# General Relativity

Cian Luke Martin

2024-10-30

## Contents

# 1 Lecture: Introduction

11/10/2024

General Relativity is our best theory of gravitation on the largest scales. It is classical, geometrical and dynamical.

## 1.1 Differentiable Manifolds

The basic object of study in differential geometry is the (differentiable) manifold. This is an object which 'locally looks like $\mathbb{R}^n$', and has enough structure to let us do calculus.

**Definition 1.1:** A **differentiable manifold** of dimension $n$ is a set $M$, together with a collection of coordinate charts $(O_\alpha, \phi_\alpha)$ where

- $O_\alpha \subset M$ are subsets of $M$ such that $\cup_\alpha O_\alpha = M$,
- $\phi_\alpha$ is a bijective map (one to one and onto) from $O_\alpha \to U_\alpha$, an open subset of $\mathbb{R}^n$,
- If $O_\alpha \cap O_\beta \neq \emptyset$, then $\phi_\beta \circ \phi_\alpha^{-1}$ is a smooth (infinitely differentiable) map from $\phi_\alpha (O_\alpha \cap O_\beta) \subset U_\alpha$ to $\phi_\beta (O_\alpha \cap O_\beta) \subset U_\beta$.

**Note.** We could replace smooth with finite differentiability (*e.g.* $k$-differentiable) but it is not particularly interesting.

Further, these charts define a topology of $M$, $\mathcal{R} \subset M$ is open iff $\phi_\alpha (\mathcal{R} \cap O_\alpha)$ is open in $\mathbb{R}^n$ for all $\alpha$.

Every open subset of $M$ is itself a manifold (restrict charts to $\mathcal{R}$).

**Definition 1.2:** The collection $\{(O_\alpha, \phi_\alpha)\}$ is called an **atlas**. Two atlases are **compatible** if their union is an atlas. An atlas $A$ is **maximal** if there exists no atlas $B$ with $A \subsetneq B$.

Every atlas is contained in a maximal atlas (consider the union of all compatible atlases). We can assume without loss of generality we are working with the maximal atlas.

**Examples.**

i) If $U \subset \mathbb{R}^n$ is open, we can take $O = U$ and

$$\phi : O \to U \tag{1}$$

$$\phi \left(x^i\right) = x^i, \tag{2}$$

and $\{(U, \phi)\}$ is an atlas.

ii) $S^1 = \{\mathbf{p} \in \mathbb{R}^2 \mid |p| = 1\}$. If $\mathbf{p} \in S^1 \setminus \{(-1,0)\} = \mathcal{O}_1$, there is a unique $\theta_1 \in (-\pi, \pi)$ such that $\mathbf{p} = (\cos\theta_1, \sin\theta_1)$.

If $\mathbf{p} \in S^1 \setminus \{(1,0)\} = \mathcal{O}_2$, then there is a unique $\theta_2 \in (0, 2\pi)$ such that $\mathbf{p} = (\cos\theta_2, \sin\theta_2)$ such that

$$\phi_1 : \mathbf{p} \to \theta_1, \text{ for } \mathbf{p} \in \mathcal{O}_1, U_1 = (-\pi, \pi), \tag{3}$$

$$\phi_2 : \mathbf{p} \to \theta_2, \text{ for } \mathbf{p} \in \mathcal{O}_2, U_2 = (0, 2\pi). \tag{4}$$

We have that $\phi_1(\mathcal{O} \cap \mathcal{O}_2) = (-\pi, 0) \cup (0, \pi)$ and

$$\phi_2 \circ \phi_1^{-1}(\theta) = \begin{cases} \theta, & \theta \in (0, \pi), \\ \theta + 2\pi, & \theta \in (-\pi, 0). \end{cases} \tag{5}$$

This is smooth where defined and similarly for $\phi_1 \circ \phi_2^{-1}$ and thus $S_1$ is a 1-manifold.

iii) $S^n = \{\mathbf{p} \in \mathbb{R}^{n+1} \,\big|\, |\mathbf{p}| = 1\}$. We define charts by stereographic projection if $\{\mathbf{E}_1, \cdots, \mathbf{E}_{n+1}\}$ is a standard basis for $\mathbb{R}^{n+1}$ and $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is a standard basis for $\mathbb{R}^n$, we write

$$\mathbf{p} = p^i \mathbf{e}_i. \tag{6}$$

We set $\mathcal{O}_1 = S^n \setminus \{E_{n+1}\}$ and

$$\phi_1(\mathbf{p}) = \frac{1}{1 - p^{n+1}} \left(p^i \mathbf{e}_i\right), \tag{7}$$

and $\mathcal{O}_2 = S^n \setminus \{-E_{n+1}\}$ such that

$$\phi_2(\mathbf{p}) = \frac{1}{1 + p^{n+1}} \left(p^i \mathbf{e}_i\right). \tag{8}$$

We have $\phi_1(\mathcal{O}_1 \cap \mathcal{O}_2) = \mathbb{R}^n \setminus \{0\}$ and $\phi_2 \circ \phi_1^{-1}(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|^2}$.

***Proof.*** Take $\mathbf{x} \in \phi_1(\mathcal{O}_1 \cap \mathcal{O}_2) \subset \mathbb{R}^n$. We have that $\phi_1^{-1}(\mathbf{x}) = \frac{1}{1 + x_j x^j}\left(2x^i, x^j x_j - 1\right)$ which satisfies $\left|\phi_1^{-1}(\mathbf{x})\right| = 1$ and is an inverse as

$$\phi_1 \circ \phi_1^{-1}(x_i) = \frac{1}{1 - \frac{x^j x_j - 1}{1 + x_j x^j}} \frac{2x^i}{1 + x_j x^j} \tag{9}$$

$$= \frac{1 + x_j x^j}{1 + x_j x^j - (x^j x_j - 1)} \frac{2x^i}{1 + x_j x^j} \tag{10}$$

$$= \frac{1}{2} 2x^i = x^i. \tag{11}$$

Similarly, we have

$$\phi_2 \circ \phi_1^{-1}(x_i) = \frac{1}{1 + \frac{x^j x_j - 1}{1 + x_j x^j}} \frac{2x^i}{1 + x_j x^j} \tag{12}$$

$$= \frac{1 + x_j x^j}{1 + x_j x^j + (x^j x_j - 1)} \frac{2x^i}{1 + x_j x^j} \tag{13}$$

$$= \frac{1}{2x_j x^j} 2x^i = \frac{x^i}{|x|^2}, \tag{14}$$

which is well defined on $\mathbb{R}^n \setminus \{0\}$ as desired. $\square$

This is smooth on $\mathbb{R}^n \setminus \{0\}$ and similarly for $\phi_1 \circ \phi_2^{-1}$. Thus $S^n$ is an $n$-manifold.

# 2 Lecture: Smooth Functions on Manifolds

14/10/2024

## 2.1 Smooth Functions

Suppose $M, N$ are manifolds of $\dim n, n'$ respectively. Let $f : M \to N$ and $p \in M$. We pick charts $(\mathcal{O}_\alpha, \phi_\alpha)$ for $M$ and $\left(\mathcal{O}'_\beta, \phi'_\beta\right)$ for $N$ with $p \in \mathcal{O}_\alpha$ and $f(p) \in \mathcal{O}_\beta$.

Then $\phi'_\beta \circ f \circ \phi_\alpha^{-1}$ maps an open neighbourhood of $\phi_\alpha(p)$ in $U_\alpha \subset \mathbb{R}^n$ to $U'_\beta \subset \mathbb{R}^{n'}$.

**Definition 2.1:** If $\phi'_\beta \circ f \circ \phi_\alpha^{-1} : (U_\alpha \subset \mathbb{R}^n) \to \left(U'_\beta \subset \mathbb{R}^{n'}\right)$ is smooth for all possible choices of charts, we say $f : M \to N$ is **smooth**.

**Note.** A smooth map $\Psi : M \to N$ which has a smooth inverse $\Psi^{-1}$ is called a **diffeomorphism** and this implies $n = n'$.

Also, if $N = \mathbb{R}$ or $\mathbb{C}$, we sometimes call $f$ a **scalar field**. Further if $M$ is an (open) interval such that $M = I \subset \mathbb{R}$, then $f : I \to N$ is a smooth curve in $N$.

Lastly, if $f$ is smooth in one atlas, it is smooth with respect to all compatible atlases.

**Examples.**

1) Recall $S^1 = \{\mathbf{x} \in \mathbb{R}^2 \mid |x| = 1\}$. Let $f(x, y) = x$, $f : S^1 \to \mathbb{R}$.

   Using previous charts,

$$f \circ \phi_1^{-1} : (-\pi, \pi) \to \mathbb{R} \tag{15}$$

$$f \circ \phi_1^{-1}(\theta_1) = \cos\theta_1, \tag{16}$$

   and similarly,

$$f \circ \phi_2^{-1} : (0, 2\pi) \to \mathbb{R} \tag{17}$$

$$f \circ \phi_2^{-1}(\theta_2) = \cos\theta_2. \tag{18}$$

   In both cases, $f$ is smooth.

2) If $(\mathcal{O}, \phi)$ is a coordinate chart on $M$, write for $\mathbf{p} \in \mathcal{O}$,

$$\phi(\mathbf{p}) = \left(x^1(\mathbf{p}), x^2(\mathbf{p}), \cdots, x^n(\mathbf{p})\right), \tag{19}$$

   then $x^i(\mathbf{p})$ defines a map from $\mathcal{O}$ to $\mathbb{R}$. This is a smooth map for each $i = 1, \cdots, n$. If $(\mathcal{O}', \phi')$ is another overlapping coordinate chart, then $x^i \circ \phi'^{-1}$ is the $i$th component of $\phi \circ \phi'^{-1}$, which is smooth.

3) We can define a smooth function chart by chart. For simplicity, we take $N = \mathbb{R}$. Let $\{(\mathcal{O}_\alpha, \phi_\alpha)\}$ be an atlas on $M$. Define smooth functions $F_\alpha : U_\alpha \to \mathbb{R}$, and suppose that

$$F_\alpha \circ \phi_\alpha = F_\beta \circ \phi_\beta, \tag{20}$$

   on $\mathcal{O}_\alpha \cap \mathcal{O}_\beta$ for all $\alpha, \beta$. Then for $\mathbf{p} \in M$, we can define $f(\mathbf{p}) = F_\alpha \circ \phi_\alpha(\mathbf{p})$ where $(\mathcal{O}_\alpha, \phi_\alpha)$ is any chart with $\mathbf{p} \in \mathcal{O}_\alpha$ as this is constant by construction of $F$. $f$ is smooth as

$$f \circ \phi_\beta^{-1} = F_\alpha \circ \underbrace{\phi_\alpha \circ \phi_\beta^{-1}}_{\text{always smooth}}. \tag{21}$$

In practice, we often don't distinguish between $f$ and its **coordinate chart representation** $F_\alpha$. This coordinate chart representation $F_\alpha$ captures $f$ but maps from $U_\alpha \subset \mathbb{R}^n$ rather than from subsets of $M$. One can think of $F_\alpha = f \circ \phi_\alpha^{-1}$ as finding the point on $M$ that $\phi_\alpha$ mapped from and evaluating $f$ at that point.

## 2.2 Curves and Vectors

For a surface in $\mathbb{R}^3$, we have a notion of a tangent space at a point, consisting of all vectors tangent to the surface. Such tangent spaces are vector spaces (copies of $\mathbb{R}^2$). Different points have different tangent spaces.

In order to define the tangent space for a manifold, we first consider tangent vectors of a curve.

Recall that a smooth map from an interval $\lambda : I \subset \mathbb{R} \to M$ is a smooth curve in $M$.

If $\lambda(t)$ is a smooth curve in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function, then for $f(\lambda(t)) : \mathbb{R} \to \mathbb{R}$, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[f(\lambda(t))\right] = \mathbf{X}(t) \cdot \boldsymbol{\nabla} f(\lambda(t)), \tag{22}$$

where $\mathbf{X}(t) = \frac{\mathrm{d}\lambda(t)}{\mathrm{d}t}$ is the **tangent vector** to $\lambda$ at $t$. The idea is that we identify the tangent vector $\mathbf{X}(t)$ with the differential operator $\mathbf{X}(t) \cdot \boldsymbol{\nabla}$.

**Definition 2.2:** Let $\lambda : I \to M$ be a smooth curve with $\lambda(0) = \mathbf{p}$. The **tangent vector** to $\lambda$ at $\mathbf{p}$ is the linear map $X_\mathbf{p}$ from the space of smooth functions, $f : M \to \mathbb{R}$ given by

$$X_\mathbf{p}(f) = \frac{\mathrm{d}}{\mathrm{d}t} f(\lambda(t))\bigg|_{t=0}. \tag{23}$$

We observe a few things of note.

1) $X_\mathbf{p}$ is linear such that $X_\mathbf{p}(f + ag) = X_\mathbf{p}(f) + aX_\mathbf{p}(g)$ for $f, g$ smooth and $a \in \mathbb{R}$.
2) $X_\mathbf{p}$ satisfies the Leibniz rule,

$$X_\mathbf{p}(fg) = (X_\mathbf{p}(f))\, g + f X_\mathbf{p}(g). \tag{24}$$

3) If $(\mathcal{O}, \phi)$ is a chart with $\mathbf{p} \in \mathcal{O}$, we write

$$\phi(\mathbf{p}) = \left(x^1(\mathbf{p}), \cdots, x^n(\mathbf{p})\right). \tag{25}$$

Let $F = f \circ \phi^{-1}$, $x^i(t) = x^i(\lambda(t))$ and $\mathbf{x}(t) = \phi(\lambda(t))$. Then we have

$$f \circ \lambda(t) = f \circ \phi^{-1} \circ \phi \circ \lambda(t) = F \circ \mathbf{x}(t), \tag{26}$$

and thus the tangent vector defined in Eq. (23) can also be written as

$$X_\mathbf{p}(f) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(f(\lambda(t))\right)\bigg|_{t=0} = \frac{\partial F(x)}{\partial x^\mu} \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\bigg|_{t=0}, \tag{27}$$

where $\frac{\partial F}{\partial x^\mu}$ depends on $f$ and $\phi$ and $\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$ depends on $\lambda$ and $\phi$.

# 3 Lecture: Tangent Spaces

16/10/2024

## 3.1 The Tangent Space is a Vector Space

**Proposition 3.1:** The set of tangent vectors to curves at $\mathbf{p}$ forms a vector space, $T_{\mathbf{p}}M$ of dimension $n = \dim M$. We call $T_{\mathbf{p}}M$, the **tangent space** to $M$ at $\mathbf{p}$.

***Proof.*** Given $X_{\mathbf{p}}, Y_{\mathbf{p}}$ are tangent vectors, we need to show that $\alpha X_{\mathbf{p}} + \beta Y_{\mathbf{p}}$ is a tangent vector for $\alpha, \beta \in \mathbb{R}$. Let $\lambda, \kappa$ be smooth curves with $\lambda(0) = \kappa(0) = \mathbf{p}$ and whose tangent vectors at $\mathbf{p}$ are $X_{\mathbf{p}}$ and $Y_{\mathbf{p}}$ respectively. Let $(\mathcal{O}, \phi)$ be a chart with $p \in \mathcal{O}$ such that $\phi(\mathbf{p}) = 0$. We call this a *chart centered at* $\mathbf{p}$.

Let $\nu(t) = \phi^{-1}[\alpha\phi(\lambda(t)) + \beta\phi(\kappa(t))]$ where notice $\nu(0) = \phi^{-1}(0) = \mathbf{p}$.

From Eq. (27), we have that if $Z_p$ is the tangent to $\nu$ at $\mathbf{p}$, we have

$$Z_{\mathbf{p}}(f) = \frac{\mathrm{d}}{\mathrm{d}t}(f(\nu(t)))\bigg|_0 \tag{28}$$

$$= \frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t}[\alpha x^\mu(\lambda(t)) + \beta x^\mu(\kappa(t))]\bigg|_{t=0} \tag{29}$$

$$= \alpha\frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t}x^\mu(\lambda(t))\bigg|_{t=0} + \beta\frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t}x^\mu(\kappa(t))\bigg|_{t=0} \tag{30}$$

$$= \alpha X_{\mathbf{p}}(f) + \beta X_{\mathbf{p}}(f), \tag{31}$$

as desired. Therefore $T_{\mathbf{p}}M$ is a vector space. □

To see that $T_{\mathbf{p}}M$ is $n$-dimensional, consider the curves

$$\lambda_\mu(t) = \phi^{-1}\left(0, \cdots, 0, \underbrace{t}_{\mu\text{th component}}, 0, \cdots, 0\right). \tag{32}$$

We denote the tangent vector to $\lambda_\mu$ at $\mathbf{p}$ by $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$.

**Note.** This is **not** a differential operator.

However observe that by definition, we have

$$\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}(f) = \frac{\partial F}{\partial x^\mu}\bigg|_{\phi(\mathbf{p})=0}, \tag{33}$$

and thus it acts like a differential operator in $\mathbb{R}^n$ on the coordinates of the chart.

The vectors $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$ are linearly independent. Otherwise $\exists \alpha^\mu \in \mathbb{R}$ not all zero such that

$$\alpha^\mu \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} = 0, \tag{34}$$

which implies

$$\alpha^\mu \frac{\partial F}{\partial x^\mu} = 0, \tag{35}$$

for all $F$. Setting $F = x^\nu$ gives $\alpha^\nu = 0$ and thus these vectors are linearly independent and spanning. They are therefore a basis for the vector space.

Further one can see that $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$ form a basis for $T_{\mathbf{p}}M$, since if $\lambda$ is any curve with tangent $X_{\mathbf{p}}$ at $\mathbf{p}$, we have

$$X_{\mathbf{p}}(f) = \left.\frac{\partial F}{\partial x^\mu}\right|_{x=0} \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t)) = X^\mu \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} (f), \tag{36}$$

where $X^\mu = \left.\frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\right|_{t=0}$ are the **components** of $X_{\mathbf{p}}$ with respect to the basis $\{\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}\}_{\mu=1,\cdots,n}$ for $T_{\mathbf{p}}M$.

**Note.** The basis $\{\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}\}_{\mu=1,\cdots,n}$ depends on the coordinate chart $\phi$.

Suppose we choose another chart $(\mathcal{O}', \phi')$, again centered at $\mathbf{p}$. We write $\phi' = \left(\left(x'\right)^1, \cdots, (x')^n\right)$. Then if $F' = f \circ \phi'^{-1}$, we have

$$\begin{aligned} F(x) &= f \circ \phi^{-1}(x) && (37)\\ &= f \circ \phi'^{-1} \circ \phi' \circ \phi^{-1}(x) && (38)\\ &= F'(x'(x)). && (39) \end{aligned}$$

Therefore,

$$\begin{aligned} \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} (f) &= \left.\frac{\partial F}{\partial x^\mu}\right|_{\phi(\mathbf{p})} && (40)\\ &= \left.\left(\frac{\partial x'^\nu}{\partial x^\mu}\right)\right|_{\phi(\mathbf{p})} \left.\left(\frac{\partial F'}{\partial x'^\nu}\right)\right|_{\phi'(\mathbf{p})} && (41)\\ &= \left.\left(\frac{\partial x'^\nu}{\partial x^\mu}\right)\right|_{\phi(\mathbf{p})} \left(\frac{\partial}{\partial x'^\nu}\right)_{\mathbf{p}} (f). && (42) \end{aligned}$$

We then deduce that

$$\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} = \left.\left(\frac{\partial x'^\nu}{\partial x^\mu}\right)\right|_{\phi(\mathbf{p})} \left(\frac{\partial}{\partial x'^\nu}\right)_{\mathbf{p}}. \tag{43}$$

Let $X^\mu$ be components of $X_{\mathbf{p}}$ with respect to the basis $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$, and $X'^\mu$ be components of $X_{\mathbf{p}}$ with respect to the basis $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$ such that

$$\begin{aligned} X_{\mathbf{p}} = X^\mu \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} &= X'^\mu \left(\frac{\partial}{\partial x'^\mu}\right)_{\mathbf{p}} && (44)\\ &= X^\mu \left(\frac{\partial x'^\sigma}{\partial x^\mu}\right) \left(\frac{\partial}{\partial x'^\sigma}\right)_{\mathbf{p}}, && (45) \end{aligned}$$

and therefore

$$X'^\mu = \left(\frac{\partial x'^\mu}{\partial x^\nu}\right) X^\nu. \tag{46}$$

**Note.** We do note have to choose a coordinate basis such as $\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}$. With respect to a general basis $\{e_\mu\}$, for $T_{\mathbf{p}}M$, we can write $X_{\mathbf{p}} = X^\mu e_\mu$ for $X^\mu \in \mathbb{R}$.

We always use summation convention, contracting covariant indices with contravariant indices.

## 3.2 Covectors

Recall that if $V$ is a vector space over $\mathbb{R}$, the dual space $V^*$ is the space of linear maps $\phi : V \to \mathbb{R}$. If $V$ is $n$-dimensional then so is $V^*$ (the spaces are then isomorphic). Given a basis $\{e_\mu\}$ for $V$, we can define the dual basis $\{f^\mu\}$ for $V^*$ by requiring that

$$f^\mu\left(e_\nu\right) = \delta^\mu{}_\nu = \begin{cases} 1, & \text{if } \mu = \nu, \\ 0, & \text{if } \mu \neq \nu. \end{cases} \tag{47}$$

If $V$ is finite dimensional, then $V^{**} = \left(V^*\right)^* \simeq V$. Namely, to an element $X \in V$, we assign the linear map

$$\Lambda_X : V^* \to \mathbb{R}, \tag{48}$$

$$\Lambda_X\left(\omega\right) = \omega\left(X\right), \tag{49}$$

for $\omega \in V^*$.

**Definition 3.1:** The dual space of $T_{\mathbf{p}}M$ is denoted $T^*_{\mathbf{p}}M$ and is called the **cotangent space** to $M$ at $\mathbf{p}$. An element of this space is a **covector** at $\mathbf{p}$. If $\{e_\mu\}$ is a basis for $T_{\mathbf{p}}M$ and $\{f^\mu\}$ is the dual basis for $T^*_{\mathbf{p}}M$, we can expand a covector $\eta$ as

$$\eta = \eta_\mu f^\mu, \tag{50}$$

for **components** $\eta_\mu \in \mathbb{R}$.

# 4 Lecture: Tensors

18/10/2024

## 4.1 Tangent bundle

Notice that

$$\eta\left(e_\nu\right) = \eta_\mu f^\mu\left(e_\nu\right) = \eta_\mu \delta^\mu{}_\nu = \eta_\nu, \tag{51}$$

and thus we can get the components of $\eta$ by acting it on basis vectors in the tangent space. Further as we have $X = X^\mu e_\mu$,

$$\eta\left(X\right) = \eta\left(X^\mu e_\mu\right) \tag{52}$$

$$= X^\mu \eta\left(e_\mu\right) \tag{53}$$

$$= X^\mu \eta_\mu, \tag{54}$$

and thus the action of the covector $\eta$ on the vector $X$ is essentially a contraction between the components.

Recall that a vector $X$ is defined by its action on a function $f$, $X : f \to \mathbb{R}$, eating a smooth function and returning the rate of change as one moves in the direction of $X$.

Analogously, given a function $f$, one can consider a linear operator of that function being eaten by a generic vector $X$.

**Definition 4.1:** If $f : M \to \mathbb{R}$ is a smooth function, then we can define a covector $(df)_{\mathbf{p}} \in T^*_{\mathbf{p}}M$, the **differential** of $f$ at $\mathbf{p}$, by

$$(df)_{\mathbf{p}}(X) = X(f), \tag{55}$$

for any $X \in T_{\mathbf{p}}M$. This is also sometimes called the **gradient** of $f$ at $\mathbf{p}$.

If $f$ is constant, $X(f) = 0$ which implies $(df)_{\mathbf{p}} = 0$.

If $(\mathcal{O}, \phi)$ is a coordination chart with $\mathbf{p} \in \mathcal{O}$ and $\phi = \left(x^1, \cdots, x^n\right)$ then we can set $f = x^\mu$ to find $(dx^\mu)_{\mathbf{p}}$. Observe

$$(dx^\mu)_{\mathbf{p}} \left(\frac{\partial}{\partial x^\nu}\right)_{\mathbf{p}} = \left(\frac{\partial x^\mu}{\partial x^\nu}\right)_{\phi(\mathbf{p})} = \delta^\mu{}_\nu. \tag{56}$$

Therefore the coordinate differentials $\{(dx^\mu)_{\mathbf{p}}\}$ is the dual basis to $\{\left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}}\}$.

In this basis, we can compute

$$\left[(df)_{\mathbf{p}}\right]_\mu = (df)_{\mathbf{p}} \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} = \left(\frac{\partial}{\partial x^\mu}\right)_{\mathbf{p}} f = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(\mathbf{p})}. \tag{57}$$

This justifies the language of *gradient*.

**Exercise 1:** Show that if $(\mathcal{O}', \phi')$ is another chart with $\mathbf{p} \in \mathcal{O}'$, then

$$(dx^\mu)_{\mathbf{p}} = \left(\frac{\partial x^\mu}{\partial (x')^\nu}\right)_{\phi'(\mathbf{p})} \left(d(x')^\nu\right)_{\mathbf{p}}, \tag{58}$$

where $x(x') = \phi \circ (\phi')^{-1}$, and hence if $\eta_\mu, \eta'_\mu$ are components with respect to these bases,

$$\eta'_\mu = \left(\frac{\partial x^\nu}{\partial (x')^\mu}\right)_{\phi'(\mathbf{p})} \eta_\nu. \tag{59}$$

***Proof.***

□

**Definition 4.2 (Tangent bundle):** We can glue together the tangent spaces $T_{\mathbf{p}}M$ as $\mathbf{p}$ varies to get a new $2n$ dimensional manifold $TM$, the **tangent bundle**. We define the tangent bundle to be

$$TM := \bigcup_{\mathbf{p} \in M} \{\mathbf{p}\} \times T_{\mathbf{p}}M. \tag{60}$$

Namely, it is the set of ordered pairs $(\mathbf{p}, X)$, with $\mathbf{p} \in M$, $X \in T_{\mathbf{p}}M$.

If $\{(\mathcal{O}_\alpha, \phi_\alpha)\}$ is an atlas on $M$, we obtain an atlas for $TM$ by setting

$$\mathcal{O}_\alpha = \bigcup_{\mathbf{p} \in \mathcal{O}_\alpha} \{\mathbf{p}\} \times T_{\mathbf{p}}M, \tag{61}$$

and

$$\widetilde{\phi}_\alpha\left((p, X)\right) = \left(\phi\left(\mathbf{p}\right), X^\mu\right) \in \mathcal{U}_\alpha \times \mathbb{R}^n = \widetilde{\mathcal{U}}_2, \tag{62}$$

where $X^\mu$ are the components of $X$ with respect to the coordinate basis of $\phi_\alpha$.

**Exercise 2:** If $(\mathcal{O}, \phi)$ and $(\mathcal{O}', \phi')$ are two charts on $M$, show that on $\widetilde{U} \cap \widetilde{U}'$, if we write $\phi' \circ \phi^{-1}(x) = x'(x)$, then

$$\widetilde{\phi}' \circ \widetilde{\phi}^{-1}\left(x, X^\mu\right) = \left(x'\left(x\right), \left(\frac{\partial \left(x'\right)^\mu}{\partial x^\nu}\right)_x X^\nu\right). \tag{63}$$

Deduce that $TM$ is a (differentiable) manifold.

***Proof.***

□

A similar construction permits us to define the cotangent bundle $T^*M = \cup_{\mathbf{p}\in M}\{\mathbf{p}\} \times T^*_{\mathbf{p}}M$.

**Exercise 3:** Show that the map $\Pi : TM \to M$ which projects onto the point such that

$$(\mathbf{p}, X) \mapsto \mathbf{p}, \tag{64}$$

is smooth.

***Proof.***

□

## 4.2 Abstract Index Notation

We have used Greek letters $\mu, \nu$ etc. to label components of vectors (or covectors) with respect to the basis $\{e_\mu\}$ (respectively $\{f^\mu\}$). Equations involving these quantities refer to a specific basis.

**Example.** Taking $X^\mu = \delta^\mu$, this says $X^\mu$ only has one non-zero component in the current basis. This won't be true in other bases as $X^\mu$ transforms.

We know some equations do hold in all bases, for example,

$$\eta\left(X\right) = X^\mu \eta_\mu. \tag{65}$$

To capture this, we use *abstract index notation.* We denote a vector with $X^a$, where the Latin index $a$ does not denote a component, rather it tells us $X^a$ is a vector. Similarly, we denote a covector $\eta$ by $\eta_a$.

If an equation is true in all bases, we can replace Greek indices by Latin indices, for example

$$\eta\left(X\right) = X^a \eta_a = \eta_a X^a, \tag{66}$$

or

$$X\left(f\right) = X^a \left(df\right)_a. \tag{67}$$

Such notation tells us that the expression is **independent** of the choice of basis. One can go back to Greek letter by picking a basis and swapping $a \to \mu$.

### 4.3 Tensors

**Definition 4.3:** A **tensor** of type $(r, s)$ at $p$ is a multilinear map

$$T : \underbrace{T^*_{\mathbf{p}}(M) \times \cdots \times T^*_{\mathbf{p}}(M)}_{r \text{ factors}} \times \underbrace{T_{\mathbf{p}}(M) \times \cdots \times T_{\mathbf{p}}(M)}_{s \text{ factors}} \to \mathbb{R}, \tag{68}$$

where multilinear map means linear in each argument.

**Examples.**

- A tensor of type $(0, 1)$ is a linear map $T_{\mathbf{p}}(M) \to \mathbb{R}$, i.e. it is a covector.
- A tensor of type $(1, 0)$ is a linear map from $T^*_{\mathbf{p}}(M) \to \mathbb{R}$, i.e. an element of $\left(T^*_{\mathbf{p}}(M)\right)^* \simeq T_{\mathbf{p}}(M)$ thus it is a vector.
- We can define a $(1, 1)$ tensor, $\delta$ by $\delta(\omega, X) = \omega(X)$ for any covector $\omega$ and vector $X$.

**Definition 4.4:** If $\{e_\mu\}$ is a basis for $T_{\mathbf{p}}M$ and $\{f^\mu\}$ is the dual basis, the components of an $(r, s)$ tensor $T$ are

$$T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s} = T\left(f^{\mu_1}, \cdots, f^{\nu_r}, e_{\nu_1} \cdots e_{\nu_s}\right). \tag{69}$$

In abstract index notation we write $T$ as $T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$.

**Note.** Tensors of type $(r, s)$ at $p$ form a vector space over $\mathbb{R}$ of dimension $n^{r+s}$.

**Examples.**

1) Consider the $\delta$ tensor above. It has components

$$\delta^\mu{}_\nu := \delta(X, \omega) = f^\mu(e_\nu), \tag{70}$$

which recovers our expected Kronecker delta $\delta^\mu_\nu$.

2) Consider a $(2, 1)$ tensor $T$. If $\omega, \eta \in T^*_{\mathbf{p}}M$, $X \in T_{\mathbf{p}}M$,

$$\begin{aligned} T(\omega, \eta, X) &= T\left(\omega_\mu f^\mu, \eta_\nu f^\nu, X^\sigma e_\sigma\right) && (71) \\ &= \omega_\mu \eta_\nu X^\sigma T\left(f^\mu, f^\nu, e^\sigma\right) && (72) \\ &= \omega_\mu \eta_\nu X^\sigma T^{\mu\nu}{}_\sigma. && (73) \end{aligned}$$

which in abstract index notation is $T(\omega, \eta, X) = \omega_a \eta_b X^c T^{ab}{}_c$. This generalised to higher ranks.

# 5 Lecture: Tensor Fields

21/10/2024

I now drop the bold face on $\mathbf{p} \in M \to p \in M$.

### 5.1 Change of Bases

We've seen how components of $X$ or $\eta$ change with respect to a coordinate basis ($X^\mu, \eta_\nu$, respectively). Under a change of coordinates, we don't only have to consider coordinate bases.

Suppose $\{e_\mu\}$ and $\{e'_\mu\}$ are two bases for $T_pM$ with dual bases $\{f^\mu\}$ and $\{f'^\mu\}$.

We can expand

$$f'^\mu = A^\mu{}_\nu f^\nu \text{ and } e'_\mu = B^\nu{}_\mu e_\nu, \tag{74}$$

but

$$\begin{aligned}
\delta^\mu_\nu &= f'^\mu\left(e'_\nu\right) && (75)\\
&= A^\mu{}_\tau f^\tau\left(B^\sigma{}_\nu e_\sigma\right) && (76)\\
&= A^\mu{}_\tau B^\sigma{}_\nu f^\tau\left(e_\sigma\right) && (77)\\
&= A^\mu{}_\tau B^\sigma{}_\nu \delta^\tau{}_\sigma && (78)\\
&= A^\mu{}_\sigma B^\sigma{}_\nu, && (79)
\end{aligned}$$

Thus $B^\mu{}_\nu = \left(A^{-1}\right)^\mu{}_\nu$.

If $e_\mu = \left(\frac{\partial}{\partial x^\mu}\right)_p$ and $e'_\mu = \left(\frac{\partial}{\partial x^\mu}\right)_p$. We've already seen

$$A^\mu{}_\nu = \left(\frac{\partial x'^\mu}{\partial x^\nu}\right)_{\phi(p)} \qquad B^\mu{}_\nu = \left(\frac{\partial x^\mu}{\partial x'^\nu}\right)_{\phi(p)}. \tag{80}$$

Therefore we see that a change of bases induces a transformation of tensor components. For example, if $T$ is a $(1,1)$-tensor,

$$\begin{aligned}
T^\mu{}_\nu &= T\left(f^\mu, e_\nu\right) && (81)\\
T'^\mu{}_\nu &= T\left(f'^\mu, e'_\nu\right) && (82)\\
&= T\left(A^\mu{}_\sigma f^\sigma, \left(A^{-1}\right)^\tau{}_\nu e_\tau\right) && (83)\\
&= A^\mu{}_\sigma \left(A^{-1}\right)^\tau{}_\nu T\left(f^\sigma, e_\tau\right) && (84)\\
&= A^\mu{}_\sigma \left(A^{-1}\right)^\tau{}_\nu T^\sigma{}_\tau. && (85)
\end{aligned}$$

## 5.2 Tensor operations

**Definition 5.1:** Given an $(r,s)$ tensor, we can form an $(r-1,s-1)$ tensor by **contraction**.

For simplicity assume $T$ is a $(2,2)$ tensor. Define a $(1,1)$ tensor $S$ by

$$S\left(\omega, X\right) = T\left(\omega, f^\mu, X, e_\mu\right). \tag{86}$$

To see that this is independent of the choice of basis, observe that a different basis would give

$$\begin{aligned}
S\left(\omega, X\right) = T\left(\omega, f'^\mu, X, e'_\mu\right) &= T\left(\omega, A^\mu{}_\sigma f^\sigma, X, \left(A^{-1}\right)^\tau{}_\mu e_\tau\right) && (87)\\
&= A^\mu{}_\sigma \left(A^{-1}\right)^\tau{}_\mu T\left(\omega, f^\sigma, X, e_\tau\right) && (88)\\
&= \delta^\tau_\sigma T\left(\omega, f^\sigma, X, e_\tau\right) && (89)\\
&= T\left(\omega, f^\sigma, X, e_\sigma\right) = S\left(\omega, X\right), && (90)
\end{aligned}$$

and thus we have basis independence as desired. Thus we write the components of these tensors as

$$S^\mu{}_\nu = T^{\mu\sigma}{}_{\nu\sigma}, \tag{91}$$

which in abstract index notation, is written

$$S^a{}_b = T^{ac}{}_{bc} \, . \tag{92}$$

This can be generalized to contract any pair of covariant (lower) and contravariant (upper) indices on an arbitrary tensor.

Another way to form new tensors is to use a *tensor product*.

**Definition 5.2:** If $S$ is a $(p,q)$ tensor and $T$ is an $(r,s)$ tensor then $S \otimes T$ is a $(p+r, q+s)$ tensor given by

$$S \otimes T \left(\omega^1, \cdots, \omega^p, \eta^1, \cdots, \eta^r, X_1, \cdots, X_q, Y_1, \cdots, Y_s\right), \tag{93}$$

which in abstract index notation can be written

$$(S \otimes T)^{a_1 \cdots a_p b_1 \cdots b_r}{}_{c_1 \cdots c_q d_1 \cdots d_s} = S^{a_1 \cdots a_p}{}_{c_1 \cdots c_q} T^{b_1 \cdots b_r}{}_{d_1 \cdots d_s} \, . \tag{94}$$

**Exercise 4:** For any $(1,1)$ tensor $T$, in a basis we have

$$T = T^\mu{}_\nu \, e_\mu \otimes f^\nu . \tag{95}$$

***Proof.***

□

The final tensor operations we require are anti-symmetrization and symmetrization.

**Definition 5.3:** If $T$ is a $(0,2)$ tensor, we can define two new tensors

$$S(X,Y) = \frac{1}{2}\left(T(X,Y) + T(Y,X)\right) \tag{96}$$

$$A(X,Y) = \frac{1}{2}\left(T(X,Y) - T(Y,X)\right), \tag{97}$$

which in abstract index notation become

$$S_{ab} = \frac{1}{2}\left(T_{ab} + T_{ba}\right) \tag{98}$$

$$A_{ab} = \frac{1}{2}\left(T_{ab} - T_{ba}\right), \tag{99}$$

one also writes $S_{ab} = T_{(ab)}$ and $A_{ab} = T_{[ab]}$ to denote symmetrization and antisymmetrization respectively.

These operations can be applied to any pair of matching indices. Similarly, to symmetrize over $n$ indices we sum over all permutations and divide by $n!$, and identically to antisymmetrize, with the addition of a minus sign for odd permutations.

For example,

$$T^{(abc)} = \frac{1}{3!}\left(T^{abc} + T^{bca} + T^{cab} + T^{acb} + T^{cba} + T^{bac}\right) \tag{100}$$

$$T^{[abc]} = \frac{1}{3!}\left(T^{abc} + T^{bca} + T^{cab} - T^{acb} - T^{cba} - T^{bac}\right). \tag{101}$$

Lastly, to exclude indices from symmetrization, we use vertical lines such that

$$T^{(a|b|c)} = \frac{1}{2}\left(T^{abc} + T^{cba}\right). \tag{102}$$

## 5.3 Tensor Bundles

**Definition 5.4:** The space of $(r,s)$ tensors at a point $p$ is the vector space $\left(T^r{}_s\right)_p M$. These can be glued together to form the **bundle** of $(r,s)$-tensors, which we write

$$T^r{}_s M = \bigcup_{p\in M} \{p\} \times \left(T^r{}_s\right)_p M. \tag{103}$$

If $(\mathcal{O},\phi)$ is a coordinate chart on $M$, set

$$\widetilde{\mathcal{O}} = \bigcup_{p\in \mathcal{O}} \{p\} \times \left(T^r{}_s\right)_p M \subset T^r{}_s M, \tag{104}$$

where $\widetilde{\phi}\left(p, S_p\right) = \left(\phi\left(p\right), S^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\right)$

$T^r{}_s M$ is a manifold with a natural smooth map $\Pi : T^r{}_s M \to M$ such that $\Pi\left(p, S_p\right) = p$.

**Definition 5.5:** An $(r,s)$ tensor field is a smooth map $T : M \to T^r{}_s M$ such that $\Pi \circ T = \mathrm{id}$ (namely, that $T : p \mapsto (p, S_p)$). If $(\mathcal{O},\phi)$ is a coordinate chart on $M$ then

$$\widetilde{\phi} \circ T \circ \phi^{-1}\left(x\right) = \left(x, T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\left(x\right)\right), \tag{105}$$

which is smooth provided the components $T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\left(x\right)$ are smooth functions of $x$.

One can think of a tensor field as defining a tensor at every point with respect to the coordinate basis at that point.

If $T^r{}_s M = T^1{}_0 M \sim TM$, the tensor field is called a **vector field**. In a local coordinate patch, if $X$ is a vector field, we can write

$$X\left(p\right) = \left(p, X_p\right), \tag{106}$$

with $X_p = X^\mu\left(x\right)\left(\frac{\partial}{\partial x^\mu}\right)_p$.

In particular, $\frac{\mathrm{d}}{\mathrm{d}x^\mu}$ are always smooth but only defined locally.

# 6 Lecture: The metric tensor

23/10/2024

A vector field can be thought of as we usually do, as placing a vector at every point on the manifold. A vector field can also act on a function $f : M \to \mathbb{R}$ to give a new function

$$Xf\left(p\right) = X_p\left(f\right), \tag{107}$$

which in a coordinate basis becomes

$$Xf(p) = X^{\mu}(\phi(p)) \left.\frac{\partial F}{\partial x^{\mu}}\right|_{\phi(p)}, \tag{108}$$

which we now think of as a function of $p$ across the manifold.

## 6.1 Integral curves

**Definition 6.1:** Given a vector field $X$ on $M$, we say a curve $\lambda : I \to M$, is an **integral curve** of $X$ if its tangent vector at every point along it is $X$. Namely, denote the tangent vector to $\lambda$ at $t$ by $\frac{\mathrm{d}\lambda}{\mathrm{d}t}(t)$, then

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) = X_{\lambda(t)}, \tag{109}$$

$\forall t \in I$.

Through each point $p$, an integral curve passes, and is unique up to reparametrization or curve extension.

To see that this is true, pick a chart $\phi$ with $\phi = \left(x^1, \cdots, x^n\right)$ and assume $\phi(p) = 0$. In this chart, Eq. (109) becomes

$$\frac{\mathrm{d}x^{\mu}}{\mathrm{d}t}(t) = X^{\mu}(x(t)), \tag{110}$$

where $x^{\mu}(t) = x^{\mu}(\lambda(t))$. Assuming without loss of generality that $\lambda(0) = p$, we get an initial condition that $x^{\mu}(0) = 0$.

Standard ODE theory gives us that Eq. (110) with an initial condition has a solution unique up to extension.

## 6.2 Commutators

Suppose $X$ and $Y$ are two vector fields and $f : M \to \mathbb{R}$ is smooth. Then $X(Y(f))$ is a smooth function. Is it of the form $K(f)$ for some vector field $K$? No, as

$$\begin{aligned} X(Y(fg)) = X(fX(g) + gY(f)) &= X(Y(fg)) && (111)\\ &= X(fX(g)) + X(gY(f)) && (112)\\ &= fX(Y(g)) + gX(Y(f)) + X(f)Y(g) + X(g)Y(f), && (113) \end{aligned}$$

and thus the Leibniz rule does not hold, implying this cannot be a vector field. However notice that the last two terms that ruin this are symmetric in $f$ and $g$, and thus if we instead consider

$$[X, Y](f) = X(Y(f)) - Y(X(f)), \tag{114}$$

then the Leibniz rule will hold and (while we have not show it explicitly) this does in fact define a vector field.

To see this, use coordinate bases such that

$$[X, Y](f) = X\left(Y^{\nu}\frac{\partial F}{\partial x^{\nu}}\right) - Y\left(X^{\mu}\frac{\partial F}{\partial x^{\mu}}\right) \tag{115}$$

$$= X^\mu \frac{\partial}{\partial x^\mu}\left(Y^\nu \frac{\partial F}{\partial x^\nu}\right) - Y^\nu \frac{\partial}{\partial x^\nu}\left(X^\mu \frac{\partial F}{\partial x^\mu}\right) \tag{116}$$

$$= X^\mu Y^\mu \frac{\partial^2 F}{\partial x^\mu \partial x^\nu} - Y^\nu X^\mu \frac{\partial^2 F}{\partial x^\nu \partial x^\mu} + X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\frac{\partial F}{\partial x^\mu}. \tag{117}$$

As mixed partials on smooth functions in $\mathbb{R}^n$ commute, the first two terms cancel leaving

$$[X,Y](f) = X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\frac{\partial F}{\partial x^\mu} \tag{118}$$

$$= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\nu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial F}{\partial x^\nu} \tag{119}$$

$$= [X,Y]^\nu \frac{\partial F}{\partial x^\nu}, \tag{120}$$

where $[X,Y]^\nu = X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}$ are the components of the commutator.

Since $f$ is arbitrary, the expression

$$[X,Y] = [X,Y]^\nu \frac{\partial}{\partial x^\nu}, \tag{121}$$

is valid only once one has chosen a coordinate basis.

## 6.3 The metric tensor

We're familiar from Euclidean geometry (and special relativity) with the fact that the fundamental object of when discussing distance and angles (time intervals/rapidity) is an inner product between vectors.

**Definition 6.2:** A **metric tensor** at $p \in M$ is a $(0,2)$-tensor $g$, satisfying two conditions:

i) $g$ is *symmetric* such that $g(X,Y) = g(Y,X)$, $\forall X, Y \in T_pM$, i.e. $g_{ab} = g_{ba}$,

ii) $g$ is *non-degenerate*, $G(X,Y) = 0$, $\forall Y \in T_pM \Leftrightarrow X = 0$.

Sometimes we write $g(X,Y) = \langle X,Y\rangle = \langle X,Y\rangle_g = X \cdot Y$.

By adapting the Gram-Schmidt algorithm, we can always find a basis $\{e_\mu\}$ for the tangent space at $p$, $T_pM$, such that

$$g(e_\mu, e_\nu) = \begin{cases} 0, & \mu \neq \nu, \\ +1 \text{ or } -1, & \mu = \nu. \end{cases} \tag{122}$$

Note this basis is not unique, but the **signature** (the number of +1's and −1's) does not depend on the choice of basis (Sylvestre's Law of inertia).

If $g$ has signature $(+ + \cdots +)$ we say it is **Riemannian**.

If $g$ has signature $(- + \cdots +)$, we say it is **Lorentzian**.

**Definition 6.3:** A **Riemannian manifold** (or respectively a Lorentzian manifold) is a pair $(M, g)$ where $M$ is a manifold and $g$ is a Riemannian (or respectively Lorentzian) metric tensor field.

On a Riemannian manifold, the norm of a vector is

$$|X| = \sqrt{g\left(X, X\right)}, \tag{123}$$

and the angle between $X, Y \in T_pM$, is given by

$$\cos\theta = \frac{g\left(X, Y\right)}{|X|\,|Y|}. \tag{124}$$

The length $\ell$ of a curve $\lambda : (a, b) \to M$ is given by

$$\ell\left(\lambda\right) = \int_a^b \left|\frac{\mathrm{d}\lambda}{\mathrm{d}t}\left(t\right)\right| \mathrm{d}t\,. \tag{125}$$

**Exercise 5:** If $\tau : (c, d) \to (a, b)$ with $\frac{\mathrm{d}t}{\mathrm{d}\tau} > 0$ and $\tau\left(c\right) = a$, $\tau\left(d\right) = b$, then

$$\widetilde{\lambda} = \lambda \circ \tau : (c, d) \to M, \tag{126}$$

is a reparametrization of $\lambda$ such that $\ell\left(\widetilde{\lambda}\right) = \ell\left(\lambda\right)$.

***Proof.***

□

# 7 Lecture: Proper time

25/10/2024

In a coordinate basis, $g = g_{\mu\nu}\,\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$. We often write

$$\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \frac{1}{2}\left(\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu + \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu\right), \tag{127}$$

and by convention often write $g = \mathrm{d}s^2$ so that

$$g = \mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\,. \tag{128}$$

**Examples.**

i) $\mathbb{R}^n$ with $g = \mathrm{d}s^2 = \left(\mathrm{d}x^1\right)^2 + \cdots + \left(\mathrm{d}x^n\right)^2 = \delta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$ is called **Euclidean space**. Any chart covering $\mathbb{R}^n$ in which the metric takes this form is called **Cartesian**.

ii) $\mathbb{R}^{1+3} = \{\left(x^0, x^1, x^2, x^3\right)\}$ with

$$\begin{aligned} g = \mathrm{d}s^2 &= -\left(\mathrm{d}x^0\right)^2 + \left(\mathrm{d}x^1\right)^2 + \left(\mathrm{d}x^2\right)^2 + \left(\mathrm{d}x^3\right)^2 && (129)\\ &= \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\,, && (130)\end{aligned}$$

is **Minkowski space**. A coordinate chart covering $\mathbb{R}^{1+3}$ in which the metric takes this form is called an **inertial frame**.

iii) On $S^2 = \{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x}| = 1\}$. Define a chart by

$$\phi^{-1} : (0, \pi) \times (-\pi, \pi) \to S^2 \tag{131}$$
$$(\theta, \phi) \mapsto (\cos\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)\,, \tag{132}$$

In this chart, the **round metric** is

$$g = \mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2\theta\, \mathrm{d}\phi^2\,. \tag{133}$$

This covers $S^2 \setminus \{|\mathbf{x}| = 1,\ x^2 = 0,\ x' \leq 0\}$. To cover the rest, let

$$\widetilde{\phi}^{-1} : (0, \pi) \times (-\pi, \pi) \to S^2 \tag{134}$$
$$(\theta', \phi') \mapsto (-\sin\theta' \cos\phi', \cos\theta', \sin\theta' \sin\phi')\,. \tag{135}$$

This covers $S^2 \setminus \{|\mathbf{x}| = 1,\ x^3 = 0,\ x' \geq 0\}$ and thus setting

$$g = \mathrm{d}\theta'^2 + \sin^2\theta'\, \mathrm{d}\phi'^2\,. \tag{136}$$

Defines a metric on all of $S^2$.

Since $g_{ab}$ is non-degenerate, it is invertible as a matrix in any basis. We can check that the inverse defines a symmetric $(2, 0)$ tensor, $g^{ab}$ satisfying

$$g^{ab} g_{bc} = \delta^a_c. \tag{137}$$

**Example.** In the $\phi$ coordinates of the $S^2$ example.

$$g^{\mu\nu} = \left(1, \frac{1}{\sin^2\theta}\right). \tag{138}$$

An important property of the metric is that it induces a canonical identification of $T_pM$ and $T_p^*M$. Given $X^a \in T_pM$, we define a covector $g_{ab}X^b = X_a$ and given $\eta_a \in T_p^*M$ we define a vector $g^{ab}\eta_b = \eta^a$.

In Euclidean space $\left(\mathbb{R}^3, \delta\right)$ we often do this without realising.

More generally, this allows us to raise tensor indices with $g^{ab}$ and lower them with $g_{ab}$. Namely, if $T^{ab}{}_c$ is a $(2, 1)$ tensor, then $T_a{}^{bc}$ is the $(2, 1)$ tensor given by

$$T_a{}^{bc} = g_{ad} g^{ce} T^{db}{}_e\,. \tag{139}$$

## 7.1 Lorentzian signature

At any point $p$ in a Lorentzian manifold we can find a basis $\{e_\mu\}$ such that

$$g\left(e_\mu, e_\nu\right) = \eta_{\mu\nu} = \mathrm{diag}\left(-1, 1, \cdots, 1\right). \tag{140}$$

This basis is not unique. Namely, if $e'_\mu = \left(A^{-1}\right)^\nu{}_\mu e_\nu$ is another such basis, then

$$\eta_{\mu\nu} = g\left(e'_\mu, e'_\nu\right) = \left(A^{-1}\right)^\sigma{}_\mu \left(A^{-1}\right)^\tau{}_\nu\, g\left(e_\sigma, e_\tau\right) \left(A^{-1}\right)^\sigma{}_\mu \left(A^{-1}\right)^\tau{}_\nu \tag{141}$$
$$= \left(A^{-1}\right)^\sigma{}_\tau \left(A^{-1}\right)^\tau{}_\nu\, \eta_{\sigma\tau} \tag{142}$$
$$\Rightarrow A^\mu{}_\kappa A^\nu{}_\rho \eta_{\mu\nu} = \eta_{\kappa\rho}, \tag{143}$$

which is the condition that $A^\mu{}_\nu$ is a **Lorentz transformation**.

The tangent space at $p$ has $\eta_{\mu\nu}$ as a metric tensor (in this basis) so has the structure of Minkowski space.

**Definition 7.1:** $X \in T_pM$ is

$$\begin{cases} \text{spacelike,} & \text{if } g(X,X) > 0, \\ \text{null-like/light-like,} & \text{if } g(X,X) = 0, \\ \text{timelike,} & \text{if } g(X,X) < 0. \end{cases} \tag{144}$$

A curve $\lambda : I \to M$ in a Lorentzian manifold is spacelike/timelike/null if the tangent vector is spacelike/timelike/null everywhere respectively.

A spacelike curve has a well-defined **length**, given by the same formula as in the Riemannian case. For a timelike curve $\lambda : (a,b) \to M$, the relevant quantity is the **proper time**

$$\tau(\lambda) = \int_a^b = \sqrt{-g_{ab}\frac{\mathrm{d}\lambda^a}{\mathrm{d}u}\frac{\mathrm{d}\lambda^b}{\mathrm{d}u}}\,\mathrm{d}u\,. \tag{145}$$

If $g_{ab}\frac{\mathrm{d}\lambda^a}{\mathrm{d}u}\frac{\mathrm{d}\lambda^b}{\mathrm{d}u} = -1$ for all $u$, then $\lambda$ is parametrised by proper time.

In this case we call the tangent vector

$$u^a \equiv \frac{\mathrm{d}\lambda^a}{\mathrm{d}u}, \tag{146}$$

the **4-velocity** of $\lambda$.

## 7.2 Curves of extremal proper time

Suppose $\lambda : (0,1) \to M$ is timelike, satisfies $\lambda(0) = p$, $\lambda(1) = q$ and extremizes proper time among all such curves. This is a variational problem, associated to (in a coordinate chart),

$$\tau[\lambda] = \int_0^1 G(x^\mu(u), \dot{x}^\mu(u))\,\mathrm{d}u\,, \tag{147}$$

with

$$G(x^\mu(u), \dot{x}^\mu(u)) = \sqrt{-g_{\mu\nu}(x(u))\,\dot{x}^\mu(u)\,\dot{x}^\nu}, \tag{148}$$

where $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}u}$. The Euler Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial G}{\partial \dot{x^\mu}}\right) = \frac{\partial G}{\partial x^\mu}. \tag{149}$$

We can compute

$$\frac{\partial G}{\partial \dot{x}^\mu} = -\frac{1}{G} g_{\mu\nu}\dot{x}^\nu \tag{150}$$

$$\frac{\partial G}{\partial x^\mu} = -\frac{1}{2G}\frac{\partial}{\partial x^\mu}(g_{\sigma\tau})\,\dot{x}^\sigma\dot{x}^\tau \tag{151}$$

$$= -\frac{1}{2G} g_{\sigma\tau,\mu}\dot{x}^\sigma\dot{x}^\tau. \tag{152}$$

This does not have a unique solution as one can re-parametrize the curve without changing the proper time $\tau$.

# 8 Lecture: Christoffel Symbols

28/10/2024

## 8.1 Geodesic Equation

Therefore we fix the parameterisation such that the curve is parameterized by the proper time $\tau$ itself. Doing this, since

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \dot{x}^\mu \frac{\mathrm{d}u}{\mathrm{d}\tau} \text{ and } -1 = g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}, \tag{153}$$

we have that

$$-1 = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \left( \frac{\mathrm{d}u}{\mathrm{d}\tau} \right)^2 \Rightarrow \frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{1}{\sqrt{G}}, \tag{154}$$

which then implies

$$\frac{1}{G} \frac{\mathrm{d}}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}\tau}. \tag{155}$$

Returning to the Euler Lagrange equation, we find that we can write it as

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \left( g_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \right) = \frac{1}{2} g_{\nu\rho,\mu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}. \tag{156}$$

This then becomes

$$g_{\mu\nu} \frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\tau^2} + g_{\mu\nu,\rho} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} - \frac{1}{2} g_{\sigma\rho,\mu} \frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0. \tag{157}$$

Where notice that we can replace

$$g_{\mu\nu,\rho} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = g_{\mu(\nu,\rho)} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}, \tag{158}$$

as it is symmetric in $\nu$ and $\rho$.

Thus, notice that we can write this as

$$\frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\tau^2} + \Gamma_\mu{}^\nu{}_\rho \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0, \tag{159}$$

where

$$\Gamma_\mu{}^\nu{}_\rho \equiv \frac{1}{2} g^{\nu\sigma} \left( g_{\mu\sigma,\rho} + g_{\sigma\rho,\mu} - g_{\nu\rho,\sigma} \right), \tag{160}$$

are the **Christoffel symbols** of $g$.

**Notes.**

- These symbols have a symmetry such that

$$\Gamma_\nu{}^\mu{}_\rho = \Gamma_\rho{}^\mu{}_\nu. \tag{161}$$

- Christoffel symbols are **not** tensor components as they do not transform desirably under coordinate transformations.
- Solutions to Eq. (159) are obtainable with standard ODE theory. Such solutions are called **geodesics**.
- The same equation governs curves of extremal length in a Riemannian manifold (or spacelike curves in a Lorentzian manifold) parameterized by arc-length.

**Exercise 6:** Show that Eq. (159) can be obtained as the Euler-Lagrange equation for the Lagrangian

$$L = -g_{\mu\nu}\left(x\left(\tau\right)\right)\dot{x}^{\mu}\left(\tau\right)\dot{x}^{\nu}\left(\tau\right). \tag{162}$$

**Examples.**

1) In Minkowski space, in an inertial frame $g_{\mu\nu} = \eta_{\mu\nu}$ so $\Gamma_{\mu}{}^{\nu}{}_{\rho} = 0$ and the geodesic equation is

$$\frac{\mathrm{d}^2 x^{\mu}}{\mathrm{d}\tau^2} = 0, \tag{163}$$

which has geodesics (solutions) which are straight lines.

2) The Schwarzschild metric in Schwarzschild coordinates is a metric on $M = \mathbb{R}_t \times (2m, \infty)_r \times S^2_{\theta,\phi}$ given by

$$\mathrm{d}s^2 = -f\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right), \tag{164}$$

where $f = 1 - \frac{2m}{r}$.

One can then write the Lagrangian as

$$L = f\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \frac{1}{f}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}\tau}\right)^2 - r^2\sin^2\theta\left(\frac{\partial\phi}{\partial t}\right)^2. \tag{165}$$

The Euler-Lagrange equation for $t\left(\tau\right)$ is

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial t'}\right) = \frac{\partial L}{\partial t}, \tag{166}$$

where $t' = \frac{\mathrm{d}t}{\mathrm{d}\tau}$. This gives us

$$2\frac{\mathrm{d}}{\mathrm{d}\tau}\left(f\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) = 0, \tag{167}$$

which implies

$$f\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} + \frac{\mathrm{d}f}{\mathrm{d}r}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) = 0. \tag{168}$$

Comparing this to the geodesic equation Eq. (159), we see

$$\Gamma_{1}{}^{0}{}_{0} = \Gamma_{0}{}^{0}{}_{1} = \frac{1}{2}\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}r}, \tag{169}$$

and $\Gamma_{\mu}{}^{0}{}_{\nu} = 0$ otherwise. The rest of the symbols can be found from the other Euler Lagrange equations.

## 8.2 Covariant Derivative

For a function $f : M \to \mathbb{R}$, we know that

$$\frac{\partial f}{\partial x^\mu} \text{ are the components of a covector } (df)_a. \tag{170}$$

For a vector field we can't just differentiate it's components as the basis vectors themselves can have spatial dependence.

**Exercise 7:** Show that if $V$ is a vector field, then

$$T^\mu{}_\nu := \frac{\partial V^\mu}{\partial x^\nu}, \tag{171}$$

are not the components of a $(1, 1)$ tensor.

**Definition 8.1:** A **covariant derivative** $\nabla$ on a manifold $M$ is a map sending smooth vector fields $X$, $Y$ to a vector field $\nabla_X Y$ satisfying

i) linearity in the first vector such that

$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z, \tag{172}$$

ii) linearity in the second such that

$$\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z, \tag{173}$$

iii) and a Leibniz rule such that

$$\nabla_X (fY) = f\nabla_X Y + (\nabla_X f) Y, \tag{174}$$

where $\nabla_X f = X(f)$ and $X, Y, Z$ are smooth vector fields and $f, g$ are functions.

**Note.** This first condition implies that $\nabla Y : X \mapsto \nabla_X Y$ is a linear map of $T_pM$ to itself and so defines a $(1, 1)$ tensor, which we call the covariant derivative of $Y$.

In abstract index notation, one can write

$$(\nabla Y)^a{}_b = \nabla_b Y^a \text{ or } Y^a{}_{;b}\,. \tag{175}$$

**Definition 8.2:** In a basis $\{e_\mu\}$ the **connection components** $\Gamma_\nu{}^\mu{}_\rho$ are defined by

$$\nabla_{e_\rho} e_\nu = \Gamma_\nu{}^\mu{}_\rho e_\mu. \tag{176}$$

Once we know these connection components, they completely determine $\nabla$. Namely, take

$$\nabla_X Y = \nabla_{X^\mu e_\mu} (Y^\nu e_\nu) \tag{177}$$

$$\overset{\text{i)}}{=} X^\mu \nabla_{e_\mu} (Y^\nu e_\nu) \tag{178}$$

$$\overset{\text{ii) \& iii)}}{=} X^\mu \left(e_\mu (Y^\nu) e_\nu + Y^\sigma \nabla_{e_\mu} e_\sigma\right) \tag{179}$$

$$= \left(X^{\mu} e_{\mu}\left(Y^{\nu}\right) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma} X^{\mu}\right) e_{\nu}. \tag{180}$$

Hence the components of the covariant derivative can be written as

$$\left(\nabla_X Y\right)^{\nu} = X^{\mu}\left(e_{\mu}\left(Y^{\nu}\right) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma}\right), \tag{181}$$

or identically, in different notation,

$$Y^{\nu}{}_{;\mu} = e_{\mu}\left(Y^{\nu}\right) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma} \tag{182}$$

$$= Y^{\nu}{}_{,\mu} + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma}, \tag{183}$$

where $Y^{\nu}{}_{,\mu} = \frac{\partial Y^{\nu}}{\partial x^{\mu}}$.

**Note.** Remember that $\Gamma_{\mu}{}^{\nu}{}_{\sigma}$ are not the components of a tensor, hence we call them *symbols*, like the Levi-Civita symbol $\varepsilon_{\mu\nu\rho\tau}$.

We extend $\nabla$ to arbitrary tensor field by requiring the Leibniz property holds.

**Example.** For a tensor field $\eta$, we define

$$\left(\nabla_X \eta\right)(Y) := \nabla_X\left(\eta(Y)\right) - \eta\left(\nabla_X Y\right). \tag{184}$$

In component form, we can write this as

$$\left(\nabla_X \eta\right) Y = X^{\mu} e_{\mu}\left(\eta_{\sigma} Y^{\sigma}\right) - \eta_{\sigma}\left(\nabla_X Y\right)^{\sigma} \tag{185}$$

$$= X^{\mu} e_{\mu}\left(\eta_{\sigma}\right) Y^{\sigma} + X^{\mu} \eta_{\sigma} e_{\mu}\left(Y^{\sigma}\right) - \eta_{\sigma}\left(X^{\nu} e_{\nu}\left(Y^{\sigma}\right) + X^{\nu} \Gamma_{\tau}{}^{\sigma}{}_{\nu} Y^{\tau}\right) \tag{186}$$

$$= \left(e_{\mu}\left(\eta_{\sigma}\right) - \Gamma_{\sigma}{}^{\nu}{}_{\mu} \eta_{\nu}\right) X^{\mu} Y^{\sigma}, \tag{187}$$

and thus as $\boldsymbol{\nabla}\eta$ is linear in both $X$ and $Y$, it is a (0,2) tensor (it also transforms appropriately). Therefore, with respect to our basis, we have

$$\nabla_{\mu} \eta_{\sigma} = e_{\mu}\left(\eta_{\sigma}\right) = \Gamma_{\sigma}{}^{\nu}{}_{\mu} \eta_{\nu} =: \eta_{\sigma;\mu} \tag{188}$$

$$\Rightarrow \eta_{\sigma;\mu} = \eta_{\sigma,\mu} - \Gamma_{\sigma}{}^{\nu}{}_{\mu} \eta_{\nu}. \tag{189}$$

# 9 Lecture

30/10/2024

**Exercise 8:** In a coordinate basis

$$T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s;\rho} = T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s,\rho} + \Gamma^{\mu_1}{}_{\rho\sigma} T^{\sigma\mu_2 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s} + \cdots + \Gamma^{\mu_r}{}_{\rho\sigma} T^{\mu_1 \cdots \mu_{r-1}\sigma}{}_{\nu_1 \cdots \nu_s} \tag{190}$$

$$- \Gamma_{\nu_1}{}^{\sigma}{}_{\rho} T^{\mu_1 \cdots \mu_r}{}_{\sigma\nu_2 \cdots \nu_s} - \cdots - \Gamma_{\nu_s}{}^{\sigma}{}_{\rho} T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_{s-1}\sigma}. \tag{191}$$

**Remark.** If $T^a{}_b$ is a $(1,1)$ tensor, then $T^a{}_{b;c}$ is a $(1,2)$ tensor and we can take further covariant derivatives,

$$\left(T^a{}_{b;c}\right)_{;d} = T^a{}_{b;cd} = \nabla_d \nabla_c T^a{}_b, \tag{192}$$

In general $T^a{}_{b;cd} \neq T^a{}_{b;dc}$. If $f$ is a function $f_{;a} = (df)_a$ is a covector. In a coordinate basis $f_{;\mu} = f_{,\mu}$ which implies

$$f_{;\mu\nu} = f_{,\mu\nu} - \Gamma_{\mu}{}^{\sigma}{}_{\nu} f_{,\sigma} \tag{193}$$

$$\Rightarrow f_{;[\mu\nu]} = -\Gamma_{[\mu}{}^{\sigma}{}_{\nu]} f_{,\sigma}. \tag{194}$$

**Definition 9.1:** A connection (eq. covariant derivative) is **torsion free** or symmetric if

$$\nabla_a \nabla_b f - \nabla_b \nabla_a f = 0. \tag{195}$$

For any $f$, in a coordinate basis, this is equivalent to

$$\Gamma_{[\mu}{}^{\rho}{}_{\nu]} = 0 \Leftrightarrow \Gamma_{\mu}{}^{\rho}{}_{\nu} = \Gamma_{\nu}{}^{\rho}{}_{\mu}. \tag{196}$$

**Lemma 9.1:** If $\nabla$ is torsion free, then for $X, Y$ vector fields

$$\nabla_X Y - \nabla_Y X = [X, Y]. \tag{197}$$

***Proof.*** In a coordinate basis,

$$\begin{aligned}
(\nabla_X Y - \nabla_Y X)^\mu &= X^\sigma Y^\mu{}_{;\sigma} - Y^\sigma X^\mu{}_{;\sigma} && (198)\\
&= X^\sigma \left(Y^\mu{}_{,\sigma} + \Gamma_{\rho}{}^{\mu}{}_{\sigma} Y^\rho\right) - Y^\sigma \left(X^\mu{}_{,\sigma} + \Gamma_{\rho}{}^{\mu}{}_{\sigma} X^\rho\right) && (199)\\
&= [X, Y]^\mu + 2 X^\sigma Y^\rho \Gamma_{[\rho}{}^{\mu}{}_{\sigma]}. && (200)
\end{aligned}$$

This is a tensor equation so if it is true in one basis, it is true in all. □

**Note.** Even if $\nabla$ is torsion free, $\nabla_a \nabla_b X^c \neq \nabla_b \nabla_a X^c$ in general.

## 9.1 The Levi-Civita Connection

For a manifold with metric, there is a preferred connection.

**Theorem 9.1 (Fundamental Theorem of Riemannian geometry):** If $(M, g)$ is a manifold with a metric, there is a unique torsion free connection $\nabla$ satisfying $\nabla g = 0$. This is called the **Levi-Civita connection**.

***Proof.*** Suppose such a connection exists. By the Leibniz rule, if $X, Y, Z$ are smooth vector fields, then

$$\begin{aligned}
X\left(g\left(Y, Z\right)\right) &= \nabla_X \left(g\left(Y, Z\right)\right) && (201)\\
&= \left(\nabla_X g\right)\left(Y, Z\right) + g\left(\nabla_X Y, Z\right) + g\left(Y, \nabla_X Z\right) && (202)\\
X\left(g\left(Y, Z\right)\right) &= g\left(\nabla_X Y, Z\right) + g\left(Y, \nabla_X Z\right) && (203)\\
\Rightarrow Y\left(g\left(Z, X\right)\right) &= g\left(\nabla_Y Z, X\right) + g\left(Z, \nabla_Y X\right) && (204)\\
\Rightarrow Z\left(g\left(X, Y\right)\right) &= g\left(\nabla_Z X, Y\right) + g\left(X, \nabla_Z Y\right). && (205)
\end{aligned}$$

Taking Eq. (203) + Eq. (204) - Eq. (205),

$$X\left(g\left(Y, Z\right)\right) + Y\left(g\left(Z, X\right)\right) - Z\left(g\left(X, Y\right)\right) = g\left(\nabla_X Y + \nabla_Y X, Z\right) + g\left(\nabla_X Z - \nabla_Z X, Y\right) \tag{206}$$

$$+ g\left(\nabla_X Z - \nabla_Z X, Y\right) + g\left(\nabla_Y Z - \nabla_Z Y, X\right). \tag{207}$$

As $\nabla_X Y - \nabla_Y X = [X, Y]$, this becomes

$$X\left(g\left(Y,Z\right)\right) + Y\left(g\left(Z,X\right) - Z\left(g\left(X,Y\right)\right)\right) = 2g\left(\nabla_X Y, Z\right) - g\left(\left[X,Y\right], Z\right) \tag{208}$$

$$- g\left(\left[Z,X\right], Y\right) + g\left(\left[Y,Z\right], X\right). \tag{209}$$

Therefore

$$g\left(\nabla_X Y, Z\right) = \frac{1}{2}\Bigg(X\left(g\left(y,Z\right)\right) + Y\left(g\left(Z,X\right)\right) - Z\left(g\left(X,Y\right)\right) \tag{210}$$

$$+ g\left(\left[X,Y\right], Z\right) + g\left(\left[Z,X\right] Y\right) - g\left(\left[Y,Z\right], X\right)\Bigg), \tag{211}$$

and therefore $\nabla_X Y$ is uniquely determined since $g$ is non-degenerate and $X, Y$ and $Z$ are general. □

Conversely, we can use this expression to define $\nabla_X Y$. We then need to check the properties of a symmetric connection hold.

**Example.** Observe that

$$g\left(\nabla_{fX} Y, Z\right) = \frac{1}{2}\Bigg(fX\left(g\left(y,Z\right)\right) + Y\left(fg\left(Z,X\right)\right) - Z\left(fg\left(X,Y\right)\right)$$

$$+ g\left(\left[fX,Y\right], Z\right) + g\left(\left[Z,fX\right] Y\right) - g\left(\left[Y,Z\right], fX\right)\Bigg) \tag{212}$$

$$= \frac{1}{2}\Bigg(fXg\left(Y,Z\right) + fY\left(g\left(Z,X\right) - fZ\left(g\left(X,Y\right)\right)\right) + \left(Y\left(f\right)g\left(Z,X\right) - Z\left(f\right)g\left(X,Y\right)\right)$$

$$+ g\left(f\left[X,Y\right] - Y\left(f\right)X, Z\right) + g\left(f\left[Z,X\right] + Z\left(f\right)X, Y\right) - fg\left(\left[Y,Z\right], X\right)\Bigg) \tag{213}$$

$$= g\left(f\nabla_X Y, Z\right), \tag{214}$$

which implies

$$g\left(\nabla_{fX} Y - f\nabla_X Y\right) = 0, \tag{215}$$

$\forall Z$. Therefore $\nabla_{fX} Y = f\nabla_X Y$ as $g$ is non-degenerate.

**Exercise 9:** Check the other properties.

In a coordinate basis, we can compute

$$g\left(\nabla_{e_\mu} e_\nu, e_\sigma\right) = \frac{1}{2}\Bigg(e_\mu\left(g\left(e_\nu, e_\sigma\right)\right) + e_\nu\left(g\left(e_\sigma, e_\mu\right) - e_\sigma\left(g\left(e_\mu, e_\nu\right)\right)\right)\Bigg) \tag{216}$$

$$g\left(\Gamma_\nu{}^\tau{}_\mu e_\tau, e_\sigma\right) = \Gamma_\nu{}^\tau{}_\mu g_{\tau\sigma} = \frac{1}{2}\left(g_{\sigma\nu,\mu} + g_{\sigma\mu,\nu} - g_{\mu\nu,\sigma}\right). \tag{217}$$

This provides

$$\Gamma_\nu{}^\tau{}_\mu = \frac{1}{2} g^{\sigma\tau}\left(g_{\sigma\nu,\mu} + g_{\mu\sigma,\nu} - g_{\mu\nu,\sigma}\right), \tag{218}$$

which is exactly the form of the Christoffel symbols.

Thus if $\nabla$ is a Levi-Civita connection, we can raise/lower indices and it commutes with covariant differentiation.

## 9.2 Geodesics

We found that a curve extremizing proper time $\tau$ satisfies

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma_\nu{}^\mu{}_\rho\left(x\left(\tau\right)\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0. \tag{219}$$

The tangent vector $X^a$ to the curve has components $X^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$, we get a vector field of which the geodesic is an integral curve. We note that

$$\begin{aligned} \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) && (220)\\ &= \frac{\mathrm{d}X^\mu}{\mathrm{d}x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} && (221)\\ &= X^\mu{}_{,\nu}X^\nu. && (222) \end{aligned}$$

Using the geodesic equation, Eq. (219) we have

$$X^\mu{}_{,\nu}X^\nu + \Gamma_\nu{}^\mu{}_\rho X^\nu X^\rho = 0 \Leftrightarrow X^\nu X^\mu{}_{;\nu} = 0 \Leftrightarrow \nabla_X X = 0. \tag{223}$$

We can extend this to any connection.

**Definition 9.2:** Let $M$ be a manifold with connection $\nabla$. An **affinely parameterized geodesic** satisfies

$$\nabla_X X = 0, \tag{224}$$

where $X$ is the tangent vector to the curve.