Gen Rel

See youtube playlist: International Winter School on Gravity and Light

Lecture 1: Topology

Definition 1 (Topology)**.** Given a set M, a subset of the power set $\mathcal{O} \subseteq \mathcal{P}(M)$ is called a topology if it satisfies the following axioms:

(T1) $M, \emptyset \in \mathcal{O}$.

(T2) If $U, V \in \mathcal{O}$, then $U \cap V \in \mathcal{O}$.

(T3) If I is an index set over elements in $\mathcal{O}$, then $(\bigcup_I U) \in \mathcal{O}$.

Note: The intersection in (T2) must be finite, but the index set in (T3) may be infinite, allowing infinite unions.

Definition 2 (Open Set)**.** If $U \in \mathcal{O}$, then U is called an *open set.*

Definition 3 (Topological Space)**.** A set M with the topology $\mathcal{O}$ is called a topological space, and is denoted $(M, \mathcal{O})$.

Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be topological spaces.

Definition 4. A map $f : M \to N$ is continuous if for all $V \in \mathcal{O}_N$, $f^{-1}(V) \in \mathcal{O}_M$. That is, the preimage of every open set must be an open set.

Theorem 1. The composition of continuous maps is continuous. That is, if $f : M \to N$, $g : N \to P$ are continuous maps, then $g \circ f : M \to P$ is continuous.

Proof. Easy. □

Definition 5 (Standard topology)**.** In lieu of an alternative, we may use the so-called standard topology for $\mathbb{R}^n$. The standard topology is one for which the 'base' is simply open balls $B(x_0, r) \equiv \{x : d_2(x, x_0) < r\}$.

Definition 6 (Subset Topology)**.** A subset $S \subseteq M$ may inherit the topology $\mathcal{O}_S \equiv \mathcal{O}|_S \equiv \{U \cap S : U \in \mathcal{O}\}$.

Proof. To show $\mathcal{O}_S$ is a topology: check axioms. □

Lecture 2: Topological Manifolds

Definition 7 (Homeomorphism)**.** A function $x : M \to N$ between $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ is called homeomorphic if

1. x is invertible,
2. x is continuous,
3. x^{-1} is continuous.

Definition 8 (d-Dimensional Topological Manifold)**.** The topological space $(M, \mathcal{O})$ is called a d-dimensional topological manifold if $\forall p \in M$ there exists a $U \in \mathcal{O}$ containing p for which there exists a homeomorphism $x : U \to x(U) \subseteq \mathbb{R}^d$ (we may take the subset topology for U and the standard topology for $\mathbb{R}^n$).

Definition 9 (Chart)**.** The pair (U, x) in the above definition is called a chart. That is to say, a chart is a homeomorphism from an open subset U of M to an open subset of euclidean space.

Definition 10 (Chart Transition Map)**.** If we have two charts (U, x), (V, y) with a nonzero overlap $U \cap V \neq \emptyset$, then we may define the chart transition map $y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$. As it is a composition of continuous maps, $x \circ y^{-1}$ is continuous.

Definition 11 (Atlas)**.** The collection $\mathcal{A} := \{(U_\alpha, x_\alpha) | \alpha \in A\}$ (indexed by a set A) of charts is called an atlas of $(M, \mathcal{O})$ if $M = \cup_{\alpha \in A} U_\alpha$.

Lecture 3: Multilinear Algebra

Recall the usual definition of a vector space V over a field $\mathbb{F}$. We will take $\mathbb{F}$ to be $\mathbb{R}$.

Definition 12 (Dual Space)**.** The dual space of V is $V^* := \left\{\phi : V \xrightarrow{\sim} \mathbb{R}\right\} \equiv \mathrm{Hom}(V, \mathbb{R})$ – that is, the set of all linear maps from V to $\mathbb{R}$. V^* is a vector space under the obvious choices for addition and scalar multiplication. Elements of V^* are called covectors.

Note: $\xrightarrow{\sim}$ means the map is linear.

Definition 13 (Tensor)**.** Let V be a vector space. An (r,s)-tensor T over V is a multi-linear map

$$T : \underbrace{V^* \times \ldots \times V^*}_{r \text{ times}} \times \underbrace{V \times \ldots \times V}_{s \text{ times}} \xrightarrow{\sim} \mathbb{R}.$$

(r,s) is called the **valence** of T.

Example 1. Let T be a (1,1)-tensor. Then

$$T(\phi + \psi, v) = T(\phi, v) + T(\psi, v), \qquad T(\lambda\phi, v) = \lambda T(\phi, v),$$

$$T(\phi, v + w) = T(\phi, v) + T(\phi, w), \qquad T(\phi, \lambda v) = \lambda T(\phi, v).$$

Remark 1. A map from V to V contains the same data as a map from $V^* \times V$ to $\mathbb{R}$. (Why?)

Example 2. Consider the inner product in P_n, $\phi(p, q) = \int p(x)q(x)\,\mathrm{d}x$. ϕ is a (0,2)-tensor.

Corollary 1. ϕ is a covector $\iff \phi \in V^* \iff \phi : V \xrightarrow{\sim} \mathbb{R} \iff \phi$ is a (0,1)-tensor.

Theorem 2. If $\dim(V) < \infty$, then $V = (V^*)^*$.

Proof. ¯_(ツ)_/¯ □

Corollary 2. v is a vector $\iff v \in V = (V^*)^* \iff v : V^* \xrightarrow{\sim} \mathbb{R} \iff v$ is a (1,0)-tensor.

Definition 14 (Dual Basis)**.** Suppose we have a basis $e_1, \ldots, e_n$ for V and we want a basis $\epsilon^1, \ldots, \epsilon^n$ for V^*. We typically choose our ϵ's such that $\epsilon^a(e_b) = \delta^a_b$. This uniquely determines our ϵ's.

Example 3. In P_3 let us choose $e_0 = 1, e_1 = x, e_2 = x^2, e_3 = x^3$ as our basis. Then $\epsilon^a = \frac{1}{a!}\partial^a|_{x=0}$ – that is, $\epsilon^a(f) = \frac{1}{a!}\frac{\mathrm{d}^a f}{\mathrm{d}x^a}(0)$.

Definition 15 (Components of a Tensor)**.** Let T be an (r,s)-tensor over V, $\dim(V) < \infty$. Then we may define the $(r+s)^{\dim(V)}$-many real numbers

$$T^{i_1,\ldots,i_r}{}_{j_1,\ldots,j_s} := T(\epsilon^{i_1}, \ldots, \epsilon^{i_r}, e_{j_1}, \ldots, e_{j_s}). \tag{1}$$

Example 4. Let T be a (1,1)-tensor. Then $T^i{}_j = T(\epsilon^i, e_j)$ and

$$\begin{aligned}
T(\phi, v) &= T\left(\sum_{i=1}^{\dim V} \phi_i \epsilon^i, \sum_{j=1}^{\dim V} v^j e_j\right) \\
&= \sum_{i=1}^{\dim V}\sum_{j=1}^{\dim V} \phi_i v^j T(\epsilon^i, e_j) \\
&= \sum_{i=1}^{\dim V}\sum_{j=1}^{\dim V} \phi_i v^j T^i{}_j \\
&\equiv \phi_i v^j T^i{}_j \qquad\qquad \text{(Einstein summation)}
\end{aligned} \tag{2}$$

(Note $\phi_i, v^j \in \mathbb{R}$ are the real-numbered components of ϕ and v.)

From now on we will exclusively use Einstein summation notation.

Lecture 4: Differentiable Manifolds

Suppose on our topological manifold $(M, \mathcal{O})$ we have a curve $\gamma : \mathbb{R} \to U \subseteq M$, and we want to know if γ is differentiable. To do this we can take a chart (U, x) and look at $x \circ \gamma : \mathbb{R} \to \mathbb{R}^d$ and utilise regular undergraduate analysis. However, differentiability is a property that should be independent of our choice of chart, and so for any other chart (V, y) we must also have $y \circ \gamma$ is differentiable. In particular, if we restrict the codomain of γ to $U \cap V$, then from the chart transition map

$$y \circ \gamma = \underbrace{\overbrace{(y \circ x^{-1})}^{\mathbb{R}^d \to \mathbb{R}^d}}_{\text{chart trans: cts}} \circ \underbrace{\overbrace{(x \circ \gamma)}^{\mathbb{R} \to \mathbb{R}^d}}_{\text{differentiable}} .$$

We see that differentiability is not guaranteed since the chart transition map is not necessarily differentiable.

Definition 16 (Maximal Atlas)**.** The maximal atlas $\mathcal{A}_{\max}$ is the atlas containing all possible charts.

Definition 17 (Compatibility)**.** Two charts (U, x) and (V, y) of a topological manifold are called $\square$-compatible if either

(a) $U \cap V = \emptyset$, or

(b) $U \cap V \neq \emptyset$ and both chart transition maps

$$y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$$
$$x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$$

satisfy the $\square$ property.

Remark 2. Lecturer used a flower as his generic placeholder symbol instead of $\square$.

Definition 18. An atlas $\mathcal{A}_\square$ is a $\square$-compatible atlas if any two charts in $\mathcal{A}_\square$ are $\square$-compatible.

Example 5. All atlas' are C^0-compatible since all transition maps are continuous.

Definition 19. A $\square$-manifold is a triple $(M, \mathcal{O}, \mathcal{A}_\square)$, where $(M, \mathcal{O})$ is a topological manifold and $\mathcal{A}_\square \subseteq \mathcal{A}_{\max}$.

We tabulate some types of compatibility here:

$\square$	Description
C^0	Continuous maps under $\mathcal{O}_{\text{standard}}$.
C^k	k-times continuously differentiable.
D^k	k-times differentiable, not necessarily continuously.
C^∞	Infinitely differentiable.
C^ω	Functions for which $\exists$ a multidimensional Taylor series. Note $C^\infty \subsetneq C^\omega$.
$\mathbb{C}^\infty$	Only works in even dimensions. The CRE must apply to the transition functions.

Theorem 3 (Whitney's Theorem)**.** Any $C^{k \geq 1}$-atlas $\mathcal{A}_{C^k}$ of a topological manifold contains a C^∞ atlas.

Proof. Not given. ☐

Remark 3. This theorem means that all the hard work is in finding a C^1 atlas – from there we can throw away charts until we end up with a C^∞ atlas.

Definition 20 (Smooth Manifold)**.** A C^∞ manifold $(M, \mathcal{O}, \mathcal{A})$ is called a smooth manifold.

Definition 21 (Homeomorphism of Topological Spaces)**.** Two topological spaces $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ are called topologically isomorphic (or *homeomorphic*) if there exists a bijection $\phi : M \to N$, where ϕ and ϕ^{-1} are both continuous. We write $(M, \mathcal{O}_M) \cong (N, \mathcal{O}_N)$.

Remark 4. The continuity is the 'structure-preserving' part of the whole 'structure-preserving map' schtick.

Definition 22 (Diffeomorphism)**.** Two C^∞-manifolds $(M, \mathcal{O}_M, \mathcal{A}_M)$ and $(N, \mathcal{O}_N, \mathcal{A}_N)$ are said to be diffeomorphic if there exists a bijection $\phi : M \to N$ such that both ϕ and ϕ^{-1} are C^∞.

Remark 5. As per usual we really need to check that for charts (U, x) (where $x : U \to \mathbb{R}^d$) of M and (V, y) (where $y : V \to \mathbb{R}^e$) of N that $y^{-1} \circ \phi' \circ x$ is a diffeomorpism, and once again this involve making sure our chart transition maps are C^∞ etc. so that the diffeomorphism is independent of our choice of chart.

Theorem 4. The number of C^∞ manifolds one can make out of a given C^0 manifold (if any) up to a diffeomorphism are:

- 1 for dimensions 1,2,3.
- ? for dimension 4 (at least one, possibly infinitely many).
- Finitely many for dimensions > 4.

See wikipedia page on Differential Structure for more information.

Lecture 5: Tangent Spaces

Lead question: what is the velocity of a curve γ at the point p?

Definition 23 (Velocity)**.** Let $(M, \mathcal{O}, \mathcal{A})$ be a smooth manifold, and let $\gamma : \mathbb{R} \to M$ be at least C^1. Suppose $\gamma(x_0) = p$. The velocity of γ at p is the linear map

$$v_{\gamma,p} : C^\infty(M) \xrightarrow{\sim} \mathbb{R}$$

defined by

$$f \mapsto (f \circ \gamma)'(x_0),$$

where $C^\infty(M) = \{f : M \to \mathbb{R} \mid f \text{ is a smooth function}\}$ is a vector space equipped with the obvious addition and scalar multiplication.

Remark 6. Let's try to develop some intuition. Suppose f is the temperature on the manifold, and we run along the path γ, so that we detect $f \circ \gamma : \mathbb{R} \xrightarrow{\gamma} M \xrightarrow{f} \mathbb{R}$. As we run along we can observe how the temperature changes and obtain a directional derivative for it. If we step out of diff geo for a second this directional derivative could be expressable in terms of some chart/basis as $\vec{\mathbf{v}} \cdot \vec{\boldsymbol{\nabla}} f = v^i(\partial_i f) = (v^i \partial_i) f$. The rebracketed expression $(v^i \partial_i)$ may be thought of as the linear map v defined above – our velocity "vector" – which is being applied to f (we will make this construction precise below). Since we can do this for any f, we can in particular do this to our basis vectors (since $C^\infty(M)$ is a vector space), and recover the components of v (see Eq. (1)). In this way **vectors in differential geometry survive as the directional derivatives they induce.**

Definition 24 (Tangent Vector Space)**.** For each point $p \in M$ define the set called the "tangent space to M at p"

$$T_pM := \{v_{\gamma,p} \mid \gamma \text{ is a smooth curve}\}$$

– that is, the set of all velocities to smooth curves at p.

Remark 7. Our common intuition for the tangent space is as the plane tangent to some surface, but just note that this is not what our definition is! Defining it as such a plane technically means it would embedded in some higher dimension (which physically would mean we're embedding spacetime in something outside the universe), which we don't want. However there are embedding theorems which state it doesn't matter anyway.

Theorem 5. *T_pM forms a vector space.*

Proof. This proof may be omitted unless you really care to know. We check closure under addition and scalar multiplication. We'll do the latter first since it is easier.

$\odot$: (Scalar multiplication)

$$\mathbb{R} \times T_pM \to \mathrm{Hom}(C^\infty(M), \mathbb{R})$$
$$(a \odot v_{\gamma,p})(f) := a \cdot_{\mathbb{R}} v_{\gamma,p}(f)$$

Now need to show that $\exists \tau$ such that $a \odot v_{\gamma,p} = v_{\tau,p}$.

Claim: $\tau : \mathbb{R} \to M$ given by $\lambda \mapsto \gamma(a\lambda + \lambda_0) \equiv (\gamma \circ \mu_a)(x)$, where ($\mu_a : \mathbb{R} \to \mathbb{R}$ is given by $\lambda \mapsto a\lambda + \lambda_0$) does the trick. Now

$$\begin{aligned} \tau(0) &= \gamma(\lambda_0) = p, \text{ and so} \\ v_{\tau,p}(f) &= (f \circ \tau)'(0) \\ &= (f \circ \gamma \circ \mu_a)'(0) \\ &= (f \circ \gamma)'(\lambda_0) \cdot a \\ &= a \cdot v_{\gamma,p} \\ \implies v_{\tau,p} &= a \odot v_{\gamma,p} \end{aligned}$$

$\oplus$: (Addition)

$$T_pM \times TpM \to \mathrm{Hom}(C^\infty(M), \mathbb{R})$$
$$(v_{\gamma,p} \oplus v_{\delta,p})(f) := v_{\gamma,p}(f) +_{\mathbb{R}} v_{\delta,p}(f)$$

Now need to show that $\exists \sigma$ such that $v_{\gamma,p} \oplus v_{\delta,p} = v_{\sigma,p}$.

We make a choice of chart (U, x), for which $p \in U$. We also define λ_0 and λ_1 such that $\gamma(\lambda_0) = p = \delta(\lambda_1)$.

Claim: $\sigma_x : \mathbb{R} \to M$ given by

$$\lambda \mapsto x^{-1}((x \circ \gamma)(\lambda + \lambda_0) +_{\mathbb{R}^d} (x \circ \delta)(\lambda + \lambda_1) -_{\mathbb{R}^d} (x \circ \gamma)(\lambda_0))$$

does the trick. Now

$$
\begin{aligned}
\sigma_x(0) &= x^{-1}((x\circ\gamma)(\lambda_0) + (x\circ\delta)(\lambda_1) - (x\circ\gamma)(\lambda_0)) \\
&= \delta(\lambda_1) = p, \text{ and so} \\
v_{\sigma_x,p}(f) &= (f\circ\sigma_x)'(0) \\
&= (\underbrace{(f\circ x^{-1})}_{\mathbb{R}^d\to\mathbb{R}} \circ \underbrace{(x\circ\sigma_x)}_{\mathbb{R}\to\mathbb{R}^d})'(0) \\
&= (x\circ\sigma_x)^{i\,\prime}(0)\cdot(\partial_i(f\circ x^{-1}))(x(\sigma_x(0))) && \text{(chain rule)} \\
&= ((x\circ\gamma)^{i\,\prime}(\lambda_0) + (x\circ\delta)^{i\,\prime}(\lambda_1))\cdot(\partial_i(f\circ x^{-1}))(x(\sigma_x(0))) \\
&= ((x\circ\gamma)^{i\,\prime}(\lambda_0)\cdot(\partial_i(f\circ x^{-1}))(x(p)) + (x\circ\delta)^{i\,\prime}(\lambda_1))\cdot(\partial_i(f\circ x^{-1}))(x(p)) \\
&= (f\circ\gamma)'(\lambda_0) + (f\circ\delta)'(\lambda_1) && \text{(chain rule in reverse)} \\
&= v_{\gamma,p}(f) + v_{\delta,p}(f) \\
\implies v_{\sigma,p} &= v_{\gamma,p}\oplus v_{\delta,p}.
\end{aligned}
$$

Note that while we used the chart (U,x) along the way, the result obtained is chart-independent. $\square$

Definition 25 (Components of a Vector w.r.t a Chart)**.** Let $(U,x)\in\mathcal{A}_{smooth}$, and let $\gamma:\mathbb{R}\to U$ with $\gamma(0)=p$. We calculate

$$
\begin{aligned}
v_{\gamma,p}(f) &= (f\circ\gamma)'(0) \\
&= (\underbrace{(f\circ x^{-1})}_{\mathbb{R}^d\to\mathbb{R}} \circ \underbrace{(x\circ\gamma)}_{\mathbb{R}\to\mathbb{R}^d})'(0) \\
&= \underbrace{(x\circ\gamma)^{i\,\prime}(0)}_{=:\dot\gamma_x^i(0)}\cdot\underbrace{(\partial_i(f\circ x^{-1}))(x(p))}_{=:\left(\frac{\partial f}{\partial x^i}\right)_p} \\
&\equiv \dot\gamma_x^i(0)\cdot\left(\frac{\partial}{\partial x^i}\right)_p f.
\end{aligned}
$$

Following tradition we delete the brackets so that $v_{\gamma,p}(f)\to v_{\gamma,p}f$ to follow the notion of a vector acting on whatever is to its right. Thus we ignore the f and write

$$v_{\gamma,p} = \dot\gamma_x^i(0)\left(\frac{\partial}{\partial x^i}\right)_p.$$

We understand the $\dot\gamma_x^i(0)$ to be the components of the velocity vector and $\left(\frac{\partial}{\partial x^i}\right)_p$ to be the basis elements of the tangent space T_pM with respect to which the components are understood – the so-called "chart-induced basis of T_pM".

Remark 8. We need to be careful with $\left(\frac{\partial}{\partial x^i}\right)_p$ to ensure we don't confuse it with the usual partial derivative. We must always keep in mind its true meaning $(\partial_i(f\circ x^{-1}))(x(p))$. However, it turns out that it behaves as one would expect a partial derivative to, so the notation is handy nonetheless.

Definition 26 (Chart-induced Basis)**.** Let $(U,x)\in\mathcal{A}_{smooth}$. Then

$$\left(\frac{\partial}{\partial x^1}\right)_p,\ldots,\left(\frac{\partial}{\partial x^d}\right)_p\in T_pU\subseteq T_pM$$

constitutes a basis for T_pU (*not* T_pM).

Proof. We have already seen that these generate T_pU since we can represent any vector in this basis. It remains to be checked that they are linearly independent. Recall the relevant equation for linear independence:

$$\lambda^i\left(\frac{\partial}{\partial x^i}\right)_p = 0.$$

We apply this to the j^{th} coordinate function $x^j:U\to\mathbb{R}$ (which we recall is differentiable):

$$
\begin{aligned}
\lambda^i\left(\frac{\partial}{\partial x^i}\right)_p(x^j) &= \lambda^i(\partial_i(x^j\circ x^{-1}))(x(p)) \\
&= \lambda^i\delta_i^j = \lambda^j \\
\implies \lambda^j &= 0, \qquad j=1,\ldots,d,
\end{aligned}
$$

where the second line follows from the fact that $x^j\circ x^{-1}:\mathbb{R}^d\to\mathbb{R}$ is given by $(\alpha^1,\ldots,\alpha^d)\mapsto\alpha^j$, so that the derivative is δ_i^j. $\square$

Corollary 3. $\dim(T_pM) = d = \dim(M)$.

Note: just to remind ourselves, if $X \in T_pM$, then

1. $\exists \gamma : \mathbb{R} \to M$ such that $X = v_{\gamma,p}$,
2. $\exists X^1, \dots, X^d$ such that $X = X^i \left(\frac{\partial}{\partial x^i}\right)_p$ in terms of a chosen chart.

We now try to see how the vector components change under a change of chart. Let (U, x) and (V, y) be overlapping charts and $p \in U \cap V$. Let $X \in T_pM$ so that

$$X = X^i_{(x)} \left(\frac{\partial}{\partial x^i}\right)_p = X^i_{(y)} \left(\frac{\partial}{\partial y^i}\right)_p \tag{3}$$

Now

$$\begin{aligned}
\left(\frac{\partial}{\partial x^i}\right)_p f &= \partial_i (f \circ x^{-1})(x(p)) \\
&= \partial_i (\underbrace{(f \circ y^{-1})}_{\mathbb{R}^d \to \mathbb{R}} \circ \underbrace{(y \circ x^{-1})}_{\mathbb{R}^d \to \mathbb{R}^d})(x(p)) \\
&= (\partial_i (y \circ x^{-1}))^j (x(p)) \cdot (\partial_j (f \circ y^{-1}))(y(p)) \\
&= (\partial_i (y^j \circ x^{-1}))(x(p)) \cdot (\partial_j (f \circ y^{-1}))(y(p)) \\
&= \left(\frac{\partial y^j}{\partial x^i}\right)_p \cdot_{\mathbb{R}} \left(\frac{\partial f}{\partial y^j}\right)_p \\
&= \left(\frac{\partial y^j}{\partial x^i}\right)_p \cdot \left(\frac{\partial}{\partial y^j}\right)_p f \\
\implies \left(\frac{\partial}{\partial x^i}\right)_p &= \left(\frac{\partial y^j}{\partial x^i}\right)_p \cdot \left(\frac{\partial}{\partial y^j}\right)_p .
\end{aligned}$$

Feeding this into Eq. (3) gives

$$\begin{aligned}
X^i_{(x)} \left(\frac{\partial y^j}{\partial x^i}\right)_p \cdot \left(\frac{\partial}{\partial y^j}\right)_p &= X^j_{(y)} \left(\frac{\partial}{\partial y^j}\right)_p \\
\implies X^j_{(y)} &= \left(\frac{\partial y^j}{\partial x^i}\right)_p X^i_{(x)}
\end{aligned}$$

Definition 27 (Cotangent Space)**.** The cotangent space of a given tangent space is simply the dual:

$$(T_pM)^* := \{\phi : T_pM \xrightarrow{\sim} \mathbb{R}\}.$$

Remark 9. T_p^*M is a common alternative notation for $(T_pM)^*$.

Example 6. Let $f \in C^\infty(M)$. Define $(df)_p : T_pM \xrightarrow{\sim} \mathbb{R}$ by $X \mapsto Xf$, i.e., $(df)_p \in (T_pM)^*$, which means it is a (0,1)-tensor.

$(df)_p$ is called the gradient of f at $p \in M$. Let us calculate the components of the gradient wrt the chart-induced basis (U, x).

$$\begin{aligned}
((df)_p)_j &= (df)_p \left(\left(\frac{\partial}{\partial x^j}\right)_p \right) \\
&= \left(\frac{\partial f}{\partial x^j}\right)_p \quad (= \partial_j (f \circ x^{-1})(x(p))).
\end{aligned}$$

Theorem 6. Consider the chart $(U, x) \implies x^i : U \to \mathbb{R}$. Then

$$(dx^1)_p, (dx^2)_p, \dots, (dx^d)_p$$

is a basis of T_p^*M – in fact, it is the *dual* basis as we see that

$$(dx^a)_p \left(\left(\frac{\partial}{\partial x^b}\right)_p \right) = \left(\frac{\partial x^a}{\partial x^b}\right)_p = \dots = \delta^a_b .$$

Let us now see how the components of a covector change under the change of a chart. Let $\omega \in T_p^*M$ so that $\omega = \omega_{(x)i}(dx^i)_p = \omega_{(y)j}(dy^j)_p$. Following the same procedure as above, we end up with

$$\omega_{(y)i} = \left(\frac{\partial x^j}{\partial y^i}\right)_p \omega_{(x)j} .$$

Proof. Left as an exercise for the reader. ☐

Remark 10. If we view these transformations as matrix multiplications, we see the transformation of a covector is the inverse of that of a vector.

Lecture 6: Fields

Note: a lot of the intuitive motivations for the definitions and whatnot come from diagrams. However I can't be bothered making LaTeXdiagrams, so rip.

Definition 28 (Bundle)**.** A bundle is a triple (E, π, M), written $E \xrightarrow{\pi} M$, where E is a smooth manifold called the "total space", M is a smooth manifold called the "base space", and π is a smooth surjective map called the "projection map".

Example 7. The prototypical example the lecturer used is as follows. E is a hollow cylinder and M is a circle that we can take to coincide with a cross-section of the cylinder. π simply projects points on the cylinder onto the circle.

Definition 29 (Fibre)**.** Let $E \xrightarrow{\pi} M$ be a bundle, and $p \in M$. We define the *fibre over* p to be $\pi^{-1}(p)$ – that is, the preimage of p.

Definition 30 (Section)**.** A section σ of a bundle $E \xrightarrow{\pi} M$ is a map $\sigma : M \to E$ satisfying $\pi \circ \sigma = 1_M$ – that is, each $p \in M$ is mapped to an element of its fibre.

Remark 11. In what follows we will see that the base space M plays the role of our manifold, the fibres play the role of the tangent vector space, and the sections play the role of the field. Depending on the type of vector space, it can be a vector field, a covector field, a tensor field, etc.

Definition 31 (Tangent Bundle of Smooth Manifold)**.** Let $(M, \mathcal{O}, \mathcal{A})$ be a smooth manifold. Define

(a) as a set, the *tangent bundle*

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

where $\dot{\cup}$ means disjoint union.

(b) the surjective map $\pi : TM \to M$, which maps $X \mapsto p$, the unique point $p \in M$ such that $X \in T_pM$.

Our situation is now

$$\underset{\text{set}}{TM} \quad \underset{\text{surjective map}}{\xrightarrow{\pi}} \quad \underset{\text{smooth mfold}}{M}.$$

To make this a bundle we need to make TM a smooth manifold. To this end, we...

(c) construct the topology on TM that is the coarsest topology such that π becomes (just) continuous (the so-called "initial topology wrt π"). That is,

$$\mathcal{O}_{TM} := \{\pi^{-1}(U) \mid U \in \mathcal{O}\},$$

which can be shown to in fact be a topology. We now need to construct a C^∞ atlas. Define

$$\mathcal{A}_{TM} := \{(TU, \xi_x) \mid (U, x) \in \mathcal{A}\},$$

where $\xi_x : TU \to \mathbb{R}^{2 \cdot \dim M}$ (since the tangent space has $\dim M$ and the space itself has $\dim M$, having a tangent space for each point gives $2 \dim M$) is defined by

$$X \mapsto (\underbrace{(x^1 \circ \pi)(X), \ldots, (x^d \circ \pi)(X)}_{(U,x)\text{-coords of } p=\pi(X)}, \underbrace{(dx^1)_{\pi(X)}(X), \ldots, (dx^d)_{\pi(X)}(X)}_{\text{coords of } X \text{ wrt the chosen chart}}).$$

(We recall that if $X = X^i_{(x)} \left(\frac{\partial}{\partial x^i}\right)_{\pi(X)}$, then $(dx^j)_{\pi(X)}(X) = X^j_{(x)}$.) Note that $\xi_x^{-1} : \xi_x(TU) \to TU$ maps $(\alpha^1, \ldots, \alpha^d, \beta^1, \ldots, \beta^d) \mapsto \beta^i \left(\frac{\partial}{\partial x^i}\right)_{x^{-1}(\alpha_1, \ldots, \alpha^d)} = \beta^i \left(\frac{\partial}{\partial x^i}\right)_{\pi(X)}$.

We need to check smoothness by verifying that the transition map between (U, x) and (V, y) with $U \cap V \neq \emptyset$ is smooth.

$$\begin{aligned}
(\xi_y \circ \xi_x^{-1})(\alpha^1, \ldots, \alpha^d, \beta^1, \ldots, \beta^d) &= \xi_y \left(\beta^i \left(\frac{\partial}{\partial x^i} \right)_{x^{-1}(\alpha_1, \ldots, \alpha^d)} \right) \\
&= \left(\ldots, (y^i \circ \pi) \left(\beta^m \left(\frac{\partial}{\partial x^m} \right)_{x^{-1}(\alpha_1, \ldots, \alpha^d)} \right), \ldots, \right. \\
&\qquad \left. \ldots, (dy^i)_{x^{-1}(\alpha \cdot \ldots, \alpha^d)} \left(\beta^m \left(\frac{\partial}{\partial x^m} \right)_{x^{-1}(\alpha_1, \ldots, \alpha^d)} \right) \right) \\
&= \left(\underbrace{\ldots, (y^i \circ x^{-1})(\alpha^1, \ldots, \alpha^d), \ldots}_{\text{the familiar chart transition map}}, \underbrace{\ldots, \beta^m \left(\frac{\partial y^i}{\partial x^m} \right)_{x^{-1}(\alpha^1, \ldots, \alpha^d)}, \ldots}_{\text{the familiar change of components}} \right) \\
&= (\ldots, (y^i \circ x^{-1})(\alpha^1, \ldots, \alpha^d), \ldots, \ldots, \beta^m \partial_m (y^i \circ x^{-1})(\alpha^1, \ldots, \alpha^d), \ldots)
\end{aligned}$$

We see that the chart transition map and its first derivatives show up, and since the chart transition map is smooth, $\xi_y \circ \xi_x^{-1}$ is therefore smooth too.

Remark 12. That was a lot of work.

Definition 32 (Smooth Vector Field/Smooth Section)**.** A *smooth vector field* χ is a smooth map $\chi : M \to TM$ that is a section of the tangent bundle (i.e., $\pi \circ \chi = 1_M$). (Note the map goes in the opposite direction to the projection map.)

Remark 13. The reason we put in so much effort ensuring everything is smooth for the tangent bundle is because as a consequence our vector field is now smooth.

Definition 33 (The $C^\infty(M)$-module $\Gamma(TM)$)**.** We first note that $C^\infty(M)$ forms a ring when equipped with addition and multiplication (it is almost a field, but doesn't have multiplicative inverses).

We define the set $\Gamma(TM) := \{\chi : M \to TM \mid \chi \text{ is a smooth section}\}$. Let us equip this set with the following operations (recall that the resultant addition/multiplication in the RHS occurs in $C^\infty(M)$, see remark below):

$$(\chi \oplus \widetilde{\chi})(f) := \chi(f) + \widetilde{\chi}(f),$$
$$(g \odot \chi)(f) := g \cdot \chi(f), \qquad g \in C^\infty(M).$$

Then $(\Gamma(TM), \oplus, \odot)$ forms a *module* over $C^\infty(M)$ (recall a module is like a vector space, but over a ring instead of over a field).

Remark 14 (My Comment). We see that in the above definition we applied our vector fields to some f (a la $\chi(f)$) rather than to a point p (a la $\chi(p)$) despite the fact that if $\chi \in \Gamma(TM)$, it should be a function on M, not $C^\infty(M)$. Let's examine this. Since an element of TM is an element of some T_pM and therefore a function on $C^\infty(M)$, we in fact have

$$\chi : M \to TM = M \to (C^\infty(M) \to \mathbb{R}) = M \to C^\infty(M) \to \mathbb{R}.$$

Essentially what we did was decide to go for the middle man, so that

$$\chi(f) : M \to \cancel{C^\infty(M)} \to \mathbb{R} = M \to \mathbb{R}, \text{ or in other words } \chi(f) \in C^\infty(\mathbb{R}),$$

which is perfectly valid if you think about it. We just need to be careful in ensuring that we know in which space each object lies, so that we don't confuse ourselves.

We now have that the set of all smooth vector fields can be made into a $C^\infty(M)$-module.

Facts: 1) if you accept the axiom of choice, every vector space has a basis, 2) no such result holds for modules. This is a shame, as otherwise we could have chosen (for any manifold) vector fields $\chi_{(1)}, \dots, \chi_{(d)} \in \Gamma(TM)$ and be able to write every vector field χ as $\chi = f^i \chi_{(i)}$.

Example 8. See the problem of combing a sphere. That is, if one tries to concoct a smooth vector field over a sphere, there will always be at least one point where the field is zero. This means trying to find a basis set fails at this point.

We can however take a chart map and locally construct a vector space of fields. That is, we have the basis fields $\frac{\partial}{\partial x^i} : U \xrightarrow{\text{smooth}} TU$ given by $p \mapsto \left(\frac{\partial}{\partial x^i}\right)_p$. Just remember that this isn't a thing you can do globally, so this is a blow to the idea that we can do everything by going to a chart.

Let us move on to tensor fields. So far: $\Gamma(TM)$ = "set of vector fields" ($C^\infty(M)$-module), and $\Gamma(T^*M)$ = "set of covector fields" ($C^\infty(M)$-module).

Definition 34 (Tensor Field)**.** An (r,s)-tensor field T is a multi-linear map

$$T : \underbrace{\Gamma(T^*M) \times \ldots \times \Gamma(T^*M)}_{r \text{ times}} \times \underbrace{\Gamma(TM) \times \ldots \times \Gamma(TM)}_{s \text{ times}} \xrightarrow{\sim} C^\infty(M)$$

Example 9. For $f \in C^\infty(M)$, let us define $df : \Gamma(TM) \xrightarrow{\sim} C^\infty(M)$ by $\chi \mapsto \chi f$, where we define χf by: for $p \in M$, $(\chi f)(p) := (\chi(p))(f)$ (recall $\chi(p) \in T_pM$). Then df is a (0,1)-tensor field. One can also check that df is C^∞-linear in χ.

Remark 15. The above example is precisely analogous to Example 6, where we defined $(df)_p$.

Remark 16. If T is a (0,0)-tensor field, then, having no inputs, it is completely characterised by its output. As such we simply write $T \in C^\infty(M)$ – that is, T is a scalar field. This is analogous to how a (0,0)-tensor is just a scalar.

Lecture 7: Connections

Remark 17. In the context of this course, *connections* may also be referred to as *covariant derivatives*. Though the distinction is unimportant to us because it does not arise, it is worth noting that in a general setting, connections are more general than covariant derivatives.

Remark 18. Prior to now we used X to denote vectors and χ to denote vector fields. From now on we will use X for vector fields.

So far: we've seen that a vector field X can be used to provide a directional derivative $\nabla_X f := Xf = (df)(X)$ of a smooth function $f \in C^\infty(M)$.

Is this notational overkill?

$$\nabla_X f = Xf = (df)(X)$$

We seem to have three ways of denoting the same thing. Not quite, for each of the objects are somewhat different:

$$\begin{aligned} X &: C^\infty(M) \to C^\infty(M) \\ df &: \Gamma(TM) \to C^\infty(M) \\ \nabla_X &: C^\infty(M) \to C^\infty(M). \end{aligned}$$

In particular, ∇_M may be extended to take an arbitrary (p,q)-tensor field so that $\nabla_X : (p,q)$-field $\to (p,q)$-field.

Directional derivatives of tensor fields

Now, in the form of a definition we formulate a wishlist of properties which the ∇_X acting on a tensor field should have. (Any remaining freedom in choosing ∇ will need to be provided as additional structure beyond $(M, \mathcal{O}, \mathcal{A})$.)

Definition 35 (Connection)**.** A connection ∇ (aka. linear connection, covariant derivative, affine connection) on a smooth manifold $(M, \mathcal{O}, \mathcal{A})$ is a map that takes a pair consisting of a vector (field) X and a (p,q)-tensor field T and sends them to a (p,q)-tensor (field) $\nabla_X T$ satisfying:

1. (Extension rule) $\nabla_X f = Xf$, $\forall f \in C^\infty(M)$ (i.e., for all (0,0)-tensor fields)
2. (Additivity rule) $\nabla_X(T + S) = \nabla_X T + \nabla_X S$ (of course, T and S must have the same valence)
3. (Leibniz rule) $\nabla_X(\underbrace{T(\omega, Y)}_{\in C^\infty(M)}) = (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y)$. (Here T is a (1,1)-tensor field.) This should hold analogously for an arbitrary (p,q)-tensor field.
4. (C^∞-linearity in the lower entry) $\nabla_{fX+Z} T = f \cdot \nabla_X T + \nabla_Z T$, where $f \in C^\infty(M)$.

Remark 19. Since in axiom 3, $T(\omega, Y) \in C^\infty(M)$, the behaviour of $\nabla_X(T(\omega, V))$ should be completely characterised by axiom 1.

Remark 20. Axiom 3 may be equivalently thought of as the familiar product rule. This becomes more apparent if we use the tensor product formulation instead of the multilinear map formulation, e.g., $V \otimes W$. Alternatively, this equivalence follows easily from looking at Eq. (2), where we went to a chart and wrote $T(\phi, v) = \phi_i v^j T^i{}_j$.

Definition 36 (Manifold with Connection)**.** A manifold with connection (or affine manifold) is a quadruple of structures $(M, \mathcal{O}, \mathcal{A}, \nabla)$.

Remark 21 (Lecturer's remark). If ∇_X is the extension of X, then ∇ is the extension of d. We could indeed have formulated the connection without the attached X, but this would've simply yielded another level of abstraction and obscurity which we do not need.

We now ask what structure on $(M, \mathcal{O}, \mathcal{A})$ is required to fix ∇. How much freedom do we have in choosing such a structure?

We shall study the simple case of vector fields and see what we learn. Consider vector fields X, Y, and let us also use the chart (U, x). Then

$$\begin{aligned} \nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left(Y^j \frac{\partial}{\partial x^j} \right) \\ &= X^i \left(\nabla_{\frac{\partial}{\partial x^i}} Y^j \right) \frac{\partial}{\partial x^j} + X^i Y^j \nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j} \right) && \text{(axioms 3, 4 (and 2 implicitly))} \\ &= X^i \left(\frac{\partial}{\partial x^i} Y^j \right) \frac{\partial}{\partial x^j} + X^i Y^j \Gamma^q{}_{ji} \frac{\partial}{\partial x^q} && \text{(axiom 1, also } \nabla_{\frac{\partial}{\partial x^i}} \left(\tfrac{\partial}{\partial x^j} \right) \text{ is a vector field)} \end{aligned}$$

The last line follows from the fact that though we do not know much about $\nabla_{\frac{\partial}{\partial x^i}}\left(\frac{\partial}{\partial x^j}\right)$, we know that it is at least a vector field, and so we expand it as $\Gamma^q{}_{ji}\frac{\partial}{\partial x^q}$. We call the $\Gamma^q{}_{ji}$ the connection coefficient functions (on M) of ∇ wrt (U, x).

Definition 37. Let $(M, \mathcal{O}, \mathcal{A}, \nabla)$ be a manifold with connection and $(U, x) \in \mathcal{A}$ be a chart. Then the connection coefficient functions (the "Γs") wrt (U, x) are the $(\dim M)^3$-many functions

$$\begin{aligned} \Gamma^i{}_{jk} : U &\to \mathbb{R} \\ p &\mapsto \left(dx^i\left(\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j}\right)\right)(p). \end{aligned}$$

Remark 22. The expression may seem complicated but we can break it down: $\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j}$ is a vector field, and we extract the i^{th} component by acting on it with dx^i.

Remark 23. Just think about the definition of the gammas a bit, and the notion that they be thought of as a way to describe curvature starts to make a lot of sense.

Now we may obtain the components of $\nabla_X Y$:

$$(\nabla_X Y)^i = X^m\left(\frac{\partial}{\partial x^m}Y^i\right) + \Gamma^i{}_{jm}Y^jX^m = X(Y^i) + \Gamma^i{}_{jm}Y^jX^m$$

Remark 24 (Lecturer's Remark). On a chart domain U, choice of the $(\dim M)^3$ functions $\Gamma^i{}_{jk}$ suffices to fix the action of ∇ on a *vector* field. Fortunately, as it so happens, the same $(\dim M)^3$ functions fix the action of ∇ on any tensor field.

Key observation: if we act now on a basis covector dx^i, we cannot say for sure what we get, only that it may be expanded in terms of some other coefficients

$$\nabla_{\frac{\partial}{\partial x^m}}(dx^i) = \Sigma^i{}_{jm}dx^j.$$

We ask if we can express these coefficients in terms of the Gammas. The answer is yes. Observe

$$\nabla_{\frac{\partial}{\partial x^m}}\left(dx^i\left(\frac{\partial}{\partial x^j}\right)\right) = \nabla_{\frac{\partial}{\partial x^m}}(\delta^i_j) = \frac{\partial}{\partial x^m}(\delta^i_j) = 0,$$

whereas using axiom 3

$$\begin{aligned} 0 = \nabla_{\frac{\partial}{\partial x^m}}\left(dx^i\left(\frac{\partial}{\partial x^j}\right)\right) &= \left(\nabla_{\frac{\partial}{\partial x^m}}dx^i\right)\left(\frac{\partial}{\partial x^j}\right) + dx^i\left(\nabla_{\frac{\partial}{\partial x^m}}\frac{\partial}{\partial x^j}\right) \\ &= \left(\nabla_{\frac{\partial}{\partial x^m}}dx^i\right)\left(\frac{\partial}{\partial x^j}\right) + dx^i\left(\Gamma^q{}_{jm}\frac{\partial}{\partial x^q}\right) \\ \implies \left(\nabla_{\frac{\partial}{\partial x^m}}dx^i\right)_j &= -\Gamma^i{}_{jm} \\ \implies \Sigma^i{}_{jm} &= -\Gamma^i{}_{jm}. \end{aligned}$$

Summary (*commit this to memory!*):

$$\begin{aligned} \Gamma^i{}_{jk} &= \left(\nabla_{\frac{\partial}{\partial x^k}}\frac{\partial}{\partial x^j}\right)^i = -\left(\nabla_{\frac{\partial}{\partial x^k}}dx^i\right)_j \\ (\nabla_X Y)^i &= X(Y^i) + \Gamma^i{}_{jm}Y^jX^m \\ (\nabla_X \omega)_i &= X(\omega_i) - \Gamma^j{}_{im}\omega_jX^m \end{aligned}$$

Similarly, by further application of Leibnitz for an e.g., (1,2)-tensor field:

$$(\nabla_X T)^i{}_{jk} = X(T^i{}_{jk}) + \Gamma^i{}_{sm}T^s{}_{jk}X^m - \Gamma^s{}_{jm}T^i{}_{sk}X^m - \Gamma^s{}_{km}T^i{}_{js}X^m.$$

We need to be able to do this ourselves!

Definition 38 (Divergence)**.** Let X be a vector field on $(M, \mathcal{O}, \mathcal{A}, \nabla)$. Then the divergence of X is the function

$$\vec{\boldsymbol{\nabla}} \cdot X := \nabla_{\frac{\partial}{\partial x^i}}X^i.$$

Claim without proof: the divergence is chart-independent.

Change of Gammas under change of chart

Let $(U,x),(V,y)\in\mathcal{A}$ with $U\cap V\neq\emptyset$. Now

$$\begin{aligned}
{\Gamma_{(y)}}^i{}_{jk} &= dy^i\left(\nabla_{\frac{\partial}{\partial y^k}}\frac{\partial}{\partial y_j}\right)\\
&= \frac{\partial y^i}{\partial x^q}dx^q\left(\nabla_{\frac{\partial x^p}{\partial y^k}\frac{\partial}{\partial x^p}}\frac{\partial x^s}{\partial y^j}\frac{\partial}{\partial x^s}\right)\\
&= \frac{\partial y^i}{\partial x^q}dx^q\left(\frac{\partial x^p}{\partial y^k}\left[\left(\nabla_{\frac{\partial}{\partial x^p}}\frac{\partial x^s}{\partial y^j}\right)\frac{\partial}{\partial x^s}+\frac{\partial x^s}{\partial y^j}\left(\nabla_{\frac{\partial}{\partial x^p}}\frac{\partial}{\partial x^s}\right)\right]\right)\\
&= \frac{\partial y^i}{\partial x^q}\underbrace{\frac{\partial x^p}{\partial y^k}\frac{\partial}{\partial x^p}}_{\frac{\partial}{\partial y^k}}\left(\frac{\partial x^s}{\partial y^j}\right)\delta^q_s+\frac{\partial y^i}{\partial x^q}\frac{\partial x^s}{\partial y^j}\frac{\partial x^p}{\partial y^k}{\Gamma_{(x)}}^q{}_{sp}\\
&= \frac{\partial y^i}{\partial x^q}\frac{\partial x^s}{\partial y^j}\frac{\partial x^p}{\partial y^k}{\Gamma_{(x)}}^q{}_{sp}+\frac{\partial y^i}{\partial x^q}\frac{\partial^2 x^q}{\partial y^k\partial y^j}.
\end{aligned}$$

Remark 25. The second term destroys what would've otherwise been your usual transformation of tensor components. If the Gammas are totally zero in one chart it might not be in another one!

Remark 26. This is called the change of connection coefficient functions under change of chart $(U\cap V,x)\to(U\cap V,y)$ and should be committed to memory.

Definition 39 (Symmetrisation & Antisymmetrisation)**.** For any tensor T we can symmetrise it as $T_{(ab)}:=\frac{1}{2}(T_{ab}+T_{ba})$ and antisymmetrise it as $T_{[ab]}:=\frac{1}{2}(T_{ab}-T_{ba})$.

Proposition 1 (Normal coordinates)**.** Let $p\in M$ of $(M,\mathcal{O},\mathcal{A},\nabla)$. Then one can construct a chart (U,x) with $p\in U$ such that

$${\Gamma_{(x)}}^i{}_{(jk)}(p)=0$$

at the point p, not necessarily in any neighbourhood.

Proof. Let (V,y) be any chart with $p\in V$. In general ${\Gamma_{(y)}}^i{}_{jk}\neq 0$ (unless we're lucky). Then consider a new chart (U,x) to which one transits by virtue of

$$(x\circ y^{-1})^i(\alpha^1,\dots,\alpha^d):=\alpha^i-\frac{1}{2}{\Gamma_{(y)}}^i{}_{jk}(p)\alpha^j\alpha^k=\alpha^i-\frac{1}{2}{\Gamma_{(y)}}^i{}_{(jk)}(p)\alpha^j\alpha^k.$$

(Note the jk is symmetrised automatically by the $\alpha^j\alpha^k$.) We have $p=y^{-1}(\alpha^1,\dots,\alpha^d)$, and WLOG we can set all $\alpha^i=0$. Then

$$\begin{aligned}
\left(\frac{\partial x^i}{\partial y^j}\right)_p &= \partial_j(x^i\circ y^{-1})=\delta^i_j-{\Gamma_{(y)}}^i{}_{(mj)}(p)\alpha^m\Big|_{\alpha=0}=\delta^i_j\\
\frac{\partial^2 x^i}{\partial y^k\partial y^j} &= -{\Gamma_{(y)}}^i{}_{(kj)}(p)\\
\implies {\Gamma_{(x)}}^i{}_{jk}(p) &= {\Gamma_{(y)}}^i{}_{jk}-{\Gamma_{(y)}}^i{}_{(kj)}(p)={\Gamma_{(y)}}^i{}_{[jk]}(p)=:{T_{(y)}}^i{}_{jk},
\end{aligned}$$

where T is called the torsion and can be made to vanish (apparently). ☐

Terminology: (U,x) is called a *normal coordinate chart* of ∇ at $p\in M$.

Remark 27. I don't understandu, and as of this writing I'm too lazy to read through the proof with an updated understanding of other things.

Lecture 8: Parallel Transport & Curvature

Parallelity of Vector Fields

Let $(M, \mathcal{O}, \mathcal{A}, \nabla)$ be a manifold with connection.

Definition 40 (Parallely Transported)**.** A vector field X on M is said to be **parallely transported** along a smooth curve $\gamma : \mathbb{R} \to M$ if

$$\nabla_{v_\gamma} X = 0.$$

We can also define a weaker notion of a parallel vector field.

Definition 41 (Parallel)**.** A vector field X is said to be **parallel** along a curve γ if

$$\left(\nabla_{v_{\gamma,\gamma(\lambda)}} X\right)_{\gamma(\lambda)} = \mu(\lambda) X_{\gamma(\lambda)}$$

for some $\mu : \mathbb{R} \to \mathbb{R}$.

Remark 28. The latter definition is probably better formulated pointwise, so we did that. We could've equally well done so for the former definition.

We now consider what happens if we were to "follow our nose" – that is, simply move "forwards".

Definition 42 (Autoparellely Transported Curves)**.** A curve $\gamma : \mathbb{R} \to M$ is called *autoparallely transported* if

$$\nabla_{v_\gamma} v_\gamma = 0, \text{ or}$$
$$\left(\nabla_{v_{\gamma,\gamma(\lambda)}} v_\gamma\right)_{\gamma(\lambda)} = 0.$$

Remark 29 (Lecturer's Remark). As above, we can define a weaker notion of an autoparallel curve: $\nabla_{v_\gamma} v_\gamma = \mu v_\gamma$. Warning: some literature does not make a distinction and refers to what we call "autoparallely transported" as simply "autoparallel".

Example 10. Consider $(\mathbb{R}^2, \mathcal{O}, \mathcal{A}, \nabla_E)$ (∇_E is the Euclidean connection). A curve is autoparallely transported if it is for *constant* straight line motion, and autoparallel if it is for straight line motion.

Autoparallel Equation

Take an autoparalelly transported curve γ and consider that portion of the curve that lies in U, where (U, x) is a chart. We would like to express the autoparallel condition $\nabla_{v_\gamma} v_\gamma = 0$ in terms of the chart representation.

$$\begin{aligned}
\nabla_{v_\gamma} v_\gamma &= \nabla_{\dot\gamma^m_{(x)} \frac{\partial}{\partial x^m}} \dot\gamma^n_{(x)} \frac{\partial}{\partial x^n} \\
&= \underbrace{\dot\gamma^m \frac{\partial \dot\gamma^q}{\partial x^m}}_{\ddot\gamma^q_{(x)}} \cdot \frac{\partial}{\partial x_q} + \dot\gamma^m \dot\gamma^n \Gamma^q{}_{nm} \frac{\partial}{\partial x^q} \qquad \text{(relabelling of dummy indices)} \\
\implies (\nabla_{v_\gamma} v_\gamma)^q &= \ddot\gamma^q + \dot\gamma^m \dot\gamma^n \Gamma^q{}_{nm}
\end{aligned}$$

Remark 30. Lecturer says he could show that $\dot\gamma^m \frac{\partial \dot\gamma^q}{\partial x^m} = \ddot\gamma^q$ over two blackboards (possible exaggeration) but didn't want to bother. TODO: I've tried and I can't get this result.

The autoparallel equation thus becomes

$$\ddot\gamma^m_{(x)}(\lambda) + \Gamma_{(x)}{}^m{}_{ab}(\gamma(\lambda))\, \dot\gamma^a_{(x)}(\lambda)\, \dot\gamma^b_{(x)}(\lambda) = 0, \tag{4}$$

or more succintly,

$$\ddot\gamma^m + \Gamma^m{}_{ab} \dot\gamma^a \dot\gamma^b = 0. \tag{5}$$

We call this the 'chart expression of the condition that γ be autoparallely transported'.

Remark 31. We can ask about the transformation behaviour of the expression on the LHS. Is it a vector? That is, does it transform like a vector? As it so happens, the part of Γ that fails to transform like a tensor is exactly cancelled by the part of $\ddot\gamma$ that fails to transform like a vector, so that the overall quantity is a vector. We may call it the acceleration a and write $a^m = (\nabla_{v_\gamma} v_\gamma)^m = \ddot\gamma^m + \Gamma^m{}_{ab} \dot\gamma^a \dot\gamma^b$.

Example 11. Some autoparallels.

(a) Euclidean plane, $U = \mathbb{R}^2$, $x = \mathbb{I}_{\mathbb{R}^2}$, ${\Gamma_{(x)}}^i{}_{jk} = 0$. Then

$$\ddot{\gamma}^m_{(x)} = 0 \implies \gamma^m_{(x)}(\lambda) = a^m\lambda + b^m, \qquad a, b \in \mathbb{R}^d.$$

This is a straight line! Tada!

(b) Round sphere $(S^2, \mathcal{O}, \mathcal{A}, \nabla_{\text{round}})$. Consider a chart $x(p) = (\theta_p, \phi_p)$, $\theta \in (0, \pi)$, $\phi \in (0, 2\pi)$. For a round sphere we set

$$\begin{aligned} {\Gamma_{(x)}}^1{}_{22}(x^{-1}(\theta, \phi)) &:= -\sin\theta\cos\theta, \\ {\Gamma_{(x)}}^2{}_{12} = {\Gamma_{(x)}}^2{}_{21} &:= \cot\theta, \end{aligned}$$

and all other Gammas to 0. If we use the notation from classical mechanics of $x^1(p) = \theta(p)$, $x^2(p) = \phi(p)$, then the autoparallel equation is

$$\begin{aligned} \ddot{\theta} + \Gamma^1{}_{22}\dot{\phi}\dot{\phi} &= \ddot{\theta} - \sin(\theta)\cos(\theta)\dot{\phi}\dot{\phi} = 0 \\ \ddot{\phi} + 2\Gamma^2{}_{12}\dot{\theta}\dot{\phi} &= \ddot{\phi} + 2\cot(\theta)\dot{\theta}\dot{\phi} = 0. \end{aligned}$$

We observe that we could've obtained these equations using the Euler-Lagrange equation, however its formulation requires the notion of a metric for the kinetic energy term.

For an example, consider moving around the equator of a sphere, with $\theta(\lambda) = \pi/2$, $\phi(\lambda) = \omega\lambda + \phi_0$. This motion satisfies the autoparallel equations, which can be easily verified. That is, running around the equator at constant velocity is an autoparallel curve.

Question: Can we use ∇ to define tensors on $(M, \mathcal{O}, \mathcal{A}, \nabla)$? Well, as it so happens, we may define...

Definition 43 (Torsion)**.** The torsion of a connection ∇ is the (1,2)-tensor field

$$T(\omega, X, Y) := \omega(\nabla_X Y - \nabla_Y X - [X, Y]).$$

Comment: the commutator $[X, Y]$ is defined by $[X, Y]f := X(Yf) - Y(Xf)$.

Proof. Need to check that T is in fact a tensor field – i.e., it is C^∞-linear in each entry. We have

$$\begin{aligned} T(f \cdot \omega, X, Y) &= f \cdot \omega(\ldots) = fT(\omega, X, Y) \\ T(\omega + \phi, X, Y) &= \ldots = T(\omega, X, Y) + T(\phi, X, Y) \\ T(\omega, fX, Y) &= \omega(\nabla_{fX} Y - \nabla_Y(fX) - [fX, Y]) \\ &= \omega(f\nabla_X Y - (Yf)X - f\nabla_Y X - f[X, Y] + (Yf)X) \qquad (6) \\ &= f \cdot \omega(\nabla_X Y - \nabla_Y X - [X, Y]) \\ &= fT(\omega, X, Y) \end{aligned}$$

Eq. (6) follows from

$$[fX, Y]g = fX(Yg) - Y(fXg) = fX(Yg) - (Yf)(Xg) - fY(Xg) = f[X, Y]g - (Yf)Xg.$$

Now since $T(\omega, Y, X) = -T(\omega, X, Y)$ we don't need to check scaling of the last factor; furthermore, out of laziness we don't bother checking additivity of the last two factors. □

Definition 44 (Torsion-free)**.** A manifold $(M, \mathcal{O}, \mathcal{A}, \nabla)$ is called torsion-free if $T = 0$.

In a chart, this becomes

$$T^i{}_{ab} = T\left(dx^i, \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right) = dx^i(\cdots) = \Gamma^i{}_{ab} - \Gamma^i{}_{ba} = 2\Gamma^i{}_{[ab]} = 0.$$

Warning: From now on, in these lectures, we only use torsion-free connections.

Remark 32 (Lecturer's Remark). Schrödinger has an interesting book on general relativity in which he discusses the physical meaning behind torsion in spacetime.

Curvature

Definition 45 (Riemann Curvature)**.** The Riemann Curvature of a connection ∇ is the (1,3)-tensor field

$$\mathrm{Riem}(\omega, Z, X, Y) := \omega\big(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z\big)$$

Proof. To prove C^∞-linearity: justdoit.gif $\square$

It is a straightforward exercise to show that the components of Riem in a chart are

$$\mathrm{Riem}^i{}_{jkl} = \frac{\partial \Gamma^i{}_{jl}}{\partial x^k} - \frac{\partial \Gamma^i{}_{jk}}{\partial x^l} + \Gamma^q{}_{jl}\Gamma^i{}_{qk} - \Gamma^q{}_{jk}\Gamma^i{}_{ql} \equiv \Gamma^i{}_{jl,k} - \Gamma^i{}_{jk,l} + \Gamma^q{}_{jl}\Gamma^i{}_{qk} - \Gamma^q{}_{jk}\Gamma^i{}_{ql},$$

where we introduce a new notation as a convenient way of denoting the derivative. Note this means that if the Γs vanish everywhere in any one chart then $\mathrm{Riem} = 0$.

The algebraic relevance of Riem is the following. Consider

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

which in *one* chart (where we adopt the shorthand notation $\nabla_a := \nabla_{\frac{\partial}{\partial x^a}}$) is

$$(\nabla_a \nabla_b Z)^m - (\nabla_b \nabla_a Z)^m = \mathrm{Riem}^m{}_{nab} Z^n + \cancelto{0}{\nabla_{[\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}]} Z}.$$

This tells us that in a coordinate-induced chart, the Riemann tensor components contain all the information about how the ∇_a's fail to commute when acting on a vector.

As for the geometric significance of Riem: needs diagrams, so I'm not doing it here. See ∼1:19:00 of the lecture video. Essentially, it has to do with parallel transport. If you paralelly transport a vector field along some path from A to B, and along some different path from A to B, and at point B the vector field doesn't end up the same, then Riem essentially tells you all about this.

Lecture 9: Newtonian Spacetime is Curved!

We begin by restating Newton's laws.

Newton I: A body on which no force acts moves uniformly along a straight line.

Newton II: Deviation of a body's motion from such uniform straight motion is effected by a force, reduced by a factor of the body's reciprocal mass.

Remark 33. (Lecturer's remarks)

1. The first axiom – in order to be relevant – must be read as a measurement prescription for the geometry of space.
2. Since gravity universally acts on every particle, in a universe with at least two particles, gravity must not be considered a force if Newton I is supposed to remain applicable.

Laplace first asked the question "Can gravity be encoded in a curvature of space such that its effects show if particles under the influence of (no other) force are postulated to move along straight lines in this curved space?"
Answer: No! However, it is instructive to see why this doesn't work.

Proof. Under the gravity-is-a-force point of view we have

$$m\ddot{x}^\alpha(t) = \underbrace{mf^\alpha}_{F^\alpha}(x(t)),$$

where f satisfies Poisson's equation $-\partial_\alpha f^\alpha = 4\pi G\rho$ (ρ is the mass density of matter in the universe). We ask if this can be interpreted as an autoparallel equation a la

$$\ddot{x}^\alpha(t) + \Gamma^\alpha{}_{\beta\gamma}(x(t))\dot{x}^\beta(t)\dot{x}^\gamma(t) = 0.$$

Clearly it can not due to the $\dot{x}$ factors. We conclude that one can't find Γs such that Newton's equation takes the form of an autoparallel. $\square$

The problem lies in that Laplace didn't heed the full wisdom of Newton I, which states that particles under no (other) force move *uniformly*. We tried to implement autoparallel motion but failed because we ignored uniformity under the assumption that it would be a weaker case.

Let us introduce the appropriate setting to talk about the difference easily. (Once again I am too lazy to make diagrams in latex, so use your imagination.) Previously we may have drawn lines by segmenting them such that the same amount of time is elapsed over each segment – in this way we may draw accelerating curves by making the segments progressively longer. Let us, however, use the standard high school strategy of (for one-dimensional motion) graphing the motion by putting time on one axis and displacement on another. We see that uniform motion is a straight line on this graph, but non-uniform motion isn't (e.g., it could be a parabola $x(t) = t^2$). This yields the insight: uniform straight motion in space becomes straight motion in **spacetime**.

Let us therefore reattempt our construction in spacetime. First, however, we need to convert the curve data to the right form.

Let $x : \mathbb{R} \to \mathbb{R}^3$ be a particle's trajectory in space. We convert this to a 'worldline' $X : \mathbb{R} \to \mathbb{R}^4$ given by

$$t \mapsto (X^0(t), X^1(t), X^2(t), X^3(t)) := (t, x^1(t), x^2(t), x^3(t)).$$

Since $\dot{X}^0 = 1$, Newton's equation $\ddot{x}^\alpha - f^\alpha(x(t)) = 0$ now becomes

$$\begin{aligned} &\ddot{X}^0 &&= 0 \\ &\ddot{X}^\alpha - f^\alpha(X(t)) \cdot \dot{X}^0 \cdot \dot{X}^0 &&= 0 \qquad (\alpha = 1, 2, 3), \end{aligned}$$

which may be written more concisely as

$$\ddot{X}^a + \Gamma^a{}_{bc}\dot{X}^b\dot{X}^c = 0, \qquad (a = 0, 1, 2, 3).$$

This is exactly an autoparallel equation in Newtonian spacetime under the choices

$$\begin{aligned} \Gamma^0{}_{ab} &= 0 \\ \Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{0\beta} = \Gamma^\alpha{}_{\beta 0} &= 0 \\ \Gamma^\alpha{}_{00} &= -f^\alpha. \end{aligned}$$

Remark 34. The nonzero Gammas ${\Gamma^\alpha}_{00} = -f^\alpha$ feature time indices, and as such ${\mathrm{Riem}^\alpha}_{\beta\gamma\delta} = 0$ for $\alpha, \beta, \gamma, \delta = 1, 2, 3$. This tells us that space itself is still flat; curvature only appears when time comes into the fray.

Question: Is this an artefact of our coordinate choice? No, since the only non-vanishing components of the Riemann curvature are ${\mathrm{Riem}^\alpha}_{0\beta 0} = -\frac{\partial}{\partial x^\beta} f^\alpha$. (This is known as the 'tidal force tensor', and is minus the hessian of the gravitational potential.) Let's jump ahead a bit for fun. We can compute the Ricci tensor $R_{00} = {\mathrm{Riem}^m}_{0m0} = -\partial_\alpha f^\alpha = 4\pi G\rho$ (due to Poisson's equation). Writing an energy-momentum tensor $T_{00} = \frac{1}{2}\rho$ we arrive at $R_{00} = 8\pi G T_{00}$, which is one of ten gravitational equations that Einstein postulated in 1912.

Conclusion: Laplace's idea works in spacetime. We have thus reformulated classical gravity as a curvature in Newtonian spacetime.

The Foundations of the Geometric Formulation of Newton's Axioms

We note that we did everything above in terms of chart representations, so let's take a step back into the abstract.

Definition 46. A **Newtonian spacetime** is a quintuple $(M, \mathcal{O}, \mathcal{A}, \nabla, t)$, where $(M, \mathcal{O}, \mathcal{A})$ is a 4-dimensional smooth manifold and $t : M \to \mathbb{R}$ is a smooth function called the absolute time. Newton stated the following, which we write as axioms:

(i) "There is an absolute space". That is, $(dt)_p \neq 0$ for all $p \in M$.

(ii) "Absolute time flows uniformly". That is, $\nabla dt = 0$ everywhere.

(iii) ∇ is torsion-free.

To expand on axiom (i), we may define:

Definition 47 (Absolute space)**.** The absolute space at time τ is $S_\tau := \{p \in M \mid t(p) = \tau\}$. Since $dt \neq 0$ this implies $M = \dot{\cup} S_\tau$.

Definition 48. A vector $X \in T_pM$ is called

(a) future-directed if $dt(X) > 0$,

(b) spatial if $dt(X) = 0$, and

(c) past-directed if $dt(X) < 0$.

We may now update Newton's laws in our new machinery.

Newton I: The worldline of a particle under the influence of no force (gravity isn't one, anyway) is a future-directed autoparallel.

Newton II: $\nabla_{v_X} v_X = F/m$, or alternatively $ma = F$, where F is a spatial vector field (and a is future-directed).

We have the following incredibly practical convention. We restrict our attention to the atlas' $A_{\text{stratified}}$ whose charts (U, x) have the property $x^0 = t|_U$. In such a stratified atlas we have that

$$\begin{aligned} 0 &= \nabla dt \\ &= \left(\nabla_{\frac{\partial}{\partial x^a}} dx^0\right)_b = -{\Gamma^0}_{ba}, \qquad a, b = 0, 1, 2, 3. \end{aligned}$$

Let's evaluate Newton II in a chart (U, x) of a stratified atlas A_{strat}. We have

$$X^{0\prime\prime} + {\Gamma^0}_{ab} X^{a\prime} X^{b\prime} = 0, \text{ and} \tag{6a}$$

$$X^{\alpha\prime\prime} + {\Gamma^\alpha}_{\gamma\delta} X^{\gamma\prime} X^{\delta\prime} + {\Gamma^\alpha}_{00} X^{0\prime} X^{0\prime} + 2{\Gamma^\alpha}_{\gamma 0} X^{\gamma\prime} X^{0\prime} = \frac{F^\alpha}{m}, \tag{6b}$$

where the $2{\Gamma^\alpha}_{\gamma 0} = 2{\Gamma^\alpha}_{(\gamma 0)}$ comes from ∇ being torsion-free. Since in our stratified atlas ${\Gamma^0}_{ab} = 0$, Eq. (6a) becomes $X^{0\prime\prime} = 0$, which implies that $(t \circ X)(\lambda) = (x^0 \circ X)(\lambda) = X^0(\lambda) = a\lambda + b$.

By convention we parametise the worldline by the absolute time, so that $\frac{\mathrm{d}}{\mathrm{d}\lambda} = a\frac{\mathrm{d}}{\mathrm{d}t}$. Then Newton II becomes

$$\ddot{X}^\alpha + {\Gamma^\alpha}_{\gamma\delta} \dot{X}^\gamma \dot{X}^\delta + {\Gamma^\alpha}_{00} + 2{\Gamma^\alpha}_{\gamma 0} \dot{X}^\gamma = \frac{1}{a^2} \frac{F^\alpha}{m},$$

though we usually absorb the $\frac{1}{a^2}$ into F^α.

Remark 35. Suppose we are in a chart that represents a rotating frame – i.e., the spatial component rotates as the time component increases. Then the term ${\Gamma^\alpha}_{00}$ becomes the centrifugal force and $2{\Gamma^\alpha}_{\gamma 0} \dot{X}^\gamma$ becomes the coriolis force.

Lecture 10: Metric Manifolds

We shall now establish a structure on a smooth manifold that allows us to assign vectors in a tangent space a length (and an angle between vectors in the same tangent space). From this structure we can then define the length of a curve. Then we may look at shortest curves. Requiring then that the shortest curves coincide with straight curves (wrt a ∇) will result in ∇ being determined by the metric structure.

Metrics

Definition 49 (Metric)**.** A metric g on a smooth manifold $(M, \mathcal{O}, \mathcal{A})$ is a (0,2)-tensor field satisfying:

(i) (Symmetry) $g(X,Y) = g(Y,X)$ for all vector fields X, Y.

(ii) (Non-degeneracy) The so-called 'musical map' $\flat$ is a C^∞-isomorphism – that is, it is invertible. ($\flat : \Gamma(TM) \to \Gamma(T^*M)$ is defined by $\flat(X)(Y) := g(X,Y)$.)

Remark 36. A familiar example of the 'musical isomorphism' is quantum mechanics, where we may write $g(\psi, \phi) = \langle \psi | \phi \rangle$ and $\flat(|\psi\rangle) = \langle \psi |$.

Component-wise we may write $(\flat(X))_a := g_{am} X^m = g_{ma} X^m$ and similarly $(\flat^{-1}(\omega))^a := (g^{-1})^{am} \omega_m$. This leads us to define...

Definition 50. The (2,0)-tensor field g^{-1} wrt a metric g is the symmetric tensor

$$\begin{aligned} g^{-1} : \Gamma(T^*M) \times \Gamma(T^*M) &\xrightarrow{\sim} C^\infty(M) \\ (\omega, \sigma) &\mapsto \omega(\flat^{-1}(\sigma)). \end{aligned}$$

(Note that even though we denote it as such, strictly speaking g^{-1} is not the inverse of g.)

In a chart we have $g_{ab} = g_{ba}$ and $(g^{-1})^{am} g_{mb} = \delta^a_b$.

Convention: We often write $g^{ab} := (g^{-1})^{ab}$, $X_a := g_{am} X^m$, and $\omega^a := g^{am} \omega_m$, and so on for higher order tensors.

Example 12. Take $(S^2, \mathcal{O}, \mathcal{A})$ and the chart (U, x) defined in an earlier example. We define the metric

$$g_{ij}(x^{-1}(\theta, \phi)) = \begin{bmatrix} R^2 & 0 \\ 0 & R^2 \sin^2\theta \end{bmatrix}_{ij}, \qquad R \in \mathbb{R}^+.$$

We call this "the metric of the round sphere of radius R".

We now step back to linear algebra and recall the eigenvalue equation $A^a{}_m v^m = \lambda v^a$. Thus we may say that any (1,1) tensor has eigenvalues (recall that a linear map is equivalently a (1,1) tensor). In its eigenbasis we may write A as a diagonal matrix with its eigenvalues as its entries. Now, we can't do the same thing with a (0,2)-tensor as it does not have eigenvalues. However, as it so happens we may write it as a diagonal matrix with some number of 1s, -1s and 0s. Thus we define the **signature** of a (0,2)-tensor to be (p, q), where p is the number of 1s, q is the number of -1s, and $\dim V - p - q = d - p - q$ is the number of 0s. (The non-degeneracy condition of a metric ensures that $p + q = \dim V$ and so there will in fact be no 0s.)

Remark 37 (My Remark)*.* For my own sake I'm going to formalise the above paragraph a little better.

Definition 51. A metric is called *positive definite* if $g(X,X) \geq 0$ for all vector fields X at all points on the manifold, and *negative definite* if $g(X,X) \leq 0$.

Proposition 2. For any metric we may in each chart choose a set of basis vectors $\{X_i\}$ such that

$$g(X_i, X_i) = \begin{cases} 1, & i \leq p \\ -1, & i > p \end{cases}.$$

Definition 52 (Metric Signature)**.** We define the signature of a metric to be the pair $(p, q) := (p, d - p)$ in the above definition.

Definition 53. A metric is called Riemannian if its signature is $(d, 0)$ or $(0, d)$; otherwise it is called pseudo-Riemannian. A special case of pseudo-Riemannian is $(d-1, 1)$ or $(1, d-1)$, in which case we call it Lorentzian.

Lengths

Now that we have a metric it is sensible to define the length of a curve.

Definition 54. On a Riemannian metric manifold $(M, \mathcal{O}, \mathcal{A}, g)$ the speed of a curve γ at $\gamma(\lambda)$ is the number

$$s(\lambda) := \left(\sqrt{g(v_\gamma, v_\gamma)}\right)_{\gamma(\lambda)}.$$

Remark 38. In a much earlier lecture the lecturer claimed that velocity has units $[v] = 1/T$. We can reconcile this now. Since $[g_{ab}] = L^2$ we now have $\left[\sqrt{g_{ab}v^a v^b}\right] = \sqrt{L^2/T^2} = L/T$.

Definition 55. Let $\gamma : (0,1) \to M$ be a smooth curve. Then the length of γ is the number

$$L[\gamma] := \int_0^1 s(\lambda)\, d\lambda = \int_0^1 \sqrt{g(v_\gamma, v_\gamma)}_{\gamma(\lambda)}\, d\lambda.$$

Example 13. Reconsider the round sphere of radius R. Consider its equator parametised as

$$\begin{aligned}\theta(\lambda) &= (x^1 \circ \gamma)(\lambda) = \pi/2 \\ \phi(\lambda) &= (x^2 \circ \gamma)(\lambda) = 2\pi\lambda^3 + \lambda_0.\end{aligned}$$

Then

$$\begin{aligned}L[\gamma] &= \int_0^1 \sqrt{g_{ij}(x^{-1}(\theta(\lambda), \phi(\lambda))) \cdot (x^i \circ \gamma)'(\lambda) \cdot (x^j \circ \gamma)'(\lambda)}\, d\lambda \\ &= \int_0^1 \sqrt{R^2 \cdot 0 + R^2 \sin^2(\theta(\lambda))(6\pi\lambda^2)^2}\, d\lambda \\ &= 6\pi R \int_0^1 \lambda^2\, d\lambda \\ &= 2\pi R.\end{aligned}$$

Theorem 7. If $\gamma : (0,1) \to M$ is a smooth map and $\sigma : (0,1) \to (0,1)$ is smooth, bijective and monotonically increasing. Then $L[\gamma] = L[\gamma \circ \sigma]$. That is, the length of a curve is independent of parametisation.

Proof. We'll use a chart (U, x) in the interim but find the result to be chart-independent.

$$\begin{aligned}L[\gamma \circ \sigma] &= \int_0^1 \left[g_{ij}(\gamma(\sigma(\lambda))) \cdot (x^i \circ \gamma \circ \sigma)'(\lambda) \cdot (x^j \circ \gamma \circ \sigma)'(\lambda)\right]^{1/2} \mathrm{d}\lambda \\ &= \int_0^1 \left[g_{ij}(\gamma(u)) \cdot (x^i \circ \gamma \circ \sigma)'(\sigma^{-1}(u)) \cdot (x^j \circ \gamma \circ \sigma)'(\sigma^{-1}(u))\right]^{1/2} \left(\frac{\partial \sigma}{\partial \lambda}\right)^{-1} \mathrm{d}u \qquad \text{(set } u = \sigma(\lambda)\text{)} \\ &= \int_0^1 \left[g_{ij}(\gamma(u)) \cdot (x^i \circ \gamma)'(u) \cdot \frac{\mathrm{d}\sigma}{\mathrm{d}\lambda}(\sigma^{-1}(\lambda)) \cdot (x^j \circ \gamma)'(u) \cdot \frac{\mathrm{d}\sigma}{\mathrm{d}\lambda}(\sigma^{-1}(u))\right]^{1/2} \left(\frac{\partial \sigma}{\partial \lambda}\right)^{-1} \mathrm{d}u \qquad \text{(chain rule)} \\ &= \int_0^1 \left[g_{ij}(\gamma(u)) \cdot (x^i \circ \gamma)'(u) \cdot (x^j \circ \gamma)'(u)\right]^{1/2} \mathrm{d}u \\ &= L[\gamma]\end{aligned}$$

□

Definition 56 (Geodesic)**.** A curve $\gamma : (0,1) \to M$ is called a geodesic on a Riemannian manifold $(M, \mathcal{O}, \mathcal{A}, g)$ if it is a stationary curve wrt the length functional L.

Theorem 8. γ is geodesic iff it satisfies the Euler-Lagrange equations for the Lagrangian

$$\begin{aligned}\mathcal{L} : TM &\to \mathbb{R} \\ X &\mapsto \sqrt{g(X, X)}.\end{aligned}$$

In a chart, the Lagrangian becomes

$$\mathcal{L}(\gamma, \dot{\gamma}) = \sqrt{g_{ij}(\gamma(\lambda))\dot{\gamma}^i(\lambda)\dot{\gamma}^j(\lambda)}$$

and the Euler-Lagrange equation is written

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial \mathcal{L}}{\partial \dot{x}^m} - \frac{\partial \mathcal{L}}{\partial x^m} = 0.$$

Let us evaluate it. We have

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} &= \frac{1}{\sqrt{\cdots}} g_{mj}(\gamma(\lambda))\dot{\gamma}^j(\lambda), \\
\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} &= \left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{1}{\sqrt{\cdots}}\right) g_{mj}(\gamma(\lambda))\dot{\gamma}^j(\lambda) + \frac{1}{\sqrt{\cdots}}\big(g_{mj}(\gamma(\lambda))\ddot{\gamma}^j(\lambda) + \dot{\gamma}^i(\partial_i g_{mj})\dot{\gamma}^j(\lambda)\big), \\
\frac{\partial \mathcal{L}}{\partial \gamma^m} &= \frac{1}{2\sqrt{\cdots}} \partial_m g_{ij}(\gamma(\lambda))\dot{\gamma}^i(\lambda)\dot{\gamma}^j(\lambda).
\end{aligned}
$$

Here we make use of the above reparametisation theorem by parametising our curve such that the velocity is a constant 1 (that is, $g(\dot{\gamma}, \dot{\gamma}) = 1$); in this way $1/\sqrt{\cdots} = 1$ and the $\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{1}{\sqrt{\cdots}}$ term vanishes. The Euler-Lagrange equation thus becomes

$$
\begin{aligned}
g_{mj}\ddot{\gamma}^j + \partial_i g_{mj}\dot{\gamma}^i\dot{\gamma}^j - \frac{1}{2}\partial_m g_{ij}\dot{\gamma}^i\dot{\gamma}^j &= 0 \\
\ddot{\gamma}^q + (g^{-1})^{qm}\left(\partial_i g_{mj} - \frac{1}{2}\partial_m g_{ij}\right)\dot{\gamma}^i\dot{\gamma}^j &= 0 \qquad (\times (g^{-1})^{qm}) \\
\ddot{\gamma}^q + (g^{-1})^{qm}\frac{1}{2}(\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij})\dot{\gamma}^i\dot{\gamma}^j &= 0, \qquad (7)
\end{aligned}
$$

where in the last line we have symmetrised the i and j, which is afforded by the presence of $\dot{\gamma}^i\dot{\gamma}^j$. Equation (7) is called the geodesic equation for γ in a chart.

Levi-Civita Connection

Comparing Eq. (7) to Eq. (5) leads us to define the following.

Definition 57 (Christoffel Symbols)**.** We define

$$
{}^{L.C}\Gamma^q{}_{ij}(\gamma(\lambda)) := (g^{-1})^{qm}\frac{1}{2}(\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij}). \qquad (8)
$$

These are the connection coefficient functions of the so-called Levi-Civita connection ${}^{L.C}\nabla$.

Insisting that the geodesics defined by Eq. (7) coincide with autoparallels by introduction of the Christoffel symbols thus yields us the Levi-Civita connection. We usually make this choice of ∇ if g is given.

We may also prescribe the Levi-Civita connection in the following abstract way without making reference to a chart.

Theorem 9. If ∇ satisfies $\nabla g = 0$ and $T = 0$ then $\nabla = {}^{L.C}\nabla$.

Proof. Left as an exercise. Hints: Expand $(\nabla_a g)_{bc}$, $(\nabla_b g)_{ca}$, and $(\nabla_c g)_{ab}$ in terms of the Γs and then, recalling that $T = 0 \implies \Gamma^a{}_{[bc]} = 0$, add/subtract the resultant equations in a clever way to obtain Eq. (8). □

Some Definitions

We chuck these out now because why not.

Definition 58. The Riemann-Christoffel curvature is defined by

$$
R_{abcd} := g_{am}\mathrm{Riem}^m{}_{bcd}.
$$

Remark 39. Note that we defined this in terms of a chart; ordinarily we would want to define it in the abstract but in this case it isn't worth it.

Definition 59. The Ricci tensor is defined by

$$
R_{ab} := R^m{}_{amb}.
$$

Definition 60. The (Ricci) scalar curvature is defined by

$$
R := (g^{-1})^{ab}R_{ab}.
$$

Definition 61. The Einstein curvature is defined by

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

Lecture 11: Symmetry

Recall the round sphere $(S^2, \mathcal{O}, \mathcal{A}, g_{\text{round}})$. We feel that it ought to have rotational symmetry, while the potato $(S^2, \mathcal{O}, \mathcal{A}, g_{\text{potato}})$ should not. This leads us to the question: how do we describe symmetries of a metric? This turns out to be useful down the line since supposedly we can't solve Einstein's equations without knowledge of symmetries in the solution.

Let M, N be smooth manifolds and $\phi : M \to N$ be a smooth map.

Definition 62 (Push-forward map)**.** The push-forward is the map $\phi_* : TM \to TN$, where we define $\phi_*(X)$ by – for any $f \in C^\infty(N)$ –

$$\phi_*(X) := (f \mapsto X(f \circ \phi)).$$

Note: from this construction $\phi_*(T_pM) \subseteq T_{\phi(p)}N$.

We have the following mnemonic: "vectors are pushed forward".

We may compute the components of ϕ_* wrt two charts $(U, x) \in \mathcal{A}_M$, $(V, y) \in A_N$:

$$\phi_*{}^a{}_i := dy^a\left(\phi_*\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\right) := \phi_*\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)y^a = \left(\frac{\partial}{\partial x^i}\right)_p (y^a \circ \phi) \equiv \left(\frac{\partial \hat{\phi}^a}{\partial x^i}\right)_p, \qquad \hat{\phi} := y \circ \phi.$$

Note that since $a = 1, \ldots, \dim N$, $i = 1, \ldots, \dim M$, these are not really tensor components.

Let us try to develop some intuition. If we have a curve γ on M, then we may also draw it on N as per $\phi \circ \gamma$. We also have the tangent vector $v_{\gamma,p}$ on M. The push-forward of $v_{\gamma,p}$ is then $\phi_*(v_{\gamma,p}) = v_{\phi\circ\gamma,\phi(p)}$ in N, as we would hopefully expect.

Theorem 10. $\phi_*(v_{\gamma,p}) = v_{\phi\circ\gamma,\phi(p)}$.

Proof. Observe

$$\begin{aligned}
\phi_*(v_{\gamma,p})f &= v_{\gamma,p}(f \circ \phi) \\
&= ((f \circ \phi) \circ \gamma)'(\lambda_0) \qquad \text{(reminder: } \gamma(\lambda_0) = p) \\
&= (f \circ (\phi \circ \gamma))'(\lambda_0) \qquad \text{(associativity of } \circ) \\
&= v_{\phi\circ\gamma,\phi(p)}f.
\end{aligned}$$

□

A particularly interesting case is when M has lower dimension than N. In this case ϕ may be thought of as an embedding. Then the push-forward map transforms a tangent vector to a vector embedded in a higher dimension, tangent to some surface $\phi(M)$. Once again a diagram would probably help. ¯_(ツ)_/¯

We may now consider the pull-back of $\phi : M \to N$.

Definition 63 (Pull-back map)**.** The pull-back is the map $\phi^* : T^*N \to T^*M$, where we define $\phi^*(\omega)$ by – for all vectors $X \in TM$ –

$$\phi^*(\omega) := (X \mapsto \omega(\phi_*(X))).$$

Our mnemonic here is: "covectors are pulled back".

The components of the pull-back wrt the same charts as above are

$$\phi^{*a}{}_i := \phi^*((dy^a)_{\phi(p)})\left(\left(\frac{\partial}{\partial x^i}\right)_p\right) = dy^a\left(\phi_*\left(\frac{\partial}{\partial x^i}\right)_p\right) = \phi_*{}^a{}_i = \left(\frac{\partial \hat{\phi}^a}{\partial x^i}\right)_p,$$

the exact same as the push-forward! As such we may write

$$\begin{aligned}
(\phi_*(X))^a &= \phi_*{}^a{}_i X^i, \\
(\phi^*(\omega))_i &= \phi^{*a}{}_i \omega_a.
\end{aligned}$$

We may similarly construct a picture for the pull-back. If we have a function $f \in C^\infty(N)$, then the pull-back of the gradient of f is the same as the gradient of the pull-back of f, i.e., ...

Theorem 11. $\phi^*(df) = d(\phi^* f)$.

Proof. We have

$$\begin{aligned}
\phi^*(df)X &= df(\phi_*(X)) && \text{(by definition)}\\
&= \phi_*(X)f && \text{(definition of } df\text{)}\\
&= X(f\circ\phi) && \text{(definition of push-forward)}\\
&= d(f\circ\phi)X\\
&= d(\phi^* f)X. && (\phi^* f := f\circ\phi)
\end{aligned}$$

□

If ϕ is invertible we'll similarly be able to push covectors forward and pull vectors back.

We can now talk about an important application of push-forward and pull-back: if ϕ is injective (so that $\dim M \leq \dim N$), and we have a metric on N, can we induce a metric on M?

Definition 64. If $\phi : M \to N$ is an injective map from a smooth manifold $(M, \mathcal{O}, \mathcal{A})$ to a manifold with metric $(N, \mathcal{O}, A_N, g)$, then we may define the induced metric g_M given by

$$g_M(X, Y) := g(\phi_*(X), \phi_*(Y)).$$

Component-wise we have, evaluated pointwise,

$$((g_M)_{ij})_p = (g_{ab})_{\phi(p)} \left(\frac{\partial\hat{\phi}^a}{\partial x^i}\right)_{\phi(p)} \left(\frac{\partial\hat{\phi}^b}{\partial x^j}\right)_{\phi(p)} = g_{ab}{\phi_*}^a{}_i{\phi_*}^b{}_j.$$

Remark 40 (My remark). In an analogous way we may define the pull-back of an arbitrary $(0, s)$-tensor field, and if we so wish, the push-forward of an arbitrary $(r, 0)$-tensor field.

Flow of a complete vector field

Let $(M, \mathcal{O}, \mathcal{A})$ be a smooth manifold and X be a vector field on M.

Definition 65. A curve $\gamma : I \subseteq \mathbb{R} \to M$ is called an **integral curve of** X if

$$v_{\gamma,\gamma(\lambda)} = X_{\gamma(\lambda)}.$$

Intuitively, this occurs when the tangent of γ agrees with the vector field where it is defined. For example: X could be the velocity of water molecules in a river, and γ could be the trajectory of a ship in said river.

Definition 66. A vector field X is **complete** if all integral curves have $I = \mathbb{R}$.

Example 14. We may consider a circular vector field in $\mathbb{R}^2$ which points clockwise everywhere. This is complete since we can follow each loop around and around infinitely. Completeness still holds at the origin where it is simply zero.

Theorem 12 (Not important; prompted by a question)**.** A compactly supported smooth vector field is complete.

Definition 67. The **flow of a complete vector field** X is a one-parameter family

$$\begin{aligned}
h^X : \mathbb{R} \times M &\to M\\
(\lambda, p) &\mapsto \gamma_p(\lambda),
\end{aligned}$$

where $\gamma_p : \mathbb{R} \to M$ is *the* integral curve of X with $\gamma(0) = p$.

It is tradition to write $\gamma_p(\lambda) = h^X_\lambda(p)$. For a fixed $\lambda \in \mathbb{R}$, $h^X_\lambda : M \to M$ is a smooth map.

Lie subalgebras

We now discuss Lie subalgebras of the Lie algebra $(\Gamma(TM), [\cdot,\cdot])$ of vector fields.

First recall $\Gamma(TM) = \{\text{all vector fields}\}$, and $[X, Y] \in \Gamma(TM)$, where $[X, Y]f := X(Yf) - Y(Xf)$. The bracket satisfies the properties

(i) $[X, Y] = -[Y, X]$,

(ii) $[\lambda X + Z, Y] = \lambda[X, Y] + [Z, Y]$, where $\lambda \in \mathbb{R}$, and

(iii) $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$.

We say that $(\Gamma(TM), [\cdot,\cdot])$ is a Lie algebra.

Now let $X_1, \dots, X_s$ be s-many vector fields on M where $[X_i, X_j] = C^k{}_{ij} X_k$ for some "structure constants" $C^k{}_{ij} \in \mathbb{R}$.

Definition 68. $(\mathrm{span}_{\mathbb{R}}\{X_1, \dots, X_s\}, [\cdot,\cdot])$ is called a Lie subalgebra. We may for convenience write $L := \mathrm{span}_{\mathbb{R}}\{X_1, \dots, X_s\}$.

Example 15. Take some vector fields X_1, X_2, X_3 on the *merely* smooth manifold (i.e., no metric!) $(S^2, \mathcal{O}, \mathcal{A})$ and suppose they satisfy $[X_1, X_2] = X_3, [X_2, X_3] = X_1, [X_3, X_1] = X_2$. Then $(\mathrm{span}_{\mathbb{R}}\{X_1, X_2, X_3\}, [\cdot,\cdot]) = \mathrm{SO}(3)$. Indeed:

$$\begin{aligned} X_1(p) &= -\sin\phi(p)\frac{\partial}{\partial\theta} - \cot\theta(p)\cos\phi(p)\frac{\partial}{\partial\phi}, \\ X_2(p) &= \cos\phi(p)\frac{\partial}{\partial\theta} - \cot\theta(p)\sin\phi\frac{\partial}{\partial\phi}, \\ X_3(p) &= \frac{\partial}{\partial\phi}. \end{aligned}$$

Symmetry

Definition 69. A finite-dimensional Lie subalgebra $(L, [\cdot,\cdot])$ is said to be a **symmetry** of a metric tensor field g if for every vector field $X \in L$, $\forall\lambda \in \mathbb{R}$ and for all vector fields A, B,

$$g\big((h^X_\lambda)_*(A), (h^X_\lambda)_*(B)\big) = g(A, B) \tag{8a}$$

$$\Longleftrightarrow$$

$$(h^X_\lambda)^* g = g. \tag{8b}$$

Lie derivative

We now introduce the Lie derivative, which makes checking for symmetries very easy.

It should be obvious from Eq. (8b) that for $\forall X \in L$, if

$$\mathcal{L}_X g := \lim_{\lambda\to 0} \frac{(h^X_\lambda)^* g - g}{\lambda} = 0,$$

then L is a symmetry of g. It can be shown that $\mathcal{L}$ is precisely the so-called Lie derivative of g wrt the vector field X (we will not demonstrate this).

Definition 70. The **Lie derivative** $\mathcal{L}$ on a smooth manifold $(M, \mathcal{O}, \mathcal{A})$ sends a pair of a vector field X and a (p,q)-tensor field T to a (p,q)-tensor field such that

(i) $\mathcal{L}_X f = Xf$,

(ii) $\mathcal{L}_X Y = [X, Y]$,

(iii) $\mathcal{L}_X(T + S) = \mathcal{L}_X T + \mathcal{L}_X S$,

(iv) $\mathcal{L}_X(T(\omega, Y)) = (\mathcal{L}_X T)(\omega, Y) + T(\mathcal{L}_X\omega, Y) + T(\omega, \mathcal{L}_X Y)$, and

(v) $\mathcal{L}_{X+Y} = \mathcal{L}_X T + \mathcal{L}_Y T$.

Remark 41. There are a few remarks we may make here.

1. Axioms (i), (iii), and (iv) match axioms of the covariant derivative. Axiom (v) is weaker as it does not feature the C^∞-linearity that the lower slot of the covariant derivative did. We see that we do not do this because of axiom (ii), since in the proof of the linearity of the torsion we saw that $[fX, Y] = f[X, Y] - (Yf)X \neq f[X, Y]$.
2. Unlike the covariant derivative, which required specification of the Γs to fix it, axiom (ii) fully specifies the Lie derivative. The price of this is seen in the next remark.
3. Recall that at a point p the covariant derivative $\nabla_X Y$ needs to know X at p and Y in a neighbourhood of p. The Lie derivative, however, needs to know X in the neighbourhood of p too. This is reflected in the lack of C^∞-linearity, where we see that the commutator bracket detects this in the additional $-(Yf)X$ term.

In a chart (U, x) we have:

$$(\mathcal{L}_X Y)^i = X(Y^i) - \frac{\partial X^i}{\partial x^s} Y^s \tag{8c}$$

$$(\mathcal{L}_X \omega)_i = X(\omega_i) + \frac{\partial X^s}{\partial x^i}\omega_s \tag{8d}$$

$$(\mathcal{L}_X T)^i{}_j = X(T^i{}_j) - \frac{\partial X^i}{\partial x^s} T^s{}_j + \frac{\partial X^s}{\partial x^j} T^i{}_s. \tag{8e}$$

Higher rank tensors follow the pattern accordingly.

Proof. Since it is an instructive exercise let's check the results. For the first case,

$$
\begin{aligned}
(\mathcal{L}_X Y)f &= \left[X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^j}\right] f && \text{(axiom (ii))} \\
&= X^i \frac{\partial}{\partial x^i}\left(Y^j \frac{\partial f}{\partial x^j}\right) - Y^j \frac{\partial}{\partial x^j}\left(X^i \frac{\partial f}{\partial x^i}\right) \\
&= X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial f}{\partial x^j} + \cancel{X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j}} - Y^j \frac{\partial X^i}{\partial x^j}\frac{\partial f}{\partial x^i} - \cancel{Y^j X^i \frac{\partial^2 f}{\partial x^j \partial x^i}} \\
&= \left(X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}\frac{\partial}{\partial x^i}\right) f,
\end{aligned}
$$

and the result follows. For a covector field ω,

$$
\begin{aligned}
X(\omega_i) = \mathcal{L}_X \omega_i = \mathcal{L}_X\left(\omega\left(\frac{\partial}{\partial x^i}\right)\right) &= (\mathcal{L}_X \omega)\left(\frac{\partial}{\partial x^i}\right) + \omega\left(\mathcal{L}_X \frac{\partial}{\partial x^i}\right) && \text{(axioms (i), (iv))} \\
&= (\mathcal{L}_X \omega)_i + \omega\left(-\frac{\partial X^s}{\partial x^i}\frac{\partial}{\partial x^s}\right) && \text{(using the first case)} \\
&= (\mathcal{L}_X \omega)_i - \frac{\partial X^s}{\partial x^i}\omega_s.
\end{aligned}
$$

Finally, for an e.g., (1,1)-tensor field, we see from the previous results that

$$
\begin{aligned}
X(T^i{}_j) = \mathcal{L}_X\left(T^i{}_j\right) = \mathcal{L}_X\left(T\left(dx^i, \frac{\partial}{\partial x^j}\right)\right) &= (\mathcal{L}_X T)^i{}_j + T\left(\mathcal{L}_X dx^i, \frac{\partial}{\partial x^j}\right) + T\left(dx^i, \mathcal{L}_X \frac{\partial}{\partial x^j}\right) && \text{(axiom (iv))} \\
&= (\mathcal{L}_X T)^i{}_j + T\left(\frac{\partial X^i}{\partial x^s} dx^s, \frac{\partial}{\partial x^j}\right) + T\left(dx^i, -\frac{\partial X^s}{\partial x^i}\frac{\partial}{\partial x^s}\right) \\
&= (\mathcal{L}_X T)^i{}_j + \frac{\partial X^i}{\partial x^s} T^s{}_j - \frac{\partial X^s}{\partial x^i} T^i{}_s,
\end{aligned}
$$

and we are done. □

We remind ourselves of the application for which we introduced the Lie derivative: if for all $X \in L$, $0 = (\mathcal{L}_X g)_{ij} = \ldots$, then L is a symmetry of g. This means that it is very easy to check whether a metric features a symmetry!

Lecture 12: Integration

This topic marks the completion of our "lift" of analysis on the charts to the manifold level.

The aim of this lecture is to be able to write $\int_M f$, and we will see this requires mild new structure – the choice of a so-called volume form, and a restriction of $\mathcal{A}$ (orientation).

Review of Integration in $\mathbb{R}^d$

We begin with a quick review of integration in $\mathbb{R}^d$ to set the stage.

(a) Let $F : \mathbb{R} \to \mathbb{R}$ be a function and assume a notion of integration is known. Then $\int_{(a,b)} F := \int_a^b \mathrm{d}x\, F(x)$.

(b) Let $F : \mathbb{R}^d \to R$ be a function. Then we can integrate

  (i) on a box-shaped domain $(a,b) \times (c,d) \times \cdots \times (u,v) \subseteq \mathbb{R}^d$ as

$$\int_{(a,b)\times\cdots\times(u,v)} \mathrm{d}^d x\, F(x) := \int_{(a,b)} \mathrm{d}x^1 \cdots \int_{(u,v)} \mathrm{d}x^d\, F(x^1, \ldots, x^d).$$

  (ii) on other domains $G \subseteq \mathbb{R}^d$ by defining the indicator function $\mu_G : \mathbb{R}^d \to \mathbb{R}$ by

$$x \mapsto \begin{cases} 1 & x \in G \\ 0 & \text{otherwise} \end{cases}.$$

  Then

$$\int_G \mathrm{d}^d x\, F(x) := \int_{-\infty}^{\infty} \mathrm{d}x^1 \cdots \int_{-\infty}^{\infty} \mathrm{d}x^d\, \mu_G(x) F(x).$$

Change of Variables

Theorem 13. When performing a change of variables $x = \phi(y)$ we had

$$\int_G \mathrm{d}^d x\, F(x) = \int_{\phi^{-1}(G)} \mathrm{d}^d y\, \left|\det\big(\partial_a \phi^b\big)(y)\right| \cdot (f \circ \phi)(y). \tag{9}$$

Integration on One Chart

Let $(M, \mathcal{O}, \mathcal{A})$ be a smooth manifold and $f : M \to \mathbb{R}$. Choose the charts $(U,x), (V,y) \in \mathcal{A}$. In each chart we write $f_{(x)} := f \circ x^{-1}$ and $f_{(y)} := f \circ y^{-1}$.

We first try to implement the integral naively in one chart and see why it won't work. Here ϕ will be our chart transition map as we change charts/variables as per Eq. (9).

$$\begin{aligned}
\int_{y(U)} \mathrm{d}^d \beta\, f_{(y)}(\beta) &= \int_{x(U)} \mathrm{d}^d \alpha\, \left|\det\big(\partial_a (y^b \circ x^{-1})(\alpha)\big)\right| (f_{(y)} \circ (y \circ x^{-1}))(\alpha) \\
&= \int_{x(U)} \mathrm{d}^d \alpha \left| \det \left[\frac{\partial y^b}{\partial x^a}\right]_{x^{-1}(\alpha)} \right| \underbrace{(f \circ \cancel{y^{-1} \circ y} \circ x^{-1})}_{f_{(x)}}(\alpha) \\
&\neq \int_{x(U)} \mathrm{d}^d \alpha\, f_{(x)}(\alpha)
\end{aligned}$$

We observe that our attempt to define $\int_U f$ as $\int_{x(U)} \mathrm{d}^d\alpha\, f_{(x)}(\alpha)$ fails as it is ill-defined. However, we we may be able to fix the problem if we introduce a term which transforms as to cancel out the Jacobian factor. Unfortunately no such object exists (idk why), and so this is where we need to introduce the additional structure.

Volume Forms

We don't go down the road of determining this extra structure, we will just introduce it.

Definition 71. On a smooth manifold $(M, \mathcal{O}, \mathcal{A})$ a $(0,d)$-tensor field Ω is called a **volume form** if

(a) Ω vanishes nowhere

(b) it is totally antisymmetric, i.e.:

$$\Omega(\ldots, \underbrace{X}_{i^{th}}, \ldots, \underbrace{Y}_{j^{th}}, \ldots) = -\Omega(\ldots, \underbrace{Y}_{i^{th}}, \ldots, \underbrace{X}_{j^{th}}, \ldots).$$

In a chart, $\Omega_{i_1 \cdots i_j} = \Omega_{[i_1 \cdots i_j]}$.

Suppose we are given a Riemannian metric manifold $(M, \mathcal{O}, \mathcal{A}, g)$. We can construct a volume form Ω from the metric. In *any* chart (U, x):

$$\Omega_{(x)i_1 \cdots i_d} := \sqrt{\det\big((g_{(x)})_{ij}\big)}\epsilon_{i_1 \cdots i_d},$$

where $\epsilon_{i_1 \cdots i_d}$ is the familiar Levi-Civita symbol. We claim this is well-defined and transforms like a tensor, with only a slight caveat.

Proof. We have

$$\begin{aligned}
\Omega_{(y)_{i_1 \cdots i_d}} &= \sqrt{\det\big(g_{(y)ij}\big)}\epsilon_{i_1 \cdots i_d} \\
&= \sqrt{\det\left(g_{(x)mn}\frac{\partial x^m}{\partial y^i}\frac{\partial x^n}{\partial y^j}\right)}\frac{\partial y^{m_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{m_d}}{\partial x^{i_d}}\epsilon_{m_1 \cdots m_d} \\
&= \sqrt{\det\big(g_{(x)ij}\big)}\underbrace{\sqrt{\det^2\left[\frac{\partial x}{\partial y}\right]}}_{\left|\det\left[\frac{\partial x}{\partial y}\right]\right|}\det\left[\frac{\partial y}{\partial x}\right]\epsilon_{i_1 \cdots i_d} \\
&= \sqrt{\det\big(g_{(x)ij}\big)}\epsilon_{i_1 \cdots i_d}\operatorname{sgn}\left(\det\left[\frac{\partial x}{\partial y}\right]\right).
\end{aligned}$$

We see that there is a troublesome term! Ω can only be well-defined if $\det\left[\frac{\partial y}{\partial x}\right] > 0$ for any pair of charts (U, x) and (V, y). □

Alright then, let's require that we restrict the smooth atlas $\mathcal{A}$ to a subatlas $\mathcal{A}^{\uparrow} \subseteq \mathcal{A}$ such that any two charts (U, x), (V, y) have chart transition maps $y \circ x^{-1}$, $x \circ y^{-1}$ satisfying $\det\left[\frac{\partial y}{\partial x}\right] > 0$. Such an $\mathcal{A}^{\uparrow}$ is called an **oriented** atlas.

Remark 42. Strictly speaking, the Levi-Civita symbol is a "tensor density" (which is why it doesn't transform as we'd expect of a tensor), and to properly define it we apparently need to define the "tensor density bundle of the frame bundle", which is too much work for so little payoff that we don't bother.

Remark 43. It might seem a bit silly to write $\det(g_{ij})$ since g is not a matrix but a bilinear form. However we can do it nonetheless by defining it in a natural way. After all, in a chart we can do whatever we want with the numbers, as long as we find a way to ensure the resultant quantity is chart-independent.

Let Ω be a volume form on $(M, \mathcal{O}, A^{\uparrow})$ and consider a chart (U, x).

Definition 72. We define in the chart

$$\omega_{(x)} := \Omega_{i_1 \cdots i_d}\epsilon^{i_1 \cdots i_d},$$

where $\epsilon^{i_1 \cdots i_d}$ behaves identically to $\epsilon_{i_1 \cdots i_d}$. One can show that

$$\omega_{(y)} = \det\left[\frac{\partial x}{\partial y}\right]\omega_{(x)},$$

and we call such an object a scalar density.

Integration on one Chart Domain U

Definition 73.

$$\int_U f := \int_{x(U)} \mathrm{d}^d\alpha\, \omega_{(x)}(x^{-1}(\alpha))f_{(x)}(\alpha)$$

Proof. Proof of well-definedness.

$$\begin{aligned}
\int_U f &= \int_{y(U)} \mathrm{d}^d\beta\, \omega_{(y)}(y^{-1}(\beta))f_{(y)}(\beta) \\
&= \int_{x(U)} \mathrm{d}^d\alpha\, \omega_{(x)}(x^{-1}(\alpha))\det\left[\frac{\partial x}{\partial y}\right]\left|\det\left[\frac{\partial y}{\partial x}\right]\right|f_{(x)}(\beta) \\
&= \int_{x(U)} \mathrm{d}^d\alpha\, \omega_{(x)}(x^{-1}(\alpha))f_{(x)}(\alpha) \qquad \text{(using oriented atlas)}
\end{aligned}$$

□

On an oriented metric manifold $(M, \mathcal{O}, \mathcal{A}^{\uparrow}, g)$ this becomes

$$\int_U f := \int_{x(U)} \mathrm{d}^d\alpha \, \sqrt{\det(g_{(x)ij})(x^{-1}(\alpha))} f_{(x)}(\alpha),$$

which, with increasing practice, we will eventually write as

$$\int_{x(U)} \mathrm{d}^d\alpha \, \sqrt{g} f_{(x)}(\alpha).$$

Integration on the Entire Manifold

In general, integrating over the entire manifold entails integrating over multiple overlapping chart domains. We can't just cut the intersection out since it isn't an open set.

Idea: require that the manifold admit a so-called *partition of unity*.

Roughly: for any *finite* subatlas $\mathcal{A}' = \{(U_1, x_1), \ldots, (U_n, x_n)\} \subseteq \mathcal{A}^{\uparrow}\}$, there exist continuous functions $\rho_i : U_i \to \mathbb{R}$ such that $\forall p \in M$,

$$\sum_{U_i \ni p} \rho_i(p) = 1.$$

Remark 44. A weaker but sufficient case is simply to ensure for each $p \in M$ that we have only finitely many charts containing p, as opposed to having finitely many charts altogether. In this way we still avoid having to worry about convergence in the above equation.

Definition 74 (Integration over M).

$$\int_M f := \sum_{i=1}^{\text{finite}} \int_{U_i} (\rho_i \cdot f)$$

Lecture 13 - Relativistic Spacetime

Recall from lecture 9 the definition of Newtonian spacetime $(M, \mathcal{O}, \mathcal{A}, \nabla, t)$. Here, relativistic spacetime will be $(M, \mathcal{O}, A, \nabla, g, T)$, where g is a Lorentzian metric, and T is a time orientation.

Time orientation

Definition 75. Let $(M, \mathcal{O}, \mathcal{A}^{\uparrow}, g)$ be a Lorentzian manifold. Then a time-orientation is given by a vector field T that

(a) does not vanish anywhere

(b) satisfies $g(T, T) > 0$.

Remark 45. Recall that the Lorentzian metric defines two cones satisfying $g(X, X) > 0$. The purpose of T is to select one of the cones as going 'forward in time'.

This definition of (relativistic) spacetime has been made to enable the following physical postulates.

(P1) The worldline γ of a *massive* particle satisfies

(i) $g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) > 0$

(ii) $g_{\gamma(\lambda)}(T, v_{\gamma,\gamma(\lambda)}) > 0$.

(P2) Worldlines of *massless* particles satisfy

(i) $g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) = 0$

(ii) $g_{\gamma(\lambda)}(T, v_{\gamma,\gamma(\lambda)}) > 0$.

From here on we will always deal with a spacetime $(M, \mathcal{O}, \mathcal{A}^{\uparrow}, \nabla, g, T)$.

Observers

Definition 76. An **observer** is a worldline γ with $g(v_\gamma, v_\gamma) > 0$ and $g(T, v_\gamma) > 0$ together with a choice of basis

$$e_0(\lambda), e_1(\lambda), e_2(\lambda), e_3(\lambda)$$

of each $T_{\gamma(\lambda)}M$ where the observer worldline passes. The chosen basis must satisfy $e_0(\lambda) \equiv v_{\gamma,\gamma(\lambda)}$ and

$$g(e_a(\lambda), e_b(\lambda)) = \eta_{ab} := \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}_{ab}.$$

(Aside: To be precise, an observer is a smooth curve in the frame bundle LM over M. A frame bundle is one in which the fibres are no longer all tangent vectors, but the quadruplets of vectors which constitute a basis.)

We now have two more physical postulates.

(P3) A *clock* carried by a specific observer (γ, e) will measure a *time*

$$\tau := \int_{\lambda_0}^{\lambda_1} \mathrm{d}\lambda \sqrt{g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)})}$$

between the two "*events*" $\gamma(\lambda_0)$ ("start the clock") and $\gamma(\lambda_1)$ ("stop the clock"). We may call τ the "proper time" or "eigentime".

Remark 46. P3 is the first time we make reference to a notion of 'time'.

Example 16. Take $M = \mathbb{R}^4$, $\mathcal{O} = \mathcal{O}_{\text{st}}$, $\mathcal{A}^{\uparrow} \ni (\mathbb{R}^4, \mathbb{I}_{\mathbb{R}^4})$, g such that $g_{(x)ij} = \eta_{ij}$ (which $\implies \Gamma_{(x)}{}^{i}{}_{jk} = 0$ everywhere $\implies$ Riem $= 0 \implies$ spacetime is flat); $T^i_{(x)} = (1, 0, 0, 0)^i$. This situation is called special relativity.

Consider two observers $\gamma : (0, 1) \to M$ defined by

$$\gamma^i_{(x)} = (\lambda, 0, 0, 0)^i$$

and $\delta : (0, 1) \to M$ defined by

$$\delta^i_{(x)} = \begin{cases} (\lambda, \lambda\alpha, 0, 0)^i & \lambda \leq \frac{1}{2} \\ (\lambda, (1-\lambda)\alpha, 0, 0)^i & \lambda > \frac{1}{2} \end{cases}.$$

We see that this is the setup of the twin paradox. Let's calculate:

$$\tau_\gamma = \int_0^1 \mathrm{d}\lambda \sqrt{g_{(x)ij}\dot{\gamma}^i_{(x)}\dot{\gamma}^j_{(x)}} = \int_0^1 \mathrm{d}\lambda = 1,$$

and

$$\tau_\delta = \int_0^{1/2} \mathrm{d}\lambda \sqrt{1-\alpha^2} + \int_{1/2}^1 \mathrm{d}\lambda \sqrt{1-(-\alpha)^2} = \int_0^1 \mathrm{d}\lambda \sqrt{1-\alpha^2} = \sqrt{1-\alpha^2}.$$

We can't claim that if we switch frames the paradox will emerge, since we did this entirely on the global chart level, *not* in the frame of either observer.

Note: Taking the clock postulate (P3) seriously, one better come up with a realistic clock design that supports the postulate. One such idea is to have two mirrors with a photon bouncing between them, which we can draw in a simple spacetime diagram to illustrate the effect.

Now for our next postulate.

(P4) Let (γ, e) be an observer, and δ be a *massive* particle worldline that is parametised such that $g(v_\delta, v_\delta) = 1$. Suppose the observer and the particle meet somewhere in spacetime $p = \gamma(\tau_1) = \delta(\tau_2)$. This observer measures the 3-velocity (spatial velocity) of this particle as

$$v := \epsilon^\alpha(v_{\delta,\delta(\tau_2)})e_\alpha \qquad (\alpha = 1, 2, 3),$$

(where $\epsilon^0, \ldots, \epsilon^3$ is the unique dual basis of $e_0, \ldots, e_1$).

A consequence is the following. An observer (γ, e) will always in this manner extract quantities measurable in his/her laboratory for objective spacetime quantities.

Example 17. Take the F, the Faraday (0,2)-tensor of electromagnetism

$$F(e_a, e_b) = F_{ab} = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{bmatrix}_{ab}.$$

Then $E_\alpha = F(e_0, e_\alpha)$, $B^\gamma = F(e_\alpha, e_\beta)\epsilon^{\alpha\beta\gamma}$. This is most definitely observer-dependent.

Role of the Lorentz Transformations

Lorentz transformations emerge as follows: Let (γ, e) and $(\tilde{\gamma}, \tilde{e})$ be observers with $\gamma(0) = \tilde{\gamma}(0)$. Now $e_0, \ldots, e_3$ at $\tau_1 = 0$ and $\tilde{e}_0, \ldots, \tilde{e}_3$ at $\tau_2 = 0$ are both bases for the same $T_{\gamma(0)}M$. Thus

$$\tilde{e}_a = \Lambda^b{}_a e_b, \qquad \Lambda \in \mathrm{GL}(4).$$

Now

$$\eta_{ab} = g(\tilde{e}_a, \tilde{e}_b) = g(\Lambda^m{}_a e_m, \Lambda^n{}_b e_n) = \Lambda^m{}_a \Lambda^n{}_b \underbrace{g(e_m, e_n)}_{\eta_{mn}},$$

i.e., $\Lambda \in \mathrm{O}(1,3)$ (the Lorentz group). (TODO: why?)

Result: Lorentz transformations relate the *frames* of any two observers at the same point. (Note it is nonsense to talk about taking Lorentz transformations at different points in spacetime.)

Lecture 14 - Matter

Matter is a broad topic for which we can't hope to do justice in only one lecture; however we will nonetheless press on and hope to cover the salient points.

There are two types of (classical) matter: point matter (particles) and field matter (electromagnetic fields). The former has a phenomonological importantance (e.g., massive particles), while the latter is more fundamental from a GR point of view.

Point Matter

Our postulates (P1) and (P2) already constrain the possible particle worldlines. But what is their precise law of motion, possibly in the presence of "forces"?

Without external forces

We have the action $S_{\text{massive}}[\gamma] := m \int \mathrm{d}\lambda \sqrt{g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)})}$[1], and we also require that $g_{\gamma(\lambda)}(T_{\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) > 0$. Our dynamical law is the Euler-Lagrange Equation.

Similarly $S_{\text{massless}}[\gamma, \mu] = \int \mathrm{d}\lambda \, \mu_\lambda g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)})$. The Euler-Lagrange equations of the additional Lagrange multiplier μ enforces $g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)}) = 0$, while the ELE for γ yields the equations of motion.

Our reason for describing equations of motion by actions is that composite systems have an action that is the sum of the actions of the parts of that system, possibly including "interaction terms". (For example, $S = S[\gamma] + S[\delta] + S_{\text{int}}[\gamma, \delta]$.)

In the presence of external forces

Or rather: the presence of *fields* to which a particle "couples".

Example 18. $S[\gamma; A] = \int \mathrm{d}\lambda \left(m\sqrt{g_{\gamma(\lambda)}(v_{\gamma,\gamma(\lambda)}, v_{\gamma,\gamma(\lambda)})} + q \cdot A(v_{\gamma,\gamma(\lambda)}) \right)$, where A is a *covector field* on M (e.g., the electromagnetic potential).

Consider the Euler-Lagrange equations with $\mathscr{L}_{\text{int}} = qA_{(x)m}\dot{\gamma}^m_{(x)}$:

$$\left(\frac{\partial \mathscr{L}_{\text{int}}}{\partial \dot{\gamma}^a}\right)^{\cdot} = \left(q \cdot A_{(x)a}\right)^{\cdot} = q \cdot \frac{\partial}{\partial x^m}\left(A_{(x)a}\right)\dot{\gamma}^m_{(x)},$$
$$\frac{\partial \mathscr{L}_{\text{int}}}{\partial \gamma^a} = q \cdot \frac{\partial}{\partial x^a}\left(A_{(x)m}\right)\dot{\gamma}^m.$$

The equation of motion is then

$$\begin{aligned}
m(\nabla_{v_\gamma} v_\gamma)_a + \left(\frac{\partial \mathscr{L}_{\text{int}}}{\partial \dot{\gamma}^a_{(x)}}\right)^{\cdot} - \frac{\partial \mathscr{L}_{\text{int}}}{\partial \gamma^a_{(x)}} &= 0 \\
m(\nabla_{v_\gamma} v_\gamma)_a + \left(q \cdot \frac{\partial}{\partial x^m}\left(A_{(x)a}\right)\dot{\gamma}^m_{(x)} - q \cdot \frac{\partial}{\partial x^a}\left(A_{(x)m}\right)\dot{\gamma}^m_{(x)}\right) &= 0 \\
m(\nabla_{v_\gamma} v_\gamma)_a + q\left(\frac{\partial A_a}{\partial x^m} - \frac{\partial A_m}{\partial x^a}\right)\dot{\gamma}^m_{(x)} &= 0 \\
m(\nabla_{v_\gamma} v_\gamma)_a + q \cdot F_{(x)am}\dot{\gamma}^m_{(x)} &= 0 \\
\implies m(\nabla_{v_\gamma} v_\gamma)^a &= -qF^a{}_m\dot{\gamma}^m,
\end{aligned}$$

where F is the Faraday tensor, and we have raised the index in the last line. $-qF^a{}_m\dot{\gamma}^m$ is the Lorentz force on a charged particle in an electromagnetic field.

Field Matter

It is very difficult to give an abstract definition that does not understate or overstate what field matter is.

Definition 77. Classical field matter is any tensor field on spacetime whose equations of motion derive from an action.

Example 19.

$$S_{\text{Maxwell}}[A; g] = \frac{1}{4}\int_M \mathrm{d}^4x \underbrace{\sqrt{-g}F_{ab}F_{cd}g^{ac}g^{bd}}_{\text{Lagrangian density } \mathcal{L}},$$

where A is a (0,1)-tensor field, g is fixed for the time being, and we assume for simplicity that one chart covers all of M. We remind outselves that $F_{ab} := 2\partial_{[a}A_{b]} = 2(\nabla_{[a}A)_{b]}$.

[1] For now the m factor is unimportant, but when we add in additional interaction terms it will become relatively important.

Euler-Lagrange equations for fields

$$0 = \frac{\partial \mathcal{L}}{\partial A_m} - \frac{\partial}{\partial x^s}\left(\frac{\partial \mathcal{L}}{\partial \partial_s A_m}\right)$$

If the Lagrangian density depends on higher derivatives of A, the equation picks up more terms with alternating signs:

$$0 = \frac{\partial \mathcal{L}}{\partial A_m} - \frac{\partial}{\partial x^s}\left(\frac{\partial \mathcal{L}}{\partial \partial_s A_m}\right) + \frac{\partial}{\partial x^s}\frac{\partial}{\partial x^t}\left(\frac{\partial \mathcal{L}}{\partial \partial_t \partial_s A_m}\right) - \cdots.$$

Example 20. For S_{Maxwell}, when we plug in (which we won't) we find

$$\cdots = (\nabla_{\frac{\partial}{\partial x^m}} F)^{ma} \equiv (\nabla_m F)^{ma} = 0,$$

and we call these the inhomogeneous Maxwell equations. If we coupled to a current a la $\mathcal{L} \to \mathcal{L} + A(j)$, it would become

$$(\nabla_m F)^{ma} = j^a.$$

(For example, if the current came from a point particle, then we'd have $j = qv_\gamma$.) (The homogeneous Maxwell equations are $\partial_{[a} F_{bc]} = (\nabla_{[a} F)_{bc]} = 0$.)

Example 21. Another example well-liked by textbooks is

$$S_{\text{Klein-Gordon}}[\phi] = \int_M \mathrm{d}^4 x\, \sqrt{-g}\left[g^{ab}(\partial_a \phi)(\partial_b \phi) - m^2 \phi^2\right],$$

where ϕ is a (0,0)-tensor field.

Energy-Momentum tensor of matter fields

At some point we will want to write down an action for the metric tensor field itself. This action $S_{\text{grav}}[g]$ will then be added to any $S_{\text{matter}}[A, \phi, \ldots]$ in order to describe the total system.

Example 22.

$$S_{\text{total}}[g, A] = S_{\text{grav}}[g] + S_{\text{matter}}[A, g]$$

Then we will see that

$$\begin{aligned} \delta A &\implies \text{Maxwells eqns} \\ \delta g &\implies \frac{1}{16\pi G_N} G^{ab} + (-\frac{1}{2} T^{ab}) = 0. \end{aligned}$$

As a general feature S_{grav} will always be there so we will get the $\frac{1}{16\pi G_N} G^{ab}$ term; however the latter term depends on what we use for S_{matter}. However it proves convenient to always associate the object T^{ab} with whatever comes out of S_{matter}, so that the resultant equation will always be of the form

$$G^{ab} = 8\pi G_N T^{ab}.$$

Definition 78. If $S_{\text{matter}}[\Phi, g]$ is a matter action for any matter field Φ, the so-called **energy-momentum tensor** is

$$T^{ab} := -\frac{2}{\sqrt{-g}}\left(\frac{\partial \mathcal{L}_{\text{matter}}}{\partial g_{ab}} - \partial_s \frac{\partial \mathcal{L}_{\text{matter}}}{\partial \partial_s g_{ab}} + \cdots\right).$$

Remark 47. Sign conventions here are notoriously dangerous. We choose all sign conventions such that $T(\epsilon^0, \epsilon^0) > 0$.

Example 23. For S_{Maxwell} (hoping we're consistent with our prescribed convention):

$$T_{ab} = F_{am} F_{bn} g^{mn} - \frac{1}{4} F_{mn} F^{mn} g_{ab}.$$

Then $T(e_0, e_0) = E^2 + B^2$, the energy density, and $T(e_0, e_\alpha) = (\vec{\mathbf{E}} \times \vec{\mathbf{B}})_\alpha$, the momentum density.

Fact: One often does not specify the fundamental action for some matter, but is one is rather satisfied to assume certain properties/forms of T_{ab}.

Example 24. The prime example we will look at is in cosmology. In a homogeneous and isotropic universe, one considers a perfect fluid of pressure p and density ρ which is modelled by

$$T^{ab} = (\rho + p) U^a U^b - p g^{ab},$$

where U is the flow vector field. One does not care of the action from which this comes.