# III General Relativity

Ishan Nath, Michaelmas 2024

Based on Lectures by Prof. Claude Warnick

November 8, 2024

# Contents

# 0 Introduction

Office hours: 8:40AM MWF, in MR2. Normal room E1.14. Will follow roughly Reall's course.

General relativity is our best theory of gravitation on the largest scales. It is:

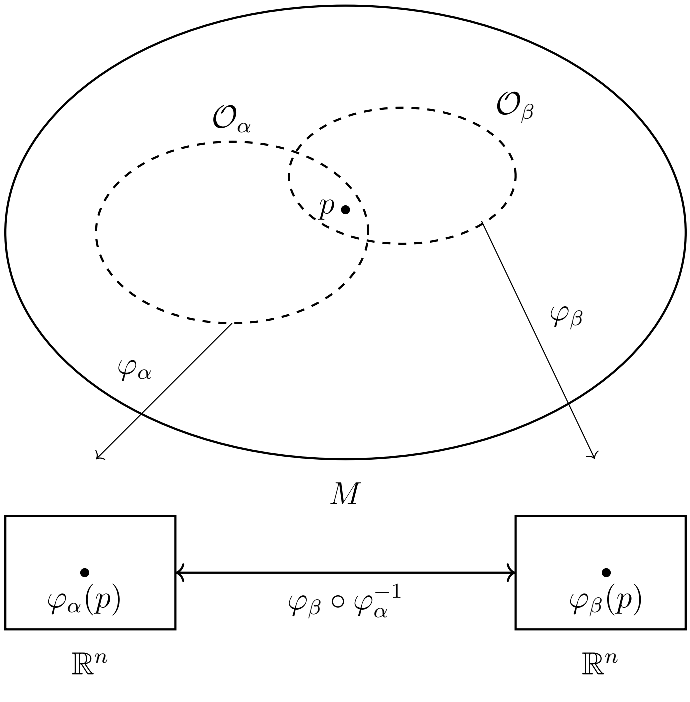
- Classical : No quantum effects.
- Geometrical: Space and time are combined in a curved spacetime.
- Dynamical: In contrast to Newton's theory of gravity, Einstein's gravitational field has its own non-trivial dynamics.

# 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:

- $O_\alpha \subseteq M$ are subsets of $M$ such that
$$\bigcup_\alpha O_\alpha = M.$$
- $\phi_\alpha$ is a bijective map 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 map from $\phi_\alpha(O_\alpha \cap O_\beta) \subseteq U_\alpha$ to $\phi_\beta(O_\alpha \cap O_\beta) \subseteq U_\beta$.




*Remark.*

- We could replace smooth with finite differentiability (e.g. $k$-times differentiable).

- The charts define a topology on $M$: $U \subseteq M$ is open if and only if $\phi_\alpha(U \cap O_\alpha)$ is open in $\mathbb{R}^n$ for all $\alpha$. Every open subset of $M$ is itself a manifold, by restricting the charts to $U$.

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 $B$ which is compatible with $A$, and strictly larger than $A$. Every atlas is contained in a maximal atlas (by taking the union of all compatible atlases). Hence we can assume without loss of generality that we work with a maximal atlas.

**Example 1.1.**

1. If $U \subseteq \mathbb{R}^n$ is open, we can take $O = U$, and $\phi : U \to \mathbb{R}^n$ to be the identity on $U$. Then $\{(O, \phi)\}$ is an atlas.
2. Take $S^1$. If $p \in S^1 \setminus \{(-1, 0)\} = O_1$, there is a unique $\theta_1 \in (-\pi, \pi)$ such that
$$p = (\cos\theta_1, \sin\theta_1).$$
If $p \in S^1 \setminus \{(1, 0)\} = O_2$, there is a unique $\theta_2 \in (0, 2\pi)$ such that
$$p = (\cos\theta_2, \sin\theta_2).$$
These maps from $(-\pi, \pi)$ and $(0, 2\pi)$ to $O_1, O_2$ give $\phi_1^{-1}, \phi_2^{-1}$ respectively. Note that $\phi_1(O_1 \cap O_2) = (-\pi, 0) \cup (0, \pi)$, and the transition function is
$$\phi_2 \circ \phi_1^{-1}(\theta) = \begin{cases} \theta & \theta \in (0, \pi), \\ \theta + 2\pi & \theta \in (-\pi, 0). \end{cases}$$
This is smooth where defined, and similarly for $\phi_2 \circ \phi_1^{-1}$. Hence $S^1$ is a manifold.
3. More generally, we can consider $S^n$, and can define charts by stereographic projections. If $\{\mathbf{E}_1, \ldots, \mathbf{E}_{n+1}\}$ is the standard basis for $\mathbb{R}^{n+1}$, and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$, write
$$\mathbf{P} = P^1\mathbf{E}_1 + \cdots + P^{n+1}\mathbf{E}_{n+1}.$$
Set $O_1 = S^n \setminus \{\mathbf{E}_{n+1}\}$, and write
$$\phi_1(\mathbf{P}) = \frac{1}{1 - P^{n+1}}(P^1\mathbf{e}_1 + \cdots + P^n\mathbf{e}_n).$$

In a similar way we may define $O_2, \phi_2$, for $-\mathbf{E}_{n+1}$. The transition map is then

$$\phi_2 \circ \phi_1^{-1}(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|^2},$$

which is smooth on $\mathbb{R}^n \setminus \{0\} = \phi_1(O_1 \cap O_2)$.

"Nice" surfaces in $\mathbb{R}^n$ are manifolds with no cusps, cornered or self-intersections, for example $S^n \subseteq \mathbb{R}^{n+1}$.

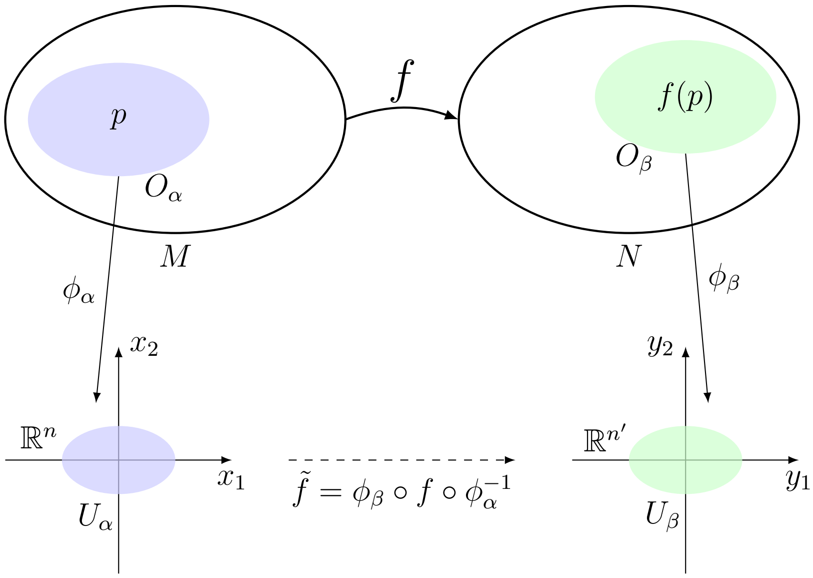
## 1.1 Smooth Functions on Manifolds

Suppose $M, N$ are manifolds of dimension $n, n'$ respectively. Let $f : M \to N$.

Then let $p \in M$, and pick charts $(O_\alpha, \phi_\alpha)$ for $M$, and $(O'_\beta, \phi'_\beta)$ for $N$, with $p \in O_\alpha$ and $f(p) \in O_\beta$. Then,

$$\phi'_\beta \circ f \circ \phi_\alpha^{-1}$$

maps a neighbourhood of $\phi_\alpha(p)$ in $U_\alpha \subseteq \mathbb{R}^n$, to $U'_\beta \in \mathbb{R}^{n'}$. If this function is smooth for all possible choices of this chart, we say that $f : M \to N$ is smooth.




*Remark.*

- A smooth map $\psi : M \to N$ which has a smooth inverse is called a *diffeomorphism*.
- If $N = \mathbb{R}$ or $\mathbb{C}$, we sometimes call $f$ a *scalar field*.
- If $M = I \subseteq \mathbb{R}$, an open interval, then $f : I \to N$ is a *smooth curve* in $N$.

- If $f$ is smooth in one atlas, then it is smooth in all compatible atlases.

**Example 1.2.**

1. Recall $S^1$. Let $f(x, y) = x$, $f : S^1 \to \mathbb{R}$. Using our previous charts,

$$\begin{aligned} f \circ \phi_1^{-1} : (-\pi, \pi) &\to \mathbb{R} \\ \theta_1 &\mapsto \cos\theta_1. \end{aligned}$$

Similarly,

$$\begin{aligned} f \circ \phi_2^{-1} : (0, 2\pi) &\to \mathbb{R} \\ \theta_2 &\mapsto \cos\theta_2. \end{aligned}$$

Hence $f$ is smooth.

2. If $(O, \phi)$ is a coordinate chart on $M$, write

$$\phi(p) = (x^1(p), x^2(p), \dots, x^n(p)),$$

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

3. We can define a smooth function chart-by-chart. For simplicity, let $N = \mathbb{R}$, and let $\{(O_\alpha, \phi_\alpha)\}$ be an atlas on $M$. Define smooth functions

$$F_\alpha : U_\alpha \to \mathbb{R},$$

and suppose that $F_\alpha \circ \phi_\alpha = F_\beta \circ \phi_\beta$ on $O_\alpha \cap O_\beta$, for all $\alpha, \beta$.

Then for $p \in M$, we can define

$$f(p) = F_\alpha \circ \phi_\alpha(p),$$

for any chart $(O_\alpha, \phi_\alpha)$ with $p \in O_\alpha$. Now $f$ is smooth, as

$$f \circ \phi_p^{-1} = F_\alpha \circ \phi_\alpha \circ \phi_\beta^{-1}.$$

In practice, we often don't distinguish between $f$ and its coordinate chart representations $F_\alpha$.

## 1.2 Curves and Vectors

For a surface in $\mathbb{R}^3$, we have a notion of the 'tangent space' at a point $p$, which consists of all vectors tangent to the surface at that point.

The 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 $\lambda : I \to M$ a smooth map, 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, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t}[f(\lambda(t))] = X(t) \cdot \nabla f(\lambda(t)),$$

where $X(t) = \mathrm{d}\lambda/\,\mathrm{d}t(t)$ is the tangent vector to $\lambda$ at $t$.

**Definition 1.2.** Let $\lambda : I \to M$ be a smooth curve with $\lambda(0) = p$. The *tangent vector* to $\lambda$ at $p$ is the linear map $X_p$ from the space of smooth functions $f : M \to \mathbb{R}$, given by

$$X_p(f) = \frac{\mathrm{d}}{\mathrm{d}t} f(\lambda(t))\bigg|_{t=0}.$$

We observe that:

(i) $X_p$ is linear: $X_p(f + ag) = X_p(f) + aX_p(g)$, for $f, g$ smooth, and $a \in \mathbb{R}$.

(ii) $X_p$ satisfies the Leibniz rule:

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g).$$

If $(O, \phi)$ is a chart and $p \in O$, write

$$\phi(p) = (x^1(p), \ldots, x^n(p)).$$

Let $F = f \circ \phi^{-1}$, and $x^i(t) = x^i(\lambda(t))$. Now $F : \mathbb{R}^n \to \mathbb{R}$, and $x : \mathbb{R} \to \mathbb{R}^n$. Then,

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

and by applying the regular chain rule for functions $\mathbb{R}^m \to \mathbb{R}^n$,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\lambda(t))\bigg|_{t=0} = \frac{\partial F}{\partial x^\mu}(x) \cdot \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\bigg|_{t=0}.$$

**Proposition 1.1.** *The set of tangent vectors to curves at $p$ forms a vector space, $T_pM$, of dimension $n = \dim M$. We call $T_pM$ the* tangent space *to $M$ at $p$.*

**Proof:** Given $X_p, Y_p$ tangent vectors, we need to show that $\alpha X_p + \beta Y_p$ is a tangent vector for $\alpha, \beta \in \mathbb{R}$.

Let $\lambda, \kappa$ be smooth curves with $\lambda(0) = \kappa(0) = p$, and whose tangent vectors at $p$ are $X_p, Y_p$, respectively.

Let $(O, \phi)$ be a chart with $p \in O$ and $\phi(p) = 0$, and define

$$\gamma(t) = \phi^{-1}(\alpha\phi(\lambda(t)) + \beta\phi(\kappa(t))).$$

This exists, as this is just the sum of elements in $\mathbb{R}^n$.

Now $\gamma(0) = \phi^{-1}(0) = p$. If $Z_p$ is the tangent to $v$ at $p$, then

$$\begin{aligned} Z_p(f) = \frac{\mathrm{d}}{\mathrm{d}t} f(v(t))\bigg|_{t=0} &= \frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t}[\alpha x^\mu(\lambda(t)) + \beta x^\mu(\kappa(t))]\bigg|_{t=0} \\ &= \alpha \frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\bigg|_{t=0} + \beta \frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\kappa(t))\bigg|_{t=0} \\ &= \alpha X_p(f) + \beta Y_p(f). \end{aligned}$$

Thus $T_pM$ is a vector space.

To see that $T_pM$ is $n$-dimensional, consider the curves

$$\lambda_\mu(t) = \phi^{-1}\underbrace{(0, \ldots, 0, t, 0, \ldots, 0)}_{\mu\text{'th component}}.$$

We denote the tangent vector to $\lambda_\mu$ at $p$ by $(\partial/\partial x^\mu)_p$. To see why, note that

$$\left(\frac{\partial}{\partial x^\mu}\right)_p f = \frac{\partial F}{\partial x^\mu}\bigg|_{\phi(p)=0}.$$

These vectors are linearly independent. Otherwise there exists $\alpha^\mu$ not all zero such that

$$\alpha^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p = 0 \implies \alpha^\mu \frac{\partial F}{\partial x^\mu}\bigg|_0 = 0,$$

for all $F$. Setting $F = x^\nu$ gives $\alpha^\nu = 0$.

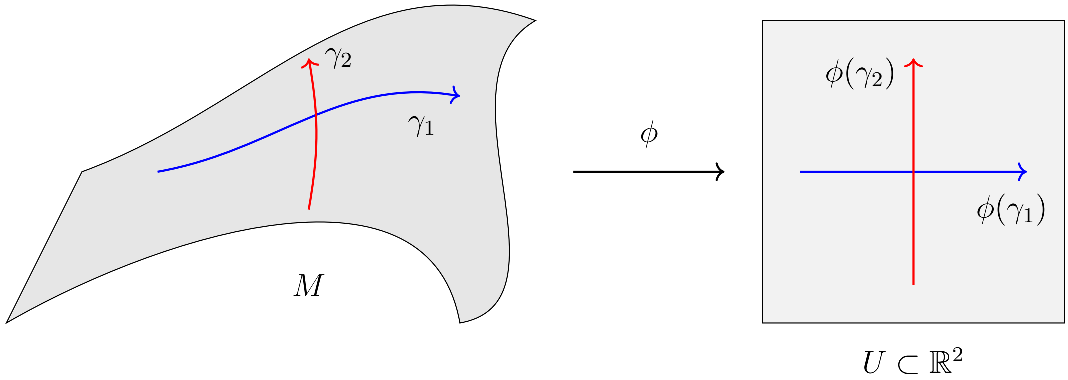
Further, these vectors form a basis for $T_pM$, since if $\lambda$ is any curve with tangent $X_p$ at $p$, then

$$X_p(f) = \frac{\partial F}{\partial x^\mu}\bigg|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\bigg|_{t=0} = X^\mu \left(\frac{\partial}{\partial x^\mu}\right) f.$$

Here, we set

$$X^\mu = \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\bigg|_{t=0}.$$

These are the components of $X_p$ with respect to the basis.




Notice that the basis $\{(\partial/\partial x^\mu)_p\}$ depends on the coordinate chart $\phi$. Suppose we choose another chart $(O', \phi')$, again centred at $p$. Write $\phi' = (x'^1, \dots, x'^n)$.

Then if $F = f \circ (\phi')^{-1}$, then

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

Therefore,

$$\left(\frac{\partial}{\partial x^\mu}\right)_p f = \frac{\partial F}{\partial x^\mu}\bigg|_{\phi(p)} = \left(\frac{\partial x'^\nu}{\partial x^\mu}\right)_{\phi(p)} \left(\frac{\partial F'}{\partial x'^\nu}\right)_{\phi'(p)} = \left(\frac{\partial x'^\nu}{\partial x^\mu}\right)_{\phi(p)} \cdot \left(\frac{\partial}{\partial x'^\nu}\right)_p f.$$

We deduce that

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

So let $X^\mu$ be the components of $X_p$ with respect to $\{(\partial/\partial x^\mu)_p\}$, and $X'^\mu$ be the components with respect to $\{(\partial/\partial x'^\mu)_p\}$. So,

$$\begin{aligned} X_p = X^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p &= X'^\mu \left(\frac{\partial}{\partial x'^\mu}\right)_p \\ &= X^\mu \left(\frac{\partial x'^\sigma}{\partial x^\mu}\right)_{\phi(p)} \left(\frac{\partial}{\partial x'^\sigma}\right)_p. \end{aligned}$$

Therefore,

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

We do not have to choose a coordinate basis such as $\{(\partial/\partial x^{\mu})_p\}$. With respect to a general basis $\{e_\mu\}$ for $T_pM$, we write $X_p = X^\mu e_\mu$, for $X^\mu \in \mathbb{R}$ the components.

We always use summation conventions: we always contract one upstairs and one downstairs index.

## 1.3 Covectors

Recall that if $V$ is a vector space over $\mathbb{R}$, the *dual space* $V^*$ is the space of linear maps from $V$ to $\mathbb{R}$. If $V$ is $n$-dimensional, then so is $V^*$.

Given a basis $\{e_\mu\}$ for $V$, we define the *dual basis* $\{f^\mu\}$ for $V^*$ by requiring that

$$f^\mu(e_\nu) = \delta^\mu{}_\nu.$$

If $V$ is finite dimensional, then $V^{**} = (V^*)^*$ is isomorphic to $V$: to an element $x$ of $V$, we assign a linear map $\Lambda_x : V^* \to \mathbb{R}$ by

$$\Lambda_x(\omega) = \omega(x).$$

**Definition 1.3.** The dual space of $T_pM$ is denoted $T_p^*M$, and is called the *cotangent space* to $M$ at $p$.

An element of this space is a covector at $p$. If $\{e_\mu\}$ is a basis for $T_pM$ and $\{f^\mu\}$ the dual basis for $T_p^*M$, then we can expand a covector $\eta$ as

$$\eta = \eta_\mu f^\mu,$$

for $\eta_\mu \in \mathbb{R}$ the components of $\eta$.

Note that:

- $\eta(e_\nu) = \eta_\mu f^\mu(e_\nu) = \eta_\mu \delta^\mu{}_\nu = \eta_\nu$.
- $\eta(X) = \eta(X^\mu e_\mu) = X^\mu \eta(e_\mu) = X^\mu \eta_\mu$.

**Definition 1.4.** If $f : M \to \mathbb{R}$ is a smooth function, define $(\mathrm{d}f)_p \in T_p^*M$, the *differential* of $f$ at $p$, by

$$(\mathrm{d}f)_p(X) = X(f),$$

for any $X \in T_pM$. $(\mathrm{d}f)_p$ is sometimes also called the *gradient* of $f$ at $p$.

If $f$ is a constant, then $X(f) = 0 \implies (\mathrm{d}f)_p = 0$.

If $(O, \phi)$ is a coordinate chart with $p \in O$ and $\phi = (x^1, \ldots, x^n)$, then we can set $f = x^\mu$ to find $(\mathrm{d}x^\mu)_p$. Now,

$$(\mathrm{d}x^\mu)_p \left(\frac{\partial}{\partial x^\nu}\right)_p = \left(\frac{\partial x^\mu}{\partial x^\nu}\right)_{\phi(p)} = \delta^\mu{}_\nu.$$

Hence, $\{(\mathrm{d}x^\mu)_p\}$ is the dual basis to $\{(\partial/\partial x^\mu)_p\}$. In this basis, we can compute

$$[(\mathrm{d}f)_p]_\mu = (\mathrm{d}f)_p \left(\frac{\partial}{\partial x^\mu}\right)_p = \left(\frac{\partial}{\partial x^\mu}\right)_p f = \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)},$$

justifying the language of the 'gradient'.

It can be shown that if $(O', \phi')$ is another chart with $p \in O'$, then

$$(\mathrm{d}x^\mu)_p = \left(\frac{\partial x^\mu}{\partial x'^\nu}\right)_{\phi'(p)} (\mathrm{d}x'^\nu)_p,$$

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

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

## 1.4 The (Co)tangent bundle

We can glue together the tangent spaces $T_pM$ as $p$ varies to get a new $2n$ dimensional manifold, $TM$, the *tangent bundle*. We let

$$TM = \bigcup_{p \in M} \{p\} \times T_pM,$$

the set of ordered pairs $(p, X)$, with $p \in M$, $X \in T_pM$. If $\{(O_\alpha, \phi_\alpha)\}$ is an atlas on $M$, we obtain an atlas for $TM$ by setting

$$\tilde{O}_\alpha = \bigcup_{p \in O_\alpha} \{p\} \times T_pM,$$

and

$$\tilde{\phi}_\alpha((p, X)) = (\phi_\alpha(p), X^\mu) \in U_\alpha \times \mathbb{R}^n = \tilde{U}_\alpha,$$

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

It can be shown that if $(O, \phi)$ and $(O', \phi')$ are two charts on $M$, on $\tilde{U} \cap \tilde{U}'$, if we write $\phi' \circ \phi^{-1}(x) = x'(x)$, then

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

This lets us deduce that $TM$ is a manifold.

A similar construction permits us to define the *cotangent bundle*

$$T^*M = \bigcup_{p \in M} \{p\} \times T_p^* M.$$

The map $\pi : TM \to M$ which takes $(p, X) \mapsto p$ is smooth (show this!)

## 1.5 Abstract Index Notation

We have used Greek letters $\mu, \nu$ to label components of vectors (or covectors) with respect to the bases $\{e_\mu\}$. Equations involving these quantities refer to the specific basis, for example if we write $X^\mu = \delta^\mu$, this says $X$ only has one non-zero component *in the current basis*, which will not be true in other bases.

However, we know that some equations hold in all bases, e.g.

$$\eta(X) = X^\mu \eta_\mu.$$

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

If an equation is true in all bases, then we can replace Greek indices by latin indices:

$$\eta(X) = X^a \eta_a = \eta_a X^a,$$

or

$$X(f) = X^a (\mathrm{d}f)_a.$$

An equation in abstract index notation can always be turned into an equation for components, by picking a basis and changing $a \to \mu$, $b \to \nu$.

## 1.6 Tensors

In Newtonian physics, we know that some quantities are described by higher rank orders, e.g. the inertia tensor or the metric.

**Definition 1.5.** A *tensor* of type $(r, s)$ is a multilinear map

$$T : (T_p^* M)^r \times (T_p M)^s \to \mathbb{R}.$$

**Example 1.3.**

1. A tensor of type $(0,1)$ is a linear map $T_pM \to \mathbb{R}$, i.e. just a covector.
2. A tensor of type $(1,0)$ is a linear map $T^*pM \to \mathbb{R}$, i.e. an element of $(T^*pM)^* \cong T_pM$, a vector.
3. We can define a $(1,1)$ tensor $\delta$ by

$$\delta(\omega, X) = \omega(X),$$

where $\omega \in T_p^*M$, $X \in T_pM$.

If $\{e_\mu\}$ is a basis for $T_pM$ and $\{f^\mu\}$ its dual basis, then the components of an $(r,s)$ tensor $T$ are

$$T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} = T(f_1^\mu, \ldots, f^{\mu_r}, e_{\nu_1}, \ldots, e_{\nu_s}).$$

In abstract index notation, we denote $T$ by $T^{a_1\cdots a_r}{}_{b_1\cdots b_s}$. Tensors at $p$ form a vector space over $\mathbb{R}$ of dimension $n^{r+s}$.

**Example 1.4.**

1. Consider $\delta$ above. Then,

$$\delta^\mu{}_\nu = \delta(f^\mu, e_\nu) = f^\mu(e_\nu) = \begin{cases} 1 & \mu = \nu, \\ 0 & \mu \neq \nu. \end{cases}$$

We can write $\delta$ as $\delta^a{}_b$ in AIN.

2. Consider a $(2,1)$ tensor $T$, and let $\omega, \eta \in T_p^*M$, $X \in T_pM$. Then,

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

In AIN,

$$T(\omega, \nu, X) = \omega_a \eta_b X^c T^{ab}{}_c.$$

This can be generalized to higher ranks.

## 1.7 Change of Bases

We have already seen how the components of $X$ or $\eta$ with respect to a coordinate basis ($X^\mu, \eta_\nu$ respectively) change under a change of coordinates.

But we do not only have to consider coordinate bases.

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

As these are bases, we can expand

$$f'^\mu = A^\mu{}_\nu f^\nu, \qquad e'_\mu = B^\nu{}_\mu e_\nu,$$

for some $A^\mu{}_\nu, B^\mu{}_\nu \in \mathbb{R}$. But,

$$\begin{aligned} \delta^\mu{}_\nu &= f'^\mu(e'_\nu) = A^\mu{}_\tau f^\tau(B^\sigma{}_\nu e_\sigma) \\ &= A^\mu{}_\tau B^\sigma{}_\nu f^\tau(e_\sigma) = A^\mu{}_\tau B^\sigma{}_\nu \delta^\tau{}_\sigma \\ &= A^\mu{}_\sigma B^\sigma{}_\nu. \end{aligned}$$

Therefore, looking at these as matrices,

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

If

$$e_\mu = \left(\frac{\partial}{\partial x^\mu}\right)_p, \qquad e'_\mu = \left(\frac{\partial}{\partial x'^\mu}\right)_p,$$

then we have already seen

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

which indeed satisfies $A^\mu{}_\sigma B^\sigma{}_\nu = \delta^\mu{}_\nu$, by the chain rule.

A chance of bases induces a transformation of tensor components, for example if $T$ is a $(1,1)$-tensor, then

$$\begin{aligned} T^\mu{}_\nu &= T(f^\mu, e_\nu) \\ T'^\mu{}_\nu &= T(f'^\mu, e'_\nu) = T(A^\mu{}_\sigma f^\sigma, (A^{-1})^\tau{}_\nu e_\tau) \\ &= A^\mu{}_\sigma (A^{-1})^\tau{}_\nu T(f^\sigma, e_\tau) \\ &= A^\mu{}_\sigma (A^{-1})^\tau{}_\nu T^\sigma{}_\tau. \end{aligned}$$

## 1.8 Tensor Operations

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

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

$$S(\omega, X) = T(\omega, f^\mu, X, e_\mu).$$

To see that this is independent of the choice of basis, note

$$\begin{aligned} T(\omega, f'^{\mu}, X, e'_{\mu}) &= T(\omega, A^{\mu}{}_{\sigma} f^{\sigma}, X, (A^{-1})^{\tau}{}_{\mu} e_{\tau}) = A^{\mu}{}_{\sigma}(A^{-1})^{\tau}{}_{\mu} T(\omega, f^{\sigma}, X, e_{\tau}) \\ &= \delta^{\tau}{}_{\sigma} T(\omega, f^{\sigma}, X, e_{\tau}) = T(\omega, f^{\tau}, X, e_{\tau}) = S(\omega, X). \end{aligned}$$

So this does not depend on the choice of basis. $S$ and $T$ have components related by

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

In any basis in AIN we write

$$S^{a}{}_{b} = T^{ac}{}_{bc}\,.$$

We can generalise to contract any upstairs index with any downstairs index in a general $(r, s)$-tensor.

Another way to make new tensors from old tensors is to form the *tensor product*. If $S$ is a $(p, q)$-tensor and $T$ is a $(r, s)$-tensor, then $S \otimes T$ is an $(p+r, q+s)$-tensor:

$$\begin{aligned} S{\otimes}T(&\omega^1, \dots, \omega^p, \eta^1, \dots, \eta^r, X_1, \dots, X_q, Y_1, \dots, Y_s) \\ &= S(\omega^1, \dots, \omega^p, X_1, \dots, X_q) T(\eta^1, \dots, \eta^r, Y_1, \dots, Y_s). \end{aligned}$$

This is independent of basis. In AIN,

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

We can show that for any $(1, 1)$-tensor $T$, in any basis we have

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

The final tensor operations we require are *(anti)symmetrization*. If $T$ is a $(0, 2)$-tensor, we can define two new tensors:

$$\begin{aligned} S(X, Y) &= \frac{1}{2}(T(X, Y) + T(Y, X)), \\ A(X, Y) &= \frac{1}{2}(T(X, Y) - T(Y, X)). \end{aligned}$$

In AIN,

$$\begin{aligned} S_{ab} &= \frac{1}{2}(T_{ab} + T_{ba}) = T_{(ab)}, \\ A_{ab} &= \frac{1}{2}(T_{ab} - T_{ba}) = T_{[ab]}. \end{aligned}$$

These operations can be applied to any pair of matching symmetries in a more general tensor, for example:

$$T^{a(bc)}{}_{de} = \frac{1}{2}(T^{abc}{}_{de} + T^{acb}{}_{de}).$$

We can also (anti)symmetrize over more than two indices.

- To symmetrize over $n$ indices, we sum over all permutations of the indices and divide by $n!$.
- To anti-symmetrize over $n$ indices, we sum over all permutation weighted by their sign, and then divide by $n!$.

For example,

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

To exclude indices from (anti)symmetrization, we use vertical lines:

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

## 1.9 Tensor Bundles

The space of $(r,s)$-tensors at a point $p$ is the vector space $(T^r{}_s)_pM$. These can be glued together to form the bundle of $(r,s)$-tensors

$$T^r{}_s M = \bigcup_{p\in M}\{p\} \times (T^r{}_s)_pM.$$

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

$$\tilde{O} = \bigcup_{p\in O}\{p\} \times (T^r{}_s)_pM \subseteq T^r{}_s M,$$

and

$$\tilde{\phi}(p, S_p) = (\phi(p), S^{\mu_1\dots\mu_r}{}_{\nu_1\dots\nu_s}).$$

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

A *tensor field* is a smooth map $T : M \to T^r{}_s M$ such that $\pi \circ T = \text{id}$.

If $(O,\phi)$ is a coordinate chart on $M$, then

$$\tilde{\phi} \circ T \circ \phi(x) = (x, T^{\mu_1\dots\mu_r}{}_{\nu_1\dots\nu_s}(x)),$$

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

**Example 1.5.**

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

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

In particular, $\partial/\partial x^\mu$ are always smooth, but only defined locally.

A vector field can act on a function $f : M \to \mathbb{R}$ to give a new function $Xf$ by

$$Xf(p) = X_p(f).$$

In coordinates,

$$Xf(p) = X^\mu(\phi(p)) \frac{\partial F}{\partial x^\mu}\bigg|_{\phi(p)},$$

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

## 1.10 Integral Curves

Given a vector field $X$ on $M$, we say a curve $\lambda : I \to M$ is an *integral curve* of $X$ if its tangent at every point $x$, which we denote by $\mathrm{d}\lambda/\,\mathrm{d}t$, then

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) = X_{\lambda(t)}.$$

Through each point $p$, there is a unique integral curve passing through it, unique up to extension or shift of the parameter.

To see this, pick a chart $\phi$ with $\phi = (x^1, \ldots, x^n)$, and $\phi(p) = 0$. In this chart, the equation becomes

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

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

By standard ODE theory, this equation with initial conditions has a solution which is unique up to extension.

## 1.11 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, however it is not of the form $K(f)$ for some vector field $K$, since

$$\begin{aligned} X(Y(fg)) &= X(fY(g) + gY(f)) = X(f(Y(g)) + X(gY(f)) \\ &= fX(Y(g)) + gX(Y(f)) + X(f)Y(g) + X(g)Y(f). \end{aligned}$$

So the Leibniz rule does not hold. But, if we consider $[X,Y](f) = X(Y(f)) - Y(X(f))$, then the Leibniz rule does hold. In fact, $[X,Y]$ defines a vector field.

To see this, we can use coordinates:

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

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,

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

is valid only in a coordinate basis.

## 1.12 Metric Tensor

We are familiar from Euclidean geometry (and special relativity) with the fact that the fundamental object when talking about distances and angles (or time intervals and rapidity) is an inner product between vectors. For example,

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

in $\mathbb{R}^3$ with Euclidean geometry, and

$$X \cdot Y = -X^0Y^0 + X^1Y^1 + X^2Y^2 + X^3Y^3$$

for $\mathbb{R}^{3+1}$ with the Minkowski geometry (and $c = 1$).

**Definition 1.6.** A *metric tensor* at $p \in M$ is a $(0,2)$-tensor satisfying:

(i) $g$ is symmetric, so $g(X,Y) = g(Y,X)$ for all $X, Y \in T_pM$. In coordinates, $g_{ab} = g_{ba}$.

(ii) $g$ is *non-degenerate*: $g(X,Y) = 0$ for all $Y \in T_pM$ if and only if $X = 0$.

We sometimes 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 $T_pM$ such that

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

The number of $-1$'s and 1's does not depend on the choice of basis, by Sylvester's law of inertia, and is called the *signature*.

If $g$ has signature that is entirely positive, we say it is *Riemannian*.

If $g$ has signature that is positive apart from one component, we say it is *Lorentzian*.

**Definition 1.7.** A *Riemannian* (resp. *Lorentzian*) manifold is a pair $(M, g)$ where $M$ is a manifold, and $g$ is a Riemannian (resp. Lorentzian) metric tensor field.

On a Riemannian manifold, the *norm* of a vector $X \in T_pM$ is

$$|X| = \sqrt{g(X,X)},$$

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

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

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

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

It is an easy exercise to show that if $\tau : (c,d) \to (a,b)$ is a reparametrization with $\mathrm{d}\tau/\mathrm{d}u > 0$, and $\tau(c) = a$, $\tau(d) = b$, then $\tilde{\lambda} = \lambda \circ \tau : (c,d) \to M$ is a reparametrization of $\lambda$, in which $\ell(\tilde{\lambda}) = \ell(\lambda)$.

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

$$\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \frac{1}{2}(\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu + \mathrm{d}x^\nu \otimes \mathrm{d}x^\mu).$$

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

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

**Example 1.6.**

(i) Consider $\mathbb{R}^n$ with

$$g = \mathrm{d}s^2 = (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^n)^2 + \delta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu.$$

This is called Euclidean space, and any chart covering $\mathbb{R}^n$ in which the metric takes this form is called *Cartesian*.

(ii) Consider $\mathbb{R}^{1+3}$ with

$$g = \mathrm{d}s^2 = -(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu.$$

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

(iii) On $S^2 = \{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x}| = 1\}$, define a chart by $\phi^{-1} : (0,\pi) \times (-\pi,\pi) \to S^2$ by

$$(\theta,\varphi) \mapsto (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta).$$

In this chart, the round metric is

$$g = \mathrm{d}s^2 = \mathrm{d}\theta^2 + \cos^2\theta\,\mathrm{d}\varphi^2.$$

This covers $S^2 \setminus \{\mathbf{x}| = 1, x^2 = 0, x^1 \leq 0\}$.

To cover the rest, let $\tilde{\phi}^{-1}$ be given by

$$(\theta',\varphi') \mapsto (-\sin\theta'\cos\varphi', \cos\theta', \sin\theta'\sin\varphi').$$

Setting $g = \mathrm{d}\theta'^2 + \sin^2\theta'\,\mathrm{d}\varphi'^2$ 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 the inverse define a symmetric $(2,0)$-tensor $g^{ab}$ satisfying

$$g^{ab}g_{bc} = \delta^a{}_c.$$

**Example 1.7.**

In the $\phi$ coordinates of the $S^2$ example, note

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

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 $(\mathbb{R}^3, \delta)$ we often do this without realizing. More generally, this allows us to raise tensor indices with $g^{ab}$ and lower with $g_{ab}$. For example if $T^{ab}{}_c$ is $(2,1)$-tensor, then $T_a{}^{bc}$ is the $(2,1)$-tensor given by

$$T_a{}^{bc} = g_{ad}g^{ce}T^{db}{}_e.$$

## 1.13 Lorentzian Signature

In Lorentzian signature, with indices $0, 1, \ldots, n$. At any point in a Lorentzian manifold we can find a basis $\{e_\mu\}$ such that

$$g(e_\mu, e_\nu) = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, \ldots, 1).$$

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

$$\eta_{\mu\nu} = g(e'_\mu, e'_\nu) = (A^{-1})^\sigma{}_\mu(A^{-1})^\tau{}_\nu g(e_\sigma, e_\tau) = (A^{-1})^\sigma{}_\mu(A^{-1})^\tau{}_\nu\eta_{\sigma\tau}.$$

This is the condition that $A$ is a *Lorentz transformation*.

The tangent space at $p$ has $\eta_{\mu\nu}$ as a metric tensor, so it has the structure of Minkowski space.

**Definition 1.8.** $X \in T_pM$ is:

- *spacelike* if $g(X, X) > 0$,
- *null* or *lightlike* if $g(X, X) = 0$,
- *timelike* if $g(X, X) < 0$.

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

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

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 parametrized by proper time. In this case, we call the tangent vector $U^a = \frac{\mathrm{d}\lambda^a}{\mathrm{d}x}$ the *four-velocity* of $\lambda$.

## 1.14 Curves of Extremal Proper Time

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

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

with

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

The Euler-Lagrange equation is

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

We can compute

$$\begin{aligned}
\frac{\partial G}{\partial \dot{x}^\mu} &= -\frac{1}{G}g_{\mu\nu}\dot{x}^\nu,\\
\frac{\partial G}{\partial x^\mu} &= -\frac{1}{2G}\frac{\partial}{\partial x^\mu}(g_{\mu\nu})\dot{x}^\sigma\dot{x}^\tau\\
&= -\frac{1}{2G}g_{\sigma\tau,\mu}\dot{x}^\sigma\dot{x}^\tau.
\end{aligned}$$

Fix the parametrization so that the curve is parametrized by proper time $\tau$. Doing this, we get

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \dot{x}^\mu\frac{\mathrm{d}u}{\mathrm{d}\tau}, \qquad -1 = g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}.$$

We deduce that

$$-1 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \left(\frac{\mathrm{d}u}{\mathrm{d}\tau}\right)^2 \implies \frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{1}{G}$$
$$\implies \frac{1}{G}\frac{\mathrm{d}}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}\tau}.$$

Returning to the Euler-Lagrange quations, we find

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

Hence expanding the LHS,

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

after symmetrizing. Thus,

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

where

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

are the *Christoffel symbols* of $g$. Note that:

- $\Gamma_{\nu\ \rho}^{\ \mu} = \Gamma_{\rho\ \nu}^{\ \mu}$.
- $\Gamma_{\nu\ \rho}^{\ \mu}$ are not tensor components.
- We can solve ($*$) with standard ODE theory, to get *geodesics*.
- The same equation governs curves of extremal length in a Riemannian manifold, or spacelike curves in a Lorentzian manifold, parametrized by arc-length.

We can show that the geodesic equation can be obtained as the Euler-Lagrange equation for the Lagrangian

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

**Example 1.8.**

1. In Minkowski space in an inertial frame, $g_{\mu\nu} = \eta_{\mu\nu}$, so $\Gamma_{\mu\ \rho}^{\ \nu} = 0$, and

the geodesic equation is

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0.$$

These are straight lines.

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

$$\mathrm{d}s^2 = -f\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2),$$

where $f = 1 - 2m/r$. Then

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

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

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

where $t' = \mathrm{d}t/\,\mathrm{d}\tau$. This gives

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

or

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

We can compare this with the geodesic equation to see that

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

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

## 1.15 Covariant Derivative

For a function $f : M \to \mathbb{R}$, we know that $\partial f/\partial x^\mu$ are components of a covector ($\mathrm{d}f$). For a vector field, we cannot just differentiate the components.

Indeed, we can show that if $V$ is a vector field, then

$$T^{\mu}{}_{\nu} = \frac{\partial V^{\mu}}{\partial x^{\nu}}$$

do not form the components of a $(1,1)$-tensor.

**Definition 1.9.** A *covariant derivative* $\nabla$ on a manifold $M$ is a map sending $X, Y$, smooth vector fields, to a vector field $\nabla_X Y$ satisfying, for $X, Y, Z$ smooth vector fields and $f, g$ functions:

(i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.

(ii) $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$.

(iii) $\nabla_X(fY) = f\nabla_X Y + (\nabla_X f)Y$, where $\nabla_X f = X(f)$.

Here a *smooth vector field* is a map $X : C^{\infty}(M) \to C^{\infty}(M)$ that satisfies the Leibniz rule.

Note the first definition implies $\nabla Y : X \mapsto \nabla_X Y$ is a linear map of $T_pM$ to itself, so it defines a $(1,1)$-tensor, the *covariant derivative* of $Y$. In AIN,

$$(\nabla Y)^{a}{}_{b} = \nabla_b Y^a,$$

or $Y^{a}{}_{;b}$ .

**Definition 1.10.** In a basis $\{e_\mu\}$, the *connection components* $\Gamma_{\nu}{}^{\mu}{}_{\rho}$ are defined by

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

These define $\nabla$, as

$$\begin{aligned}
\nabla_X Y = \nabla_{X^\mu e_\mu}(Y^\nu e_\nu) &= X^\mu \nabla_{e_\mu}(Y^\nu e_\nu) \\
&= X^\mu (e_\mu(Y^\nu) e_\nu + Y^\sigma \nabla_{e_\mu} e_\sigma) \\
&= (X^\mu e_\mu(Y^\nu) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^\sigma X^\mu) e_\nu.
\end{aligned}$$

Hence,

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

and so

$$Y^{\nu}{}_{;\mu} = e_\mu(Y^\nu) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^\sigma.$$

In the coordinate bases $e_\mu = \partial/\partial x^\mu$, then

$$Y^{\nu}{}_{;\mu} = \frac{\partial Y^\nu}{\partial x^\mu} + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^\sigma = Y^{\nu}{}_{,\mu} + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^\sigma.$$

Again note ${\Gamma_\mu}^\nu{}_\sigma$ are not components of a tensor.

We can extend $\nabla$ to arbitrary tensor field by requiring that the Leibniz property holds, for example for $\eta$ a tensor field, we define $\nabla_X \eta$ by

$$(\nabla_X \eta)(Y) = \nabla_X(\eta(Y)) - \eta(\nabla_X Y).$$

In components,

$$\begin{aligned}
(\nabla_X \eta)Y &= X^\mu e_\mu(\eta_\sigma Y^\sigma) - \eta_\sigma(\nabla_X Y)^\sigma \\
&= X^\mu e_\mu(\eta_\sigma) Y^\sigma + X^\mu \eta_\sigma e_\mu(Y^\sigma) - \eta_\sigma(X^\nu e_\nu(Y^\sigma) + X^\nu {\Gamma_\tau}^\sigma{}_\nu Y^\tau) \\
&= (e_\mu(\eta_\sigma) - {\Gamma_\sigma}^\nu{}_\mu \eta_\nu) X^\mu Y^\sigma,
\end{aligned}$$

which is linear in $X$ and $Y$. Hence $\nabla\eta$ is a $(0,2)$-tensor, with components

$$\nabla_\mu \eta_\sigma = e_\mu(\eta_\sigma) - {\Gamma_\sigma}^\nu{}_\mu \eta_\nu = \eta_{\sigma;\mu}.$$

In a coordinate basis,

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

In general, we can show, due to the Leibniz rule, that in a coordinate basis,

$$\begin{aligned}
{T^{\mu_1\ldots\mu_r}}_{\nu_1\ldots\nu_s;\rho} &= {T^{\mu_1\ldots\mu_r}}_{\nu_1\ldots\nu_s,\rho} + {\Gamma_\sigma}^{\mu_1}{}_\rho {T^{\sigma\mu_2\ldots\mu_r}}_{\nu_1\ldots\nu_s} + \cdots \\
&\quad + {\Gamma_\sigma}^{\mu_r}{}_\rho {T^{\mu_1\ldots\mu_{r-1}\sigma}}_{\nu_1\ldots\nu_s} - {\Gamma_{\nu_1}}^\sigma{}_\rho {T^{\mu_1\ldots\mu_r}}_{\sigma\nu_2\ldots\nu_s} - \cdots \\
&\quad - {\Gamma_{\nu_s}}^\sigma{}_\rho {T^{\mu_1\ldots\mu_r}}_{\nu_1\ldots\nu_{s-1}\sigma}.
\end{aligned}$$

*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 derivatives:

$$({T^a}_{b;c})_{;d} = {T^a}_{b;cd} = \nabla_d \nabla_c {T^a}_b.$$

In general, ${T^a}_{b;cd} \neq {T^a}_{b;dc}$.

If $f$ is a function, then $f_{;a} = (\mathrm{d}f)_a$ is a covector. In a coordinate basis $f_{;\mu} = f_{,\mu}$, then

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

hence

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

**Definition 1.11.** A connection (or covariant derivative) is *torsion free* (or symmetric) if

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f.$$

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

$${\Gamma_{[\mu}}^\rho{}_{\nu]} = 0 \implies {\Gamma_\mu}^\rho{}_\nu = {\Gamma_\nu}^\rho{}_\mu.$$

**Lemma 1.1.** *If $\nabla$ is torsion free, then for $X, Y$ vector fields,*

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

**Proof:** We show this is true in one basis: then as the objects are defined in a basis-free way, this must be true in all bases.

In a basis,

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

Even if $\nabla$ is torsion free, we may still not have $\nabla_a \nabla_b X^c = \nabla_b \nabla_a X^c$.

## 1.16 The Levi-Civita Connection

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

**Theorem 1.1** (Fundamental Theorem of Riemannian Geometry)**.** *If $(M, g)$ is a manifold with a metric, then there is a unique, torsion free connection $\nabla$ which satisfies*

$$\nabla g = 0.$$

This is called the *Levi-Civita connection*.

**Proof:** Suppose such a connection exists. Then by Leibniz, if $X, Y, Z$ are smooth vector fields, then

$$X(g(Y, Z)) = \nabla_X(g(Y, Z)) = (\nabla_X g)(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

If $\nabla_X g = 0$, then

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

Summing the first two equations and subtracting the third,

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

Using the fact that $\nabla_X Y - \nabla_Y X = [X, Y]$,

$$X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) = 2g(\nabla_X Y, Z) - g([X,Y],Z) - g([Z,X],Y) + g([Y,Z],X).$$

Hence we can write

$$g(\nabla_X Y, Z) = \frac{1}{2}\Big(X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) + g([X,Y],Z) + g([Z,X],Y) - g([Y,Z],X)\Big).$$

This determines $\nabla_X Y$ uniquely, since $g$ is non-degenerate. We can thus use this to define $\nabla_X Y$. Then we need to check the properties of a symmetric connection hold. For example,

$$\begin{aligned} g(\nabla_{fX} Y, Z) &= \frac{1}{2}\Big(fX(g(Y,Z)) + Y(fg(Z,X)) - Z(fg(X,Y)) \\ &\quad + g([fX,Y],Z) + g([Z,fX],Y) - g([Y,Z],fX)\Big) \\ &= \frac{1}{2}\Big(fX(g(Y,Z)) + fY(g(Z,X)) - fZ(g(X,Y)) \\ &\quad + Y(f)g(Z,X) - Z(f)g(X,Y) \\ &\quad + g(f[X,Y] - Y(f)X, Z) + g(f[Z,X] \\ &\quad + Z(f)X, Y) - fg([Y,Z],X)\Big). \end{aligned}$$

Hence by cancelling terms,

$$g(\nabla_{fX} Y, Z) = g(f\nabla_X Y, Z) \implies \nabla_{fX} Y = f\nabla_X Y.$$

We can check the other properties hold ourselves.

In a coordinate basis, we can compute

$$g(\nabla_{e_\mu} e_\nu, e_\sigma) = \frac{1}{2}\Big(e_\mu(g(e_\nu, e_\sigma)) + e_\nu(g(e_\sigma, e_\mu)) - e_\sigma(g(e_\mu, e_\nu))\Big).$$

This gives

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

or,

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

This $\nabla$ is the Levi-Civita connection, which is the same as the connection involved when extremizing proper time. We can raise or lower indices with $\nabla$, and this commutes with covariant differentiation.

## 1.17 Geodesics

We found that a curve extremizing proper time satisfies

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

where $\tau$ is the proper time. The tangent vector $X^a$ to the curve has components

$$X^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}.$$

Extending this off of the curve, we get a vector field, of which the geodesic is an integral curve. We note that

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) = \frac{\partial X^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = X^\mu{}_{,\nu}X^\nu,$$

hence the geodesic equation gives

$$X^\mu{}_{,\nu}X^\nu + \Gamma_\nu{}^\mu{}_\rho X^\nu X^\rho = 0 \iff X^\nu X^\mu{}_{;\nu} = 0 \implies \nabla_X X = 0.$$

We can extend this to any connection.

**Definition 1.12.** Let $M$ be a manifold with connection $\nabla$. An *affinely parametrized geodesic* satisfies

$$\nabla_X X = 0,$$

where $X$ is the tangent vector.

If we reparametrize $t \to t(u)$, then

$$\underbrace{\frac{\mathrm{d}x^\mu}{\mathrm{d}u}}_{Y} = \underbrace{\frac{\mathrm{d}x^\mu}{\mathrm{d}t}}_{X}\frac{\mathrm{d}t}{\mathrm{d}u},$$

so if we take $X \to Y = hX$ with $h > 0$, then

$$\nabla_Y Y = \nabla_{hX}(hX) = h\nabla_X(hX) = h^2\underbrace{\nabla_X X}_{0} + hX \cdot X(h) = fY,$$

with

$$f = X(h) = \frac{\mathrm{d}}{\mathrm{d}t}(h) = \frac{1}{h}\frac{\mathrm{d}h}{\mathrm{d}u} = \frac{1}{h}\frac{\mathrm{d}^2 t}{\mathrm{d}u^2}.$$

So $\nabla_Y Y = 0 \iff t = \alpha u + \beta$, for $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$.

**Theorem 1.2.** *Given $p \in M$, $X_p \in T_pM$, there exists a unique affinely parametrized geodesic $\lambda : I \to M$ satisfying*

$$\lambda(0) = p, \qquad \dot{\lambda}(0) = X_p.$$

**Proof:** Choose coordinates with $\phi(p) = 0$. Then $x^\mu(t) = \phi(\lambda(t))$ satisfies $\nabla_X X = 0$ with $X = X^\mu \frac{\partial}{\partial x^\mu}$, $X^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}$. This becomes

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} + \Gamma_\nu{}^\mu{}_\sigma \frac{\mathrm{d}x^\nu}{\mathrm{d}t}\frac{\mathrm{d}x^\sigma}{\mathrm{d}t} = 0.$$

This has unique solution $x^\mu : (-\varepsilon : \varepsilon) \to \mathbb{R}^n$ for $\varepsilon$ sufficiently small, given by standard ODE theory.

We just descend the condition $\nabla_X X = 0$ to $\mathbb{R}^n$.

## 1.18 Geodesic Postulate

In general relativity, free particles move along geodesics of the Levi-Civita connection. These are timelike for massive particles, and null/lightlike for massless particles.

If we fix $p \in M$, we can map $T_pM$ into $M$ by setting

$$\psi(X_p) = \lambda_{X_p}(1),$$

where $\lambda_{X_p}$ is the unique affinely parametrized geodesic with $\lambda_{X_i}(0) = p$, $\dot{\lambda}_{X_i}(0) = X_p$. Notice that

$$\lambda_{\alpha X_p}(t) = \lambda_{X_p}(\alpha t)$$

for $\alpha \in \mathbb{R}$, since if $\tilde{\lambda}(t) = \lambda_{X_p}(\alpha t)$, this is an affine reparametrization, so is still a geodesic, and

$$\dot{\tilde{\lambda}}(0) = \alpha \dot{\lambda}_{X_p}(0) = \alpha X_p.$$

Moreover $\alpha \mapsto \psi(\alpha X_p)$ is an affinely parametrized geodesic, so must be $\lambda_{X_p}(\alpha)$.

We claim that if $U \subseteq T_pM$ is a sufficiently small neighbourhood of the origin, then $\psi : T_pM \to M$ is one-to-one and onto.

**Definition 1.13.** Suppose $\{e_\mu\}$ is a basis for $T_pM$. We construct *normal coordinates* at $p$ as follows: for $q \in \psi(U) \subseteq M$, we define $\phi(q) = (X^1, \ldots, X^n)$, where $X^\mu$ are the components of the unique $X_p \in U$ with $\psi(X_p) = q$, with respect to $\{e_\mu\}$.

By our previous observation, the curve given in normal coordinates by $X^\mu(t) = tY^\mu$ for $Y^\mu$ a constant, is an affinely parametrized geodesic, so from the geodesic equation,

$$\Gamma_\nu{}^\mu{}_\sigma(tY)Y^\nu Y^\sigma = 0.$$

Set $t = 0$ to deduce that $\Gamma_{(\nu}{}^\mu{}_{\sigma)}|_p = 0$. So if $\nabla$ is torsion free, then $\Gamma_\nu{}^\mu{}_\sigma|_p = 0$ in normal coordinates.

Hence if $\nabla$ is the Levi-Civita connection of a metric, then

$$g_{\mu\nu,\rho}|_p = 0,$$

since

$$\begin{aligned} g_{\mu\nu,\rho} &= \frac{1}{2}(g_{\mu\nu,\rho} + g_{\rho\nu,\mu} - g_{\mu\rho,\nu}) + \frac{1}{2}(g_{\mu\nu,\rho} + g_{\mu\rho,\nu} - g_{\rho\nu,\mu}) \\ &= \Gamma_\mu{}^\sigma{}_\rho g_{\sigma\nu} + \Gamma_\rho{}^\sigma{}_\nu g_{\sigma\mu} = 0 \end{aligned}$$

at $p$. We can always choose the basis $\{e_\mu\}$ for $T_pM$ on which we base the normal coordinates to be orthonormal. We have:

**Lemma 1.2.** *On a Riemannian or Lorentzian field, we can choose normal coordinates at $p$ such that $g_{\mu\nu,\rho}|_p = 0$ and*

$$g_{\mu\nu}|_p = \begin{cases} \delta_{\mu\nu} & \text{Riemannian,} \\ \eta_{\mu\nu} & \text{Lorentzian.} \end{cases}$$

**Proof:** The curve given in normal coordinates by $t \mapsto (t, 0, \ldots, 0)$ is the APG with $\lambda(0) = p$, and $\dot\lambda(0) = e_1$, by the previous argument. But by definition of the coordinate basis, this vector is

$$\left(\frac{\partial}{\partial x^1}\right)_p.$$

So if $\{e_\mu\}$ is orthonormal, then at $p$

$$\left(\frac{\partial}{\partial x^\mu}\right)_p$$

forms an orthonormal basis.

# 2 Curvature

## 2.1 Parallel Transport

Suppose $\lambda : I \to M$ is a curve with tangent vector $\dot{\lambda}(t)$. We say that a tensor field $T$ is *parallely transported* or *propagated* along $\lambda$ if

$$\nabla_{\dot{\lambda}} T = 0$$

on $\lambda$. In particular, if $\lambda$ is an APG, then $\dot{\lambda}$ is parallely transported along $\lambda$.

A parallely propagated tensor is determined everywhere on $\lambda$ by its value at one point.

**Example 2.1.**

If $T$ is a $(1,1)$ tensor, then in coordinates, parallel propagation becomes

$$0 = \frac{\mathrm{d}x^\mu}{\mathrm{d}t} T^\nu{}_{\sigma;\mu} = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}(T^\nu{}_{\sigma,\mu} + \Gamma_\rho{}^\nu{}_\mu T^\rho{}_\sigma - \Gamma_\sigma{}^\rho{}_\mu T^\nu{}_\rho).$$

But

$$T^\nu{}_{\sigma,\mu} \frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(T^\nu{}_\sigma),$$

so

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} T^\nu{}_\sigma + (\Gamma_\rho{}^\nu{}_\mu T^\rho{}_\sigma - \Gamma_\sigma{}^\rho{}_\mu T^\nu{}_\rho)\frac{\mathrm{d}x^\mu}{\mathrm{d}t}.$$

This is a first-order linear ODE for $T^\nu{}_\sigma(x(t))$, so ODE theory gives us a unique solution, once $T^\nu{}_\sigma(x(0))$ is specified.

Note that parallel transport along a curve from $p$ to $q$ gives an isomorphism between tensors at $p$ and $q$, however this isomorphism depends on the choice of curve in general.

## 2.2 The Riemann Tensor

The Riemann tensor captures the extent to which parallel transport depends on the curve.

**Lemma 2.1.** *Given $X, Y, Z$ smooth vector fields, and $\nabla$ a connection, define*

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

*Then we claim*

$$(R(X,Y)Z)^a = R^a{}_{bcd} X^c Y^d Z^b,$$

*for a* $(1,3)$*-tensor* $R^a{}_{bcd}$*, the* Riemann tensor*.*

**Proof:** Suppose $f$ is a smooth function. Then,

$$\begin{aligned} R(fX,Y)Z &= \nabla_{fX}\nabla_Y Z - \nabla_Y \nabla_{fX} Z - \nabla_{[fX,Y]}Z \\ &= f\nabla_X \nabla_Y Z - \nabla_Y(f\nabla_X Z) - \nabla_{f[X,Y]-Y(f)X} Z \\ &= f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z - f\nabla_{[X,Y]}Z + Y(f)\nabla_X Z \\ &= fR(X,Y)Z. \end{aligned}$$

Since $R(X,Y)Z = -R(Y,X)Z$, we have $R(X,fY)Z = fR(X,Y)Z$. Similarly, we can check $R(X,Y)(fZ) = fR(X,Y),Z$.

Now suppose we pick a basis $\{e_\mu\}$ with dual basis $\{f^\mu\}$. Then

$$\begin{aligned} R(X,Y)Z &= R(X^\rho e_\rho, Y^\sigma e_\sigma)(Z^\nu e_\nu) = X^\rho Y^\sigma Z^\nu R(e_\rho, e_\sigma) e_\nu \\ &= (R^\mu{}_{\nu\rho\sigma} X^\rho Y^\sigma Z^\nu) e_\mu, \end{aligned}$$

where $R^\mu{}_{\nu\rho\sigma} = f^\mu(R(e_\rho, e_\sigma)e_\nu)$ are the components of $R^a{}_{bcd}$ in this basis. Since this results holds in one basis, it hold in all, since the object was defined tensorially.

In a coordinate basis $e_\mu = \frac{\partial}{\partial x^\mu}$, with $[e_\mu, e_\nu] = 0$, we have

$$\begin{aligned} R(e_\rho, e_\sigma)e_\nu &= \nabla_{e_\rho}(\nabla_{e_\sigma} e_\nu) - \nabla_{e_\sigma}(\nabla_{e_\rho} e_\nu) - \nabla_{e_\rho}(\Gamma_\nu{}^\tau{}_\sigma e_\tau) - \nabla_{e_\sigma}(\Gamma_\nu{}^\tau{}_\rho e_\tau) \\ &= \partial_\rho(\Gamma_\nu{}^\tau{}_\sigma)e_\tau + \Gamma_\nu{}^\tau{}_\sigma \Gamma_\tau{}^\mu{}_\rho e_\mu - \partial_\sigma(\Gamma_\nu{}^\tau{}_\rho)e_\tau - \Gamma_\nu{}^\tau{}_\rho \Gamma_\tau{}^\mu{}_\sigma e_\mu. \end{aligned}$$

Hence the components are

$$T^\mu{}_{\nu\rho\sigma} = \partial_\rho(\Gamma_\nu{}^\mu{}_\sigma) - \partial_\sigma(\Gamma_\nu{}^\mu{}_\rho) + \Gamma_\nu{}^\tau{}_\sigma \Gamma_\tau{}^\mu{}_\rho - \Gamma_\nu{}^\tau{}_\rho \Gamma_\tau{}^\mu{}_\sigma.$$

In normal coordinates, we can drop the last two terms, since the connection is zero at $p$. However the derivative is not 0.

**Example 2.2.**

For the Levi-Civita connection of Minkowski space in an intertial frame, $\Gamma_\mu{}^\nu{}_\sigma = 0$. Hence in these coordinates $R^\mu{}_{\nu\sigma\tau} = 0$. Thus

$$R^a{}_{bcd} = 0.$$

Such a connection is called *flat*.

Conversely, for a Lorentzian spacetime with flat Levi-Civita connection, we can locally find coordinates such that $g_{\mu\nu} = \mathrm{diag}(-1, 1, \ldots, 1)$.

Here is a note of caution:

$$(\nabla_X \nabla_Y Z)^c = X^a \nabla_a (Y^b \nabla_b Z^c) \neq X^a Y^b \nabla_a \nabla_b Z^c.$$

Hence

$$\begin{aligned}(R(X,Y)Z)^c &= X^a \nabla_a (Y^b \nabla_b Z^c) - Y^a \nabla_a (X^b \nabla_b Z^c) - [X,Y]^b \nabla_b Z^c \\ &= X^a Y^b \nabla_a \nabla_b Z^c - Y^a X^b \nabla_a \nabla_b Z^c + (\nabla_X Y - \nabla_Y X - [X,Y])^b \nabla_b Z^c \\ &= X^a Y^b R^c{}_{dab} Z^d.\end{aligned}$$

So if $\nabla$ is torsion free,

$$\nabla_a \nabla_b Z^c - \nabla_b \nabla_a Z^c = R^c{}_{dab} Z^d.$$

This is the *Ricci identity*. We can generalise this to an expression for

$$\nabla_{[a} \nabla_{b]} T^{c_1 \ldots c_r}{}_{d_1 \ldots d_s}.$$

We can construct a new tensor from $R^a{}_{bcd}$ by contraction.

**Definition 2.1.** The *Ricci tensor* is the $(0, 2$-tensor defined by

$$R_{ab} = R^c{}_{acb}.$$

Suppose $X, Y$ are vector fields satisfying $[X, Y] = 0$.

Insert picture of vector curves going from $A \to B \to C \to D \to A$.

Consider going from $A$ to $B$ flowing a distance $\varepsilon$ along the integral curve of $X$, then $B$ to $C$ distance $\varepsilon$ along the integral curve of $Y$, then $C$ to $D$ along $X$, and $D$ to $A$ along $Y$.

Since $[X, Y] = 0$, we return to the starting location. We claim the following of a propagated tensor:

**Proposition 2.1.** *If $Z$ is parallely transported around $ABCD$ to a vector $Z'$, then*

$$(Z - Z')^\mu = \varepsilon^2 R^\mu{}_{\nu\rho\sigma} Z^\nu X^\rho Y^\sigma + \mathcal{O}(\varepsilon^3).$$

## 2.3 Geodesic Deviation

Let $\nabla$ be a symmetric connection, and suppose $\lambda : I \to M$ is an affinely parametrized geodesic through $p$. We can pick normal coordinates centred at $p$, such at $\lambda$ is given by $t \mapsto (t, 0, \ldots, 0)$. Suppose we start a geodesic with

$$x_s^\mu(0) = s x_0^\mu, \qquad \dot{x}_s^\mu(0) = s x_1^\mu + (1, 0, \ldots, 0).$$

Then we find that

$$x_s^\mu(t) = x^\mu(s,t) = (t, 0, \ldots, 0) + sY^\mu(t) + \mathcal{O}(s^2),$$

where

$$Y^\mu(t) = \left.\frac{\partial x^\mu}{\partial s}\right|_{s=0}$$

are the components of a vector field along $\lambda$ measuring the (infinitesimal) deviation of the geodesics. We have, by the geodesic equation,

$$\frac{\partial^2 x^\mu}{\partial t^2} + \Gamma_{\nu\ \sigma}^{\ \mu}(x^\mu(s,t)) \frac{\partial x^\nu}{\partial t} \frac{\partial x^\mu}{\partial t} = 0.$$

Take the derivative with respect to $s$, at $s = 0$. Then,

$$\frac{\partial^2 Y^\mu}{\partial t^2} + \partial_\mu (\Gamma_{\nu\ \sigma}^{\ \mu})_{s=0} T^\nu T^\sigma Y^\rho + Z \Gamma_{\rho\ \sigma}^{\ \mu} \frac{\partial}{\partial t} Y^\rho T^\sigma = 0,$$

which gives

$$T^\nu (T^\sigma Y^\mu_{\ ,\sigma})_{,\nu} + \partial_\rho (\Gamma_{\nu\ \sigma}^{\ \mu})_{s=0} T^\nu T^\sigma Y^\rho + Z \Gamma_{\rho\ \sigma}^{\ \mu} \frac{\partial Y^\rho}{\partial t} T^\sigma = 0.$$

At $p = 0$, $\Gamma = 0$ so

$$T^\nu (T^\sigma Y^\mu_{\ ;\sigma} - \Gamma_{\rho\ \sigma}^{\ \mu} T^\sigma Y^\rho)_{;\nu} + (\partial_\rho \Gamma_{\nu\ \sigma}^{\ \mu})_p T^\nu T^\sigma Y^\rho = 0.$$

We can write this as

$$\begin{aligned}
T^\nu (T^\sigma Y^\mu_{\ ;\sigma})_{;\nu} + (\partial_\rho \Gamma_{\nu\ \sigma}^{\ \mu} - \partial_\nu \Gamma_{\rho\ \sigma}^{\ \mu}) T^\nu T^\sigma Y^\rho &= 0 \\
\implies (\nabla_T \nabla_T Y)^\mu + R^\mu_{\ \omega\rho\nu} T^\nu T^\sigma Y^\rho &= 0 \\
\implies \nabla_T \nabla_T Y + R(Y,T)T &= 0.
\end{aligned}$$

This is the Jacobi equation for geodesic deviation.

## 2.4 Riemann Tensor Symmetries

From the definition, it is clear that

$$R^a{}_{bcd} = -R^a{}_{bdc} \iff R^a{}_{b(cd)} = 0.$$

**Proposition 2.2.** *If $\nabla$ is torsion free, then*

$$R^a{}_{[bcd]} = 0.$$

**Proof:** Fix $p \in M$, and choose normal coordinates at $p$, and work in the coordinate basis. Then $\Gamma$ vanishes, and $\Gamma_\mu{}^\sigma{}_\nu = \Gamma_\nu{}^\sigma{}_\mu$ everywhere. Then,

$$\begin{aligned} R^\mu{}_{\nu\rho\sigma}|_p &= \partial_\rho(\Gamma_\nu{}^\mu{}_\sigma)|_p - \partial_\sigma(\Gamma_\nu{}^\mu{}_\rho)|_p \\ \implies R^\mu{}_{[\nu\rho\sigma]}|_p| &= 0, \end{aligned}$$

as by symmetry

$$\partial_\rho \Gamma_\nu{}^\mu{}_\sigma|_p = \partial_\rho \Gamma_\sigma{}^\mu{}_\nu|_p$$

Here $p$ is arbitrary, so $R^\mu{}_{[\nu\rho\sigma]} = 0$ everywhere.

**Proposition 2.3.** *If $\nabla$ is torsion free, then the Bianchi identity holds:*

$$R^a{}_{b[cd;e]} = 0.$$

**Proof:** Choose normal coordinates. Then

$$R^\mu{}_{\nu\rho\sigma;\tau}|_p = R^\mu{}_{\nu\rho\sigma,\tau}|_p.$$

Schematically, $R \sim \partial\Gamma + \Gamma^2$, so $\partial R \sim \partial^2\Gamma + \partial\Gamma \cdot \Gamma$, and since $\Gamma|_p = 0$, we deduce

$$R^\mu{}_{\nu\rho\sigma,\tau}|_p = \partial_\tau\partial_\rho\Gamma_\nu{}^\mu{}_\sigma|_p - \partial_\tau\partial_\sigma\Gamma_\nu{}^\mu{}_\rho|_p.$$

By symmetry of mixed partial derivatives, we see

$$R^\mu{}_{\nu[\rho\sigma;\tau]}|_p = 0.$$

Since $p$ was arbitrary, the result follows.

Suppose $\nabla$ is the Levi-Civita connection of a manifold with a metric $g$. We can lower an index with $g_{ab}$, and consider $R_{abcd}$.

**Proposition 2.4.** $R_{abcd}$ *satisfies*

$$R_{abcd} = R_{cdab} \implies R_{(ab)cd} = 0.$$

**Proof:** Pick normal coordinates at $p$ so that $\partial_\mu g_{\nu\rho}|_p = 0$. We notice that

$$0 = \partial_\mu \delta^\nu{}_\sigma|_p = \partial_\mu(g^{\nu\tau} g_{\tau\sigma}) = (\partial_\mu g^{\nu\tau}) g_{\tau\sigma}|_p.$$

The derivative $\partial_\mu^{\nu\tau}|_p = 0$, hence

$$\begin{aligned}\partial_\rho(\Gamma_\nu{}^\mu{}_\sigma)|_p &= \partial_\rho\left(\frac{1}{2}g^{\mu\tau}(g_{\tau\sigma,\nu} + g_{\nu\tau,\sigma} - g_{\nu\sigma,\tau})\right)|_p \\ &= \frac{1}{2}g^{\mu\tau}(g_{\tau\sigma,\nu\rho} + g_{\nu\tau,\sigma\rho} - g_{\nu\sigma,\tau\rho})|_p.\end{aligned}$$

We have then that

$$\begin{aligned}R_{\mu\nu\rho\sigma}|_p &= g_{\mu\kappa}(\partial_\rho \Gamma_\nu{}^\kappa{}_\sigma - \partial_\sigma \Gamma_\nu{}^\kappa{}_\rho)|_p \\ &= \frac{1}{2}(g_{\mu\sigma,\nu\rho} + g_{\nu\rho,\mu\sigma} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma})|_p.\end{aligned}$$

This satisfies

$$R_{\mu\nu\rho\sigma}|_p = R_{\rho\sigma\mu\nu}|_p,$$

hence this is true everywhere.

**Corollary 2.1.** *The Ricci tensor is symmetric:*

$$R_{ab} = R_{ba}.$$

**Proof:** We have

$$R_{ab} = R^c{}_{acb} = g^{cd} R_{cadb} = g^{cd} R_{dbca} = R_{ba}.$$

**Definition 2.2.** The *Ricci scalar* (scalar curvature) is

$$R = R^a{}_a = g^{ab} R_{ab}.$$

The *Einstein tensor* is

$$G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R.$$

We can show that the Bianchi identity implies

$$\nabla_a G^a{}_b = 0.$$

This is the contracted Bianchi identity.

# 3 Diffeomorphisms and the Lie Derivative

Suppose $\varphi : M \to N$ is a smooth map. Then $\varphi$ induces maps between corresponding vector or covector bundles.

**Definition 3.1.** Given $f : N \to \mathbb{R}$, the *pullback* of $f$ by $\varphi$ is the map $\varphi^* f : M \to \mathbb{R}$ given by

$$\varphi^* f(p) = f(\varphi(p)).$$

**Definition 3.2.** Given $X \in T_p M$, we define the *push forward* of $X$ by $\varphi$, say $\varphi_* X \in T_{\varphi(p)} N$ as follows:

Let $\lambda : I \to M$ be a curve with $\lambda(0) = p$, $\dot{\lambda}(0) = X$. Then set $\tilde{\lambda} = \varphi \circ \lambda$. $\tilde{\lambda} : I \to N$ gives a curve in $N$ with $\tilde{\lambda}(0) = \varphi(p)$.

We set $\varphi_* X = \dot{\tilde{\lambda}}(0)$.

Note that if $f : N \to \mathbb{R}$, then

$$\begin{aligned} \varphi_* X(f) &= \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \tilde{\lambda}(t))|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \varphi \circ \lambda(t))|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(\varphi^* f \circ \lambda(t))|_{t=0} = X(\varphi^* f). \end{aligned}$$

We can show that if $x^\mu$ are coordinates on $M$ near $p$, and $y^\alpha$ are coordinates on $N$ near $\varphi(p)$, then $\varphi$ gives a map $y^\alpha(x^\mu)$, and we can show in a coordinate basis that

$$(\varphi_* X)^\alpha = \left(\frac{\partial y^\alpha}{\partial x^\mu}\right)_p X^\mu.$$

On the contangent bundle, we go backwards.

**Definition 3.3.** If $\eta \in T^*_{\varphi(p)} N$, then the *pullback* of $\eta$, $\varphi^* \eta \in T^*_p M$ is defined by

$$\varphi^* \eta(X) = \eta(\varphi_* X),$$

for all $X \in T_p M$.

Note that if $f : N \to \mathbb{R}$, then

$$\begin{aligned} \varphi^*(\mathrm{d}f)[X] &= \mathrm{d}f[\varphi_* X] = \varphi_* X(f) = X(\varphi^* f) \\ &= \mathrm{d}(\varphi^* f)[X], \end{aligned}$$

hence $\varphi^* \, \mathrm{d}f = \mathrm{d}(\varphi^* f)$. We can also show that in a basis,

$$(\varphi^* \eta)_\mu = \left(\frac{\partial y^\alpha}{\partial x^\mu}\right)_p \eta_\alpha.$$

We can extend the pullback to map a $(0,s)$-tensor $T$ at $\varphi(p) \in N$, to a $(0,s)$-tensor $\varphi^*T$ at $p \in M$, by

$$\varphi^*T(X_1, \dots, X_s) = T(\varphi_* X_1, \dots, \varphi_* X_s)$$

for all $X_i \in T_pM$. Similarly, we can push forward a $(s,0)$-tensor $S$ at $p \in M$ to a $(S,0)$-tensor $\varphi_*S$ at $\varphi(p)$ by

$$\varphi_* S(\eta_1, \dots, \eta_s) = S(\varphi^*\eta_1, \dots, \varphi^*\eta_s),$$

for all $\eta_1, \dots, \eta_s \in T^*_{\varphi(p)}N$.

If $\varphi : M \to N$ has the property that $\varphi_* : T_pM \to T_{\varphi(p)}N$ is injective, we say that $\varphi$ is an *immersion*. In particular $\dim N \geq \dim M$.

If $N$ is a manifold with metric $g$ and $\varphi : M \to N$ is an immersion, we can consider $\varphi^*g$. If $g$ is Riemannian, then $\varphi^*g$ is non-degenerate and positive definite, so it defines a metric on $M$, the *induced metric*.

**Example 3.1.**

Let $(N, g) = (\mathbb{R}^3, \delta)$, and $M = S^2$. Let $\varphi$ be the map taking a point on $S^2$ with spherical coordinates $(\theta, \phi)$ to

$$(x^1, x^2, x^3) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

Then

$$\varphi^*((\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2) = \mathrm{d}\theta^2 + \sin^2\theta \,\mathrm{d}\phi^2.$$

If $\varphi$ is an immersion and $(N, g)$ is Lorentzian, then $\varphi^*g$ is not in general a metric on $M$. There are three important cases:

- $\varphi^*g$ is a Riemannian metric $\implies$ $\varphi(M)$ is spacelike.
- $\varphi^*g$ is a Lorentian metric $\implies$ $\varphi(M)$ is timelike.
- $\varphi^*g$ is everywhere degenerate $\implies$ $\varphi(M)$ is null.

Recall that $\varphi : M \to N$ is a *diffeomorphism* if it is bijective with a smooth inverse. If we have a diffeomorphism, we can push forward a general $(r,s)$-tensor at $p$ to an $(r,s)$-tensor at $\varphi(p)$ by'

$$\varphi_*T(\eta^1, \dots, \eta^r, X_1, \dots, X_s) = T(\varphi^*\eta^1, \dots, \varphi^*\eta^r, \varphi_*^{-1}X_1, \dots, \varphi_*^{-1}X_s),$$

for all $\eta_i \in T^*_{\varphi(p)}N$, and $X_i \in T_{\varphi(p)}N$. We can similarly define a pullback by $\psi_*^{-1} = \varphi^*$.

If $M, N$ are diffeomorphic, we often do not distinguish between them, and can think of $\varphi : M \to M$.

We say that a diffeomorphism $\varphi : M \to M$ is a symmetry of $T$ if $\varphi_* T = T$. If $T$ is the metric, we say $\varphi$ is an *isometry*.

**Example 3.2.**

In a Minkowski space with an interial frame,

$$\varphi(x^0, x^1, \dots, x^n) = (x^0 + 1, x^1, \dots, x^n)$$

is a symmetry of $g$.

An important class of diffeomorphisms are those that are generated by a vector field. If $X$ is a smooth vector field, we associate to each point $p \in M$ the point $\varphi_t^X(p) \in M$ given by flowing a parameter distance $t$ along the integral curve of $X$ starting at $p$.

Suppose $\varphi_t^X(p)$ is well defined for all $t \in I \subseteq \mathbb{R}$ for each $p \in M$. Then $\varphi_t^X : M \to M$ is a diffeomorphism for all $t \in I$. Further,

- If $t, s, t+s \in I$, then $\varphi_t^X \circ \varphi_s^X = \varphi_{t+s}^X$, and $\varphi_0^X = \mathrm{id}$.
- If $I = \mathbb{R}$, this gives $\{\varphi_t^X\}$ the structure of a one-parameter abelian group.
- If $\varphi_t$ is any smooth family of diffeomorphisms satisfying the group property, we can define a vector field by

  $$X_p = \frac{\mathrm{d}}{\mathrm{d}t}(\phi_t(p))\bigg|_{t=0}.$$

  Then $\varphi_t = \varphi_t^X$.

We can use $\varphi_t^X$ to compare tensors at different points. As $t \to 0$, this gives a new notion of derivative.

## 3.1 The Lie Derivative

Suppose $\varphi_t^X : M \to M$ is the smooth one-parameter family of diffeomorphisms generated by a vector field $X$.

**Definition 3.4.** For a tensor field $T$, the *Lie derivative* of $T$ with repsect to $X$ is

$$(\mathcal{L}_X T)_p = \lim_{t \to 0} \frac{((\varphi_T^X)^* T)_p - T_p}{t}.$$

This is the push-forward of the tensor field minus the original tensor field.

It is easy to see that for constants $\alpha, \beta$ and $(r, s)$-tensors $S, T$,

$$\mathcal{L}_X(\alpha S + \beta T) + \alpha \mathcal{L}_X S + \beta \mathcal{L}_X T.$$

To see how $\mathcal{L}_X$ acts in components, it is useful to construct coordinates adapted to $X$. Near $p$, we can construct an $(n-1)$-surface $\Sigma$ which is transverse to $X$. Pick coordinates $x^i$ on $\Sigma$ and assign the coordinate $(t, x^i)$ to the point a parameter distance $t$ along the integral curve of $X$ starting at $x^i$ on $\Sigma$.

In these coordinates $X = \frac{\partial}{\partial t}$, and $\varphi_t^X(\tau, x^i) = (\tau + t, x^i)$. So if $y^\mu = \varphi_t^X(x^\mu)$, then

$$\frac{\partial y^\mu}{\partial x^\nu} = \delta^\mu{}_\nu.$$

Hence we find

$$[(\varphi_t^X)^* T]^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_{(t, x^i)} = T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_{(\tau + t, x^i)}.$$

Thus we find

$$(\mathcal{L}_X T)^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_p = \frac{\partial T^{\mu_1 \dots \mu_r}}{\partial t}{}_{\nu_1 \dots \nu_s}|_p.$$

So in these coordinates, $\mathcal{L}_X$ acts on components by $\frac{\partial}{\partial t}$. In particular, we immediately see that

$$\mathcal{L}_X(S \otimes T) = (\mathcal{L}_X S) \otimes T + S \otimes \mathcal{L}_X T.$$

So $\mathcal{L}_X$ commutes with contraction. To write $\mathcal{L}_X$ in a coordinate free fashion, we can simply seek a basis independent expression that agrees with $\mathcal{L}_X$ in these coordinates.

For example, for a function

$$\mathcal{L}_X f = \frac{\partial f}{\partial t} = X(f)$$

in these coordinates. For a vector field $Y$, we observe that

$$\frac{\partial Y^\mu}{\partial t} = X^\sigma \frac{\partial}{\partial x^\sigma}(Y^\mu) - Y^\sigma \frac{\partial}{\partial x^\sigma} X^\mu = [X, Y]^\mu.$$

So $\mathcal{L}_X Y = [X, Y]$. For a covector field, we can show

$$(\mathcal{L}_X \omega)_\mu = X^\sigma \partial_\sigma w_\mu + w_\sigma \partial_\mu X^\sigma.$$

If $\nabla$ is torsion-free, then

$$(\mathcal{L}_X \omega)_a = X^b \nabla_b \omega_a + \omega_b \nabla_a X^b.$$

If $g_{ab}$ is a metric tensor, and $\nabla$ is Levi-Civita, then

$$(\mathcal{L}_X g)_{ab} = \nabla_a X_b + \nabla_b X_a.$$

If $\varphi_t^X$ is a one-parameter family of isometries for a manifold with metric $g$, then $\mathcal{L}_X g = 0$. Conversely if $\mathcal{L}_X g = 0$, then $X$ generates a one-parameter family of isometries.

**Definition 3.5.** A vector field $K$ satisfies $\mathcal{L}_K g = 0$ is called a *Killing vector*. It satisfies Killing's equation:

$$\nabla_a K_b + \nabla_b K_a = 0.$$

**Lemma 3.1.** *Suppose $K$ is Killing and $\lambda : T \to M$ is a geodesic of the Levi-Civita connection then $g_{ab}\dot{\lambda}^a K^b$ is constant along $\lambda$.*

**Proof:** We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(K_b\dot{\lambda}^b) = \dot{\lambda}^a\nabla_a(K_b\dot{\lambda}^b) &= (\nabla_{(a}K_{b)})\dot{\lambda}^a\dot{\lambda}^b + K_b\dot{\lambda}^a\nabla_a\dot{\lambda}^b \\ &= 0,\end{aligned}$$

for the first part as $K$ is Killing and the latter as $\nabla_\lambda \lambda = 0$ for a geodesic.

# Index