

III General Relativity

Ishan Nath, Michaelmas 2024

Based on Lectures by Prof. Claude Warnick

October 25, 2024

Contents

0	Introduction	2
1	Differentiable Manifolds	3
1.1	Smooth Functions on Manifolds	5
1.2	Curves and Vectors	7
1.3	Covectors	10
1.4	The (Co)tangent bundle	11
1.5	Abstract Index Notation	12
1.6	Tensors	12
1.7	Change of Bases	13
1.8	Tensor Operations	14
1.9	Tensor Bundles	16
1.10	Integral Curves	17
1.11	Commutators	18
1.12	Metric Tensor	18
1.13	Lorentzian Signature	21
1.14	Curves of Extremal Proper Time	22
	Index	23

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_α, ϕ_α) , where:

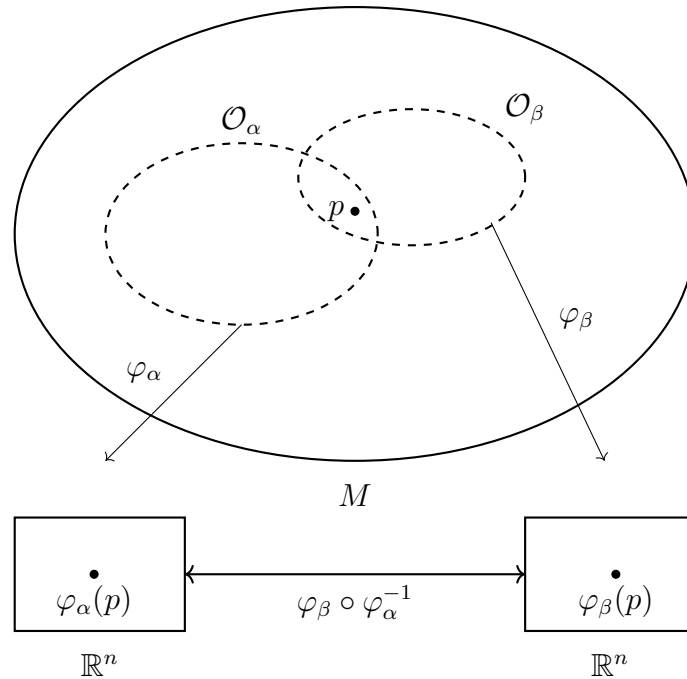
- $O_\alpha \subseteq M$ are subsets of M such that

$$\bigcup_{\alpha} O_\alpha = M.$$

- ϕ_α is a bijective map from O_α to U_α , 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 α . 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 $\phi : U \rightarrow \mathbb{R}^n$ 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 ϕ_1^{-1}, ϕ_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 $\{\mathbf{E}_1, \dots, \mathbf{E}_{n+1}\}$ is the standard basis for \mathbb{R}^{n+1} , and $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis for \mathbb{R}^n , write

$$\mathbf{P} = P^1 \mathbf{E}_1 + \dots + P^{n+1} \mathbf{E}_{n+1}.$$

Set $O_1 = S^n \setminus \{\mathbf{E}_{n+1}\}$, and write

$$\phi_1(\mathbf{P}) = \frac{1}{1 - P^{n+1}}(P^1 \mathbf{e}_1 + \dots + P^n \mathbf{e}_n).$$

In a similar way we may define O_2, ϕ_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 $S^n \subseteq \mathbb{R}^{n+1}$.

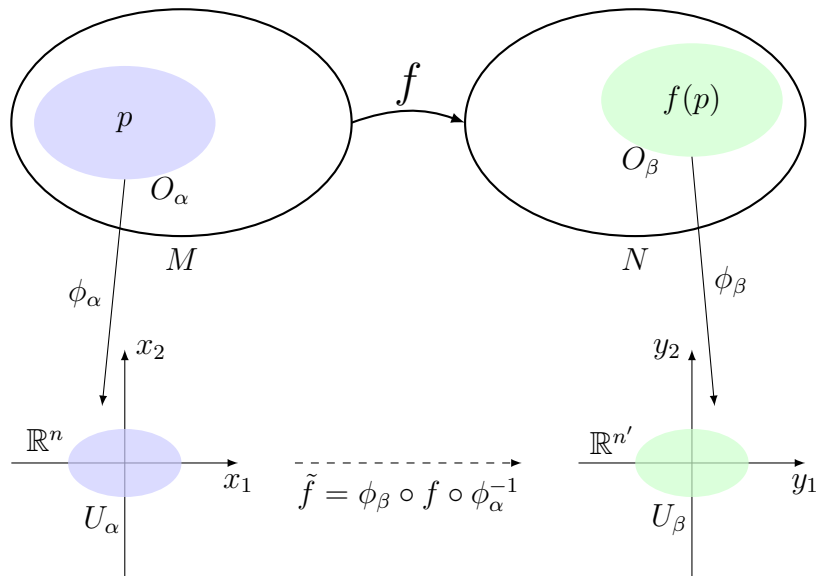
1.1 Smooth Functions on Manifolds

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

Then let $p \in M$, and pick charts (O_α, ϕ_α) for M , and (O'_β, ϕ'_β) 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 $U'_\beta \subseteq \mathbb{R}^{n'}$. If this function is smooth for all possible choices of this chart, we say that $f : M \rightarrow N$ is smooth.



Remark.

- A smooth map $\psi : M \rightarrow N$ 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 $f : I \rightarrow N$ 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 $f(x, y) = x$, $f : S^1 \rightarrow \mathbb{R}$. Using our previous charts,

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

Similarly,

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

Hence f is smooth.

2. If (O, ϕ) 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', ϕ') 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

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

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

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

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

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

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

In practice, we often don't distinguish between f and its coordinate chart representations F_α .

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 $\lambda : I \rightarrow M$ a smooth map, is a smooth curve in M .

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

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

where $X(t) = d\lambda/dt(t)$ is the tangent vector to λ at t .

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

$$X_p(f) = \left. \frac{d}{dt} f(\lambda(t)) \right|_{t=0}.$$

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, ϕ) is a chart and $p \in O$, write

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

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

$$\left. \frac{d}{dt} f(\lambda(t)) \right|_{t=0} = \frac{\partial F}{\partial x^\mu}(x) \cdot \left. \frac{dx^\mu}{dt} \right|_{t=0}.$$

Proposition 1.1. *The set of tangent vectors to curves at p forms a vector space, $T_p M$, of dimension $n = \dim M$. We call $T_p M$ 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 λ, κ be smooth curves with $\lambda(0) = \kappa(0) = p$, and whose tangent vectors at p are X_p, Y_p , respectively.

Let (O, ϕ) 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 γ at p , then

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

Thus $T_p M$ is a vector space.

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

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

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

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

These vectors are linearly independent. Otherwise there exists α^μ not all zero such that

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

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

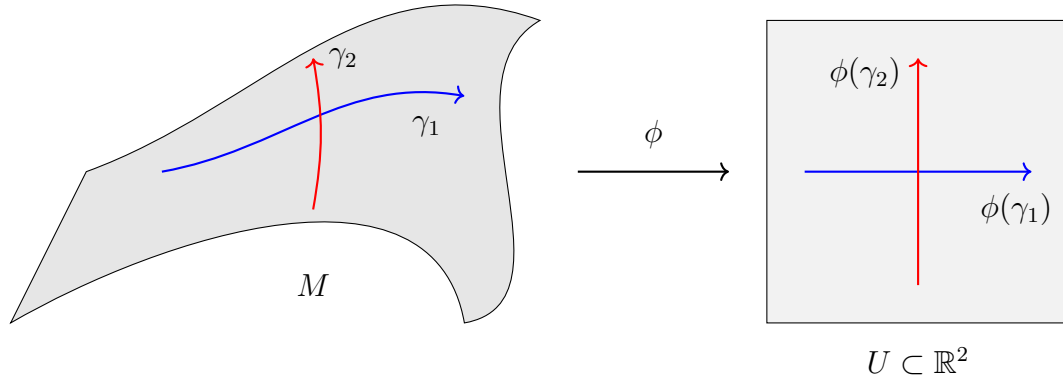
Further, these vectors form a basis for $T_p M$, since if λ is any curve with tangent X_p at p , then

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

Here, we set

$$X^\mu = \left. \frac{d}{dt} x^\mu(\lambda(t)) \right|_{t=0}.$$

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 ϕ . Suppose we choose another chart (O', ϕ') , 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,

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

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^μ be the components of X_p with respect to $\{(\partial/\partial x^\mu)_p\}$, and X'^μ 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 $T_p M$, 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 $\Lambda_x : V^* \rightarrow \mathbb{R}$ by

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

Definition 1.3. The dual space of $T_p M$ 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 $T_p M$ and $\{f^\mu\}$ the dual basis for $T_p^* M$, then we can expand a covector η as

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

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

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 $f : M \rightarrow \mathbb{R}$ is a smooth function, define $(df)_p \in T_p^* M$, the *differential* of f at p , by

$$(df)_p(X) = X(f),$$

for any $X \in T_p M$. $(df)_p$ is sometimes also called the *gradient* of f at p .

If f is a constant, then $X(f) = 0 \implies (df)_p = 0$.

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

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

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

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

justifying the language of the ‘gradient’.

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

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

where $x(x') = \phi \circ \phi'^{-1}$, and hence if η_μ, η'_μ 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 $T_p M$ as p varies to get a new $2n$ dimensional manifold, TM , the *tangent bundle*. We let

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

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

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

and

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

where X^μ are the components of X with respect to the coordinate basis of ϕ_α .

It can be shown that if (O, ϕ) and (O', ϕ') 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 $\pi : TM \rightarrow M$ which takes $(p, X) \mapsto p$ is smooth (show this!)

1.5 Abstract Index Notation

We have used Greek letters μ, ν 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 η by η_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

$$X(f) = X^a (df)_a.$$

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

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

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

Example 1.3.

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

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

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

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

$$T^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} = T(f_1^{\mu_1}, \dots, f_r^{\mu_r}, e_{\nu_1}, \dots, e_{\nu_s}).$$

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

Example 1.4.

1. Consider δ 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 δ as δ^a_b in AIN.

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

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

This can be generalized to higher ranks.

1.7 Change of Bases

We have already seen how the components of X or η with respect to a coordinate basis (X^μ, η_ν) 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 $T_p M$ with dual bases $\{f^\mu\}$ and $\{f'^\mu\}$.

As these are bases, we can expand

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

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

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

then we have already seen

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

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

A change 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,

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

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 $(T_s^r)_p M$. These can be glued together to form the bundle of (r, s) -tensors

$$T_s^r M = \bigcup_{p \in M} \{p\} \times (T_s^r)_p M.$$

If (O, ϕ) is a coordinate chart on M , set

$$\tilde{O} = \bigcup_{p \in O} \{p\} \times (T_s^r)_p M \subseteq T_s^r M,$$

and

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

$T_s^r M$ is a manifold, with a natural smooth map $\Pi : T_s^r M \rightarrow M$ such that $\pi(p, S_p) = p$.

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

If (O, ϕ) 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 $f : M \rightarrow \mathbb{R}$ to give a new function Xf by

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

In coordinates,

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

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

1.10 Integral Curves

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

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

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 ϕ with $\phi = (x^1, \dots, x^n)$, and $\phi(p) = 0$. In this chart, the equation becomes

$$\frac{dx^\mu}{dt}(t) = X^\mu(x(t)),$$

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 $f : M \rightarrow \mathbb{R}$ 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:

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

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 $\lambda : (a, b) \rightarrow M$ is given by

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

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

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

$$dx^\mu dx^\nu = \frac{1}{2}(dx^\mu \otimes dx^\nu + dx^\nu \otimes dx^\mu).$$

And by convention we often write $g = ds^2$, so that

$$g = ds^2 = g_{\mu\nu} dx^\mu dx^\nu.$$

Example 1.6.

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

$$g = ds^2 = (dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2 + \delta_{\mu\nu} dx^\mu dx^\nu.$$

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

$$g = ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \eta_{\mu\nu} dx^\mu dx^\nu.$$

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

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

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

In this chart, the round metric is

$$g = ds^2 = d\theta^2 + \cos^2 \theta d\varphi^2.$$

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

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 $g = d\theta'^2 + \sin^2 \theta' d\varphi'^2$ 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 ϕ 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 $T_p M$ and $T_p^* M$.

Given $X^a \in T_p M$, 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, δ) 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 $0, 1, \dots, n$. At any point in a Lorentzian manifold we can find a basis $\{e_\mu\}$ such that

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

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. $X \in T_p M$ is:

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

A curve $\lambda : I \rightarrow M$ 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 $\lambda : (a, b) \rightarrow M$, the relevant quantity is the *proper time*

$$\tau(\lambda) = \int_a^b \sqrt{-g_{ab} \frac{dx^a}{du} \frac{dx^b}{du}} du.$$

If

$$g_{ab} \frac{d\lambda^a}{du} \frac{d\lambda^b}{du} = -1$$

for all u , then λ is parametrized by proper time. In this case, we call the tangent vector $U^a = \frac{d\lambda^a}{dx}$ the *four-velocity* of λ .

1.14 Curves of Extremal Proper Time

Suppose $\lambda : (0, 1) \rightarrow M$ 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

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

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

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

We can compute

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

Index

abstract index notation, 12
angle, 19
antisymmetrization, 15
atlas, 4

Cartesian, 20
compatible atlases, 4
contraction, 14
cotangent bundle, 12
cotangent space, 10

diffeomorphism, 5
differentiable manifold, 3
differential, 10
dual basis, 10
dual space, 10

four-velocity, 22

gradient, 10

integral curve, 17
inertial frame, 20

length, 19, 22
lightlike, 21
Lorentz transformation, 21
Lorentzian, 19

manifold, 3
maximal atlas, 4
metric tensor, 19
Minkowski space, 20

non-degenerate, 19
norm, 19
null, 21

proper time, 22

Riemannian, 19

scalar field, 5
signature, 19
spacelike, 21
symmetrization, 15

tangent bundle, 11
tangent space, 7
tangent vector, 7
tensor, 12
tensor field, 16
tensor product, 15
timelike, 21

vector field, 17