MA3429–Differential Geometry

Peter Taylor BSc MASt PhD

June 9, 2014

Contents

Chapter 1

Differentiable Manifolds

Informally, an n-dimensional manifold, which we will denote throughout by $\mathcal{M}$, is a set of points that locally looks like a subset of $\mathbb{R}^n$, the usual n-dimensional space of vector algebra (not to be confused with the Euclidean space which we denote by $\mathbb{E}^n$).

1.1 Topological Spaces

We cannot begin to discuss differentiability in a space of interest, without first considering continuity, which is made mathematically precise in the study of its (local) topology. The fundamental concept of interest in topology is that of a *neighbourhood*.

Definition 1.1.1. Let $\mathcal{M}$ be a set and $p \in \mathcal{M}$ (the elements of the set are usually called *points* but can be any mathematical object, e.g., a function). A neighbourhood of p is a subset U containing an open set which includes p, or equivalently, p is an element of the interior of $U \subset \mathcal{M}$.

In $\mathbb{R}^n$, there is a 'natural' topology induced by the Euclidean distance function

$$d(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + \cdots + (x^n - y^n)^2} \tag{1.1}$$

and a neighbourhood of a point p of radius r is an open solid sphere centered at p defined by the locus of points satisfying

$$U(p) = \{x : d(p, x) < r, \text{where } r > 0\}. \tag{1.2}$$

The space $\mathbb{R}^n$ endowed with this distance function (1.1) defines a metric space, the familiar Euclidean space $\mathbb{E}^n$.

Topology, however, is a more a primitive notion than distance and the topology of $\mathbb{R}^n$ does not depend on the above definition of distance. Indeed, the definition of a neighbourhood above contains no mention of distance. For example, if we have another distance function, $\bar{d}(x, y)$ say, then if we have a neighbourhood at p with respect to $d(x, y)$, then there exists a neighbourhood at p with respect to $\bar{d}(x, y)$ entirely contained within the d-type neighbourhood, and vice-versa.

We may now define a *topological space* as follows:

Definition 1.1.2. A topological space is a set of points $\mathcal{M}$ along with an assignment to each $p \in \mathcal{M}$ of neighbourhoods, $N(p)$ say, satisfying the following four axioms:

(i) p belongs to any neighbourhood of p.

(ii) If U is a neighbourhood of p and $U \subset V$, then V is a neighbourhood of p.

(iii) If U, V are neighbourhoods of p, then $U \cap V$ is a neighbourhood of p.

(iv) If U is a neighbourhood of p, then there is a neighbourhood V of p such that $V \supset U$ and V is a neighbourhood of each point of U.

A collection of neighbourhoods satisfying these four axioms define a neighbourhood topology. In short then, a topological space is a set $\mathcal{M}$ along with a neighbourhood topology N. There are several other equivalent definitions of a topological space, for example, in terms of open sets, but we find it most intuitive to discuss in terms of neighbourhoods.

1.2 Charts

Let $\mathcal{M}$ be a topological space, let $p \in \mathcal{M}$ and U an open neighbourhood of p. A chart on U is a one-to-one map

$$\phi : U \to \phi(U) \subset \mathbb{R}^n. \tag{1.3}$$

If these maps are bicontinuous (i.e., continuous bijective maps with a continuous inverse), then we say that the map is a **homeomorphism** and U is **homeomorphic** to $\phi(U)$. From a topological viewpoint, two topological spaces that are homeomorphic are equivalent. If we have a chart ϕ defined on U which is homeomorphic to $\phi(U) \subset \mathbb{R}^n$, then there is a neighbourhood of the point p that is topologically equivalent to a neighbourhood of $\mathbb{R}^n$ containing $\phi(p)$.

The n-tuple $\phi(p) \in \mathbb{R}^n$ constitute a local coordinate system defined in the open neighbourhood U, and we usually write

$$\phi(p) = \{x^\mu(p)\} = (x^1(p), x^2(p), ..., x^n(p)). \tag{1.4}$$

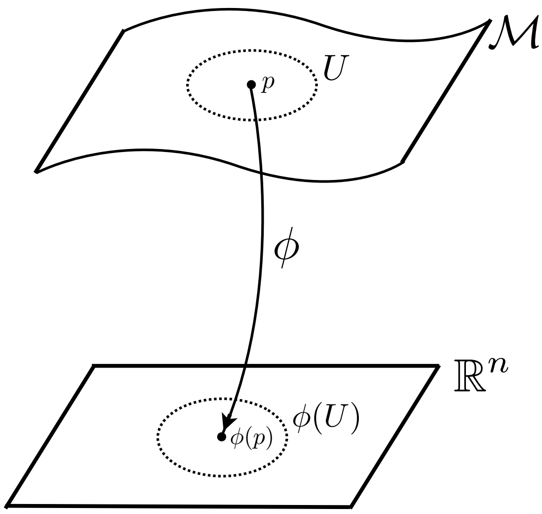

Figure 1.1:

Suppose now that we have two such charts on $U \subset \mathcal{M}$, say ϕ_1 and ϕ_2. Since these are injective maps, they both possess inverses, for instance,

$$\phi_1^{-1} : \phi_1(U) \subset \mathbb{R}^n \to U. \tag{1.5}$$

We may therefore define a map

$$\phi_2 \circ \phi_1^{-1} : \mathbb{R}^n \to \mathbb{R}^n,\ \phi_1(U) \to \phi_2(U). \tag{1.6}$$

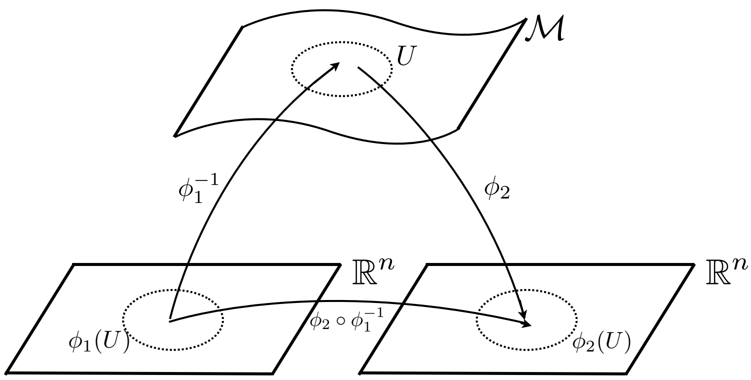

Figure 1.2:

The **meshing condition** is a requirement that each chart be C^1-related (i.e. smooth) to every other chart that it overlaps with. If ϕ_1 and ϕ_2 are C^1-related, we call them meshing charts.

The map $\phi_2 \circ \phi_1^{-1}$ takes us from the chart ϕ_1 to the chart ϕ_2. For a point $p \in U$, the map $\phi_2 \circ \phi_1^{-1}$ defines a coordinate transformation from the coordinates $\phi_1(p) = (x^1(p), ..., x^n(p))$ to the coordinates $\phi_2(p) = (X^1(p), ..., X^n(p))$.

The smoothness and overlap region of the map $\phi_2 \circ \phi_1^{-1}$ may be checked by computing the Jacobian matrix $D(\phi_2 \circ \phi_1^{-1})$ and its determinant J (also called the Jacobian),

$$J \equiv \det(D(\phi_2 \circ \phi_1^{-1})) = \begin{vmatrix} \dfrac{\partial X^1}{\partial x^1} & \dfrac{\partial X^1}{\partial x^2} & \cdots & \dfrac{\partial X^1}{\partial x^n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial X^n}{\partial x^1} & \dfrac{\partial X^n}{\partial x^2} & \cdots & \dfrac{\partial X^n}{\partial x^n} \end{vmatrix} \tag{1.7}$$

Example 1.2.1. Let $\mathcal{M} = \mathbb{R}^2$, and let ϕ_1 map the point p to standard Cartesian coordinates (x, y) and ϕ_2 map p to another set of Cartesian coordinates (X, Y) obtained from the first set by a rotation through the constant angle α. We have,

$$\phi_1(p) = (x, y), \qquad \phi_2(p) = (X, Y) \tag{1.8}$$

and it is easy to see that $\phi_2 \circ \phi_1^{-1}$ maps

$$(x, y) \to (X = x\cos\alpha + y\sin\alpha, Y = -x\sin\alpha + y\cos\alpha). \tag{1.9}$$

The Jacobian is therefore

$$J = \det(D(\phi_2 \circ \phi_1^{-1})) = \begin{vmatrix} \dfrac{\partial X}{\partial x} & \dfrac{\partial X}{\partial y} \\ \dfrac{\partial Y}{\partial x} & \dfrac{\partial Y}{\partial y} \end{vmatrix} = \begin{vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{vmatrix} = 1. \tag{1.10}$$

Since the Jacobian is non-zero and non-singular everywhere, ϕ_1 and ϕ_2 are meshing charts on the entire manifold. In fact, the charts are C^∞-related on the entire manifold.

We could introduce an alternative chart ϕ_3 which maps p to the standard polar coordinates (r, θ). Then $\phi_3 \circ \phi_1^{-1}$ maps

$$(x, y) \to (r = \sqrt{x^2 + y^2}, \theta = \tan^{-1}(y/x)). \tag{1.11}$$

One can easily check that the Jacobian for this coordinate transformation is

$$J = 1/r. \tag{1.12}$$

Therefore, ϕ_1, ϕ_3 are meshing except at $r = 0$. Hence to cover all of $\mathbb{R}^2$ in polar coordinates, we would need at least two sets of polar coordinates with different origins. ∎

Example 1.2.2. Consider the two-sphere $\mathbb{S}^2$ embedded in 3-dimensional Euclidean space which is the locus of points satisfying $x^2 + y^2 + z^2 = 1$. Let us construct a chart from an open set U_1 defined by the entire sphere minus the north pole

$$U_1 = \mathbb{S}^2 - \{0, 0, 1\}$$

via the so-called "stereographic projection". Each point in U_1 is surjectively mapped to a point in the plane $z = -1$ by drawing a straight line through the north pole intersecting $\mathbb{S}^2$ and the plane and assigning to the point on $\mathbb{S}^2$ the coordinates (w^1, w^2) of the point in the plane. It is

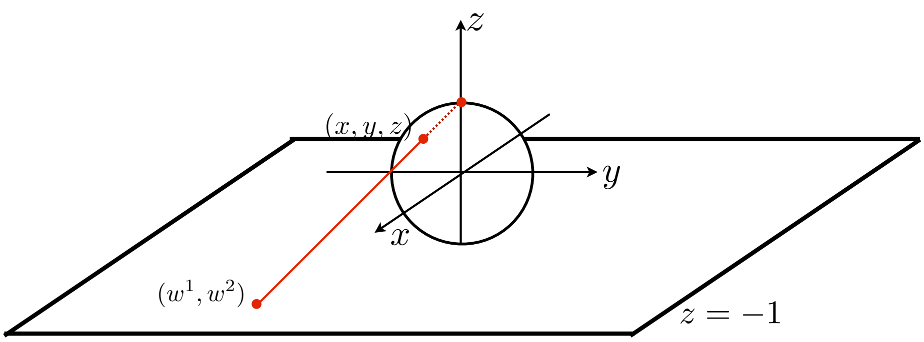

Figure 1.3:

straight-forward to show that the map is explicitly given by

$$\phi_1(x, y, z) \equiv (w^1, w^2) = \left(\frac{2x}{1-z}, \frac{2y}{1-z}\right).$$

The inverse map is also clearly continuous and hence the two-sphere with any point removed is homeomorphic to $\mathbb{R}^n$. We could have also considered the chart (U_2, ϕ_2) which projects the two-sphere minus the south-pole

$$U_2 = \mathbb{S}^2 - \{0, 0, -1\}$$

onto the plane $z = 1$ where the projection is given by

$$\phi_2(x, y, z) \equiv (u^1, u^2) = \left(\frac{2x}{1+z}, \frac{2y}{1+z}\right).$$

Combined these two charts cover the entire two-sphere and overlap in the region excluding the two poles. The coordinate transformation is given by

$$u^i = \frac{4w^i}{((w^1)^2 + (w^2)^2)}$$

and hence the Jacobian is

$$J = -\frac{16}{((w^1)^2 + (w^2)^2)^2},$$

which implies that the charts are meshing in the overlap region. ∎

1.3 Differentiable Manifolds

A set of charts such that every point $p \in \mathcal{M}$ lies in the domain of at least one chart is known as an **atlas** for $\mathcal{M}$. The union of all such atlases is the maximal atlas.

Definition 1.3.1. An n-dimensional manifold is an n-dimensional topological space $\mathcal{M}$ along with a maximal atlas.

If each chart in the atlas are C^1-related, then we say the manifold is differentiable. Similarly, we say the manifold is a C^k manifold if each chart in the atlas is C^k-related to every other chart it overlaps with.

This definition naturally leads to the intuitive notion of a manifold as a space that locally looks like $\mathbb{R}^n$ since for every point p, there is a neighbourhood containing p on which a chart is defined, which implies that this patch on which the chart is defined is topologically equivalent to a subset of $\mathbb{R}^n$.

The definition is not restrictive to intuitive geometrical ideas of a manifold, however, and includes any abstract space whose "points" can be continuously parametrized by n-tuples we call coordinates, e.g., the set of all rotations of a rigid object in three dimensions, which is a manifold since it is continuously parametrized by the three so-called 'Euler' angles.

Chapter 2

Tangent Vectors and Tangent Spaces

In our familiar treatment of vectors, they represent "directed magnitudes" stretching from one point to another, which may even be trivially transported from point to point. Such concepts, however, are no longer useful in curved manifolds.

In a two dimensional manifold $\mathcal{M}$, a tangent space at a point p, which we denote by $T_p(\mathcal{M})$, is easy to visualize by embedding the manifold in a three-dimensional Euclidean space. Then the tangent space at p is just the two-dimensional plane tangent to the point p in the usual way (see figure). Tangent vectors at p can now be visualized as "directed magnitudes" living in this plane.

This picture of course becomes impossible in higher dimensions. Moreover, it relies on a higher-dimensional embedding, i.e., one has to embed the manifold in a Euclidean space of higher dimensions in order to be able to visualize/define the tangent vectors. We would prefer if we could define the tangent vectors and the tangent space at p in terms of intrinsic properties of the manifold, and which do not require any embedding in a higher-dimensional space. We must first build up some notions of smooth functions and curves on a manifold.

2.1 Smooth Functions

Let $\mathcal{M}$ be a manifold and f be a real function with

$$f : \mathcal{M} \to \mathbb{R}. \tag{2.1}$$

How do we define the "smoothness" of f since this requires a notion of differentiability. We use the fact that $\mathcal{M}$ is locally like $\mathbb{R}^n$ and we understand calculus in $\mathbb{R}^n$.

As before, we introduce a chart, ϕ say.

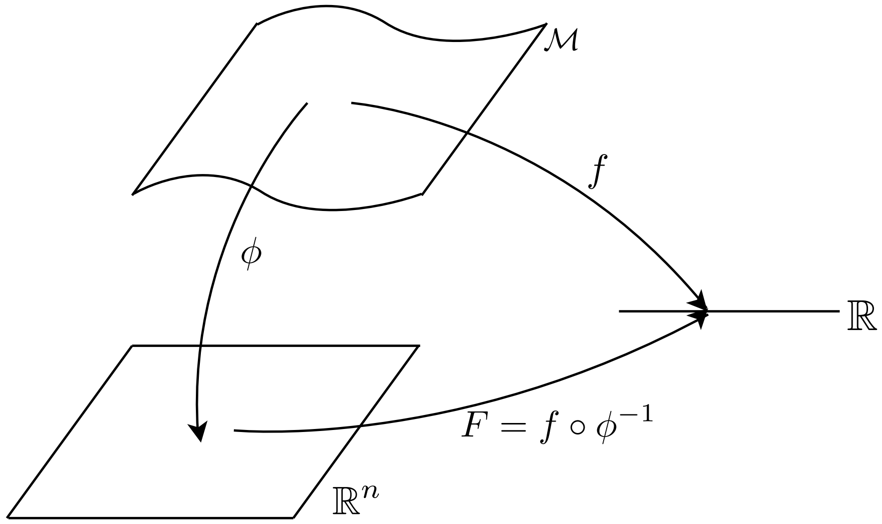

Figure 2.1:

We define a function F to be the coordinate image of f under the chart ϕ:

$$F : \mathbb{R}^n \to \mathbb{R}, \qquad F = f \circ \phi^{-1}. \tag{2.2}$$

We say that f is smooth if and only if F is smooth in the familiar sense. (We can comfortably talk about the smoothness of F since F is a function on $\mathbb{R}^n$.) However, this definition of smoothness appears, at least naively, to be chart dependent. However, the following theorem proves that smoothness as defined above is a chart independent concept.

Theorem 2.1.1. The smoothness of f is chart independent.

Proof. Let ϕ_1 and ϕ_2 be two meshing charts and $F_i = f \circ \phi_i^{-1}$ for $i = 1, 2$. Then

$$\begin{aligned} F_1 &= f \circ \phi_1^{-1} \\ &= f \circ \phi_2^{-1} \circ \phi_2 \circ \phi_1^{-1} \\ &= F_2 \circ \phi_2 \circ \phi_1^{-1}. \end{aligned} \tag{2.3}$$

Since $\phi_2 \circ \phi_1^{-1}$ is smooth, F_1 and F_2 have the same smoothness properties. Therefore, the smoothness of f is chart independent. ∎

This definition of a smooth function may be generalized to a function mapping a manifold $\mathcal{M}$ to another manifold $\mathcal{N}$,

$$f : \mathcal{M} \to \mathcal{N}. \tag{2.4}$$

Let ϕ_1 be a chart on a manifold $\mathcal{M}$ of dimension n_1, ϕ_2 a chart on a manifold $\mathcal{N}$ of dimension n_2. We define

$$F : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}, \qquad F = \phi_2 \circ f \circ \phi_1^{-1}. \tag{2.5}$$

As before, we say f is smooth if and only if F is smooth. Analogous to the proof above, it can be shown that this definition is chart independent.

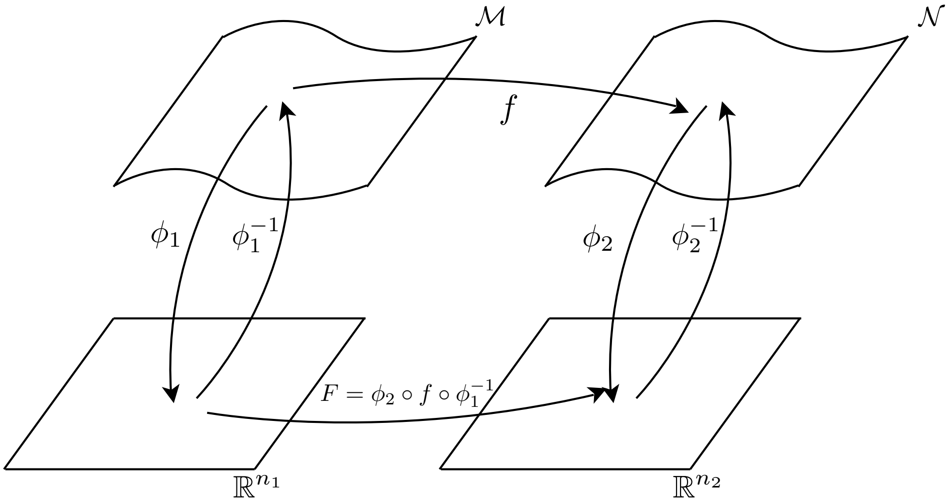

Figure 2.2:

2.2 Smooth Curves

Let $I = (a, b)$ be an open interval of $\mathbb{R}$. Then we define a curve in $\mathcal{M}$ as a map

$$\gamma : \mathbb{R} \to \mathcal{M}, \qquad \gamma(s) = p. \tag{2.6}$$

The curve γ is smooth if its coordinate image

$$\phi \circ \gamma : I \to \mathbb{R}^n, \qquad \phi \circ \gamma(s) = (x^1(s), x^2(s), ..., x^n(s)) \tag{2.7}$$

is smooth in the usual sense. It is straight-forward to verify that this definition is chart independent.

2.3 Coordinate Induced Basis for the Tangent Space

We wish to construct a basis for the tangent space at $p \in \mathcal{M}$, $T_p(\mathcal{M})$, using only intrinsic properties of the manifold. We begin by combining the concepts of smooth functions and

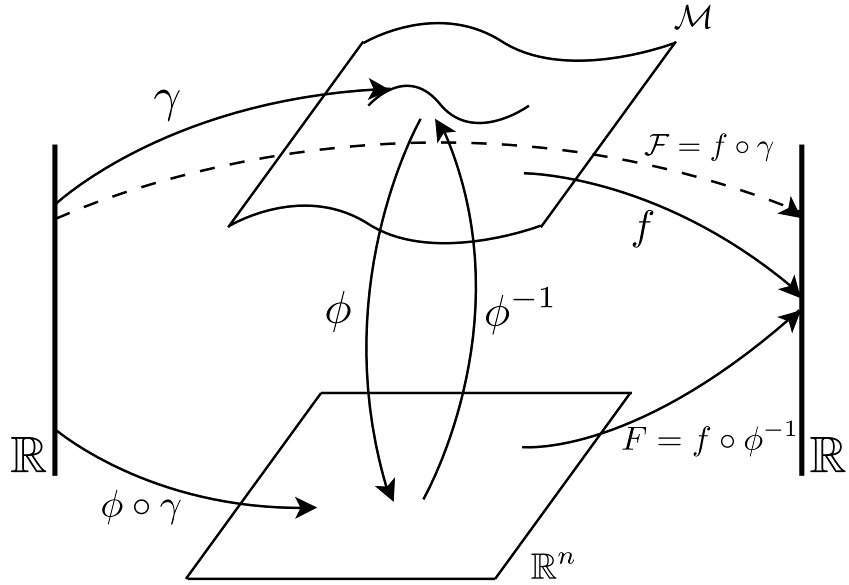

Figure 2.3:

smooth curves to define the function

$$\mathcal{F} : I \to \mathbb{R}, \qquad s \to \mathcal{F}(s) = f \circ \gamma(s) = f(\gamma(s)), \tag{2.8}$$

i.e., $\mathcal{F}$ evaluates the function f at points along the curve γ. Therefore, the rate at which $\mathcal{F}$ changes, $d\mathcal{F}/ds$ gives the rate of change of f following the curve $\gamma(s)$.

Hence, for a particular smooth curve γ at p (where without loss of generality we have parametrized such that $s = 0$ at p), there is a map from the set of real functions to $\mathbb{R}$ defined by

$$\dot{\gamma}_p : f \to \dot{\gamma}_p f \equiv \dot{\gamma}_p(f) = \left[\frac{d}{ds} f \circ \gamma\right]_{s=0} \equiv \left[\frac{d\mathcal{F}}{ds}\right]_{s=0}. \tag{2.9}$$

Note that this map is explicitly chart independent.

It is clear that each distinct curve passing through p defines a distinct map. Let us denote the set of all such maps by $T_p(\mathcal{M})$, which correspond to the directional derivatives along every smooth curve passing through p, parametrized such that $s = 0$ at p.

Claim 2.3.1. Let ϕ be a chart such that

$$\phi : p \to x^\mu(p). \tag{2.10}$$

Then

$$\dot{\mathcal{F}}(0) \equiv \left[\frac{d}{ds} f \circ \gamma\right]_{s=0} = \sum_{\mu=1}^{n} \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{dx^\mu(\gamma(s))}{ds}\right]_{s=0} \tag{2.11}$$

where we recall that $F = f \circ \phi^{-1}$.

Proof.

$$\mathcal{F}(s) = f \circ \gamma = f \circ \phi^{-1} \circ \phi \circ \gamma = F \circ \phi \circ \gamma. \tag{2.12}$$

The function $\phi \circ \gamma$ maps s to the coordinates of the curve $\gamma(s)$ and hence

$$\mathcal{F}(s) = F(x^1(s), x^2(s), ..., x^n(s)). \tag{2.13}$$

The chain rule now gives

$$\begin{aligned}\Big(\frac{d\mathcal{F}(s)}{ds}\Big)_{s=0} &= \Big(\frac{d}{ds}F(x^1(s), x^2(s), ..., x^n(s))\Big)_{s=0} \\ &= \sum_{\mu=1}^{n}\Big(\frac{\partial F}{\partial x^\mu}\Big)_{\phi(p)}\Big[\frac{dx^\mu(\gamma(s))}{ds}\Big]_{s=0} \\ &= \Big(\frac{\partial F}{\partial x^\mu}\Big)_{\phi(p)}\Big[\frac{dx^\mu(\gamma(s))}{ds}\Big]_{s=0},\end{aligned} \tag{2.14}$$

where in the last line we have adopted the **Einstein Summation Convention** which is a convenient notation whereby summation over repeat indices is implied. ∎

We have shown how elements of $T_p(\mathcal{M})$ act on functions in a given chart. We will now show that these elements are vectors living in an n-dimensional vector space and we will also obtain a basis for this vector space.

Theorem 2.3.1. $T_p(\mathcal{M})$ is a vector space at $p \in \mathcal{M}$, i.e., this set is closed under addition and scalar multiplication.

Proof.

- **Closure Under Addition**: The elements of $T_p(\mathcal{M})$ are defined by their action on a function, which gives the rate of change of a function evaluated along a smooth curve, $\gamma(s)$ say, passing through p, parametrized such that $\gamma(0) = p$. We let $\gamma(s)$, $\sigma(s)$ be two smooth curves through $p \in \mathcal{M}$ such that $\gamma(0) = \sigma(0) = p$ with which we associate the maps $\dot{\gamma}_p$ and $\dot{\sigma}_p$, respectively. Two curves in $\mathcal{M}$ may be added by adopting a chart ϕ which gives coordinate representations $\{x^\mu(\gamma(s))\}$, $\{x^\mu(\sigma(s))\}$ of the curves in $\mathbb{R}^n$, yielding a new curve

$$\hat{\nu} : s \to \{x^\mu(\gamma(s)) + x^\mu(\sigma(s)) - x^\mu(p)\} \tag{2.15}$$

  which is the parametric representation in $\mathbb{R}^n$ of a curve ν in $\mathcal{M}$ with $\nu(0) = p$. Now for an arbitrary function f, we have

$$\dot{\nu}_p f = \Big(\frac{\partial F}{\partial x^\mu}\Big)_{\phi(p)}\Big[\frac{d}{ds}\big(x^\mu(\gamma(s)) + x^\mu(\sigma(s)) - x^\mu(p)\big)\Big]_{s=0} \tag{2.16}$$

$$= \Big(\frac{\partial F}{\partial x^\mu}\Big)_{\phi(p)}\Big[\frac{d}{ds}\big(x^\mu(\gamma(s)) + x^\mu(\sigma(s)\big)\Big]_{s=0} \tag{2.17}$$

$$= \dot{\gamma}_p f + \dot{\sigma}_p f. \tag{2.18}$$

Since this is true for arbitrary f, we identify $\dot{\nu}_p = \dot{\gamma}_p + \dot{\sigma}_p$. It is clear too that $\dot{\nu}_p$ is an element of the set $T_p(\mathcal{M})$, since its action on a function gives the rate of change at p of the function evaluated along a smooth curve which has been defined so that $\nu(0) = p$. Hence, the set is closed under addition.

- **Closure Under Scalar Multiplication**: Next, we require closure under scalar multiplication. We may multiply a curve γ in $\mathcal{M}$ by a scalar, α, by considering its coordinate image in $\mathbb{R}^n$ given by

$$\hat{\nu} : s \to \alpha x^\mu(\gamma(s)) + (1 - \alpha)x^\mu(p). \tag{2.19}$$

We note again that this is the parametric representation in $\mathbb{R}^n$ of a curve ν in $\mathcal{M}$ parametrized such that $\nu(0) = p$. For an arbitrary function f, we have

$$\dot{\nu}_p f = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{d}{ds}(\alpha x^\mu(\gamma(s)) + (1 - \alpha)x^\mu(p))\right]_{s=0} \tag{2.20}$$

$$= \alpha \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{dx^\mu(\gamma(s))}{ds}\right]_{s=0} \tag{2.21}$$

$$= \alpha \dot{\gamma}_p f. \tag{2.22}$$

Since this is true for arbitrary f, we identify $\dot{\nu}_p = \alpha\dot{\gamma}_p$ and since it is also clear that $\dot{\nu}_p$ is in $T_p(\mathcal{M})$, the set is closed under scalar multiplication.

Hence $T_p(\mathcal{M})$ is a vector space, which we call the tangent space at p and elements of the **tangent space** are called **tangent vectors**. ▮

Let us now show that the tangent space at each point of the manifold has the same dimension as the manifold. We will prove this by obtaining a basis for the space.

Theorem 2.3.2. $\dim(T_p(\mathcal{M})) = \dim(\mathcal{M}) = n$

Proof. We must show that the dimension of the tangent space at p is $n = \dim(\mathcal{M})$. We consider a chart ϕ

$$\phi : p \to x^\mu(p) \tag{2.23}$$

such that $x^\mu(p) = 0$ for all $\mu = 1, ..., n$, i.e., p gets mapped to the origin of $\mathbb{R}^n$. We consider the curve $\gamma_\nu(s)$ through p such that its coordinate representation in $\mathbb{R}^n$ is $\hat{\gamma}_\nu(s) = (0, ..., s, ..., 0)$, the ν^{th}-axis, i.e., $\hat{\gamma}_1 = (s, 0, 0, ..., 0)$, $\hat{\gamma}_2 = (0, s, 0, ..., 0)$ etc. Obviously, there are n such curves and they satisfy

$$x^\mu \circ \gamma_\nu(s) = \begin{cases} s & \mu = \nu \\ 0 & \text{otherwise.} \end{cases} \tag{2.24}$$

Therefore

$$\left[\frac{d}{ds}(x^\mu \circ \gamma_\nu(s))\right]_p = \delta^\mu{}_\nu. \tag{2.25}$$

Moreover,

$$(\dot{\gamma}_\nu)_p f = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{d}{ds}(x^\mu \circ \gamma_\nu(s))\right]_{s=0} = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \delta^\mu{}_\nu = \left(\frac{\partial F}{\partial x^\nu}\right)_{\phi(p)} \tag{2.26}$$

so that $(\dot{\gamma}_\nu)_p$ is a vector at p mapping

$$f \to \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}. \tag{2.27}$$

We will now show that the set $\{(\dot{\gamma}_\nu)_p\}$ form a basis for the tangent space $T_p(\mathcal{M})$ and since there are n such vectors, this will prove the result $\dim(T_p(\mathcal{M})) = n = \dim(\mathcal{M})$.

We first prove that $\{(\dot{\gamma}_\nu)_p\}$ span the tangent space at p. Let $\dot{\lambda}_p$ be a tangent vector at p, then

$$\dot{\lambda}_p(f) = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{d(x^\mu \circ \lambda(s))}{ds}\right]_{s=0} \qquad = c^\mu (\dot{\gamma}_\mu)_p(f) \tag{2.28}$$

where we have set

$$c^\mu = \left[\frac{d(x^\mu \circ \lambda(s))}{ds}\right]_{s=0}. \tag{2.29}$$

This holds for all f so that

$$\dot{\lambda}_p = c^\mu (\dot{\gamma}_\mu)_p, \tag{2.30}$$

and hence any element of the tangent space at p may be written as a linear combination of $\{(\dot{\gamma}_\mu)_p\}$, i.e., the set $\{(\dot{\gamma}_\mu)_p\}$ span the tangent space at p.

Next we show that $\{(\dot{\gamma}_\mu)_p\}$ are linearly independent. Suppose that

$$a^\mu (\dot{\gamma}_\mu)_p = 0 \tag{2.31}$$

where a^μ are constants. Consider the function $x^\nu : \mathcal{M} \to \mathbb{R}$, then

$$0 = a^\mu (\dot{\gamma}_\mu)_p (x^\nu) \tag{2.32}$$

$$= a^\mu \left(\frac{\partial x^\nu}{\partial x^\mu}\right)_{\phi(p)} \tag{2.33}$$

$$= a^\mu \delta^\mu{}_\nu \tag{2.34}$$

$$= a^\nu. \tag{2.35}$$

Hence the $\{(\dot{\gamma}_\mu)_p\}$ are linearly independent; they form a basis for the tangent space and since there are n elements of this basis, we must have that $\dim(T_p(\mathcal{M})) = n$. ∎

We usually write this basis as

$$(\dot{\gamma}_\mu)_p = \left(\frac{\partial}{\partial x^\mu}\right)_p = (\partial_\mu)_p. \tag{2.36}$$

This is known as a **coordinate induced basis** (or just coordinate basis or natural basis) for the tangent space $T_p(\mathcal{M})$, it corresponds to setting up the basis vectors so that they point along the coordinate axes. Hence, a change of coordinates induces a change of basis.

2.4 Transformation Rule for Vectors

We have shown that any vector in the tangent space may be naturally decomposed with respect to a coordinate basis, i.e., for $X_p \in T_p(\mathcal{M})$, we have

$$X_p = X_p^\mu \Big(\frac{\partial}{\partial x^\mu}\Big)_p \tag{2.37}$$

where X_p^μ are the components of vector X_p with respect to a coordinate basis $\{x^\mu\}$.

If we now introduce a change of coordinates

$$x^{\mu'} = x^{\mu'}(x^\mu), \tag{2.38}$$

then

$$\begin{aligned} X_p = X_p^\mu \Big(\frac{\partial}{\partial x^\mu}\Big)_p &= X_p^{\mu'} \Big(\frac{\partial}{\partial x^{\mu'}}\Big)_p \\ &= X_p^{\mu'} \Big(\frac{\partial x^\mu}{\partial x^{\mu'}}\Big)_p \Big(\frac{\partial}{\partial x^\mu}\Big)_p \end{aligned} \tag{2.39}$$

and comparing the two we obtain

$$\boxed{X_p^\mu = X_p^{\mu'} \Big(\frac{\partial x^\mu}{\partial x^{\mu'}}\Big)_p \quad \text{or} \quad X_p^{\mu'} = X_p^\mu \Big(\frac{\partial x^{\mu'}}{\partial x^\mu}\Big)_p} \tag{2.40}$$

This is the vector component transformation rule in a coordinate induced basis.

Note: Vectors are invariant under a coordinate transformation or a change of basis but the components are not.

Of course, we are not limited to coordinate bases for $T_p(\mathcal{M})$ and any linearly independent set $\{e_\mu\}$ spanning the space is an appropriate basis. In general then, we have

$$X = X^\mu e_\mu \tag{2.41}$$

where X^μ are the components of the vector X with respect to the basis $\{e_\mu\}$. This basis is related to any other basis by a non-singular set of transformations

$$e_{\mu'} = \Lambda^\mu{}_{\mu'} e_\mu \tag{2.42}$$

where Λ and its inverse satisfy

$$\Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu = \delta^\mu{}_\nu. \tag{2.43}$$

We then have

$$X = X^\mu e_\mu = X^{\mu'} e_{\mu'} = X^{\mu'} \Lambda^\mu{}_{\mu'} e_\mu \tag{2.44}$$

and hence

$$\boxed{X^\mu = \Lambda^\mu{}_{\mu'} X^{\mu'} \quad \text{or} \quad X^{\mu'} = \Lambda^{\mu'}{}_\mu X^\mu} \tag{2.45}$$

This is the vector component transformation rule in an arbitrary basis.

Chapter 3

Covectors and Tensors

3.1 Covectors and the Cotangent Space

All vector spaces have a corresponding dual space of equal dimension, comprising the set of all *linear* maps from the vector space to the real line.

The dual space of $T_p(\mathcal{M})$ is denoted by $T_p^*(\mathcal{M})$, the cotangent vector space with elements

$$\lambda : T_p(\mathcal{M}) \to \mathbb{R} \tag{3.1}$$

Elements of the cotangent space are known as covectors or one-forms (or covariant vectors in older textbooks). The action of a one-form λ on a tangent vector X_p is written as $\lambda(X_p)$ or $\langle \lambda, X_p \rangle$.

These maps are by definition linear, and hence

$$\lambda(a\, X_p + b\, Y_p) = a\, \lambda(X_p) + b\, \lambda(Y_p), \tag{3.2}$$

where X_p, $Y_p \in T_p(\mathcal{M})$ and a, $b \in \mathbb{R}$. It is also straight-forward to define addition and multiplication by real numbers, i.e.,

$$\begin{aligned} (a\, \lambda)(X_p) &= a\, \lambda(X_p) \\ (\lambda + \omega)(X_p) &= \lambda(X_p) + \omega(X_p), \end{aligned} \tag{3.3}$$

and hence one-forms at p form a vector space. One may now regard the tangent vectors at p as linear real-valued maps over the one-forms by taking a vector X_p and associating to each one-form the value $\lambda(X_p) \equiv X_p(\lambda)$ where the latter notation is adopted to explicitly convey that vectors are functions over covectors. That this map is linear follows directly from Eqs. (3.3).

Hence, vectors are real-valued functions over covectors and covectors are real-valued functions over vectors and we say that vectors and covectors are dual to each other. The duality is best represented in the notation $\langle \lambda, X_p \rangle$ which is bi-linear and emphasizes the fact that it can be thought of as a map from $T_p^*(\mathcal{M})$ to $\mathbb{R}$ or from $T_p(\mathcal{M})$ to $\mathbb{R}$.

For a tangent space $T_p(\mathcal{M})$ with basis $\{e_\mu\}$, there is a corresponding dual basis $\{e^\mu\}$ of $T_p^*(\mathcal{M})$ defined by

$$e^\mu(e_\nu) = \langle e^\mu, e_\nu \rangle = \delta^\mu{}_\nu. \tag{3.4}$$

3.2 The Gradient/Differential

A particularly important covector is the so-called gradient or differential of a function, defined as follows:

Definition 3.2.1. Let $p \in \mathcal{M}$ and $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. For $X_p \in T_p(\mathcal{M})$, we define the the gradient (or differential) df to be

$$df : T_p(\mathcal{M}) \to \mathbb{R}, \qquad df(X_p) = X_p(f). \tag{3.5}$$

For a particular chart, one may associate with the coordinate x^μ, the covector dx^μ defined by

$$dx^\mu(X_p) = X_p(x^\mu), \tag{3.6}$$

which gives the μ^{th} component of the vector X_p in a coordinate induced basis.

Claim 3.2.1. The set $\{dx^\mu\}$ form a basis for $T_p^*(\mathcal{M})$ dual to the coordinate basis $\{(\partial/\partial x^\mu)_p\}$ of $T_p(\mathcal{M})$.

Proof. We must prove that the set $\{dx^\mu\}$ form a basis and also that the duality condition is satisfied. The latter part is trivial,

$$\langle dx^\mu, (\partial/\partial x^\nu)_p \rangle = \left(\frac{\partial}{\partial x^\nu}\right)_p x^\mu = \delta^\mu{}_\nu, \tag{3.7}$$

as required. To show that $\{dx^\mu\}$ form a basis for the cotangent space, we must prove that this set spans the space and are linearly independent. Now, suppose there exists a set of coefficients $\{\omega_\mu\}$ such that

$$\omega \equiv \omega_\mu \, dx^\mu = 0. \tag{3.8}$$

Then

$$\begin{aligned} 0 = \langle \omega, (\partial/\partial x^\nu)_p \rangle &= \omega_\mu \langle dx^\mu, (\partial/\partial x^\nu)_p \rangle \\ &= \omega_\mu \delta^\mu{}_\nu \\ &= \omega_\nu, \end{aligned} \tag{3.9}$$

and hence the elements of $\{dx^\mu\}$ are linearly independent.

To show the set spans the space, we let $\omega \in T_p^*(\mathcal{M})$, and define

$$\lambda = \omega - \langle \omega, (\partial/\partial x^\mu)_p \rangle \, dx^\mu. \tag{3.10}$$

Letting λ act on the vector $(\partial/\partial x^\nu)_p$ gives

$$\begin{aligned} \lambda((\partial/\partial x^\nu)_p) = \langle \lambda, (\partial/\partial x^\nu)_p \rangle &= \langle \omega, (\partial/\partial x^\nu)_p \rangle - \langle \omega, (\partial/\partial x^\mu)_p \rangle \langle dx^\mu, (\partial/\partial x^\nu)_p \rangle \\ &= \langle \omega, (\partial/\partial x^\nu)_p \rangle - \langle \omega, (\partial/\partial x^\mu)_p \rangle \, \delta^\mu{}_\nu \\ &= 0, \end{aligned} \tag{3.11}$$

for all ν. Hence $\lambda \equiv 0$ and

$$\omega = \langle \omega, (\partial/\partial x^\mu)_p \rangle \, dx^\mu, \tag{3.12}$$

i.e., the set $\{dx^\mu\}$ spans $T_p^*(\mathcal{M})$. ∎

It is clear from Eq.(3.12) that $\langle \omega, (\partial/\partial x^\mu)_p \rangle$ gives the μ^{th} component of a covector ω in a coordinate induced basis. For example, the gradient is

$$\begin{aligned} df &= \langle df, (\partial/\partial x^\mu)_p \rangle \, dx^\mu \\ &= \left(\frac{\partial}{\partial x^\mu} \right)_p f \, dx^\mu \\ &= \left(\frac{\partial F}{\partial x^\mu} \right)_{\phi(p)} dx^\mu. \end{aligned} \tag{3.13}$$

3.3 Transformation Rule for Covectors

As with the tangent space, a coordinate transformation induces a change in the coordinate induced basis of the cotangent space. if we introduce new coordinates

$$x^{\mu'} = x^{\mu'}(x^\nu), \tag{3.14}$$

then we have

$$dx^{\mu'} = \left(\frac{\partial x^{\mu'}}{\partial x^\mu} \right)_{\phi(p)} dx^\mu. \tag{3.15}$$

Hence, for $\omega \in T_p^*(\mathcal{M})$, we have

$$\begin{aligned} \omega = \omega_\mu dx^\mu &= \omega_{\mu'} dx^{\mu'} \\ &= \omega_{\mu'} \left(\frac{\partial x^{\mu'}}{\partial x^\mu} \right)_{\phi(p)} dx^\mu, \end{aligned} \tag{3.16}$$

which implies the components of a covector in a corodinate basis transform as

$$\boxed{\omega_\mu = \omega_{\mu'} \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right)_{\phi(p)} \qquad \text{or} \qquad \omega_{\mu'} = \omega_\mu \left(\frac{\partial x^{\mu}}{\partial x^{\mu'}}\right)_{\phi(p)}.} \tag{3.17}$$

As before, we are not restricted to a coordinate basis. If we let $w = \omega_\mu \, e^\mu = \omega_{\mu'} \, e^{\mu'}$ where $\{e^\mu\}$ and $\{e^{\mu'}\}$ are two bases for $T_p^*(\mathcal{M})$ related by

$$e^\mu = e^{\mu'} \Lambda^\mu{}_{\mu'}, \tag{3.18}$$

then the components in these bases transform according to

$$\boxed{\omega_{\mu'} = \omega_\mu \, \Lambda^\mu{}_{\mu'} \qquad \text{or} \qquad \omega_\mu = \omega_{\mu'} \, \Lambda^{\mu'}{}_\mu.} \tag{3.19}$$

3.4 Tensors

We can generalize the linear maps we have called tangent vectors and tangent covectors to more general multi-linear maps, known as tensors. For example

$$T : \underbrace{T_p^*(\mathcal{M}) \times T_p^*(\mathcal{M}) \times \cdots T_p^*(\mathcal{M})}_{r} \times \underbrace{T_p(\mathcal{M}) \times T_p(\mathcal{M}) \times \cdots T_p(\mathcal{M})}_{s} \to \mathbb{R} \tag{3.20}$$

is known as a tensor of type (rank) $\binom{r}{s}$ at p which takes as its arguments r one-forms and s vectors. Special cases of tensors are $\binom{1}{0}$ tensors which are simply vectors while $\binom{0}{1}$ tensors are one-forms. We define scalar functions to be $\binom{0}{0}$ tensors.

A $\binom{2}{2}$ tensor takes two one-forms λ, ω and two vectors X, Y and outputs the real number

$$T(\lambda, \omega; X, Y), \tag{3.21}$$

where we separate the covector arguments from a vector one by a semi-colon. This is linear in each argument, i.e.

$$\begin{aligned}
T(a\,\lambda + b\,\sigma, \omega; X, Y) &= a\,T(\lambda, \omega; X, Y) + b\,T(\sigma, \omega; X, Y) \\
T(\lambda, a\,\omega + b\,\sigma; X, Y) &= aT(\lambda, \omega; X, Y) + b\,T(\lambda, \sigma; X, Y) \\
T(\lambda, \omega; a\,X + b\,Z, Y) &= a\,T(\lambda, \omega; X, Y) + b\,T(\lambda, \omega; Z, Y) \\
T(\lambda, \omega, X, a\,Y + b\,Z) &= a\,T(\lambda, \omega; X, Y) + b\,T(\lambda, \omega; X, Z),
\end{aligned} \tag{3.22}$$

where a, $b \in \mathbb{R}$ and $\sigma \in T_p^*(\mathcal{M})$, $Z \in T_p(\mathcal{M})$. We can also combine tensors with vectors and one-forms to define new tensors of different rank. For example, a $\binom{1}{1}$ tensor takes a one-form and a vector and outputs a real number, $T(\lambda; X)$ say. However, if we regard λ as fixed, then $T(\lambda; \quad)$ maps a vector X to the real number $T(\lambda; X)$ and hence it is a one-form. Similarly, if we regard X as fixed, then $T(\quad; X)$ is a vector since it is a map from one-forms to the reals.

Two tensors may be multiplied by defining the **tensor product** as follows: If T_1 is a $\binom{r_1}{s_1}$ tensor and T_2 is a $\binom{r_2}{s_2}$ tensor, then the tensor product is

$$\begin{aligned} &T_1 \otimes T_2(\omega_1, \omega_2, ..., \omega_{r_1+r_2}; X_1, X_2, ..., X_{s_1+s_2}) \\ &\equiv T_1(\omega_1, \omega_2, ..., \omega_{r_1}; X_1, X_2, ..., X_{s_1})\, T_2(\omega_{r_1+1}, \omega_{r_1+1}, ..., \omega_{r_1+r_2}; X_{s_1+1}, X_{s_2+2}, ..., X_{s_1+s_2}). \end{aligned} \tag{3.23}$$

An appropriate coordinate basis for the space of $\binom{r}{s}$ tensors at p is therefore

$$\{\partial_{\mu_1} \otimes \partial_{\mu_2} \otimes \cdots \otimes \partial_{\mu_r} \otimes dx^{\nu_1} \otimes dx^{\nu_2} \otimes \cdots \otimes dx^{\nu_s}\}. \tag{3.24}$$

Decomposed in this basis, an $\binom{r}{s}$ tensor is

$$T = T^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s}\, \partial_{\mu_1} \otimes \partial_{\mu_2} \otimes \cdots \otimes \partial_{\mu_r} \otimes dx^{\nu_1} \otimes dx^{\nu_2} \otimes \cdots \otimes dx^{\nu_s}. \tag{3.25}$$

Since we know how both the basis vectors and covectors in the tensor basis transform under a coordinate transformation, it is straightforward to show that the components in the chart $\{x^{\mu'}\}$ are related to those in the chart $\{x^{\mu}\}$ by

$$\boxed{T^{\mu_1'\mu_2'...\mu_r'}{}_{\nu_1'\nu_2'...\nu_s'} = \frac{\partial x^{\mu_1'}}{\partial x^{\mu_1}} \frac{\partial x^{\mu_2'}}{\partial x^{\mu_2}} \cdots \frac{\partial x^{\mu_r'}}{\partial x^{\mu_r}} \frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}} \frac{\partial x^{\nu_2}}{\partial x^{\nu_2'}} \cdots \frac{\partial x^{\nu_s}}{\partial x^{\nu_s'}}\, T^{\mu_1\mu_2...\mu_r}{}_{\nu_1\nu_2...\nu_s}.} \tag{3.26}$$

This is the tensor component transformation rule in a coordinate induced basis, with the obvious generalization to an arbitrary basis.

Example 3.4.1. Show that the Kronecker delta $\delta^\mu{}_\nu$ represents the components of a $\binom{1}{1}$ tensor.

To prove this, we must show that they transform in the appropriate way, i.e.,

$$\begin{aligned} \delta^\mu{}_\nu = \frac{\partial x^\mu}{\partial x^\nu} &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^{\mu'}}{\partial x^{\nu'}} \\ &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \delta^{\mu'}{}_{\nu'}. \end{aligned} \tag{3.27}$$

Hence $\delta^\mu{}_\nu$ transforms like a $\binom{1}{1}$ tensor under a change of coordinates. ∎

From an $\binom{r}{s}$ tensor, one can obtain a tensor of type $\binom{r-1}{s-1}$ by contraction which corresponds to multiplying by the Kronecker delta. This is known as a **contraction**. For example, given a $\binom{1}{3}$ tensor, we can form a $\binom{0}{2}$ tensor by **contracting** over the first and third index,

$$T^\mu{}_{\nu\lambda\rho} \quad \xrightarrow{\text{contraction}} \quad T^\mu{}_{\nu\mu\rho} = \delta^\lambda{}_\mu T^\mu{}_{\nu\lambda\rho}. \tag{3.28}$$

We will also make use of the Bach bracket notation where round parentheses about a set of indices denotes symmetrization of indices and square brackets denotes anti-symmetrization. For example, if T is a $\binom{1}{2}$ tensor, then

$$\begin{aligned} T^\mu{}_{(\nu\lambda)} &= \frac{1}{2}(T^\mu{}_{\nu\lambda} + T^\mu{}_{\lambda\nu}) \\ T^\mu{}_{[\nu\lambda]} &= \frac{1}{2}(T^\mu{}_{\nu\lambda} - T^\mu{}_{\lambda\nu}). \end{aligned} \tag{3.29}$$

In general, we have

$$
\begin{aligned}
T_{(\mu_1\mu_2...\mu_r)} &= \frac{1}{r!}(\text{sum of all permutations of the indices}) \\
T_{[\mu_1\mu_2...\mu_r]} &= \frac{1}{r!}(\text{alternating sum of all permutations of the indices}),
\end{aligned} \tag{3.30}
$$

where for the anti-symmetric case, the terms which pick up a minus sign are those whose indices are an odd permutation of $\{\mu_1\mu_2...\mu_r\}$. For example,

$$
T_{[\mu\nu\lambda]} = \frac{1}{6}\Big(T_{\mu\nu\lambda} - T_{\mu\lambda\nu} + T_{\lambda\mu\nu} - T_{\lambda\nu\mu} + T_{\nu\lambda\mu} - T_{\nu\mu\lambda}\Big). \tag{3.31}
$$

Chapter 4

The Tangent Bundle and Vector Fields

4.1 Tangent Bundle as a Fiber Bundle

The union of all tangent vector spaces for each point of the manifold defines the tangent bundle, denoted by $T(\mathcal{M})$,

$$T(\mathcal{M}) = \bigcup_{p \in \mathcal{M}} T_p(\mathcal{M}). \tag{4.1}$$

This is itself a manifold of dimension $2n$. To see this, consider a 1-dimensional differentiable manifold (i.e., a smooth curve) and its tangent spaces (which are simply tangent lines at each point of the curve).

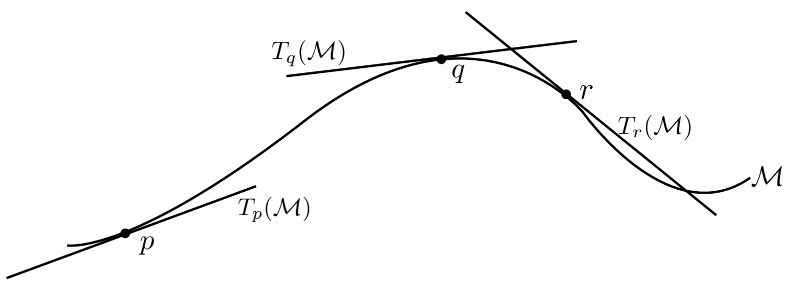

Figure 4.1:

The intersections of the tangent lines with one another and with the manifold in the figure above are meaningless and in general, there is no natural way to compare vectors at different points of the manifold since they live in different tangent spaces. This picture becomes somewhat clearer if we rotate each line so that the tangent spaces are nowhere intersecting and they cross $\mathcal{M}$ only at the point where they are defined, as in the figure below.

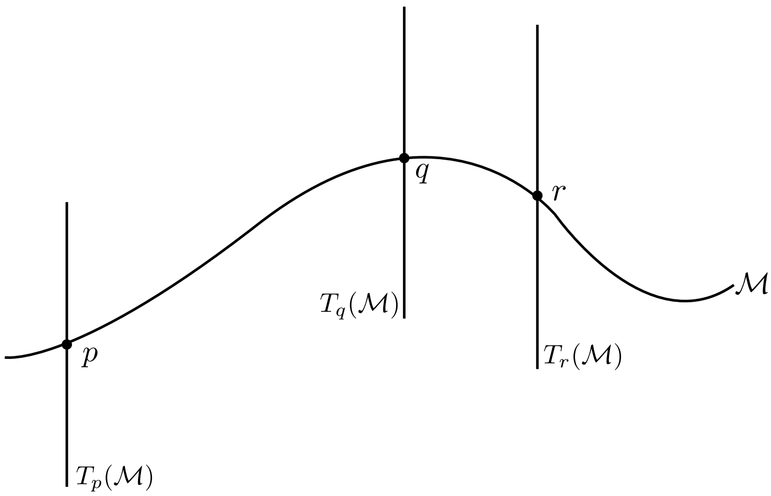

Figure 4.2:

When the tangent bundle is drawn in this way, every point along $T_p(\mathcal{M})$ represents a vector in that tangent space and every point in the figure is a point of one and only one tangent space. That this is a two-dimensional manifold can be seen by constructing a chart for $T(\mathcal{M})$ as follows: Take x to be a coordinate over $\mathcal{M}$, then every vector in each tangent space may be decomposed in a coordinate induced basis by $X = y\,\partial/\partial x$. So y is a coordinate for $T_p(\mathcal{M})$. Each tangent space has fixed x-value and so the coordinates (x, y) pick out a particular vector y tangent to a particular point x. Since every point $T(\mathcal{M}$ lies in a region of this sort, we have constructed an atlas for $T(\mathcal{M})$, hence it is a manifold. This construction generalizes to higher dimensions, though it is impossible to visualize.

The tangent bundle is an example of a more general construct called a **Fiber Bundle**, which consists of two spaces B and E and a projection map $\pi : E \to B$ which associates to each element of E a point in the space B, known as the base space. For each point $b \in B$, the space $\pi^{-1}(b) \in E$ is known as a fiber over b. A fiber bundle is "locally trivial", i.e., it is locally the (Cartesian) product space $B \times E$ but its global properties are determined by a structure group on E.

In the language of fiber bundles, the base space of the tangent bundle is the differentiable manifold $\mathcal{M}$, the fibers are the tangent spaces at each point $T_p(\mathcal{M})$ and the structure group is $GL(n, \mathbb{R})$, which is the set of $n \times n$ matrices with non-zero determinant, reflecting the fact that tangent vector components are related by invertible $n \times n$ basis transformation matrices as we saw in the previous chapter.

4.2 Vector Fields

A vector field on $\mathcal{M}$ is a map that specifies a unique vector at each point of the manifold,

$$X : \mathcal{M} \to T(\mathcal{M}), \qquad p \to X_p \in T_p(\mathcal{M}). \tag{4.2}$$

In the context of the fiber bundle picture in Fig. 4.2, a vector field is any curve which intersects each fiber (tangent space) at one and only one point, i.e., a curve which is nowhere parallel to a fiber. Or put another way, a vector field is a *cross-section* of $T(\mathcal{M})$.

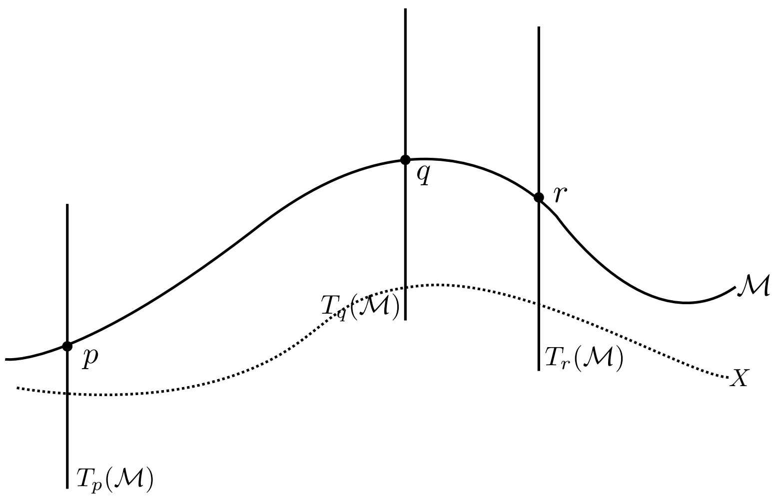

Figure 4.3:

In a coordinate basis $\{x^\mu\}$, a vector field is defined by taking the vector X_p to be

$$X_p = X_p^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p \qquad \text{where} \qquad X_p^\mu = X_p(x^\mu). \tag{4.3}$$

Henceforth, we can tentatively drop the p subscript since there is a unique vector at each point of the manifold.

We saw previously that a vector at p maps a function (scalar) to a real number $X_p(f)$. A vector field then acts on a scalar function to give a scalar function whose value at each point $p \in \mathcal{M}$ is $X_p(f)$,

$$X(f) : \mathcal{M} \to \mathbb{R}, \qquad p \to X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}. \tag{4.4}$$

It is straightforward to see that vector fields satisfy the Leibniz Rule (generalization of the product rule)

$$X(f \cdot g) = X(f) \cdot g + f \cdot X(g) \tag{4.5}$$

and since we already know they are linear maps, we have that vector fields are *derivations* over the smooth real functions.

4.3 Covector and Tensor Fields

The definition of a vector field above naturally extends to covector fields and tensor fields. A covector field, or field of one-forms, is an assignment to each point of $p \in \mathcal{M}$ a unique covector. We can also define a covector field as a cross-section of the cotangent bundle $T^*(\mathcal{M})$ where the fiber at p is the cotangent spaces $T^*_p(\mathcal{M})$. Recall that a covector at p maps a vector at p to a real number. A covector field maps a vector at each point to a real number, i.e., $\omega(X)$ is a scalar function over the reals. Similarly, linearity of a covector at p implies

$$\omega(a\,X_p + b\,Y_p) = \langle \omega, a\,X_p + b\,Y_p \rangle = a\,\omega(X_p) + b\,\omega(Y_p) \tag{4.6}$$

where a and b are real numbers. For a covector field, a and b can take on different values at each point, i.e., they are scalar functions. Hence linearity of a covector field is with respect to scalar multiplication

$$\omega(f\,X + g\,Y) = f\,\omega(X) + g\,\omega(Y) \tag{4.7}$$

which is itself a scalar function over the reals.

A tensor field assigns a unique tensor of equal rank to each point of a manifold. Again, we can formulate this in the language of fiber bundles where each fiber is an appropriate product of the tangent and cotangent spaces, depending on the rank of the tensor field. Linearity of a tensor field is with respect to multiplication by scalar functions. For example, for a $\binom{1}{1}$ tensor,

$$T(f\,\omega + g\,\eta; u\,X + v\,Y) = f\,u\,T(\omega;X) + f\,v\,T(\omega;Y) + g\,u\,T(\eta;X) + g\,v\,T(\eta;Y), \tag{4.8}$$

where ω, $\eta \in T^*_p(\mathcal{M})$, X, $Y \in T_p(\mathcal{M})$ and f, g, u, v are real-valued functions.

4.4 The Commutator (Lie Bracket)

Let X, Y be vector fields and f an arbitrary scalar, then we can consider the composition

$$\begin{aligned} X(Y(f)) &= X^\mu \frac{\partial}{\partial x^\mu}\Big(Y^\nu \frac{\partial F}{\partial x^\nu}\Big) \\ &= X^\mu Y^\nu \frac{\partial^2 F}{\partial x^\mu \partial x^\nu} + X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu}. \end{aligned} \tag{4.9}$$

The commutator (or Lie Bracket) is a vector field defined by

$$[X,Y](f) = X(Y(f)) - Y(X(f)). \tag{4.10}$$

In a coordinate basis, we have

$$\begin{aligned} [X,Y](f) &\equiv [X,Y]^\mu \frac{\partial F}{\partial x^\mu} \\ &= X^\mu Y^\nu \frac{\partial^2 F}{\partial x^\mu \partial x^\nu} + X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu} - Y^\mu X^\nu \frac{\partial^2 F}{\partial x^\mu \partial x^\nu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu} \\ &= \Big(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\Big)\frac{\partial F}{\partial x^\mu}, \end{aligned} \tag{4.11}$$

where we have re-labelled some dummy indices to obtain the last line. Since f was arbitrary, we must have that the components of the commutator in a coordinate basis are

$$\begin{aligned}[X,Y]^\mu &= X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu} \\ &= X^\nu Y^\mu{}_{,\nu} - Y^\nu X^\mu{}_{,\nu}.\end{aligned} \tag{4.12}$$

One can show that the commutator satisfies the Leibniz Rule

$$[X,Y](f \cdot g) = [X,Y](f) \cdot g + f[X,Y](g). \tag{4.13}$$

One may also verify the following properties:

1. $[X,X] = 0$,
2. $[X,Y] = -[Y,X]$,
3. $[X,Y+Z] = [X,Y] + [X,Z]$,
4. $[X,f\,Y] = f[X,Y] + X(f)\,Y$,
5. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$, (The Jacobi Identity).

We will return to the geometric interpretation of the commutator later, we must first develop the formalism of maps between manifolds.

Chapter 5

Lie Derivatives

5.1 Maps Between Manifolds

We say a map $h : \mathcal{M} \to \mathcal{N}$ is smooth if for every smooth function $f : \mathcal{N} \to \mathbb{R}$, the function $f \circ h : \mathcal{M} \to \mathbb{R}$ is also smooth.

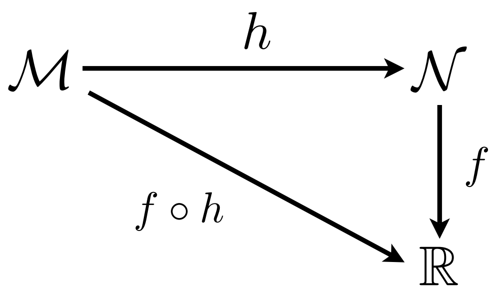

Figure 5.1:

Assuming h is smooth, then it maps a smooth curve γ in $\mathcal{M}$ to a smooth curve $h \circ \gamma$ in $\mathcal{N}$.

Let X_p be the tangent vector to γ at p. Then there exists a map

$$h_* : T_p(\mathcal{M}) \to T_{h(p)}(\mathcal{N}) \tag{5.1}$$

known as the **push-forward** to $h \circ \gamma$ and it maps the tangent vector at p in $\mathcal{M}$ to the tangent vector at $h(p)$ in $\mathcal{N}$. Then for a smooth function $f : \mathcal{N} \to \mathbb{R}$, we have

$$(h_* X_p)(f) = X_p(f \circ h). \tag{5.2}$$

This is clearly linear since X_p is linear.

Analogously, there is a map between covectors that goes in the opposite direction

$$h^* : T^*_{h(p)}(\mathcal{N}) \to T^*_p(\mathcal{M}) \tag{5.3}$$

known as the **pull-back**, which maps tangent covectors at $h(p)$ in $\mathcal{N}$ to the tangent covector at p in $\mathcal{M}$. If $\omega \in T^*_{h(p)}(\mathcal{N})$ and X_p is any vector in $T_p(\mathcal{M})$, then $h^*\omega \in T^*_p(\mathcal{M})$ is defined by

$$h^*\omega(X_p) = \omega(h_* X_p), \tag{5.4}$$

i.e., the pull-back of a covector acting on a vector is equal to the covector acting on the push-forward of the vector.

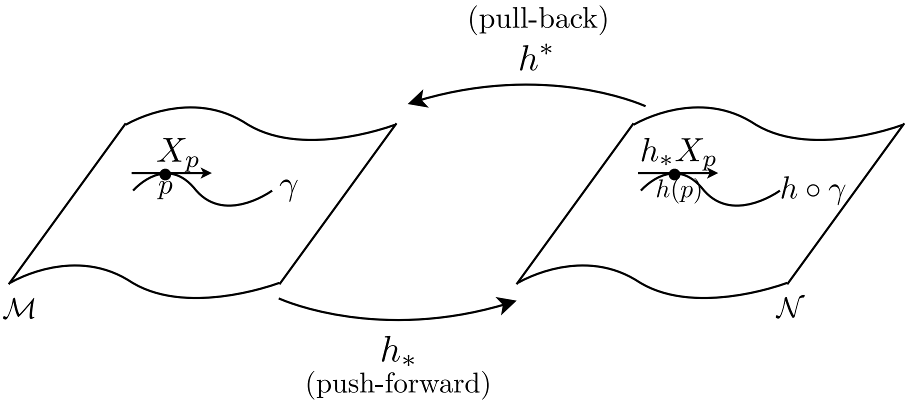

Figure 5.2:

Example 5.1.1. Let $h : \mathbb{R}^2 \to \mathbb{R}^3$ be a map defined by

$$h : (x^1, x^2) = (x, y) \to (x^2 y + y^2, x - 2y^3, ye^{2x-2}) = (y^1, y^2, y^3). \tag{5.5}$$

Let $X_p \in T_p(\mathbb{R}^2)$ be a vector with components $(1, 0)$. Calculate the push-forward $h_* X_p$ and hence show that, for the point p whose coordinate image in the chart (x, y) is $(1, 1)$, we have

$$h_* X_p = 2\frac{\partial}{\partial y^1} + \frac{\partial}{\partial y^2} + 2\frac{\partial}{\partial y^3} \tag{5.6}$$

where the tangent vector on the right-hand side is evaluated at $h(p)$ whose coordinate image in the chart (y^1, y^2, y^3) is $(2, -1, 1)$.

Solution: We define the push-forward by its action on an arbitrary function g as

$$(h_* X_p)g = X_p(g \circ h) = X_p(g(x^2 y + y^2, x - 2y^3, ye^{2x-2})) \tag{5.7}$$

$$= (1)\frac{\partial g}{\partial x} + (0)\frac{\partial g}{\partial y} \tag{5.8}$$

$$= \frac{\partial g}{\partial y^1}\frac{\partial y^1}{\partial x} + \frac{\partial g}{\partial y^2}\frac{\partial y^2}{\partial x} + \frac{\partial g}{\partial y^3}\frac{\partial y^3}{\partial x} \tag{5.9}$$

$$= 2xy\frac{\partial g}{\partial y^1} + 1\frac{\partial g}{\partial y^2} + 2ye^{2x-2}\frac{\partial g}{\partial y^3}. \tag{5.10}$$

Hence in the coordinate induced chart $\{y^i\}$, $(i = 1, 2, 3)$, the push-forward of X_p, $h_* X_p$ has components $(2xy, 1, 2ye^{2x-2})$. At the point p, whose coordinate image is $(x, y) = (1, 1)$, $h_* X_p$ has components $(2, 1, 2)$, i.e.,

$$h_* X_p = 2\frac{\partial}{\partial y^1} + \frac{\partial}{\partial y^2} + 2\frac{\partial}{\partial y^3} \tag{5.11}$$

where the vector on the right-hand side is evaluated at $h(p) = h(1, 1) = (2, -1, 1)$. ∎

5.2 Integral Curves

Let X be a vector field on $\mathcal{M}$. An integral curve of X in $\mathcal{M}$ is a curve γ such that at each point p of γ, the tangent vector is X_p, i.e., $\gamma(s)$ is an integral curve of X if and only if

$$\dot{\gamma}_p = X_{p=\gamma(s)}. \tag{5.12}$$

To see that such integral curves exist and are unique (at least locally), we work in coordinates $\{x^\mu\}$ an we let f be a smooth function, then

$$\begin{aligned} & \dot{\gamma}_p(f) = X_p(f) \\ \iff & \left(\frac{d}{ds}(f \circ \gamma)\right)_p = X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \\ \iff & \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \frac{dx^\mu(\gamma(s))}{ds} = X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}, \end{aligned} \tag{5.13}$$

which is true for arbitrary f. Hence

$$\frac{dx^\mu(s)}{ds} = X_p^\mu(x^1(s), x^2(s), ..., x^n(s)), \tag{5.14}$$

along with the initial condition $x^\mu(s = 0) = x^\mu(p)$. This is a set of n ordinary differential equations with initial conditions which have unique solutions by the standard theory of ODEs.

Example 5.2.1. Let $V = x\partial/\partial y - y\partial/\partial x$ be a vector field on $\mathbb{R}^2$. Plot the integral curves of V.

Solution The integral curves are obtained by solving

$$\frac{dx^\mu}{ds} = V^\mu(x^\nu) \tag{5.15}$$

which gives the set of coupled ODEs

$$\frac{dx}{ds} = -y \qquad \frac{dy}{ds} = x. \tag{5.16}$$

This is trivially solved yielding

$$x^2 + y^2 = C \tag{5.17}$$

where C is an arbitrary constant. Hence the integral curves for each point of the manifold are circles centred on the origin (Fig. 5.3). ∎

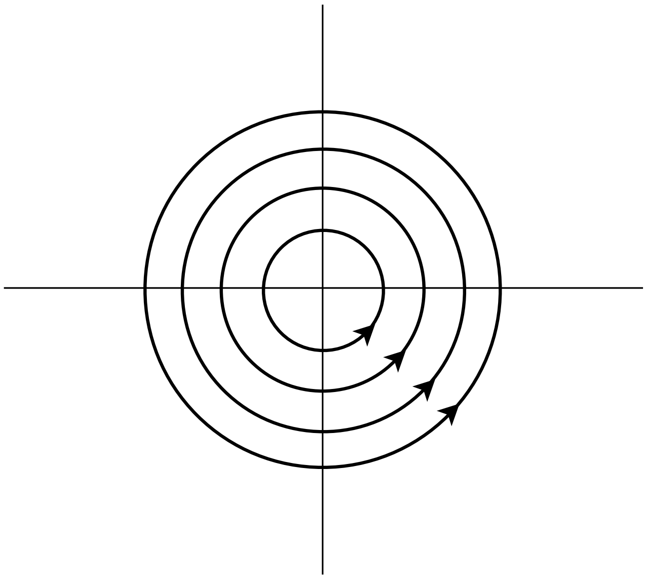

Figure 5.3:

If Eq. (5.14) is globally valid (holds for all s), we say that the integral curve is complete. The set of complete integral curves is a **congruence**. A congruence 'fills' $\mathcal{M}$ in that there is an integral curve passing through each point of $\mathcal{M}$, and in fact the congruence is itself a manifold.

Given a congruence, we may define a one-parameter family of transformations

$$h_s : \mathcal{M} \to \mathcal{M} \tag{5.18}$$

such that $h_s(p)$ is a point on the integral curve through p a parameter "distance" s from p. Then

$$h_s(h_t(p)) = h_{s+t}(p) = h_t(h_s(p)). \tag{5.19}$$

We also have an identity map h_0 such that

$$h_0(p) = p \tag{5.20}$$

and an inverse map $h_s^{-1} = h_{-s}$ such that

$$h_s(h_s^{-1}(p)) = h_{s-s}(p) = h_0(p) = p. \tag{5.21}$$

Therefore, the family of maps h_s forms an Abelian group of transformations $\mathcal{M} \to \mathcal{M}$, known as the **flow** generated by the vector field.

5.3 The Commutator Revisited: A Geometric Interpretation

Let X, and Y be vector fields on $\mathcal{M}$ with which we may associate the flows h_s and k_t, respectively, where $h_s(p)$ is the point along the integral curve of X through p a parameter distance s from p, and similarly for k_t. Starting from a point p, we move a distance ds along the integral curve of X to a point $r = h_{ds}(p)$ followed by moving a distance dt along the integral curve of Y to reach the point $v = k_{dt}(r)$. Starting again from p, we move in the opposite direction by first moving a distance dt along the integral curve of Y to the point $q = k_{dt}(p)$ followed by moving a distance ds along the integral curve of X to the point $u = h_{ds}(q)$. This is represented in the figure below.

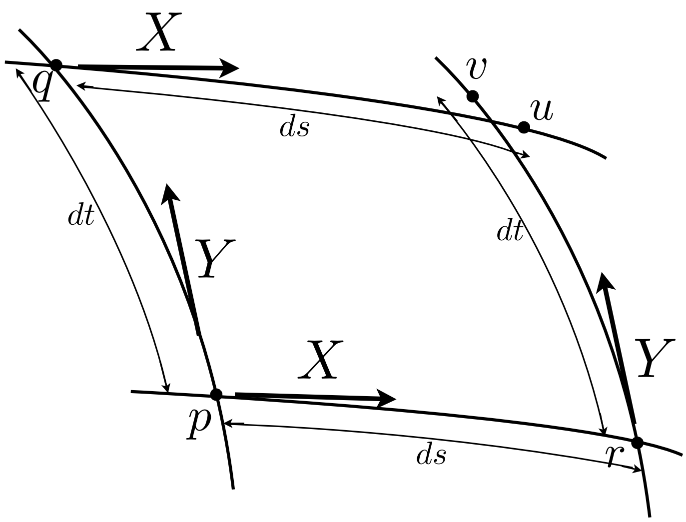

Figure 5.4:

Then if the flow of the vector fields X and Y don't commute, i.e.,

$$k_{dt} \circ h_{ds}(p) \neq h_{ds} \circ k_{dt}(p), \tag{5.22}$$

the coordinate difference, $x^\mu(v) - x^\mu(u)$, will be non-zero. Since these are just functions in $\mathbb{R}^n$, we can perform a Taylor expansion about $x^\mu(p)$ in the usual way, so the coordinate image at p and $r = h_{ds}(p)$ are related by

$$\begin{aligned} x^\mu(r) &= x^\mu(p) + \left(\frac{dx^\mu}{ds}\right)_p ds + \frac{1}{2}\left(\frac{d^2x^\mu}{ds^2}\right)_p ds^2 + \mathcal{O}(ds^3) \\ &= x^\mu(p) + (X^\mu)_p ds + \tfrac{1}{2}(X^\mu{}_{,\nu}X^\nu)_p ds^2 + \mathcal{O}(ds^3) \end{aligned} \tag{5.23}$$

where we have used the fact that if $\{x^\mu(s)\}$ is the coordinate image of the integral curve of the vector field X^μ, then

$$X^\mu = \frac{dx^\mu}{ds} \qquad \text{and} \qquad \frac{d^2x^\mu(s)}{ds^2} = \frac{dX^\mu}{ds} = \frac{\partial X^\mu}{\partial x^\nu}\frac{dx^\nu}{ds} = X^\mu{}_{,\nu}X^\nu. \tag{5.24}$$

Moreover, we have

$$
\begin{aligned}
x^\mu(v) &= x^\mu(r) + \Big(\frac{dx^\mu}{dt}\Big)_r dt + \frac{1}{2}\Big(\frac{d^2x^\mu}{dt^2}\Big)_r dt^2 + \mathcal{O}(dt^3) \\
&= x^\mu(r) + (Y^\mu)_r dt + \tfrac{1}{2}(Y^\mu{}_{,\nu}Y^\nu)_r dt^2 + \mathcal{O}(dt^3) \\
&= x^\mu(p) + (X^\mu)_p ds + \tfrac{1}{2}(X^\mu{}_{,\nu}X^\nu)_p ds^2 + \mathcal{O}(ds^3) + \Big[(Y^\mu)_p + \Big(\frac{dY^\mu}{ds}\Big)_p ds + \mathcal{O}(ds^2)\Big] dt \\
&\quad + \tfrac{1}{2}\Big[(Y^\mu{}_{,\nu}Y^\nu)_p + \mathcal{O}(ds)\Big] dt^2 + \mathcal{O}(dt^3) \\
&= x^\mu(p) + (X^\mu)_p ds + (Y^\mu)_p dt + (Y^\mu{}_{,\nu}X^\nu)_p ds\, dt + \tfrac{1}{2}(X^\mu{}_{,\nu}X^\nu)_p ds^2 + \tfrac{1}{2}(Y^\mu{}_{,\nu}Y^\nu)_p dt^2 \\
&\quad + \mathcal{O}(ds^3 + dt^3). \qquad (5.25)
\end{aligned}
$$

Similarly, by moving in the other direction, we obtain (interchanging X and Y, s and t)

$$
\begin{aligned}
x^\mu(u) &= x^\mu(p) + (Y^\mu)_p dt + (X^\mu)_p ds + (X^\mu{}_{,\nu}Y^\nu)_p ds\, dt + \tfrac{1}{2}(Y^\mu{}_{,\nu}Y^\nu)_p dt^2 + \tfrac{1}{2}(X^\mu{}_{,\nu}X^\nu)_p ds^2 \\
&\quad + \mathcal{O}(ds^3 + dt^3). \qquad (5.26)
\end{aligned}
$$

Subtracting yields a coordinate image of $k_{dt} \circ h_{ds}(p) - h_{ds} \circ k_{dt}(p)$ which is the commutation of the flows generated by the vector fields X and Y through infinitesimal distances ds and dt, respectively:

$$
\begin{aligned}
x^\mu(v) - x^\mu(u) &= (Y^\mu{}_{,\nu}X^\nu - X^\mu{}_{,\nu}Y^\nu)_p ds\, dt + \mathcal{O}(ds^3 + dt^3) \\
&= [X, Y]^\mu_p ds\, dt + \mathcal{O}(ds^3 + dt^3), \qquad (5.27)
\end{aligned}
$$

from which we obtain the following representation of the components of the commutator

$$
[X, Y]^\mu_p = \lim_{ds, dt \to 0} \frac{x^\mu(v) - x^\mu(u)}{ds\, dt}. \qquad (5.28)
$$

Hence, the commutator $[X, Y]$ measures the discrepancy between the point u and v obtained by following the integral curves of the vector fields X and Y in different orders, starting from a point p and moving infinitesimal distances along the curves. Or put more succinctly, the Lie Bracket of two vector fields measures the lack of commutation of their flows.

If the Lie Bracket of two vector fields vanishes, then we say that the vectors commute. It is clear then that the vectors comprising a coordinate induced basis commute since partial derivatives commute. It turns out that the converse is also true, i.e., if all the elements of a basis for vector fields commute, then the basis is coordinate induced. (See Q. 6 on Assignment 1)

5.4 Lie Derivative of a Function

Given a congruence, we can define a derivation over real smooth functions by taking the difference of the function evaluated at two points along an integral curve divided by the parameter distance between them, and taking the limit as the distance vanishes. This is called the **Lie derivative** of a function. It is clearly dependent on the vector field whose integral curves the function is evaluated along. In particular, if X is a vector field and f is a function, then the Lie derivative of f with respect to X is

$$(\mathcal{L}_X f)_p = \lim_{dt \to 0} \left[\frac{f(h_{dt}(p)) - f(p)}{dt} \right]. \tag{5.29}$$

If the Lie derivative of f with respect to X vanishes, then we say that f is 'Lie dragged' along integral curves of X.

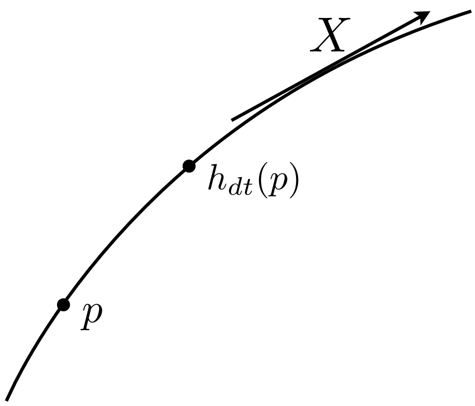

Figure 5.5:

We can recast the Lie derivative of a function in another form by introducing a chart. A particularly useful coordinate chart for computing Lie derivatives is obtained by taking one of the coordinate lines, y^1 say, to be the integral curve of the vector field. In this chart the vector field is simply $\partial/\partial y^1$. We claim that such a chart always exists in a neighbourhood of p.

Claim 5.4.1. Suppose X is a smooth vector field and $X_p \neq 0$. The there exists a coordinate chart $\{y^\mu\}$ defined in a neighbourhood U of p, such that $X = \partial/\partial y^1$ in U.

Proof. We proof our claim by explicitly constructing the chart. By continuity, there exists a neighbourhood U of p such that $X \neq 0$ in U. Choose a hypersurface (an $(n-1)$-dimensional submanifold) Σ in U nowhere tangent to X and arbitrary coordinates $(y^2, y^3, ..., y^n)$ on Σ. There exists a unique integral curve through each point of Σ, parameterized by t with $t = 0$ at Σ. Define y^1 to be the parameter distance t along such curves and $(y^2, y^3, ..., y^n)$ to be constant on each curve. Then $\{y^\mu\}$ is the required chart. (see Fig. 5.6) ∎

Now returning to the definition of the Lie derivative of a function, in this adapted chart $\{y^\mu\}$, we have

$$h_{dt}(p) = (y^1(p) + dt, y^2(p), y^3(p), ..., y^n(p)) \tag{5.30}$$

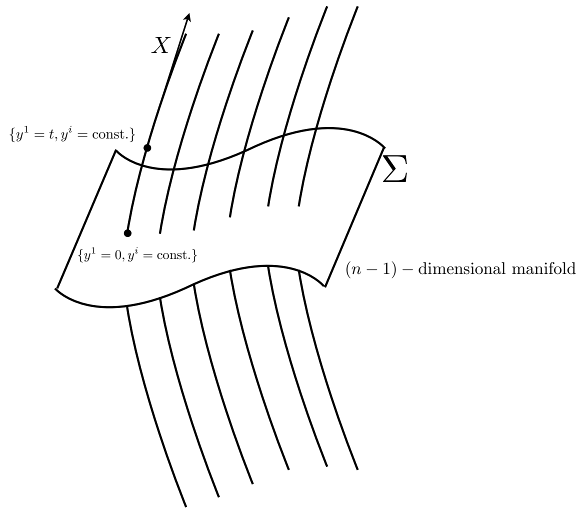

Figure 5.6:

and hence the right-hand side of Eq.(5.29) reduces to the partial derivative

$$\left(\frac{\partial F}{\partial y^1}\right)_{\phi(p)} = X_p(f). \tag{5.31}$$

Hence we have shown that, in a particular chart, the Lie derivative of a function with respect to a vector field is equal to the vector field acting on a function. However, $X_p(f)$ is chart-invariant and so this is true in all charts,

$$\boxed{(\mathcal{L}_X f)_p = X_p(f).} \tag{5.32}$$

5.5 Lie Derivative of a Vector Field

Given two smooth vector fields, X and Y, we wish to define the Lie derivative of Y with respect to X which gives the rate of change of the vector field Y following the integral curves of X. Given the definition of the Lie derivative of a function, it is very tempting to write

$$\lim_{dt\to 0}\left[\frac{Y^\mu(h_{dt}(p)) - Y^\mu(p)}{dt}\right], \tag{5.33}$$

but this is **not well-defined** since vectors at different points live in different tangent spaces. In general, there is no unique or natural way to transport tangent vectors at one point of a manifold to tangent vectors at another so that they can be compared. In fact defining a notion of parallel transport is an additional structure on the manifold called a connection, which is the subject of the next chapter.

Notwithstanding this fact, when the points at which the tangent vectors are defined are on an integral curve of a smooth vector field, then we can use the push-forward $(h_{dt})_* : T_p(\mathcal{M}) \to$

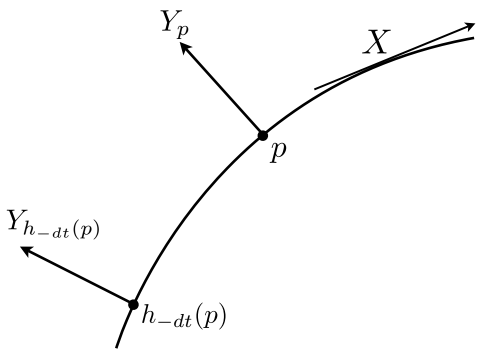

Figure 5.7:

$T_{h_{dt}(p)}(\mathcal{M})$ which maps vectors at p to vectors at $h_{dt}(p)$ which is a parameter distance dt from p along the integral curve of X.

Hence if Y_p is a vector at p, then so is the vector $(h_{dt})_* Y_{h_{-dt}(p)}$ (see Fig. 5.7), and hence we define

$$(\mathcal{L}_X Y)_p = \lim_{dt \to 0} \left[\frac{Y_p - (h_{dt})_* Y_{h_{-dt}(p)}}{dt} \right]. \tag{5.34}$$

In the adapted chart introduced in the previous section, only the coordinate y^1 changes along the integral curves of X, so that the components of the push-forward of $Y_{h_{dt}(p)}$ are the set of functions whose value at p is just the functions Y_p^μ evaluated at $h_{-dt}(p) = (y^1 - dt, y^2(p), ..., y^n(p))$,

$$\left((h_{dt})_* Y_{h_{dt}(p)}\right)^\mu = Y_p^\mu(y^1(p) - dt, y^2(p), ..., y^n(p)). \tag{5.35}$$

Hence the right-hand side of Eq. (5.34) is simply

$$\frac{\partial Y^\mu}{\partial y^1} = Y^\mu{}_{,\nu} X^\nu \tag{5.36}$$

where we have used the fact that $X^\nu = \delta^\nu{}_1$ in this chart. Now this expression does not transform like a vector and hence is not the appropriate expression for the Lie derivative in any chart. But we note that

$$X^\mu{}_{,\nu} Y^\nu = 0 \tag{5.37}$$

in this chart and hence we may subtract to get

$$\begin{aligned} (\mathcal{L}_X Y)_p^\mu &= Y^\mu{}_{,\nu} X^\nu - X^\mu{}_{,\nu} Y^\nu \\ &= [X, Y]_p^\mu. \end{aligned} \tag{5.38}$$

So we have shown that, in this particular chart, the components of the Lie derivative of a vector field Y with respect to a vector field X are the same as the components of the Lie bracket of the two vector fields. However, if two tensors have equal components in one chart, they are equal in all charts, and hence the following equality is chart-invariant:

$$\boxed{(\mathcal{L}_X Y)_p = [X, Y]_p}. \tag{5.39}$$

5.6 Lie Derivatives of Covectors and Tensors

Once we have defined the Lie derivative of a function and of a vector field, we can obtain the Lie derivative of covectors and tensors uniquely by requiring the Lie derivative to satisfy the Leibniz rule. In particular, let us obtain the components of the Lie derivative of a covector field ω with respect to the vector field X. Let us introduce an arbitrary smooth vector field Y. Then $\omega(Y)$ is a scalar function whose Lie derivative with respect to X is simply $X(\omega(Y))$. But requiring the Lie derivative to satisfy the Leibniz rule implies

$$\mathcal{L}_X(\omega(Y)) = (\mathcal{L}_X\omega)(Y) + \omega\, \mathcal{L}_X(Y), \tag{5.40}$$

and re-arranging gives

$$(\mathcal{L}_X\omega)(Y) = \mathcal{L}_X(\omega(Y)) - \omega\, \mathcal{L}_X(Y). \tag{5.41}$$

The left-hand side here is just the contraction $(\mathcal{L}_X(\omega))_\mu Y^\mu$ where $(\mathcal{L}_X\omega)_\mu$ and Y^μ are components in a coordinate induced basis. Hence

$$\begin{aligned}
(\mathcal{L}_X(\omega))_\mu Y^\mu &= \mathcal{L}_X(\omega(Y)) - \omega\, \mathcal{L}_X(Y) \\
&= X(\omega(Y)) - \omega([X, Y]) \\
&= X^\mu \partial_\mu(\omega_\nu Y^\nu) - \omega_\nu(Y^\nu{}_{,\mu} X^\mu - X^\nu{}_{,\mu} Y^\mu) \\
&= X^\mu \omega_{\nu,\mu} Y^\nu + X^\mu \omega_\nu Y^\nu{}_{,\mu} - \omega_\nu Y^\nu{}_{,\mu} X^\mu + \omega_\nu X^\nu{}_{,\mu} Y^\mu \\
&= X^\mu \omega_{\nu,\mu} Y^\nu + \omega_\nu X^\nu{}_{,\mu} Y^\mu \\
&= (X^\nu \omega_{\mu,\nu} + \omega_\nu X^\nu{}_{,\mu}) Y^\mu.
\end{aligned} \tag{5.42}$$

Since this is true for arbitrary Y, we have

$$\boxed{(\mathcal{L}_X\omega)_\mu = X^\nu \omega_{\mu,\nu} + \omega_\nu X^\nu{}_{,\mu}.} \tag{5.43}$$

One can then continue to obtain the components of Lie derivatives of all higher rank tensors uniquely simply by enforcing the Leibniz rule. The next highest rank tensor would be a $\binom{1}{1}$ tensor, whose Lie derivative we derive in the following example:

Example 5.6.1. Compute the components of the Lie derivative of a $\binom{1}{1}$ tensor with respect to X in a coordinate induced basis.

Solution: We introduce an arbitrary covector field ω and vector field Y. Then $T(\omega;Y)$ is a scalar and the Leibniz rule implies

$$\mathcal{L}_X(T(\omega;Y)) = (\mathcal{L}_X T)(\omega;Y) + T(\mathcal{L}_X\omega;Y) + T(\omega;\mathcal{L}_X Y) \tag{5.44}$$

and re-arranging gives

$$(\mathcal{L}_X T)(\omega;Y) = \mathcal{L}_X(T(\omega;Y)) - T(\mathcal{L}_X\omega;Y) - T(\omega;\mathcal{L}_X Y). \tag{5.45}$$

Hence, we have

$$\begin{aligned}(\mathcal{L}_X T)^{\mu}{}_{\nu}\omega_\mu Y^\nu &= X(T(\omega;Y)) - T(\mathcal{L}_X\omega;Y) - T(\omega;[X,Y]) \\ &= X^\mu\partial_\mu(T^{\lambda}{}_{\nu}\omega_\lambda Y^\nu) - T^{\mu}{}_{\nu}(\omega_{\mu,\lambda}X^\lambda + \omega_\lambda X^{\lambda}{}_{,\mu})Y^\nu \\ &\quad - T^{\mu}{}_{\nu}\omega_\mu(Y^{\nu}{}_{,\lambda}X^\lambda - X^{\nu}{}_{,\lambda}Y^\lambda) \\ &= X^\mu(T^{\lambda}{}_{\nu,\mu}\omega_\lambda Y^\nu + T^{\lambda}{}_{\nu}\omega_{\lambda,\mu}Y^\nu + T^{\lambda}{}_{\nu}\omega_\lambda Y^{\nu}{}_{,\mu}) \\ &\quad - T^{\mu}{}_{\nu}\omega_{\mu,\lambda}X^\lambda Y^\nu - T^{\mu}{}_{\nu}\omega_\lambda X^{\lambda}{}_{,\mu}Y^\nu \\ &\quad - T^{\mu}{}_{\nu}\omega_\mu Y^{\nu}{}_{,\lambda}X^\lambda + T^{\mu}{}_{\nu}\omega_\mu X^{\nu}{}_{,\lambda}Y^\lambda. \end{aligned} \tag{5.46}$$

Now after rearranging indices, the second and fourth terms cancel, as do the third and sixth terms, yielding

$$\begin{aligned}(\mathcal{L}_X T)^{\mu}{}_{\nu}\omega_\mu Y^\nu &= T^{\lambda}{}_{\nu,\mu}X^\mu\omega_\lambda Y^\nu - T^{\mu}{}_{\nu}\omega_\lambda X^{\lambda}{}_{,\mu}Y^\nu + T^{\mu}{}_{\nu}\omega_\mu X^{\nu}{}_{,\lambda}Y^\lambda \\ &= (T^{\mu}{}_{\nu,\lambda}X^\lambda - T^{\lambda}{}_{\nu}X^{\mu}{}_{,\lambda} + T^{\mu}{}_{\lambda}X^{\lambda}{}_{,\nu})\omega_\mu Y^\nu, \end{aligned} \tag{5.47}$$

where again we have rearranged the indices to factor out the $\omega_\mu Y^\nu$. Since ω and Y were arbitrary, we must have

$$\boxed{(\mathcal{L}_X T)^{\mu}{}_{\nu} = T^{\mu}{}_{\nu,\lambda}X^\lambda + T^{\mu}{}_{\lambda}X^{\lambda}{}_{,\nu} - T^{\lambda}{}_{\nu}X^{\mu}{}_{,\lambda}} \tag{5.48}$$

∎

In general, the Lie derivative of an arbitrary $\binom{k}{l}$ tensor field with respect to the vector field X is

$$\begin{aligned}(\mathcal{L}_X T)^{\mu_1\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\nu_l} &= X^\lambda T^{\mu_1\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\nu_l,\lambda} - X^{\mu_1}{}_{,\lambda}T^{\lambda\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\nu_l} - \cdots - X^{\mu_k}{}_{,\lambda}T^{\mu_1\mu_2\ldots\lambda}{}_{\nu_1\nu_2\ldots\nu_l} \\ &\quad + X^{\lambda}{}_{,\nu_1}T^{\mu_1\mu_2\ldots\mu_k}{}_{\lambda\nu_2\ldots\nu_l} + \cdots + X^{\lambda}{}_{,\nu_l}T^{\mu_1\mu_2\ldots\mu_k}{}_{\nu_1\nu_2\ldots\lambda}, \end{aligned} \tag{5.49}$$

where we pick up a negative sign every time the sum is over an upper index of the tensor and a plus sign whenever the sum is over a lower index of the tensor.

Chapter 6

Linear (Affine) Connections and Covariant Differentiation

6.1 Linear Connections

A well-defined notion of differentiation of a vector field on a curved manifold requires a comparison of two vectors living in different tangent spaces. This issue is circumvented in $\mathbb{R}^n$ since there is a natural isomorphism between each tangent space, i.e., we can trivially parallel transport one of the vectors so that they are defined at the same point, i.e., they live in the same tangent space. However, there is no natural way to define parallel transport in curved manifolds.

To see this, consider transporting a tangent vector, X, in $\mathbb{S}^2$ as follows: Assuming our vector is both tangent to the sphere and also perpendicular to the equator as in Fig. 6.1, we start from a point on the equator, A say, and move through an angle ϕ to the point B by keeping our vector tangent and also by not allowing any rotations so that it also remains perpendicular to the equator. Hence the vector is parallel transported in this intuitive sense, i.e., it remains fixed with respect to the curve along which it is being carried. Now 'parallel transport' from B along the line of longitude to the northpole N which gives a tangent vector which we call X'. Now parallel transporting in the opposite direction around ADN, we obtain a tangent vector X''. It is clear from the figure that the two vectors at N are not identical.

The problem arises since there is no global well-defined notion of parallel transport and in the above example, we have effectively defined a rule for parallel transport by requiring the vector to remain fixed with respect to the curve along which it is transported. A connection, which we denote by ∇, is a **rule** for parallel transport. In the language of fibre bundles, a linear (affine) connection is a map that 'connects' or identifies different fibres in the tangent bundle.

Given a smooth curve $\gamma(s)$ whose tangent vector at each point on the curve is the vector field $X = \dot{\gamma}(s)$ and a connection ∇, we can pick an arbitrary vector at some point p along this curve,

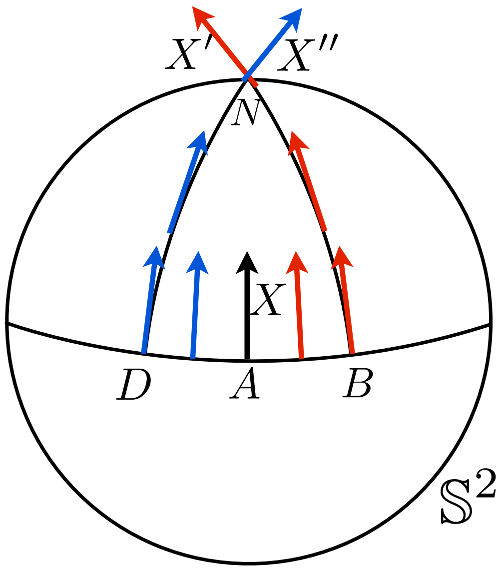

Figure 6.1:

Y_p say, and define a vector field Y along γ by parallel transporting Y_p along the curve. Hence, we can associate with a connection a derivation, ∇_X, known as the **covariant derivative** along X for which Y has zero rate of change

$$\nabla_X Y = 0 \qquad \Longleftrightarrow \qquad \text{Y is parallel transported along } X. \tag{6.1}$$

Then if Z is a vector field along γ, we can define its covariant derivative with respect to X at a point $p = \gamma(0)$ along the curve by parallel transporting the vector Z at $\gamma(ds)$ along X to the vector Z at p and taking the appropriate limit as the parameter distance between them vanishes, much in the same way we did for Lie derivatives.

Though the above exposition gives a geometric intuition to the concept of a connection, it is perhaps clearest to define a linear connection axiomatically as a map which sends a pair of smooth vector fields to a smooth vector field:

$$\nabla : X, Y \to \nabla_X Y \tag{6.2}$$

which is linear in the sense that it is required to satisfy

$$\begin{aligned} \nabla_X(Y+Z) &= \nabla_X Y + \nabla_X Z \\ \nabla_{f\,X+Y} Z &= f \nabla_X Z + \nabla_Y Z \qquad \forall f : \mathcal{M} \to \mathbb{R}. \end{aligned} \tag{6.3}$$

Since the connection also defines a derivation, we have

$$\nabla_X f = X(f) \tag{6.4}$$

and the Leibniz rule

$$\nabla_X (f\,Y) = f\,\nabla_X Y + X(f)\,Y. \tag{6.5}$$

It is clear from the Leibniz rule above that ∇ is NOT a tensor since it is evidently not linear in Y. However, for fixed Y, we can define a map

$$\nabla Y : T_p(\mathcal{M}) \to T_p(\mathcal{M}), \qquad X \to \nabla_X Y, \tag{6.6}$$

which maps a smooth vector field to a smooth vector field. By Eqs. (6.3), it is linear in its argument and hence ∇Y is a $\binom{1}{1}$ tensor known as the covariant derivative of Y.

6.2 The Covariant Derivative of a Vector Field

In an arbitrary basis $\{e_\mu\}$, we have a map ∇e_ν taking a basis vector e_μ to a vector field $\nabla_{e_\mu} e_\nu \equiv \nabla_\mu e_\nu$. Since this is a vector field for each e_μ, we may write $\nabla_\mu e_\nu$ as a linear combination of basis vectors

$$\nabla_\mu e_\nu = \Gamma^\lambda_{\nu\mu} e_\lambda, \tag{6.7}$$

where $\Gamma^\lambda_{\nu\mu}$ are known as the *connection coefficients* (note the order of the indices above) and are not the components of a tensor. Given two smooth vector fields $X = X^\mu e_\mu$ and $Y = Y^\mu e_\mu$, we decompose in the basis $\{e_\mu\}$ as follows:

$$\nabla_X Y = \nabla_{X^\nu e_\nu} Y = X^\nu \nabla_\nu Y = X^\nu (\nabla_\nu Y)^\mu e_\mu = X^\nu Y^\mu{}_{;\nu} e_\mu. \tag{6.8}$$

We could also apply the Leibniz rule to $Y = Y^\mu e_\mu$,

$$\begin{aligned}
\nabla_X Y &= \nabla_{X^\nu e_\nu}(Y^\mu e_\mu) = X^\nu \nabla_\nu (Y^\mu e_\mu) \\
&= X^\nu e_\nu(Y^\mu) e_\mu + X^\nu Y^\mu \nabla_\nu e_\mu \\
&= X^\nu e_\nu(Y^\mu) e_\mu + X^\nu Y^\mu \Gamma^\lambda_{\mu\nu} e_\lambda \\
&= (e_\nu(Y^\mu) + Y^\lambda \Gamma^\mu_{\lambda\nu}) X^\nu e_\mu.
\end{aligned} \tag{6.9}$$

Comparing Eq. (6.8) with Eq. (6.9) we obtain the components of the covariant derivative of Y:

$$Y^\mu{}_{;\nu} = e_\nu(Y^\mu) + \Gamma^\mu_{\lambda\nu} Y^\lambda. \tag{6.10}$$

Finally, in a coordinate basis, this yields

$$\boxed{Y^\mu{}_{;\nu} = Y^\mu{}_{,\nu} + \Gamma^\mu_{\lambda\nu} Y^\lambda.} \tag{6.11}$$

From this expression, we can see that the connection measures the failure of the partial derivative to be a well-defined tensor. To be more explicit, the partial derivative transforms as

$$\begin{aligned}
Y^\mu{}_{,\nu} &= \frac{\partial Y^\mu}{\partial x^\nu} \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial}{\partial x^{\nu'}} \left(\frac{\partial x^\mu}{\partial x^{\mu'}} Y^{\mu'} \right) \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial Y^{\mu'}}{\partial x^{\nu'}} + \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial^2 x^\mu}{\partial x^{\nu'} \partial x^{\mu'}} Y^{\mu'}.
\end{aligned} \tag{6.12}$$

The second term here means that the components do not transform like a $\binom{1}{1}$ tensor. If we now look at how the connection transforms under a change of basis:

$$\begin{aligned}
\nabla_{e_\nu} e_\lambda = \Gamma^\mu_{\lambda\nu} e_\mu &= \Gamma^\mu_{\lambda\nu} e_{\mu'} \Lambda^{\mu'}{}_\mu \\
&= \nabla_{\Lambda^{\nu'}{}_\nu e_{\nu'}} (\Lambda^{\lambda'}{}_\lambda e_{\lambda'}) \\
&= \Lambda^{\nu'}{}_\nu \nabla_{\nu'} (\Lambda^{\lambda'}{}_\lambda e_{\lambda'}) \\
&= \Lambda^{\nu'}{}_\nu e_{\nu'}(\Lambda^{\lambda'}{}_\lambda) e_{\lambda'} + \Lambda^{\nu'}{}_\nu \Lambda^{\lambda'}{}_\lambda \Gamma^{\mu'}_{\lambda'\nu'} e_{\mu'} \\
&= \Lambda^{\nu'}{}_\nu e_{\nu'}(\Lambda^{\mu'}{}_\lambda) e_{\mu'} + \Lambda^{\nu'}{}_\nu \Lambda^{\lambda'}{}_\lambda \Gamma^{\mu'}_{\lambda'\nu'} e_{\mu'},
\end{aligned} \tag{6.13}$$

where we changed indices ($\lambda' \to \mu'$) in the first term. Comparing with the first line above, we have

$$\Gamma^\mu_{\lambda\nu} \Lambda^{\mu'}{}_\mu = \Lambda^{\nu'}{}_\nu e_{\nu'}(\Lambda^{\mu'}{}_\lambda) + \Lambda^{\nu'}{}_\nu \Lambda^{\lambda'}{}_\lambda \Gamma^{\mu'}_{\lambda'\nu'}. \tag{6.14}$$

Multiplying by $\Lambda^\sigma{}_{\mu'}$ and using the fact that

$$\Lambda^\sigma{}_{\mu'} \Lambda^{\mu'}{}_\mu = \delta^\sigma{}_\mu \qquad \text{or} \qquad \Lambda^{\sigma'}{}_\lambda \Lambda^\lambda{}_{\mu'} = \delta^{\sigma'}{}_{\mu'} \tag{6.15}$$

gives

$$\Gamma^\sigma_{\lambda\nu} = \Lambda^{\nu'}{}_\nu \Lambda^\sigma{}_{\mu'} e_{\nu'}(\Lambda^{\mu'}{}_\lambda) + \Lambda^{\nu'}{}_\nu \Lambda^\sigma{}_{\mu'} \Lambda^{\lambda'}{}_\lambda \Gamma^{\mu'}_{\lambda'\nu'}. \tag{6.16}$$

Since vector fields are derivations, their action on the components of the Kronecker delta vanishes since all components are 0 or 1. Hence, letting $e_{\nu'}$ act on Eq. (6.15) implies

$$\Lambda^\sigma{}_{\mu'} e_{\nu'}(\Lambda^{\mu'}{}_\lambda) = -\Lambda^{\mu'}{}_\lambda e_{\nu'}(\Lambda^\sigma{}_{\mu'}), \tag{6.17}$$

so we have that the connection coefficients transform as

$$\Gamma^\sigma_{\lambda\nu} = -\Lambda^{\nu'}{}_\nu \Lambda^{\mu'}{}_\lambda e_{\nu'}(\Lambda^\sigma{}_{\mu'}) + \Lambda^{\nu'}{}_\nu \Lambda^\sigma{}_{\mu'} \Lambda^{\lambda'}{}_\lambda \Gamma^{\mu'}_{\lambda'\nu'}. \tag{6.18}$$

Moreover, since Y^λ are the components of a vector, we know they transform like $Y^\lambda = \Lambda^\lambda{}_{\sigma'} Y^{\sigma'}$ and therefore

$$\Gamma^\sigma_{\lambda\nu} Y^\lambda = -\Lambda^{\nu'}{}_\nu e_{\nu'}(\Lambda^\sigma{}_{\sigma'}) Y^{\sigma'} + \Lambda^{\nu'}{}_\nu \Lambda^\sigma{}_{\mu'} \Gamma^{\mu'}_{\sigma'\nu'} Y^{\sigma'} \tag{6.19}$$

where we have again used Eq. (6.15). Then in a coordinate basis, we have

$$\Gamma^\sigma_{\lambda\nu} Y^\lambda = -\frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial^2 x^\sigma}{\partial x^{\nu'} \partial x^{\sigma'}} Y^{\sigma'} + \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\sigma}{\partial x^{\mu'}} \Gamma^{\mu'}_{\sigma'\nu'} Y^{\sigma'}. \tag{6.20}$$

Finally, from Eq. (6.20) and Eq. (6.12), we have that the covariant derivative transforms like

$$\begin{aligned}
Y^\mu{}_{;\nu} &= Y^\mu{}_{,\nu} + \Gamma^\mu_{\lambda\nu} Y^\lambda \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial Y^{\mu'}}{\partial x^{\nu'}} + \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial^2 x^\mu}{\partial x^{\nu'} \partial x^{\mu'}} Y^{\mu'} - \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial^2 x^\mu}{\partial x^{\nu'} \partial x^{\mu'}} Y^{\mu'} + \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \Gamma^{\mu'}_{\lambda'\nu'} Y^{\lambda'} \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} \left(Y^{\mu'}{}_{,\nu'} + \Gamma^{\mu'}_{\lambda'\nu'} Y^{\lambda'} \right) \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} Y^{\mu'}{}_{;\nu'},
\end{aligned} \tag{6.21}$$

which is precisely the transformation rule for a $\binom{1}{1}$ tensor in a coordinate basis. Although neither the partial derivative operator or the connection are tensors, their combination that appears in the covariant derivative is a well-defined tensor and is the appropriate generalization on a manifold of the partial derivative in $\mathbb{R}^n$. In fact, it is only for Cartesian coordinates in $\mathbb{R}^n$ that the components of the partial derivative transform as a tensor, if one works in curvilinear coordinates such as polar coordinates, then the connection coefficients are non-zero and the covariant derivative is the appropriate tensorial directional derivative.

We have explicitly shown that the connection is not a tensor by examining how the connection coefficients transform under a change of basis. Notwithstanding this fact, the difference of two connections is a tensor.

Claim 6.2.1. Let ∇ and $\tilde{\nabla}$ be two connections on $\mathcal{M}$, then their difference is a tensor.

Proof. We define the map D to be

$$D(X,Y) = \nabla_X Y - \tilde{\nabla}_X Y \tag{6.22}$$

which is the difference between the connections which maps two smooth vector fields to a smooth vector field. Rather than introducing a basis and showing that the components satisfy an appropriate tensor transformation rule, a more elegant approach is to simply check multilinearity. Since we already know that $\nabla_X Y$ is linear in X, and that $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$, we need only check linearity in the second argument with respect to multiplication by scalar functions.

$$\begin{aligned} D(X, fY) &= \nabla_X(fY) - \tilde{\nabla}_X(fY) \\ &= (\nabla_X f)(Y) + f(\nabla_X Y) - (\tilde{\nabla}_X f)(Y) - f(\tilde{\nabla}_X Y) \\ &= X(f)\,Y + f\,(\nabla_X Y) - X(f)\,Y - f\,(\tilde{\nabla}_X Y) \\ &= f(\nabla_X Y - \tilde{\nabla}_X Y) \\ &= f\,D(X,Y). \end{aligned} \tag{6.23}$$

∎

6.3 The Covariant Derivative of Covectors and Tensors

Armed with the covariant derivative of a vector field and a scalar, we can generalize to arbitrary tensors by insisting that the covariant derivative satisfy the Leibniz rule, analogous to how we obtained the components of the Lie derivative of tensors.

Let us explicitly obtain the components of the covariant derivative of a covector field ω. We start by introducing an arbitrary vector field, Y say, and we know that $\omega(Y)$ is a scalar whose covariant derivative with respect to a basis vector e_ν is simply the action of the vector on the scalar. But we also have the Leibniz rule which states that

$$\begin{aligned}\nabla_\nu(\omega(Y)) &= (\nabla_\nu\omega)(Y) + \omega(\nabla_\nu Y)\\ &= (\nabla_\nu\omega)_\mu Y^\mu + \omega_\mu(\nabla_\nu Y)^\mu\\ &= \omega_{\mu;\nu} Y^\mu + \omega_\mu Y^\mu{}_{;\nu}.\end{aligned} \tag{6.24}$$

Rearranging gives

$$\begin{aligned}\omega_{\mu;\nu}Y^\mu &= \nabla_\nu(\omega_\mu Y^\mu) - \omega_\mu Y^\mu{}_{;\nu}\\ &= e_\nu(\omega_\mu Y^\mu) - \omega_\mu Y^\mu{}_{;\nu}.\end{aligned} \tag{6.25}$$

If the basis is coordinate induced then $e_\nu = \partial_\nu$ and we have

$$\begin{aligned}\omega_{\mu;\nu}Y^\mu &= \omega_{\mu,\nu}Y^\mu + \omega_\mu Y^\mu{}_{,\nu} - \omega_\mu(Y^\mu{}_{,\nu} + \Gamma^\mu_{\lambda\nu}Y^\lambda)\\ &= \omega_{\mu,\nu}Y^\mu - \omega_\mu\Gamma^\mu_{\lambda\nu}Y^\lambda\\ &= (\omega_{\mu,\nu} - \omega_\lambda\Gamma^\lambda_{\mu\nu})Y^\mu,\end{aligned} \tag{6.26}$$

where we have exchanged indices to arrive at the last line. Now since Y was arbitrary, we must have

$$\boxed{\omega_{\mu;\nu} = \omega_{\mu,\nu} - \omega_\lambda\Gamma^\lambda_{\mu\nu}} \tag{6.27}$$

Repeating this procedure, we can obtain the components of the covariant derivative of arbitrary tensors, once we know the components for all lower rank tensors. In the following example, we compute the components of the covariant derivative of a $\binom{1}{1}$ tensor using our knowledge of the components of a vector and covector, and insisting that the covariant derivative satisfy the Leibniz rule.

Example 6.3.1. Compute the components of the covariant derivative of a $\binom{1}{1}$ tensor.

Solution: Let T be a $\binom{1}{1}$ tensor field and ω, Y arbitrary covector and vector fields, respectively. Then $T(\omega;Y)$ is a scalar and the Leibniz rule gives

$$\begin{aligned}&\nabla_\lambda(T(\omega;Y)) = \nabla_\lambda T(\omega;Y) + T(\nabla_\lambda\omega;Y) + T(\omega;\nabla_\lambda Y)\\ \implies\quad &\nabla_\lambda T(\omega;Y) = \nabla_\lambda(T(\omega;Y)) - T(\nabla_\lambda\omega;Y) - T(\omega;\nabla_\lambda Y).\end{aligned} \tag{6.28}$$

Introducing a coordinate basis gives

$$\begin{aligned}(\nabla_\lambda T)^\mu{}_\nu\omega_\mu Y^\nu &\equiv T^\mu{}_{\nu;\lambda}\omega_\mu Y^\nu\\ &= \partial_\lambda(T^\mu{}_\nu\omega_\mu Y^\nu) - T(\omega_{\mu;\lambda}dx^\mu;Y^\nu\partial_\nu) - T(\omega_\mu dx^\mu;Y^\nu{}_{;\lambda}\partial_\nu)\\ &= T^\mu{}_{\nu,\lambda}\omega_\mu Y^\nu + T^\mu{}_\nu\omega_{\mu,\lambda}Y^\nu + T^\mu{}_\nu\omega_\mu Y^\nu{}_{,\lambda}\\ &\quad - T^\mu{}_\nu(\omega_{\mu,\lambda} - \Gamma^\rho_{\mu\lambda}\omega_\rho)Y^\nu - T^\mu{}_\nu\omega_\mu(Y^\nu{}_{,\lambda} + \Gamma^\nu_{\rho\lambda}Y^\rho).\end{aligned} \tag{6.29}$$

Now, all terms involving derivatives of ω and Y cancel and we are left with (after re-labelling some dummy indices)

$$T^{\mu}{}_{\nu;\lambda}\omega_{\mu}Y^{\nu} = (T^{\mu}{}_{\nu,\lambda} + T^{\rho}{}_{\nu}\Gamma^{\mu}_{\rho\lambda} - T^{\mu}{}_{\rho}\Gamma^{\rho}_{\nu\lambda})\omega_{\mu}Y^{\nu}, \tag{6.30}$$

and since ω, Y were arbitrary, we have

$$\boxed{T^{\mu}{}_{\nu;\lambda} = T^{\mu}{}_{\nu,\lambda} + T^{\rho}{}_{\nu}\Gamma^{\mu}_{\rho\lambda} - T^{\mu}{}_{\rho}\Gamma^{\rho}_{\nu\lambda}} \quad \blacksquare \tag{6.31}$$

In general then we have

$$\begin{aligned}\nabla_{\lambda}T^{\mu_1\mu_2\ldots}{}_{\nu_1\nu_2\ldots} = T^{\mu_1\mu_2\ldots}{}_{\nu_1\nu_2\ldots,\lambda} + T^{\rho\mu_1\ldots}{}_{\nu_1\nu_2\ldots}\Gamma^{\mu_1}_{\rho\lambda} + T^{\mu_1\rho\ldots}{}_{\nu_1\nu_2\ldots}\Gamma^{\mu_2}_{\rho\lambda} + \cdots \\ - T^{\mu_1\mu_2\ldots}{}_{\rho\nu_2\ldots}\Gamma^{\rho}_{\nu_1\lambda} - T^{\mu_1\mu_2\ldots}{}_{\nu_1\rho\ldots}\Gamma^{\rho}_{\nu_2\lambda} - \cdots\end{aligned} \tag{6.32}$$

where every connection term where the sum is over the lower index picks up a plus sign whereas every connection term where the sum is over its upper index picks up a minus sign. (Or equivalently, where the sum is over the upper index of the tensor components, we pick up a plus whereas a sum over the lower index of the tensor components picks up a minus).

6.4 Geodesics

Since the concept of covariant differentiation naturally extends from vector fields to tensor fields, so to does the concept of parallel transport on which the notion of differentiation relies. In particular, we say that a tensor T is parallel transported along a curve whose tangent vector is X (or parallel transported with respect to X) if

$$\nabla_X T = 0. \tag{6.33}$$

If X is the tangent to the curve $\gamma(\tau)$ whose coordinate image in $\mathbb{R}^n$ is $\phi \circ \gamma = \{x^{\mu}(\tau)\}$, then $X^{\mu} = dx^{\mu}/d\tau$ and the parallel transport equation reads

$$\frac{dx^{\mu}}{d\tau}\nabla_{\mu}T \equiv \frac{DT}{d\tau} = 0. \tag{6.34}$$

In particular, the components of the parallel transport equation for a vector field Z is

$$\frac{dx^{\mu}}{d\tau}Z^{\nu}{}_{;\mu} = \frac{dx^{\mu}}{d\tau}\left(Z^{\nu}{}_{,\mu} + \Gamma^{\nu}_{\rho\mu}Z^{\rho}\right) = 0, \tag{6.35}$$

which is most succinctly written as

$$\boxed{\frac{dZ^{\nu}}{d\tau} + \Gamma^{\nu}_{\rho\mu}\frac{dx^{\mu}}{d\tau}Z^{\rho} = 0.} \tag{6.36}$$

Turning our attention now to *geodesic curves*, which intuitively we think of as curves representing the shortest distance between two points of a manifold, e.g., geodesics in Euclidean space are simply straight lines while geodesics on the two-sphere are so-called great circles. There is an alternative definition of geodesic curves which in the absence of a precise notion of distance afforded only by a metric structure, is the one we must adopt here. The definition is as follows: *A geodesic is a curve that parallel transports its own tangent vector*. It is not completely obvious that the two definitions are equivalent but we shall return to this in the next chapter when we endow the manifold with an additional metric structure.

Now if the curve is geodesic, then the tangent vector components $dx^\mu/d\tau$ must satisfy Eq. (6.36), which yields

$$\boxed{\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0.} \tag{6.37}$$

It is worth noting that we also have the initial conditions $x^\mu(0) = p$ and $\dot{x}^\mu(0) = X^\mu_p$, so that locally there is a unique solution to the geodesic equation. This, of course, is not true globally since, for example, in $\mathbb{S}^2$, there are many geodesics (great circles) connecting two points on opposite sides of the sphere. The neighbourhood p for which p is connected by a unique geodesic to each point in the neighbourhood is known as the *normal neighbourhood* of p.

Finally, we note that the parametrization of a curve has always been considered as part of the definition of the curve itself, so different parametrizations correspond to different curves (and hence different tangent vectors at points along the curve) even if they represent the same set of points. A natural question then to ask then is if there is any re-parametrization freedom for geodesic curves. To address this, suppose we re-parametrize from τ to $s(\tau)$, then the geodesic equation becomes

$$\left(\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{ds}\frac{dx^\lambda}{ds}\right)\left(\frac{ds}{d\tau}\right)^2 + \left(\frac{d^2s}{d\tau^2}\right)\left(\frac{dx^\mu}{ds}\right) = 0. \tag{6.38}$$

Hence we retrieve the standard form of the geodesic equation if and only if

$$\begin{aligned} & \frac{ds}{d\tau} \neq 0 \qquad \text{and} \qquad \frac{d^2s}{d\tau^2} = 0 \\ \implies \quad & s = a\,\tau + b \qquad a, b \text{ constants and } a \neq 0. \end{aligned} \tag{6.39}$$

Therefore, all parameters that leave the geodesic equation in this standard form are linearly related, such parameters are called **affine parameters**.

6.5 Normal Coordinates

Given a point $p \in \mathcal{M}$ and a chart $\{x^\mu\}$, we may find a new chart $\{\hat{x}^\mu\}$ such that $\hat{\Gamma}^\mu_{(\nu\lambda)}(p) = 0$, i.e., the geodesics at p satisfy $\ddot{\hat{x}} = 0$ so that they are locally linear functions of τ. These coordinates are valid in the normal neighbourhood of p.

To prove that such a chart exists, let us explicitly construct it. We take $x^\mu(p) = 0$ and set

$$\hat{x}^\mu = x^\mu + \tfrac{1}{2} Q^\mu_{\nu\lambda} x^\nu x^\lambda, \tag{6.40}$$

where $Q^\mu_{\nu\lambda} = Q^\mu_{(\nu\lambda)}$ are constants. Let

$$|x|^2 = |x^1|^2 + |x^2|^2 + \cdots + |x^n|^2 \tag{6.41}$$

and hence we have

$$\hat{x}^\mu = x^\mu + \mathcal{O}(|x|^2) \qquad \text{as} \quad |x| \to 0. \tag{6.42}$$

We may invert Eq. (6.40) by solving perturbatively

$$x^\mu = \hat{x}^\mu - \tfrac{1}{2} Q^\mu_{\nu\lambda} x^\nu x^\lambda \tag{6.43}$$

which yields

$$\begin{aligned} x^\mu &= \hat{x}^\mu - \tfrac{1}{2} Q^\mu_{\nu\lambda} (\hat{x}^\nu - \tfrac{1}{2} Q^\nu_{\rho\sigma} x^\rho x^\sigma)(\hat{x}^\lambda - \tfrac{1}{2} Q^\lambda_{\alpha\beta} x^\alpha x^\beta) \\ &= \hat{x}^\mu - \tfrac{1}{2} Q^\mu_{\nu\lambda} \hat{x}^\nu \hat{x}^\lambda + \mathcal{O}(|x|^3). \end{aligned} \tag{6.44}$$

Therefore, differentiating Eq. (6.40) gives

$$\begin{aligned} \frac{\partial \hat{x}^\mu}{\partial x^\rho} &= \delta^\mu{}_\rho + \tfrac{1}{2} Q^\mu_{\nu\lambda} \delta^\nu{}_\rho x^\lambda + \tfrac{1}{2} Q^\mu_{\nu\lambda} x^\nu \delta^\lambda{}_\rho \\ &= \delta^\mu{}_\rho + Q^\mu_{\rho\lambda} x^\lambda, \end{aligned} \tag{6.45}$$

while differentiating Eq. (6.44) gives

$$\frac{\partial x^\mu}{\partial \hat{x}^\rho} = \delta^\mu{}_\rho - Q^\mu_{\rho\lambda} \hat{x}^\lambda + \mathcal{O}(|x|^2). \tag{6.46}$$

Now, recall that we have shown how the connection coefficients transform under a change of coordinates, so from Eq. (6.16), we have

$$\hat{\Gamma}^\mu_{\nu\lambda} = \frac{\partial \hat{x}^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial \hat{x}^\lambda} \frac{\partial^2 x^\rho}{\partial x^\sigma \partial \hat{x}^\nu} + \frac{\partial \hat{x}^\mu}{\partial x^\rho} \frac{\partial x^\sigma}{\partial \hat{x}^\lambda} \frac{\partial x^\gamma}{\partial \hat{x}^\nu} \Gamma^\rho_{\gamma\sigma}. \tag{6.47}$$

It is clear from Eqs. (6.45)-(6.46) that at the point p, we have

$$\left.\frac{\partial \hat{x}^\mu}{\partial x^\rho}\right|_p = \delta^\mu{}_\rho, \qquad \left.\frac{\partial x^\mu}{\partial \hat{x}^\rho}\right|_p = \delta^\mu{}_\rho, \qquad \left.\frac{\partial^2 x^\mu}{\partial x^\nu \partial \hat{x}^\rho}\right|_p = -Q^\mu_{\rho\nu}, \tag{6.48}$$

and hence from the transformation rule, we obtain

$$\begin{aligned} \hat{\Gamma}^\mu_{\nu\lambda}(p) &= -\delta^\mu{}_\rho \delta^\sigma{}_\lambda Q^\rho_{\nu\sigma} + \delta^\mu{}_\rho \delta^\sigma{}_\lambda \delta^\gamma{}_\nu \Gamma^\rho_{\gamma\sigma} \\ &= -Q^\mu_{\nu\lambda} + \Gamma^\mu_{\nu\lambda}(p). \end{aligned} \tag{6.49}$$

Since $Q^\mu_{\nu\lambda}$ are an arbitrary set of constants, we choose

$$Q^\mu_{\nu\lambda} = \Gamma^\mu_{(\nu\lambda)}(p) \tag{6.50}$$

which yields

$$\hat{\Gamma}^\mu_{(\nu\lambda)}(p) = 0, \tag{6.51}$$

as required.

6.6 Torsion

The torsion tensor is a $\binom{1}{2}$ tensor field defined by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{6.52}$$

It is straight-forward to verify that this is indeed linear in both arguments. To obtain the components of the torsion tensor, it is useful to expand the commutator vector field as a linear combination of basis vectors

$$[e_\nu, e_\lambda] = \gamma^\mu_{\nu\lambda} e_\mu \tag{6.53}$$

where $\gamma^\mu_{\nu\lambda}$ are the so-called commutator coefficients and are NOT the components of a tensor. Now

$$\begin{aligned} T(e_\nu, e_\lambda) &= T^\mu{}_{\nu\lambda} e_\mu \\ &= \nabla_\nu e_\lambda - \nabla_\lambda e_\nu - [e_\nu, e_\lambda] \\ &= \Gamma^\mu_{\lambda\nu} e_\mu - \Gamma^\mu_{\nu\lambda} e_\mu - \gamma^\mu_{\nu\lambda} e_\mu. \end{aligned} \tag{6.54}$$

Hence we write the components of the torsion tensor in terms of commutator coefficients and the anti-symmetric part of the connection:

$$\begin{aligned} T^\mu{}_{\nu\lambda} &= \Gamma^\mu_{\lambda\nu} - \Gamma^\mu_{\nu\lambda} - \gamma^\mu_{\nu\lambda} \\ &= -2\Gamma^\mu_{[\nu\lambda]} - \gamma^\mu_{\nu\lambda}. \end{aligned} \tag{6.55}$$

Claim 6.6.1. If the connection in a coordinate basis is symmetric, then the torsion vanishes.

Proof. Introducing a coordinate basis $\{\partial_\mu\}$, we have

$$[\partial_\mu, \partial_\nu] = 0 \qquad \Longrightarrow \qquad \gamma^\mu_{\nu\lambda} = 0. \tag{6.56}$$

Moreover, if the connection is symmetric in this coordinate basis, then $\Gamma^\mu_{[\nu\lambda]} = 0$, and hence from the definition of torsion above, we have

$$T^\mu{}_{\nu\lambda} = 0. \tag{6.57}$$

∎

One might be tempted to conclude that the torsion is always zero since although we introduced a coordinate chart to arrive at the result, the result itself is a tensorial equation which ought to be valid in all charts. However, we have also assumed that the connection is symmetric, which is a chart-dependent statement.

It's straight-forward to see that the converse is also true, i.e., if the torsion is zero and we work in a coordinate basis, then the connection is symmetric in these coordinates.

Chapter 7

Curvature

7.1 Riemann Curvature Tensor

The Riemann curvature tensor is a $\binom{1}{3}$ tensor field defined by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \tag{7.1}$$

where X, Y, Z are smooth vector fields.

To see that this is a tensor, we must check linearity in each of its arguments. It is clear from the definition that $R(X,Y)Z = -R(Y,X)Z$ so we need only check linearity in the first and third arguments. We shall show that it is linear only in the first argument, leaving the proof of linearity in the third argument to the reader. Consider

$$\begin{aligned}
R(f\,X + W, Y)Z &= \nabla_{(f\,X+W)} \nabla_Y Z - \nabla_Y \nabla_{(f\,X+W)} Z - \nabla_{[f\,X+W,Y]} Z \\
&= f\,\nabla_X \nabla_Y Z + \nabla_W \nabla_Y Z - \nabla_Y (f\,\nabla_X Z + \nabla_W Z) - \nabla_{(f\,[X,Y]+[W,Y]-Y(f)X)} Z \\
&= f\,\nabla_X \nabla_Y Z + \nabla_W \nabla_Y Z - f\,\nabla_Y \nabla_X Z - Y(f)\nabla_X Z - \nabla_Y \nabla_W Z \\
&\qquad\qquad\qquad - f\,\nabla_{[X,Y]} Z - \nabla_{[W,Y]} Z + Y(f)\,\nabla_X Z \\
&= f\,(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) + (\nabla_W \nabla_Y Z - \nabla_Y \nabla_W Z - \nabla_{[W,Y]} Z) \\
&= f\,R(X,Y)Z + R(W,Y)Z \qquad \text{as required.}
\end{aligned} \tag{7.2}$$

Similarly, we find that

$$R(X,Y)(fZ + W) = f\,R(X,Y)Z + R(X,Y)W, \tag{7.3}$$

and hence R is a tensor.

Let us obtain an expression for the components of the Riemann curvature tensor. In an arbitrary basis $\{e_\mu\}$, we have

$$\begin{aligned}
R(e_\nu, e_\lambda)e_\rho &= R^\mu{}_{\rho\nu\lambda} e_\mu && (7.4)\\
&= \nabla_\nu \nabla_\lambda e_\rho - \nabla_\lambda \nabla_\nu e_\rho - \nabla_{[e_\nu, e_\lambda]} e_\rho && (7.5)\\
&= \nabla_\nu(\Gamma^\mu_{\rho\lambda} e_\mu) - \nabla_\lambda(\Gamma^\mu_{\rho\nu} e_\mu) - \nabla_{\gamma^\mu_{\nu\lambda} e_\mu}(e_\rho) && (7.6)\\
&= e_\nu(\Gamma^\mu_{\rho\lambda})e_\mu + \Gamma^\mu_{\rho\lambda}\Gamma^\sigma_{\mu\nu} e_\sigma - e_\lambda(\Gamma^\mu_{\rho\nu})e_\mu - \Gamma^\mu_{\rho\nu}\Gamma^\sigma_{\mu\lambda} e_\sigma - \gamma^\mu_{\nu\lambda}\Gamma^\sigma_{\rho\mu} e_\sigma && (7.7)\\
&= e_\nu(\Gamma^\mu_{\rho\lambda})e_\mu - e_\lambda(\Gamma^\mu_{\rho\nu})e_\mu + \Gamma^\sigma_{\rho\lambda}\Gamma^\mu_{\sigma\nu} e_\mu - \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda} e_\mu - \gamma^\sigma_{\nu\lambda}\Gamma^\mu_{\rho\sigma} e_\mu && (7.8)
\end{aligned}$$

Since this is true for all e_μ, we have

$$R^\mu{}_{\rho\nu\lambda} = e_\nu(\Gamma^\mu_{\rho\lambda}) - e_\lambda(\Gamma^\mu_{\rho\nu}) + \Gamma^\sigma_{\rho\lambda}\Gamma^\mu_{\sigma\nu} - \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda} - \gamma^\sigma_{\nu\lambda}\Gamma^\mu_{\rho\sigma}. \tag{7.9}$$

In a coordinate induced basis, $e_\mu = \partial_\mu$ and all the commutator coefficients $\gamma^\mu_{\nu\lambda}$ vanish since partial derivatives commute. Hence

$$\boxed{R^\mu{}_{\rho\nu\lambda} = \Gamma^\mu_{\rho\lambda,\nu} - \Gamma^\mu_{\rho\nu,\lambda} + \Gamma^\sigma_{\rho\lambda}\Gamma^\mu_{\sigma\nu} - \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda}.} \tag{7.10}$$

Claim 7.1.1. The curvature tensor measures the lack of commutation of covariant differentiation according to the **Ricci Identity**

$$X^\mu{}_{;\nu\lambda} - X^\mu{}_{;\lambda\nu} = -R^\mu{}_{\rho\nu\lambda} X^\rho. \tag{7.11}$$

Proof. In an arbitrary basis, we consider the second order covariant derivative of a vector field X,

$$\begin{aligned}
\nabla_\lambda \nabla_\nu X &= (\nabla_\lambda \nabla_\nu X)^\mu e_\mu \equiv X^\mu{}_{;\nu\lambda} e_\mu\\
&= \nabla_\lambda \nabla_\nu (X^\mu e_\mu)\\
&= \nabla_\lambda (e_\nu(X^\mu)\, e_\mu + X^\mu \nabla_\nu e_\mu)\\
&= \nabla_\lambda (e_\nu(X^\mu)\, e_\mu + X^\mu \Gamma^\sigma_{\mu\nu} e_\sigma)\\
&= e_\lambda(e_\nu(X^\mu))\, e_\mu + e_\nu(X^\mu)\nabla_\lambda e_\mu + e_\lambda(X^\mu)\Gamma^\sigma_{\mu\nu} e_\sigma + X^\mu e_\lambda(\Gamma^\sigma_{\mu\nu}) e_\sigma + X^\mu \Gamma^\sigma_{\mu\nu} \nabla_\lambda e_\sigma\\
&= e_\lambda(e_\nu(X^\mu))\, e_\mu + e_\nu(X^\mu)\Gamma^\sigma_{\mu\lambda} e_\sigma + e_\lambda(X^\mu)\Gamma^\sigma_{\mu\nu} e_\sigma + X^\mu e_\lambda(\Gamma^\sigma_{\mu\nu}) e_\sigma + X^\mu \Gamma^\sigma_{\mu\nu}\Gamma^\rho_{\sigma\lambda} e_\rho\\
&= (e_\lambda(e_\nu(X^\mu)) + e_\nu(X^\sigma)\Gamma^\mu_{\sigma\lambda} + e_\lambda(X^\sigma)\Gamma^\mu_{\sigma\nu} + X^\sigma e_\lambda(\Gamma^\mu_{\sigma\nu}) + X^\rho \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda}) e_\mu && (7.12)
\end{aligned}$$

which is to be compared with the top line. Then in a coordinate basis we have

$$X^\mu{}_{;\nu\lambda} = X^\mu{}_{,\nu\lambda} + X^\sigma{}_{,\nu}\Gamma^\mu_{\sigma\lambda} + X^\sigma{}_{,\lambda}\Gamma^\mu_{\sigma\nu} + X^\sigma \Gamma^\mu_{\sigma\nu,\lambda} + X^\rho \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda}. \tag{7.13}$$

Covariantly differentiating in the opposite order gives

$$X^\mu{}_{;\lambda\nu} = X^\mu{}_{,\lambda\nu} + X^\sigma{}_{,\lambda}\Gamma^\mu_{\sigma\nu} + X^\sigma{}_{,\nu}\Gamma^\mu_{\sigma\lambda} + X^\sigma \Gamma^\mu_{\sigma\lambda,\nu} + X^\rho \Gamma^\sigma_{\rho\lambda}\Gamma^\mu_{\sigma\nu}, \tag{7.14}$$

and hence subtracting yields (after some re-labelling of dummy indices)

$$\begin{aligned}
X^\mu{}_{;\nu\lambda} - X^\mu{}_{;\lambda\nu} &= (\Gamma^\mu_{\rho\nu,\lambda} + \Gamma^\sigma_{\rho\nu}\Gamma^\mu_{\sigma\lambda} - \Gamma^\mu_{\rho\lambda,\nu} - \Gamma^\sigma_{\rho\lambda}\Gamma^\mu_{\sigma\nu}) X^\rho\\
&= R^\mu{}_{\rho\lambda\nu} X^\rho\\
&= -R^\mu{}_{\rho\nu\lambda} X^\rho. && (7.15)
\end{aligned}$$

∎

7.2 Geometric Interpretation of Riemann Tensor

We let X and Y be two vector fields on a manifold $\mathcal{M}$ whose integral curves are parametrized by s and t, respectively. We assume that the Lie Bracket of these vector fields vanishes, i.e., $[X, Y] = 0$, so that these vector fields span a closed infinitesimal quadrilateral.

Starting from a point $p \in \mathcal{M}$, we parallel transport a vector Z a parameter distance ds along the integral curve of X to a point r, followed by parallel transport a parameter distance dt along the integral curve of Y reaching a point u, followed again by parallel transport a parameter distance ds along the integral curve X to a point q and finally returning to the point p by parallel transporting a parameter distance dt along the integral curve of Y. This is represented schematically in the figure below.

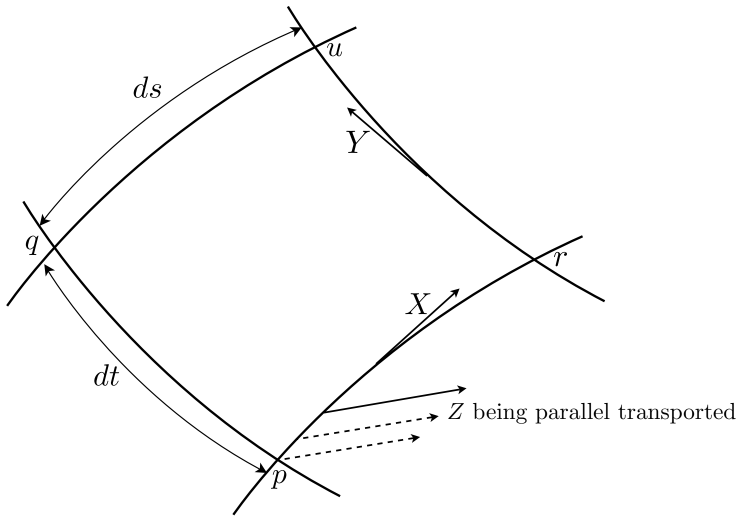

Figure 7.1:

We wish to compute the change in Z, ΔZ^μ, as it is parallel transported about the closed infinitesimal quadrilateral *pruqp*.

We work in normal coordinates constructed at the point p such that the connection coefficients vanish at the point p, i.e.

$$\Gamma^\mu_{\nu\lambda}(p) = 0. \tag{7.16}$$

The vector Z is parallel transported along pr and hence satisfies the parallel transport equation $\nabla_X Z = 0$, which in component form is

$$\frac{dZ^\mu}{ds} + \Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda = 0 \tag{7.17}$$

$$\implies \quad \left(\frac{dZ^\mu}{ds}\right)_p = -\Gamma^\mu_{\nu\lambda}(p) Z^\nu_p X^\lambda_p = 0 \qquad \text{(in normal coords).} \tag{7.18}$$

Differentiating both sides of the parallel transport equation gives

$$\frac{d^2 Z^\mu}{ds^2} = -\frac{d}{ds}(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda) \tag{7.19}$$

$$= -(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda)_{,\rho} X^\rho \qquad (\text{using } dx^\rho/ds = X^\rho) \tag{7.20}$$

$$= -\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho - \Gamma^\mu_{\nu\lambda}(Z^\nu X^\lambda)_{,\rho} X^\rho, \tag{7.21}$$

and hence in normal coordinates this reduces to

$$\left(\frac{d^2 Z^\mu}{ds^2}\right)_p = -(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p. \tag{7.22}$$

The components Z^μ are really the coordinate image in $\mathbb{R}^n$ of the components of the vector field Z in the chosen chart, and hence we may Taylor expand the components at the point r about the components at the point p using standard calculus in $\mathbb{R}^n$, yielding

$$Z^\mu_r = Z^\mu_p + \left(\frac{dZ^\mu}{ds}\right)_p ds + \tfrac{1}{2}\left(\frac{d^2 Z^\mu}{ds^2}\right)_p ds^2 + \mathcal{O}(ds^3) \tag{7.23}$$

$$= Z^\mu_p - \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p ds^2 + \cdots \tag{7.24}$$

where we have used the corresponding expressions from above for these derivatives in normal coordinates. Now along ru, Z is parallel propagated along integral curves of Y, and hence $\nabla_Y Z = 0$. As before, the parallel transport equation leads to the following expressions for the first and second derivative of Z^μ with respect to t which parametrizes the integral curves of Y,

$$\frac{dZ^\mu}{dt} = -\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda \tag{7.25}$$

$$\frac{d^2 Z^\mu}{dt^2} = -(\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)_{,\rho} Y^\rho. \tag{7.26}$$

Again, we can Taylor expand the components Z^μ at u about the components at r

$$Z^\mu_u = Z^\mu_r + \left(\frac{dZ^\mu}{dt}\right)_r dt + \tfrac{1}{2}\left(\frac{d^2 Z^\mu}{dt^2}\right)_r dt^2 + \cdots \tag{7.27}$$

$$= Z^\mu_r - (\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)_r dt - \tfrac{1}{2}((\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)_{,\rho} Y^\rho)_r dt^2 + \cdots \tag{7.28}$$

and further expanding these expressions about p and using the fact that the connection coefficients vanish at p in normal coordinates, we obtain

$$Z^\mu_u = Z^\mu_p - \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p ds^2 \tag{7.29}$$

$$- \left[(\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)_p + \left(\frac{d}{ds}(\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)\right)_p ds + \cdots\right] dt \tag{7.30}$$

$$- \tfrac{1}{2}\left[((\Gamma^\mu_{\nu\lambda} Z^\nu Y^\lambda)_{,\rho} Y^\rho)_p + \cdots\right] dt^2 + \cdots \tag{7.31}$$

$$= Z^\mu_p - \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p ds^2 - (\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda Y^\rho)_p ds\, dt \tag{7.32}$$

$$- \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda Y^\rho)_p dt^2 + \cdots \tag{7.33}$$

Interchanging X with Y and s with t in this expression gives the expression we would obtain by parallel propagating Z in the opposite direction pqu:

$$Z^\mu_u = Z^\mu_p - \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda Y^\rho)_p dt^2 - (\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda X^\rho)_p ds\, dt \tag{7.34}$$
$$- \tfrac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p ds^2 + \cdots \tag{7.35}$$

The difference in these expressions represents the change in Z as it is parallel propagated around the closed infinitesimal loop $pruqp$ and is given by

$$\Delta Z^\mu = -(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu)_p (Y^\lambda X^\rho - X^\lambda Y^\rho)_p ds\, dt + \cdots \tag{7.36}$$
$$= [(\Gamma^\mu_{\nu\lambda,\rho} X^\lambda Y^\rho Z^\nu)_p - (\Gamma^\mu_{\nu\rho,\lambda} X^\lambda Y^\rho Z^\nu)_p] ds\, dt + \cdots \tag{7.37}$$
$$= (\Gamma^\mu_{\nu\lambda,\rho} - \Gamma^\mu_{\nu\rho,\lambda})_p X^\lambda_p Y^\rho_p Z^\nu_p \, ds\, dt + \cdots \tag{7.38}$$
$$= (R^\mu{}_{\nu\rho\lambda} X^\lambda Y^\rho Z^\nu)_p ds\, dt + \cdots \qquad \text{(since } \Gamma\Gamma \text{ terms vanish in normal coords.)} \tag{7.39}$$

Hence, we have

$$R^\mu{}_{\nu\rho\lambda} X^\lambda Y^\rho Z^\nu = \lim_{ds,dt\to 0} \left(\frac{\Delta Z^\mu}{ds\, dt} \right), \tag{7.40}$$

which is a tensor equation valid in all charts. Therefore, we interpret the quantity $R(X, Y)Z$ as a measure of the change in Z after parallel transporting around the closed quadrilateral spanned by the vector fields X and Y.

7.3 Geodesic Deviation

We let γ_1 and γ_2 be two neighbouring integral curves of a vector field X which are parametrized by t. We let $\{\sigma_t\}$ be curves parametrized by s intersecting γ_1 and γ_2 at the parameter value t along these integral curves. Let Z be the tangent vector to σ_t such that $s = 0$ on γ_1 and $s = 1$ on γ_2.

It is clear from Fig. 7.2 that the quadrilateral spanned by the integral curves of X and the curves $\{\sigma_t\}$ form a closed infinitesimal quadrilateral and hence $[X, Z] = 0$. The tangent vector Z is a vector which points form one integral curve to another and hence is known as a *connecting vector* to the congruence of curves.

Claim 7.3.1. Assuming the torsion vanishes, let X be tangent to a congruence of geodesics and let Z be a connecting vector, then the acceleration of Z along the congruence, which is defined by $\nabla_X \nabla_X Z$, is given by

$$\nabla_X \nabla_X Z = R(X, Z)X. \tag{7.41}$$

This is the geodesic deviation equation since it gives a measure of the relative acceleration between neighbouring geodesics.

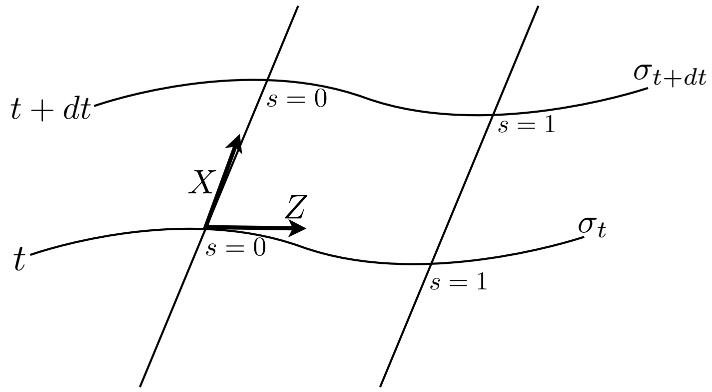

Figure 7.2:

Proof. Since the torsion vanishes, we have

$$T(X, Z) = \nabla_X Z - \nabla_Z X - [X, Z] = 0. \tag{7.42}$$

However, since $[X, Z] = 0$ by construction, it follows that

$$\nabla_X Z = \nabla_Z X. \tag{7.43}$$

Then the acceleration of the connecting vector along the curve is given by

$$\begin{aligned} \nabla_X \nabla_X Z &= \nabla_X \nabla_Z X \\ &= R(X, Z)X + \nabla_Z \nabla_X X + \nabla_{[X,Z]} X. \end{aligned} \tag{7.44}$$

Since X is tangent to a geodesic, we have $\nabla_X X = 0$ and by cpnstruction the commutator vanishes and hence we obtain

$$\nabla_X \nabla_X Z = R(X, Z)X. \tag{7.45}$$

∎

In component form the geodesic deviation equation reads

$$\frac{D^2 Z^\mu}{dt^2} \equiv (Z^\mu{}_{;\nu} X^\nu)_{;\lambda} X^\lambda = R^\mu{}_{\nu\lambda\rho} X^\nu X^\lambda Z^\rho. \tag{7.46}$$

Therefore, we see that the relative acceleration between two neighbouring geodesics is proportional to that part of the curvature tensor that is symmetric in the middle two indices $R^\mu{}_{(\nu\lambda)\rho}$, since the contraction over $X^\nu X^\lambda$ picks out this symmetric part of the Riemann tensor.

7.4 Symmetries of the Riemann Tensor

From Eq. (7.10), it is easy to see that the components are anti-symmetric in the last two indices,

$$R^{\mu}{}_{\nu\lambda\rho} = -R^{\mu}{}_{\nu\rho\lambda}$$

$$\Longleftrightarrow \quad \boxed{R^{\mu}{}_{\nu(\lambda\rho)} = 0.} \tag{7.47}$$

Furthermore, if the torsion vanishes, then in a coordinate basis we have a symmetric connection, in which case the Riemann tensor satisfies

$$R^{\mu}{}_{\nu\lambda\rho} + R^{\mu}{}_{\rho\nu\lambda} + R^{\mu}{}_{\lambda\rho\nu} = 0. \tag{7.48}$$

In order to prove this, we introduce normal coordinates at a point p, where the connection coefficients (but not its derivatives) vanish. Then

$$\begin{aligned} R^{\mu}{}_{\nu\lambda\rho} &= \Gamma^{\mu}_{\nu\rho,\lambda} - \Gamma^{\mu}_{\nu\lambda,\rho}, \\ R^{\mu}{}_{\rho\nu\lambda} &= \Gamma^{\mu}_{\rho\lambda,\nu} - \Gamma^{\mu}_{\rho\nu,\lambda}, \\ R^{\mu}{}_{\lambda\rho\nu} &= \Gamma^{\mu}_{\lambda\nu,\rho} - \Gamma^{\mu}_{\lambda\rho,\nu}, \end{aligned} \tag{7.49}$$

and hence summing and using the fact that the connection is symmetric gives

$$R^{\mu}{}_{\nu\lambda\rho} + R^{\mu}{}_{\rho\nu\lambda} + R^{\mu}{}_{\lambda\rho\nu} = 0. \tag{7.50}$$

Now although this has been derived by introducing a particular chart (i.e. normal coordinates), the resultant equation is a tensor equation and hence (7.50) holds in all charts. This equation is equivalent to the vanishing of the anti-symmetric part of the lower three indices,

$$\boxed{R^{\mu}{}_{[\nu\lambda\rho]} = 0.} \tag{7.51}$$

The covariant derivative of the Riemann tensor $R^{\mu}{}_{\nu\lambda\rho;\sigma}$ is a $\binom{1}{4}$ tensor which satisfies the symmetry

$$R^{\mu}{}_{\nu[\lambda\rho;\sigma]} = 0. \tag{7.52}$$

These are the so-called **Bianchi Identities**.

To prove this, we first note that the Bianchi Identities are equivalent to

$$R^{\mu}{}_{\nu\lambda\rho;\sigma} - R^{\mu}{}_{\nu\lambda\sigma;\rho} + R^{\mu}{}_{\nu\sigma\lambda;\rho} - R^{\mu}{}_{\nu\sigma\rho;\lambda} + R^{\mu}{}_{\nu\rho\sigma;\lambda} - R^{\mu}{}_{\nu\rho\lambda;\sigma} = 0, \tag{7.53}$$

and using the symmetry (7.47), this is equivalent to

$$R^{\mu}{}_{\nu\lambda\rho;\sigma} + R^{\mu}{}_{\nu\sigma\lambda;\rho} + R^{\mu}{}_{\nu\rho\sigma;\lambda} = 0. \tag{7.54}$$

Now adopting normal coordinates and assuming zero torsion (and hence a symmetric connection), then schematically we have

$$\begin{aligned} & R = \partial\Gamma - \partial\Gamma + \Gamma\,\Gamma - \Gamma\,\Gamma \\ \implies \quad & \partial R = \partial\partial\Gamma - \partial\partial\Gamma + (\partial\Gamma)\Gamma + \Gamma(\partial\Gamma) - (\partial\Gamma)\Gamma - \Gamma(\partial\Gamma) \\ & \quad = \partial\partial\Gamma - \partial\partial\Gamma \qquad \text{in normal coordinates.} \end{aligned} \tag{7.55}$$

Restoring indices, we have

$$R^{\mu}{}_{\nu\lambda\rho;\sigma} = \Gamma^{\mu}_{\nu\rho,\lambda\sigma} - \Gamma^{\mu}_{\nu\lambda,\rho\sigma}. \tag{7.56}$$

Similarly,

$$\begin{aligned} R^{\mu}{}_{\nu\sigma\lambda;\rho} &= \Gamma^{\mu}_{\nu\lambda,\sigma\rho} - \Gamma^{\mu}_{\nu\sigma,\lambda\rho}, \\ R^{\mu}{}_{\nu\rho\sigma;\lambda} &= \Gamma^{\mu}_{\nu\sigma,\rho\lambda} - \Gamma^{\mu}_{\nu\rho,\sigma\lambda}. \end{aligned} \tag{7.57}$$

Hence summing and using the fact that partial derivatives commute yields

$$R^{\mu}{}_{\nu\lambda\rho;\sigma} + R^{\mu}{}_{\nu\sigma\lambda;\rho} + R^{\mu}{}_{\nu\rho\sigma;\lambda} = 0, \tag{7.58}$$

which is equivalent to the more succinct expression

$$\boxed{R^{\mu}{}_{\nu[\lambda\rho;\sigma]} = 0.} \tag{7.59}$$

Chapter 8

The Metric

Thus far, we have given a broad description of the local differential structure of a smooth manifold. This description, for the most part, was in terms of intrinsic properties of the manifold itself, with the exception of the covariant derivative which requires an additional connection structure. Moreover, none of our concepts have relied on the notion of the length of a vector or angles between vectors. These can only be defined by endowing the manifold with an additional metric structure which is analogous to the dot product of ordinary vector calculus. A metric structure subsumes a connection structure and we shall see that given a metric tensor, there is a unique metric preserving connection.

8.1 The Metric Tensor

A metric tensor, usually denoted by g, is a $\binom{0}{2}$ non-degenerate, symmetric tensor such that:

1. The magnitude of the vector X is $|g(X,X)|^{1/2}$.
2. the angle between two vectors X, Y is

$$\cos^{-1}\left(\frac{g(X,Y)}{|g(X,X)|^{1/2}|g(Y,Y)|^{1/2}}\right), \qquad \text{for } g(X,X) \neq 0, \quad g(Y,Y) \neq 0. \tag{8.1}$$

   If $g(X,Y) = 0$, the X and Y are orthogonal vectors.

3. The length of a curve whose tangent vector is X between t_1 and t_2 is

$$\int_{t_1}^{t_2} |g(X,X)|^{1/2} dt. \tag{8.2}$$

In a particular basis, the components, $\{e_\mu\}$ say, the components of the metric tensor are denoted by

$$g_{\mu\nu} = g(e_\mu, e_\nu) = g(e_\nu, e_\mu) = g_{\nu\mu}. \tag{8.3}$$

The fact that g is non-degenerate implies that if

$$g(X, Y) = 0 \quad \forall Y, \quad \text{then} X = 0. \tag{8.4}$$

Therefore g has an inverse whose components we denote by

$$g^{\mu\nu} = (g^{-1})^{\mu\nu}. \tag{8.5}$$

These are the components of a $\binom{2}{0}$ tensor satisfying

$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu{}_\lambda. \tag{8.6}$$

The metric and its inverse define a natural isomorphism between the tangent space and its dual, the cotangent space, as follows:

$$T_p(\mathcal{M}) \leftrightarrow T_p^*(\mathcal{M}), \qquad \begin{aligned} X^\mu &\to g_{\mu\nu} X^\nu \equiv X_\mu \\ \eta_\mu &\to g^{\mu\nu} \eta_\nu \equiv \eta^\mu, \end{aligned} \tag{8.7}$$

i.e., the metric and inverse metric components are used to **lower and raise indices** of vectors and covectors, respectively. This naturally generalizes to arbitrary rank tensors so we can raise or lower indices by an appropriate contraction with the metric, e.g.,

$$T^{\mu\nu}{}_{\lambda\rho} = g^{\nu\sigma} T^\mu{}_{\sigma\lambda\rho}. \tag{8.8}$$

The metric gives us a way of characterizing the distance between neighbouring points of a manifold. In a particular chart $\{x^\mu\}$, the coordinate image in $\mathbb{R}^n$ of two nearby points in $\mathcal{M}$ is x^μ and $x^\mu + dx^\mu$. Then the length-squared of the interval between these points is called the line-element (or sometimes also called the metric), and is given by

$$ds^2 = g_{\mu\nu} dx^\mu \otimes dx^\mu = g_{\mu\nu} dx^\mu dx^\nu. \tag{8.9}$$

The line-element ds^2 is of course explicitly chart-invariant, but the metric components are not. For example, the line-element in three-dimensional Euclidean space $\mathbb{E}^3$ in Cartesian coordinates is

$$ds^2 = dx^2 + dy^2 + dz^2, \tag{8.10}$$

which is simply Pythagoras' Theorem. Comparing with (8.9), we read off the metric components as $g_{\mu\nu} = \text{diag}(1, 1, 1) = \delta_{\mu\nu}$ as expected for Euclidean space. If we now introduce polar coordinates $\{\hat{x}^\mu\} = \{r, \theta, \phi\}$, it is straight-forward to apply the transformation properties of the metric tensor to show that $\hat{g}_{\mu\nu} = \text{diag}(1, r^2, r^2 \sin^2\theta)$. Hence, in these coordinates, the line-element is expressed as

$$ds^2 = \hat{g}_{\mu\nu} d\hat{x}^\mu d\hat{x}^\nu = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2. \tag{8.11}$$

8.2 Riemannian and Pseudo-Riemannian Metrics

The metric components in a particular basis can be represented as a square symmetric matrix which has n eigenvalues. In general, these eigenvalues may be positive or negative (but not zero since the metric is non-degenerate) so let's say in the chosen basis there are p positive eigenvalues and $q = n - p$ negative ones. Then by *Sylvester's Law of Inertia* (also known as *Sylvester's Rigidity Theorem*), the number of positive and negative eigenvalues is independent of the basis chosen, or equivalently if the basis is coordinate induced, then the number of positive and negative eigenvalues is chart invariant.

The difference $(p-q)$ is known as the **signature** of the metric, i.e. the signature is the number of positive eigenvalues of the metric tensor minus the number of negative ones.

If $q = 0$, then all eigenvalues are positive and we say the metric is **Riemannian**. The length-squared of every non-zero vector in a Riemannian metric is positive-definitie, i.e. $g(X, X) > 0$.

If $q \neq 0$, then the absolute value of the signature is strictly less than the dimension of the manifold, and the metric is said to be **pseudo-Riemannian**.

One particularly important sub-class of pseudo-Riemannian metrics are the **Lorentzian** metrics. These are metrics with either $p = 1$ or $q = 1$, i.e. the absolute value of the signature is $n - 2$. It is the four-dimensional Lorentzian metrics that form the geometric setting for Einstein's Theory of Relativity, where the signature is often written explicitly as $(-, +, +, +)$ with the negative eigenvalue associated with the temporal direction and the positive eigenvalues with the spatial directions (conventions vary in the literature, e.g., another possibility is $(+, -, -, -)$ with the positive eigenvalue now corresponding to the temporal direction). In this case, we characterize vectors as being time-like, space-like and null depending on whether the length-squared of the vector is negative, positive or zero, respectively, i.e.

$$\begin{aligned} g(X_p, X_p) < 0 &\iff X_p \text{ is time-like,} \\ g(X_p, X_p) > 0 &\iff X_p \text{ is space-like,} \\ g(X_p, X_p) = 0 &\iff X_p \text{ is null.} \end{aligned} \tag{8.12}$$

Similarly, if X is tangent to a curve $\gamma(s)$, then we characterize $\gamma(s)$ as time-like, space-like or null depending on whether X is time-like, space-like or null.

8.3 The Levi-Civita or Metric Connection

Up until now, the connection has been defined by some rule for parallel transport, with some results depending on restricting the connection to be symmetric. In general, however, it is not uniquely defined. On the other hand, there is a unique torsion-free connection that preserves the metric ($\nabla g = 0$), known as the **Levi-Civita** or metric connection.

Theorem 8.3.1. If a manifold possesses a metric, then there is a unique, torsion-free metric satisfying

$$\nabla g = 0. \tag{8.13}$$

Proof. We prove existence and uniqueness by explicitly constructing the unique connection. We let X, Y, Z be smooth vector fields, then $g(Y,Z)$ is a scalar and

$$\begin{aligned} X(g(Y,Z)) &= \nabla_X(g(Y,Z)) \qquad \text{since } \nabla_\mu f = \partial_\mu f \text{ for } f \text{ a scalar.} \\ &= (\nabla_X g)(Y,Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \qquad \text{since } \nabla g = 0. \end{aligned} \tag{8.14}$$

Similarly, by a permutation of the vector fields, we have

$$\begin{aligned} Y(g(Z,X)) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \\ Z(g(X,Y)) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y). \end{aligned} \tag{8.15}$$

Adding these in the following combination:

$$\begin{aligned} \tfrac{1}{2}[X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y))] =& \tfrac{1}{2}g(\nabla_X Y, Z) + \tfrac{1}{2}g(Y, \nabla_X Z) \\ &+ \tfrac{1}{2}g(\nabla_Y Z, X) + \tfrac{1}{2}g(Z, \nabla_Y X) \\ &- \tfrac{1}{2}g(\nabla_Z X, Y) - \tfrac{1}{2}g(X, \nabla_Z Y) \\ =& \tfrac{1}{2}g(\nabla_X Y, Z) + \tfrac{1}{2}g(Z, \nabla_Y X) \\ &+ \tfrac{1}{2}g(Y, \nabla_X Z - \nabla_Z X) \\ &+ \tfrac{1}{2}g(X, \nabla_Y Z - \nabla_Z Y). \end{aligned} \tag{8.16}$$

Assuming vanishing torsion, we have

$$\begin{aligned} T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y] = 0 \\ \implies &\quad \nabla_X Y - \nabla_Y X = [X,Y]. \end{aligned} \tag{8.17}$$

Using this result to eliminate all the connection terms except $\nabla_X Y$ yields

$$\begin{aligned} \tfrac{1}{2}[X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y))] =& \tfrac{1}{2}g(\nabla_X Y, Z) + \tfrac{1}{2}g(Z, \nabla_X Y - [X,Y]) \\ &+ \tfrac{1}{2}g(Y, [X,Z]) + \tfrac{1}{2}g(X, [Y,Z]). \end{aligned} \tag{8.18}$$

Simplifying we obtain

$$\begin{aligned} g(\nabla_X Y, Z) = \tfrac{1}{2}[& X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) - g(Z,[X,Y]) \\ & - g(Y,[Z,X]) + g(X,[Y,Z])]. \end{aligned} \tag{8.19}$$

This expression defines the unique *Levi-Civita* connection. ∎

To obtain an expression for the components of this connection in a coordinate-induced basis, we take $X = \partial_\mu$, $Y = \partial_\nu$, $Z = \partial_\lambda$. The left-hand side of our defining equation above yields

$$g(\nabla_\mu \partial_\nu, \partial_\lambda) = g(\Gamma^\sigma_{\nu\mu}\partial_\sigma, \partial_\lambda) = \Gamma^\sigma_{\nu\mu} g_{\sigma\lambda} \tag{8.20}$$

while the right-hand side becomes

$$\tfrac{1}{2}(\partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu} - g(\partial_\lambda, [\partial_\mu, \partial_\nu]) - g(\partial_\nu, [\partial_\lambda, \partial_\mu]) + g(\partial_\mu, [\partial_\nu, \partial_\lambda]). \tag{8.21}$$

However, the Lie bracket of two coordinate-induced basis vectors vanish since partial derivatives commute and therefore, we have

$$\Gamma^\sigma_{\nu\mu} g_{\sigma\lambda} = \tfrac{1}{2}(g_{\nu\lambda,\mu} + g_{\mu\lambda,\nu} - g_{\mu\nu,\lambda}). \tag{8.22}$$

Finally, multiplying across by the metric inverse $g^{\lambda\rho}$, we arrive at

$$\boxed{\Gamma^\rho_{\nu\mu} = \Gamma^\rho_{\mu\nu} = \tfrac{1}{2} g^{\lambda\rho}(g_{\lambda\mu,\nu} + g_{\nu\lambda,\mu} - g_{\mu\nu,\lambda}).} \tag{8.23}$$

The Levi-Civita or metric connection coefficients are often referred to as the **Christoffel symbols** or Christoffel coefficients. One may also encounter the notation

$$\Gamma^\lambda_{\mu\nu} = g^{\lambda\rho}[\mu\nu, \rho] \tag{8.24}$$

where

$$[\mu\nu, \rho] \equiv \tfrac{1}{2}(g_{\rho\mu,\nu} + g_{\nu\rho,\mu} - g_{\mu\nu,\rho}). \tag{8.25}$$

Example 8.3.1. Compute the Christoffel symbols for the three-dimensional Euclidean space in Cartesian and polar coordinates.

Solution: In Cartesian coordinates, the metric is simply $g_{\mu\nu} = \mathrm{diag}(1, 1, 1)$ and since Eq. (8.23) involves derivatives of the metric, we have

$$\Gamma^\mu_{\nu\lambda} = 0. \tag{8.26}$$

Turning now to polar coordinates where we know the metric and its inverse are

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \qquad g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2}\csc^2\theta \end{pmatrix}. \tag{8.27}$$

For each fixed upper index, the Christoffel symbols can be thought of as a symmetric $n \times n$ array which has $\frac{1}{2}n(n+1)$ independent components. And since there are n choices for the upper index, we have that there are $\frac{1}{2}n^2(n+1)$ independent Christoffel symbols to be computed. In this particular case, $n = 3$ and there are 18 components.

Fixing the upper index to be r, we have

$$\Gamma^r_{\nu\lambda} = \tfrac{1}{2} g^{rr}(g_{r\nu,\lambda} + g_{\lambda r,\nu} - g_{\nu\lambda,r}). \tag{8.28}$$

Then it is clear to see that they are all zero unless the two lower indices are both θ or are both ϕ, whence

$$\begin{array}{lll} \Gamma^r_{rr} = 0 & \Gamma^r_{r\theta} = 0 & \Gamma^r_{r\phi} = 0 \\ & \Gamma^r_{\theta\theta} = -r & \Gamma^r_{\theta\phi} = 0 \\ & & \Gamma^r_{\phi\phi} = -r\sin^2\theta. \end{array} \tag{8.29}$$

Now fixing the upper index to be θ gives

$$\Gamma^\theta_{\nu\lambda} = \tfrac{1}{2} g^{\theta\theta}(g_{\theta\nu,\lambda} + g_{\lambda\theta,\nu} - g_{\nu\lambda,\theta}) \tag{8.30}$$

from which we compute the coefficients to be

$$\begin{array}{lll} \Gamma^\theta_{rr} = 0 & \Gamma^\theta_{r\theta} = 1/r & \Gamma^\theta_{r\phi} = 0 \\ & \Gamma^\theta_{\theta\theta} = 0 & \Gamma^\theta_{\theta\phi} = 0 \\ & & \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta. \end{array} \tag{8.31}$$

Finally fixing the upper index to be ϕ gives

$$\Gamma^\phi_{\nu\lambda} = \tfrac{1}{2} g^{\phi\phi}(g_{\phi\nu,\lambda} + g_{\lambda\phi,\nu} - g_{\nu\lambda,\phi}) \tag{8.32}$$

from which we compute the coefficients to be

$$\begin{array}{lll} \Gamma^\phi_{rr} = 0 & \Gamma^\phi_{r\theta} = 0 & \Gamma^\phi_{r\phi} = 1/r \\ & \Gamma^\phi_{\theta\theta} = 0 & \Gamma^\phi_{\theta\phi} = \cot\theta \\ & & \Gamma^\phi_{\phi\phi} = 0. \qquad \blacksquare \end{array} \tag{8.33}$$

8.4 Geodesics Revisited

Recall that we previously derived the geodesic equation directly from the parallel transport equation by using the definition of a geodesic as a curve that parallel transports its own tangent vector. Let us now rederive the geodesic equation using the definition of the length of a curve afforded us by the metric and show that geodesic are curves that extremize this length. (The intuitive idea that a geodesic between two points is the curve that minimizes the length is only true for space-like curves. On the other hand, time-like geodesics are curves which *maximize* the time between two points.)

If X is tangent to the curve $\gamma(s)$, then the tangent vector components are $X^\mu = dx^\mu/ds$ where $\{x^\mu(s)\}$ is the parametric representation of the curve $\gamma(s)$ in the chart $\{x^\mu\}$. Then the length of the curve between the fixed points s_1 and s_2 is

$$S = \int_{s_1}^{s_2} \sqrt{\epsilon g(X,X)}\, ds = \int_{s_1}^{s_2} \sqrt{\epsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}\, ds \tag{8.34}$$

where $\epsilon = \pm 1$ for space-like/time-like curves. The action principle $\delta S = 0$ gives rise to the Euler-Lagrange equations

$$\frac{d}{ds}\Big(\frac{\partial L}{\partial \dot{x}^\lambda}\Big) - \frac{\partial L}{\partial x^\lambda} = 0 \tag{8.35}$$

where the Lagrangian for this action is

$$L = L(x, \dot{x}) = \sqrt{\epsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}. \tag{8.36}$$

Now

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}^\lambda} &= \tfrac{1}{2}(\epsilon g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta)^{-1/2}(\epsilon g_{\mu\nu}\delta^\mu{}_\lambda \dot{x}^\nu + \epsilon g_{\mu\nu}\dot{x}^\mu \delta^\nu{}_\lambda \\ &= \tfrac{1}{2}L^{-1}(\epsilon g_{\lambda\nu}\dot{x}^\nu + \epsilon g_{\mu\lambda}\dot{x}^\mu) \\ &= L^{-1}\epsilon g_{\lambda\nu}\dot{x}^\nu. \end{aligned} \tag{8.37}$$

Hence

$$\frac{d}{ds}\Big(\frac{\partial L}{\partial \dot{x}^\lambda}\Big) = \frac{1}{L}\Big(- \frac{1}{L}\frac{dL}{ds}\epsilon g_{\lambda\nu}\dot{x}^\nu + \epsilon g_{\lambda\nu,\mu}\dot{x}^\mu \dot{x}^\nu + \epsilon g_{\lambda\nu}\ddot{x}^\nu\Big), \tag{8.38}$$

and since

$$\frac{\partial L}{\partial x^\lambda} = \frac{1}{2L}\epsilon g_{\mu\nu,\lambda}\dot{x}^\mu \dot{x}^\nu, \tag{8.39}$$

the Euler-Lagrange equations yield

$$g_{\lambda\nu}\ddot{x}^\nu + g_{\lambda\nu,\mu}\dot{x}^\mu \dot{x}^\nu - \tfrac{1}{2}g_{\mu\nu,\lambda}\dot{x}^\mu \dot{x}^\nu = g_{\lambda\nu}\dot{x}^\nu \frac{1}{L}\frac{dL}{ds}. \tag{8.40}$$

Multiplying by the metric inverse components $g^{\lambda\sigma}$ and rearranging some indices we obtain

$$\ddot{x}^\sigma + \tfrac{1}{2}g^{\lambda\sigma}(g_{\lambda\mu,\nu} + g_{\nu\lambda,\mu} - g_{\mu\nu,\lambda})\dot{x}^\mu \dot{x}^\nu = \dot{x}^\sigma \frac{1}{L}\frac{dL}{ds}. \tag{8.41}$$

We recognize the middle term here as the Christoffel symbols $\Gamma^\sigma_{\mu\nu}$ and re-writing the right-hand side gives

$$\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = \tfrac{1}{2}\dot{x}^\sigma \frac{d}{ds}(\ln(\epsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu)). \tag{8.42}$$

Hence we see that the standard from of the geodesic equation is retrieved only when the right-hand side is zero, which will occur in general if and only if

$$g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = g(X, X) = K, \qquad \text{a constant.} \tag{8.43}$$

Since parametrizations of the curve that yield the standard form of the geodesic equation are so-called affine parameters, this tells us that in the context of the metric structure, the affine parameters are those for which the length of the tangent vectors remain constant as it is parallel transported along the geodesic. So if we choose such a parametrization, then we have

$$\ddot{x}^\sigma + \Gamma^\sigma_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0. \tag{8.44}$$

There is a natural choice of affine parametrization that normalizes the magnitude of the tangent vector to be unity. If the curve is time-like then this parameter is known as the **proper time** whereas for space-like curves this parametrization is known as the **proper distance**, i.e.,

$$g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \begin{cases} -1 & \text{curve is time-like and parametrized by proper time} \\ +1 & \text{curve is space-like and parametrized by proper distance} \\ 0 & \text{curve is null.} \end{cases} \tag{8.45}$$

In relativistic physics, the proper time is the time elapsed along a trajectory as measured by an observer along this trajectory.

Example 8.4.1. Given that the Euler-Lagrange equations are the same for the Lagrangian L as they are for some smooth function of L, $f(L)$ say. Choose $f(L) = \epsilon L^2$ and rederive the geodesic equations from the new action.

Solution: The action principle $\delta S = 0$ leads to the Euler-Lagrange equations

$$\frac{d}{ds}\Big(\frac{\partial L_{\text{new}}}{\partial \dot{x}^\mu}\Big) - \frac{\partial L_{\text{new}}}{\partial x^\mu} = 0, \tag{8.46}$$

where $L_{\text{new}} = \epsilon L^2 = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ is the new Lagrangian. These equations give

$$\begin{aligned} & \frac{d}{ds}\Big(2g_{\mu\nu}\dot{x}^\nu\Big) - g_{\lambda\nu,\mu}\dot{x}^\lambda\dot{x}^\nu = 0 \\ \Longrightarrow\quad & 2g_{\mu\nu}\ddot{x}^\nu + 2\frac{dg_{\mu\nu}}{ds}\dot{x}^\nu - g_{\lambda\nu,\mu}\dot{x}^\lambda\dot{x}^\nu = 0 \\ \Longrightarrow\quad & 2g_{\mu\nu}\ddot{x}^\nu + (g_{\mu\nu,\lambda} + g_{\mu\lambda,\nu} - g_{\lambda\nu,\mu})\dot{x}^\nu\dot{x}^\lambda = 0 \\ \Longrightarrow\quad & \ddot{x}^\mu + \tfrac{1}{2}g^{\mu\rho}(g_{\rho\nu,\lambda} + g_{\rho\lambda,\nu} - g_{\lambda\nu,\rho})\dot{x}^\nu\dot{x}^\lambda = 0 \\ \Longrightarrow\quad & \ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0, \end{aligned} \tag{8.47}$$

which is the standard geodesic equation. ∎

In practice, it is far easier to derive and solve the geodesic equations directly from the Euler-Lagrange equations with the Lagrangian taken to be $L = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$. If we parametrize by proper time or distance, it is also typically easier to eliminate one of the geodesic equations in favour of the equation (8.45). This equation is sometimes referred to as the **First Integral**.

8.5 Metric Curvature

Now assuming the manifold is endowed with a metric, we can define a metric curvature tensor by lowering the upper index $R_{\mu\nu\lambda\rho} = g_{\mu\sigma}R^\sigma{}_{\nu\lambda\rho}$. If there exists a chart on $\mathcal{M}$ such that the line-element is given by

$$ds^2 = \epsilon_1(dx^1)^2 + \epsilon_2(dx^2)^2 + \cdots + \epsilon_n(dx^n)^2, \tag{8.48}$$

where $\epsilon_i = \pm 1$ for each $i = 1, .., n$, then $R_{\mu\nu\lambda\rho} = 0$ and we say the manifold is **locally flat**. The converse is also true, i.e., if $R_{\mu\nu\lambda\rho} = 0$ and there exists a chart for which the line-element takes the form (8.48).

Moreover, when the manifold is endowed with a metric, the curvature satisfies additional symmetries. We recall that the Riemann curvature tensor satisfies the following algebraic symmetries

$$\begin{aligned} R^{\mu}{}_{\nu(\lambda\rho)} &= 0 \\ R^{\mu}{}_{[\nu\lambda\rho]} &= 0. \end{aligned} \tag{8.49}$$

It is straight-forward to verify that $R_{\mu\nu\lambda\rho}$ satisfies the following symmetries:

$$\begin{aligned} R_{\mu\nu(\lambda\rho)} &= 0 \\ R_{(\mu\nu)\lambda\rho} &= 0 \\ R_{\mu[\nu\lambda\rho]} &= 0 \\ R_{\mu\nu\lambda\rho} &= R_{\lambda\rho\mu\nu}. \end{aligned} \tag{8.50}$$

The additional symmetries satisfied by $R_{\mu\nu\lambda\rho}$ reduce the number of linearly independent components. A natural question that arises is how many linearly independent components does the metric curvature tensor have in arbitrary dimensions?

Claim 8.5.1. The number of linearly independent components of $R_{\mu\nu\lambda\rho}$ is $\frac{1}{12}n^2(n^2-1)$.

Proof. The first, second and fourth of the symmetries in Eq. (8.50) imply we can think of the curvature tensor as

$$R_{[\mu\nu][\lambda\rho]},$$

a symmetric $m \times m$ matrix where the pairs of indices $\mu\nu$ and $\lambda\rho$ are treated as individual indices. Such an $m \times m$ matrix has $\frac{1}{2}m(m+1)$ independent components where each of these components are themselves $n \times n$ antisymmetric matrices with $\frac{1}{2}n(n-1)$ independent components, where n is the dimension of the manifold. So with $m = \frac{1}{2}n(n-1)$, we have

$$\tfrac{1}{2}m(m+1) = \tfrac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) \tag{8.51}$$

components. Not all of these are independent since we still haven't used the second symmetry $R_{\mu[\nu\lambda\rho]} = 0$. We note that

$$R_{[\mu\nu\lambda\rho]} = \tfrac{1}{4}(R_{\mu[\nu\lambda\rho]} + R_{\rho[\mu\nu\lambda]} + R_{\lambda[\rho\mu\nu]} + R_{\nu[\lambda\rho\mu]}) \tag{8.52}$$

and therefore $R_{\mu[\nu\lambda\rho]} = 0$ implies that $R_{[\mu\nu\lambda\rho]} = 0$. It turns out that imposing $R_{[\mu\nu\lambda\rho]} = 0$ is equivalent to imposing $R_{\mu[\nu\lambda\rho]} = 0$ once the other three symmetries have been accounted for. Since a totally anti-symmetric 4-index tensor has

$$\tfrac{1}{4!}n(n-1)(n-2)(n-3) \tag{8.53}$$

independent components, the total number of components of the curvature tensor should be reduced by this amount. Hence the number of independent components of the curvature tensor in n dimensions is

$$\tfrac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) - \tfrac{1}{4!}n(n-1)(n-2)(n-3) = \tfrac{1}{12}n^2(n^2-1). \tag{8.54}$$

$\blacksquare$

In particular, this implies that metric curvature vanishes in one dimension. In two dimensions, there is only one linearly independent component, i.e., the curvature is effectively described by a scalar. In three dimensions, we have six linearly independent components. In four dimensions such as the space-time manifold of General Relativity, there are 20 linearly independent components.

There are several other metric curvature tensors that can be defined from various combinations of the Riemann tensor and its contractions. Contracting over the first and third indices of the Riemann tensor defines the **Ricci Curvature tensor**

$$\boxed{R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu} = g^{\lambda\rho}R_{\rho\mu\lambda\nu} = R_{\nu\mu}.} \tag{8.55}$$

Contracting the Ricci tensor defines the **Ricci scalar**

$$\boxed{R = g^{\mu\nu}R_{\mu\nu} = R^{\mu}{}_{\mu}.} \tag{8.56}$$

Another particularly important tensor, especially in the context of General Relativity, is the **Einstein tensor** which is defined as

$$\boxed{G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R.} \tag{8.57}$$

Finally, we introduce the **Weyl tensor** which is defined by

$$\boxed{C_{\mu\nu\lambda\rho} = R_{\mu\nu\lambda\rho} - \tfrac{2}{n-2}(g_{\mu[\lambda}R_{\rho]\nu} - g_{\nu[\lambda}R_{\rho]\mu}) + \tfrac{2}{(n-1)(n-2)}g_{\mu[\lambda}g_{\rho]\nu}R.} \tag{8.58}$$

This is sometimes called the conformal tensor since it is invariant under conformal transformations

$$g_{\mu\nu} \to \Omega^2 g_{\mu\nu}. \tag{8.59}$$

Moreover, as well as satisfying the same symmetries as $R_{\mu\nu\lambda\rho}$, all contractions over the Weyl tensor vanish.