General Relativity Tutorial Notes
Differential Geometry of Manifolds

Jeffrey C. Y. Teo

Nov. 2005

This tutorial notes is a supplement to the course PHYS3033 General Relativity on a more detail and mathematically flavored introduction on differential geometry of real smooth n-dimensional manifold. The first section is a review on the basics of a smooth manifold, including tangent and cotangent spaces in the form of differential operators and differential forms, tensor bundles, tensor fields and their local index representation as well as the rule of transformation. It is followed by the second section, which is the main focus of this notes, discussing the differential aspect of geometry in the language of a tensor operator known as the connection or equivalently, the Christoffel symbols. Topics such as covariant differentiation, horizontal vector fields, parallelism, and geodesic are described. The concept of torsion and curvature in both operator and tensorial form are defined, some of their algebraic and differential properties are shown, and the Ricci tensor and scalar curvature are derived from the curvature, which is essential in the study of relativity, since they help defining the Einstein tensor in the gravitational field equation. Finally, the third section of Riemannian geometry is a study of length and angle measurments on (the tangent space of) a manifold in the language of a (pseudo-) metric. Its implication, namely the unique Riemannian or Levi-Civita connection, on the geometry of the manifold is mentioned. The section will end by a simple illustration of the metric and the geometry of the sphere $\mathbb{S}^2$ in $\mathbb{R}^3$. [1, 2, 3] are standard reference texts for a more complete and general understanding on differential geometry.

1 Differentiable Manifolds

The objects studied in differentiable geometry are smooth manifolds, for example a sphere, a torus (do-nut), a saddle or a hyperboloid. Before we could investigate the geometric characteristics of these objects, we would need to develop a rigorous language describing a manifold. Like most objects in mathematics, manifolds are sets. For instance, the 2-dimensional unit sphere embedded in our 3-dimensional space is the set $\mathbb{S}^2 = \left\{(x, y, z) : x^2 + y^2 + z^2 = 1\right\}$. In order to discuss the shape of a manifold, it is assumed to be a topological space. For example, the open sets of the sphere $\mathbb{S}^2$ are in the form $\Omega = \Omega_0 \cap \mathbb{S}^2$, where Ω_0

is an open set in $\mathbb{R}^3$. For a quick review on the basic of topology, the readers are referred to the standard undergraduate textbook [4]. In general relativity, we rarely have to worry about the topology of space-time since the emphasis is on local properties (like curvature) and a local domain is usually topological trivial (except around a singularity like a blackhole).

Obviously, a manifold should be able to be characterized by its dimension. For example, a sphere, a saddle and a torus are of dimension two since they are surfaces. Any curves are 1-dimensional manifolds. The space-time manifold is of dimension four. Dimension is a concept originated from linear algebra, where the objects of consideration are vector spaces such as the linear spaces $\mathbb{R}^n$. In geometry however, manifolds are curved space. Binary operations like vector addition or scalar multiplication are not built-in properties of a manifold, and hence it lacks the concept of a basis. To generalized the notion of dimensionality, we first have to visualize a manifold locally as an open domain inside a linear space of the same dimension. Take the sphere again as an example. We may project the upper sphere $\mathbb{S}^2_+ = \left\{(x, y, z) \in \mathbb{S}^2 : z > 0\right\}$ onto the flat xy-plane by the bijective bicontinuous function (homeomorphism) $\mathbb{S}^2_+ \to D^2$ mapping $(x, y, z) \mapsto (x, y)$, where $D^2 = \left\{(x, y) : x^2 + y^2 < 1\right\}$ is the unit disc. This projection is a visulization of the curved upper hemisphere as a flat disc. In other words, we need a coordinate system to describe a manifold. However, there may not exist a coordinate system that can cover the whole manifold due to some topological reasons. For instance, the sphere $\mathbb{S}^2$ could not be visualized by a single coordinate system. The largest covering area is the stereographic projection which includes the whole sphere except one point. Thus in general, we usually need more than one coordinate system to parametrize a manifold, each system is defined on a open patch on the manifold, and the collection of all coordinate systems covers the whole sphere. The mathematical definition of a manifold is defined in the following.

Definition of a topological manifold: A topological space M is a (real topological) manifold of dimension n if there is a collection of homeomorphisms $\left\{x_j : U_j \to x_j(U_j) \subseteq \mathbb{R}^n\right\}_j$ where U_j are open sets in M such that the collection $\left\{U_j\right\}_j$ covers the whole topological space.

Each coordinate system $x_j : U_j \to \mathbb{R}^n$ is called a chart, and the collection of charts $\left\{x_j : U_j \to \mathbb{R}^n\right\}_j$ is known as an atlas.

Of course, it is possible for a local domain to be parametrized by more than one chart. For example, the right hemisphere $\mathbb{S}^2_R = \left\{(x, y, z) \in \mathbb{S}^2 : y > 0\right\}$ could be parametrized by the coordinate charts $p : (x, y, z) \mapsto (x, z)$ and $s : (x, y, z) \mapsto (\theta, \phi)$, where θ, ϕ are angles in the spherical coordinate system, i.e. $\theta = \cos^{-1} z$ and $\phi = \tan^{-1} y/x$. The coordinate transformation between two charts is called a transition. In this case, transitions are the homeomorphisms $p \circ s^{-1} : (0, \pi) \times (-\pi/2, \pi/2) \to D^2$ mapping $(\theta, \phi) \mapsto (x, y) = (\sin\theta\cos\phi, \sin\theta\sin\phi)$ or its inverse $s \circ p^{-1} : (x, y) \mapsto (\theta, \phi) = (\cos^{-1}\sqrt{1 - x^2 - y^2}, \tan^{-1} y/x)$. Note that all transitions are maps between flat domains, and since charts are assumed to be homeomorphisms, all transitions must also be homeomorphisms. Differentiability of a manifold is defined by the differentiability of the transitions.

Definition of a differentiable manifold: A (real topological) manifold M is said to be differentiable or smooth if it is equipped with an atlas $\{x_j : U_j \to x_j(U_j) \subseteq \mathbb{R}^n\}_j$ such that all transistions $t_i^j = x_i \circ x_j^{-1} : x_j(U_i \cap U_j) \to x_i(U_i \cap U_j)$ are differentiable for non-empty $U_i \cap U_j$.

The corresponding atlas is called a differentiable structure of the manifold. Here like most textbook on differential geometry, differentiability and smoothness means the capability of differentiation for an indefinite number of times.

No matter in physics or mathematics, eventually we want to study functions on differentiable manifolds, such as the energy density distribution in our curved space-time, which is a real-valued function on a manifold. In most situations, we would expect the maps we encounter are smooth or differentiable (or at least mollifiable by differentiable approximations). A real-valued function $f : M \to \mathbb{R}$ on a smooth manifold M is said to be differentiable if for any chart $x_j : U_j \to \mathbb{R}^n$ of M, $f \circ x_j^{-1} : x_j(U_j) \to \mathbb{R}$ is differentiable. A function between manifolds $f : M \to N$ is said to be differentiable if for any charts $x_i : U_i \to \mathbb{R}^m$ and $y_j : V_j \to \mathbb{R}^n$ of M and N respectively, suppose $f(U_i) \cap V_j$ is non-empty, $y_j \circ f \circ x_i^{-1} : x_i(U_i \cap f^{-1}(V_j)) \to y_j(f(U_i) \cap V_j)$ is differentiable. A bijective bi-differentiable map between manifolds $f : M \to N$ is called a diffeomorphism, and in this case, the manifolds M and N are said to be diffeomorphic, denoted by $M \approx N$.

1.1 Tangent and cotangent Space

The tangent space of any point on a sphere can be easily found out by using the normal of the sphere at that point. The tangent plane would then be the plane of the same normal containing the particular point. However, this conventional concept of tangent space is not useful in general relativity or instrinsic geometry. The existence of normal vectors is a result from the fact that the sphere is embedded inside a larger space, namely the 3-dimensional Euclidean space. Einstein theory of relativity does not assume our space-time to be a subspace of a larger dimension object, and hence a normal is not inherited as in the case of a sphere in $\mathbb{R}^3$. Besides, tangent space is actually an intrinsic property.

Suppose there is a smooth curve $\alpha : (-\epsilon, \epsilon) \to \mathbb{S}^2$ parametrized by time t on the unit sphere. A tangent vector of the curve at the point $\underline{x_0} = \alpha(0)$ could be the differential $\dot{\alpha}(0) = \frac{d\alpha}{dt}|_{t=0}$. Suppose there is a coordinate chart $X : (x, y, z) \mapsto (u, v)$ around $\underline{x_0}$. Without loss of generality, assume $X(\underline{x_0}) = (0, 0)$ Then, α induces a curve $\tilde{\alpha} = X \circ \alpha : t \mapsto (u(t), v(t))$ on the flat space $\mathbb{R}^2$. By chain rule,

$$\dot{\alpha}(0) = \frac{dX^{-1} \circ \tilde{\alpha}}{dt}|_{t=0} = \frac{du}{dt}|_{t=0}\partial_u X^{-1}|_{u=v=0} + \frac{dv}{dt}|_{t=0}\partial_v X^{-1}|_{u=v=0},$$

where $\frac{du}{dt}|_{t=0}$, $\frac{dv}{dt}|_{t=0}$ are real numbers and $\partial_u X^{-1}|_{u=v=0}$, $\partial_v X^{-1}|_{u=v=0}$ are tangent vectors. Since all tangent vectors must be in the form of $A\partial_u X^{-1}|_{u=v=0} + B\partial_v X^{-1}|_{u=v=0}$, where A, B are real scalars, and the (affined) tangent space at $\underline{x_0}$ is the span of $\left\{\partial_u X^{-1}|_{u=v=0}, \partial_v X^{-1}|_{u=v=0}\right\}$. Since the tangent space is of dimension two, $\partial_u X^{-1}|_{u=v=0}, \partial_v X^{-1}|_{u=v=0}$ must form a basis of the tangent space.

To generalized a tangent space to an intrinsic concept, we could replace $\partial_u X^{-1}|_{u=v=0}, \partial_v X^{-1}|_{u=v=0}$ by the differential operators $\partial_u|_{\underline{x_0}}, \partial_v|_{\underline{x_0}}$ respectively, and the (generalized) tangent space is the span of the differential operators. (For a rigorous definition of these differential operators, please refer to the Appendix.) The linear independence of $\partial_u|_{\underline{x_0}}, \partial_v|_{\underline{x_0}}$ can be checked easily. Moreover, we do not have to worry whether this definition of tangent space depends on the choice of coordinate charts or not. For instance, if there is another chart $Y : (x, y, z) \mapsto (u', v')$ around $\underline{x_0}$, then by chain rule,

$$\frac{\partial}{\partial u'}|_{\underline{x_0}} = \frac{\partial u}{\partial u'}|_{\underline{x_0}} \frac{\partial}{\partial u}|_{\underline{x_0}} + \frac{\partial v}{\partial u'}|_{\underline{x_0}} \frac{\partial}{\partial v}|_{\underline{x_0}},$$

$$\frac{\partial}{\partial v'}|_{\underline{x_0}} = \frac{\partial u}{\partial v'}|_{\underline{x_0}} \frac{\partial}{\partial u}|_{\underline{x_0}} + \frac{\partial v}{\partial v'}|_{\underline{x_0}} \frac{\partial}{\partial v}|_{\underline{x_0}}.$$

And hence, $span\left\{\partial_u|_{\underline{x_0}}, \partial_v|_{\underline{x_0}}\right\} = span\left\{\partial_{u'}|_{\underline{x_0}}, \partial_{v'}|_{\underline{x_0}}\right\}$.

So in general for a smooth manifold, the definition of tangent space at a point is described by the following.

Definition of tangent space: For M is a smooth manifold, p is a point in M, suppose $x = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a coordinate chart on a neighbourhood of p, the tangent space at p is defined to be the vector space, $T_pM = span\left\{\frac{\partial}{\partial x^1}|_p, \ldots, \frac{\partial}{\partial x^n}|_p\right\}$.

(Again, for a rigorous definition of the differential operators, please refer to Appendix.) If $\alpha : (-\epsilon, \epsilon) \to M$ is a curve on a smooth manifold M, $p = \alpha(0)$, $x = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a coordinate chart on a neighbourhood of p, then the tangent vector of α at the point $p = \alpha(0)$ is

$$\dot{\alpha}(0) = \frac{dx^j \circ \alpha}{dt}|_{t=0} \frac{\partial}{\partial x^j}|_p,$$

where Einstein summation notation is adopted for j running from 1 to n. And this definition is independent from the choice of coordinate charts. Suppose there is another chart $y = (y^1, \ldots, y^n) : U \to \mathbb{R}^n$. By chain rule, for $i = 1, \ldots, n$,

$$\frac{\partial}{\partial y^i}|_p = \frac{\partial x^j}{\partial y^i}|_p \frac{\partial}{\partial x^j}|_p$$

(see Appendix). And hence, the tangent vector of the curve α is

$$\dot{\alpha}(0) = \frac{dx^j \circ \alpha}{dt}|_{t=0} \frac{\partial}{\partial x^j}|_p$$

$$= \frac{d(x^j \circ y^{-1} \circ y \circ \alpha)}{dt}|_{t=0} \frac{\partial}{\partial x^j}|_p = \partial_i (x^j \circ y^{-1})|_{y(p)} \frac{dy^i \circ \alpha}{dt}|_{t=0} \frac{\partial}{\partial x^j}|_p$$

$$= \frac{dy^i \circ \alpha}{dt}|_{t=0} \frac{\partial x^j}{\partial y^i}|_p \frac{\partial}{\partial x^j}|_p = \frac{dy^i \circ \alpha}{dt}|_{t=0} \frac{\partial}{\partial y^i}|_p,$$

or equivalently $\frac{dx^j\circ\alpha}{dt}|_{t=0} = \frac{dy^i\circ\alpha}{dt}|_{t=0}\frac{\partial x^j}{\partial y^i}|_p$. Similarly, we obtained the contravariant rule of transformation of a tangent vector (contravariant vector). By writing $x' = y$, a tangent vector X_p at the point p can be written as the following forms.

$$X_p = X_p^i\frac{\partial}{\partial x^i}|_p = X_p^i\frac{\partial x'^j}{\partial x^i}|_p\frac{\partial}{\partial x'^j}|_p = {X'}^j_p\frac{\partial}{\partial x'^j}|_p,$$

where ${X'}^j_p = X_p^i\frac{\partial x'^j}{\partial x^i}|_p$.

Suppose V is a real vector space with a basis $\{e_1,\dots,e_n\}$. The dual space of V, denoted by V^* is defined to be the space of linear functionals $L : V \to \mathbb{R}$. For $j = 1,\dots,n$, we could define the functional $e^j : V \to \mathbb{R}$ mapping $v \mapsto v^j$, for any vector $v = v^j e_j$ in V, where v^j are real scalars. In other words, $e^j(e_i) = \delta_i^j$. Then, every functional $L : V \to \mathbb{R}$ can be expressed as $L = L(e_j)e^j$, since for any vector v in V, $L(v) = L(v^j e_j) = L(e_j)v^j = L(e_j)e^j(v)$. Therefore, the dual space V^* is the span of $\{e^1,\dots,e^n\}$. Clearly, $e^1,\dots,e^n$ are linearly independent, and thus form a basis of V^*. This basis is called the canonical basis of the dual space with respect to the basis $\{e_1,\dots,e_n\}$ of V.

The cotangent space at a point p of a smooth manifold is defined to be the dual space of the tangent space T_pM, and is denoted by T_p^*M. Suppose $x = (x^1,\dots,x^n) : U \to \mathbb{R}^n$ is a coordinate chart on a neighbourhood of p, then $\left\{\frac{\partial}{\partial x^1}|_p,\dots,\frac{\partial}{\partial x^n}|_p\right\}$ is a basis of T_pM. The canonical basis of T_p^*M is denoted by $\left\{dx_p^1,\dots,dx_p^n\right\}$. This means $dx_p^j\left(\frac{\partial}{\partial x^i}|_p\right) = \delta_i^j$. There is one remark about dual spaces. The dual space of V^* is V itself, and hence the dual space of T_p^*M is T_pM.

Now, suppose there is another coordinate chart $x' = (x'^1,\dots,x'^n) : U \to \mathbb{R}^n$. We have another basis $\left\{\frac{\partial}{\partial x'^j}|_p\right\}_j$ of T_pM, and its corresponding canonical basis $\left\{dx'^j_p\right\}_j$ of T_p^*M. For any tangent vector X_p in T_pM, for $j = 1,\dots,n$,

$$d{x'}^j_p(X_p) = d{x'}^j_p\left({X'}^k_p\frac{\partial}{\partial x'^k}|_p\right) = {X'}^k_p\delta_k^j = {X'}^j_p = X_p^i\frac{\partial x'^j}{\partial x^i}|_p.$$

Since $X_p^i = dx_p^i(X_p)$, we have

$$d{x'}^j_p(X_p) = \frac{\partial x'^j}{\partial x^i}|_p dx_p^i(X_p),$$

in other words, we obtained the chain rule of the differential form

$$d{x'}^j_p = \frac{\partial x'^j}{\partial x^i}|_p dx_p^i.$$

If ω_p is a cotangent vector (covariant vector) in T_p^*M, it can be expressed in either coordinate systems as the following.

$$\omega_p = {\omega'}_{p_j} d{x'}^j_p = {\omega'}_{p_j}\frac{\partial x'^j}{\partial x^i}|_p dx_p^i = \omega_{p_i} dx_p^i,$$

where $\omega_{p_i} = {\omega'}_{p_j}\frac{\partial x'^j}{\partial x^i}|_p$, which is the covariant rule of transformation.

Suppose V and W are vector spaces with bases $\{e_1, \ldots, e_m\}$ and $\{f_1, \ldots, f_n\}$ respectively. Their tensor space is defined to be the real vector space $V \otimes W = span\{e_i \otimes f_j : i = 1, \ldots, m, j = 1, \ldots, n\}$. Hence, $V \otimes W$ is of dimension mn. The tensor product $\otimes : V \times W \to V \otimes W$ is defined to be the multi-linear (but not linear) function mapping $(v, w) \mapsto v \otimes w = v^i w^j e_i \otimes f_j$. By iteration, we may define the tensor space $V_1 \otimes \ldots \otimes V_k$ of k vector spaces, and the corresponding tensor product $(v_1, \ldots, v_k) \mapsto v_1 \otimes \ldots \otimes v_k$. Applying the definition of tensor product, we may derive from the tangent and cotangent space, the tensor space $\otimes_s T_p M \otimes_r T_p^* M = \underbrace{T_p M \otimes \ldots \otimes T_p M}_{s} \otimes \underbrace{T_p^* M \otimes \ldots \otimes T_p^* M}_{r}$.

A vector inside the tensor space $\otimes_s T_p M \otimes_r T_p^* M$ is in the form of

$$T_p = T_{p\,i_1 \ldots i_r}^{\;j_1 \ldots j_s} \frac{\partial}{\partial x^{j_1}}|_p \otimes \ldots \otimes \frac{\partial}{\partial x^{j_s}}|_p \otimes dx_p^{i_1} \otimes \ldots \otimes dx_p^{i_r}.$$

Under a change of coordinate $x \to x'$, the chain rule implies

$$T_p = T_{p\,i_1 \ldots i_r}^{\;j_1 \ldots j_s} \frac{\partial}{\partial x^{j_1}}|_p \otimes \ldots \otimes dx_p^{i_r}$$

$$= T_{p\,i_1 \ldots i_r}^{\;j_1 \ldots j_s} \frac{\partial x'^{l_1}}{\partial x^{j_1}}|_p \ldots \frac{\partial x'^{l_s}}{\partial x^{j_s}}|_p \frac{\partial x^{i_1}}{\partial x'^{k_1}}|_p \ldots \frac{\partial x^{i_r}}{\partial x'^{k_r}}|_p \frac{\partial}{\partial x'^{l_1}}|_p \otimes \ldots \otimes dx'^{k_r}_p$$

$$= T'^{\;l_1 \ldots l_s}_{p\,k_1 \ldots k_r} \frac{\partial}{\partial x'^{l_1}}|_p \otimes \ldots \otimes dx'^{k_r}_p,$$

where $T'^{\;l_1 \ldots l_s}_{p\,k_1 \ldots k_r} = T_{p\,i_1 \ldots i_r}^{\;j_1 \ldots j_s} \frac{\partial x'^{l_1}}{\partial x^{j_1}}|_p \ldots \frac{\partial x'^{l_s}}{\partial x^{j_s}}|_p \frac{\partial x^{i_1}}{\partial x'^{k_1}}|_p \ldots \frac{\partial x^{i_r}}{\partial x'^{k_r}}|_p$.

After discussing tangent and cotangent spaces, let us turn to functions between manifolds, and see how they induce maps between tangent spaces. Suppose $f : M \to N$ is a differentiable function between smooth manifolds, p is a point in M, V is an open domain in N containing $f(p)$, $U = f^{-1}(V)$, $x = (x^1, \ldots, x^m) : U \to \mathbb{R}^m$ is a chart of M, and $y = (y^1, \ldots, y^m) : V \to \mathbb{R}^n$ is a chart of N. Then, the map $F : y \circ f \circ x^{-1} : x(U) \to y(V)$ is a differentiable function between flat spaces, for which $x(U) \subseteq \mathbb{R}^m$ and $y(V) \subseteq \mathbb{R}^n$. Let $F = (F^1, \ldots, F^n)$, where $F^1, \ldots, F^n$ are real-valued functions, the differential of F at $x(p)$ is a matrix

$$dF|_{x(p)} = \begin{pmatrix} \partial_1 F^1|_{x(p)} & \cdots & \partial_m F^1|_{x(p)} \\ \vdots & & \vdots \\ \partial_1 F^n|_{x(p)} & \cdots & \partial_m F^n|_{x(p)} \end{pmatrix}_{n \times m},$$

which may be viewed as a linear map $dF|_{x(p)} : \mathbb{R}^m \to \mathbb{R}^n$ mapping $v = v^i e_i \mapsto dF|_{x(p)}(v) = v^i \partial_i F^j e'_j$, where $\{e_i\}_i$ is the usual basis for $\mathbb{R}^m$ and $\{e'_j\}_j$ is the usual basis for $\mathbb{R}^n$. Using this, we may define a linear function $f_{*p} : T_p M \to T_{f(p)} N$ such that for any tangent vector $X_p = X_p^i \frac{\partial}{\partial x^i}|_p$,

$$f_{*p} X_p = X_p^i \partial_i F^j|_{x(p)} \frac{\partial}{\partial y^j}|_{f(p)}.$$

By chain rule, one may straightforwardly check that the definition of f_{*p} is independent from the choice of coordinate charts x and y. This linear map $f_{*p} : T_pM \to T_{f(p)}N$ is known as the differential or push-forward of f at p.

Let $\alpha : (-\epsilon, \epsilon) \to M$ be a curve on M that generates X_p at $\alpha(0) = p$, i.e. $\dot{\alpha}(0) = X_p$. It induces a new curve $\beta = f \circ \alpha$ on N. Then,

$$f_{*p}(\dot{\alpha}(0)) = f_{*p}X_p = X_p^i \partial_i F^j|_{x(p)} \frac{\partial}{\partial y^j}|_{f(p)}$$

$$= \frac{dx^i \circ \alpha}{dt}|_{t=0} \partial_i (y^j \circ f \circ x^{-1})|_{x(p)} \frac{\partial}{\partial y^j}|_{f(p)} = \frac{dx^i \circ \alpha}{dt}|_{t=0} \frac{\partial (y^j \circ f)}{\partial x^i}|_p \frac{\partial}{\partial y^j}|_{f(p)}$$

$$= \frac{dy^j \circ f \circ \alpha}{dt}|_{t=0} \frac{\partial}{\partial y^j}|_{f(p)} = \frac{dy^j \circ \beta}{dt}|_{t=0} \frac{\partial}{\partial y^j}|_{f(p)} = \dot{\beta}(0).$$

The differential can also be extended to the linear map between tensor spaces $f_{*p} : \otimes_r T_pM \to \otimes_r T_{f(p)}N$ mapping

$$X_p \otimes Y_p \otimes \ldots \otimes Z_p \mapsto f_{*p}X_p \otimes f_{*p}Y_p \otimes \ldots \otimes f_{*p}Z_p.$$

The differentiable $f : M \to N$ can also induce a linear map $f^* : T^*_{f(p)}N \to T^*_pM$ (beware of the order) known as the pull-back of f. The definition is for any cotangent vector $\omega_{f(p)}$ in $T^*_{f(p)}N$, the pull-back map this to the cotangent vector $(f^*\omega)_p$ in T^*_pM, which is a linear functional mapping

$$X_p \mapsto (f^*\omega)_p(X_p) = \omega_{f(p)}(f_{*p}X_p).$$

By using the definition of the push-forward of f,

$$(f^*\omega)_p(X_p) = \omega_{f(p)}(f_{*p}X_p) = \omega_{f(p)_j} dy^j_{f(p)} \left(X_p^i \partial_i F^k|_{x(p)} \frac{\partial}{\partial y^k}|_{f(p)} \right)$$

$$= \omega_{f(p)_j} X_p^i \partial_i F^k|_{x(p)} dy^j_{f(p)} \left(\frac{\partial}{\partial y^k}|_{f(p)} \right) = \omega_{f(p)_j} X_p^i \partial_i F^k|_{x(p)} \delta^j_k$$

$$= \omega_{f(p)_j} X_p^i \partial_i F^j|_{x(p)} = \omega_{f(p)_j} \partial_i F^j|_{x(p)} dx^i_p(X_p).$$

In other words, the pull-back of $\omega_{f(p)}$ is

$$(f^*\omega)_p = \omega_{f(p)_j} \partial_i F^j|_{x(p)} dx^i_p.$$

Similarly, the pull-back $f^* : \otimes_r T^*_{f(p)}N \to \otimes_r T *_p M$ can also be defined on tensor spaces mapping

$$\omega_{f(p)} \otimes \xi_{f(p)} \otimes \ldots \otimes \psi_{f(p)} \mapsto (f^*\omega)_p \otimes (f^*\xi)_p \otimes \ldots \otimes (f^*\psi)_p.$$

Their is one remark about the order of composition of push-forward and pull-back. If $f : M \to N$ and $g : N \to K$ are differential functions between smooth manifolds, then for any point p in M, $(g \circ f)_{*p} = g_{*f(p)} \circ f_{*p}$ but $(g \circ f)^* = f^* \circ g^*$. The difference in ordering of composition is the distinction between contravariant functors (e.g. push-forward) and covariant functors (e.g. pull-back) in categorial algebra.

Lastly, may be not in the content of the general theory of relativity, the concept of a submanifold is commonly encountered in many fields in physics, such as the Liouville's theorem of an integrable classical mechanical system, some constraint problems in Statistical mechanics and quantum mechanics. A differentiable map $f : M \to N$ between smooth manifolds is said to be an immersion if for any point p in M, the push-forward $f_{*p} : T_pM \to T_{f(p)}N$ is injective. (As a direct result, N must be of larger dimension than M.) $f : M \to N$ is said to be an imbedding if it is an injective immersion, and is denoted by $f : M \hookrightarrow N$. In this case, M (usually treated as a subset of N) is called a submanifold of N. For example, a sphere, a saddle and a torus is a submanifold inside $\mathbb{R}^3$.

1.2 Tensors

The tangent and cotangent bundle of a smooth manifold M is defined to be the disjoint union of tangent and cotangent spaces over the whole manifold respectively, and are denoted by

$$TM = \bigcup_{p\in M} T_pM, \quad T^*M = \bigcup_{p\in M} T_p^*M.$$

The (r,s)-type tensor bundle is defined to be the disjoint union of tensor spaces

$$\otimes_r TM \otimes_s T^*M = \bigcup_{p\in M} \otimes_r T_pM \otimes_s T_p^*M.$$

For instance, TM is the (1,0)-type tensor bundle and T^*M is the (0,1)-type tensor bundle. For any point p in M, the fibre space of the bundle TM at p is the tangent space T_pM, and similarly the fibre spaces of T^*M and $\otimes_r TM \otimes_s T^*M$ at p are T_p^*M and $\otimes_r T_pM \otimes_s T_p^*M$ respectively. For every tensor bundle, there is a natural projection $\pi : \otimes_r TM \otimes_s T^*M \to M$ mapping $T_p \mapsto p$, for any T_p inside the fibre space $\otimes_r T_pM \otimes_s T_p^*M$. In other words, the fibre space could be expressed as $\pi^{-1}(p) = \otimes_r T_pM \otimes_s T_p^*M$.

Let us explore some properties of the tensor bundle. Suppose $\pi : TM \to M$ is the natural projection. For any coordinate chart $x : U \to \mathbb{R}^n$ of M, a canonical bijection $\Phi_x : \pi^{-1}(U) \to U \times \mathbb{R}^n$ can be induced such that for any point p in U, for any tangent vector $X_p = X_p^i \frac{\partial}{\partial x^i}|_p$ in T_pM, $\Phi_x(X_p) = \big(p, X_p^i e_i\big)$, where $\{e_i\}_i$ is the usual basis of $\mathbb{R}^n$. This canonical bijection is known as a trivialization of the tangent bundle over U, and defines a topology and differentiable structure for $\pi^{-1}(U)$, which is diffeomorphic to $U \times \mathbb{R}^n$. Note that the topology and differentiable structure is independent from the choice of coordinate charts. Suppose there is another chart $y : U \to \mathbb{R}^n$. Then, the corresponding canonical trivialization is $\Phi_y : \pi^{-1}(U) \to U \times \mathbb{R}^n$ mapping $X_p = {X'}_p^j \frac{\partial}{\partial y^j}|_p \mapsto \Phi_y(X_p) = \Big(p, {X'}_p^j e_j\Big)$. Combine with Φ_x, we have the bijection $\Phi_y \circ {\Phi_x}^{-1} : U \times \mathbb{R}^n \to$

$U \times \mathbb{R}^n$. For any vector $v = v^i e_i$ in $\mathbb{R}^n$, for any point p in U,

$$\Phi_y \circ {\Phi_x}^{-1}(p,v) = \Phi_y \left(v^i \frac{\partial}{\partial x^i}|_p \right) = \Phi_y \left(v^i \frac{\partial y^j}{\partial x^i}|_p \frac{\partial}{\partial y^j}|_p \right) = \left(p, v^i \frac{\partial y^j}{\partial x^i}|_p e_j \right),$$

and hence $\Phi_y \circ {\Phi_x}^{-1}$ must be diffeomorphic since the partial derivatives $\frac{\partial y^j}{\partial x^i}$ are differentiable functions on U and the differential of the transition function $t = y \circ x^{-1}$ is the matrix

$$t_* = dt = \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial y^n}{\partial x^1} & \cdots & \frac{\partial y^n}{\partial x^n} \end{pmatrix}_{n \times n},$$

which is always invertible. This implies the differentiable structures of $\pi^{-1}(U)$ induces by x and y are the same. Finally, combining all differentiable patches $\pi^{-1}(U)$ for any charts $x : U \to \mathbb{R}^n$ of M, this gives a differentiable structure to the whole bundle TM, which makes it a smooth manifold of dimensional $2n$, and the natural projection $\pi : TM \to M$ is differentiable.

Similar differentiable structure can be defined on the cotangent bundles T^*M and all tensor bundles $\otimes_r TM \otimes_s T^*M$, which are of manifold dimension $2n$ and $n(1+r+s)$ respectively. These objects are prototypes of vector bundle over M.

Definition of tensor fields: Suppose M is a smooth manifold. A (r,s)-type tensor field over M is a differentiable map $T : M \to \otimes_r TM \otimes_s T^*M$ such that $\pi \circ T$ is the identity map, where $\pi : \otimes_r TM \otimes_s T^*M \to M$ is the natural projection, i.e. $T : p \mapsto T_p \in \otimes_r T_pM \otimes_s T_p^*M$.

For instance, (1,0)-type tensors are vector fields (contravariant) and (0,1)-type tensors are differential 1-forms (covariant). Suppose $x : U \to \mathbb{R}^n$ is a coordinate chart of M. Since $\left\{ \frac{\partial}{\partial x^{i_1}}|_p \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}|_p \otimes dx_p^{j_1} \otimes \ldots \otimes dx_p^{j_s} \right\}_{i_1 \ldots i_r j_1 \ldots j_s}$ is a basis of the tensor space $\otimes_r T_pM \otimes_s T_p^*M$ for any point p in U, the tensor T at p can be expressed as

$$T_p = {T_p}^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}}|_p \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}|_p \otimes dx_p^{j_1} \otimes \ldots \otimes dx_p^{j_s}.$$

If we let $T^{i_1 \ldots i_r}_{j_1 \ldots j_s} : U \to \mathbb{R}$ be smooth functions mapping $p \mapsto {T_p}^{i_1 \ldots i_r}_{j_1 \ldots j_s}$, and $\frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}$ be the tensor field mapping $p \mapsto \frac{\partial}{\partial x^{i_1}}|_p \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}}|_p \otimes dx_p^{j_1} \otimes \ldots \otimes dx_p^{j_s}$, then we may write the tensor T as a summation of the basis tensors

$$T = T^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}. \tag{1}$$

There is one remark on convention in physics and mathematics. In physics, especially general relativity, the geometric objects are usually topologically trivial or only local properties are studied, and this means a single coordinate is already sufficient for discussion. To avoid the complication of definition, tensors are referred as a collection of indexed functions $T^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ instead of a function from the manifold to the tensor bundle as described above. In mathematics however, most interesting geometric object are topologically non-trivial, and there is no way a single coordinate chart could cover the whole manifold. Say we need u number of charts $x_1, x_2, \ldots, x_u$. For each chart x_α, we need a set of functions ${}^\alpha T^{i_1 \cdots i_r}_{j_1 \cdots j_s}$ representing the tensor on the domain of x_α. As you can see, if u is large, the indexed convention would become very messy, and therefore the bundle section definition $T : M \to \otimes_r TM \otimes_s T^*M$ is adopted. A second advantage the mathematics convention is the tensor rule of transformation is already included in Eq.(1), and can be deduced by the chain rule of the differential operators. Besides, comparing the two notations, the physical indexed convention seems inappropriate because a tensor T is a single entity and should not be separated into individual functions $T^{i_1 \cdots i_r}_{j_1 \cdots j_s}$.

The set of (r,s)-type tensors on M is denoted by $\Gamma(M, \otimes_r TM \otimes_s T^*M)$ or $D^r_s(M)$. Denote $D^0_0(M) = C^\infty(M)$, $D^r(M) = D^r_0(M)$, and $D_s(M) = D^0_s(M)$. Clearly, $D^r_s(M)$ is a real vector space, i.e. vector summation $S + T$ and scalar multiplication αT are defined in the usual way, and satisfy the necessary commutative, associative and distributive rules. (Actually, it is also a $C^\infty(M)$-module, where $C^\infty(M)$ is the ring of smooth real-valued functions on M. A module is a vector space, except the underlying field becomes an underlying ring with identity. See [5] for details.) $D^r_s(M)$ is the prime object of study in characterizing the geometric properties of a manifold in the differential point of view.

2 Differential Geometry

2.1 Connection and Covariant Differentiation

The mathematical language that describes the (differential) geometry of a smooth manifold M is the connection operator.

Definition of connection: A function $\nabla : D^1(M) \times D^1(M) \to D^1(M)$ mapping $(X, Y) \mapsto \nabla_X Y$ is said to be a connection for M if it satisfies

1. ($C^\infty(M)$-linear in the 1st argument) for any vector fields X, Y, Z, for any smooth functions f, g on M, $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.
2. ($\mathbb{R}$-linear in the 2nd argument) for any vector fields X, Y, Z on M, $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$.
3. (Leibnitz's product rule) for any vector fields X, Y, for any smooth functions f on M, $\nabla_X(fY) = f\nabla_X Y + X(f)Z$, where $X(f)$ is the smooth function on M mapping $p \mapsto X_p(f)$.

(Recall $X = X^i \frac{\partial}{\partial x^i}$ is a differential operator, and thus $X(f) = X^i \frac{\partial f}{\partial x^i} = df(X)$, where $df = \frac{\partial f}{\partial x^i} dx^i$.)

Suppose $x : U \to \mathbb{R}^n$ is a coordinate chart of M. For $i, j = 1, \ldots, n$, denote $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}$, where Γ^k_{ij} are smooth functions on U. This collection of indexed maps Γ_{ij}^k are known as the Riemann-Christoffel symbols. They are not tensors as they obey a different rule of coordinate transformation. Suppose there is another chart $x' : U \to \mathbb{R}^n$. Then,

$$\Gamma_{ij}^k \frac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial x'^s}{\partial x^j} \frac{\partial}{\partial x'^s} \right)$$

$$= \frac{\partial x'^s}{\partial x^j} \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x'^s} + \frac{\partial^2 x'^s}{\partial x^i \partial x^j} \frac{\partial}{\partial x'^s} = \frac{\partial x'^s}{\partial x^j} \nabla_{\frac{\partial x'^r}{\partial x^i} \frac{\partial}{\partial x'^r}} \frac{\partial}{\partial x'^s} + \frac{\partial^2 x'^s}{\partial x^i \partial x^j} \frac{\partial}{\partial x'^s}$$

$$= \frac{\partial x'^s}{\partial x^j} \frac{\partial x'^r}{\partial x^i} \nabla_{\frac{\partial}{\partial x'^r}} \frac{\partial}{\partial x'^s} + \frac{\partial^2 x'^s}{\partial x^i \partial x^j} \frac{\partial}{\partial x'^s} = \frac{\partial x'^s}{\partial x^j} \frac{\partial x'^r}{\partial x^i} {\Gamma'}_{rs}^{t} \frac{\partial}{\partial x'^t} + \frac{\partial^2 x'^s}{\partial x^i \partial x^j} \frac{\partial}{\partial x'^s}$$

$$= \left(\frac{\partial x'^s}{\partial x^j} \frac{\partial x'^r}{\partial x^i} {\Gamma'}_{rs}^{t} + \frac{\partial^2 x'^t}{\partial x^i \partial x^j} \right) \frac{\partial}{\partial x'^t} = \left(\frac{\partial x'^s}{\partial x^j} \frac{\partial x'^r}{\partial x^i} {\Gamma'}_{rs}^{t} + \frac{\partial^2 x'^t}{\partial x^i \partial x^j} \right) \frac{\partial x^k}{\partial x'^t} \frac{\partial}{\partial x^k}.$$

Comparing coefficient of $\frac{\partial}{\partial x^k}$, we have

$$\Gamma_{ij}^k = {\Gamma'}_{rs}^{t} \frac{\partial x'^r}{\partial x^i} \frac{\partial x'^s}{\partial x^j} \frac{\partial x^k}{\partial x'^t} + \frac{\partial^2 x'^t}{\partial x^i \partial x^j} \frac{\partial x^k}{\partial x'^t}. \tag{2}$$

There is an alternative but equivalent version of Eq.(2) involving differential form and no indices. The connection 1-form of the connection ∇ on U with respect to the chart $x : U \to \mathbb{R}^n$ is defined to be the $n \times n$-matrix valued differential 1-form denoted by $\Gamma = \left(\Gamma_j^i\right)_{n\times n} : U \to gl(n;\mathbb{R}) \otimes T^*U$ (here $gl(n;\mathbb{R}) = \mathbb{R}^{n\times n}$ is the space (or Lie algebra) of $n \times n$-real matrices, T^*U is the cotangent bundle over U), where $\Gamma_j^i = \Gamma_{sj}^i dx^s$ are the real-valued differential 1-form generated by the Chridtoffel symbols Γ_{ij}^k. Note that the differential 1-form Γ is still not a tensor since it is defined only on the local domain U, and there is no guarantee that it could be defined globally on M.

Consider the transition function $t = x \circ x'^{-1} : x'(U) \to x(U)$ between the charts x and x'. Its differential t_* can be treated as a non-singular $n \times n$-matrix valued function $t_* : U \to GL(n;\mathbb{R})$ (here $GL(n;\mathbb{R})$ is the (Lie) group of non-singular $n \times n$-matrices) mapping

$$p \mapsto t_{*p} = \begin{pmatrix} \frac{\partial x^1}{\partial x'^1}|_p & \cdots & \frac{\partial x^1}{\partial x'^n}|_p \\ \vdots & & \vdots \\ \frac{\partial x^n}{\partial x'^1}|_p & \cdots & \frac{\partial x^n}{\partial x'^n}|_p \end{pmatrix}_{n\times n}.$$

Then by a direct manipulation, Eq.(2) can be rewritten as

$$\Gamma = t_* \cdot \Gamma' \cdot t_*^{-1} + t_* \cdot d(t_*^{-1}), \tag{3}$$

where $d(t_*^{-1}) = \frac{\partial t_*^{-1}}{\partial x^k} dx^k$. Or written the matrix entries explicitly,

$$\Gamma_j^i = {\Gamma'}_s^r \frac{\partial x^i}{\partial x'^r} \frac{\partial x'^s}{\partial x^j} + \frac{\partial x^i}{\partial x'^r} d\left(\frac{\partial x'^r}{\partial x^i}\right).$$

Suppose X and Y are vector fields on U, using the Christoffel symbols, we may write

$$\nabla_X Y = \nabla_{X^i \frac{\partial}{\partial x^i}} \left(Y^j \frac{\partial}{\partial x^j} \right) = X^i \left(\frac{\partial Y^k}{\partial x^i} + Y^j \Gamma_{ij}^k \right) \frac{\partial}{\partial x^k}. \tag{4}$$

For instance, if the manifold is the flat space $\mathbb{R}^n$, the flat geometry is described by the vanishing of the Christoffel symbols, i.e. $\Gamma_{ij}^k = 0$. This means for any vector field X, Y, $\nabla_X Y = X^i \frac{\partial Y^k}{\partial x^i} \frac{\partial}{\partial x^k}$. If we adopt the usual basis $\{e_j\}_j$, and perform the transformation $\frac{\partial}{\partial x^j} \mapsto e_j$, then $\nabla_X Y = X^i \frac{\partial Y^k}{\partial x^i} e_k$ is just the direction derivative of the vector field $Y = Y^i e_i$ along the direction $X = X^i e_i$. However, in a curved space, the concept of parallelism differs from the Euclidean sence. And the change of the vector field Y would not be the sole contributor to its direction derivative. The curved geometry would also affect the value. The extra term $X^i Y^j \Gamma_{ij}^k \frac{\partial}{\partial x^k}$ measure the geometric contribution. (Details about parallelism would be further discussed.)

Let X be a fixed vector field on M. Using the Leibnitz product rule, the differential operator $\nabla_X : D^1(M) \to D^1(M)$ is $\mathbb{R}$-linear (i.e. treating $D^1(M)$ to be a $\mathbb{R}$-vector space rather than a $C^\infty(M)$-module). We could generalized it to be a linear map $\nabla_X : D_s^r(M) \to D_s^r(M)$ on higher order tensors. Firstly, on the space of smooth functions $C^\infty(M)$ (i.e. zero order tensors), the differentiation $\nabla_X : C^\infty(M) \to C^\infty(M)$ is defined to mapping $f \mapsto \nabla_X f = X(f)$. Recall $X = X^i \frac{\partial}{\partial x^i}$ is a differential operator, and hence $\nabla_X f = X^i \frac{\partial f}{\partial x^i} = df(X)$, where $df = \frac{\partial f}{\partial x^i} dx^i$ is a also a smooth function in $C^\infty(M)$. There is no contribution from the curved geometry as the definition is independent from the connection 1-form Γ. Secondly, the definition of the differentiation $\nabla_X : D_1(M) \to D_1(M)$ on the space of first order covariant tensors $D_1(M)$ can actually be deduced from the differentiation on smooth functions and vector fields. For any differential form ω in $D_1(M)$, for any vector field Y in $D^1(M)$, $\omega(Y)$ is a smooth function mapping $p \mapsto \omega_p(Y_p)$. Since ∇_X is a differential operator, it should satisfies the Leibnitz product rule, i.e.

$$\nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y).$$

Since by definition $\nabla_X(\omega(Y)) = X(\omega(Y))$, $\nabla_X \omega$ must satisfies

$$(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y),$$

(written explicitly, $X(\omega(Y)) = X(\omega_i Y^i) = X^j \frac{\partial \omega_i}{\partial x^j} Y^i + X^j \omega_i \frac{\partial Y^i}{\partial x^j}$) and this may treated as the definition of the covariant tensor $\nabla_X \omega$. On a local domain U covered by the chart x, we could use the Christoffel symbols to write down the differentiation in index form. For any vector field Y,

$$(\nabla_X \omega)(Y) = X^j \frac{\partial \omega_i}{\partial x^j} Y^i + X^j \omega_i \frac{\partial Y^i}{\partial x^j} - \omega(\nabla_X Y)$$

$$= X^j \frac{\partial \omega_i}{\partial x^j} Y^i + X^j \omega_i \frac{\partial Y^i}{\partial x^j} - \omega \left(X^i \left(\frac{\partial Y^k}{\partial x^i} + Y^j \Gamma_{ij}^k \right) \frac{\partial}{\partial x^k} \right)$$

$$= X^j \frac{\partial \omega_i}{\partial x^j} Y^i + X^j \omega_i \frac{\partial Y^i}{\partial x^j} - \omega_k X^i \left(\frac{\partial Y^k}{\partial x^i} + Y^j \Gamma^k_{ij} \right)$$

$$= X^i Y^j \frac{\partial \omega_j}{\partial x^i} - \omega_k X^i Y^j \Gamma^k_{ij} = X^i \left(\frac{\partial \omega_j}{\partial x^i} - \omega_k \Gamma^k_{ij} \right) dx^j(Y).$$

And hence, by eliminating the arbitrary vector field Y,

$$\nabla_X \omega = X^i \left(\frac{\partial \omega_j}{\partial x^i} - \omega_k \Gamma^k_{ij} \right) dx^j. \tag{5}$$

When compare with the differentiation of a (contravariant) vector field Eq.(4), there is a change of sign in the term involving the Christoffel symbols.

In general, by using Leibnitz product rule, the differentiation of the tensor $Y \otimes \ldots \otimes Z \otimes \omega \otimes \ldots \otimes \xi$ in $D^r_s(M)$ is defined to be

$$\nabla_X (Y \otimes \ldots \otimes \xi) = (\nabla_X Y) \otimes \ldots \otimes \xi + \ldots + Y \otimes \ldots \otimes (\nabla_X \xi)$$

This means $\nabla_X : D^r_s(M) \to D^r_s(M)$ operates on tensors of any contravariant and covariant order. Written locally in index form, the differentiation of a (r,s)-type tensor T is $(\nabla_X T)^{i_1 \ldots i_r}_{j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes dx^{j_s}$, where

$$(\nabla_X T)^{i_1 \ldots i_r}_{j_1 \ldots j_s} = X^k \frac{\partial T^{i_1 \ldots i_r}_{j_1 \ldots j_s}}{\partial x^k} + X^k \left(T^{\mu i_2 \ldots i_r}_{j_1 \ldots j_s} \Gamma^{i_1}_{k\mu} + \ldots + T^{i_1 \ldots i_{r-1}\mu}_{j_1 \ldots j_s} \Gamma^{i_r}_{k\mu} \right)$$
$$- X^k \left(T^{i_1 \ldots i_r}_{\mu j_2 \ldots j_s} \Gamma^{\mu}_{k j_1} + \ldots + T^{i_1 \ldots i_r}_{j_1 \ldots j_{s-1}\mu} \Gamma^{\mu}_{k j_s} \right). \tag{6}$$

The tensor product $\otimes$ maps lower order tensors to higher order one. There is another binary operator $\rfloor$ known as the contraction mapping higher order tensors to lower order one. Consider the two tensors spaces $D^r_s(M)$ and $D^t_u(M)$. If $s \geq t$, then the contraction is the map $\rfloor : D^r_s(M) \times D^t_u(M) \to D^r_{s-t+u}(M)$ mapping $(S,T) \mapsto S \rfloor T = (S \rfloor T)^{i_1 \ldots i_r}_{j_1 \ldots j_{s-t+u}} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes dx^{j_{s-t+u}}$, where

$$(S \rfloor T)^{i_1 \ldots i_r}_{j_1 \ldots j_{s-t+u}} = S^{i_1 \ldots i_r}_{\mu_1 \ldots \mu_t j_1 \ldots j_{s-t}} T^{\mu_1 \ldots \mu_t}_{j_{s-t+1} \ldots j_{s-t+u}},$$

or if $s \leq t$, then $\rfloor : D^r_s(M) \times D^t_u(M) \to D^{r-s+t}_u(M)$ mapping $(S,T) \mapsto S \rfloor T$,

$$(S \rfloor T)^{i_1 \ldots i_{r-s+t}}_{j_1 \ldots j_u} = S^{i_1 \ldots i_r}_{\mu_1 \ldots \mu_s} T^{\mu_1 \ldots \mu_s i_{r+1} \ldots i_{r-s+t}}_{j_1 \ldots j_u}.$$

This means the lower indices of S contract with the upper indices of T untill one of them has run out of indices. One can check that the definition is invariant under change of coordinate charts, and hence the contraction is still a tensor. For example, if ω is in $D_1(M)$ and X is in $D^1(M)$, then $\omega \rfloor X = \omega_i X^i = \omega(X)$, if ξ is in $D_2(M)$, then $\xi \rfloor X = \xi_{ij} X^i dx^j$, or if Y is in $D^2(M)$, then $\omega \rfloor Y = \omega_i Y^{ij} \frac{\partial}{\partial x^j}$ and $\xi \rfloor Y = \xi_{ij} Y^{ij}$.

Using the contraction, we could define the covariant differential operator of a given connection.

Definition of covariant differentiation: The covariant differential is defined to be the operator $D : D^r_s(M) \to D^r_{s+1}(M)$ mapping $T \mapsto DT$ such that $DT \rfloor X = \nabla_X T$ for any vector field X in $D^1(M)$.

Note that the definition is unique since for any local chart $x : U \to \mathbb{R}^n$, the covariant differential can be written locally in index form, $DT = (DT)^{i_1 \ldots i_r}_{k j_1 \ldots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^k \otimes dx^{j_1} \ldots \otimes dx^{j_s}$, where

$$T^{i_1 \ldots i_r}_{j_1 \ldots j_s ; k} = (DT)^{i_1 \ldots i_r}_{k j_1 \ldots j_s} = \frac{\partial T^{i_1 \ldots i_r}_{j_1 \ldots j_s}}{\partial x^k} + \Gamma^{i_1}_{k\mu} T^{\mu i_2 \ldots i_r}_{j_1 \ldots j_s} + \ldots + \Gamma^{i_r}_{k\mu} T^{i_1 \ldots i_{r-1} \mu}_{j_1 \ldots j_s}$$
$$- \Gamma^{\mu}_{k j_1} T^{i_1 \ldots i_r}_{\mu j_2 \ldots j_s} + \ldots + \Gamma^{\mu}_{k j_s} T^{i_1 \ldots i_r}_{j_1 \ldots j_{s-1} \mu}. \tag{7}$$

For instance, for any vector field X, $DX = \left(\frac{\partial X^k}{\partial x^i} + X^j \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k} \otimes dx^i$, and for any covector field ω, $D\omega = \left(\frac{\partial \omega_j}{\partial x^i} - \Gamma^k_{ij} \omega_k \right) dx^i \otimes dx^j$.

For readers who are familiar with exterior algebra, the subspace of alternating (or skew-symmetric) tensors in $\otimes_s T^*_p M$ is denoted by $\bigwedge_s T^*_p M$ for any point p in M. Similar to the tensor bundle $\otimes_s T^* M$ the wedge bundle is defined to be the disjoint union $\bigwedge_s T^* M = \bigcup_{p \in M} \bigwedge_s T^* M$, and $\Gamma\left(M, \bigwedge_s T^* M\right)$ is denoted as the subspace of alternating tensor fields (or differential s-form) inside $D_s(M) = \Gamma\left(M, \otimes_s T^* M\right)$. The exterior differential operator $d : \Gamma\left(M, \bigwedge_s T^* M\right) \to \Gamma\left(M, \bigwedge_{s+1} T^* M\right)$ is defined to be $d = (s+1) Alt \circ d$, where $Alt : \Gamma\left(M, \otimes_{s+1} T^* M\right) \to \Gamma\left(M, \bigwedge_{s+1} T^* M\right)$ is the alternation, the d on the left hand side is the exterior differential, and the d on the right hand side is the differential of tensors mapping $\omega \mapsto d\omega = \frac{\partial \omega_{j_1 \ldots j_s}}{\partial x^k} dx^k \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s}$. In other words, for any alternting differential form η,

$$d\eta = (s+1) Alt(d\eta) = (s+1) Alt \left(\frac{\partial \eta_{j_1 \ldots j_s}}{\partial x^k} dx^k \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s} \right)$$
$$= (s+1) \frac{\partial \eta_{j_1 \ldots j_s}}{\partial x^k} Alt\left(dx^k \otimes dx^{j_1} \otimes \ldots \otimes dx^{j_s} \right)$$
$$= \frac{s+1}{(s+1)!} \frac{\partial \eta_{j_1 \ldots j_s}}{\partial x^k} dx^k \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_s}$$
$$= \sum_{j_1 < \ldots < j_s} \sum_{k=1}^{n} \frac{\partial \eta_{j_1 \ldots j_s}}{\partial x^k} dx^k \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_s}.$$

Now, we could define the covariant differential operator D on $\Gamma\left(M, \bigwedge_s T^* M\right)$ similarly mapping $\eta \mapsto D\eta = (s+1) Alt(D\eta)$, where the D on the left hand side of the equation is the exterior covariant differential operator on differential forms, and the D on the right hand side of the equation is the covariant differential operator on tensors described before. For a simple review on exterior algebra, differential forms, alternation and wedge product, the readers could refer to [6].

2.2 Parallel Displacement and Geodesic

Suppose X is a fixed vector field on the smooth manifold M equipped with a connection ∇. A tensor field T in $D_s^r(M)$ is said to be horizontal with respect to X if it satisfies the linear differential equation $DT\rfloor X = \nabla_X T = 0$. This means Eq.(6) is zero. There are in a total n^{r+s} equations and T has n^{r+s} components, and hence $\nabla_X T = 0$ has a non-empty solution set for T. In this section, we shall focus only on vector fields for simplicity, but similar arguments could be applied to tensors of arbitrary orders.

Let $\alpha : (-t_0, t_0) \to M$ is a smooth curve on M. For every t in $(-t_0, t_0)$, $\dot{\alpha}(t)$ is a tangent vector inside $T_{\alpha(t)}M$, and hence $\dot{\alpha} : (-t_0, t_0) \to TM$ is a vector field on M mapping $t \mapsto \dot{\alpha}(t)$.

Definition of horizontal vector field: A vector Y in $D^1(M)$ is said to be horizontal with respect to the curve α if $\frac{DY}{dt} = \nabla_{\dot{\alpha}} Y = 0$ along the curve.

For simplicity, let assume the curve could be covered by a single coordinate chart $x : U \to \mathbb{R}^n$ (otherwise coordinate transformation should be applied whenever necessary). The tangent can be written as $\dot{\alpha}(t) = \frac{dx^i \circ \alpha}{dt}(t) \frac{\partial}{\partial x^i}|_{\alpha(t)}$. And hence, the vanishing of the covariant derivative means (for simplicity, write $x^i = x^i \circ \alpha$ and $Y^i = Y^i \circ \alpha$ as functions of t)

$$\frac{DY}{dt} = \frac{dx^i}{dt} \left(\frac{\partial Y^k}{\partial x^i} + Y^j \Gamma^k_{ij} \right) \frac{\partial}{\partial x^k} = \left(\frac{dY^k}{dt} + \frac{dx^i}{dt} \Gamma^k_{ij} Y^j \right) \frac{\partial}{\partial x^k} = 0, \quad (8)$$

i.e. we have the first order differential system $\left(\frac{d}{dt} + \frac{dx^i}{dt} \Gamma^k_{ij} \right) Y^j = 0$.

By the fundamental theorem of differential system, given any tangent vector Y_0 at the point $\alpha(0)$ (tangential to the manifold but not necessarily to the curve α), there is a unique vector field $Y(t)$ satisfying Eq.(8) and the initial condition $Y(0) = Y_0$. This curve $Y : t \mapsto Y(t)$ on the tangent bundle TM is known as the horizontal lifting of α with respect to the fixed initial vector Y_0. The terminology horizontal refers to the fact that Y satisfies the geometric governing equation Eq.(8), and lifting means the curves α and Y are related by $\pi \circ Y = \alpha$, where $\pi : TM \to M$ is the canonical projection.

Definition of parallel displacement: The parallel displacement by the curve α is defined to be the vector space isomorphism, also denoted by the symbol of the curve, $\alpha_{0\to t} : T_{\alpha(0)}M \to T_{\alpha(t)}M$ mapping $Y_0 \mapsto Y(t)$ as described above.

The parallel displacement operator acts as a transporting agent that transports an initial vector Y_0 at $\alpha(0)$ parallelly to a final vector $Y(t)$ at $\alpha(t)$. From Eq.(8), a horizontal vector field in a flat space is a constant vector field. This reduces to the common knowledge of parallelism in Euclidean space, i.e. the parallel displacement of a vector along a curve is the transportation of vector *parallel* to the starting one. Except in non-Euclidean geometry, due to the non-vanishing Christoffel symbols, parallelism is not characterized by the constancy of a vector.

Definition of geodesic curves: A curve α is a geodesic if its tangent is horizontal along itselt, i.e. $\frac{D\dot{\alpha}}{dt} = \nabla_{\dot{\alpha}}\dot{\alpha} = 0$.

Written in local coordinate $x = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$, this means (again for simplicity, write $x^i = x^i \circ \alpha$ as a function of t)

$$\frac{D\dot{\alpha}}{dt} = \left(\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)\frac{\partial}{\partial x^k} = 0, \tag{9}$$

i.e. the second order differential system $\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0$. In flat space, Eq.(9) implies the tangent of a geodesic is constant, i.e. it is a straight line.

Although this geodesic description seems differs from the common calculus of variation interpretation of the shortest curve between two points, this differential geometric definition is a more general one. In the later Riemannian geometry section, when the manifold equips a metric, the length of a curve could be discussed. One may check a geodesic with respect to the Riemannian connection is (locally) the shortest path between two points.

2.3 Lie differentiation, Torsion and Curvature

Suppose X, Y are vector fields in $D^1(M)$. Recall they are differential operators, and hence we can consider their composition

$$X \circ Y = X^i\frac{\partial}{\partial x^i}\left(Y^j\frac{\partial}{\partial x^j}\right) = X^iY^j\frac{\partial^2}{\partial x^i\partial x^j} + X^i\frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial x^j},$$

$$Y \circ X = X^iY^j\frac{\partial^2}{\partial x^i\partial x^j} + Y^i\frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j}.$$

These are not vector fields since they are second order differential operators. However, their difference is first order and obey the tensor rule of transformation, and hence is a vector field.

Definition of Lie bracket: The Lie bracket on vector field is defined to be the $\mathbb{R}$-bilinear map $[*, *] : D^1(M) \times D^1(M) \to D^1(M)$ mapping $(X, Y) \mapsto [X, Y] = X \circ Y - Y \circ X$.

Written in index form,

$$[X, Y] = \left(X^i\frac{\partial Y^j}{\partial x^i} - Y^i\frac{\partial X^j}{\partial x^i}\right)\frac{\partial}{\partial x^j}. \tag{10}$$

The covariance of the Lie bracket is already guaranteed by its definition. If one prefers to check it explicitly, for each coefficient,

$$X^i\frac{\partial Y^j}{\partial x^i} - Y^i\frac{\partial X^j}{\partial x^i} = X'^i\frac{\partial}{\partial x'^i}\left(Y'^k\frac{\partial x^j}{\partial x'^k}\right) - Y'^i\frac{\partial}{\partial x'^i}\left(X'^k\frac{\partial x^j}{\partial x'^k}\right)$$

$$= \left(X'^i\frac{\partial Y'^k}{\partial x'^i} - Y'^i\frac{\partial X'^k}{\partial x'^i}\right)\frac{\partial x^j}{\partial x'^k} + X'^iY'^k\frac{\partial^2 x^j}{\partial x'^i\partial x'^k} - X'^kY'^i\frac{\partial^2 x^j}{\partial x'^i\partial x'^k}$$

$$= \left(X'^i \frac{\partial Y'^k}{\partial x'^i} - Y'^i \frac{\partial X'^k}{\partial x'^i} \right) \frac{\partial x^j}{\partial x'^k}.$$

Below are some properties of the Lie bracket that the reader could check straightforwardly.

1. As stated in the definition, the Lie bracket is $\mathbb{R}$-bilinear, i.e. for any vector fields X, Y and Z, for any real scalars a and b, $[aX + bY, Z] = a[X, Z] + b[Y, Z]$ and $[X, aY + bZ] = a[X, Y] + b[X, Z]$.

2. It is alternating, i.e. for any vector fields X and Y, $[X, Y] = -[Y, X]$.

3. (Jacobi's identity) For any vector fields X, Y and Z, $[[X, Y], Z]+[[Y, Z], X]+[[Z, X], Y] = 0$.

4. (Leibnitz product rule) For any vector fields X and Y, for any smooth functions f, $[fX, Y] = f[X, Y] - Y(f)X$ and $[X, fY] = f[X, Y] + X(f)Y$.

(Recall X, Y are differential operators and $X(f) = df(X)$, $Y(f) = df(Y)$.) The first three identities are the properties that any Lie bracket should possess in any Lie algebra. The last identity implies the operator $L_X : D^1(M) \to D^1(M)$ mapping $Y \mapsto L_X Y = [X, Y]$ is a derivation, i.e. is $\mathbb{R}$-linear and satisfies the Leibniz product rule. The differential operator L_X is known as the Lie differentiation. (The reason for the name differentiation would be skipped as it inolves the concept of a 1-parameter group of transformation generated by X, which is not discussed in this notes. For detail, please refer to [3].)

Note that the Lie bracket and Lie differentiation can be defined in any smooth manifold independent from a connection since the Christoffel symbols are not used in the definition, and hence these are not relevant to the differential geometric property of a manifold. However, with the help of the Lie bracket, two important (particularly the latter one) characterizations of the geometry of a manifold can be defined.

Definition of torsion: With respect to a connection ∇, the torsion is an operator $T : D^1(M) \times D^1(M) \to D^1(M)$ such that for any vector fields X and Y, $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$

Write $End_{C^\infty(M)}\left(D^1(M)\right)$ to be the space of all endomorphisms on $D^1(M)$ (i.e. $C^\infty(M)$-linear $D^1(M) \to D^1(M)$ maps).

Definition of curvature: With respect to a connection ∇, the curvature is an operator $R : D^1(M) \times D^1(M) \to End_{C^\infty(M)}\left(D^1(M)\right)$ mapping $(X, Y) \mapsto R(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$.

In other words, for any vector field Z, $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$. Let us explore some properties of the torsion and curvature operator.

Theorem 2.1 *The torsion operator $T : D^1(M) \times D^1(M) \to D^1(M)$ is skew-symmetric and $C^\infty(M)$-bilinear.*

Proof: By a simple checking, it is easy to see T is skew-symmetric, i.e. $T(X,Y) = -T(Y,X)$. Clearly, by the property of the connection $T(X+Y,Z) = T(X,Z) + T(Y,Z)$. For any smooth function f in $C^\infty(M)$,

$$T(fX,Y) = \nabla_{fX}Y - \nabla_Y(fX) - [fX,Y]$$

$$= f\nabla_X Y - f\nabla_Y X - Y(f)X - f[X,Y] + Y(f)X$$

$$= f\left(\nabla_X Y - \nabla_Y X - [X,Y]\right) = fT(X,Y).$$

Thus, T is $C^\infty(M)$-bilinear. Q.E.D.

This theorem implies T can be treated as a (1,2)-type tensor in $D^1_2(M)$. Define $T^k_{ij}\frac{\partial}{\partial x^k} = T\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$, or equivalently, $T^k_{ij} = dx^k \rfloor T\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$. Since it satisfies the transformation rule

$$T^k_{ij} = dx^k\left(T\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\right) = \frac{\partial x^k}{\partial x'^t}dx'^t\left(T\left(\frac{\partial x'^r}{\partial x^i}\frac{\partial}{\partial x'^r}, \frac{\partial x'^s}{\partial x^j}\frac{\partial}{\partial x'^s}\right)\right)$$

$$= \frac{\partial x'^r}{\partial x^i}\frac{\partial x'^s}{\partial x^j}\frac{\partial x^k}{\partial x'^t}dx'^t\left(T\left(\frac{\partial}{\partial x'^r}, \frac{\partial}{\partial x'^s}\right)\right) = \frac{\partial x'^r}{\partial x^i}\frac{\partial x'^s}{\partial x^j}\frac{\partial x^k}{\partial x'^t}{T'}^t_{rs},$$

if we write $T = T^k_{ij}\frac{\partial}{\partial x^k} \otimes dx^i \otimes dx^j$ on each local chart x of M, then it is a (1,2)-type tensor defined globlally on M. (To be more specific, T is a vector-valued differential 2-form since $T^k_{ij} = -T^k_{ji}$, i.e. $T = \frac{\partial}{\partial x^k} \otimes T^k$, for which each component is a differential 2-form $T^k = \sum_{i<j} T^k_{ij} dx^i \wedge dx^j$, where $dx^i \wedge dx^j = dx^i \otimes dx^j - dx^j \otimes dx^i$.) This tensor is called the torsion tensor.

Written in terms of the Christoffel symbols,

$$T^k_{ij}\frac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^i} - \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right]$$

$$= \left(\Gamma^k_{ij} - \Gamma^k_{ji}\right)\frac{\partial}{\partial x^k},$$

since $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$. This means

$$T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}. \tag{11}$$

As expected, torsion vanishes in flat spaces.

After studying the torsion tensor, let turn to curvature. While the torsion operator is a first order differential of vector fields, curvature is a second order operation. The reader may check straightforwardly that for any vector fields X, Y and Z, $R(X,Y)Z = DDZ\rfloor(X \otimes Y - Y \otimes X)$, where D is the covariant differential operator. Hence, we may interpret the curvature as D^2.

Theorem 2.2 *For any vector fields X and Y, $R(X,Y): D^1(M) \to D^1(M)$ is $C^\infty(M)$-linear.*

Proof: For any smooth function f in $C^\infty(M)$, for any vector field Z,

$$R(X,Y)fZ = \nabla_X\nabla_Y(fZ) - \nabla_Y\nabla_X(fZ) - \nabla_{[X,Y]}(fZ)$$

$$= \nabla_X\left(f\nabla_Y Z + Y(f)Z\right) - \nabla_Y\left(f\nabla_X Z + X(f)Z\right) - f\nabla_{[X,Y]}Z - \left([X,Y](f)\right)Z$$

$$= f\nabla_X\nabla_Y Z + X(f)\nabla_Y Z + Y(f)\nabla_X Z + X\circ Y(f)Z - f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z$$

$$-X(f)\nabla_Y Z - Y\circ X(f)Z - f\nabla_{[X,Y]}Z - X\circ Y(f)Z + Y\circ X(f)Z$$

$$= f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - f\nabla_{[X,Y]}Z = fR(X,Y)Z.$$

Thus, $R(X,Y): D^1(M)\to D^1(M)$ is an endomorphism of the $C^\infty(M)$-module $D^1(M)$. Q.E.D.

Note that every endomorphism E in $End_{C^\infty(M)}\left(D^1(M)\right)$ corresponds to a unique (1,1)-type tensor in $D^1_1(M)$. Define $E^i_j\frac{\partial}{\partial x^i} = E\left(\frac{\partial}{\partial x^j}\right)$, i.e. $E^i_j = dx^i\rfloor E\left(\frac{\partial}{\partial x^j}\right)$. Since E is $C^\infty(M)$-linear, E^i_j satisfies the tensor transformation

$$E^i_j = dx^i\left(E\left(\frac{\partial}{\partial x^j}\right)\right) = \frac{\partial x^i}{\partial x'^k}dx'^k\left(E\left(\frac{\partial x'^l}{\partial x^j}\frac{\partial}{\partial x'^l}\right)\right)$$

$$= \frac{\partial x^i}{\partial x'^k}\frac{\partial x'^l}{\partial x^j}dx'^k\left(E\left(\frac{\partial}{\partial x'^l}\right)\right) = \frac{\partial x^i}{\partial x'^k}\frac{\partial x'^l}{\partial x^j}dx'^k\left(E'^u_l\frac{\partial}{\partial x'^u}\right)$$

$$= \frac{\partial x^i}{\partial x'^k}\frac{\partial x'^l}{\partial x^j}E'^u_l\delta^k_u = \frac{\partial x^i}{\partial x'^k}\frac{\partial x'^l}{\partial x^j}E'^k_l,$$

if we set $E = E^i_j\frac{\partial}{\partial x^i}\otimes dx^j$ on all local charts x, this defines a global (1,1)-tensor E in $D^1_1(M)$. Conversely, for any (1,1)-tensor T in $D^1_1(M)$, it defines an endomorphism $T: X\mapsto T(X) = T\rfloor X = T^i_j X^j\frac{\partial}{\partial x^i}$ on $D^1(M)$. Hence, there is a natural isomorphism $End_{C^\infty(M)}\left(D^1(M)\right)\cong D^1_1(M)$, and the two spaces can be treated to be the same. Apply to the curvature operator, $R: D^1(M)\times D^1(M)\to End_{C^\infty(M)}\left(D^1(M)\right) = D^1_1(M)$. Write $R(X,Y) = R(X,Y)^i_j\frac{\partial}{\partial x^i}\otimes dx^j$, and define the correspondent $R^i_j: D^1(M)\times D^1(M)\to C^\infty(M)$ mapping $(X,Y)\mapsto R^i_j(X,Y) = R(X,Y)^i_j$.

Theorem 2.3 *$R^i_j: D^1(M)\times D^1(M)\to C^\infty(M)$ is skew-symmetric and $C^\infty(M)$-bilinear, i.e. R^i_j is a differential 2-form.*

Proof: By a direct checking of the definition of the curvature operator, it is obvious that $R(X,Y) = -R(Y,X)$. For any smooth function f in $C^\infty(M)$,

$$R(fX,Y) = \nabla_{fX}\nabla_Y - \nabla_Y\nabla_{fX} - \nabla_{[fX,Y]}$$

$$= f\nabla_X\nabla_Y - \nabla_Y\left(f\nabla_X\right) - \nabla_{(f[X,Y]-Y(f)X)}$$

$$= f\nabla_X\nabla_Y - f\nabla_Y\nabla_X - Y(f)\nabla_X - f\nabla_{[X,Y]} + Y(f)\nabla_X$$

$$= f\nabla_X\nabla_Y - f\nabla_Y\nabla_X - f\nabla_{[X,Y]} = fR(X,Y).$$

Since R is alternating, its components R^i_j must also be alternating, in other words $R^i_j(X,Y) = -R^i_j(Y,X)$. Write $R^i_{jkl} = R^i_j(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l})$. Due to its $C^\infty(M)$-bilinearity, the covariance of R^i_{jkl} can be checked in a similar manner as T^k_{ij}, and hence if we define $R^i_j = R^i_{jkl}dx^k \otimes dx^l$ on any local chart x, R^i_j is a (0,2)-type tensor. Since $R^i_{jkl} = -R^i_{jlk}$, $R^i_j = \sum_{k<l} R^i_{jkl}dx^k \wedge dx^l$ is a differential 2-form. Q.E.D.

This means the curvature is $R = \frac{\partial}{\partial x^i} \otimes dx^j \otimes R^i_j$. Written explicitly, $R = R^i_{jkl}\frac{\partial}{\partial x^i} \otimes dx^j \otimes dx^k \otimes dx^l$ on any chart x of M, i.e. the curvature R can be treated as a global (1,3)-type tensor in $D^1_3(M)$. This tensor is known as the curvature tensor.

Written in terms of Christoffel symbols,

$$R^i_{jkl}\frac{\partial}{\partial x^i} = R\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right)\frac{\partial}{\partial x^j}$$

$$= \nabla_{\frac{\partial}{\partial x^k}}\nabla_{\frac{\partial}{\partial x^l}}\frac{\partial}{\partial x^j} - \nabla_{\frac{\partial}{\partial x^l}}\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j} - \nabla_{[\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}]}\frac{\partial}{\partial x^j}$$

$$= \nabla_{\frac{\partial}{\partial x^k}}\left(\Gamma^t_{lj}\frac{\partial}{\partial x^t}\right) - \nabla_{\frac{\partial}{\partial x^l}}\left(\Gamma^t_{kj}\frac{\partial}{\partial x^t}\right)$$

$$= \Gamma^t_{lj}\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^t} + \frac{\partial\Gamma^t_{lj}}{\partial x^k}\frac{\partial}{\partial x^t} - \Gamma^t_{kj}\nabla_{\frac{\partial}{\partial x^l}}\frac{\partial}{\partial x^t} - \frac{\partial\Gamma^t_{kj}}{\partial x^l}\frac{\partial}{\partial x^t}$$

$$= \left(\frac{\partial\Gamma^t_{lj}}{\partial x^k} - \frac{\partial\Gamma^t_{kj}}{\partial x^l} + \Gamma^m_{lj}\Gamma^t_{km} - \Gamma^m_{kj}\Gamma^t_{lm}\right)\frac{\partial}{\partial x^t}.$$

Or in each component,

$$R^i_{jkl} = \frac{\partial\Gamma^i_{lj}}{\partial x^k} - \frac{\partial\Gamma^i_{kj}}{\partial x^l} + \Gamma^m_{lj}\Gamma^i_{km} - \Gamma^m_{kj}\Gamma^i_{lm}. \tag{12}$$

Recall locally on the domain U of the chart x, we could define a $n \times n$-matrix valued connection 1-form $\Gamma = \left(\Gamma^i_j\right)_{n\times n} : U \to gl(n;\mathbb{R}) \otimes T^*U$, where $\Gamma^i_j = \Gamma^i_{sj}dx^s$. For curvature, we could treat it as a $n\times n$-matrix valued differential 2-form $R = \left(R^i_j\right)_{n\times n} : U \to gL(n;\mathbb{R}) \otimes \bigwedge_2 T^*M$, where $R^i_j = R^i_{jkl}dx^k \otimes dx^l = \sum_{k<l} R^i_{jkl}dx^k \wedge dx^l$. This matrix valued differential 2-form is known as the curvature 2-form. Hence by a straightforward computation, Eq.(12) is equivalent to the Cartan's structure equation

$$R^i_j = d\Gamma^i_j + \Gamma^i_m \wedge \Gamma^m_j, \tag{13}$$

for which d is the exterior differential operator. Or in matrix form, it is simply $R = d\Gamma + \Gamma \wedge \Gamma$, where $\Gamma \wedge \Gamma = \left(\Gamma^i_m \wedge \Gamma^m_j\right)_{n\times n}$. Clearly, curvature vanishes in flat space as $\Gamma = 0$.

Theorem 2.4 *If the manifold M is equipped with a torsion free ($T = 0$) connection, then*

1. *(Jacobi's identity) for any vector fields X, Y and Z, $R(X,Y)Z+R(Y,Z)X+R(Z,X)Y=0$, and*

2. *(Bianchi's identity) $DR^i_j=0$, where R^i_j is the curvature 2-form and D is the exterior covariant differential operator.*

Proof: For the Jacobi's identity,

$$R(X,Y)Z+R(Y,Z)X+R(Z,X)Y=\nabla_X\nabla_Y Z-\nabla_Y\nabla_X Z-\nabla_{[X,Y]}Z$$

$$+\nabla_Y\nabla_Z X-\nabla_Z\nabla_Y X-\nabla_{[Y,Z]}X+\nabla_Z\nabla_X Y-\nabla_X\nabla_Z Y-\nabla_{[Z,X]}Y$$

$$=\nabla_X\left(\nabla_Y Z-\nabla_Z Y\right)+\nabla_Y\left(\nabla_Z X-\nabla_X Z\right)+\nabla_Z\left(\nabla_X Y-\nabla_Y X\right)$$

$$-\nabla_{[X,Y]}Z-\nabla_{[Y,Z]}X-\nabla_{[Z,X]}Y$$

$$=\nabla_X[Y,Z]+\nabla_Y[Z,X]+\nabla_Z[X,Y]-\nabla_{[X,Y]}Z-\nabla_{[Y,Z]}X-\nabla_{[Z,X]}Y$$

$$=[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0,$$

where the last three equalities holds because of the vanishing torsion and the Jacobi's identity of Lie bracket. Next, for the Bianchi's identity, for any point p in the manifold M, let $x=(x^1,\ldots,x^n):U\to\mathbb{R}^n$ be a local chart around P such that $x(p)=0$, if we perform the coordinate transformation $x\mapsto x'$ such that $x^k={x'}^k-\frac{1}{2}\Gamma^k_{ij}(p){x'}^i{x'}^j$, where Γ^k_{ij} is the Christoffel symbols with respect to the chart x (note that the transformation may not be invertible on the whole domain U, but it must be invertible on a smaller neighbourhood around p), then by the transformation Eq.(2)

$${\Gamma'}^k_{ij}(p)=\Gamma^t_{rs}(p)\frac{\partial x^r}{\partial {x'}^i}|_p\frac{\partial x^s}{\partial {x'}^j}|_p\frac{\partial {x'}^k}{\partial x^t}|_p+\frac{\partial^2 x^t}{\partial {x'}^i\partial {x'}^j}|_p\frac{\partial {x'}^k}{\partial x^t}|_p$$

$$=\left(\Gamma^t_{ij}(p)+\frac{\partial^2 x^t}{\partial {x'}^i\partial {x'}^j}|_p\right)\frac{\partial {x'}^k}{\partial x^t}|_p=\left(\Gamma^t_{ij}(p)-\frac{1}{2}\Gamma^t_{ij}(p)-\frac{1}{2}\Gamma^t_{ji}(p)\right)\frac{\partial {x'}^k}{\partial x^t}|_p=0.$$

Now with respect with the new coordinate x', $D{R'}^i_j|_p=d{R'}^i_j|_p$ since $\Gamma'(p)=0$ and by the Cartan's structure equation Eq.(13),

$$d{R'}^i_j|_p=d\left(d{\Gamma'}^i_j+{\Gamma'}^i_m\wedge{\Gamma'}^m_j\right)|_p=dd{\Gamma'}^i_j|_p=0.$$

Thus, $D{R'}^i_j|_p=0$. Since this is true for any point p in M, and R is a tensor, $DR^i_j=0$. Q.E.D.

There are two remerks about the Bianchi's identity. Firstly, the identity holds even for a connection with non-vanishing torsion, but this is not the focus in general relativity, since all Riemannian connection (will be discussed in a later section) are torsion free. For details, please refer to [3]. Secondly, the above Bianchi's identity is actually the second Bianchi's identity, and the Cartan structure equation is one of many structure equations. The other identities could also be found in [3]. These identities are description of the relationship between the matrix-valued connection 1-form, the torsion and curvature 2-forms, and another differential 1-form called the canonical 1-form, which is skipped in this notes.

In physics, these identities are usual expressed in index form. The Jacobi's identity is clearly equivalent to

$$R^i_{jkl} + R^i_{klj} + R^i_{ljk} = 0. \tag{14}$$

To illustrate the Bianchi's identity in index form, if we let $DR^i_j = R^i_{jkl;\sigma} dx^\sigma \otimes dx^k \otimes dx^l$, where D in here is the covariant differential operator (NOT the exterior one). Then the exterior covariant derivative of the curvature 2-form is

$$DR^i_j = 3Alt\left(DR^i_j\right) = 3R^i_{jkl;\sigma} Alt\left(dx^\sigma \otimes dx^k \otimes dx^l\right) = \frac{1}{2} R^i_{jkl;\sigma} dx^\sigma \wedge dx^k \wedge dx^l$$

$$= \frac{1}{2} R^i_{jkl;\sigma}(dx^\sigma \otimes dx^k \otimes dx^l + dx^k \otimes dx^l \otimes dx^\sigma + dx^l \otimes dx^\sigma \otimes dx^k$$

$$-dx^\sigma \otimes dx^l \otimes dx^k - dx^k \otimes dx^\sigma \otimes dx^l - dx^l \otimes dx^k \otimes dx^\sigma)$$

$$= \frac{1}{2}\left(R^i_{jkl;\sigma} + R^i_{jl\sigma;k} + R^i_{j\sigma k;l} - R^i_{jlk;\sigma} - R^i_{j\sigma l;k} - R^i_{jk\sigma;l}\right) dx^\sigma \otimes dx^k \otimes dx^l$$

Since R^i_j is skew-symmetric, i.e. $R^i_{jkl;\sigma} = -R^i_{jlk;\sigma}$, $R^i_{jl\sigma;k} = -R^i_{j\sigma l;k}$ and $R^i_{j\sigma k;l} = -R^i_{jk\sigma;l}$,

$$DR^i_j = \left(R^i_{jkl;\sigma} + R^i_{jl\sigma;k} + R^i_{j\sigma k;l}\right) dx^\sigma \otimes dx^k \otimes dx^l.$$

And the Bianchi's identity implies

$$R^i_{jkl;\sigma} + R^i_{jl\sigma;k} + R^i_{j\sigma k;l} = 0. \tag{15}$$

Back to the curvature operator, for any fixed vector fields Y and Z, the map $X \mapsto R(X,Y)Z$ is an endomorphism on $D^1(M)$, and we could consider the trace of this endomorphism, denoted by $r(Y,Z) = Tr\left(R(*,Y)Z\right)$. [Note that the $C^\infty(M)$-module $End_{C^\infty(M)}\left(D^1(M)\right) = D^1_1(M)$ is (locally) generated by the $C^\infty(M)$-basis $\left\{\frac{\partial}{\partial x^i} \otimes dx^j\right\}_{ij}$, and the trace of any (1,1)-type tensor $T = T^i_j \frac{\partial}{\partial x^i} \otimes dx^j$ is defined to be the summation $Tr(T) = T^i_i$. Clearly, this definition is independent from the choice of coordinate charts, which makes the trace the correpondent $Tr : End_{C^\infty(M)}\left(D^1(M)\right) \to C^\infty(M)$.] This operator $r : D^1(M) \times D^1(M) \to C^\infty(M)$ mapping $(Y,Z) \mapsto r(Y,Z)$ is known as the Ricci curvature. By the properties of the curvature operator theorem (2.2, 2.3), the Ricci curvature operator is $C^\infty(M)$-bilinear, and since the space of $C^\infty(M)$-bilinear functionals $D^1(M) \times D^1(M) \to C^\infty(M)$ is isomorphic to the tensor space $D_2(M)$, we may treat the Ricci curvature operator as a (0,2)-type tensor. The curvature tensor is written as $R = R^i_{jkl} \frac{\partial}{\partial x^i} \otimes dx^j \otimes dx^k \otimes dx^l$, and its trace is $r = R_{jl} = R^m_{jml} dx^j \otimes dx^l$. This tensor is called the Ricci tensor, and can be easily checked to obey the tensor rule of transformation.

3 Riemannian Geometry

In the previous section, we have seen the curvature of a manifold can be studied using a connection, which is an operator form describing parallelism. In this section, we shall see how a connection could be induced by the measurement of length and angle between tangent vectors. In linear spaces, these measurements are concluded by a binary operator called an inner product. Recall an inner product of a real vector space V is a map $\langle *, * \rangle : V \times V \to \mathbb{R}$ mapping $(u, v) \mapsto \langle u, v \rangle$ and satisfies

1. (positive definite) $\langle u, u \rangle \geq 0$, and equality holds if and only if $u = 0$,
2. (symmetric) $\langle u, v \rangle = \langle v, u \rangle$, and
3. (linear in the first argument) $\langle au + bv, w \rangle = a \langle u, w \rangle + b \langle v, w \rangle$.

(Note that the inner product in quantum mechanics is a complex inner product, i.e. it is conjugate symmetric rather than symmetric, and the Dirac notation obeys linearity on the second argument rather than on the first one.) The length of a vector u is defined to be the norm $\|u\| = \sqrt{\langle u, u \rangle}$ and the angle between the non-zero vectors u and v is defined to be $\theta = \cos^{-1}\left(\frac{\langle u,v \rangle}{\|u\|\|v\|}\right)$.

Suppose V is finite dimensional. There exists an orthonormal basis $\{e_i\}_i$, which is a basis of V such that $\langle e_i, e_j \rangle = \delta_{ij}$, and every vector can be expressed as $v = v^i e_i$, where $v^i = \langle v, e_i \rangle$. Hence, $\langle u, v \rangle = \delta_{ij} u^i v^j$. In other words, the inner product can be viewed, under this orthonormal basis, as the identity matrix $1 = (\delta_{ij})_{n \times n}$. If we perform a change of coordinate basis $e \mapsto f$, where $f_j = B^i_j e_i$. Here $B = \left(B^i_j\right)_{n \times n}$ is a non-singulatr matrix (not necessarily unitary). Under the new basis, every vector can be expressed as $v = {v'}^j f_j = {v'}^j B^i_j e_i$, and the inner product becomes $\langle u, v \rangle = {u'}^k {v'}^l B^i_k B^j_l \delta_{ij}$. Hence, the inner product can be viewed, under the basis $\{f_j\}_j$, as the positive definite symmetric matrix $g = B^T B$, where $g_{ij} = B^k_i B^l_j \delta_{kl}$, which implies $\langle u, v \rangle = g_{ij} {u'}^i {v'}^j$.

In special relativity, the concept of *distance* is not positive definite. The space-time metric is the matrix $\eta = diag(1, -1, -1, -1)$, which is not positive definite. Hence, we have to adopt a modified version of inner product called the pseudo-inner product, which is non-degenerate (instead of positive definite), symmetric and linear in the first argument, where non-degenerate means the product satisfies '$u = 0$ if and only if $\langle u, v \rangle = 0$ for any vector v' (in simplier terms, this means the matrix representing the inner product is non-singular). Again, if we perform a change of basis $f_j = B^i_j e_i$, the inner product becomes the invertible symmetric matrix $g = B^T \eta B$, where $g_{ij} = B^k_i B^l_j \eta_{kl}$. The (Lie) group of Lorentz transformation, denoted by $O(1, 3)$, is defined to be set of matrices B that leave η invariant, i.e. $\eta = B^T \eta B$.

At each point p in a smooth manifold M, we have the tangent space $T_p M$, and suppose $x : U \to \mathbb{R}^n$ is a chart around p, $\left\{\frac{\partial}{\partial x^j}|_p\right\}_j$ is a basis of $T_p M$. The essence of a (pseudo-) metric is to define a (pseudo-) inner product at each tangent space T_pM smoothly over all points p in the manifold. A pseudo-inner product on T_pM, denoted by $g_p : (X_p, Y_p) \mapsto g_p(X_p, Y_p)$, can be written in form of a $n \times n$-invertible symmetric matrix $g_p = \left(g_{p_{ij}}\right)_{n\times n}$, where $g_{p_{ij}} = g_p\left(\frac{\partial}{\partial x^i}|_p, \frac{\partial}{\partial x^j}|_p\right)$. In other words, $g_p\left(X_p, Y_p\right) = g_{p_{ij}} X_p^i Y_p^j$. Unlike the flat space-time, where the metric η is uniform, the matrix g_p may varies from point to point throughout the manifold M, i.e. we have a set of functions $g_{ij} : p \mapsto g_{ij}(p) = g_{p_{ij}}$. However, these functions are defined only locally on U by the coordinate chart x. Suppose we perform a change of coordinate $x \mapsto x'$. Then,

$$
\begin{aligned}
{g'}_{ij} &= g\left(\frac{\partial}{\partial x'^i}, \frac{\partial}{\partial x'^j}\right) = g\left(\frac{\partial x^k}{\partial x'^i}\frac{\partial}{\partial x^k}, \frac{\partial x^l}{\partial x'^j}\frac{\partial}{\partial x^l}\right) \\
&= \frac{\partial x^k}{\partial x'^i}\frac{\partial x^l}{\partial x'^j} g\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right) = \frac{\partial x^k}{\partial x'^i}\frac{\partial x^l}{\partial x'^j} g_{kl},
\end{aligned}
$$

i.e. g_{ij} satisfies the tensor rule of transformation. In other words, we could defined a global (0,2)-type non-degenerate symmetric tensor g in $D_2(M)$ that is expressed locally as $g_{ij} dx^i \otimes dx^j$ and at each point p in M, $g_p = \left(g_{ij}(p)\right)_{n\times n}$ serves as a pseudo-inner product of the tangent space T_pM. Since the $C^\infty(M)$-module $D_2(M)$ is isomrophic to the module of functionals $D^1(M) \times D^1(M) \to C^\infty(M)$, we may treat g as a non-degenerate symmetric bilinear operator $g : D^1(M) \times D^1(M) \to C^\infty(M)$ mapping $(X, Y) \mapsto g(X, Y) = g\rfloor\left(X \otimes Y\right)$ or in local coordinate, $g(X, Y) = g_{ij} X^i Y^j$. The definition of a metric tensor can be summarized by the following.

Definition of a pseudo-Riemannian metric: A pseudo-Riemannian metric is a non-degenerate symmetric (0,2)-type tensor in $D_2(M)$, i.e. could be written locally as $g = g_{ij} dx^i \otimes dx^j$, where g_{ij} are smooth real-valued functions, the matrix $\left(g_{ij}\right)_{n\times n}$ is non-singular and $g_{ij} = g_{ji}$. Or equivalently, it could be treated as a map $g : D^1(M) \times D^1(M) \to C^\infty(M)$ that satisfies

1. (non-degenerate) a vector field X is zero if and only if for any vector field Y, $g(X, Y) = 0$,
2. (symmetric) $g(X, Y) = g(Y, X)$ for any vector field X and Y, and
3. (linear) $g(\alpha X + \beta Y, Z) = \alpha g(X, Z) + \beta g(Y, Z)$ for any smooth functions α and β on M, and any vector fields X, Y and Z.

A smooth manifold is called a (pseudo-) Riemann manifold if it is equipped with a (pseudo-) Riemannian metric.

Denote $Sym(n)$ be the space of $n \times n$-symmetric matrices. For any matrix S in $Sym(n)$, since S is diagonalizable, let $P(S)$ be the number of positive diagonal entries, $N(S)$ be the number of negative diagonal entries. The rank of S is then $P(S) + N(S)$. The signature of S is defined to be $sig(S) = P(S) - N(S)$. Two matrices S and S' are said to be congruent if there is a non-singular matrix B such that $S' = B^T S B$. Clearly, the ranks of two congruent matrices are the same. The Sylvester's theorem [7] guarantees the signatures of congruent matrices must also be the same. The converse is also true, i.e. matrices of the same rank and signature are congruent, since they are congruent to the diagonal matrix $diag(1, \ldots, 1, 0 \ldots, 0, -1, \ldots, -1)$ of the same rank and signature. In general relativity, we would expect any pseudo-Riemannian metric g to have a constant signature throughout the whole space, i.e. $sig(g(p)) = sig(\eta) = -2$ (of course g has full rank 4 as it is non-singular) for each point p in the space-time manifold. Thus, the space-time manifold is a pseudo-Riemann manifold of signature -2.

3.1 Riemannian Connection

Definition of the Riemannian Connection: A connection of a (pseudo-) Riemann manifold equipped with the (pseudo-) metric g is said to be Riemannian if it is torsion free and compatible with the metric, i.e. $Dg = 0$, where D is the covariant differential operator.

The Riemannian connection is also known as the Levi-Civita connection. A (pseudo-) Riemann manifold must exist a unique Riemann connection, where the Christoffel symbols are

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l} \right), \tag{16}$$

where $g^{-1} = \left(g^{ij}\right)_{n \times n}$ is the inverse of the metric matrix g. The proof can be found in the course lecture notes or other standard textbooks on differential geometry such as [3, 1, 2]. Note that the inverse of g is upper indexed, and it is easy to check the corresponding contravariant tensor rule of transformation since

$$\left(\frac{\partial x'^i}{\partial x^j} \right)^{-1}_{n \times n} = \left(\frac{\partial x^i}{\partial x'^j} \right)_{n \times n}.$$

Hence, g^{-1} is a global (2,0)-type tensor in $D^2(M)$. Written in index form, $g^{-1} = g^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$.

Recall the curvature tensor, written in $R = R^i_{jkl} \frac{\partial}{\partial x^i} \otimes dx^j \otimes dx^k \otimes dx^l$, can be contracted by a trace to give the Ricci tensor $r = R_{jl} dx^j \otimes dx^l$, where $R_{jl} = R^m_{jml}$. By the metric g^{-1}, we may define a smooth function on M by $s = r \rfloor g^{-1} = R_{ij} g^{ij}$ known as the scalar curvature. In general relativity, we define the Einstein tensor be the (0,2)-type tensor $G = r - \frac{1}{2} sg$, i.e. $G_{ij} = R_{ij} - sg_{ij}$.

Next, let us study geodesics on a Riemann manifold M. Suppose $\alpha : (-\epsilon, \epsilon) \to M$ is a smooth curve mapping $\tau \mapsto \alpha(\tau)$. A reparametrization is an invertible change of parameter $\varphi : (-\epsilon, \epsilon) \to (-\varepsilon, \varepsilon)$ mapping $\tau \mapsto \varphi(\tau) = t$. Under the new parameter, the reparametrized curve is $\tilde{\alpha} : (-\varepsilon, \varepsilon) \to M$ mapping $t \mapsto \tilde{\alpha}(t) = \alpha(\tau)$, where $\tau = \varphi^{-1}(t)$. (We always assume the curve starts at the same position, i.e. $\varphi(0) = 0$.) By an abuse of notation, we usually write $\tilde{\alpha}$ simply as α. We could always think the parameters τ or t as the time when a particle travels along the curve. So the geometric meaning of a reparametrization is to alter the speed of the propagation without changing the trace of the curve.

By using the metric g of the Riemann manifold, we could define the arc length of the curve from 0 to τ_0 to be

$$s(\tau_0) = \int_0^{\tau_0} \sqrt{|g(\dot{\alpha}(\tau), \dot{\alpha}(\tau))|} d\tau.$$

Hence, there is a correspondence $s : \tau \mapsto s(\tau)$. If the curve has non-vanishing $g(\dot{\alpha}, \dot{\alpha})$, then the map $s : \tau \mapsto s(\tau)$ is invertible, and thus this defines a reparametrization of the curve known as the parametrization by arc length. By a change of variable in the integral, the readers may check very easily the parametrization of arc length will still be the same if we use the parameter t instead of τ.

Recall the curve α is a geodesic if it satisfies the geodesic equation $\nabla_{\dot{\alpha}} \dot{\alpha} = 0$, or in index form

$$\frac{d^2 x^k}{d\tau^2} + \Gamma^k_{ij} \frac{dx^i}{d\tau} \frac{dx^j}{d\tau} = 0,$$

where x^k is the abreviation for $x^k \circ \alpha$, and $x = (x^1, \ldots, x^n)$ is a coordinate chart around $\alpha(0)$. The question is will the geodesic equation be invariant under a reparametrization of the curve. The answer, in general, is no. Indeed, we could prove that on a Riemann manifold, if the curve $\alpha : \tau \mapsto \alpha(\tau)$ is a geodesic, then the parameter τ must be proportional to the parametrization of arc length, i.e. $s = v\tau$, where v is a constant. In other words, all geodesics propagate in constant speed. To prove this, we have to make use of the fact that the Riemann connection is compatible with the metric, i.e. $Dg = 0$.

$$\frac{d}{d\tau} g(\dot{\alpha}, \dot{\alpha}) = \nabla_{\dot{\alpha}} \left(g(\dot{\alpha}, \dot{\alpha})\right) = \left(\nabla_{\dot{\alpha}} g\right)(\dot{\alpha}, \dot{\alpha}) + g\left(\nabla_{\dot{\alpha}} \dot{\alpha}, \dot{\alpha}\right) + g\left(\dot{\alpha}, \nabla_{\dot{\alpha}} \dot{\alpha}\right) = 0.$$

Denote the constant $v = \sqrt{|g(\dot{\alpha}, \dot{\alpha})|}$. Hence,

$$s = \int_0^{\tau} \sqrt{|g(\dot{\alpha}, \dot{\alpha})|} = v\tau.$$

This implies the geodesic equation can be rewritten under the parametrization of arc length as

$$\frac{d^2 x^k}{ds^2} + \Gamma^k_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0. \tag{17}$$

In special relativity, the space-time metric is $g = dx^0 \otimes dx^0 - dx^1 \otimes dx^1 - dx^2 \otimes dx^2 - dx^3 \otimes dx^3$. If t is time, i.e. $x^0 = ct$, then the space-time interval (or the arc length parametrization) of a time-like trajectory is

$$s = \int_0^t \sqrt{c^2 - \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2 - \left(\frac{dz}{dt}\right)^2} dt.$$

In general relativity, for a massive object traveling around a gravitational field, there is no guarantee that $g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}$ would be constant, and hence the interval $s = \int_0^t \sqrt{g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}}dt$ may not be proportional to time t. This means the equation $\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0$ is NOT a geodesic equation.

3.2 The Sphere $\mathbb{S}^2$

Recall $\mathbb{S}^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ is the unit sphere in $\mathbb{R}^3$. Consider the chart $(x, y, z) \mapsto (\theta, \phi)$ such that $x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$, $z = \cos\theta$. Then, a tangent vector is in the form $X = X^\theta\frac{\partial}{\partial\theta} + X^\phi\frac{\partial}{\partial\phi}$. Transform back to the container space $\mathbb{R}^3$, $X = X^x\hat{\imath} + X^y\hat{\jmath} + X^z\hat{k}$, where

$$X^x = X^\theta\frac{\partial x}{\partial\theta} + X^\phi\frac{\partial x}{\partial\phi} = X^\theta\cos\theta\cos\phi - X^\phi\sin\theta\sin\phi$$

$$X^y = X^\theta\frac{\partial y}{\partial\theta} + X^\phi\frac{\partial y}{\partial\phi} = X^\theta\cos\theta\sin\phi + X^\phi\sin\theta\cos\phi$$

$$X^z = X^\theta\frac{\partial z}{\partial\theta} + X^\phi\frac{\partial z}{\partial\phi} = -X^\theta\sin\theta$$

The Riemannian metric g is the one induced by the usual inner product of $\mathbb{R}^3$

$$g(X, Y) = X^xY^x + X^yY^y + X^zY^z = X^\theta Y^\theta + X^\phi Y^\phi\sin^2\theta.$$

If we take $(x^1, x^2) = (\theta, \phi)$, then the metric is

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\sin^2\theta \end{pmatrix},$$

and the Christoffel symbols of the Riemannian connection are

$$\Gamma^1_{11} = \Gamma^2_{22} = \Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{11} = 0, \quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{\cos\theta}{\sin\theta}, \quad \Gamma^1_{22} = -\sin\theta\cos\theta.$$

Put in differential form, the connection 1-form is

$$\Gamma = \begin{pmatrix} 0 & -\sin\theta\cos\theta d\phi \\ \frac{\cos\theta}{\sin\theta}d\phi & \frac{\cos\theta}{\sin\theta}d\theta \end{pmatrix},$$

and the curvature 2-form is

$$R = d\Gamma + \Gamma\wedge\Gamma = \begin{pmatrix} 0 & \sin^2\theta \\ -1 & 0 \end{pmatrix} d\theta\wedge d\phi,$$

which means the only non-zero components of the curvature tensor in index form are

$$R^1_{212} = -R^1_{221} = \sin^2\theta, \quad R^2_{112} = -R^2_{121} = -1.$$

The Ricci tensor is $r = d\theta\otimes d\theta + \sin^2\theta d\phi\otimes d\phi$, and the scalar curvature is $s = 2$, which means the sphere has constant positive curvature.

Suppose α is a geodesic on the unit sphere. The geodesic equations Eq.(17) are

$$\frac{d^2\theta}{ds^2} - \sin\theta\cos\theta\frac{d\phi}{ds}\frac{d\phi}{ds} = 0, \quad \frac{d^2\phi}{ds^2} + \frac{2\cos\theta}{\sin\theta}\frac{d\theta}{ds}\frac{d\phi}{ds} = 0.$$

Hence, longitudes ϕ = constant, $\theta = \omega s$ and the equator $\theta = \pi/2$, $\phi = \omega t$ are geodesics. By the symmetry of the sphere, all great circles are geodesics, and by the fundamental theorem of differential equations, the solution of the geodesic equations are unique if the initial position and velocity are specified, which implies all geodesics must be arcs of great circles.

3.3 The Poincare's half plane H

The Poincare's half plane is defined to be the upper half plane $H = \{(x,y) : y > 0\}$ in $\mathbb{R}^2$ equipped with the metric $g = y^{-2}(dx\otimes dx + dy\otimes dy)$. Write $x^1 = x$ and $x^2 = y$. In matrix form,

$$g = \begin{pmatrix} 1/y^2 & 0 \\ 0 & 1/y^2 \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix}.$$

The Riemann connection has Christoffel symbols $\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^2_{12} = \Gamma^2_{21} = 0$, $\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -1/y$, and $\Gamma^2_{11} = 1/y$, i.e. the connection 1-form is the matrix

$$\Gamma = \begin{pmatrix} -y^{-1}dy & -y^{-1}dx \\ y^{-1}dx & -y^{-1}dy \end{pmatrix}.$$

The curvature 2-form is the matrix

$$R = d\Gamma + \Gamma\wedge\Gamma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\frac{1}{y^2}dx\wedge dy.$$

The Ricci tensor is $r = -y^{-2}(dx\otimes dx + dy\otimes dy)$, and the scalar curvature is $s = -2$, which means the Poincare's half plane has constant negative curvature.

The geodesic equations Eq.(17) are

$$\frac{d^2x}{ds^2} - \frac{2}{y}\frac{dx}{ds}\frac{dy}{ds} = 0, \quad \frac{d^2y}{ds^2} + \frac{1}{y}\frac{dx}{ds}\frac{dx}{ds} - \frac{1}{y}\frac{dy}{ds}\frac{dy}{ds} = 0.$$

Since the speed of a geodesic is constant, let it be v, we have

$$v^2 = g\left(\frac{dx}{ds}, \frac{dy}{ds}\right) = \frac{1}{y^2}\left[\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2\right].$$

Denote $\dot{x} = \frac{dx}{ds}$ and $\dot{y} = \frac{dy}{ds}$.

Firstly, it is easy to see vertical straightlines are geodesics. Suppose $\dot{x} = 0$, then $\dot{y} = \pm vy$. Hence, the vertical line x = constant, $y = Ae^{\pm vs}$ is a geodesic, where A is positive. Secondly, we could show all other geodesics are arcs of half circles centered at points on the x-axis. Suppose $\dot{x} \neq 0$.

$$\frac{d}{ds}\left(\frac{x\dot{x}+y\dot{y}}{\dot{x}}\right) = \frac{\dot{x}(x\ddot{x}+\dot{x}^2+\dot{y}^2+y\ddot{y}) - \ddot{x}(x\dot{x}+y\dot{y})}{\dot{x}^2}$$

$$= \frac{\dot{x}^3+\dot{x}\dot{y}^2+\dot{x}y\ddot{y}-\ddot{x}y\dot{y}}{\dot{x}^2} = \frac{\dot{x}^3+\dot{x}\dot{y}^2+\dot{x}y\frac{\dot{y}^2-\dot{x}^2}{y}-y\dot{y}\frac{2\dot{x}\dot{y}}{y}}{\dot{x}^2} = 0,$$

where the geodesic equations are used in the second last equality. This means there is some constant x_0 such that $x\dot{x}+y\dot{y} = x_0\dot{x}$. Hence,

$$\frac{d}{ds}\left[(x-x_0)^2+y^2\right] = 2\dot{x}(x-x_0)+2y\dot{y} = 0,$$

which implies there is some positive constant r such that $(x-x_0)^2+y^2 = r^2$, i.e. it is a half circle with radius r centered at $(x_0, 0)$.

4 Appendix

Suppose M is a smooth manifold, p is a point in M, and $x = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ is a coordinate chart on a neighbourhood around p. For any differentiable function $f : U \to \mathbb{R}$, $f \circ x^{-1} : x(U) \to \mathbb{R}$ is a differentiable function defined on a flat n-dimensional domain $x(U)$. Thus, partial derivatives $\partial_j(f \circ x^{-1})|_{x(p)}$ is well-defined, for $j = 1, \ldots, n$. Using this, the partial derivative of f at p with respect to x^j is defined to be

$$\frac{\partial f}{\partial x^j}|_p = \partial_j(f \circ x^{-1})|_{x(p)}.$$

And the differential operator is defined to be the functional $\frac{\partial}{\partial x^j}|_p : C^\infty(p) \to \mathbb{R}$ mapping $f \mapsto \frac{\partial f}{\partial x^j}|_p$, where $C^\infty(p)$ is the space of (to be precise, the germs of) smooth functions around p.

Suppose there is another chart $y = (y^1, \ldots, y^n) : U \to \mathbb{R}^n$. The chain rule can be rewritten in terms of the differential operators according to the new definition as the following. For $j = 1, \ldots, n$

$$\frac{\partial f}{\partial y^j}|_p = \partial_j(f \circ y^{-1})|_{y(p)} = \partial_j(f \circ x^{-1} \circ x \circ y^{-1})|_{y(p)}$$

$$= \partial_i(f \circ x^{-1})|_{x(p)}\partial_j(x^i \circ y^{-1})|_{y(p)} = \frac{\partial f}{\partial x^i}|_p \frac{\partial x^i}{\partial y^j}|_p,$$

where the third identity is the original chain rule, and the first and last identity are by the new definitions of the differential operators. And hence,

$$\frac{\partial}{\partial y^j} = \frac{\partial x^i}{\partial y^j}\frac{\partial}{\partial x^i}.$$

References

[1] M. P. do Carmo, *Riemannian geometry* (1992), Birkhauser, Boston.

[2] S. Helgason, *Differential geometry and symmetric spaces* (1962), Academic Press, New York.

[3] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* Vol I (1991), John Wiley and Sons.

[4] M. A. Armstrong, *Basic topology* (1983), Springer, New York.

[5] S. Lang, *Algebra* 2nd ed. (1984), Addison-Wesley.

[6] M. Spivak, *Calculus on Manifold* (1965), Benjamin, New York.

[7] S. Lipschutz and M. L. Lipson, *Schaum's outline of Linear Algebra* 3rd ed. (2001), McGraw-Hill.