

Cool stuff in General Relativity

Afiq Hatta

November 19, 2019

Contents

1	Starting off with non-relativistic particles	2
1.1	Exploring different Lagrangians	3
2	Special relativity	4
3	Introducing Differential Geometry for General Relativity	6
3.1	Tangent vectors	7
3.2	Treating tangent vectors as derivatives of curves on the manifold	9
3.3	Vector fields	9
3.4	Integral curves	10
3.5	Differentiating vector fields with respect to other vector fields	12
3.5.1	Push-forwards and Pull-backs	12
3.5.2	Components for Push-forwards and Pull-backs	12
3.5.3	Introducing the Lie derivative	13
4	Tensors	14
4.1	A note about dual spaces	14
4.1.1	Vector and Covector spaces	15
4.2	Taking the Lie Derivative of Covectors	16
4.3	Tensor fields	18
4.4	Differential forms	20
4.4.1	Wedge products	20
4.4.2	Properties of wedge products	20
4.5	The exterior derivative	21
4.5.1	Examples	23
4.6	Integration	24
4.6.1	Integrating over submanifolds	24
4.6.2	Stokes' theorem	25
5	Introducing Riemannian Geometry	27
5.1	The metric	27
5.1.1	Riemannian Manifolds	27
5.1.2	Riemannian Geometry	28
5.1.3	The Joys of a metric	28
5.2	Connections and Curvature	31

5.3	The Levi-Civita Connection	33
5.3.1	The Divergence Theorem	34
5.4	Parallel Transport	35
5.4.1	Geodesics	35
5.5	Normal coordinates	35
5.5.1	Constructing Normal Coordinates	37
5.6	Path Deviations and Curvature	37
5.7	Derived tensors from the Riemann Tensor	39
5.8	Connection 1-forms	39
6	The Einstein Equations	42
6.0.1	The Einstein-Hilbert Action	42
6.0.2	Dimensional analysis	44
6.0.3	The Cosmological Constant	44
6.1	Diffeomorphisms Revisited	44
6.2	Some simple solutions	45
6.2.1	De Sitter Spacetime	45
6.2.2	Anti de-Sitter Space-time	48
6.3	Symmetries	49
6.3.1	More examples	50
6.4	Conserved quantities	51
7	Example Sheet 1	52
7.1	Question 1	53
7.2	Question 2	54
7.3	Question 3	55
7.4	Question 5	56
7.5	Question 6	58
7.6	Question 7	59
7.7	Question 8	61
7.8	Question 9	63
7.9	Question 10	64
7.10	Question 11	65
7.11	Being careful about how we expand things	66
7.12	Tricks in the Levi-Civita connection	66
7.13	Integrating over volume forms	66
7.14	Showing that magnitude is constant	67
7.15	Symmetries of the Riemann Tensor	67
7.16	General techniques	67

1 Starting off with non-relativistic particles

Deriving the Euler Lagrange equations. We perturb our action

$$S = \int_{t_1}^{t_2} dt L$$

If we perturb our action slightly,

$$\begin{aligned}
S[x^i + \delta x^i] &= \int_{t_1}^{t_2} dt L(x^i + \delta x^i, \dot{x}^i + \delta \dot{x}^i) \\
&= S[x^i] + \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial x^i} \delta x^i + \frac{\partial L}{\partial \dot{x}^i} \delta \dot{x}^i \right) \\
&= S[x^i] + \int_{t_1}^{t_2} dt \delta x^i \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) \right)
\end{aligned}$$

This implies that the integrand goes to zero.

1.1 Exploring different Lagrangians

Consider the Lagrangian in Euclidean coordinates

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

The E-L equations imply that $\ddot{x} = 0$. So we have a constant velocity. Now in different coordinates,

$$L = \frac{1}{2}g_{ij}(x)\dot{x}^i\dot{x}^j$$

This is a metric. Our distance in these general coordinates between $x^i \rightarrow x^i + \delta x^i$ is now

$$ds^2 = g_{ij}dx^i dx^j$$

Some g_{ij} do not come from \mathbb{R}^3 , and these spaces are **curved**. This means there is no smooth map back into \mathbb{R}^3 . Our equations of motion that comes from the Euler Lagrange equations are the geodesic equations.

Observe that

$$\begin{aligned}
\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) &= \frac{d}{dt} g_{ik} \dot{x}^k \\
&= g_{ik} \ddot{x}^k + \dot{x}^k \dot{x}^l \partial_l g_{ik}
\end{aligned}$$

And, differentiating the lagrangian with ∂_i , we get

$$\frac{\partial L}{\partial x^i} = \frac{1}{2} \partial_i g_{kl} \dot{x}^k \dot{x}^l$$

Substituting this into the EL equations, this reads

$$g_{ik} \ddot{x}^k + \dot{x}^k \dot{x}^l \partial_l g_{ik} - \frac{1}{2} \partial_i g_{kl} \dot{x}^k \dot{x}^l = 0$$

Now, the second term is symmetric in k, l , so we can split this term in two. We also multiply by the inverse metric to cancel out this annoying factor of g that we have in front of everything. This gives us the final expression that

$$\ddot{x}^i + \frac{1}{2} g^{il} (\partial_k g_{lj} + \partial_j g_{lk} - \partial_l g_{jk}) \dot{x}^j \dot{x}^k = 0$$

2 Special relativity

We'll now put time and space on the same 'footing' per se, and talk about special relativity. In special relativity, instead of time being its own separate variable, we have that dynamic events take place in 4 spacetime coordinates denoted $x^\mu = (t, x, y, z)$, where we now use greek indices to denote four components $\mu = 0, 1, 2, 3$. Now, we wish to construct an action and extremize this path, but since our t variable is already taken, we need to parametrise paths in spacetime by a different parameter. We'll call this parameter σ , and show that there's a natural choice for this, something called 'proper time', later.

We define our metric, the Minkowski metric, on this spacetime to be $\eta^{\mu\nu} = \text{diag}(-1, +1, +1, +1)$. Thus, distances in Minkowski spacetime are denoted as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2$$

We have names for different events based on their infinitesimal distance. Since our metric is no longer positive definite, we have that events can have a distance of any sign.

- If $ds^2 < 0$, events are called timelike.
- If $ds^2 = 0$, events are called null.
- If $ds^2 > 0$, events are called spacelike.

Our action, then, should look like (now with the use of an alternate parameter σ to parametrize our paths)

$$S[x^\mu(\sigma)] = \int_{\sigma_1}^{\sigma_2} \sqrt{-ds^2}$$

Now, we can parametrise the integrand with sigma to get

$$S[x^\mu(\sigma)] = m \int_{\sigma_1}^{\sigma_2} d\sigma \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma}}$$

In this case, our Lagrangian $L = m \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma}}$. Now, before we begin analysing what this equation gives us, there are two symmetries we'd like to take note of. One of our symmetries is invariance under Lorentz transformations. This means, if we boost our frame with a Lorentz transformation $x^\mu \rightarrow \Lambda^\mu{}_\nu x^\nu$, one can easily verify, using the condition that

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu}$$

that the Lagrangian remains invariant under this. One can also verify that this action is invariant under reparametrisation of the curve via a new function $\sigma' = \sigma'(\sigma)$.

Using the chain rule, we reparametrise by rewriting the action as

$$S = m \int \frac{d\sigma}{d\sigma'} d\sigma' \sqrt{-\eta_{\mu\nu} \left(\frac{d\sigma'}{d\sigma} \right)^2 \frac{dx^\mu}{d\sigma'} \frac{dx^\nu}{d\sigma'}} = \int d\sigma' \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\sigma'} \frac{dx^\nu}{d\sigma'}}$$

In this case, we're just applying the chain rule but factoring out the $\frac{d\sigma'}{d\sigma}$ term. But this is exactly the same as what we had before. Thus, we have reparametrisation invariance. In analogy

with classical mechanics, we compute the conjugate momentum

$$p_\mu = \frac{\partial L}{\partial \dot{x}^\mu}$$

3 Introducing Differential Geometry for General Relativity

Our main mathematical objects of interest in general relativity are manifolds. Manifolds are topological spaces which, at every point, has a neighbourhood which is homeomorphic to a subset of \mathbb{R}^n , where we call n the dimension of the manifold. In plain English, manifolds are spaces in which, locally at a point, look like a flat plane. This can be made more rigorous by the creation of maps, which we call 'charts', that take an open set around a point (a neighbourhood), and mapping this to a subset of \mathbb{R}^n .

Precisely, for each $p \in \mathcal{M}$, there exists a map $\phi : \mathcal{O} \rightarrow \mathcal{U} \subset \mathbb{R}^n$, where $p \in \mathcal{O} \subset \mathcal{M}$, and \mathcal{O} is an open set of M defined by the topology. Think of ϕ as a set of local coordinates, assigning a coordinate system to p . We will write $\phi(p) = (x_1, \dots, x_n)$ in this regard.

This map must be a 'homeomorphism', which is a continuous, invertible map with a continuous inverse. In this sense, our idea of assigning local coordinates to a point in \mathcal{M} becomes even more clear.

We can define different charts to different regions, but we need to ensure that they're well behaved on their intersections. Suppose we had two charts and two open sets defined on our manifold, and looked at how we transfer from one chart to another. For charts to be compatible, we require that the map

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(\mathcal{O}_\alpha \cap \mathcal{O}_\beta) \rightarrow \phi_\alpha(\mathcal{O}_\alpha \cap \mathcal{O}_\beta)$$

is also smooth (infinitely differentiable).

A collection of these maps (charts) which cover the manifold is called an atlas, and the maps taking one coordinate system to another ($\phi_\alpha \circ \phi_\beta^{-1}$), are called transition functions.

Some examples of manifolds include

- \mathbb{R}^n and all subsets of \mathbb{R}^n are n -dimensional manifolds, where the identity map serves as a sufficient chart.
- S^1, S^2 are manifolds, with modified versions of polar coordinates patched together forming a chart (as we'll see in the case of S^1).

Let's start simple and try to construct a chart for S^1 . Our normal intuition would be to use a single chart $S^1 \rightarrow [0, 2\pi)$, which indeed covers S^1 but doesn't satisfy the condition that the target set is an open subset of \mathbb{R} . This yields problems in terms of differentiation functions at the point $0 \in \mathbb{R}$, because the interval is closed there, not open. One way to remedy this is to define two coordinate charts then patch them together to form an atlas. Our first open set will be the set of points on the circle which exclude the rightmost point on the diameter, a set denoted by \mathcal{O}_1 , and our second open set is the whole sphere excluding the leftmost point. We'll denote this \mathcal{O}_2 .

We assign the following charts which are inline with this geometry

$$\begin{aligned}\phi_1 : \mathcal{O}_1 &\rightarrow \theta_1 \in (0, 2\pi) \\ \phi_2 : \mathcal{O}_2 &\rightarrow \theta_2 \in (-\pi, \pi)\end{aligned}$$

It's easy to verify that if we take a point on the manifold, our transition matrix reads that

$$\theta_2 = \phi_2(\phi_1^{-1}(\theta_1)) = \begin{cases} \theta_1, & \theta_1 \in (0, \pi) \\ \theta_1 - 2\pi, & \theta_1 \in (\pi, 2\pi) \end{cases}$$

Now that we have coordinate charts, we can do things that we usually do on functions described in \mathbb{R}^n , like differentiate. Furthermore, we can define maps between manifolds (which don't necessarily have the same dimension), where smoothness is defined via smoothness on coordinate charts. These are called diffeomorphisms. A function

$$f : \mathcal{M} \rightarrow \mathcal{N}$$

is a diffeomorphism if the corresponding map between $\mathbb{R}^{\dim \mathcal{M}}$ and $\mathbb{R}^{\dim \mathcal{N}}$ is smooth:

$$\psi \circ f \circ \phi^{-1} : U_1 \rightarrow U_2$$

for all coordinate charts $\phi : \mathcal{O}_1 \rightarrow U_1$ and $\psi : \mathcal{O}_2 \rightarrow U_2$ defined on the manifolds \mathcal{M} and \mathcal{N} respectively.

3.1 Tangent vectors

Throughout our whole lives, we've been thinking of a 'vector' as a way to denote some position in space. However, this idea of a vector is only really unique to the manifold \mathbb{R}^n . A much more universal concept of a vector is the idea of 'velocity', the idea of movement and direction at a given point. A tangent vector is a 'derivative' form at a given point in the manifold. This means that we define it to obey properties that one might expect in our usual notion of a derivative for functions in \mathbb{R} . We denote a vector at a point $p \in \mathcal{M}$ as X_p . This means that a vector is simply a map $X_p : C^\infty \rightarrow \mathbb{R}$, which satisfies

- Linearity:

$$X_p(\alpha f + \beta g) = \alpha X_p(f) + \beta X_p(g), \quad \forall f, g \in C^\infty(\mathcal{M}), \alpha, \beta \in \mathbb{R}$$

- $X_p(f) = 0$ for constant functions on the manifold.
- Much like the product rule in differentiation, tangent vectors should also obey the Leibniz rule where

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

Remember that with the Leibniz rule, the functions which are not differentiated are evaluated at p ! This is useful for our theorem afterwards.

This next proof is about showing that tangent vectors can be built from differential operators in the n dimensions of the manifold. We will now show that all tangent vectors X_p have the property that they can be written out as

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

What we're saying here is that ∂_μ at the point $p \in \mathcal{M}$ forms a basis for the space of tangent vectors at a point. To do this, take your favourite arbitrary function $f : \mathcal{M} \rightarrow \mathbb{R}$. Since this is defined on the manifold, to make our lives easier we'll define $F = f \circ \phi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}$, which we know how to differentiate. The first thing we'll show is that we can locally move from $F(x(p)) \rightarrow F(x(q))$ by doing something like a Taylor expansion:

$$F(x(q)) = F(x(p)) + (x^\mu(q) - x^\mu(p))F_\mu(x(p))$$

Here, we're fixing $p \in \mathcal{M}$ and F_μ is some collection of n functions. One can easily verify that F can be written in this way by precisely doing a Taylor expansion then factorising out the factors of

$(x^\mu(q) - x^\mu(p))$. We can find an explicit expression for $F_\mu(x(p))$ by differentiating both sides and then evaluating at $x(p)$. We have that

$$\left. \frac{\partial F}{\partial x^\nu} \right|_{x(p)} = \delta^\mu_\nu F_\mu + (x^\mu(p) - x^\mu(p)) \left. \frac{\partial F_\mu}{\partial x^\nu} \right|_{x(p)} = F_\nu$$

The second term goes to zero since we're evaluating at $x(p)$, and our delta function comes from differentiating a coordinate element. Our initial $F(x(p))$ term goes to zero since it was just a constant. Recalling that $\phi^{-1} \circ x^\mu(p) = p$, we can just rewrite this whole thing as

$$f(q) = f(p) + (x^\mu(q) - x^\mu(p)) f_\mu(p)$$

where in this case we've defined that $f_\mu(p) = F_\mu \circ \phi^{-1}$. However, we can figure out what this is explicitly

$$f_\mu(p) = F_\mu \circ \phi(p) = F_\mu(x(p)) = \frac{\partial F(x(p))}{\partial x^\mu} = \frac{f \circ \phi^{-1}(x(p))}{\partial x^\mu} := \left. \frac{\partial f}{\partial x^\mu} \right|_p$$

Now, its a matter of applying our tangent vector to our previous equation, recalling that $X_p(k) = 0$ for constant k , and that all functions are evaluated at the point p . We have that, upon application of the Leibniz rule

$$\begin{aligned} X_p(f(q)) &= X_p(f(p)) + X_p(x^\mu(q) - x^\mu(p)) f_\mu(p) + (x^\mu(p) - x^\mu(p)) X_p(f_\mu(p)) \\ &= X_p(x^\mu(p)) f_\mu(p) \\ &= X^\mu f_\mu(p) \\ &= X^\mu \left. \frac{\partial f}{\partial x^\mu} \right|_p \end{aligned}$$

In the first line we've replaced q with p in the last term since Leibniz rule forces evaluation at p . We've declared $X^\mu = X_p(x^\mu)$ as our components. Since f was arbitrary, we have now written that

$$X_p = X^\mu \frac{\partial}{\partial x^\mu}$$

To show that $\{\partial_\mu\}$ forms a basis for all tangent vectors, since we've already shown that they span the space we need to show they're linearly independent. Suppose that

$$0 = X^\mu \partial_\mu$$

Then, this implies that if we take $f = x^\nu$, then $0 = X^\nu$ for any value of the index ν we take. So, we have linear independence.

Tangent vectors should be basis invariant objects

A tangent vector is a physical thing. However, so far we've expressed it in terms of the basis objects $\{\partial_\mu\}$ which are chart dependent. So, suppose we use a different chart which is denoted by coordinates \tilde{x}^μ . This means that our new tangent vector needs to satisfy the condition that

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

This relation allows us to appropriately relate the components X^μ to that of \tilde{X}^μ , in what is called a contravariant transformation. Using the chain rule, we have that

$$X^\mu \frac{\partial}{\partial x^\mu} \Big|_p = X^\mu \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \Big|_{\phi(p)} \frac{\partial}{\partial \tilde{x}^\nu} \Big|_p$$

Notice that when differentiating a coordinate chart with respect to another, we're evaluating at the coordinate chart of the point. This is why we subscript with $\phi(p)$ in the terms. Comparing coefficients, we have that

$$\tilde{X}^\nu = X^\mu \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \Big|_{\phi(p)}$$

3.2 Treating tangent vectors as derivatives of curves on the manifold

We present a different way to think of tangent vectors, which is viewing them as 'differential operators along curves'. Consider a smooth curve along our manifold, which we can parametrise from on an open interval $I = (0, 1)$, and define the starting point of this curve at $p \in \mathcal{M}$;

$$\lambda : (0, 1) \rightarrow \mathcal{M}, \quad \lambda(0) = p$$

We now ask the question, how do we differentiate along this thing? To do this, we'll have to apply coordinate charts so that we can make sense of differentiation. So, suppose we would like to differentiate a function f along this manifold. We apply our chart ϕ to λ to get a new function $\phi \circ \lambda : \mathbb{R} \rightarrow \mathbb{R}^n$, which we'll suggestively write as $x^\mu(t)$. In addition, to be able to differentiate f in a sensible way we also construct the function $F = f \circ \phi^{-1}$. Thus, differentiating a function along a curve $x^\mu(t) = \phi \circ \lambda(t)$ should look like

$$\begin{aligned} \frac{d}{dt} f(t) &= \frac{d}{dt} (F \circ \phi^{-1} \circ \phi \circ \lambda(t)) \\ &= \frac{d}{dt} (F \circ \phi^{-1} \circ x^\mu(t)) \Big|_{t=0} \\ &= \frac{dx^\mu}{dt} \frac{\partial F \circ \phi^{-1}}{\partial x^\mu} \Big|_{\phi(p)} \\ &= \frac{dx^\mu}{dt} \Big|_{t=0} \frac{\partial f}{\partial x^\mu} \Big|_p \\ &= X^\mu \partial_\mu(f) \\ &= X_p(f) \end{aligned}$$

Thus, differentiating along a curve gives rise to a tangent vector acting on f .

3.3 Vector fields

Thus far we've defined tangent spaces at only a specific point in the manifold, but we'd like to know how we can extend this notion more generally. A vector field X is an object which takes a function, and then assigns it a vector at any given point $p \in \mathcal{M}$. So, we're taking

$$X : C^\infty(\mathcal{M}) \rightarrow C^\infty(\mathcal{M}), \quad f \mapsto X(f)$$

$X(f)$ is a function on the manifold which takes a point, and then differentiates f according to the tangent vector at that point.

$$X(f)(p) = X_p(f), X = X^\mu \partial_\mu$$

In this case, X^μ is a smooth function which takes points on the manifold to the components of the tangent vector X_p^μ at p . We call the space of vector fields X as $\mathcal{X}(\mathcal{M})$. So, since $X(f)$ is now also a smooth function on the manifold, we can apply another vector field Y to it, for example. However, is the object XY a vector field on it's own? The answer is no, because vector fields also have to obey the Leibniz identity at any given point, ie the condition that

$$X(fg) = fX(g) + gX(f)$$

However, the object XY does not obey this condition since

$$\begin{aligned} XY(fg) &= X(fY(g) + gY(f)) \\ &= X(f)Y(g) + fXY(g) + X(g)Y(f) + gXY(f) \\ &\neq gXY(f) + fXY(g) \end{aligned}$$

We do get from this however, that

$$XY - YX := [X, Y]$$

does obey the Leibniz condition, because it removes the non-Leibniz cross terms from our differentiation. The commutator acts on a function f by

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

One can check that the components of the new vector field $[X, Y]^\mu$ are given by

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

The commutator obeys the Leibniz rule, where

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

3.4 Integral curves

We'll now do an interesting diversion to discuss flows on a manifold. Think of a body of water moving smoothly on a surface, where each point on the manifold moves to a different point after some amount of time. This is what a flow is. More specifically, it is a smooth map $\sigma_t : \mathcal{M} \rightarrow \mathcal{M}$ on the manifold (which makes it a diffeomorphism), where t is our 'time' parameter that we were talking about. As such, these flow maps actually form an abelian group, where

- $\sigma_0 = I_{\mathcal{M}}$. So, after time $t = 0$ has passed, nothing has moved so we have that this is the identity map.
- If we compose the same flow after two intervals in time, we should get the same flow when we've let the sum of those times pass over. So,

$$\sigma_s \circ \sigma_t = \sigma_{s+t}$$

If we take a flow at a given point $p \in \mathcal{M}$, we can define a curve on the manifold by setting:

$$\gamma(t) : \mathbb{R} \rightarrow (M), \quad \gamma(t) = \sigma_t(p)$$

where without loss of generality we have $\gamma(0) = p$. Since this is a curve, we can define it's associated curve in R^n space with a given coordinate chart, and hence associate with it a tangent vector X_p . We can also work backwards. From a given vector field, we can solve the differential equation

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

with the initial condition that $x^\mu(0) = \psi(p)$ at some point in the manifold, and have that this defines a unique curve. The set of curves then together form a flow. Thus, we've seen a one to one correspondence between vector fields and flows.

3.5 Differentiating vector fields with respect to other vector fields

3.5.1 Push-forwards and Pull-backs

A sensible question to now ask is that, since we have these smooth vector fields, how do we differentiate a vector field with respect to another one? For example, if we have $X, Y \in \mathcal{M}$, what constitutes the notion of a change in X with respect to Y . The notion of derivatives on manifolds is difficult because we can't compare tangent spaces at different points in the manifold, for example $T_p(M)$ and $T_q(M)$ are tangent spaces at different points, and we could define different charts for each space, hence we have some degrees of freedom (and our derivative wouldn't make sense). To make sense of comparing different tangent spaces, we need to create way to compare the same functions, but on different manifolds. These are called push forwards and pull backs.

Let's start by defining a smooth map between manifolds $\phi : M \rightarrow N$ (ϕ is not a chart here). We're not assuming that M and N are even the same dimension here, and so we can't assume ϕ^{-1} doesn't even exist.

Suppose we have a function $f : N \rightarrow \mathbb{R}$. How can we define a new function based on f that makes sense, which goes from $M \rightarrow \mathbb{R}$? We define the pull back of a function f , denoted $(\psi^*f) : M \rightarrow \mathbb{R}$ as

$$(\psi^*f)(p) = f(\psi(p)), \quad p \in M$$

So, we've converted this thing nicely.

Our next question then is how, from a vector field $Y \in \mathcal{X}(M)$, can we make a new vector field in $X \in \mathcal{X}(N)$? We can, and this is called the push-forward of a vector field, denoted $\phi_*Y \in \mathcal{X}(N)$. We define that object as the vector field which takes

$$(\phi_*Y)(f) = Y(\phi^*f)$$

This makes sense because $\phi^*f \in C^\infty(M)$, so applying Y makes sense. Now, to show $\phi_*Y \in \mathcal{X}(N)$, we should verify that

$$\phi_*Y : C^\infty(N) \rightarrow C^\infty(N), \quad f \mapsto C^\infty$$

Well, this object philosophically maps

$$f \mapsto \phi_*Y(f) = Y(\phi^*f)$$

But the object on the left hand side is a vector field ready to be turned into a tangent vector when we assign it to a point on the manifold:

$$p \mapsto Y_p(\phi^*f)$$

Hence this object agrees with our definition. The fact that we have a map $\phi : M \rightarrow N$ and are pushing the vector field from $\mathcal{X}(M)$ to $\mathcal{X}(N)$ is the reason why we call this new mapping a push forward.

3.5.2 Components for Push-forwards and Pull-backs

Now, since ψ_*Y is a vector field, it's now in our interest to find out about what the components are for this thing. We want to find that the components $(\psi_*Y)^\nu$ such that

$$\psi_*Y = (\psi_*Y)^\nu \partial_\nu$$

We can work first by assigning coordinates to $\phi(x)$, which we denote by $y^\alpha(x) = \phi(x)$, $x \in M, \alpha = 1, \dots, \dim(N)$. If we write out our vector field Y as $Y = Y^\mu \partial_\mu$, then our push-forward map in summation convention looks like

$$(\phi_* Y)f = Y^\mu \frac{\partial f(y(x))}{\partial x^\mu} = Y^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial f}{\partial y^\alpha}$$

In the second equality, we've applied the chain rule. Remember, y pertains to coordinates in the manifold N , so on our push-forward, we have that our new components on the manifold N , we have that

$$(\phi_* Y)^\alpha = Y^\mu \frac{\partial y^\alpha}{\partial x^\mu}$$

3.5.3 Introducing the Lie derivative

In what we've just presented, some objects are naturally pulled back and some are naturally pushed forward. However, things become when our map between manifolds is a diffeomorphism and hence invertible; which means we can pull back and push forward with whatever objects we want. We can use this idea to differentiate vector fields now. Recall that if we've given a vector field, $X \in \mathcal{X}(M)$, we can define a flow map $\sigma_t : M \rightarrow M$. This flow map diffeomorphism allows us to push vectors along flow lines in the manifolds, from the tangent spaces

$$T_p(M) \rightarrow T_{\sigma_t(p)}(M)$$

This is called the Lie derivative \mathcal{L}_X , a derivative which is induced by our flow map generated by X . For functions, we have that

$$\mathcal{L}_X f = \lim_{t \rightarrow 0} \frac{f(\sigma_t(x)) - f(x)}{t} = \left. \frac{df(\sigma_t(x))}{dt} \right|_{t=0}$$

However, the effect of doing this is exactly the same as if we were to apply the vector field X to the function:

$$\left. \frac{df}{dx^\mu} \frac{dx^\mu}{dt} \right|_{t=0} = X^\mu \frac{\partial f}{\partial x^\mu} = X(f)$$

Thus, a Lie derivative specialised to the case of functions just gives us $\mathcal{L}_X(f) = X(f)$. Now, the question is about how we can do this differentiation on vector fields. What we need to do is to 'flow' vectors at a point back in time to where they originally started, and look at this difference.

$$(\mathcal{L}_X Y)_p = \lim_{t \rightarrow 0} \frac{(\sigma_{-t}^*(Y)_p - Y_p)}{t}$$

Let's try and compute the most basic thing first, the Lie derivative of a basis element of the tangent space ∂_μ :

$$\sigma_{-t}^* \partial_\mu = (\sigma_{-t}^* \partial_\mu)^\nu \partial_\nu$$

Let's try and figure out what $(\sigma_{-t}^* \partial_\mu)^\nu$ is. Because of the fact that the diffeomorphism is induced by the vector field X , we have that

$$(\sigma_{-t}^*(x))^\nu = x^\nu - tX^\nu + \dots$$

Thus our components of a push-forward of an arbitrary vector field are given by

$$(\sigma_{-t}^* Y)^\nu = Y^\sigma \frac{\partial(\sigma_{-t}(x))^\nu}{\partial x^\sigma}, \quad \text{in our case } Y^\sigma = \delta^\sigma_\mu$$

Substituting the expressions above with one another gives us that

$$(\sigma_{-t}(x)\partial_\mu)^\nu = \delta^\nu_\mu - t \frac{\partial X^\nu}{\partial x^\mu} +$$

Contracting this with ∂_ν , and then subtracting off ∂_μ , we have that

$$\mathcal{L}_X \partial_\mu = -\frac{\partial X^\nu}{\partial x^\nu} \partial_\nu$$

We require that a Lie derivative obeys the Leibniz rule, so we have that applying on a general vector field Y ,

$$\begin{aligned} \mathcal{L}_X(Y) &= \mathcal{L}_X(Y^\mu \partial_\mu) \\ &= \mathcal{L}_X(Y^\mu) \partial_\mu + Y^\mu \mathcal{L}_X(\partial_\mu) \\ &= X^\nu \frac{\partial Y^\mu}{\partial x^\nu} \partial_\mu - Y^\mu \frac{\partial X^\nu}{\partial x^\mu} \partial_\nu \end{aligned}$$

We realise that this however are just components of the commutator! So

$$L_X(Y) = [X, Y]$$

4 Tensors

4.1 A note about dual spaces

In linear algebra, if we have a vector space which we call V , then we can define a natural object which we call it's dual space, denoted by V^* . The dual space is the space of linear functions which takes $V \rightarrow \mathbb{R}$:

$$V^* = \{f \mid f : V \rightarrow \mathbb{R}, \quad f \text{ is linear} \}$$

Now it may seem from first glance that the space of all functions is a lot larger than our original vector space, so it's counter intuitiveto call it the 'dual'. However, we can prove that these vector spaces are isomorphic. Suppose that $\{e_\mu\}$ is a basis of V . Then we pick what we call a dual basis of V^* , by choosing the

$$\mathcal{B}(V^*) = \{f^\mu \mid f^\mu(e_\nu) = \delta^\mu_\nu\}$$

One can show that this set indeed forms a basis of V^* .

Theorem. The above basis forms a basis of V^* .

Proof. First we need to show that the linear maps above, span our space. This means we need to be able to write any linear map, say ω , as a linear combination

$$\omega = \sum \omega_\mu f^\mu, \omega_\mu \in F$$

To do this, we appeal to the fact that if two linear maps agree on the vector space's basis, then they agree. So, let the values that ω takes on the basis be $\omega_\mu = \omega(e_\mu)$. Then, taking

$$\Omega = \sum \omega_\mu f^\mu$$

We find that Ω also satisfies $\Omega(e_\mu) = \omega_\mu$. Thus, the maps are the same. Hence, ω can be written as the span of our dual basis vectors. To show that these basis vectors are linearly independent, we assume that there exists a non trivial sum such that they add to the zero map.

$$0 = \sum \lambda_\mu f^\mu$$

If we apply this map to an arbitrary basis vector e_i , then we get

$$0 = \sum \lambda_\mu f^\mu(e_i) = \lambda_i$$

for arbitrary i . Hence, we must have that all λ_i are zero. Thus the basis vectors are independent. \square

Now, assuming that our original vector space V had finite dimension, the way we've defined the basis of V^* means that we had the same number of basis elements. This means that V and V^* have the same dimension. One can prove that vector spaces with the same dimension are isomorphic, so we have that

$$V \simeq V^*$$

Think of a dual space as a 'flip' of a vector space. We can identify the dual of a dual space as the original space itself, so that

$$(V^*)^* = V$$

This is because given an object in the dual space ω , we can define a natural map from $V : V^* \rightarrow \mathbb{R}$ given by

$$V(\omega) = \omega(V) \in \mathbb{R}$$

4.1.1 Vector and Covector spaces

Now, since we've identified tangent spaces as vector spaces, we can proceed to construct its dual. If we have a tangent space $T_p(\mathcal{M})$ with a basis $\{e_\mu\}$, our natural dual basis is given by

$$\mathcal{B}(T_p^*(\mathcal{M})) = \{f^\mu \mid f^\mu(e_\nu) = \delta^\mu_\nu\}$$

The corresponding dual space is denoted as $T_p^*(\mathcal{M})$, and is known as the cotangent vector space. For brevity, elements in this space are called **covectors**. In this basis, the elements $\{f^\mu\}$ have the effect of 'picking' out components of a vector $V = V^\mu e_\mu$.

$$f^\nu(V) = f^\nu(V^\mu e_\mu) = V^\mu f^\nu(e_\mu) = V^\nu$$

There's a different way to pick elements of this dual space in a smooth way. They're chosen by picking elements of a set called the set of 'one forms'. We denote the set of one forms, with an index 1 as $\Lambda^1(\mathcal{M})$. We can construct elements from this set by taking elements from $C^\infty(\mathcal{M})$. Suppose that we have an $f \in C^\infty(\mathcal{M})$, then the corresponding one-form is a map

$$df : T_p(\mathcal{M}) \rightarrow \mathbb{R} \quad V \mapsto V(f)$$

From one the set of one forms, we then have an obvious way to get the dual basis. The dual basis is obtained by just taking the coordinate element of the manifold, so that our one form is

$$dx^\nu : T_p(\mathcal{M}) \rightarrow \mathbb{R}$$

This satisfies the property of a dual basis, since

$$dx^\nu(e_\mu) = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu$$

With this convention, we can check that $V(f)$ is what its supposed to be by observing that

$$df(X) = \frac{\partial f}{\partial x_\mu} dx^\mu (X^\nu \partial_\nu) = X^\nu \frac{\partial f}{\partial x^\nu} = X(f)$$

So we recover what we expect by setting this as a basis. Now, we should check whether a change in coordinates leaves our properties invariant.

Suppose we change our basis from $x^\mu \rightarrow \tilde{x}^\mu(x)$, then, we know that our basis vector transforms like

$$\frac{\partial}{\partial \tilde{x}^\mu} = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu$$

We guess that our basis of one forms should transform as

$$d\tilde{x}^\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu$$

This ensures that, when we contract a transformed basis one form with a transformed basis vector, that

$$\begin{aligned} d\tilde{x}^\mu \frac{\partial}{\partial \tilde{x}^\nu} &= \frac{\partial \tilde{x}^\mu}{\partial x^\rho} dx^\rho \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\sigma} \\ &= \frac{\partial \tilde{x}^\mu}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} dx^\rho \left(\frac{\partial}{\partial dx^\sigma} \right) \\ &= \frac{\partial \tilde{x}^\mu}{\partial x^\rho} \frac{\partial x^\rho}{\partial \tilde{x}^\nu} = \delta^\mu_\nu \end{aligned}$$

It's no coincidence that this looks like a Jacobian!

Now, as with basis elements in our vector space, we need to determine how these objects transform under a change of basis. *Need to finish this section on basis transformations for covectors*

4.2 Taking the Lie Derivative of Covectors

We would like to repeat what we did for vectors, and take derivatives of covectors. To do this, we need to define pull-backs for covectors. Suppose we had a covector ω living in the tangent space of some manifold N , in $T_p^*(N)$. We can then define the pull back of this vector field based on first pushing forward the vector field X . Thus, if $\phi : M \rightarrow N$, then we define the pullback $\phi^*\omega$ as the covector field in $T_p(M)$ as

$$(\phi^*\omega)(X) = \omega(\phi_*X)$$

What information can we glean from this? Well, we can try to figure out what the components of $\phi^*\omega$ are. If we let $\{y^\alpha\}$ to be coordinates on N , then we expand this covector as

$$\omega = \omega_\mu dy^\mu$$

Also recall that the components of a pushed forward vector field are

$$(\phi_* X)^\mu = X^\nu \frac{\partial y^\mu}{\partial x^\nu}$$

Now, if we take the equation

$$(\phi^* \omega) = \omega(\phi_* X)$$

Then, expanding out in terms of components, we have that

$$(\phi^* \omega)^\mu dx_\mu (X^\nu e_\nu) = \omega^\mu dy_\mu (\phi_* X)^\nu \frac{\partial}{\partial y^\nu}$$

Remember, we have to expand in the correct coordinates. The object $\phi^* \omega$ lives in the space M so we expand in the dx^μ basis. On the other hand we have that ω originally lives in N so we expand in dy^μ . Substituting our expression for our push forward vector field, and we get that

$$(\phi^* \omega)_\mu X^\mu = \omega_\nu \frac{\partial y^\nu}{\partial x^\mu} X^\mu$$

This step requires a bit of explanation. After we substitute in the components for the pushed forward vector field, we then use the fact that on both manifolds, our basis vectors and our basis covectors contract to give a delta function

$$dx^\mu \left(\frac{\partial}{\partial x^\nu} \right) = \delta^\mu_\nu, \quad dy^\mu \left(\frac{\partial}{\partial y^\nu} \right) = \delta^\mu_\nu$$

This implies that the components of our pulled back vector field are

$$(\phi^* \omega)_\mu = \omega_\nu \frac{\partial y^\nu}{\partial x^\mu}$$

As in the case of vectors, we can also make rigorous the definition of a Lie derivative with respect to a covector field. This is denoted $\mathcal{L}_X \omega$, where X is our underlying vector field we're differentiating with. If our vector field X imposes a flow map which we label as σ_t , then our corresponding Lie derivative is defined as

$$\mathcal{L}_X \omega = \lim_{t \rightarrow 0} \frac{(\sigma_t^* \omega) - \omega}{t}$$

There's an important point to be made here. In our previous definition of a Lie derivative for a vector field, we took the inverse diffeomorphism σ_{-t} . But in this case, we need to take t positive since we're doing a **pull-back** instead of a pushforward, like we did with vector fields.

Let's go slow and try to compute the components of this derivative. Recall that for a flow map, we have that infinitesimally,

$$y^\nu = x^\nu + tX^\nu, \implies \delta^\nu_\mu + t \frac{dX^\nu}{dx^\mu}$$

Thus for a general covector field, our components for the pull back are

$$(\sigma_t^* \omega)_\mu = \omega_\mu + t\omega_\nu \frac{dX^\nu}{dx^\mu}$$

Hence, the components of a basis element under this flow becomes

$$(\sigma_t^* dx^\nu) = dx^\nu + t dx^\mu \frac{dX^\nu}{dx^\mu}$$

So, taking the limit, we have that our components of our Lie derivative are given by

$$\lim_{t \rightarrow 0} \frac{\sigma_t^* dx^\nu - dx^\nu}{t} = dx^\mu \frac{dX^\nu}{dx^\mu}$$

Now, as before, we impose the Liebniz property of Lie derivatives, and expand out a general covector. Hence, we have that

$$\begin{aligned} \mathcal{L}_X(\omega_\mu dx^\mu) &= \omega_\mu \mathcal{L} dx^\mu + dx^\mu \mathcal{L}_X(\omega_\mu) \\ &= \omega_\mu dx^\nu \frac{dX^\nu}{dx^\mu} + dx^\nu X^\mu \frac{d\omega_\nu}{dx^\mu} \end{aligned}$$

This implies that our components of our Lie derivative can be written nicely as

$$\mathcal{L}_X(\omega)_\mu = (X^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu X^\nu)$$

4.3 Tensor fields

Now, we can combine both maps from tangent and cotangent spaces to create tensors. A tensor of rank (r, s) is a **multilinear** map from

$$T : T_p^*(M) \times \dots T_p^*(M) \times T_p(M) \times \dots T_p(M) \rightarrow \mathbb{R}$$

where we have r copies of our cotangent field and s copies of our tangent field. We define the total rank of this multilinear map as $r + s$. Since a cotangent vector is a map from vectors to the reals, this is a rank $(0, 1)$ tensor. Also, a tangent vector has rank $(1, 0)$ since it's a map from the cotangent space. A tensor field is the smooth assignment of a rank (r, s) tensor to a point on the manifold $p \in M$. We can write the components of a tensor object by writing down a basis, then sticking this into the object. If $\{e_\nu\}$ was a basis of $T_p(M)$, and $\{f^\nu\}$ was the basis of the dual space, then the tensor has components

$$T^{\mu_1 \mu_2 \dots \mu_r}_{\nu_1 \nu_2 \dots \nu_s} = T(f^{\mu_1}, f^{\mu_2}, \dots, f^{\mu_r}, e_{\nu_1}, \dots e_{\nu_s})$$

For example, a $(2, 1)$ tensor acts as

$$T(\omega, \epsilon; X) = T(\omega_\mu f^\mu, \epsilon_\nu f^\nu, X^\rho e_\rho)$$

We have that the covectors $\omega, \epsilon \in \Lambda^1(M)$, and $X \in \mathcal{X}(M)$. The object above is then equal by multilinearity to

$$= \omega_\mu \epsilon_\nu X^\rho T^{\mu\nu}_\rho$$

Under a change of coordinates, we have that

$$\tilde{e}_\nu = A^\mu_\nu e_\mu, \quad A^\mu_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}$$

Similarly, we have that for covectors we transform as

$$\tilde{f}^\rho = B^\rho_\sigma f^\sigma, \quad B^\rho_\sigma = \frac{\partial \tilde{x}^\rho}{\partial x^\sigma}$$

Thus, a rank $(2, 1)$ tensor transforms as

$$\tilde{T}^{\mu\nu}_{\rho} = B^{\mu}_{\sigma} B^{\nu}_{\tau} A^{\lambda}_{\rho} T^{\sigma\tau}_{\lambda}$$

There are a number of operations which we can perform on tensors. We can add or subtract tensors. We can also take the tensor product. If S has rank (p, q) , and T has rank (r, s) , then we can construct $T \otimes S$, which has rank $(p + r, q + s)$.

$$S \otimes T(\omega_1, \dots, \omega_p, \nu_1, \dots, \nu_r, X_1, \dots, X_q, Y_1, \dots, Y_s) = S(\omega_1, \dots, \omega_p, X_1, \dots, X_q)T(\nu_1, \dots, \nu_r, Y_1, \dots, Y_s)$$

Our components of this are

$$(S \otimes T)^{\mu_1 \dots \mu_p \nu_1 \dots \nu_r}_{\rho_1 \dots \rho_l \sigma_1 \dots \sigma_s} = S^{\mu_1 \dots \mu_p}_{\rho_1 \dots \rho_q} T^{\nu_1 \dots \nu_q}_{\sigma_1 \dots \sigma_s}$$

We can also define a contraction. We can turn a (r, s) tensor into an $(r - 1, s - 1)$ tensor. If we have T a $(2, 1)$ tensor, then we can define a

$$S(\omega) = T(\omega, f^{\mu}, e_{\mu})$$

The sum over μ is basis independent. This has components

$$S^{\mu} = T^{\mu\nu}_{\nu}$$

This is different from $(S')^{\mu} = T^{\nu\mu}_{\nu}$. We can also symmetrise and anti symmetrise. Given a $(0, 2)$ tensor, we can define

$$S(X, Y) = \frac{1}{2}(T(X, Y) + T(Y, X)), \quad A(X, Y) = \frac{1}{2}(T(X, Y) - T(Y, X))$$

This has components which we write as

$$T_{(\mu\nu)} := \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu})$$

$$T_{[\mu\nu]} := \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

We can also symmetrise or anti symmetrise over multiple indices. So

$$T^{\mu}_{(\nu\rho\sigma)} = \frac{1}{3!}(T^{\mu}_{\nu\rho\sigma} + 5 \text{ perms})$$

We can also anti symmetrise by multiplying by the sign of permutations.

$$T^{\mu}_{[\nu\rho\sigma]} = \frac{1}{3!}(T^{\mu}_{\nu\rho\sigma} + \text{sgn}(\text{perm}) \text{ for 5 perms})$$

4.4 Differential forms

Differential forms are totally antisymmetric $(0, p)$ tensors, and are denoted $\Lambda^p(M)$. 0-forms are functions. If $\dim(M) = n$, then p-forms have n choose p components by anti-symmetry. n-forms are called top-forms.

4.4.1 Wedge products

Given a $\omega \in \Lambda^p(M)$ and $\epsilon \in \Lambda^q(M)$, we can form a $(p + q)$ form by taking the tensor product and antisymmetrising. This is the wedge product. Our components are given by

$$(\omega \wedge \epsilon)_{\mu_1 \dots \mu_p \nu_1 \dots \nu_q} = \frac{(p+q)!}{p!q!} \omega_{[\mu_1 \dots \mu_p} \epsilon_{\nu_1 \dots \nu_q]}$$

An intuitive way to think about this is that we are simply just adding anti-symmetric combinations of forms, without dividing (other than to make up for the previous anti-symmetry). So, we have that, for example, when we wedge product the forms dx^1 with dx^2 , that

$$dx^1 \wedge dx^2 = dx^1 \otimes dx^2 - dx^2 \otimes dx^1$$

In terms of components one can check that, for example, for one forms we have that

$$(\omega \wedge \epsilon)_{\mu\nu} = \omega_\mu \epsilon_\nu - \omega_\nu \epsilon_\mu$$

We can iteratively wedge the basis of forms $\{dx^\mu\}$ together to find that

$$dx^1 \wedge \dots \wedge dx^n = \sum_{\sigma \in S_n} \epsilon(\sigma) dx^{\sigma(1)} \otimes \dots \otimes dx^{\sigma(n)}$$

To show this, we use an example. Note that the components of $dx^1 \wedge dx^2$ are

$$dx^1 \wedge dx^2 = 2\delta_{[\mu}^1 \delta_{\nu]}^2 dx^\mu dx^\nu$$

Now, this means that wedging this with dx^3 gives components

$$(dx^1 \wedge dx^2) \wedge dx^3 = \frac{3!}{2} 2\delta_{[\mu}^1 \delta_{\nu]}^2 \delta_{\rho]}^3 dx^\mu dx^\nu dx^\rho$$

But this is the sum of all permutations multiplied by the sign, since a set of antisymmetrised indices nested in a bigger set is the original set.

4.4.2 Properties of wedge products

Our antisymmetry property of forms gives it properties we might expect. One of these is that switching a p and q form picks up a sign: we have that

$$\omega \wedge \epsilon = (-1)^{pq} \epsilon \wedge \omega$$

In general, for an odd form we have that

$$\omega \wedge \omega = 0$$

For the manifold $M = \mathbb{R}^3$, with $\omega, \epsilon \in \Lambda^1(M)$, we have that

$$(\omega \wedge \epsilon) = (\omega_1 dx^1 + \omega_2 dx^2 + \omega_3 dx^3) \wedge (\epsilon_1 dx^1 + \epsilon_2 dx^2 + \epsilon_3 dx^3)$$

expanding this thing, we have that

$$\begin{aligned} \omega \wedge \epsilon &= (\omega_1 \epsilon_2 - \epsilon_2 \omega_1) dx^1 \wedge dx^2 \\ &\quad + (\omega_2 \epsilon_3 - \omega_3 \epsilon_2) dx^2 \wedge dx^3 \\ &\quad + (\omega_3 \epsilon_1 - \omega_1 \epsilon_3) dx^3 \wedge dx^1 \end{aligned}$$

These are the components of the cross product. The cross product is really just a wedge product between forms. In a coordinate basis, we write that

$$\omega = \frac{1}{p!} w_{\mu_1 \dots \mu_p} dx^{\mu_1} \wedge \dots \wedge dx^{\mu_p}, \quad \omega = w_{\mu_1 \dots \mu_p} dx^{\mu_1} \otimes \dots \otimes dx^{\mu_p}$$

This is useful because writing out forms in terms of wedge products as their basis turns out to make calculations a lot easier.

4.5 The exterior derivative

Notice that given a function f , we can construct a 1-form

$$df = \frac{\partial f}{\partial x^\mu}$$

In general, there exists a map $d : \Lambda^p(M) \rightarrow \Lambda^{p+1}(M)$. this is the exterior derivative. In coordinates, we have that

$$dw = \frac{1}{p!} \frac{\partial \omega_{\mu_1 \dots \mu_p}}{\partial x^\nu} dx^\nu \wedge \dots \wedge dx^{\mu_p}$$

One should view this as a generalised curl of some vector field. In components, we have that

$$(d\omega)_{\mu_1 \dots \mu_{p+1}} = (p+1) \partial_{[\mu_1} \omega_{\mu_2 \dots \mu_{p+1}]}$$

Let's try to gain an intuition about why these two definitions are equivalent. If we contract our component definition with the tensor product $dx^1 \otimes \dots \otimes dx^{p+1}$, then we are summing over

$$d\omega = \frac{1}{p!} \sum_{\sigma \in S_n} \epsilon(\sigma) \partial_{\sigma(\mu_1)} w_{\sigma(\mu_2) \dots \sigma(\mu_{p+1})} dx^{\mu_1} \otimes \dots \otimes dx^{\mu_{p+1}}$$

However we can transfer our permutations to permutations on tensor product, but by definition this would just be the wedge product on our basis one-forms.

By antisymmetry, we have a very significant identity that

$$d(d\omega) = 0$$

We write this as $d^2 = 0$. To show this, the easiest way is not to use our definition of $d\omega$ in components but rather to use our definition in terms of wedge product basis vectors. Let's think carefully about how the exterior derivative acts on some on p form but with 'components' in our

wedge product basis $\{dx^1 \wedge \cdots \wedge dx^p\}$. In our wedge product basis, our components are $\frac{1}{p!}\omega_{\mu_1 \dots \mu_p}$ since

$$\omega = \frac{1}{p!}\omega_{\mu_1 \dots \mu_p}$$

We have that under the exterior derivative, we are mapping

$$d : \frac{1}{p!}\omega_{\mu_1 \dots \mu_p} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p} \mapsto \frac{1}{p!}\partial_\nu \omega_{\mu_1 \dots \mu_p} dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p}$$

So, in our new fancy wedge product basis, we are mapping

$$w_{\mu_1 \dots \mu_p} \mapsto \partial_\nu \omega_{\mu_1 \dots \mu_p}$$

Hence, we have that, in our wedge product basis, our components of $d(d\omega)$ are

$$d(d\omega) = \frac{1}{p!}\partial_\rho \partial_\nu w_{\mu_1 \dots \mu_p} dx^\rho \wedge dx^\nu \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p} = 0$$

since we have symmetry of mixed partial derivatives in ρ, ν contracted with the wedge product which is antisymmetric in those indices.

It's also simple to show that

- $d(\omega \wedge \epsilon) = d\omega \wedge \epsilon + (-1)^p \omega \wedge d\epsilon$
- For pull backs, $d(\phi^* \omega) = \phi^*(d\omega)$
- $\mathcal{L}_X(d\omega) = d(\mathcal{L}_X \omega)$

A p-form is closed if $d\omega = 0$ everywhere. A p form is exact if $\omega = d\epsilon$ everywhere for some ϵ . We have that

$$d^2 = 0 \implies \text{exact} \implies \text{closed}$$

Poincare's lemma states that on \mathbb{R}^n , or locally on \mathcal{M} , exact implies closed.

4.5.1 Examples

Consider a one form $\omega = \omega_\mu(x)dx^\mu$. Using our formula for the exterior derivative:

$$(d\omega)_{\mu\nu} = \partial_\mu\omega_\nu - \partial_\nu\omega_\mu$$

Or, in terms of our form basis,

$$d\omega = \frac{1}{2}(\partial_\mu\omega_\nu - \partial_\nu\omega_\mu)dx^\mu \wedge dx^\nu$$

In three dimensions,

$$d\omega = (\partial_1\omega_2 - \partial_2\omega_1)dx^1dx^2 + (\partial_2\omega_3 - \partial_3\omega_2)dx^2 \wedge dx^3 + (\partial_3\omega_1 - \partial_1\omega_3)dx^3 \wedge dx^1$$

These are the components of $\nabla \times \omega$. The exterior derivative of a 1 form is a 2 form, but 2 forms have just three components in \mathbb{R}^3 by anti symmetry (with the components shown there). So we think of it has another vector field, if we identify the basis vectors $\{dx^i \wedge dx^j\}$ with components in \mathbb{R}^3 !. However, getting another 'vector field' by doing an exterior derivative on the same type of object we had before is not the case in general.

Consider $B \in \Lambda^2(\mathbb{R}^3)$. So, we have a 2-form in a 3 dimensional manifold. Let's label the components out explicitly here.

$$B = B_1(x)dx^2 \wedge dx^3 + B_2(x)dx^3dx^1 + B_3(x)dx^1 \wedge dx^2$$

Before we do any explicit calculation, we know that the exterior derivative pushes this up to a 3-form, which only has one component in a three dimensional manifold. We compute the exterior derivative explicitly by differentiating each component and then adding on the wedge.

$$\begin{aligned} dB &= \partial_1 B_1 dx^1 \wedge dx^2 \wedge dx^3 + \partial_2 B_2 dx^2 \wedge dx^3 \wedge dx^1 + \partial_3 B_3 dx^3 \wedge dx^1 \wedge dx^2 \\ &= (\partial_1 B_1 + \partial_2 B_2 + \partial_3 B_3) dx^1 \wedge dx^2 \wedge dx^3 \end{aligned}$$

In the last step, we permuted the indices cyclically so that we don't have a sign change. Note that we get our components of a grad operator acting on \mathbf{B} !

For our final example, we take something from electromagnetism. The gauge field, or perhaps more commonly known as our vector potential $A \in \Lambda^1(\mathbb{R}^4)$, can be written out as a one-form in \mathbb{R}^4 .

If we expand this as a one form, we can write

$$A = A_\mu dx^\mu$$

What happens when we take the exterior derivative of this thing? We get that

$$\begin{aligned} dA &= \partial_\nu A_\mu dx^\nu \wedge dx^\mu \\ &= \frac{1}{2} \partial_{[\mu} A_{\mu]} dx^\nu \wedge dx^\mu \\ &= \frac{1}{2} (\partial_\nu A_\mu - \partial_\mu A_\nu) dx^\nu \wedge dx^\mu \\ &= \frac{1}{2} F_{\mu\nu} dx^\mu dx^\nu \end{aligned}$$

We identify here that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is our electromagnetic field strength tensor! We also have that since $dF = d^2A = 0$, we get for free what one may recognise as the Bianchi identities.

We can also introduce gauge transformations which act on our electromagnetic vector potential. These act as

$$A \rightarrow A + d\alpha \implies F \rightarrow d(Ad\alpha) = dA \text{ invariant}$$

where we treat $\alpha \in \Lambda^0(\mathcal{M}) = C^\infty(\mathcal{M})$ We also get Maxwell's equations for free!

$$F = dA \implies dF = d^2A = 0$$

There's are two of Maxwell's equations.

4.6 Integration

On a manifold, we integrate functions

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

with the help of a special kind of a special kind of top form. The kind of form we need is called a volume form or orientation. This is a nowhere vanishing top form. Locally, it can be written as

$$v = v(x)dx^1 \wedge \dots \wedge dx^n, \quad v(x) \neq 0!$$

For some manifolds, globally we may not be able to glue volume forms together over the whole manifold. If such a form exists, the manifold is said to be orientable. Not all manifolds are orientable, for example the Mobius strip. This says that $v(x)$ must change direction and hence be zero. Or, \mathbb{RP}^n . Given a volume form, we can integrate any function $f : M \rightarrow \mathbb{R}$ over \mathcal{M} . In chart $\mathcal{O} \subset \mathcal{M}$, we define

$$\int_{\mathcal{O}} f v = \int_{\mathcal{U}} dx^1 \dots dx^n f(x) v(x)$$

Now, this tells us how to integrate a patch. Then, summing over patches gives us the whole integral. $v(x)$ can be thought of as our measure - 'the volume of some part of the manifold'. There is freedom in our choice of volume form here, we could've chosen lots of different volume forms which satisfy our condition above.

4.6.1 Integrating over submanifolds

We haven't defined how to integrate, say a function over a p -form on an n dimensional manifold, where $p < n$ (since so far our definition of integration has only pertained to top forms).

So to do things like this, we need to find a way to 'shift down' into a lower dimensional subspace and do things there. This is why we define the concept of a submanifold.

A new manifold Σ of dimension $k < N$ is called a submanifold of \mathcal{M} if there exists an injective map $\phi : \Sigma \rightarrow \mathcal{M}$ such that $\phi^* : T_p(\Sigma) \rightarrow T_p(\mathcal{M})$, the **pullback** of ϕ , is also injective. We require the condition of injectivity so that our submanifold doesn't intersect itself when we embedded it in our larger manifold. The first condition is so that there are no crossings. The second condition is there so that we have no cusps in our tangent space.

We're now fully equipped to integrate over some portion of a submanifold of \mathcal{M} . You can think of this as a 'surface' or 'line' of some sort embedded in our manifold. We can integrate any $\omega \in \Lambda^k(\mathcal{M})$

over Σ by first identifying it with the embedded portion of Σ in \mathcal{M} , and then pulling it back into Σ itself. Now, we're in a p dimensional space, and we know how to integrate here since $\phi^*\omega$ is now a top-form

$$\int_{\phi(\Sigma)} \omega = \int_{\Sigma} \phi^*\omega$$

Let's do an example where we integrate say over a line embedded in a bigger manifold. We define an injective map σ which takes our line C into our manifold.

$$\sigma : C \rightarrow \mathcal{M}$$

defines a non intersecting curve in \mathcal{M} . Then, for $A \in \Lambda^1(\mathcal{M})$, we have, integrating over our embedding of our line in \mathcal{M} ,

$$\int_{\sigma(C)} A = \int_C \sigma^* A = \int d\tau A_{\mu}(x) \frac{dx^{\mu}}{d\tau}$$

The last equality comes from the fact that the components of a one-form pulled back transforms as

$$(\sigma^* A)_{\nu} = A_{\mu} \frac{dx^{\mu}}{dt^{\nu}}$$

where in this case, since we're pulling back to a one dimensional manifold, we only have $\nu = 0$ (which we write for brevity as τ).

4.6.2 Stokes' theorem

How does this tie in to the bigger picture? What use does Stoke's theorem have in general relativity? How do we prove Stoke's theorem?

Definition. (Boundaries). So far we've only considered manifolds which are smooth. However, we can 'chop off' a portion of our manifold to make a slightly different map that what we are used to, a new map

$$\phi_{\alpha} : \mathcal{O}_{\alpha} \rightarrow \mathcal{U}_{\alpha} \subset \frac{1}{2}\mathbb{R}^n = \{(x_1, \dots x_n), | x_1 \geq 0\}$$

Our boundary of our manifold is the set of points on \mathcal{M} which are mapped to $(0, x_2, \dots x_n)$. Boundaries are $n - 1$ dimensional manifolds of our original manifold which has dimension n

Theorem. Stokes' Theorem. Let \mathcal{M} be a manifold with boundary. This is a manifold which just stops and gets cutoff somewhere. If we call the manifold \mathcal{M} , we call the boundary $\partial\mathcal{M}$. If we take $\omega \in \Lambda^{n-1}(\mathcal{M})$. Our claim is that

$$\int_{\mathcal{M}} d\omega = \int_{\partial\mathcal{M}} \omega$$

This Stokes' theorem. It's presented in a more general form than what we're used to, but we'll see in the examples that this way of expressing this gives us both Stokes' theorem in three dimensions as well as Green's theorem.

Example. Stokes' theorem in one dimension. Take \mathcal{M} as the interval I with $x \in [a, b]$. $\omega(x)$ is a function and

$$d\omega = \frac{d\omega}{dx} \cdot dx$$

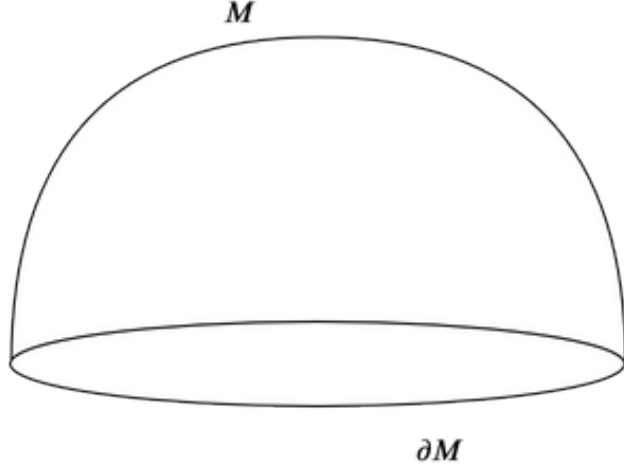


Figure 1: Here we have a manifold with boundary.

We have that

$$\int_{\mathcal{M}} d\omega = \int_a^b \frac{d\omega}{dx} dx \quad \int_{\partial\mathcal{M}} \omega = \omega(b) - \omega(a)$$

In one dimension, we've recovered integration by parts on a line.

Example. Stokes' theorem in two dimensions. In the second case, we have that

$$M \subset \mathbb{R}^2, \omega \in \Lambda^1(\mathcal{M})$$

This recovers

$$\int_{\mathcal{M}} d\omega = \int_{\mathcal{M}} \left(\frac{\partial\omega_2}{\partial x^1} - \frac{\partial\omega_1}{\partial x^2} \right) dx^1 \wedge dx^2$$

By Stokes theorem this is

$$\int_{\partial\mathcal{M}} \omega = \int_{\partial\mathcal{M}} \omega_1 dx^1 + \omega_2 dx^2$$

This equality is Green's theorem in a plane.

Example. Stokes' theorem in three dimensions. Finally, take $\mathcal{M} \subset \mathbb{R}^3$ and $\omega \in \Lambda^2(\mathcal{M})$

$$\begin{aligned} \int_{\mathcal{M}} d\omega &= \int d x^1 d x^2 d x^3 (\partial_1 \omega_1 + \partial_2 \omega_2 + \partial_3 \omega_3) \\ \int_{\partial\mathcal{M}} \omega &= \int_{\partial\mathcal{M}} \omega_1 d x^2 d x^3 + \omega_2 d x^3 d x^1 + \omega_3 d x^1 d x^2 \end{aligned}$$

Equating the two expressions above gives us our usual notion of Stokes' theorem in three dimensions, which is disguised as Gauss' divergence theorem.

5 Introducing Riemannian Geometry

5.1 The metric

Definition. (The metric tensor). We'll now do introduce a tensor object which turns our tangent space into an inner product space. We do this by introducing an object called a metric, which intuitively has been a way in which we define the notion of 'distance' in a space. A metric g is a $(0, 2)$ tensor that is

- symmetric $g(X, Y) = g(Y, X)$
- non-degenerate: $g(X, Y)_p = 0, \quad \forall Y_p \in T_p(\mathcal{M}) \implies X_p = 0$

Notice that our condition for non-degeneracy is **not** the same as having a point where $g(X, X)_p = 0$ as we shall soon see. In a coordinate basis, $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$. The components are obtained by our standard way of subbing in basis vectors into our tensor.

$$g_{\mu\nu} = g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$$

We often write this as a line element which we call

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

This is something that perhaps we're more familiar with. Since our metric is non-degenerate, it is a theorem in linear algebra that we can diagonalise this thing, and furthermore we have no zero eigenvalues. If we diagonalise $g_{\mu\nu}$, it has positive and negative elements (non are zero). The number of negative elements is called the signature of the metric. There's a theorem in linear algebra (Sylvester's law of inertia) which says that the signature is invariant, which means that it makes sense to talk about signatures in a well defined sense.

5.1.1 Riemannian Manifolds

A Riemannian manifold is a manifold with metric with signature all positive. For example, Euclidean space in \mathbb{R}^n endowed with the usual Pythagorean metric.

$$g = dx^1 \otimes dx^1 + \cdots + dx^n \otimes dx^n$$

A metric gives us a way to measure the length of a vector $X \in \mathcal{X}(\mathcal{M})$. Since our signature is positive, we have that $g(X, X)$ is a positive number, and hence we can take a square root to define a norm.

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

We can also find the angle between vectors, where

$$g(X, Y) = |X||Y| \cos \theta$$

It also gives us a way to measure the distance between two points, p, q . Along the curve

$$\sigma : [a, b] \rightarrow \mathcal{M}, \quad \sigma(a) = p, \sigma(b) = q$$

our distance is given by the integral of the metric at that point where X is the tangent to the curve,

$$s = \int_a^b dt \sqrt{g(X, X)}|_{\sigma(t)}$$

where at each point X is our tangent to the curve. If our curve has the coordinates $x^\mu(t)$, then our distance is

$$s = \int_a^b dt \sqrt{g_{\mu\nu}(x) \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}}$$

Note that this notion of distance still makes sense since and is well defined since it's easy to check that s is invariant under re-parametrisations of our curve.

5.1.2 Riemannian Geometry

A Lorentzian manifold is a manifold equipped with a metric of signature $(- + + \dots)$. For example, Minkowski space is \mathbb{R}^n but our metric is

$$\eta = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + \dots + dx^{n-1} \otimes dx^{n-1}$$

with components

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

This is slightly different to our Riemannian manifold case since now we can have vectors with negative or zero length. We classify vectors $X_p \in T_p(\mathcal{M})$ as

$$g(X_p, X_p) = \begin{cases} < 0 & \text{timelike} \\ = 0 & \text{null} \\ > 0 & \text{spacelike} \end{cases}$$

At each point $p \in \mathcal{M}$, we draw null tangent vectors called lightcones, and as we'll soon see, this region outlines our area of possible causality.

A curve is called timelike if it's tangent vector at every point is timelike. We can see this in the figure where we have two light-cones for future and past time. In this case, we can measure the distance between two points.

$$\tau = \int_a^b dt \sqrt{-g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}}$$

This object τ is called the proper time between two points. Philosophically, this is a parameter which is invariant in all frames. If we were to reparametrise this curve, our definition of τ remains invariant.

5.1.3 The Joys of a metric

Claim. Metrics induce a natural isomorphism from vectors to 1-forms. The metric gives a natural (basis independent) isomorphism

$$g : T_p(\mathcal{M}) \rightarrow T_p^*(\mathcal{M})$$

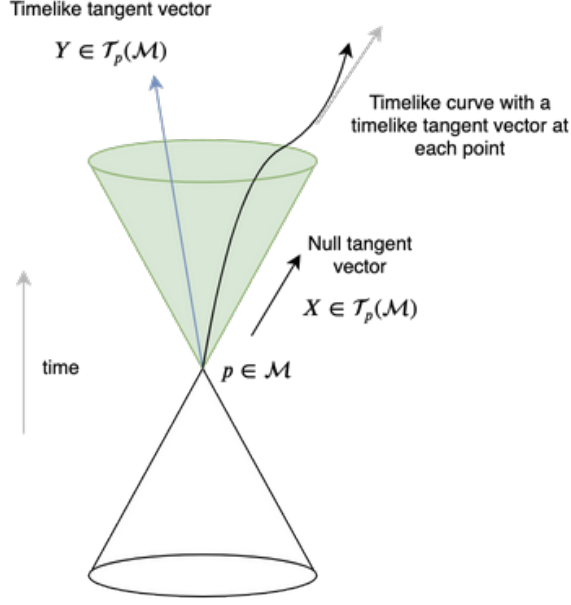


Figure 2: Here we show timelike vectors with negative norm!

Given $X \in \mathcal{X}(M)$, we can construct $g(X, \cdot) \in \Lambda^1(\mathcal{M})$. If $X = X^\mu \partial_\mu$, our corresponding one form is

$$g_{\mu\nu} X^\mu dx^\nu := X_\nu dx^\nu$$

In this formula, we've written the index on X downstairs! The metric provides a natural isomorphism between our vector space and our one-forms. Hence, this metric allows us to raise and lower indices, which means that our metric switches the mathematical space we are working in. Lowering an index is really the statement that there's a natural isomorphism. Because g is non-degenerate, there's an inverse

$$g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho$$

This defines a rank $(2, 0)$ tensor $\hat{g} = g^{\mu\nu} \partial_\mu \otimes \partial_\nu$, and we can use this to raise indices. We have

$$X^\mu = g^{\mu\nu} X_\nu$$

Claim. Metrics induce volume forms to integrate with. There's something else that the metric gives us. We also get a natural volume form. On a Riemannian manifold, our volume form is defined to be

$$v = \sqrt{\det g_{\mu\nu}} dx^1 \wedge \cdots \wedge dx^n$$

We write $g = \det g_{\mu\nu}$. On a Lorentzian manifold, $v = \sqrt{-g} dx^0 \wedge \cdots \wedge dx^{n-1}$. This is independent of coordinates. In new coordinates,

$$dx^\mu = A^\mu_\nu \frac{\partial x^\mu}{\partial \tilde{x}^\nu}$$

We see how they change.

$$\begin{aligned} dx^1 \wedge \cdots \wedge dx^n &= A^1_{\mu_1} \cdots A^n_{\mu_n} d\tilde{x}^{\mu_1} \wedge \cdots \wedge d\tilde{x}^{\mu_n} \\ &= \sum_{\text{perms } \pi} A^1_{\pi(1)} \cdots A^n_{\pi(n)} d\tilde{x}^1 \wedge \cdots \wedge d\tilde{x}^n \\ &= \det(A) d\tilde{x}^1 \wedge \cdots \wedge d\tilde{x}^n \end{aligned}$$

If we have that $\det A > 0$, coordinate change preserves orientation. Meanwhile,

$$\begin{aligned} g_{\mu\nu} &= \frac{\partial \tilde{x}^\rho}{\partial x^\mu} \frac{\partial \tilde{x}^\sigma}{\partial x^\nu} \tilde{g}_{\rho\sigma} \\ &= (A^{-1})^\rho_\mu (A^{-1})^\sigma_\nu \tilde{g}_{\rho\sigma} \end{aligned}$$

Hence,

$$\det g_{\mu\nu} = (\det A^{-1})^2 \det \tilde{g}_{\rho\sigma}$$

Thus we have that

$$v = \sqrt{|\tilde{g}|} d\tilde{x}^1 \wedge \cdots \wedge d\tilde{x}^n$$

in components, we have that

$$v = \frac{1}{n!} v_{\mu_1 \dots \mu_n} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_n}$$

where our components are given by

$$v_{\mu_1 \dots \mu_n} \epsilon_{\mu_1 \dots \mu_n}$$

we can integrate functions as

$$\int_{\mathcal{M}} f v = \int_{\mathcal{M}} d^n x \sqrt{|g|} f(x)$$

The metric provides a map from $\omega \in \Lambda^p(\mathcal{M})$ to $(*\omega) \in \Lambda^{n-p}(\mathcal{M})$ defined by

$$(*\omega)_{\mu_1 \dots \mu_{n-p}} = \frac{1}{p!} \sqrt{|g|} \epsilon_{\mu_1 \dots \mu_{n-p} \nu_1 \dots \nu_p} \omega^{\nu_1 \dots \nu_p}$$

This object is called Hodge dual. We can check that

$$*(*\omega) = \pm (-1)^{p(n-p)} \omega$$

with $+$ used in with a Riemannian metric, and $-$ used with a Lorentzian metric. We can then define an inner product on forms. Given $\omega, \eta \in \Lambda^p(\mathcal{M})$, let

$$\langle \eta, \omega \rangle = \int_{\mathcal{M}} \eta \wedge *\omega$$

The integrand is a top form so this is okay. This allows us to introduce a new object. If we have a p -form $\omega \in \Lambda^p(\mathcal{M})$, and a $p-1$ form $\alpha \in \Lambda^{p-1}(\mathcal{M})$, then

$$\langle d\alpha, \omega \rangle = \langle \alpha, d^\dagger \omega \rangle$$

when $d^\dagger : \Lambda^p(\mathcal{M}) \rightarrow \Lambda^{p-1}(\mathcal{M})$, is

$$d^\dagger = \pm (-1)^{np+n-1} * d *$$

where again our \pm signs depend on whether we have a Riemannian or Lorentzian metric. To show this, on a closed manifold Stokes' theorem implies that

$$0 = \int_{\mathcal{M}} d(\alpha \wedge *\omega) = \int_{\mathcal{M}} d\alpha \wedge *\omega + (-1)^{p-1} \alpha \wedge d*\omega$$

But the term on the right is just

$$= \langle d\alpha, \omega \rangle + (-1)^{p-1} \text{sign} \langle \alpha, *d*\omega \rangle$$

When we fix our sign, we get the result. There's actually a close relationship between forms in differential Geometry and fermionic antisymmetric fields in quantum field theory.

5.2 Connections and Curvature

What's the point of a connection? This is going to be our final way of differentiation as opposed to the things we have already written down. A connection is a map $\nabla : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \rightarrow \mathcal{X}(\mathcal{M})$

We write this as $\nabla(X, Y) = \nabla_X Y$. The purpose of doing this is to make it look more like differentiation. Here, we call ∇_X the covariant derivative, and it satisfies

- Linearity in the second argument $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$
- Linearity in the first argument $\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z \forall f, g \in C^\infty(\mathcal{M})$
- Leibniz $\nabla_X(fY) = f\nabla_X Y + (\nabla_X f)Y$, with f a function.
- In the above, we have that $\nabla_X f = X(f)$, agrees with usual differentiation.

Suppose we have a basis of vector fields $\{e_\mu\}$. We write

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

This expression is a vector field because we know the derivative spits out a vector field. We use the notation $\nabla_{e_\mu} = \nabla_\mu$, to make the connection look like a partial derivative. Then, applying the Leibniz rule we can do a derivative on a general vector to give

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

Because the vector sits out front we can write

$$\nabla_X Y = X^\nu \nabla_\nu Y$$

with

$$\nabla_\nu Y = (e_\nu(Y^\mu) + \Gamma_{\nu\rho}^\mu Y^\rho)e_\mu$$

or, we define we equivalently define

$$(\nabla_\nu Y)^\mu := \nabla_\nu Y^\mu = e_\nu(Y^\mu) + \Gamma_{\nu\rho}^\mu Y^\rho$$

Comparing this to the Lie derivative however, we have that \mathcal{L}_X depends on X and ∂X , so we can't write " $\mathcal{L}_X = X^\mu \mathcal{L}_\mu$ ". If we take a coordinate basis for the vector fields $\{e_\mu\} = \{\partial_\mu\}$, then we have that

$$\nabla_\nu Y^\mu = \partial_\nu Y^\mu + \Gamma_{\nu\rho}^\mu Y^\rho$$

In terms of notation, we can replace differentiation with punctuation. So, we have that

$$\nabla_\nu Y^\mu := Y_{;\nu}^\mu := Y_{,\nu}^\mu + \Gamma_{\nu\rho}^\mu Y^\rho$$

The connection is not a tensor! Consider a change of basis

$$\tilde{e}_\nu = A^\mu_{\ \nu} e_\mu, \text{ with } A^\mu_{\ \nu} = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}$$

We have that

$$\nabla_{\tilde{e}_\rho} \tilde{e}_\nu \tilde{\Gamma}_{\rho\nu}^\mu \tilde{e}_\mu = \nabla_{A^\sigma_{\ \rho} e_\sigma} (A^\lambda_{\ \nu} e_\lambda) = A^\sigma_{\ \rho} \nabla_\sigma (A^\lambda_{\ \nu} e_\lambda)$$

This simplifies further to give

$$\begin{aligned} &= A^\sigma_\rho (A^\lambda_\nu \Gamma^\tau_{\sigma\lambda} e_\tau + e_\lambda \partial_\sigma A^\lambda_\nu) \\ &= A^\sigma_\rho (A^\lambda_\nu \Gamma^\tau_{\sigma\lambda} + \partial_\sigma A^\tau_\lambda) e_\tau, \quad e_\tau = (A^{-1})^\mu_\tau \tilde{e}_\mu \end{aligned}$$

This implies that

$$\tilde{\Gamma}^\mu_{\rho\nu} = (A^{-1})^\mu_\tau A^\sigma_\rho A^\lambda_\nu \Gamma^\tau_{\sigma\lambda} + (A^{-1})^\mu_\tau A^\sigma_\rho \partial_\sigma A^\tau_\nu$$

So we have that an extra term is added on. We can also use the connection to differentiate other tensors. We simply ask that it obeys the Leibniz rule. For example, $\omega \in \Lambda^1(\mathcal{M}), Y \in \mathcal{X}(\mathcal{M})$

$$X(\omega(Y)) = \nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y)$$

Thus, rearranging the terms we have that

$$\nabla_X \omega(Y) = X(\omega(Y)) - \omega(\nabla_X Y)$$

So, in terms of coordinates,

$$X^\mu (\nabla_\mu \omega_\nu) Y^\nu = X^\mu \partial_\mu (\omega_\nu Y^\nu) - \omega_\nu X^\mu (\partial_\mu Y^\nu + \Gamma^\nu_{\mu\rho} Y^\rho) = X^\mu (\partial_\mu \omega_\rho - \Gamma^\nu_{\mu\rho} \omega_\nu) Y^\rho$$

Hence, we have that

$$\nabla_\mu \omega_\rho = \partial_\mu \omega_\rho - \Gamma^\nu_{\mu\rho} \omega_\nu$$

Given a connection, we can construct two tensors.

1. Torsion is a rank $(1, 2)$ tensor

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

where we have that $\omega \in \Lambda^1(\mathcal{M})$, and $X, Y \in \mathcal{X}(\mathcal{M})$. We can also think of T as a map from $\mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \rightarrow \mathcal{X}(\mathcal{M})$, with

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

2. Our second quantity that we can create is called **curvature**. This is a rank $(1, 3)$ tensor

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

This is called the Riemann tensor. We can also think of it as a map from $\mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M})$ to a differential operator which acts on $\mathcal{X}(\mathcal{M})$.

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

To show that these objects are tensors, we just need to check linearity in all the arguments. For example,

$$\begin{aligned} T(\omega, fX, Y) &= \omega(\nabla_{fX} Y - \nabla_Y(fX) - [fX, Y]) \\ &= \omega(f\nabla_X Y - f\nabla_Y X - Y(f)X - (f[X, Y] - Y(f)X)) \\ &= f\omega(\nabla_X Y - \nabla_Y X - [X, Y]) \\ &= fT(\omega, X, Y) \end{aligned}$$

Linearity is inherited from the fact that our covariant derivative is linear when you add. In a coordinate basis $\{e_\mu\} = \{\partial_\mu\}$, and our dual basis of one-forms $\{f^\mu\} = \{dx^\mu\}$, we have that the torsion in our components is

$$\begin{aligned} T_{\mu\nu}^\rho &= T(f^\rho, e_\mu, e_\nu) \\ &= f^\rho(\nabla_\mu e_\nu - \nabla_\nu e_\mu - [e_\mu, e_\nu]) \\ &= \Gamma_{\mu\nu}^\rho - \Gamma_{\nu\mu}^\rho \end{aligned}$$

A connection with $\Gamma_{\mu\nu}^\rho = \Gamma_{\nu\mu}^\rho$ has $T_{\mu\nu}^\rho = 0$ and is said to be torsion free. In addition, our curvature tensor has components

$$\begin{aligned} R_{\rho\mu\nu}^\sigma &= R(f^\sigma; e_\mu, e_\nu, e_\rho) \\ &= f^\sigma(\nabla_\mu \nabla_\nu e_\rho - \nabla_\nu \nabla_\mu e_\rho - \nabla_{[e_\mu, e_\nu]} e_\rho) \\ &= f^\sigma(\nabla_\mu(\Gamma_{\nu\rho}^\lambda e_\lambda) - \nabla_\nu(\Gamma_{\mu\rho}^\lambda e_\lambda)) \\ &= \partial_\mu \Gamma_{\nu\rho}^\sigma - \partial_\nu \Gamma_{\mu\rho}^\sigma + \Gamma_{\nu\rho}^\lambda \Gamma_{\mu\lambda}^\sigma - \Gamma_{\mu\rho}^\lambda \Gamma_{\nu\lambda}^\sigma \end{aligned}$$

Clearly, we have an antisymmetry property here, where

$$R_{\rho\mu\nu}^\sigma = -R_{\rho\nu\mu}^\sigma$$

5.3 The Levi-Civita Connection

The fundamental theorem of Riemannian Geometry is that there exists a unique, torsion-free connection with the property obeying

$$\nabla_X g = 0, \forall X \in \mathcal{X}(\mathcal{M})$$

To prove this, suppose that this object exists. Then,

$$\begin{aligned} X(g(Y, Z)) &= \nabla_X[g(Y, Z)] \\ &= \nabla_X g(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \end{aligned}$$

The fact that our torsion vanishes implies that

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

Hence, our equation on the left hand of our blackboard reads

$$X(g(Y, Z)) = g(\nabla_Y X, Z) + g(\nabla_X Z, Y) + g([X, Y], Z)$$

Now we cycle X, Y, Z , where we find that

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

If add the first two equations and then subtract by the third one we get that

$$g(\nabla_Y X, Z) = \frac{1}{2} [Xg(Y, Z) + Yg(X, Z) - Zg(X, Y) - g([X, Y], Z) - g([Y, Z], X) + g([Z, X], Y)]$$

In a coordinate basis, we have that $\{e_\mu\} = \{\partial_\mu\}$, we have that

$$g(\nabla_\nu e_\mu, e_\rho) = \Gamma_{\nu\mu}^\lambda g_{\lambda\rho} = \frac{1}{2} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

Where we have that

$$\Gamma_{\mu\nu}^\lambda = \frac{1}{2} g^{\lambda\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

This is the Levi-Civita connection, and the $\Gamma_{\mu\nu}^\lambda$ are called the Christoffel symbols. We still need to show that it transforms as a connection, which we leave as an exercise.

5.3.1 The Divergence Theorem

Consider a manifold \mathcal{M} with metric g , with boundary $\partial\mathcal{M}$, and let n^μ be an outward pointing vector orthogonal to $\partial\mathcal{M}$. Then, for any X^μ , our claim is that

$$\int_{\mathcal{M}} d^n x \sqrt{g} \nabla_\mu X^\mu = \int_{\partial\mathcal{M}} d^{n-1} x \sqrt{\gamma} n_\mu X^\mu$$

On a Lorentzian manifold, this also holds with $\sqrt{g} \rightarrow \sqrt{-g}$ and this also holds provided $\partial\mathcal{M}$ is purely timelike or purely spacelike.

First, we need a lemma, that $\Gamma_{\mu\nu}^\mu = \frac{1}{\sqrt{g}} \partial_\nu \sqrt{g}$. To prove this, we have that, writing out the definitions, that

$$\Gamma_{\mu\nu}^\mu = \frac{1}{2} g^{\mu\rho} \partial_\nu g_{\mu\rho} = \frac{1}{2} \text{tr}(\hat{g}^{-1} \partial_\nu \hat{g})$$

But from this we have that

$$\begin{aligned} \dots &= \frac{1}{2} \text{tr}(\partial_\nu \log \hat{g}) \\ &= \frac{1}{2} \partial_\nu \log \det \hat{g} \\ &= \frac{1}{2} \frac{1}{\det \hat{g}} \partial_\nu \det \hat{g} \\ &= \frac{1}{\sqrt{g}} \partial_\nu \sqrt{g} \end{aligned}$$

5.4 Parallel Transport

Now that we have a connection ∇ , we will show that we can use this object to make vectors 'travel' along the surface of a manifold. This means that our connection has given us a map from the vector space at a point $p \in \mathcal{M}$, to another point $q \in \mathcal{M}$. To set up this map, we take our favourite vector field X and then generate integral curves. At a specific point $p \in \mathcal{M}$, this specifies a unique curve \mathcal{C} . This curve is given by

$$X^\mu|_{\mathcal{C}} = \frac{dx^\mu(\tau)}{d\tau}$$

A vector field $\mathcal{X}(\mathcal{M})$ is said to be **parallel transported** along this curve if at every point, we have that the covariant derivative of Y with respect to X is zero.

$$\nabla_X Y = 0, \quad \text{at all } p \in \mathcal{C}$$

From this object we have an equation which specifies the components Y^μ at each point on the curve. Our initial condition here is what $Y^\mu(\tau = 0)$ is. The above condition reads

$$X^\mu (\partial_\mu Y^\nu + \Gamma_{\mu\rho}^\nu Y^\rho) = \frac{dY^\nu(\tau)}{d\tau} + X^\mu Y^\rho \Gamma_{\mu\rho}^\nu = 0$$

5.4.1 Geodesics

From our idea of parallel transport which we introduced above, we shall talk about a very special type of curve called a **geodesic**. Suppose we have a curve C . Now, this curve C has a tangent vector $X(\tau)$ at every point, where τ is some parameter we have used to construct the curve. If the tangent vector X is parallel transported along this curve, that is, we have that

$$\nabla_X X = 0, \quad \text{for all } p \in C$$

then, we call this object a geodesic. Our equation for our tangent vector component X^μ is

$$\frac{dX^\nu}{d\tau} + X^\mu X^\rho \Gamma_{\mu\rho}^\nu = 0$$

However, we can go a bit further here since we're on an integral curve of X^μ .

$$X^\mu = \frac{dx^\mu(\tau)}{d\tau} \implies \frac{d^2 x^\mu(\tau)}{d\tau^2} + \frac{dx^\nu}{dt} \frac{dx^\rho}{d\tau} \Gamma_{\nu\rho}^\mu = 0$$

This specifies a unique geodesic given a set of initial conditions on $x^\mu(0)$ and $\dot{x}^\mu(0)$, and is determined from our connection on our manifold.

5.5 Normal coordinates

Working with metrics is difficult, because they can be verbose and involved. What if, there was some way to change our basis such that we can work with a simple (diagonal) and flat metric at a given point. Lucky for us, there is a way! This is called switching to normal coordinates. In this bit we'll show that at a point $p \in \mathcal{M}$, we can always choose coordinates $\{x^\mu\}$ such that

$$g_{\mu\nu} = \delta_{\mu\nu}, \quad g_{\mu\nu,\rho} = 0$$

In this context our tensor δ can be either the standard Euclidean or Minkowski metric depending on whether we're on a Riemannian or Lorentzian manifold.

To show this, we count degrees of freedom (which in my opinion I think is a bit of a weird technique). Suppose we're in a coordinate system $\{\tilde{x}^\mu\}$ at a point p . We can, without loss of generality, set $\tilde{x}^\mu(p) = 0$. Our aim then is to find a change of basis which yields a new coordinate system $\{x^\mu\}$ (without loss of generality we also have that $x^\mu(p) = 0$, such that in this basis, the above conditions on the metric are satisfied).

This means that our change of basis satisfies

$$g_{\mu\nu} = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \tilde{g}_{\alpha\beta}$$

Now, we employ a trick. We Taylor expand \tilde{x}^α in terms of x^μ about the point p , which gives us the expansion

$$\tilde{x}^\mu = \left. \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \right|_{x=0} x^\nu + \frac{x^\rho x^\nu}{2} \left. \frac{\partial^2 \tilde{x}^\mu}{\partial x^\rho \partial x^\nu} \right|_{x=0}$$

When we differentiate this Taylor expansion we get that,

$$\frac{\partial \tilde{x}^\mu}{\partial x^\nu} = \left. \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \right|_{x=0}$$

Now, when we substitute this into our change of basis formula and evaluate at the point $p = 0$, we find that we need to solve the system

$$\delta_{\mu\nu} = \left(\left. \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \right|_{x=0} \left. \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \right|_{x=0} \tilde{g}_{\alpha\beta}(p) \right)$$

This equation looks like

$$I = AGA^T$$

where A represents our change of basis, and G our metric. From our change of basis, we are free to choose n^2 different components for each element. Now, since this equation is symmetric upon taking a transpose, we only have $\frac{1}{2}n(n+1)$ constraints. Thus, it's possible to choose a change of basis such that this works. Now, to second order, we have to expand further. We find that

$$g_{\mu\nu} = x^\gamma \left(\left. \frac{\partial^2 \tilde{x}^\alpha}{\partial x^\mu \partial x^\gamma} \right|_{x=0} \left. \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \right|_{x=0} + \left. \frac{\partial^2 \tilde{x}^\beta}{\partial x^\gamma \partial x^\nu} \right|_{x=0} \left. \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} \right|_{x=0} \tilde{g}_{\alpha\beta} \right)$$

Our condition for our metric to be flat means we need to differentiate this to get

$$g_{\mu\nu,\rho} = \left(\left. \frac{\partial^2 \tilde{x}^\alpha}{\partial x^\mu \partial x^\rho} \right|_{x=0} \left. \frac{\partial \tilde{x}^\beta}{\partial x^\nu} \right|_{x=0} + \left. \frac{\partial^2 \tilde{x}^\beta}{\partial x^\mu \partial x^\rho} \right|_{x=0} \left. \frac{\partial \tilde{x}^\alpha}{\partial x^\nu} \right|_{x=0} \right)$$

From this equation, we need to make sure we have enough components to specify our change of basis. Now, let's do some counting of our available degrees of freedom. $g_{\mu\nu}$ is symmetric and hence has $\frac{1}{2}n(n-1)$ free components. The addition of our derivative gives us a further n free components for each element of g . Thus, we have $\frac{1}{2}n^2(n-1)$ components. Furthermore, our second derivative term $\partial^2 \tilde{x}^\alpha / \partial x^\mu \partial x^\rho$, due to symmetry of mixed partials, has $\frac{1}{2}n(n-1)$ for each index α . Thus, we have in total also $\frac{1}{2}n^2(n-1)$ free components to choose from. This is enough to specify our change of basis!

5.5.1 Constructing Normal Coordinates

This section is going to be a short one. We basically will pull a cat of the bag and construct a set of normal coordinates with a map. The tangent space associated at a point on a manifold has a canonical map to recover points on the manifold. Consider the exponential map

$$\text{Exp} : \mathcal{T}_p(\mathcal{M}) \rightarrow \mathcal{M}$$

This map is given by picking a vector, then following the geodesic curve there for one unit of parameter distance $\tau = 1$. This means we will arrive at some point q . Now, that this point q , we define our coordinates to be

$$x^\mu(q) = X^\mu$$

where X^μ was the vector we had originally chosen.

5.6 Path Deviations and Curvature

When we parallel transport a vector on a manifold along a curve, how the vector ends up depends on the specific path taken. Specifically, on an infinitesimal 'square' path, this change in the vector is given by the Riemann curvature tensor. To illustrate this, we start by generating two curves by some vector fields X, Y , chosen such that the vector fields commute, hence $[X, Y] = 0$. We then move an infinitesimal amount in these two directions in a rectangle like fashion.

Let's take the path via q to r . The first thing we will do is to Taylor expand Z from p to q using τ as a parameter. This is a technique we shall repeatedly use, so take note!

$$Z_q^\mu = Z_p^\mu + \delta\tau \left. \frac{dZ^\mu}{d\tau} \right|_p + \frac{1}{2}(\delta\tau)^2 \left. \frac{d^2 Z^\mu}{d\tau^2} \right|_p + \dots$$

Now, by assumption of how we set up the problem, we have that Z is parallel transported along this integral curve of X , so obeys this equation

$$\frac{dZ^\mu}{d\tau} = -Z^\rho X^\nu \Gamma_{\nu\rho}^\mu$$

If we evaluate this derivative at p , it vanishes since $\Gamma_{\nu\rho}^\mu(p) = 0$ since we're working in normal coordinates! If we differentiate the above term to find our second order term, then applying the product rule, we only have one term which survives when we evaluate at p since non-derivatives of Γ vanish.

$$\left. \frac{d^2 Z^\mu}{d\tau^2} \right|_p = -Z^\rho X^\nu \left. \frac{d\Gamma_{\rho\nu}^\mu}{d\tau} \right|_p = -Z^\rho X^\nu X^\sigma \Gamma_{\nu\rho,\sigma}^\mu$$

Going into the last equality, we used the chain rule. So, we now have an expression for the Taylor expansion

$$Z_q^\mu = Z_p^\mu - \frac{1}{2} (X^\nu X^\sigma Z^\rho \Gamma_{\rho\nu,\sigma}^\mu)_p (d\tau)^2 + \dots$$

The next step to do is to go from q to r , repeating the same process. We expand Z_r^μ about Z_q^μ . Since we are Taylor expanding along the integral curve Y which is generated by the parameter λ ,

we have that

$$\begin{aligned} Z_r^\mu &= Z_q^\mu + \delta\lambda \left. \frac{dz^\mu}{d\lambda} \right|_p + \frac{1}{2}(\delta\lambda)^2 \left. \frac{d^2 Z^\mu}{d\lambda^2} \right|_p \\ &= Z_q^\mu - \delta\lambda(Y^\nu Z^\rho \Gamma_{\rho\nu}^\mu)_q - \frac{1}{2}(\delta\lambda)^2(Y^\nu Y^\sigma Z^\rho \Gamma_{\rho\nu,\sigma}^\mu)_q + \dots \end{aligned}$$

Now don't despair. You can see we're in a bit of a bind here since we have to evaluate the connection Γ at q . Thus, we can't use normal coordinates to make this term disappear. We can however, do the next best thing and Taylor expand this term about p , along the path of the integral curve generated by X

$$(\Gamma_{\rho\beta}^\mu)_q = (\Gamma_{\rho\beta}^\mu)_p + \delta\tau \left. \frac{d\Gamma_{\rho\beta}^\mu}{d\tau} \right|_p + \dots = -(Y^\nu Z^\rho X^\sigma \Gamma_{\rho\nu,\sigma}^\mu)_p d\tau$$

Note that we used the chain rule to rewrite $d/d\tau$. Now, from normal coordinates the first term here vanishes. And, since we're substituting this into an expression which is first order in $\delta\lambda$ anyway, we need only keep the derivative term. By the same logic, expanding out our second order derivative term and then Taylor expanding Γ around p gives us

$$\left. \frac{d^2 Z}{d\lambda^2} \right|_q = -(Y^\nu Y^\sigma Z^\rho \Gamma_{\rho\nu,\sigma}^\mu)_p$$

The upshot of this is that, we can now write the whole thing, from p to q , as

$$Z_r^\mu = Z_p^\mu - \frac{1}{2}(\Gamma_{\rho\nu,\sigma}^\mu)_p [X^\nu X^\sigma Z^\rho d\tau^2 + 2Y^\nu Z^\rho X^\sigma d\tau d\lambda + Y^\nu Y^\sigma Z^\rho d\lambda^2]$$

Now, if we were to go the other way and go through s , we get exactly the same result but with X, Y switched to reflect the fact that we're changing the order of vectors we're travelling with, and also a switch in λ and τ to reflect the change in the order of parametrisation. This means that only the mixed term is changed.

$$Z_r^\mu = Z_p^\mu - \frac{1}{2}(\Gamma_{\rho\nu,\sigma}^\mu)_p [X^\nu X^\sigma Z^\rho d\tau^2 + 2X^\nu Z^\rho Y^\sigma d\tau d\lambda + Y^\nu Y^\sigma Z^\rho d\lambda^2]$$

Now, subtracting one from the other gives us our Riemann curvature tensor in normal coordinates, and hence since both sides are tensors, this holds in all coordinate systems.

$$\begin{aligned} \Delta Z_r^\mu &= Z_r^\mu - (Z'_r)^\mu \\ &= (\Gamma_{\rho\nu,\sigma}^\mu - \Gamma_{\rho\sigma,\nu}^\mu) (Y^\nu Z^\rho X^\sigma)_p d\lambda d\tau \\ &= (R^\mu_{\rho\nu\sigma} Y^\nu Z^\rho X^\sigma)_p d\lambda d\tau \end{aligned}$$

Going into the last line, we've once again used the fact that since our expression is a tensor in normal coordinates, this holds in all coordinate systems.

5.7 Derived tensors from the Riemann Tensor

From our Riemann tensor, we can easily build new tensors by contracting over some indices. In this bit, we'll define some really important tensors given by our curvature.

Definition. (Ricci Tensor) Our Ricci tensor is given by contracting our first index with our third component of the Riemann tensor.

$$R_{\mu\nu} = R^\rho_{\mu\rho\nu}$$

This obeys the symmetry property that $R_{\mu\nu} = R_{\nu\mu}$.

Definition. (Ricci Scalar) From this object, we can then get our Ricci scalar which is given by

$$R = g^{\mu\nu} R_{\mu\nu}$$

Applying our Bianchi identity to the Riemann tensor, we have that

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2} \nabla_\nu R$$

This can be shown by considering the Bianchi identity

$$\nabla_{[\rho} R_{\alpha\beta]\gamma\delta} = 0, \implies \nabla_\rho R_{\alpha\beta\gamma\delta} + \nabla_\alpha R_{\beta\rho\gamma\delta} + \nabla_\beta R_{\rho\alpha\gamma\delta} = 0$$

Now, raising and lowering indices, and making use of the symmetries of the Riemann tensor should make the above result pop out. In this spirit, we define the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$$

This obeys the rule that the covariant derivative is

$$\nabla^\mu G_{\mu\nu} = 0$$

5.8 Connection 1-forms

We now will introduce a handy technology to compute our Riemann tensor, which will make it easier to do instead of computing all the Christoffel components. If we had a basis $\{e_\mu\} = \{\partial_\mu\}$, we can always introduce a different basis which is a linear sum of the coordinate induced basis. We call this basis

$$\hat{e}_a = e_a^\mu \partial_\mu$$

The upshot of this is that on a Riemannian manifold, we can pick a basis such that

$$g(\hat{e}_a, \hat{e}_b) = g_{\mu\nu} e_a^\mu e_b^\nu = \delta_{ab}$$

We call the components e_a^μ are called vielbeins. We can raise and lower indices using $g_{\mu\nu}$ and a, b using δ_{ab} . The basis of one forms $\{\hat{\theta}^a\}$ obey our standard dual basis relation

$$\hat{\theta}^a(\hat{e}_b) = \delta^a_b$$

They are $\hat{\theta}^a = e^a_\mu dx^\mu$, with $e^a_\mu e_b^\mu = \delta^a_b$. This property satisfies $e^a_\mu e_a^\nu = \delta_\mu^\nu$. Our metric is

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu = \delta_{ab} \hat{\theta}^a \otimes \hat{\theta}^b$$

This means that

$$g_{\mu\nu} = e^a{}_\mu e^b{}_\nu \delta_{ab}$$

An example would be to consider the metric (Schwarzschild metric)

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2)$$

This is $ds^2 = \eta^{ab} \hat{\theta}^a \otimes \hat{\theta}^b$, with non-coordinate 1 forms

$$\hat{\theta}^0 = f dt, \quad \hat{\theta}^1 = f^{-1} dr, \quad \hat{\theta}^2 = r d\theta, \quad \hat{\theta}^3 = r \sin \theta d\phi$$

In the basis $\{\hat{e}_a\}$, the components of our connection are defined as (not to be confused with our usual connection) that

$$\nabla_{\hat{e}_c} \hat{e}_b := \Gamma_{bc}^a \hat{e}_a$$

We define the connection 1-form to be $\omega^a{}_b = \Gamma_{bc}^a \hat{\theta}^c$, this is called the spin connection.

We have the first Cartan structure equation.

$$d\hat{\theta}^a + \omega^a{}_b \wedge \hat{\theta}^b = 0$$

In addition, we have for the Levi-Civita connection, that our Cartan structure components $\omega_{ab} = -\omega_{ba}$. In the Vielbein basis, we set

$$R^a{}_{bcd} = R(\hat{\theta}^a; \hat{e}_c, \hat{e}_d, \hat{e}_b)$$

with $R^a{}_{bcd} = -R^a{}_{bdc}$. We define the curvature 2 form to be

$$R^a{}_b = \frac{1}{2} R^a{}_{bcd} \hat{\theta}^c \wedge \hat{\theta}^d$$

This gives our second structure equation.

$$R^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b$$

We should think about what information is compressed and what is not.

Let's go back to our example. We compute $d\hat{\theta}^a$, which gives us

$$\begin{aligned} d\hat{\theta}^0 &= f' dr \wedge dt \\ d\hat{\theta}^1 &= f' dr \wedge dr = 0 \\ d\hat{\theta}^2 &= dr \wedge d\theta \\ d\hat{\theta}^3 &= \sin \theta dr \wedge d\phi - r \cos \theta d\theta \wedge d\phi \end{aligned}$$

Now, we use our Cartan structure equation to get

$$d\hat{\theta}^0 = 0 \omega^0{}_b \wedge \hat{\theta}^b \implies \omega^0{}_1 = f' f dt = f' \theta^0$$

But, we use our anti-symmetry and the Minkowski metric to get that

$$w^0{}_1 = -\omega_{01} = \omega_{10} = \omega^1{}_0$$

To check the consistency of this system, we check that our equation for

$$d\hat{\theta}^1 = 0$$

Proceeding like this, we find that

$$\begin{aligned}\omega^0_1 &= \omega^1_0 = f' \hat{\theta}^0 \\ \omega^2_1 &= -\omega^1_2 = \frac{f}{r} \hat{\theta}^2 \\ \omega^3_1 &= -\omega^1_3 = \frac{f}{r} \hat{\theta}^3 \\ \omega^3_2 &= -\omega^2_3 = \frac{\cot \theta}{r} \hat{\theta}^3\end{aligned}$$

Now the curvature tensor

$$\begin{aligned}\mathcal{R}^0_1 &= d\omega^0_1 + \omega^0_c \wedge \omega^c_1 \\ &= f' d\hat{\theta}^0 + f'' dr \wedge \hat{\theta}^0 \\ &= ((f')^2 + f'' f) dr \wedge dt \\ &= -((f')^2 + f'' f) \hat{\theta}^0 \wedge \hat{\theta}^1 \\ &\implies R_{0101} = (f f'' + f'' f)\end{aligned}$$

We can convert back by using

$$R_{\mu\nu\rho\sigma} = e^a_\mu e^\beta_\nu e^c_\rho e^d_\sigma R_{abcd}$$

This means that $R_{trtr} = f f'' + (f')^2$.

6 The Einstein Equations

Space time is a manifold \mathcal{M} equipped with a Lorentzian metric g . We view our metric as some matrix which varies from point to point on the manifold. The right way to write things down here is to write down an action principle, then vary the actions to get our equations of motion.

6.0.1 The Einstein-Hilbert Action

The dynamics is governed by the Einstein-Hilbert action. There's not very much we can write here. The only object we have to play with is our metric g , and from this we have our natural volume form we can use. The only one we have is the square root of our determinant of our metric. Our action must precede

$$S = \int d^4 \sqrt{-g}$$

Now, what scalar function can we put here? The only thing we can do is to pull out our Ricci scalar, so that

$$S = \int d^4 x \sqrt{-g} R$$

Note that schematically, $R \sim \partial\Gamma + \Gamma\Gamma$, and $\Gamma \sim g^{-1}\partial g$. Thus, R is a function of two derivatives of our metric. Note, this means that the only connection we could cook up is the Levi-Civita connection. This is the simplest thing we can do! Now, to derive the equations of motion, we need to vary this field. We take the metric and push it with a small change

$$g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$$

Now, from this we can also deduce how the inverse metric changes with this infinitesimal change.

$$g_{\rho\mu} g^{\mu\nu} = \delta^\nu_\rho \implies \delta g_{\rho\mu} g^{\mu\nu} + g_{\rho\mu} \delta g^{\mu\nu} = 0$$

This means we can read off our change $\delta g^{\mu\nu}$ as

$$\delta g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\rho\sigma}$$

When we vary the whole thing, the trick is to write out the Ricci scalar in terms of the Ricci tensor contracted with the metric. Then, we apply the product rule when looking at a small variation. When we do the small variation, we have to remember to hit the volume form as well.

$$\delta S = \int d^4 x [\delta(\sqrt{-g}) g^{\mu\nu} R_{\mu\nu} + \sqrt{-g} (\delta g^{\mu\nu} R_{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu})]$$

We need to figure out what $\delta\sqrt{-g}$ is. To do this, we prove a claim that

Claim. We claim that

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu}$$

Proof. We use the fact that $\log \det A = \text{tr} \log A$, which implies that

$$\frac{1}{\det A} \delta(\det A) = \text{tr} = \text{tr}(\delta \log A) = \text{tr}(A^{-1} \delta A) = \text{tr}(A^{-1} \delta A)$$

Now, substituting in our metric g as the matrix A in our identity above, this finally implies that

$$\begin{aligned}\delta\sqrt{-g} &= \frac{1}{2} \frac{1}{\sqrt{-g}} (-g) g^{\mu\nu} \delta g_{\mu\nu} \\ &= \frac{1}{2} \sqrt{-g} g^{\mu\nu} \delta g_{\mu\nu}\end{aligned}$$

□

Now, we compute the variation of our Ricci tensor.

Claim. We have that our variation is

$$\delta R_{\mu\nu} = \nabla_\rho \delta \Gamma_{\mu\nu}^\rho - \nabla_\nu \delta \Gamma_{\mu\rho}^\rho$$

with

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

Proof. First note that importantly, $\Gamma_{\rho\nu}^\mu$ is not a tensor. However, the difference $\delta \Gamma_{\rho\nu}^\mu$ is a tensor since it's the difference between two connections. This means that it is a well defined action to take the covariant derivative of the objects. In normal coordinates, at some point

$$\begin{aligned}\delta \Gamma_{\mu\nu}^\rho &= \frac{1}{2} g^{\rho\sigma} (\partial_\mu \delta g_{\sigma\nu} + \partial_\nu \delta g_{\sigma\mu} - \partial_\sigma \delta g_{\mu\nu}) \\ &= \frac{1}{2} g^{\rho\sigma} (\nabla_\mu \delta g_{\sigma\nu} + \nabla_\nu \delta g_{\sigma\mu} - \nabla_\sigma \delta g_{\mu\nu})\end{aligned}$$

Since we're in normal coordinates, we can now replace our partial derivatives ∂ with covariant derivatives ∇ . This gives us a tensorial relation, and hence this holds true in all coordinate frames. We can play the same game with our Riemann tensor,

$$R_{\rho\mu\nu}^\sigma = \partial_\mu \Gamma_{\nu\rho}^\sigma - \partial_\nu \Gamma_{\mu\rho}^\sigma$$

This implies that our small change is given by

$$\begin{aligned}\delta R_{\rho\mu\nu}^\sigma &= \partial_\mu \delta \Gamma_{\nu\rho}^\sigma - \partial_\nu \delta \Gamma_{\mu\rho}^\sigma \\ &= \nabla_\mu \delta \Gamma_{\nu\rho}^\sigma - \nabla_\nu \delta \Gamma_{\mu\rho}^\sigma\end{aligned}$$

Since $\Gamma = 0$ in normal coordinates.

□

Theorem. (Vacuum Einstein Equations)

Now we have that, after all the dust has settled, that

$$\delta S = \int d^4x \sqrt{-g} \left[R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right] \delta g^{\mu\nu} + \nabla_\mu X^\mu$$

Where we've written

$$X^\mu = g^{\rho\nu} \delta \Gamma_{\rho\nu}^\mu - g^{\mu\nu} \delta \Gamma_{\nu\rho}^\rho$$

When we impose the condition of stationarity for some arbitrary change in the metric, we then have a condition for our integrand. This yields the vacuum Einstein equations.

$$\delta S = 0, \forall g^{\mu\nu} \implies G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0$$

Multiplying this quantity by $g^{\mu\nu}$, we have that $R = 0$, which implies the vacuum Einstein equations

$$R_{\mu\nu} = 0$$

6.0.2 Dimensional analysis

We can use dimensional analysis to fill out the constants here. Our action S has dimension ML^2T^{-1} , Our metric is dimensionless, and R has units of L^{-2} . This implies that our full action is

$$S = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} R$$

Our Planck mass is $\mathcal{M}_{pl}^2 = \frac{\hbar c}{8\pi G}$ and this is approximately $\mathcal{M}_{pl} \propto 10^{18} GeV$. Work with units with $c = 1$, and $\hbar = 1$. Throughout the next of the section, we shall work in terms of this Planck mass.

6.0.3 The Cosmological Constant

We could add a further term to the action

$$S = \frac{1}{2} \mathcal{M}_{pl} \int d^4x \sqrt{-g} (R - 2\Lambda)$$

We will study these. We have the equation

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -\lambda g_{\mu\nu} \implies R = 4\Lambda \implies R_{\mu\nu} = \Lambda g_{\mu\nu}$$

6.1 Diffeomorphisms Revisited

Our metric has $\frac{1}{2}4 \cdot 5 = 10$ components. But two metrics related by $x^\mu \rightarrow \tilde{x}^\mu(x)$ are physically equivalent. This means that actually we only have $10 - 4 = 6$ degrees of freedom. The change of coordinates can be viewed as a diffeomorphism

$$\phi : \mathcal{M} \rightarrow \mathcal{M}$$

such 'diffeos' are the 'gauge symmetry' of general relativity. Consider diffeomorphisms which take the form

$$x^\mu \rightarrow \tilde{x}^\mu(x) = x^\mu + \delta x^\mu$$

Moreover, we can view this as a diffeomorphism generated by a vector field $X^\mu = \delta x^\mu$. The metric transforms as $g_{\mu\nu}(x) \rightarrow \tilde{g}_{\mu\nu}(\tilde{x})$ with $\tilde{g}_{\mu\nu}(\tilde{x}) = \frac{\partial \tilde{x}^\rho}{\partial x^\mu} \frac{\partial \tilde{x}^\sigma}{\partial x^\nu} g_{\rho\sigma}(\tilde{x})$. From our definition that these coordinates are generated by a vector field, this is equal to

$$\begin{aligned} \dots &= (\delta_\mu^\rho + \partial_\mu X^\rho)(\delta_\nu^\sigma + \partial_\nu X^\sigma)(g_{\rho\sigma}(x) + X^\lambda \partial_\lambda g_{\rho\sigma}(x)) \\ &= g_{\mu\nu}(x) + X^\lambda \partial_\lambda g_{\mu\nu}(x) + g_{\mu\rho}(x) \partial_\nu X^\rho + g_{\nu\rho}(x) \partial_\mu X^\rho \end{aligned}$$

This means that our infinitesimal change is

$$\begin{aligned} \delta g_{\mu\nu} &= X^\lambda \partial_\lambda g_{\mu\nu} + g_{\mu\rho} \partial_\nu X^\rho + g_{\nu\rho} \partial_\mu X^\rho \\ &= (\mathcal{L}_X g)_{\mu\nu} \end{aligned}$$

Alternatively, we can write this as

$$\begin{aligned}\delta g_{\mu\nu} &= \partial_\mu X_\nu + \partial_\nu X_\mu + X^\rho (\partial_\rho g_{\mu\nu} - \partial_\mu g_{\rho\nu} - \partial_\nu g_{\mu\rho}) \\ &= \partial_\mu X_\nu + \partial_\nu X_\mu + 2X^\rho g_{\rho\sigma} T_{\mu\nu}^\sigma \\ &= \nabla_\mu X_\nu + \nabla_\nu X_\mu\end{aligned}$$

Now, let's look back at the action

$$\sigma S = \int d^4x \sqrt{-g} G^{\mu\nu} \delta g_{\mu\nu}$$

If we restrict to these changes of coordinates, the above

$$\dots = 2 \int d^4x \sqrt{-g} G^{\mu\nu} \nabla_\mu X_\nu = 0 \forall X = -2 \int d^4x \sqrt{-g} (\nabla_\mu G^{\mu\nu}) X_\nu$$

this is because changing coordinates is a gauge symmetry. This means that our Einstein tensor obeys

$$\nabla_\mu G^{\mu\nu} = 0$$

This means that diffeomorphisms are a symmetry implies the Bianchi identity. So, not all of the $G^{\mu\nu}$ are independent. We have found four conditions, which means that the Einstein equations $G_{\mu\nu}$ is really only 6 equations, which is the right number of components to determine our metric $g_{\mu\nu}$.

6.2 Some simple solutions

Let's look at what happens when we have a vanishing cosmological constant with $\Lambda = 0$. We need to solve $R_{\mu\nu} = 0$. It's tempting to write $g_{\mu\nu} = 0$, but this isn't allowed because the metric requires an inverse! This makes gravity different from other field theories, we need to have the constraint that none of the eigenvalues are zero. This may suggest that the metric is not a fundamental field. There are lots of similarities to this in fluid mechanics. The simplest solution is Minkowski space time.

$$ds^2 = -dt^2 + d\vec{x}^2$$

6.2.1 De Sitter Spacetime

In the case, $\Lambda > 0$, we can look for solutions to $R_{\mu\nu} = \Lambda g_{\mu\nu}$ which have some spherical symmetry. We look for solutions of the form

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\psi^2)$$

We can compute our Ricci tensor to find

$$R_{tt} = -f^4 R_{rr} = f^3 \left(f'' + \frac{2f'}{r} + (f')^2 f \right), \quad R_{\psi\psi} = \sin^2 \theta R_{\theta\theta} = (1 - f^2 - 2f f' r) \sin^2 \theta$$

Now, with this ansatz we can solve for $f(r)$ by setting $R_{\mu\nu} = \Lambda g_{\mu\nu}$, and then solving the equation by substituting in specific indices. By setting $\mu, \nu = tt, rr$ we get that

$$f'' + \frac{2f'}{r} + \frac{(f')^2}{f} = -\frac{\Lambda}{f}$$

Similarly, by setting $\mu, \nu = \theta\theta, \psi\psi$, we have

$$1 - 2ff'r - f^2 = \Lambda r^2$$

This equation is solved by

$$f(r) = \sqrt{1 - \frac{r^2}{R^2}}, \quad R^2 = \frac{3}{\Lambda}$$

This means that our resulting, final metric that we get takes the form

$$ds^2 = - \left(1 - \frac{r^2}{R^2}\right)^2 dt^2 + \left(1 - \frac{r^2}{R^2}\right)^{-2} dr^2 + r^2 d\Omega_2^2$$

To save ink, we've written $r^2(d\theta^2 + \sin^2\theta d\phi^2)$ as $r^2 d\Omega_2^2$, which is the radius squared, multiplied by the familiar metric on a unit 2-sphere. In particular, if we have a set of three Cartesian coordinates such that

$$x^2 + y^2 + z^2 = r^2$$

then since this parametrises a 2-sphere of radius r , our resulting metric in spherical coordinates is

$$dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\Omega_2^2$$

This will save a lot of time later in our calculations which involve different ways to parametrise this metric.

This is de Sitter space time (or alternatively called the static patch of d S). Note that we have a valid range from $r \in [0, R]$, but the metric appears to be singular at $r = R$. Throughout this section, we will check whether this singularity comes from merely a poor choice of coordinates, or is a genuine physical space-time singularity. Recall, it's totally okay for this to be a coordinate singularity (in fact, as we shall see, it is) because when we define coordinates, we only define them in terms of coordinate patches. We can look at geodesics. We have the action, where σ represents proper time

$$S = \int d\sigma \left[-f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2) \right]$$

Here, there are two conserved quantities. Nothing depends on ϕ , so we have that $l = \frac{1}{2} \frac{\partial L}{\partial \dot{\phi}} = r^2 \sin^2\theta \dot{\phi}$. The other conserved quantity we get is the energy of the particle with

$$E = -\frac{1}{2} \frac{\partial L}{\partial \dot{t}} = f(r)^2 \dot{t}$$

For a massive particle, we also require that the trajectory is timelike. Since σ represents proper time, this means that the Lagrangian itself is equal to

$$-f^2 \dot{t}^2 + f^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2) = -1$$

Look for geodesics with $\theta = \frac{\pi}{2}$ and $\dot{\theta} = 0$. This gives

$$\dot{r}^2 + V_{eff}(r) = E^2$$

with

$$V_{\text{eff}}(r) = (1 + \frac{l^2}{r^2})(1 - \frac{r^2}{R^2})$$

For $l = 0$, we have that

$$r(\sigma) = R\sqrt{E^2 - 1} \sinh\left(\frac{\sigma}{R}\right)$$

This hits $r = R$ in finite σ . Meanwhile

$$\frac{dt}{d\sigma} = E \left(1 - \frac{r^2}{R^2}\right)^{-1}$$

Solutions to this have $t \rightarrow \infty$ as $r \rightarrow R$. To see this, suppose $r(\sigma_*) = R$ and expand $\sigma = \sigma_* - \epsilon$. We find that

$$\frac{dt}{d\epsilon} \simeq -\frac{\alpha}{\epsilon}$$

this of course means that $t \sim -\alpha \log(\frac{\epsilon}{R})$, and $t \rightarrow \infty$ as $\epsilon \rightarrow 0$. so, it takes a particle finite proper time σ but infinite coordinate time t .

In fact, the de Sitter space time can be embedded in 5 dimensional Minkowski space, $\mathbb{R}^{1,4}$. which is

$$ds^2 = -(dX^0)^2 + \sum_{i=1}^4 (dX^i)^2$$

Our surface which we are embedding is

$$-(X^0)^2 + \sum_{i=1}^4 (X^i)^2 = R^2$$

This can be viewed as a hyperbola in Minkowski space. The flat metric is inherited onto the surface. We will show that this is true. We let $r^2 = \sum_{i=1}^3 (X^i)^2$ and $X^0 = \sqrt{R^2 - r^2} \sinh(\frac{t}{R})$, and also that $X^4 = \sqrt{R^2 - r^2} \cosh(\frac{t}{R})$. This means that, we have

$$-(X^0)^2 + (X^4)^2 = R^2 - r^2$$

Now just substitute and plug those in. We can check that if you compute dX^0 and dX^4 , and plug into the metric in 5 dimensional Minkowski space, we'll recover the de Sitter metric. Note that we above is a really weird parametrisation! We only singled out X^0 and X^4 . Moreover, X^4 only runs from $-\infty$ to ∞ , but in our parametrisation we only have that $X^4 > 0$. These coordinates are not particularly symmetric, and moreover cover only $X^4 \geq 0$. A better choice of coordinates is

$$X^0 = R \sinh(\tau/R), \text{ and } X^i = \cosh(\tau/R) \cdot y^i$$

where we have a normalisation constraint on y , such that $\sum_{i=1}^4 (y^i)^2 = 1$. Another small calculation, we substitute this into our 5d Minkowski metric. We have that

$$ds^2 = -d\tau^2 + R^2 \cosh^2(\tau/R) d\Omega_3^2$$

where $d\Omega_3^2$ is a metric on S^3 . This is a hard calculation, but we've swept this under the rug and put these into the 3 sphere metric. We can see that these are a better description of de Sitter space time because they describe the whole space, and has no more coordinate singularities. These are global coordinates. This metric also solves the Einstein equations with the same cosmological constant since it's just a change of coordinates from our original solution. So now we have two metric which describes the same space. There's a natural cosmological interpretation here. We have a 3-sphere, which initially shrinks, then expands. This corresponds to a contracting / expanding universe.

6.2.2 Anti de-Sitter Space-time

The next solution we will look at is for $\Lambda < 0$. Again, looking for solutions

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2 d\Omega_2^2$$

We find that in this case,

$$f(r) = \sqrt{1 + \frac{r^2}{R^2}}, \quad R^2 = -\frac{3}{\Lambda}$$

This is called anti-de Sitter space-time (AdS). There is no coordinate singularity here! This time, massive geodesics obey the rule $\dot{r}^2 + V_{\text{eff}} = E^2$. This potential energy is what it was before but there's now a plus sign instead of a minus sign

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right) \left(1 + \frac{r^2}{R^2}\right)$$

What's happening is that this kind of looks like a gravitational potential well in which we're stuck in. In particular, massive particles seemed to be confined to the centre of Anti de-Sitter space. The origin is special here.

Now let's look at massless particles. Massless particles follow null geodesics. This means that in the derivation of our geodesics we had a constraint which was -1 , but we can change this to zero.

$$-f^2 \dot{t}^2 + f^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) = 0$$

This means that at the point $\theta = \frac{\pi}{2}, \dot{\theta} = 0$, we have that our potential obeys $\dot{r}^2 + V_{\text{eff}}(r) = E^2$, and our null geodesic obeys

$$V_{\text{null}} = \frac{l^2}{2r^2} \left(1 + \frac{r^2}{R^2}\right)$$

We introduce new coordinates, $r = R \sinh \rho$. Plugging this into our de-Sitter metric, the \sinh is chosen so the radial coordinate factors out. Thus, we get

$$ds^2 = -\cosh^2 \rho dt^2 + R^2 d\rho^2 + R^2 \sinh^2 \rho (d\theta^2 + \sin^2 \theta d\phi^2)$$

The null geodesic equation is

$$R\dot{r} = \pm \frac{E}{\cosh \rho} \implies R \sinh \rho = E(\sigma - \sigma_0)$$

Massless particles hit $\rho \rightarrow \infty$ as $\delta \rightarrow \infty$. However, $E = \cosh^2 \rho \dot{t} \implies R \sinh \rho = R \tan(t/R) = E(\sigma - \sigma_0)$. SO, $t \rightarrow \frac{\pi R}{2}$ as $\sigma \rightarrow \infty$. Massless particles reach infinity of AdS in finite coordinate time. AdS can be viewed as a hyperboloid in $\mathbb{R}^{2,3}$, as

$$-(X^0)^2 - (X^4)^2 + \sum_{i=1}^3 (X^i)^2 = R^2$$

Now, let

$$\begin{aligned} X^0 &= R \cosh \rho \sin \frac{t}{R} \\ X^4 &= R \cosh \rho \cos \frac{t}{R} \\ X^i &= R y^i \sinh \rho, \quad \sum (y^i)^2 = 1 \end{aligned}$$

and we hence recover the AdS metric. There is one last set of coordinates which we can examine.

$$X^i = \frac{\tilde{r}}{R} x^i, \quad i = 0, 1, 2$$

$$X^4 - X^3 = \tilde{r} X^4 + X^3 = \frac{R^2}{\tilde{r}} + \frac{\tilde{r}}{R^2} \eta_{ij} x^i x^j$$

The metric we get out of this is

$$ds^2 = R^2 \frac{d\tilde{r}^2}{\tilde{r}^2} + \frac{\tilde{r}^2}{R^2} \eta_{ij} dx^i dx^j$$

These don't cover all of AdS. This is called the Poincare patch.

6.3 Symmetries

Let's look at some symmetries of the metric. The first thing we'd like to do is explain what a symmetry actually is. Think about a 2 dimensional sphere. This has the symmetry group $SO(3)$, since we can rotate it in any axis. Now think of a rugby ball. This has $SO(2)$ symmetry since we can only rotate it about one axis. The correct way to think about these things is to consider flows. Consider a 1-parameter family of diffeomorphisms $\sigma_t : \mathcal{M} \rightarrow \mathcal{M}$. Recall that this is associated to a vector field

$$K^\mu = \frac{dx^\mu}{dt}$$

where $\frac{dx^\mu}{dt}$ are tangent to the flow lines. We're going to call this a symmetry if we start at any point, flow along, then our destination point looks the same.

This flow is an isometry if the metric looks the same at each point along the flow, in other words

$$\mathcal{L}_K g = 0 \iff \nabla_\mu K_\nu + \nabla_\nu K_\mu = 0$$

This equation is called the 'Killing equation' and any vector K which obeys this is called the Killing vector. This is the equation which we need to solve. These objects describe the symmetries of the metric.

Note, commuting the Lie derivatives is handy because

$$\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X = \mathcal{L}_{[X,Y]}$$

From this, we start to get a Lie algebra structure that emerges for the group of continuous symmetries of the metric, (which is exactly what we expect from continuous symmetries).

Example. Minkowski space With Minkowski space in our metric, the Killing equation implies

$$\partial_\mu k_\nu + \partial_\nu k_\mu = 0$$

In full generality, we have that

$$k_\mu = c_\mu + \omega_{\mu\nu} x^\nu, \omega_{\mu\nu} = -\omega_{\nu\mu}$$

We have that c_μ represent our translations, and $\omega_{\mu\nu}$ represent boosts or rotations depending on what indices we're choosing. We can define Killing vectors

$$P_\mu = \frac{\partial}{\partial x^\mu}, \text{ and } M_{\mu\nu} = \eta_{\mu\rho}x^\rho \frac{\partial}{\partial x^\nu} - \eta_{\nu\rho}x^\rho \frac{\partial}{\partial x^\mu}$$

Now, we find that $[P_\mu, P_\nu]$, and that

$$[M_{\mu\nu}, P_\sigma] = -\eta_{\mu\sigma}P_\nu + \eta_{\sigma\nu}P_\mu$$

in addition,

$$[M_{\mu\nu}, M_{\rho\sigma}] = \eta_{\mu\sigma}M_{\nu\rho} + \eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\rho}M_{\nu\sigma} - \eta_{\nu\sigma}M_{\mu\rho}$$

These are the commutation relations of the Poincare group.

6.3.1 More examples

The isometries of dS and AdS are inherited from the 5 dimensional embedding. de-Sitter space time has isometry group $SO(1,4)$, and Anti de-Sitter space has isometry group $SO(2,3)$. Both groups have dimension 10, same as $SO(5)$, and the same is the Poincare group as well. Minkowski and de-Sitter space are equally as symmetric. In 5d, the Killing vectors are

$$M_{AB} = \eta_{AC}X^C \frac{\partial}{\partial X^B} - \eta_{BC}X^C \frac{\partial}{\partial X^A}$$

In this case, A runs from $A = 0, 1, 2, 3, 4$, and we have that the metric are given by

$$\eta = (-, +, +, +, +)$$

$$\eta = (-, -, +, +, +)$$

The flows induced by M_{AB} map the embedding hyperboloid to itself. These are isometries of (A) dS. Let's look at an example in de Sitter space in static patch coordinates. If the metric $g_{\mu\nu}(x)$, does not depend on some coordinate y , then $K = \frac{\partial}{\partial y}$ is a Killing vector since $\mathcal{L}_{\partial_y}g = \frac{\partial g_{\mu\nu}}{\partial y} = 0$. So, for the static path, we expect $\frac{\partial}{\partial t}$ to be a Killing vector.

We had

$$X^0 = \sqrt{R^2 - r^2} \sinh\left(\frac{t}{R}\right)$$

$$X^4 = \sqrt{R^2 - r^2} \cosh\left(\frac{t}{R}\right)$$

Look at

$$\frac{\partial}{\partial t} = \frac{\partial X^A}{\partial t} \frac{\partial}{\partial X^A} = \frac{1}{R} \left(X^4 \frac{\partial}{\partial X^0} + X^0 \frac{\partial}{\partial X^4} \right)$$

It's interesting to note that timelike Killing vectors such that $g_{\mu\nu}k^\mu k^\nu < 0$ are used to define energy. Minkowski and AdS have such objects. de-Sitter space has such an object in the static patch, but not globally. For example,

$$K = X^4 \frac{\partial}{\partial X^0} + X^0 \frac{\partial}{\partial X^4}$$

Now, the first term increases X^0 then $X^4 > 0$, and decreases X^0 when $X^4 < 0$. The Killing vector is positive and timelike only in the static patch. Elsewhere, it is spacelike. Energy is a subtle concept in dS.

6.4 Conserved quantities

Consider a particle moving on a geodesic $x^\mu(\tau)$ in a spacetime with Killing vector K^μ . Then, we have that

$$Q = K_\mu \frac{dx^\mu}{d\tau} \text{ is conserved}$$

To see this, differentiate to find that

$$\begin{aligned} \frac{dQ}{d\tau} &= \partial_\nu k_\mu \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} + k_\mu \frac{d^2x^\mu}{d\tau^2} \\ &= \partial_\nu k_\mu \frac{dk^\nu}{d\tau} \frac{dx^\mu}{d\tau} - k_\mu \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} \\ &= \nabla_\nu k_\mu \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} \end{aligned}$$

We can also see this from the action

$$S = \int d\tau g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu$$

Consider $\delta x^\mu(\tau) = k^\mu(x)$. We have

$$\delta S = \int d\tau \partial_\rho g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} + 2g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dk^\nu}{d\tau}$$

Now, we use the fact that

$$\begin{aligned} g_{\mu\nu} \frac{dk^\nu}{d\tau} &= \frac{dk_\mu}{d\tau} - \frac{dg_{\mu\rho}}{d\tau} k^\rho \\ &= (\partial_\nu k_\mu - \partial_\nu g_{\mu\rho} k^\rho) \frac{dx^\nu}{d\tau} \end{aligned}$$

Substituting this into the action

$$\delta S = \int d\tau 2\nabla_\mu K_\nu \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}$$

This implies that $\delta S = 0$ iff $\nabla_{(\mu} K_{\nu)} = 0$, the Killing equation.

7 Example Sheet 1

To: João Melo. From: Afq Hatta

Questions 5, 7 (and the rest)

7.1 Question 1

If we're given components of a vector field and want to solve for its integral curve, then we need to solve the equation

$$\left. \frac{dx^\mu(t)}{dt} \right|_{\phi(p)} = X^\mu(x^\nu(t))|_{\phi(p)}$$

So for the first integral curve, we need to solve the system

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= -x \end{aligned}$$

This is made a lot easier by writing out the system in polar coordinates (which is indeed a different chart for the manifold \mathbb{R}^2 , and writing $(x, y) = (r \cos \theta, r \sin \theta)$ with the chain rule gives us

$$\begin{aligned} \dot{r} \cos \theta - r \dot{\theta} \sin \theta &= r \sin \theta \\ \dot{r} \sin \theta + r \dot{\theta} \cos \theta &= -r \cos \theta \end{aligned}$$

If we multiply the first equation by $\sin \theta$ and the second equation by $\cos \theta$, and then subtract the first equation from the second equation, we've eliminated the \dot{r} term. We're left with

$$\dot{\theta} = -1 \implies \theta = -t + C$$

for some constant C . Substituting in $\dot{\theta} = -1$ in our first equation gives the condition that $\dot{r} = 0 \implies r = R$ for R constant. Hence our integral curves are merely circles of arbitrary radius about the origin.

For our second vector field

$$X^\mu = (x - y, x + y)$$

we proceed exactly as before, with polar coordinates. One finds instead that $\dot{\theta} = 1$, and hence that $\dot{r} = r$. Thus, we have that

$$\theta = t + A, \quad r = B e^t$$

for arbitrary constants A, B . These curves are spirals.

7.2 Question 2

We're given that the map $\hat{H} : T_p(M) \rightarrow T_p^*(M)$ is a linear map. So, since \hat{H} is linear,

$$\begin{aligned} H(X, \alpha Y + \beta Z) &= \hat{H}(\alpha Y + \beta Z)(X) \\ &= (\alpha \hat{H}(Y) + \beta \hat{H}(Z))(X) \\ &= \alpha \hat{H}(Y)(X) + \beta \hat{H}(Z)(X) \\ &= \alpha H(X, Y) + \beta H(X, Z) \end{aligned}$$

Thus, H is linear in the second argument. The only fact that we've used here is that \hat{H} is a linear map. For the linearity in the first argument, we use the fact that $\hat{H}(Y) \in T_p^*(M)$, which means that it's a linear map. So

$$H(\alpha X + \beta Z, Y) = \hat{H}(Y)(\alpha X + \beta Z) = \alpha \hat{H}(Y)(X) + \beta \hat{H}(Y)(Z) = \alpha H(X, Y) + \beta H(Z, Y)$$

Thus our map is linear in the first argument. Note that

$$H : T_p(M) \times T_p(M) \rightarrow \mathbb{R}$$

and since the map is multilinear, we have a rank $(0, 2)$ tensor.

Similarly, if we had a linear map

$$\hat{G} : T_p(M) \rightarrow T_p(M)$$

we could then define a new map

$$G : T_p^*(M) \times T_p(M) \rightarrow \mathbb{R}, \quad G(\omega, X) = \omega(\hat{G}(X))$$

which is also bilinear in both arguments, and hence is a rank $(1, 1)$ tensor. If G is the identity map, then our induced function

$$\delta : T_p^*(M) \times T_p(M) \rightarrow \mathbb{R}, \quad \delta(\omega, X) = \omega(X)$$

is indeed our standard Kronecker delta function. If we set $\omega = x^\mu$, $X = e_\nu$, then $\delta^\mu_\nu = \partial_\mu(x^\mu) = \delta^\mu_\nu$.

7.3 Question 3

In this question, we show that only symmetric (antisymmetric) parts of a tensor are 'conserved' when contracted with symmetric (antisymmetric) tensors over the same indices. If $S^{\mu\nu}$ is symmetric, then

$$\begin{aligned} V^{(\mu\nu)} S_{\mu\nu} &= \frac{1}{2} (V^{\mu\nu} + V^{\nu\mu}) S_{\mu\nu} \\ &= \frac{1}{2} V^{\mu\nu} S_{\mu\nu} + \frac{1}{2} V^{\nu\mu} S_{\mu\nu} \\ &= \frac{1}{2} V^{\mu\nu} S_{\mu\nu} + \frac{1}{2} V^{\nu\mu} S_{\nu\mu} \\ &= \frac{1}{2} V^{\mu\nu} S_{\mu\nu} + \frac{1}{2} V^{\mu\nu} S_{\mu\nu} \\ &= V^{\mu\nu} S_{\mu\nu} \end{aligned}$$

Going into the second last line we've just relabelled over summed indices. Going into the third line we've used the fact that S is a symmetric tensor. The case for when we contract $V^{\mu\nu}$ for an antisymmetric tensor is entirely similar.

7.4 Question 5

We show that our components $F_{\mu\nu}$ transform appropriately under a change of coordinates. This is done with the chain rule.

$$\begin{aligned}
 F_{\mu\nu} &\rightarrow F'_{\mu\nu} \\
 &= \frac{\partial^2 f}{\partial x'^\mu \partial x'^\nu} \\
 &= \frac{\partial}{\partial x'^\mu} \left(\frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial f}{\partial x^\rho} \right) \\
 &= \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial}{\partial x^\sigma} \left(\frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial f}{\partial x^\rho} \right) \\
 &= \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial^2 x^\rho}{\partial x^\sigma \partial x'^\nu} \frac{\partial f}{\partial x^\rho} + \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial^2 f}{\partial x^\rho \partial x^\sigma} \\
 &= \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial^2 f}{\partial x^\sigma \partial x^\rho} \\
 &= \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} F_{\rho\sigma}
 \end{aligned}$$

There's a reason why we've taken the first term to zero going into the fifth line. Since $df = 0$ at p , then for an arbitrary vector A in any basis, we have that at $p \in \mathcal{M}$,

$$df(A) = A(f) = A^\mu \partial_\mu(f) = 0, \implies \partial_\mu(f) = 0 \text{ at } p, \quad \forall \mu = 1, \dots, D$$

So this term goes to zero, since we only have a single derivative acting on f . * Thus, the Hessian obeys the tensor transformation law. Since our components transform in the two lower indices with a change of coordinates, this object is basis invariant and hence is a rank $(0, 2)$ tensor.

Since this is a rank $(0, 2)$ tensor, our coordinate independent way of expressing this object would be

$$F : T_p(M) \times T_p(M) \rightarrow \mathbb{R}$$

This specific representation is

$$F(V, W) = VW(f)$$

We can show this by expanding with coordinates.

$$\begin{aligned}
 VW(f) &= V^\mu \partial_\mu (W^\nu \partial_\nu f) \\
 &= (\partial_\mu W^\nu) (\partial^\mu V_\nu) f + V^\mu W^\nu \partial_\mu \partial_\nu f \\
 &= (V^\mu \partial_\mu W^\nu) \partial_\nu f + W^\nu V^\mu \partial_\mu \partial_\nu f \\
 &= W^\nu V^\mu \partial_\mu \partial_\nu f \\
 &= W^\nu V^\mu F_{\mu\nu}
 \end{aligned}$$

Here we've used the fact that $df = 0$, which implies that for an arbitrary set of components Z^μ , we have that $Z^\mu \partial_\mu f = 0$. In the above, we identify this as $Z^\nu = V^\mu \partial_\mu W^\nu$, and hence the first term in the third line goes to zero.

This implies that $F_{\mu\nu}$ are indeed the components of F . Multi linearity in both arguments is just inherited from the linearity of V, W as vector fields.

* A different argument would be that one can note that the first term is of the form

$$(G_{\mu'\nu'})^\rho \partial_\rho f$$

Where we can view $G_{\mu'\nu'}$ as D^2 separate vectors indexed by μ' and ν' . Thus, since $df = 0$, this term goes to zero. (I like this way since it's manifestly a bit more basis invariant!)

7.5 Question 6

This question explores how the determinant of a metric transforms under coordinate transformations. For this question, we denote the determinant of a change of basis as the Jacobian:

$$\mathcal{J} = \det \left(\frac{\partial x^\rho}{\partial x'^\nu} \right)$$

Hence, when we do a coordinate transform, since $\det(AB) = \det(A)\det(B)$, we have that

$$\begin{aligned} g' &= \det(g'_{\mu\nu}) \\ &= \det \left(\frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} g_{\rho\sigma} \right) \\ &= \det(g) J^{-2} \end{aligned}$$

This is because the expression in the determinant is the inverse of what we've defined the Jacobian to be.

7.6 Question 7

Lie derivative of 1-form

Using the Leibniz rule for our Lie derivative, we consider the Lie derivative for $\mathcal{L}_X(\omega Y)$;

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

This expression is basis independent. Now, observe that $\omega(Y)$ is a function in $C^\infty(\mathcal{M})$. Thus, the Lie derivative for this term is just given by $X(\omega(Y))$. We also know that $\mathcal{L}_X(Y) = [X, Y]$. Thus,

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

Up to index relabelling of the dummy indices, the last term of the LHS and the second term of the RHS are the same, so they cancel out. Moving the negative term on the LHS to the right hand side and relabelling gives

$$Y_\nu (X^\mu \partial_\mu \omega^\nu + \omega_\mu \partial_\nu X^\mu) = (\mathcal{L}_X \omega)_\nu Y^\nu$$

However, since Y was arbitrary we can just read off the basis independent components here.

$$(X^\mu \partial_\mu \omega^\nu + \omega_\mu \partial_\nu X^\mu) = (\mathcal{L}_X \omega)_\nu$$

Lie derivative for a 2-tensor

We play exactly the same game with the rank $(0, 2)$ tensor as well.

$$\mathcal{L}_X(g(V, W)) = (\mathcal{L}_X g)(V, W) + g(\mathcal{L}_X V, W) + g(V, \mathcal{L}_X W)$$

In components, and multiplying out with the product rule, this term is

$$\begin{aligned} X^\nu (\partial_\mu g_{\alpha\beta}) V^\alpha W^\beta + X^\mu g_{\alpha\beta} W^\beta \partial_\mu V^\alpha + X^\mu g_{\alpha\beta} V^\alpha \partial_\mu W^\beta &= \\ = (\mathcal{L}_X g)_{\alpha\beta} V^\alpha W^\beta + g_{\alpha\beta} [X, V]^\alpha \partial_\mu W^\beta + g_{\alpha\beta} V^\alpha [X, W]^\beta \end{aligned}$$

The right hand side is just equal to, expanding the commutators in terms of components,

$$= (\mathcal{L}_X g)_{\alpha\beta} V^\alpha W^\beta + g_{\alpha\beta} X^\nu \partial_\nu V^\alpha W^\beta - g_{\alpha\beta} V^\nu \partial_\nu X^\alpha W^\beta + g_{\alpha\beta} V^\alpha X^\nu \partial_\nu W^\beta - g_{\alpha\beta} V^\alpha W^\beta \partial_\nu X^\beta$$

Now up to index relabelling μ and ν , the second and fourth terms of this equation cancel out with the second and third terms on the LHS of our first equation. Thus, we're left with

$$X^\mu \partial_\mu g_{\alpha\beta} V^\alpha W^\beta + g_{\alpha\beta} V^\nu (\partial_\nu X^\alpha) W^\beta + g_{\alpha\beta} V^\alpha W^\nu (\partial_\nu X^\beta) = (\mathcal{L}_X g)_{\alpha\beta} V^\alpha W^\beta$$

Now, as before, relabelling ν, α in the second term and ν, β in the third term recovers the expression in the question (after factorising out V, W).

Last part

The last part of the question is just a matter of substituting in definitions.

$$\begin{aligned}(\iota_X d\omega)_\mu &= X^\nu (d\omega)_{\nu\mu} \\ &= X^\nu 2\partial_{[\nu}\omega_{\mu]} \\ &= X^\nu \partial_\nu \omega_\mu - X^\nu \partial_\mu \omega_\nu\end{aligned}$$

Also, we have

$$\begin{aligned}d(\iota_X \omega)_\mu &= \partial_\mu (\iota_X \omega) \\ &= \partial_\mu (X^\nu \omega_\nu) \\ &= \omega_\nu \partial_\mu X^\nu + X^\nu \partial_\mu \omega_\nu\end{aligned}$$

Adding these terms together gives

$$(\iota_X d\omega)_\mu + d(\iota_X \omega)_\mu = X^\nu \partial_\nu \omega_\mu - X^\nu \partial_\mu \omega_\nu + \omega_\nu \partial_\mu X^\nu + X^\nu \partial_\mu \omega_\nu = X^\nu \partial_\nu \omega_\mu + \omega_\nu X^\nu$$

since we have cancellation with the second and last term.

7.7 Question 8

The components of the exterior derivative of a p -form consists of the antisymmetrisation of $p+1$ indices. Suppose that ω is a p -form. Then

$$(d\omega)_{\mu_1\mu_2\ldots\mu_{p+1}} = (p+1)\partial_{[\mu_1}\omega_{\mu_2\ldots\mu_{p+1}]}$$

Thus, we can expand the components of the exterior derivative of this object as

$$(d(d\omega))_{\mu_1\ldots\mu_{p+1}} = (p+2)(p+1)\partial_{[\mu_1}\partial_{\mu_2}\omega_{\mu_3\ldots\mu_{p+2}]}$$

Now, we have a tricky thing to deal with here. We have an antisymmetrisation nested inside of an antisymmetrisation. We claim that nesting an antisymmetrisation inside an antisymmetrisation is just the larger antisymmetrisation:

$$X_{[\mu_1[\mu_2\ldots\mu_p]]} = X_{[\mu_1\mu_2\ldots\mu_p]}$$

We can prove this by expanding out an antisymmetrisation based on just the μ_1 index first.

$$\begin{aligned} X_{[\mu_1\ldots\mu_p]} &= \frac{1}{p!} \left(\sum_{\sigma \in S_{p-1}} \epsilon(\sigma) X_{\mu_1\mu_{\sigma(2)}\ldots\mu_{\sigma(p)}} \right. \\ &\quad - \sum_{\sigma \in S_{p-1}} \epsilon(\sigma) X_{\mu_{\sigma(2)}\mu_1\mu_{\sigma(3)}\ldots\mu_{\sigma(p)}} \\ &\quad \vdots \\ &\quad \left. + (-1)^{p+1} \sum_{\sigma \in S_{p-1}} \epsilon(\sigma) X_{\mu_{\sigma(2)}\mu_{\sigma(3)}\ldots\mu_{\sigma(p)}\mu_1} \right) \end{aligned}$$

So, when we nest antisymmetrisations, we have terms in the sum that look like

$$X_{[\mu_1[\mu_2\ldots\mu_p]]} = \sum_{\text{similar sum as above but of}} \frac{1}{p!} \sum_{\sigma \in S_{p-1}} \epsilon(\sigma) X_{\mu_1[\mu_{\sigma(2)}\ldots\mu_{\sigma(p)}]}$$

But, expanding out the definition, we have that this term is just equal to

$$= \frac{1}{p!} \frac{1}{(p-1)!} \sum_{\sigma' \in S_{p-1}} \sum_{\sigma \in S_{p-1}} \epsilon(\sigma') \epsilon(\sigma) X_{\mu_1\mu_{\sigma'\sigma(2)}\ldots\mu_{\sigma'\sigma(p)}}$$

But, we can compose each pair of permutations and write $\sigma'' = \sigma'\sigma$, and since the sign operator for permutations is a homomorphism, we can write that $\epsilon(\sigma)\epsilon(\sigma') = \epsilon(\sigma'')$. But, we have to be careful to make sure to count twice. Hence the term above is

$$\frac{1}{p!} \frac{1}{(p-1)!} \sum_{\sigma'} \sum_{\sigma} \epsilon(\sigma'') X_{\mu_1\mu_{\sigma''(2)}\mu_{\sigma''(3)}\ldots\mu_{\sigma''(p)}}$$

Now, we can relabel the σ index as σ'' , and so we're just summing over an extra σ' . This gives

$$\frac{1}{p!} \frac{1}{(p-1)!} \sum_{\sigma'} \sum_{\sigma''} \epsilon(\sigma'') X_{\mu_1\mu_{\sigma''(2)}\mu_{\sigma''(3)}\ldots\mu_{\sigma''(p)}} = \frac{1}{p!} \sum_{\sigma''} X_{\mu_1\mu_{\sigma''(2)}\mu_{\sigma''(3)}\ldots\mu_{\sigma''(p)}}$$

But this just removes the effect of an antisymmetric tensor! Hence, given a set of indices, we have

$$[[\mu_1\mu_2\ldots\mu_p]] = [\mu_1\mu_2\ldots\mu_p]$$

So, the nested indices have no effect. Thus we have that

$$d(d\omega)_{\mu_1\ldots\mu_{p+2}} = (p+2)(p+1)\partial_{[\mu_1}\partial_{\mu_2}\omega_{\mu_3\ldots\mu_{p+2}]} = (p+2)(p+1)\partial_{[\mu_1}\partial_{\mu_2}\omega_{\mu_3\ldots\mu_{p+2}]} = 0$$

By antisymmetry of mixed partial derivatives.

Now, we'd like to show a 'product rule for the exterior derivatives and one forms.

$$d(\omega \wedge \epsilon) = d\omega \wedge \epsilon + (-1)^p \omega \wedge d\epsilon$$

The right hand side in components, by definition is

$$d(\omega \wedge \epsilon)_{\gamma\mu_1\ldots\mu_p\nu_1\ldots\nu_q} = \frac{(p+q+1)(p+q)!}{p!q!} (\partial_{[\gamma}\omega_{\mu_1\ldots\mu_p}\epsilon_{\nu_1\ldots\nu_q]} + \omega_{[\gamma\mu_1\ldots\mu_{p-1}}\partial_{\mu_p}\epsilon_{\nu_1\ldots\nu_q]})$$

Let's see what we've done here. We used the product rule to expand out the derivatives, but when doing this we need to preserve our order of our indices, which is why we kept it in this form.

Now, we reorder the indices $\gamma, \mu_1, \ldots, \mu_p$. We do the procedure

$$(\gamma, \mu_1, \mu_2, \ldots, \mu_p) \rightarrow (-1)(\mu_1, \gamma, \mu_2, \ldots, \mu_p) \rightarrow \cdots \rightarrow (-1)^p(\mu_1, \mu_2, \ldots, \mu_p, \gamma)$$

So, we've picked up a factor of $(-1)^p$. Thus, when we stick in an extra set of antisymmetric indices (which doesn't change things as we showed earlier), we get that our expression above is equal to

$$\frac{(p+q+1)!}{p!q!}\partial_{[\gamma}\omega_{\mu_1\ldots\mu_p}\epsilon_{\nu_1\ldots\nu_q]} + (-1)^p\frac{(p+q+1)!}{p!q!}\omega_{[\mu_1\ldots\mu_p}\partial_{\gamma}\epsilon_{\nu_1\ldots\nu_q]}$$

Now, we substitute our expression for an exterior derivative. The above expression is equal to

$$\begin{aligned} &= \frac{(p+q+1)!}{p!q!} \frac{1}{(p+1)} d\omega_{[\gamma\mu_1\ldots\mu_p}\epsilon_{\nu_1\ldots\nu_q]} + (-1)^p \frac{(p+q+1)!}{p!q!} \frac{1}{(p+1)!} \omega_{[\mu_1\ldots\mu_p} d\epsilon_{\nu_1\ldots\nu_q]} \\ &= \frac{(p+q+1)!}{(p+1)!q!} (d\omega)_{[\gamma\mu_1\ldots\mu_p}\epsilon_{\nu_1\ldots\nu_q]} + (-1)^p \frac{(p+q+1)!}{p!(q+1)!} \mu_{[\mu_1\ldots\mu_p} d\epsilon_{\gamma\nu_1\ldots\nu_q]} \end{aligned}$$

But these are explicitly the components of what we are looking for. Hence, we've shown that

$$d(\omega \wedge \epsilon) = d\omega \wedge \epsilon + (-1)^p \omega \wedge d\epsilon$$

Finally, we wish to show that a pull back of the one form ω , denoted as $\psi^*\omega$ from the manifold M to N commutes with our exterior derivative. In other words, we wish to show that

$$d(\psi^*\omega) = \psi^*(d\omega)$$

We do this by expanding the components explicitly first. We have that

$$\begin{aligned} d(\psi^*\omega)_{\nu\mu_1\ldots\mu_p} &= \partial_{[\nu}(\psi^*\omega)_{\mu_1\ldots\mu_p]} \\ &= \frac{\partial}{\partial x^{[\nu}} \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_p}}{\partial x^{\mu_p]}} \omega_{\alpha_1\ldots\alpha_p} \end{aligned}$$

We were careful here to ensure that, since $\psi^*\omega$ lives in the manifold M , we need to differentiate with respect to the coordinates x^α . Now, here we used the fact that for a general p-form, our components change like

$$(\psi^*\omega)_{\mu_1\dots\mu_p} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \dots \frac{\partial y^{\alpha_p}}{\partial x^{\mu_p}} \omega_{\alpha_1\dots\alpha_p}$$

Since we have one differential as $\frac{\partial}{\partial x^\nu}$, we can use the chain rule to expand this term out, giving that the above expression is equal to

$$\frac{\partial}{\partial y^\beta} \frac{\partial y^\beta}{\partial x^{[\nu}} \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \dots \frac{\partial y^{\alpha_p}}{\partial x^{\mu_p}} \omega_{\alpha_1\dots\alpha_p]} = \frac{\partial y^\beta}{\partial x^{[\nu}} \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \dots \frac{\partial y^{\alpha_p}}{\partial x^{\mu_p]} \frac{\partial}{\partial y^\beta} \omega_{\alpha_1\dots\alpha_p}$$

Now, this step deserves some explanation. By symmetry of mixed partial derivatives, we're allowed to commute the $\frac{\partial}{\partial y^\beta}$ term past everything. Because even though the product rule dictates that this has to differentiate each term in this big product, the first terms are derivatives, so by symmetry of mixed partial derivatives inside an antisymmetric tensor, all these extra terms go to zero.

Finally, due to our ability to relabel dummy indices, one can show that for a vector contraction of the form

$$X^{\mu_1}_{\nu_1} \dots X^{\mu_n}_{\nu_n} Y_{[\mu_1\dots\mu_n]} = X^{\mu_1}_{[\nu_1} \dots X^{\mu_1}_{\mu_n]} Y_{\mu_1\dots\mu_n}$$

This means that indeed, we can shift the antisymmetric terms to the right most indices, giving

$$d(\psi^*\omega)_{\nu\mu_1\dots\mu_p} = \frac{\partial y^\beta}{\partial x^\nu} \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \dots \frac{\partial y^{\alpha_p}}{\partial x^{\mu_p}} \frac{\partial}{\partial y^{[\beta}} \omega_{\alpha_1\dots\alpha_p]}$$

But indeed, these are the components of the pulled back one form

$$\psi^*d(\omega)$$

So we're done!

7.8 Question 9

This question shows the advantage of coming up with a tensorial definition of objects first, to simplify calculations for components. In the case when $p = 1$, we set the basis $X_1 = e_\mu$, $X_2 = e_\nu$. Then, our definition in tensorial form gives

$$(d\omega)_{\mu\nu} = d\omega(e_\mu, e_\nu) = e_\mu(\omega(e_\nu)) - e_\nu(\omega(e_\mu)) - \omega([e_\mu, e_\nu])$$

Now note that e_μ, e_ν aren't indexed components per say, they're just our choice of basis vector. The upshot of doing this is that by symmetry of mixed partial derivatives, we have

$$[e_\mu, e_\nu] = \frac{\partial^2}{\partial x^\mu \partial x^\nu} - \frac{\partial^2}{\partial x^\nu \partial x^\mu} = 0$$

Hence, the last term vanishes and thus

$$(d\omega)_{\mu\nu} = \partial_\mu(\omega_\nu) - \partial_\nu(\omega_\mu)$$

In the above line we've used the fact that $\omega(e_\mu) = \omega_\mu$, which can be shown by expanding ω into its components and covector basis. Once again, when $p = 3$, we can still use this trick of using

the vector basis $\{e_\mu\}$, to forget about the commutator terms in the definition. This is because our definition of $d\omega$ contains terms like

$$\omega([e_\mu, e_\nu], e_\alpha)$$

but since the commutator vanishes and ω is multilinear, $\omega(0, e_\alpha) = 0$. Thus, the only terms that are preserved from the definition is that

$$d\omega(e_\mu, e_\nu, e_\rho) = e_\mu\omega(e_\nu, e_\rho) - e_\nu\omega(e_\rho, e_\mu) + e_\rho\omega(e_\mu, e_\nu)$$

Using the fact that the basis vectors are derivative terms this becomes

$$(d\omega)_{\mu\nu\rho} = \partial_\mu\omega_{\nu\rho} - \partial_\nu\omega_{\rho\mu} + \partial_\rho\omega_{\mu\nu}$$

This is consistent with our definition that

$$(d\omega)_{\mu\nu\rho} = 3\partial_{[\mu}\omega_{\nu\rho]} = \frac{1}{2} \sum_{\text{anti symmetric perms}} \partial_\mu\omega_{\nu\rho}$$

Note the seemingly extraneous factor of two here, but this cancels out since ω is a two form and therefore we count twice the number of permutations.

7.9 Question 10

For now we'll just show that $d\sigma_1 = \sigma_2 \wedge \sigma_3$. Let's calculate the right hand side explicitly, we have that

$$\begin{aligned} \sigma_2 \wedge \sigma_3 &= (\cos\psi d\theta + \sin\psi \sin\theta d\phi) \wedge (d\psi + \cos\theta d\phi) \\ &= \cos\psi d\theta \wedge d\psi + \sin\psi \sin\theta d\phi \wedge d\psi + \cos\psi \cos\theta d\theta d\phi + \cos\psi \cos\theta d\theta \wedge d\phi \\ &\quad + \sin\psi \sin\theta \cos\theta d\phi \wedge d\phi \\ &= \cos\psi d\theta \wedge d\psi + \sin\psi \sin\theta d\phi \wedge d\psi + \cos\psi \cos\theta d\theta d\phi + \cos\psi \cos\theta d\theta \wedge d\phi \end{aligned}$$

When we do an exterior derivative in 3 dimensions on a one form, we get that in our wedge product basis

$$\begin{aligned} d\omega &= (\partial_1\omega_2 - \partial_2\omega_1)dx^1 \wedge dx^2 \\ &\quad + (\partial_2\omega_3 - \partial_3\omega_2)dx^2 \wedge dx^3 \\ &\quad + (\partial_3\omega_1 - \partial_1\omega_3)dx^3 \wedge dx^1 \end{aligned}$$

Now, if we identify $dx^1 = d\theta, dx^2 = d\psi, dx^3 = d\phi$, then our first component to calculate is

$$(\partial_\theta(\sigma_1)_\psi - \partial_\psi(\sigma_1)_\theta)d\theta \wedge d\psi = \cos\psi d\theta \wedge d\psi$$

Similarly, we find that

$$\begin{aligned} (\partial_2(\sigma_1)_3 - \partial_3(\sigma_1)_2)dx^2 \wedge dx^3 &= \sin\psi \sin\theta d\phi \wedge d\psi \\ (\partial_3(\sigma_1)_1 - \partial_1(\sigma_1)_3)dx^3 \wedge dx^1 &= \cos\psi \cos\theta d\theta \wedge d\phi \end{aligned}$$

7.10 Question 11

The point of this question is to show that a basis which is coordinate induced is equivalent to it's commutator vanishing. Showing one way is straightforward, we have that

$$[e_\mu, e_\nu] = \frac{\partial^2}{\partial x^\nu \partial x^\mu} - \frac{\partial^2}{\partial x^\mu \partial x^\nu} = 0$$

This is by the symmetry of mixed partial derivatives.

Now we go the other way. From the condition that

$$[e_\mu, e_\nu] = \gamma^\rho_{\mu\nu} e_\rho$$

We expand this out

$$[e_\mu^\rho \frac{\partial}{\partial x^\rho}, e_\nu^\lambda \frac{\partial}{\partial x^\lambda}] = \gamma^\rho_{\mu\nu} e_\rho^\lambda \frac{\partial}{\partial x^\lambda}$$

Writing this out explicitly and cancelling cross terms give

$$e_\mu^\rho \frac{\partial e_\nu^\lambda}{\partial x^\sigma} \frac{\partial}{\partial x^\lambda} - e_\nu^\lambda \frac{\partial e_\mu^\sigma}{\partial x^\sigma} \frac{\partial}{\partial x^\sigma} = \gamma^\sigma_{\mu\nu} e_\sigma^\lambda \frac{\partial}{\partial x^\lambda}$$

Upon relabelling dummy indices (for example replacing $\sigma \rightarrow \lambda$ in the second term), we end up with the expression

$$e_\mu^\sigma \frac{\partial e_\nu^\lambda}{\partial x^\sigma} \frac{\partial}{\partial x^\lambda} - e_\nu^\sigma \frac{\partial e_\mu^\lambda}{\partial x^\sigma} \frac{\partial}{\partial x^\lambda} = \gamma^\sigma_{\mu\nu} e_\sigma^\lambda \frac{\partial}{\partial x^\lambda}$$

But, we can just factor out our partials $\frac{\partial}{\partial x^\lambda}$ to get our required expression. Now, we appeal to the fact that

$$e_\mu^\rho f_\rho^\nu = \delta_\mu^\nu$$

Differentiating both sides, we have that

$$f_\rho^\nu \frac{\partial e_\mu^\rho}{\partial x^\sigma} + e_\mu^\rho \frac{\partial f_\rho^\nu}{\partial x^\sigma} = 0$$

Hence,

$$\begin{aligned} f_\rho^\nu \frac{\partial e_\mu^\rho}{\partial x^\sigma} &= -e_\mu^\rho \frac{\partial f_\rho^\nu}{\partial x^\sigma} \\ e_\nu^\tau f_\rho^\nu \frac{\partial e_\mu^\rho}{\partial x^\sigma} &= -e_\nu^\tau e_\mu^\rho \frac{\partial f_\rho^\nu}{\partial x^\sigma} \\ \frac{\partial e_\mu^\tau}{\partial x^\sigma} &= -e_\nu^\tau e_\mu^\rho \frac{\partial f_\rho^\nu}{\partial x^\sigma} \end{aligned}$$

Substituting this into the above,

$$-e_\mu^\sigma e_\alpha^\lambda e_\nu^\beta \frac{\partial f_\beta^\alpha}{\partial x^\sigma} + e_\nu^\sigma e_\alpha^\lambda e_\mu^\beta \frac{\partial f_\beta^\alpha}{\partial x^\sigma} = \gamma^\alpha_{\mu\nu} e_\alpha^\lambda$$

We can cancel out the e_α^λ . We get the

$$-e_\mu^\sigma e_\nu^\beta \frac{\partial f_\beta^\alpha}{\partial x^\sigma} + e_\nu^\sigma e_\mu^\beta \frac{\partial f_\beta^\alpha}{\partial x^\sigma} = \gamma^\alpha_{\mu\nu}$$

Contraction with $f^\mu{}_\lambda f^\nu{}_\sigma$, gives the result,

$$\frac{\partial f^\rho{}_\sigma}{\partial x^\lambda} - \frac{\partial f^\rho{}_\lambda}{\partial x^\sigma} = -\gamma^\rho{}_{\mu\nu} f^\mu{}_\lambda f^\nu{}_\sigma$$

However, this implies that each of f^μ is closed since if we have $[e_\mu, e_\nu] = 0$, then $\gamma = 0$ for all indices. So, we get that $df^\mu = 0$ for all μ by the above formula. Hence, the Poincare lemma states that we can write

$$f^\mu{}_\nu = \partial_\nu \eta^\mu$$

Hence, η^μ are a set of functions. Our condition that

$$\delta_\mu{}^\nu = e_\mu{}^\alpha f^\nu{}_\alpha = e_\mu{}^\alpha \partial_\alpha \eta^\nu = e_\mu(\eta^\nu)$$

Hence, relabelling $\eta^\nu = x^\nu$ gives us a set of coordinates given that η^ν are independent. However, we know this is the case since if we have a linear sum

$$\sum_\mu \lambda_\mu \eta^\mu = 0$$

contracting with e gives each coefficient 0. Hence, we have that n of these are linearly independent. Thus, the collection η^i is a map from $\mathcal{M} \rightarrow \mathbb{R}^n$ is injective, and since we have n of these maps, they span (by the Steinitz exchange lemma) they also span. Hence, we have a homeomorphism, and thus $\{\eta^i\}$ is a set of coordinates. Thus, the corresponding e_ν are a set of coordinate induced basis vectors.

7.11 Being careful about how we expand things

We need to be careful about trying to apply the covariant derivative to Christoffel symbols. For example, writing out

$$\nabla_\nu \Gamma^\nu_{\rho\mu}$$

doesn't make sense since Γ is not a tensor! Γ doesn't transform as a function. Thus, we need to expand out things the right way. Let's try to expand out the term

$$\nabla_{[\mu} \nabla_{\nu]} f$$

We need to work out from in, because we know that $\nabla_\nu f$ is a covector.

7.12 Tricks in the Levi-Civita connection

In the Levi-Civita connection, we have that raising indexes

7.13 Integrating over volume forms

Always integrate 'as is' - we don't have to add the metric if it's not there. When trying to calculate the hodge star, write things out in terms of the Vielbein basis.

The minus sign when taking our hodge star comes from our $\hat{\theta}^0$ has the Minkowski metric.

(Weird thought -

7.14 Showing that magnitude is constant

It's a lot easier to show things like the angle between two vectors being constant

Sometimes we need to avoid coordinate singularities, a work around is to rotate the manifold.

7.15 Symmetries of the Riemann Tensor

We have 4 main symmetries of the Riemann Tensor to worry about. Antisymmetry in the 1 and 2 indices, anti-symmetry in the 3 and 4 indices, Symmetry in switching, and Bianchi identity.

We do the first steps by writing things out in terms of an anti-symmetric basis. Then, adding on the Bianchi identity is a matter of counting the extra independent constraints we can get.

In 2D, this means that $G_{\mu\nu} = 0$, so no matter exists.

7.16 General techniques

Looking at the number of components can help simplify things greatly.