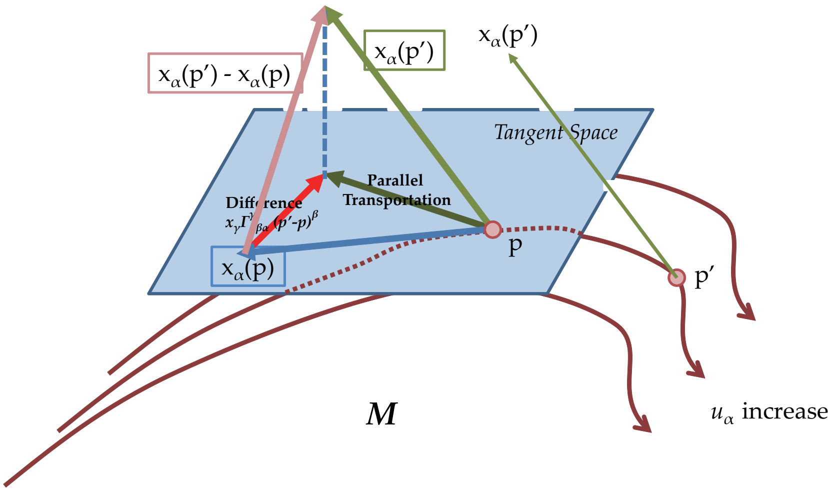

Figure 2: The illustration of connection in the ambient space.

Suppose we want to compare the vector $\mathbf{x}_\alpha(p)$ and $\mathbf{x}_\alpha(p')$ where $p' = p + \epsilon u^\beta$ is an infinitesimal change from p at coordinate u^β. Thus we have

$$\mathbf{x}_\alpha(p') - \mathbf{x}_\alpha(p) = \epsilon \mathbf{x}_{\alpha\beta} + O(\epsilon^2) \tag{26}$$

where $\mathbf{x}_{\alpha\beta} = \partial \mathbf{x}/\partial u^\alpha \partial u^\beta$. Its projection onto the tangent space can be represented by a linear combination of bases $\mathbf{x}_\gamma$ (at point p), thus we have

$$\epsilon \mathbf{x}_{\alpha\beta} = \epsilon \mathbf{x}_\gamma \Gamma^\gamma_{\beta\alpha} + < \epsilon \mathbf{x}_{\alpha\beta}, \mathbf{N} > \mathbf{N} \tag{27}$$

The component $\epsilon \mathbf{x}_\gamma \Gamma^\gamma_{\beta\alpha}$ is the desired difference in tangent space.

In general, given two close positions p and p' with $\delta p = p' - p$, the parallel transportation $\mathbf{x}_\alpha(p' \to p)$ of $\mathbf{x}_\alpha(p')$ to p is given by

$$\mathbf{x}_\alpha(p' \to p) = \mathbf{x}_\alpha + \mathbf{x}_\gamma \Gamma^\gamma_{\beta\alpha} \delta p^\beta |_p \tag{28}$$

$\Gamma^\gamma_{\beta\alpha}$ is the *Christoffel symbols*. We can compute it if we know the metric tensor $g_{\alpha\beta} \equiv < \mathbf{x}_\alpha, \mathbf{x}_\beta >$:

$$\Gamma^\tau_{\mu\beta} = \frac{1}{2} g^{\alpha\tau} \left(\frac{\partial g_{\alpha\mu}}{\partial u^\beta} + \frac{\partial g_{\beta\alpha}}{\partial u^\mu} - \frac{\partial g_{\mu\beta}}{\partial u^\alpha} \right) \tag{29}$$

where $[g^{\alpha\tau}] = [g_{\alpha\tau}]^{-1}$ (Matrix inverse).

For any vector field $\mathbf{v} = v^\gamma \mathbf{x}_\gamma$ defined on manifold M, the component difference between $\mathbf{v}(p)$ and $\mathbf{v}(p')$ is

$$\mathbf{v}(p') - \mathbf{v}(p) = v^\gamma_{,\beta} \mathbf{x}_\gamma \delta p^\beta + v^\alpha \mathbf{x}_{\alpha\beta} \delta p^\beta |_p + O(\|\delta p\|^2) \tag{30}$$

where $v^\gamma_{,\beta} \equiv \partial v^\gamma / \partial u^\beta$. Similar to Eqn. 28, Its projection on T_pM becomes $(v^\gamma_{,\beta} + \Gamma^\gamma_{\beta\alpha} v^\alpha) \mathbf{x}_\gamma \delta p^\beta$. So the parallel transportation $\mathbf{v}(p' \to p)$ from p' to p is

$$\mathbf{v}(p' \to p) = \mathbf{v} + (v^\gamma_{,\beta} + \Gamma^\gamma_{\beta\alpha} v^\alpha) \mathbf{x}_\gamma \delta p^\beta |_p \tag{31}$$

3.1.2 Without ambient space

In the case of abstract defined manifold M, one cannot build the connection this way. Instead, one can start directly with the Christoffel symbols Γ^k_{ij} (Eqn. 31). In the Riemann case, one can start with the metric tensor g_{ij} and use Eqn. 29 to build Γ^k_{ij}, which is called *Levi-Civita connection*. This connection is the unique connection that perserves inner product:

$$< \mathbf{v}(p' \to p), \mathbf{w}(p' \to p) >_p = < \mathbf{v}(p'), \mathbf{w}(p') >_{p'} \tag{32}$$

with proof in Theorem 9.18, [1].

3.2 The Covariant Differentiation

The connection enables us to define a special kind of differentiation over the vector field $\mathbf{v}$ on M, the covariant differentiation ∇.

Formally, given a vector $\mathbf{X}$ at p and a *vector field* $\mathbf{v}$ defined near p, $\nabla_{\mathbf{X}}\mathbf{v}$ is a vector at p that is the directional derivative of $\mathbf{v}$ along the direction of $\mathbf{X}$, by parallel transporting $\mathbf{v}(p')$ back to p:

$$\nabla_{\mathbf{X}}\mathbf{v}|_p = \lim_{\epsilon\to 0} \frac{\mathbf{v}(p'(\epsilon)\to p) - \mathbf{v}(p)}{\epsilon} \tag{33}$$

where $p'(\epsilon) = p + \epsilon\mathbf{X}$. Here the addition is done element-wise. Using Eqn. 31 we get

$$\nabla_{\mathbf{X}}\mathbf{v}|_p = (v^\gamma_{,\beta} + \Gamma^\gamma_{\beta\alpha}v^\alpha)\mathbf{x}_\gamma X^\beta|_p \tag{34}$$

The reason why we call it "covariant" is because $\nabla_{\mathbf{X}}\mathbf{v}$ is truly a vector, or a first-order covariant tensor.

With this concept, we can parallel transport a vector $\mathbf{v}(u(0))$ along a curve $u = u(t),\ t \in [0, t_0]$ by solving the following ODE:

$$\nabla_{\dot{u}(t)}\mathbf{v}(u(t)) = \left(v^\gamma_{,\beta}(u(t)) + \Gamma^\gamma_{\beta\alpha}(u(t))v^\alpha(u(t))\right)\mathbf{x}_\gamma(u(t))\dot{u}^\beta(t) = 0 \tag{35}$$

with a known initial condition $\mathbf{v}(u(0))$ and unknown function $\mathbf{v}(t) = \mathbf{v}(u(t))$. Note Eqn. 35 can be written component-wise and $\dot{v}^\gamma(t) \equiv v^\gamma_{,\beta}(u(t))\dot{u}^\beta(t)$ by chain rule, Eqn. 35 is actually

$$\dot{v}^\gamma + \Gamma^\gamma_{\beta\alpha}\dot{u}^\beta v^\alpha = 0 \tag{36}$$

evaulated at t (or $u(t)$).

Intuitively, Eqn. 35 means: along the direction $\dot{u}(t)$ the curve goes, $\mathbf{v}(u(t))$ should not change in the sense of parallel transportation. Note Eqn. 35 is a linear ODE w.r.t $v^\alpha(u(t))$ (Note $\dot{u}^\beta(t)$ and $\Gamma^\gamma_{\beta\alpha}(u(t))$ are known.)

As a linear ODE, there is always a solution for $\mathbf{v}(u(t_0))$, which is the transported vector from $\mathbf{v}(u(0))$. Furthermore, given $u = u(t)$, there exists a $\mathbf{v}(u(0))$-independent linear operator A so that $A^\beta_\alpha v^\alpha(u(0)) = v^\beta(u(t_0))$.

So the whole idea is, to transport a vector to a different location, first **(1)** define how to measure the difference between two vectors in two different locations, and then **(2)** fix one vector and solve an equation w.r.t another vector with the constraint that their difference defined in (1) should be zero.

4 Geodesics

As we all know, the shortest distance between two points on the Earth is not a line but the Great Circle. In general, the shortest curve on a manifold M that connects two given points is the *geodesic*.

Formally, the geodesic is the curve $u = u(t)\ t \in [0, t_0]$ that minimizes the following functional $J(u)$:

$$J(u) = \int \mathrm{d}s = \int_0^{t_0} \sqrt{g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta}\mathrm{d}t \tag{37}$$

where $g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta$ is the squared length of tangent vector $\dot{u}^\alpha(t)\mathbf{x}_\alpha$. Using Euler-Lagrange equation (in calculus of variation, very messy), we get the geodesic equation

$$\ddot{u}^\gamma + \Gamma^\gamma_{\beta\alpha}\dot{u}^\beta\dot{u}^\alpha = 0 \tag{38}$$

Geodesics can be obtained also by covariate derivative ∇:

$$\nabla_{\dot{u}(t)}\dot{u}(t) = 0 \tag{39}$$

given the initial condition $u(0)$ and $\dot{u}(0)$. Using Eqn. 36 by replacing $v = \dot{u}$ in Eqn. 39, we obtain the same equation as in Eqn. 38. Eqn. 39 means *what the curve should be so that it parallelly transports its own tangent vector along the curve.*

Note Eqn. 38 is no longer linear ODE. So given two distint points, generally there is no guarentee there would be a geodesic linking them. However, by Picard's theorem in ODE, Eqn. 38 can at least give a local geodesic by specifying a starting point $u(0)$ and a starting direction $\dot{u}(0)$ in its tangent space T_pM.

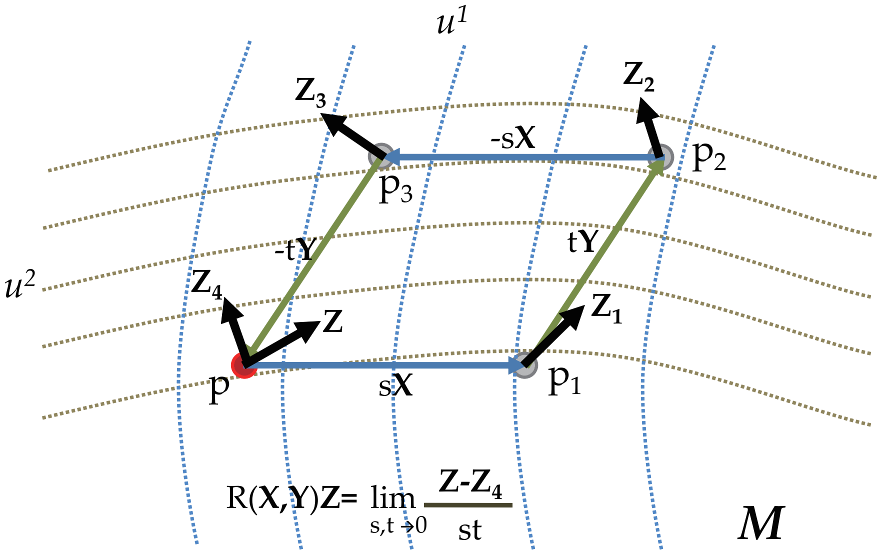

Figure 3: Riemann Curvature.

5 Curvature

5.1 General Definition

In the Euclidean space, parallelly transportign a vector from one point p to another point p' via two different paths yields the same vector at p'.

However, it is generally not the case for manifold. In particular, the parallel tranportation of a vector $\mathbf{v}$ at p on a closed curve C of M will give another tangent vector $\mathbf{v}'$ which is not the same as $\mathbf{v}$. The difference between them is the measurement of the curvature of the region $R(C)$ inside C. If $R(C)$ becomes smaller and smaller, then the measurement is more concentrated around p. In the limit of $R(C) \to 0$, we thus get a precise measurement of the curvature at the point $p \in M$.

Formally, given two vector $\mathbf{X}$ and $\mathbf{Y}$ defined at p, the *Riemann curvature tensor* $R_p(\mathbf{X}, \mathbf{Y})$ is a linear transform on the tangent space T_pM that maps any vector $\mathbf{Z} \in T_pM$ to $R_p(\mathbf{X}, \mathbf{Y})\mathbf{Z} \in T_pM$. This mapping is done by parallelly transporting $\mathbf{Z}$ first in the $\mathbf{X}$ direction, then in the $\mathbf{Y}$ direction, then in the $-\mathbf{X}$ direction, then in the $-\mathbf{Y}$ direction, and finally taking the vector difference between $\mathbf{Z}$ and transported vector $\mathbf{Z}'$.

Fig. 3 shows how this step is done. Using Eqn. 36, we have

$$Z_1^\gamma = Z^\gamma - s\Gamma^\gamma_{\beta\alpha}(p) X^\beta Z^\alpha \tag{40}$$

for small displacement $s\mathbf{X}$ from p to $p_1 = p + s\mathbf{X}$. Similarly, we have

$$\begin{aligned} Z_2^\gamma &= Z_1^\gamma - t\Gamma^\gamma_{\beta\alpha}(p_1) Y^\beta Z_1^\alpha &(41)\\ &= Z_1^\gamma - t\Gamma^\gamma_{\beta\alpha}(p + sX) Y^\beta (Z^\alpha - s\Gamma^\alpha_{\lambda\rho}(p) X^\lambda Z^\rho) &(42)\end{aligned}$$

for small displacement $s\mathbf{Y}$ from p_1 to $p_2 = p + s\mathbf{X} + t\mathbf{Y}$. The terms containing st in Eqn. 42 is:

$$-st\left(\partial_\lambda \Gamma^\gamma_{\beta\alpha} - \Gamma^\gamma_{\beta\epsilon}\Gamma^\epsilon_{\lambda\alpha}\right) X^\lambda Z^\alpha Y^\beta \tag{43}$$

Note all the zero and first-order terms will be cancelled out when the circle is finished. $R_p(\mathbf{X}, \mathbf{Y})\mathbf{Z}$ is thus defined as

$$R_p(\mathbf{X}, \mathbf{Y})\mathbf{Z} = \lim_{s\to 0, t\to 0} \frac{\mathbf{Z} - \mathbf{Z}_4}{st} = \left(\partial_\lambda \Gamma^\gamma_{\beta\alpha} - \partial_\beta \Gamma^\gamma_{\lambda\alpha} + \Gamma^\gamma_{\lambda\epsilon}\Gamma^\epsilon_{\beta\alpha} - \Gamma^\gamma_{\beta\epsilon}\Gamma^\epsilon_{\lambda\alpha}\right) X^\lambda Z^\alpha Y^\beta \tag{44}$$

Note half of the terms are from Eqn. 43, one can similarly compute the other half (which essentially permutes $\mathbf{X}$ with $\mathbf{Y}$). Eqn. 44 holds for any vector $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$, so we essentially get a fourth-order (mixed) tensor, which is the *Riemann curvature tensor*:

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \tag{45}$$

Note the dummy indices have changed to concide with wiki's. R can also be represented using covariant differentiation ∇:

$$R(\mathbf{X},\mathbf{Y})\mathbf{Z} = \left([\nabla_{\mathbf{X}},\nabla_{\mathbf{Y}}] - \nabla_{[\mathbf{X},\mathbf{Y}]}\right)\mathbf{Z} = \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{Z} - \nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z} \tag{46}$$

where $[\mathbf{X},\mathbf{Y}]$ is the Lie bracket for two vector fields $\mathbf{X}$ and $\mathbf{Y}$. Note that in our previous definition of R, vector $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are only defined at p, *not around* p, while Eqn. 46 requires $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are locally vector fields. However, Eqn. 46 yields identical answers for arbitrary smooth vector fields $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ given $\mathbf{X}(p)$, $\mathbf{Y}(p)$ and $\mathbf{Z}(p)$ fixed.

5.2 Surface in $\mathbb{R}^3$ and Gauss-Bonnet Theorem

For surface in $\mathbb{R}^3$, we can have a simpler definition of curvature.

5.2.1 The differential of the Gauss map

Consider a 2D surface $\mathbf{x} = \mathbf{x}(u^1, u^2)$ embedded in $\mathbb{R}^3$. Remind the Gauss map defined in Eqn. 25 which maps $p \in M$ to a normal vector $\mathbf{N}_p$.

Suppose we what to know how the normal vector changes when p changes. We begin with *the differential of the Gauss map* $\mathrm{d}\mathbf{N}_p$:

$$\mathrm{d}\mathbf{N}_p = \partial_\alpha \mathbf{N} \mathrm{d}u^\alpha = \frac{\partial \mathbf{N}}{\partial u^\alpha}\mathrm{d}u^\alpha \tag{47}$$

This is called *vector-value 1-form*, a linear mapping that takes $\mathbf{v} \in T_pM$ and outputs another vector in T_pM. This mapping is actually a matrix, or a (1,1) mixed tensor. Here

$$\mathrm{d}\mathbf{N}(\mathbf{v}) = \partial_\alpha \mathbf{N} \mathrm{d}u^\alpha(\mathbf{v}) = v^\alpha \partial_\alpha \mathbf{N} \tag{48}$$

i.e., $\mathrm{d}\mathbf{N}(\mathbf{v})$ is the change of normal vector along the direction of $\mathbf{v}$. Note $\mathrm{d}\mathbf{N}(\mathbf{v}) \in T_pM$ since $\mathbf{N}$ is perpendicular to the surface and $\|\mathbf{N}\| = 1$ always.

5.2.2 The Second Fundamental Form

We know the first fundamental form $I_p(\mathbf{v},\mathbf{w})$ is determined by the metric tensor g:

$$I_p(\mathbf{v},\mathbf{w}) \equiv <\mathbf{v},\mathbf{w}>_p = g_{\alpha\beta}v^\alpha w^\beta \tag{49}$$

which yields the length of a tangent vector.

The second fundamental form $II_p(\mathbf{v})$ is defined as follows:

$$II_p(\mathbf{v}) \equiv - <\mathrm{d}\mathbf{N}_p(\mathbf{v}),\mathbf{v}>_p = -g_{\beta\gamma}v^\alpha(\partial_\alpha \mathbf{N})^\beta v^\gamma \tag{50}$$

Geometrically, it is the inner-product between the change of position and the change of normal vector.

There are two special directions $\mathbf{v}$ that $\mathrm{d}\mathbf{N}_p(\mathbf{v})$ is colinear to $\mathbf{v}$, or $\mathrm{d}\mathbf{N}_p(\mathbf{v}) = k\mathbf{v}$. These directions (the eigenvectors of $\mathrm{d}\mathbf{N}_p$) are called *principle directions*, and the two ks (k_1 and k_2, the eigenvalues of $\mathrm{d}\mathbf{N}_p$) are called *principle curvatures*.

The *Gaussian curvature* $K = k_1k_2$ is thus the determinant of $\mathrm{d}\mathbf{N}_p$ (as a matrix) and the *mean curvature* $H = \frac{1}{2}(k_1 + k_2)$ is half of its trace. Both are independent of the choice of local coordinates.

5.2.3 Some interesting examples

Here is some interesting examples how the local curvature K affects the local measurement of the geometry.

In Euclidean geometry, we all know in the plane

- The summation of interior angles of a triangle is π.
- The ratio between a circle's perimeter and its radius is 2π.
- The ratio between a circle's area and its radius's square is π.

None of them are true for regions with non-vanishing K.

5.2.4 Gauss-Bonnet Theorem

Very important theorem. However too tired to type it. Please google it.

5.2.5 Advanced topic: Cartan's structural equations

Let's go back to general abstract manifold M. We can define $\mathbf{e}_i$ and σ^i similar to $\boldsymbol{\partial}_i$ and $\mathrm{d}x^i$ as the local bases of the vectors and covectors (or 1-form). Note $\mathbf{e}_i$ needs not be *coordinate-aligned*, which means the direction it points need not to be the increasing direction of a coordinate x^i, as in the case of $\boldsymbol{\partial}_i$. We call $\mathbf{e}_i$ the *coordinate frames*.

We can define similar concept of parallel transportation and covariant differentiation using $\mathbf{e}_i$ and σ^i with the aid of ω_{ij}^k that is similar to Γ_{ij}^k. For example, we now define the covariant differentiation $\nabla_{\mathbf{X}}\mathbf{v}$ for vector $\mathbf{X}$ and vector field $\mathbf{v}$ as follows:

$$\nabla_{\mathbf{X}}\mathbf{v} = \mathbf{e}_i[\mathbf{e}_k(v^i) + \omega_{kj}^i v^j]X^k \tag{51}$$

where $\mathbf{e}_k(v^i)$ is the directional derivative of function v^i in $\mathbf{e}_k$ direction. Based on this, we can define $\nabla\mathbf{v}$:

$$\nabla\mathbf{v} \equiv \mathbf{e}_i[\mathbf{e}_k(v^i) + \omega_{kj}^i v^j]\sigma^k \tag{52}$$

where σ^k extracts the k-th component under the coordinate $\mathbf{e}_k$. Particularly we have for $\mathbf{v} = \mathbf{e}_j$:

$$\nabla\mathbf{e}_j = \mathbf{e}_i\omega_{kj}^i\sigma^k \tag{53}$$

If we introduce the following (fancy) notation:

$$\mathbf{e} \equiv [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n] \tag{54}$$

$$\nabla\mathbf{e} \equiv [\nabla\mathbf{e}_1, \nabla\mathbf{e}_2, \ldots, \nabla\mathbf{e}_n] \tag{55}$$

$$\omega \equiv [\omega^i{}_j], \text{ where } \omega^i{}_j \equiv \omega_{kj}^i\sigma^k \tag{56}$$

where $\mathbf{e}$ and $\nabla\mathbf{e}$ are row vectors *whose elements are vectors*, and ω is a matrix whose elements are 1-forms, then we can write Eqn. 53 as the following compact form:

$$\nabla\mathbf{e} = \mathbf{e}\omega \tag{57}$$

where $\mathbf{e}\omega$ means multiplying a row vector with a matrix to get another row vector (symbolically). For a general vector $\mathbf{v}$, similarly we have (Note by chain rule $\mathbf{e}_k(v^i)\sigma^k = \mathrm{d}v^i$):

$$\nabla\mathbf{v} = \mathbf{e}(\mathrm{d}v + \omega v) \tag{58}$$

where $v = [v^1, v^2, \ldots, v^n]$ is a column vector whose elements are functions and $\mathrm{d}v = [\mathrm{d}v^1, \mathrm{d}v^2, \ldots, \mathrm{d}v^n]^T$ is a column vector whose elements are 1-forms.

Basically, Eqn. 57 show how the spatial variation $\nabla\mathbf{e}$ of bases can be linearly represented by the bases $\mathbf{e}$ itself. Such equations (Eqn. 27 is another example) are usually called **structure equations**. I think *this is the key idea in Differential Geometry and Riemann Geometry.*

We have similar results for the dual bases σ^i (requires a more complicated proof in Section 9.3c [1]):

$$\mathrm{d}\sigma^i = -\omega^i{}_k \wedge \sigma^k \tag{59}$$

if $\omega_{kj}^i = \omega_{jk}^i$ (Connection is symmetric). Compactly, we can write

$$\mathrm{d}\sigma = -\omega \wedge \sigma \tag{60}$$

If we take the covariant differential again from Eqn. 57, we get

$$\nabla\nabla\mathbf{e} = \nabla(\mathbf{e}\omega) = (\nabla\mathbf{e})\omega + \mathbf{e}\mathrm{d}\omega = \mathbf{e}\theta \tag{61}$$

where the curvature 2-form θ is defined as

$$\theta \equiv \mathrm{d}\omega + \omega \wedge \omega \tag{62}$$

or in full

$$\theta^i{}_j \equiv \mathrm{d}\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j \tag{63}$$

One can compute the coefficients of ${\theta^i}_j$ (which is a 2-form) and get

$$ {\theta^i}_j = \frac{1}{2} R^i_{jrs} \sigma^r \wedge \sigma^s \tag{64} $$

If $\mathbf{e} = \boldsymbol{\partial}$, then R in Eqn. 64 is the Riemann curvature tensor defined in Eqn. 45.

Sounds cool:-).

6 About "Lie Stuff"

6.1 Lie Derivatives

6.1.1 Vector Flow

Given a vector field X on M, one can establish a first-order dynamic system as follows:

$$ \dot{x}^i = X^i(x^1, x^2, \ldots, x^n) \tag{65} $$

that is, find the stream line out of a velocity field X. Given the initial condition $x(0)$, the solution defines a *flow* $\phi_t : M \mapsto M$ generated from X that maps from $x(0)$ to $x(t)$. Furthermore, ϕ_t has the following property:

$$ \phi_0 = \mathrm{Id} \tag{66} $$
$$ \phi_t \circ \phi_s(\cdot) = \phi_t(\phi_s(\cdot)) = \phi_{t+s}(\cdot) \tag{67} $$
$$ \phi_t^{-1} = \phi_{-t} \tag{68} $$

This means ϕ_t forms a *one-parameter group* acting on manifold M.

Not only $p \in M$, vector $\mathbf{v} \in T_pM$ can also be transformed by ϕ_t to $\mathbf{v}' \in T_{\phi_t(p)}M$. This transformation is defined by ϕ_{*t}:

$$ (\phi_{*t}\mathbf{v})^j \equiv \lim_{\epsilon \to 0} \frac{\phi_t^j(p + \epsilon \mathbf{v}) - \phi_t^j(p)}{\epsilon} = \partial_i \phi_t^j|_p v^i \tag{69} $$

which is linear to $\mathbf{v}$. ($\phi_t^j(p)$ is the j-th component in local coordinates).

6.1.2 Lie Derivative $L_X Y$

Given the vector field X, starting from p we define a linear transform ϕ_{*t} from the tangent space T_pM to another tangent space $T_{\phi_t(p)}M$ along the streamline of the flow. Given another vector field Y, one thus can measure *the degree of matching* between the change of Y along the streamline and change created by the transform ϕ_{*t} on a fixed vector $Y(p)$. This is the intuition behind Lie derivatives $L_X Y$.

Formally, Lie derivative $L_X Y$ is defined as follows:

$$ L_X Y|_p = \lim_{t \to 0} \frac{Y(\phi_t(p)) - \phi_{*t} Y(p)}{t} \tag{70} $$

which is another vector field. One can prove that $L_X Y = [X, Y]$ where $[X, Y]$ is the *Lie bracket*:

$$ [X, Y]f|_p = X(Y(f)) - Y(X(f))|_p \tag{71} $$

for any function f on M. Here $X(Y(f))$ means first taking directional derivative of f in Y direction, then treating this derivative as another function and taking its directional derivative again in X direction. Symbolically, we have

$$ X(Y(f)) = X^i \partial_i (Y^j \partial_j f) = X^i \frac{\partial}{\partial x^i} \left(Y^j \frac{\partial f}{\partial x^j} \right) \tag{72} $$

That is the reason why we write the bases of the tangent space as $\boldsymbol{\partial}_i$.

6.2 Lie Group and Lie Algebra

6.2.1 Lie Group

From the group point of view, Lie group is so called "continuous group" contrasted with discrete groups. From the manifold point of view, Lie Group is a special kind of manifold that has the group structure. A typical example is $S^1 = \{e^{i\theta} | \theta \in [0, 2\pi)\} \subset \mathcal{C}$. First S^1 is a 1D manifold (a close curve), second S^1 has a group structure $e^{is}e^{it} = e^{i(s+t)}$ which has the same structure ("isomorphism") as the 2D rotation ($SO(2)$).

Other Lie group includes $SO(3)$ (3D rotation), $SU(2)$ (2 by 2 unitary matrices) and many others. Many Lie groups have matrix representations, i.e, each its element can be represented by a matrix and hence we can replace group multiplication with matrix multiplication. The branch of "Group representation" aims to finding a (matrix) representation for a given group.

The combination of group and manifold makes a beautiful and simple theory.

6.2.2 "One is all, and all is one"

In the following, we denote the Lie group as G and T_eG as its tangent space around e.

This subtitle came from the animation "Fullmental Alchemist", which I think is a very good summary for Lie group. Due to its structure, the property of Lie group is completely determined by the small neighborhood of e, the identity. In other words, every neighborhood of $g \in G$ is locally similar to the neighborhood of $e \in G$ under the following mapping L_g and its associated differential L_{g*}:

$$L_g : \quad h \in G \mapsto gh \in G \tag{73}$$

$$L_{g*} : \quad T_hG \mapsto T_{gh}G \tag{74}$$

In the situation that both the group elements and the tangent vectors are represented by matricies, $L_{g*}\mathbf{X}_R = gX_R$.

Thus, given a single vector $\mathbf{X}_R \in T_eG$, one can create the *entire vector field* $\mathbf{X}$ by applying L_{g*} on $\mathbf{X}_R$ for every $g \in G$. This is called (left) invariant vector field.

Conversely, one can verify whether a vector field $\mathbf{Y}$ is invariant or not by checking whether or not $L_{g*}\mathbf{Y}(h) = \mathbf{Y}(gh)$.

6.2.3 The Exponential Map

Based on this vector field $\mathbf{X}$, one can build the *flow* $\phi_t^{\mathbf{X}}(e)$ starting from e using the same technique as in Section. 6.1.1:

$$\dot{g}(t) = \mathbf{X}(g(t)) \tag{75}$$

with initial condition $g(0) = e$. Here $\mathbf{X}(g(t))$ is the vector field evaluated at $g(t)$, which is completely determined by $g(t)$ and the tangent vector $\mathbf{X}_R \in T_eM$. In matrix representation $\mathbf{X}(g(t)) = g(t)\mathbf{X}_R$ and we have a matrix ODE:

$$\dot{g}(t) = g(t)\mathbf{X}_R \tag{76}$$

treating $g(t)$ and $\mathbf{X}_R$ as matrices. The solution to Eqn. 76 is the famous(?) *exponential map*

$$\exp : T_eG \mapsto G \tag{77}$$

that maps $\mathbf{X}_R \in T_eG$ to $\exp(t\mathbf{X}_R) = \sum_{k=0}^{\infty} \frac{t^k X_R^k}{k!} \in G$. I guess that is the reason why we have to learn it in Linear Algebra.

Similarly, for a general flow $\phi_t^{\mathbf{X}}(g)$ starting from $g_0 \in G$, we have $g(t) = g_0 \exp(t\mathbf{X}_R)$.

6.2.4 Lie Algebra

From the idea of "one is all and all is one", we see the structure of (left) invariant vector field on G is completely determined by the structure of the tangent space T_eG. If we linearly combine two vectors $\mathbf{X}_R, \mathbf{Y}_R \in T_eG$, then we get a linear combination of two associated invariant vector fields, which is also invariant.

That seems trivial. Are there any other structures? Yes. Given two invariant vector fields $\mathbf{X}$ and $\mathbf{Y}$, their *Lie bracket* $[\mathbf{X}, \mathbf{Y}]$ gives another vector field that is invariant.

So we have $\mathbf{X}, \mathbf{Y} \mapsto [\mathbf{X}, \mathbf{Y}]$, how about X_R, Y_R maps to? Naturally we can just take the vector field $[\mathbf{X}, \mathbf{Y}]$ evaluated at e as their image: $[\mathbf{X}_R, \mathbf{Y}_R] \equiv [\mathbf{X}, \mathbf{Y}]_e$. In other words, we define the following operation on the tangent space T_eG:

$$[\cdot, \cdot] : T_eG \times T_eG \mapsto T_eG \tag{78}$$

Equipped with this operation, we thus call the tangent space T_eG *Lie Algebra*, and give it a new (cool) symbol $\mathfrak{g}$. Note the original Lie bracket involves taking spatial derivatives of vector fields, but Eqn. 78 does not. So we essentially *encode the analytic structure into the algebraic structure*, which is the nice thing in Lie Group.

In matrix representation, $[\mathbf{X}_R, \mathbf{Y}_R] = \mathbf{X}_R\mathbf{Y}_R - \mathbf{Y}_R\mathbf{X}_R$.

A famous example of this bracket $[\cdot, \cdot]$ is the cross product in tangent space $\mathbb{R}^3$ with the underlying Lie Group $SO(3)$ (3D rotation).

References

[1] Theodore Frankel, *The Geometry of Physics, An Introduction.*