Differential Geometry

Fionn Fitzmaurice fionn@maths.tcd.ie

These notes correspond to the differential geometry course taught by Peter Taylor in Michælmas term 2011—essentially the first half of the general relativity course.

1 Manifolds

1.1 Topological Spaces

Let $p \in \mathbb{R}^n$. A neighbourhood of p is any set $V \subset \mathbb{R}^n$ such that V contains an open solid sphere of centre p.

Properties:

(i) $p \in$ any neighbourhood of p.

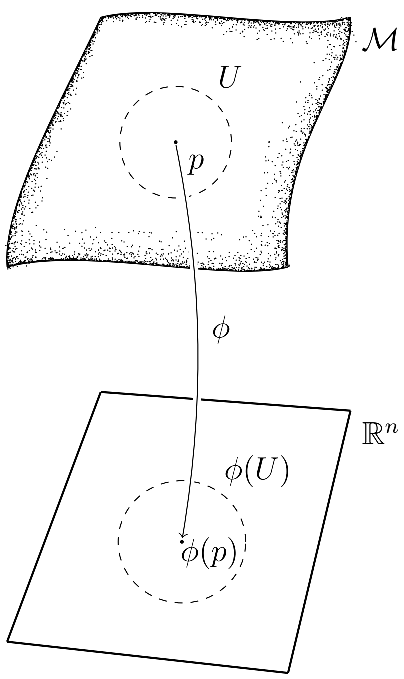
(ii) If V is a neighbourhood of p and $V \subset U$, then U is a neighbourhood of p.

(iii) If U, V are neighbourhoods of p, then $U \cap V$ is a neighbourhood of p.

(iv) If U is a neighbourhood of p, $\exists$ a neighbourhood V of p such that $V \supset U$ and V is a neighbourhood of each of its points.

Definition. A topological space is a set of points $\mathcal{M}$ along with an assignment to each $p \in \mathcal{M}$ of collections of subsets called neighbourhoods, satisfying properties (i)–(iv).

1.2 Charts



Let $\mathcal{M}$ be a topological space, $p \in \mathcal{M}$ be some point in this space and U be an open neighbourhood of p. A chart on U is an injective map

$$\phi : U \to \phi(U) \subset \mathbb{R}^n.$$

The $\phi(p) \in \mathbb{R}^n$ constitutes a local coordinate system defined in an open neighbourhood U. We usually write

$$\begin{aligned}\phi(p) &= \{x^\mu(p)\} \\ &= \Big(x^1(p), \ldots, x^n(p)\Big)\,.\end{aligned}$$

Note that the choice of chart is arbitrary, implying Einstein's equivalence principle.

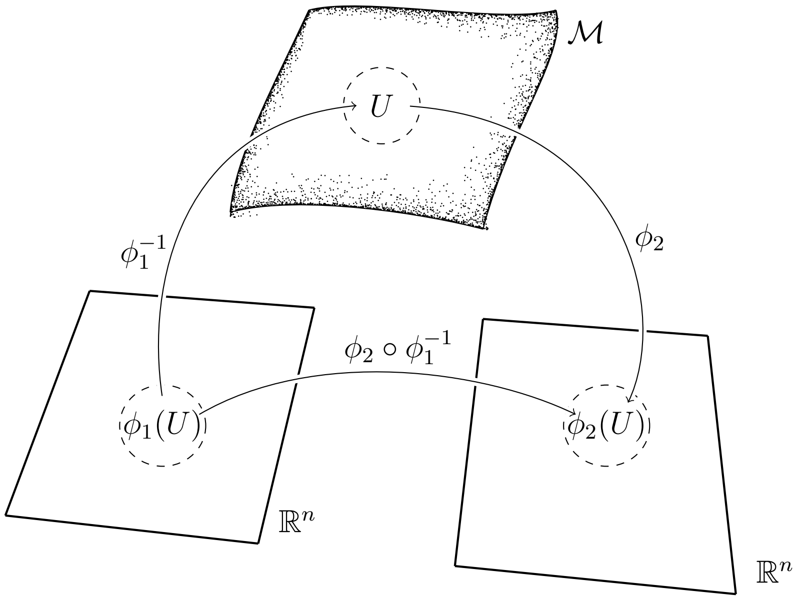
1.3 Meshing Conditions: Coordinate Transformations

Suppose we have two charts ϕ_1, ϕ_2 on $U \subset \mathcal{M}$. Since these charts are injective, they are invertible[1], e.g.

$$\phi_1^{-1} : \phi_1(U) \subset \mathbb{R}^n \to U.$$

We may define

$$\begin{aligned} \phi_2 \circ \phi_1^{-1} &: \mathbb{R}^n \to \mathbb{R}^n \\ &: \phi_1(U) \to \phi_2(U). \end{aligned}$$



We require these maps to be smooth (C^∞) where they are defined. For $p \in U$, the map $\phi_2 \circ \phi_1^{-1}(p)$ defines a coordinate transformation from the coordinates

$$\phi_1(p) = \left(x^1(p), \dots, x^n(p)\right)$$

to the coordinates

$$\phi_2(p) = \left(X^1(p), \dots, X^n(p)\right).$$

Example 1.3.1. Let $\mathcal{M} = \mathbb{R}^2$, let ϕ_1 map p to the Cartesian coordinates (x, y) and ϕ_2 map to the Cartesian coordinates (X, Y) obtained from the first set by a rotation through the constant angle α.

[1]Injections can be made bijective by replacing codomain with range.

$\phi_2 \circ \phi_1^{-1}$ maps

$$(x, y) \mapsto (X = x\cos\alpha + y\sin\alpha,\, Y = -x\sin\alpha + y\cos\alpha).$$

We can define a derivative matrix

$$D(\phi_2 \circ \phi_1^{-1}) = \begin{pmatrix} \frac{\partial X}{\partial x} & \frac{\partial X}{\partial y} \\ \frac{\partial Y}{\partial x} & \frac{\partial Y}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

The Jacobian $\mathcal{J} \equiv \det(D) = 1$. Recall that $\mathcal{J} \neq 0$ implies an invertible transformation. $\mathcal{J}$ non-singular implies ϕ_1, ϕ_2 are C^∞-related.

Introduce another chart ϕ_3 which maps p to polar coordinates (r, θ). Then

$$\phi_3 \circ \phi_1^{-1} : (x, y) \mapsto \left(r = \sqrt{x^2 + y^2},\, \theta = \tan^{-1}(\tfrac{y}{x})\right),$$
$$\mathcal{J} = \det(D) = \frac{1}{r}.$$

ϕ_1, ϕ_3 are C^∞-related except at $r = 0$. To cover all of $\mathbb{R}^2$, we would need at least two sets of polar coordinates with different origins.

1.4 Definition of a Manifold

Informally, a manifold is a set of points $\mathcal{M}$ that locally looks like a subset of $\mathbb{R}^n$. The simplest example of a curved manifold is S^2.

A set of C^∞-related charts such that every point $p \in \mathcal{M}$ lies in the domain of at least one chart is a C^∞ atlas for $\mathcal{M}$. The union of all such atlases is known as the C^∞ maximal atlas.

We define a C^∞ n-dimensional manifold by a set $\mathcal{M}$ along with a maximal atlas.

2 Tangent Vectors and Tangent Spaces

In our familiar treatment of vectors in $\mathbb{R}^n$, they represent "directed magnitudes". This is no longer a useful notion. Rather, to each point $p \in \mathcal{M}$ we have a set of all possible vectors at p known as the tangent space $T_p(\mathcal{M})$.

We prefer to describe the geometry of $\mathcal{M}$ from intrinsic properties alone; we won't rely on embedding in a higher dimensional space.

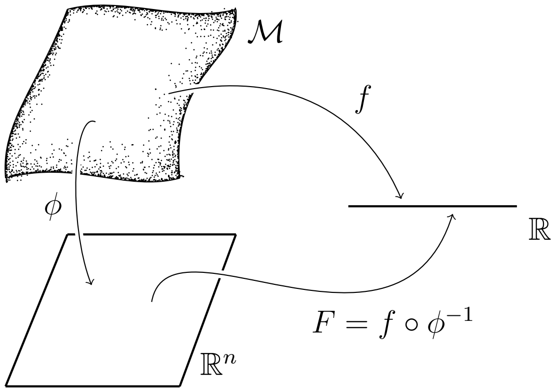
2.1 Smooth Functions

Let $\mathcal{M}$ be a manifold and f be a real function.

$$f : \mathcal{M} \to \mathbb{R}.$$

How do we define the "smoothness" of f? Introduce a chart ϕ and a new function F such that

$$F : \mathbb{R}^n \to \mathbb{R}, \quad F = f \circ \phi^{-1}.$$



We say that f is smooth iff F is smooth in the usual sense.

Theorem 2.1.1. The smoothness of f is chart independent.

Proof. Let ϕ_1, ϕ_2 be two meshing charts, $i \in \{1, 2\}$.

$$\begin{aligned} F_i &= f \circ \phi_i^{-1}, \\ F_1 &= f \circ \phi_1^{-1} \\ &= f \circ \phi_2^{-1} \circ \phi_2 \circ \phi_1^{-1} \\ &= F_2 \circ \phi_2 \circ \phi_1^{-1} \end{aligned}$$

and $\phi_2 \circ \phi_1^{-1}$ is smooth since ϕ_1, ϕ_2 are meshing charts. Hence, the smoothness properties of F_1 are the same as F_2. □

The definition of smooth functions may be generalised to a function mapping a manifold $\mathcal{M}$ to another manifold $\mathcal{N}$,

$$f : \mathcal{M} \to \mathcal{N}.$$

Let ϕ_1 be a chart in $\mathcal{M}$ ($\dim(\mathcal{M}) = n_1$) and ϕ_2 be a chart in $\mathcal{N}$ ($\dim(\mathcal{N}) = n_2$). Define $F = \phi_2 \circ f \circ \phi_1^{-1}$. f is smooth iff F is smooth. It is easy to prove that this is chart independent.

Note that the notation $\dfrac{\partial f}{\partial x^\mu}$ really means

$$\frac{\partial f}{\partial x^\mu} = \frac{\partial F}{\partial x^\mu} = \frac{\partial\, \phi_2 \circ f \circ \phi_1^{-1}}{\partial x^\mu}.$$

2.2 Smooth Curves

Let $I = (a, b)$ be an open interval of $\mathbb{R}$. We define a curve on $\mathcal{M}$ as a map

$$\gamma : \mathbb{R} \supset I \to \mathcal{M}.$$

The curve γ is smooth if its image

$$\phi \circ \gamma : I \to \mathbb{R}^n, \quad \phi \circ \gamma(s) = \left(x^1(\gamma(s)), \ldots, x^n(\gamma(s))\right)$$

is smooth.

2.3 The Tangent Space as a Space of Directional Derivatives

We wish to construct the tangent space at $p \in \mathcal{M}$ (i.e. $T_p(\mathcal{M})$) using only intrinsic properties of $\mathcal{M}$.

We combine the concept of smooth functions, f, and smooth curves, γ, and define

$$\begin{aligned} \mathcal{F} : I \to \mathbb{R}, \quad s \mapsto \mathcal{F}(s) &= f \circ \gamma(s) \\ &\equiv f\left(\gamma(s)\right), \end{aligned}$$

i.e. $\mathcal{F}$ evaluates f along the curve γ.

The rate at which $\mathcal{F}$ changes, $\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}s}$, gives the rate of change of f following the curve γ.

The tangent vector to the curve γ at p (where without loss of generality we take $s = 0$ at p) is the map from the set of real functions to $\mathbb{R}$, defined by

$$\begin{aligned} \dot{\gamma}_p : f \mapsto \dot{\gamma}_p f \equiv \dot{\gamma}_p(f) &= \left[\frac{\mathrm{d}}{\mathrm{d}s} f \circ \gamma\right]_{s=0} \\ &\equiv \left(\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}s}\right)_{s=0}. \end{aligned}$$

Claim. Let ϕ be a chart such that

$$\phi : p \mapsto x^\mu(p).$$

Then

$$\dot{\mathcal{F}}(0) = \left[\frac{\mathrm{d}}{\mathrm{d}s}\left(f \circ \gamma\right)\right]_{s=0} = \sum_{\mu=1}^{n} \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{\mathrm{d}}{\mathrm{d}s} x^\mu\left(\gamma(s)\right)\right]$$

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

Proof.

$$\begin{aligned}\mathcal{F}(s) = f\circ\gamma &= f\circ\phi^{-1}\circ\phi\circ\gamma \\ &= F\circ\phi\circ\gamma\end{aligned}$$

where the function $\phi\circ\gamma$ maps s to the coordinates of $\gamma(s)$. Identify

$$\mathcal{F}(s) = F\left(x^1(\gamma(s)),\ldots,x^n(\gamma(s))\right).$$

Then

$$\begin{aligned}\dot{\gamma}_p f &= \left[\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}s}\right]_{s=0} = \left[\frac{\mathrm{d}F\left(x^1(\gamma(s)),\ldots,x^n(\gamma(s))\right)}{\mathrm{d}s}\right]_{s=0} \\ &= \left(\frac{\partial F}{\partial x^1}\right)_{\phi(p)}\left[\frac{\mathrm{d}x^1(\gamma(s))}{\mathrm{d}s}\right]_{s=0} + \cdots + \left(\frac{\partial F}{\partial x^n}\right)_{\phi(p)}\left[\frac{\mathrm{d}x^n(\gamma(s))}{\mathrm{d}s}\right]_{s=0} \\ &= \sum_{\mu=1}^{n}\left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}\left[\frac{\mathrm{d}x^\mu(\gamma(s))}{\mathrm{d}s}\right]_{s=0}. \qquad \text{(Or use Einstein notation.)}\end{aligned}$$

□

Example 2.3.1. Let $\mathcal{M} = \mathbb{R}^2$. Take $y = 2x^2 - 3$ to be a parabola in $\mathbb{R}^2$. We parametrise this by $x = s$, $y = 2s^2 - 3$.

$$\begin{aligned}\phi\circ\gamma(s) = (x(s), y(s)) &= \left(s, 2s^2-3\right), \\ \mathcal{F}(s) &= F(s, 2s^2-3).\end{aligned}$$

$$\begin{aligned}\frac{\mathrm{d}\mathcal{F}(s)}{\mathrm{d}s} &= \frac{\partial F}{\partial x}\cdot\frac{\mathrm{d}x}{\mathrm{d}s} + \frac{\partial F}{\partial y}\cdot\frac{\mathrm{d}y}{\mathrm{d}s} \\ &= \frac{\partial F}{\partial x}\cdot 1 + \frac{\partial F}{\partial y}\cdot 4s \\ &= T\cdot\nabla F, \quad T = (1, 4s).\end{aligned}$$

$T\cdot\nabla F$ is the rate of change of F in the direction of the vector T.

The map $\dot{\gamma}_p : f \mapsto \left[\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}s}\right]_{s=0}$ we called a tangent vector at p. We must further show that these maps live in a vector space of dimension $n = \dim(\mathcal{M})$.

Theorem 2.3.1. The set of tangent vectors at p, $T_p(\mathcal{M})$, form a vector space (i.e. are closed under addition and scalar multiplication).

Proof. $X_p, Y_p \in T_p(\mathcal{M})$. We want to show that

$$X_p + Y_p \in T_p(\mathcal{M}), \qquad \alpha X_p \in T_p(\mathcal{M}).$$

With X_p, Y_p we associate smooth curves $\gamma(s)$, $\lambda(s)$ such that $\gamma(0) = \lambda(0) = p$,

$$X_p = \dot{\gamma}_p, \quad Y_p = \dot{\lambda}_p.$$

Two curves in $\mathcal{M}$ may be added by considering their coordinate image in $\mathbb{R}^n$,

$$\{x^\mu(\gamma(s))\}, \quad \{x^\mu(\lambda(s))\},$$

which gives a new curve

$$\tilde{\nu} : s \mapsto \{x^\mu(\gamma(s)) + x^\mu(\lambda(s)) - x^\mu(p)\} \qquad (\tilde{\nu} = \phi \circ \nu)$$

which is a parametric representation in $\mathbb{R}^n$ of some curve in $\mathcal{M}$ with $\nu(0) = p$.

For any function f, we have

$$\begin{aligned}
\dot{\nu}_p f &= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{\mathrm{d}}{\mathrm{d}s} x^\mu(\nu(s))\right]_{s=0} \\
&= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left\{\left[\frac{\mathrm{d}x^\mu(\gamma(s))}{\mathrm{d}s}\right]_{s=0} + \left[\frac{\mathrm{d}x^\mu(\lambda(s))}{\mathrm{d}s}\right]_{s=0}\right\} \\
&= \dot{\gamma}_p f + \dot{\lambda}_p f = X_p f + Y_p f.
\end{aligned}$$

Since f was arbitrary, we associate $\dot{\nu}_p = X_p + Y_p$. Therefore the space is closed under addition. A similar proof holds for αX_p (consider the coordinate image). Therefore, the maps $\dot{\gamma}_p$ form a vector space. □

To recap:

$$\begin{aligned}
\dot{\gamma}_p : f \mapsto \left(\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}s}\right)_{s=0} \\
= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{\mathrm{d}x^\mu(\gamma(s))}{\mathrm{d}s}\right]_{s=0}
\end{aligned}$$

and $\dot{\gamma}_p$ is closed under addition and scalar multiplication, i.e. they form a vector space.

Theorem 2.3.2. $\dim(T_p(\mathcal{M})) = \dim(\mathcal{M})$.

Proof. We consider a chart ϕ,

$$\phi : p \mapsto x^\mu(p).$$

We assume that $x^\mu(p) = 0$ for all μ (p gets mapped to the origin of $\mathbb{R}^n$). We consider curves $\gamma_\nu(s)$ through p such that its coordinate image in $\mathbb{R}^n$ is

$$\tilde{\gamma}_\nu(s) = (0, \ldots, 0, \underset{\substack{\uparrow \\ \nu^{\text{th}}\text{ component}}}{s}, 0, \ldots, 0).$$

Obviously, there are n such curves. They satisfy

$$x^\mu \circ \gamma_\nu(s) = \begin{cases} s & \mu = \nu \\ 0 & \text{otherwise} \end{cases} .$$

Therefore,

$$\left[\frac{\mathrm{d}x^\mu \circ \gamma_\nu(s)}{\mathrm{d}s}\right]_p = \delta^\mu_\nu.$$

Furthermore,

$$\begin{aligned}(\dot{\gamma}_\nu) \circ f &= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{\mathrm{d}x^\mu \circ \gamma_\nu(s)}{\mathrm{d}s}\right]_{s=0} \\ &= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \delta^\mu_\nu \\ &= \left(\frac{\partial F}{\partial x^\nu}\right)_{\phi(p)} ,\end{aligned}$$

i.e. $(\dot{\gamma}_\nu)_p$ maps

$$f \mapsto \left(\frac{\partial F}{\partial x^\nu}\right)_{\phi(p)} .$$

We will show that $\{(\dot{\gamma}_\nu)_p\}$ form a basis for $T_p(\mathcal{M})$ and, since there are n such vectors, this will prove

$$n = \dim\left(T_p(\mathcal{M})\right) = \dim\left(\mathcal{M}\right).$$

We need to show that $\{(\dot{\gamma}_\nu)_p\}$ span $T_p(\mathcal{M})$ and are linearly independent.

To determine the span, let $\dot{\lambda}_p \in T_p(\mathcal{M})$. Then

$$\begin{aligned}\dot{\lambda}_p(f) &= \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \left[\frac{\mathrm{d}x^\mu \circ \lambda(s)}{\mathrm{d}s}\right]_{s=0} \\ &= (\dot{\lambda}_\mu)_p(f) c^\mu, \quad c^\mu = \left[\frac{\mathrm{d}x^\mu \circ \lambda(s)}{\mathrm{d}s}\right]_{s=0} .\end{aligned}$$

This is true for all f, hence $\dot{\lambda}_p = c^\mu(\dot{\lambda}_\mu)_p$ and therefore $\{(\dot{\lambda}_\mu)_p\}$ span $T_p(\mathcal{M})$ (i.e. any element of $T_p(\mathcal{M})$ can be written as a linear combination of $\{(\dot{\lambda}_\mu)_p\}$).

For linear independence,

$$a^\mu(\dot{\gamma}_\mu)_p = 0.$$

Consider the function $x^\nu : \mathcal{M} \to \mathbb{R}^n$.

$$\begin{aligned} a^\mu (\dot{\gamma}_\mu)_p x^\nu &= 0, \\ a^\mu \left(\frac{\partial}{\partial x^\mu} \right)_p x^\nu &= 0, \\ a^\mu \delta^\nu_\mu &= 0, \\ a^\nu &= 0 \end{aligned}$$

and therefore $\{(\dot{\gamma}_\mu)_p\}$ are linearly independent.

Hence, $\{(\dot{\gamma}_\mu)_p\}$ form a basis for $T_p(\mathcal{M})$, so

$$\dim(T_p(\mathcal{M})) = \dim(\mathcal{M}) = n. \qquad \square$$

We usually write

$$(\dot{\gamma}_\mu)_p = \left(\frac{\partial}{\partial x^\mu} \right)_p = (\partial_\mu)_p \, .$$

This is known as a coordinate basis for the tangent space $T_p(\mathcal{M})$. It corresponds to setting up the basis vectors so that they point along the coordinate axes.

Note that a change of coordinates induces a change of basis.

2.4 Transformation Rule for Vector Components

We have shown that every element of the tangent space is naturally decomposed into a coordinate basis, i.e. if

$$V \in T_p(\mathcal{M}), \quad V = V^\mu \left(\frac{\partial}{\partial x^\mu} \right)_p , \quad V^\mu = \left[\frac{\mathrm{d}x^\mu(s)}{\mathrm{d}s} \right]_{s=0}$$

where V^μ are the components of the vector in the coordinate basis.

Introduce a change of coordinates

$$x^{\mu'} = x^{\mu'}(x^\mu)$$

(new coordinates as a function of old ones).

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

Comparing, we see that

$$V^\mu = V^{\mu'} \frac{\partial x^\mu}{\partial x^{\mu'}} \quad \Rightarrow \quad V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu .$$

This is the vector coordinate transformation rule in a coordinate basis.

Note that vectors, tensors, etc. are invariant under a change of coordinates or a change of basis, but their components are not.

We are not limited to a coordinate basis for $T_p(\mathcal{M})$. Any linearly independent set $\{e_\mu\}$ that spans the vector space is an appropriate basis.

$$V = V^\mu e_\mu$$

where V^μ are the components with respect to the basis $\{e_\mu\}$. This basis is related to any other basis $\{e_{\mu'}\}$ by a non-singular set of transformations.

$$e_{\mu'} = \Lambda^\mu{}_{\mu'} e_\mu.$$

Since $\Lambda^\mu{}_{\mu'}$ is non-singular, it is invertible and satisfies

$$\Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu = \delta^\mu_\nu \qquad (\Lambda^{-1}\Lambda = \mathbb{1}).$$

We then have

$$V = V^\mu e_\mu = V^{\mu'} e_{\mu'} = V^{\mu'} \Lambda^\mu{}_{\mu'} e_\mu,$$

giving the vector component transformation law in an arbitrary basis

$$V^{\mu'} = \Lambda^{\mu'}{}_\mu V^\mu.$$

For the coordinate basis, $\Lambda^{\mu'}{}_\mu = \frac{\partial x^{\mu'}}{\partial x^\mu}$.

Note that if $\Lambda^{\mu'}{}_\nu$ are further constrained to satisfy

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

where $\eta = \text{diag}(-1, 1, 1, 1)$, then Λ represents the Lorentz transformations of special relativity.

3 Covectors and Tensors

3.1 Co-Tangent Space

All vector spaces have a corresponding dual space of equal dimension, comprising the set of linear maps from the vector space to the real line.

The dual space of $T_p(\mathcal{M})$ we denote by $T_p^*(\mathcal{M})$, the co-tangent space consisting of maps

$$\eta : T_p(\mathcal{M}) \to \mathbb{R}.$$

Elements of $T_p^*(\mathcal{M})$ are known as covectors or 1-forms.

The action of a linear map η on $X_p \in T_p(\mathcal{M})$ is

$$\eta(X_p) = \langle \eta, X_p \rangle.$$

For a tangent space $T_p(\mathcal{M})$ with basis $\{e_\mu\}$, there is a corresponding dual basis of $T_p^*(\mathcal{M})$ satisfying

$$\omega^\mu(e_\nu) = \langle \omega^\mu, e_\nu \rangle = \delta^\mu_\nu.$$

3.2 The Gradient/Differential

Perhaps the most important example of a covector is the gradient or differential of a function f. Let $p \in \mathcal{M}$ and $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. For $X_p \in T_p(\mathcal{M})$, we define $\mathrm{d}f$ to be

$$\mathrm{d}f : T_p(\mathcal{M}) \to \mathbb{R}, \quad \mathrm{d}f(X_p) = X_p(f).$$

3.3 Coordinate Induced Dual Basis

For a particular chart, we may associate with the coordinate x^μ the covector $\mathrm{d}x^\mu$ defined by

$$\mathrm{d}x^\mu(X_p) = X_p(x^\mu).$$

Claim. The set $\{\mathrm{d}x^\mu\}$ form a basis for $T_p^*(\mathcal{M})$ dual to the coordinate basis $\left\{\left(\frac{\partial}{\partial x^\mu}\right)\right\}$ of $T_p(\mathcal{M})$.

Proof.

1. Duality.

$$\langle \mathrm{d}x^\mu, \partial_\nu \rangle = \left(\frac{\partial}{\partial x^\nu}\right)_p x^\mu = \delta^\mu_\nu.$$

2. $\{\mathrm{d}x^\mu\}$ form a basis for $T_p^*(\mathcal{M})$.

$$\begin{aligned} \eta &= \eta_\mu \mathrm{d}x^\mu = 0. \\ 0 &= \eta\left(\left(\frac{\partial}{\partial x^\nu}\right)_p\right) \\ &= \eta_\mu \mathrm{d}x^\mu\left(\left(\frac{\partial}{\partial x^\nu}\right)_p\right) \\ &= \eta_\mu \delta^\mu_\nu = \eta_\nu \end{aligned}$$

which implies that $\{\mathrm{d}x^\mu\}$ are linearly independent.

3. Take $\omega \in T_p^*(\mathcal{M})$ and set

$$\eta = \omega - \langle \omega, (\partial_\mu)_p \rangle \, \mathrm{d}x^\mu.$$

Then

$$\begin{aligned}
\eta_\nu &= \langle \eta, (\partial_\nu)_p \rangle \\
&= \langle \omega, (\partial_\nu)_p \rangle - \langle \langle \omega, (\partial_\mu)_p \rangle \, \mathrm{d}x^\mu, (\partial_\nu)_p \rangle \\
&= \langle \omega, (\partial_\nu)_p \rangle - \langle \omega, (\partial_\mu)_p \rangle \, \delta^\mu_\nu = 0
\end{aligned}$$

and

$$\eta = 0 \quad \Rightarrow \quad \omega = \langle \omega, (\partial_\mu)_p \rangle \, \mathrm{d}x^\mu.$$

Therefore $\{\mathrm{d}x^\mu\}$ spans $T_p^*(\mathcal{M})$.

$\{\mathrm{d}x^\mu\}$ forms a basis of $T_p^*(\mathcal{M})$ dual to $\left\{ \left(\frac{\partial}{\partial x^\mu} \right)_p \right\}$. □

Since

$$\omega = \langle \omega, (\partial_\mu)_p \rangle \, \mathrm{d}x^\mu,$$

we notice that

$$\begin{aligned}
\mathrm{d}f &= \langle \mathrm{d}f, (\partial_\mu)_p \rangle \, \mathrm{d}x^\mu \\
&= \left(\frac{\partial}{\partial x^\mu} \right)_p f \, \mathrm{d}x^\mu \\
&= \left(\frac{\partial F}{\partial x^\mu} \right)_{\phi(p)} \mathrm{d}x^\mu. \qquad (\circledast)
\end{aligned}$$

The components of the gradient in the basis $\{\mathrm{d}x^\mu\}$ are partial derivatives of the function.

3.4 Transformation Rule for Covector Components

As before, a coordinate transformation induces a change in coordinate basis of $T_p^*(\mathcal{M})$. We introduce

$$x^{\mu'} = x^{\mu'}(x^\mu).$$

From $(\circledast)$,

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

For $\eta \in T_p^*(\mathcal{M})$,

$$\eta = \eta_\mu \mathrm{d}x^\mu = \eta_{\mu'} \mathrm{d}x^{\mu'} = \eta_{\mu'} \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right) \mathrm{d}x^\mu.$$

$$\eta_\mu = \eta_{\mu'} \left(\frac{\partial x^{\mu'}}{\partial x^\mu}\right), \quad \eta_{\mu'} = \left(\frac{\partial x^\mu}{\partial x^{\mu'}}\right) \eta_\mu.$$

This is the covector transformation rule in a coordinate-induced basis.

As before, we are not restricted to a coordinate basis. In general,

$$\eta_{\mu'} = \Lambda^\mu{}_{\mu'} \eta_\mu, \quad \Lambda^\mu{}_{\mu'} \Lambda^{\mu'}{}_\nu = \delta^\mu_\nu,$$

the transformation rule in an arbitrary basis.

3.5 Transformation Rule for Tensor Components

We can generalise vectors and covectors to define the notion of a tensor, which is a multilinear map

$$S : \underbrace{T_p(\mathcal{M}) \times \cdots \times T_p(\mathcal{M})}_{s} \times \underbrace{T_p^*(\mathcal{M}) \times \cdots \times T_p^*(\mathcal{M})}_{r} \to \mathbb{R}.$$

This is known as a tensor of rank (or type) $\binom{r}{s}$ at p.

We define the tensor product operation. If T_1 is an $\binom{r_1}{s_1}$ tensor and T_2 is an $\binom{r_2}{s_2}$ tensor then the tensor product is defined by

$$\begin{aligned} T_1 \otimes T_2(x_1, \ldots, x_{s_1+s_2}, \eta_1, \ldots, \eta_{r_1+r_2}) &= T_1(x_1, \ldots, x_{s_1}, \eta_1, \ldots \eta_{r_1}) \\ &\quad \cdot T_2(x_{s_1+1}, \ldots, x_{s_1+s_2}, \eta_{r_1+1}, \ldots, \eta_{r_1+r_2}). \end{aligned}$$

An appropriate coordinate basis for an arbitrary $\binom{r}{s}$ tensor is

$$\left\{\mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_s} \otimes \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_r}\right\}.$$

In this basis, an arbitrary tensor

$$T = T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s} \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_s} \otimes \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_r}.$$

It is now easy to see that under a coordinate transformation, the components change according to

$$T^{\mu'_1 \cdots \mu'_r}{}_{\nu'_1 \cdots \nu'_s} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu'_r}}{\partial x^{\mu_r}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} \cdots \frac{\partial x^{\nu_s}}{\partial x^{\nu'_s}} \cdot T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s}.$$

This is the tensor component transformation rule in a coordinate basis.

Note that

- a $\binom{1}{0}$ tensor is a vector,
- a $\binom{0}{1}$ tensor is a covector (1-form),
- we can't add tensors of different types (ranks).

For example, the Kronecker delta δ^μ_ν transforms like the components of a $\binom{1}{1}$ tensor.

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

From an $\binom{r}{s}$ tensor, we can obtain an $\binom{r-1}{s-1}$ tensor by contraction. This corresponds to multiplication by a Kronecker delta. E.g.

$$T^\mu{}_{\alpha\beta\gamma} \xrightarrow{\text{contraction}} T^\mu{}_{\alpha\mu\gamma} = \delta^\beta_\mu T^\mu{}_{\alpha\beta\gamma}.$$

We introduce the Bach bracket notation:

(ab) represents symmetrisation of indices,
$[ab]$ represents antisymmetrisation of indices.

If S is a $\binom{1}{2}$ tensor, then

$$S^\mu{}_{(\nu\lambda)} = \frac{1}{2!}\left(S^\mu{}_{\nu\lambda} + S^\mu{}_{\lambda\nu}\right),$$
$$S^\mu{}_{[\nu\lambda]} = \frac{1}{2!}\left(S^\mu{}_{\nu\lambda} - S^\mu{}_{\lambda\nu}\right).$$

In general, we have

$$T_{(\mu_1\cdots\mu_r)} = \frac{1}{r!}(\text{sum over all permutations of indices}),$$
$$T_{[\mu_1\cdots\mu_r]} = \frac{1}{r!}(\text{sum over all permutations with the sign of the permutation}).$$

For example,

$$T_{[\alpha\beta\gamma]} = \frac{1}{6}\left(T_{\alpha\beta\gamma} - T_{\alpha\gamma\beta} + T_{\gamma\alpha\beta} - T_{\gamma\beta\alpha} + T_{\beta\gamma\alpha} - T_{\beta\alpha\gamma}\right).$$

4 Tensor Fields and the Commutator

4.1 The Tangent Bundle and Vector Fields

The union of all tangent vector spaces for each point of the manifold defines the tangent bundle, $T(\mathcal{M})$.

$$T(\mathcal{M}) = \bigcup_{p\in\mathcal{M}} T_p(\mathcal{M}).$$

Then a vector field on $\mathcal{M}$ is a map that specifies one vector at each point of the manifold.

$$X : \mathcal{M} \to T(\mathcal{M}), \quad p \mapsto X_p \in T_p(\mathcal{M}).$$

In a coordinate chart $\{x^\mu\}$, we take X_p to be

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

A similar definition holds for covectors and tensors.

If X is a vector field and f is some function, we define the map $X(f)$ which is a function such that

$$X(f) : \mathcal{M} \to \mathbb{R}, \quad p \mapsto X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}.$$

This is a linear map,

$$X(af + bg) = aX(f) + bX(g), \qquad a, b \in \mathbb{R}, \quad f, g : \mathcal{M} \to \mathbb{R},$$

and satisfies the Leibniz rule,

$$X(f \cdot g) = X(f) \cdot g + f \cdot X(g).$$

4.2 Commutator

Let X, Y be vector fields and f an arbitrary function. We can consider the composition

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

The Lie bracket is defined as

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

Claim. The Lie bracket is a vector field.

Proof. We just need to check that it satisfies linearity and the Liebniz rule.

1. $$[X, Y]\,(af + bg) = a\,[X, Y]\,(f) + b\,[X, Y]\,(g).$$

   This is trivial.

2. $$\begin{aligned}[X,Y]\,(f\cdot g) &= g[X,Y](f)+f[X,Y](g)\\ &= X(Y(f\cdot g))-Y(X(f\cdot g))\\ &= X(Y(f)\cdot g+f\cdot Y(g))-Y(X(f)\cdot g+f\cdot X(g))\\ &= X(Y(f))\cdot g+Y(f)\cdot X(g)+X(f)\cdot Y(g)\\ &\quad +f\cdot X(Y(g))-Y(X(f))\cdot g-X(f)\cdot Y(g)\\ &\quad -Y(f)\cdot X(g)-f\cdot Y(X(g))\\ &= g[X,Y](f)+f[X,Y](g).\end{aligned}$$ □

Since it is a vector field, it has the following representation (in a coordinate basis):

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

This implies that

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

Comparing (A) with (B), we see that

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

These are the components of the commutator in a coordinate-induced basis.

The commutator satisfies the following properties:

1. $[X,X]=0$,
2. $[X,Y]=-[Y,X]$,
3. $[X,Y+Z]=[X,Y]+[X,Z]$,
4. $[X,fY]=f[X,Y]+X(f)Y, \quad f:\mathcal{M}\to\mathbb{R}$,
5. $\Big[X,[Y,Z]\Big]+\Big[Y,[Z,X]\Big]+\Big[Z,[X,Y]\Big]=0$, the Jacobi identity.

5 Maps Between Manifolds

We now consider maps between two manifolds $\mathcal{M}$ and $\mathcal{N}$ or, in most cases, between $\mathcal{M}$ and itself.

We say a map $h : \mathcal{M} \to \mathcal{N}$ is smooth if, for every smooth function $f : \mathcal{N} \to \mathbb{R}$, the function $f \circ h : \mathcal{M} \to \mathbb{R}$ is also smooth.

$$\begin{tikzcd} \mathcal{M} \arrow[r, "h"] \arrow[dr, "f\circ h"'] & \mathcal{N} \arrow[d, "f"] \\ & \mathbb{R} \end{tikzcd}$$

Assuming h is smooth, then it maps a smooth curve γ in $\mathcal{M}$ to a smooth curve $h \circ \gamma$ in $\mathcal{N}$.

Let X_p be a tangent vector to γ at p. Then there exists a map

$$h_* : T_p\left(\mathcal{M}\right) \to T_{h(p)}(\mathcal{N})$$

known as the push-forward to $h \circ \gamma$ and it maps the tangent vector at p in $\mathcal{M}$ to the tangent vector at $h(p)$ in $\mathcal{N}$.

One can show that h_* is a linear map such that for every smooth function $f : \mathcal{N} \to \mathbb{R}$,

$$(h_* X_p)(f) = X_p(f \circ h).$$

Analogously, there is a map between covectors that goes in the opposite direction,

$$h^* : T^*_{h(p)}(\mathcal{N}) \to T^*_p\left(\mathcal{M}\right),$$

known as the pull-back. This maps the tangent covectors at $h(p)$ in $\mathcal{N}$ to the tangent covectors at p in $\mathcal{M}$.

If $\eta \in T^*_{h(p)}(\mathcal{N})$ and X_p is any vector in $T_p\left(\mathcal{M}\right)$, then $h^*\eta \in T^*_p\left(\mathcal{M}\right)$ is defined by

$$h^*\eta(X_p) = \eta(h_* X_p).$$

The pull-back of the covector acting on a vector is the same as the covector acting on the push-forward of the vector. (Again, $h^*\eta$ is a linear map since the right hand side is linear in X_p.)

Example 5.1. Let $\mathcal{M} = \mathbb{R}^3$ with coordinates (x, y, z) and $\mathcal{N} = \mathbb{R}^2$ with coordinates (x, y). Let h be the map $h(x, y, z) = (x, y)$. Then for any $f : \mathcal{N} \to \mathbb{R}$,

$$\begin{aligned} f \circ h(x, y, z) &= f(h(x, y, z)) \\ &= f(x, y). \end{aligned}$$

Suppose X_p has coordinates (X, Y, Z). Then

$$X_p(g) = X\frac{\partial g}{\partial x} + Y\frac{\partial g}{\partial y} + Z\frac{\partial g}{\partial z}, \quad g : \mathcal{M} \to \mathbb{R}$$

and

$$(h_* X_p)(g) = X_p(g \circ h) = X_p(g(x, y)) = X\frac{\partial g}{\partial x} + Y\frac{\partial g}{\partial y},$$

i.e. $h_* X_p$ has components (X, Y). If $\eta \in T^*_{h(p)}(\mathcal{N})$ such that

$$\eta = \lambda \, \mathrm{d}x + \mu \, \mathrm{d}y,$$

then

$$\begin{aligned}
h^*\eta(X_p) = \eta(h_* X_p) &= \eta(X\partial_x + Y\partial_y) \\
&= \langle \lambda \, \mathrm{d}x + \mu \, \mathrm{d}y, X\partial_x + Y\partial_y \rangle \\
&= \lambda X + \mu Y \\
&= (\lambda \, \mathrm{d}x + \mu \, \mathrm{d}y + 0 \, \mathrm{d}z) X_p,
\end{aligned}$$

i.e. components of $h^*\eta$ are $(\lambda, \mu, 0)$.

The definition of h^* as a map from tangent vectors in $\mathcal{N}$ to tangent vectors in $\mathcal{M}$ naturally extends to functions. If $f : \mathcal{N} \to \mathbb{R}$, then $h^* f$ is defined by

$$h^* f : \mathcal{M} \to \mathbb{R}, \quad f \mapsto f \circ h.$$

Claim. h^* commutes with the gradient, i.e.

$$h^*(\mathrm{d}f) = \mathrm{d}(h^* f).$$

Proof. Let $\mathrm{d}f \in T^*_{h(p)}(\mathcal{N})$ and take $X_p \in T_p(\mathcal{M})$.

$$\begin{aligned}
(h^*\mathrm{d}f)(X_p) &= \mathrm{d}f(h_* X_p) \\
&= (h_* X_p) f \\
&= X_p(f \circ h) \\
&= X_p(h^* f) \\
&= \mathrm{d}(h^* f)(X_p)
\end{aligned}$$

and therefore

$$h^*\mathrm{d}f = \mathrm{d}(h^* f).$$

□

6 Lie Derivatives and the Commutator Revisited

6.1 Integral Curves

Let X be a vector field on $\mathcal{M}$. An integral curve of X in $\mathcal{M}$ is a curve γ such that at each point p of γ, the tangent vector is X_p.

$\gamma(s)$ is an integral curve iff

$$\dot{\gamma}_p(s) = X_{p=\gamma(s)}.$$

To see existence and uniqueness of such curves (at least locally), we consider a chart $\{x^\mu\}$ and a test function f. Then

$$\begin{aligned}\dot{\gamma}_p(f) = X_p(f) \quad &\Leftrightarrow \quad \left(\frac{\mathrm{d}}{\mathrm{d}s}(f \circ \gamma)\right)_p = X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \\ &\Leftrightarrow \quad \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)} \frac{\mathrm{d}x^\mu\left(\gamma(s)\right)}{\mathrm{d}s} = X_p^\mu \left(\frac{\partial F}{\partial x^\mu}\right)_{\phi(p)}.\end{aligned}$$

Since this is true for arbitrary f,

$$\frac{\mathrm{d}x^\mu(\gamma(s))}{\mathrm{d}s} = X_p^\mu \left(x^1(\gamma(s)), \ldots, x^n(\gamma(s))\right). \tag{C}$$

We also have a set of initial conditions

$$x^\mu(\gamma(s=0)) = x^\mu(p).$$

Note that there exists a chart $\{x^\mu\}$ in a neighbourhood of p such that $X = \frac{\partial}{\partial x^1} = (1, 0, \ldots, 0)$ corresponding to integral curves of varying x^1 and constant $x^2, \ldots, x^n$.

6.2 Congruence of Curves

If equation C is globally valid (holds for all s) for an integral curve, we say the curve is complete.

The set of complete curves is a congruence (one at each spacetime point).

Given a congruence, we may define a 1-parameter family of transformations

$$h_s : \mathcal{M} \to \mathcal{M}$$

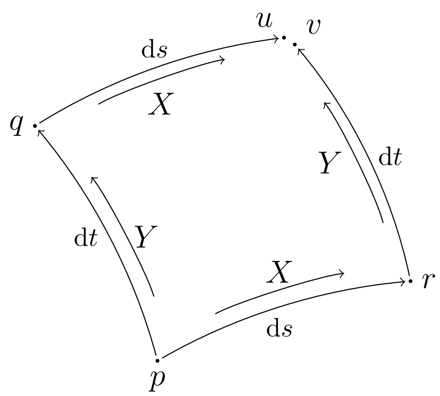
such that $h_s(p)$ is a point on the integral curve through p a parameter "distance" s from p. Then

$$h_s(h_t(p)) = h_{s+t}(p) = h_t(h_s(p)).$$

Clearly we have an identity map h_0 such that $h_0(p) = p$ and we have an inverse map $h_s^{-1} = h_{-s}$. Together these properties show that the maps h_s form an Abelian group of transformations $\mathcal{M} \to \mathcal{M}$.

6.3 The Commutator Revisited: A Geometric Interpretation

We let X, Y be vector fields on $\mathcal{M}$ with groups h_s, k_t respectively. Starting from a point p, we move a distance $\mathrm{d}s$ along the integral curve of X to a point $r = h_{\mathrm{d}s}(p)$ followed by moving a distance $\mathrm{d}t$ along the integral curve of Y to reach some point $v = k_{\mathrm{d}t}(r)$. Starting again from p, we now move a distance $\mathrm{d}t$ along the integral curve of Y to the point $q = k_{\mathrm{d}t}(p)$ followed by a distance $\mathrm{d}s$ along the integral curve of X to the point $u = h_{\mathrm{d}s}(q)$.



What can we say about the difference $x^\mu(v) - x^\mu(u)$? We choose a chart such that $\{x^\mu(p)\} = 0$. Using a Taylor expansion,

$$x^\mu(q) = x^\mu(p) + \left(\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\right)_p \mathrm{d}t + \frac{1}{2}\left(\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2}\right)_p \mathrm{d}t^2 + \cdots .$$

But $x^\mu(p) = 0$. Let $Y_p^\mu = \left(\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\right)_p$, so

$$\begin{aligned}\left(\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2}\right)_p = \left(\frac{\mathrm{d}Y^\mu}{\mathrm{d}t}\right)_p &= \left(\frac{\partial Y^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}\right)_p \\ &= \left(\frac{\partial Y^\mu}{\partial x^\nu} Y^\nu\right)_p \\ &= \left(Y^\mu{}_{,\nu} Y^\nu\right)_p\end{aligned}$$

where $\partial_\nu Y^\mu = Y^\mu{}_{,\nu}$. This gives

$$x^\mu(q) = Y^\mu_p \mathrm{d}t + \frac{1}{2}(Y^\mu{}_{,\nu} Y^\nu)_p \mathrm{d}t^2 + \cdots .$$

Furthermore, we have

$$\begin{aligned}
x^\mu(u) &= x^\mu(q) + \left(\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\right)_q \mathrm{d}s + \frac{1}{2}\left(\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2}\right)_q \mathrm{d}s^2 + \cdots \\
&= x^\mu(q) + X^\mu_q \mathrm{d}s + \frac{1}{2}(X^\mu{}_{,\nu} X^\nu)_q \mathrm{d}s^2 + \cdots \\
&= \left[Y^\mu_p \mathrm{d}t + \frac{1}{2}(Y^\mu{}_{,\nu} Y^\nu)_p \mathrm{d}t^2 + \cdots\right] \\
&\quad + \left[X^\mu_p + \left(\frac{\mathrm{d}X^\mu}{\mathrm{d}t}\right)_p \mathrm{d}t + \cdots\right] \mathrm{d}s + \frac{1}{2}(X^\mu{}_{,\nu} X^\nu)_p \mathrm{d}s^2 + \cdots .
\end{aligned}$$

Noticing that $\left(\frac{\mathrm{d}X^\mu}{\mathrm{d}t}\right)_p = \left(\frac{\partial X^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}\right)_p = (X^\mu{}_{,\nu} Y^\nu)_p$,

$$x^\mu(u) = Y^\mu_p \mathrm{d}t + X^\mu_p \mathrm{d}s + \frac{1}{2}(Y^\mu{}_{,\nu} Y^\nu)_p \mathrm{d}t^2 + (X^\mu{}_{,\nu} Y^\nu)_p \mathrm{d}s\,\mathrm{d}t + \frac{1}{2}(X^\mu{}_{,\nu} X^\nu)_p \mathrm{d}s^2 + \cdots .$$

Similarly, by interchanging $X \leftrightarrow Y$, $s \leftrightarrow t$, we have

$$x^\mu(v) = X^\mu_p \mathrm{d}s + Y^\mu_p \mathrm{d}t + \frac{1}{2}(X^\mu{}_{,\nu} X^\nu)_p \mathrm{d}s^2 + (Y^\mu{}_{,\nu} X^\nu)_p \mathrm{d}s\,\mathrm{d}t + \frac{1}{2}(Y^\mu{}_{,\nu} Y^\nu)_p \mathrm{d}t^2 + \cdots .$$

Subtracting, we get

$$\begin{aligned}
x^\mu(v) - x^\mu(u) &= \left(Y^\mu{}_{,\nu} X^\nu - X^\mu{}_{,\nu} Y^\nu\right)_p \mathrm{d}s\,\mathrm{d}t + \cdots \\
&= [X, Y]^\mu_p \,\mathrm{d}s\,\mathrm{d}t + \cdots .
\end{aligned}$$

The commutator $[X, Y]$ measures the discrepancy between the points u and v obtained by following the integral curves of the vector fields X and Y in different orders, starting from p and moving infinitesimal distances along the curves. Now $[X, Y] = 0$ implies $u = v$.

We say a basis $\{e_\mu\}$ is coordinate-induced if $[e_\mu, e_\nu]_p = 0$ for all μ, ν. For the coordinate-induced basis $\{\partial_\mu\}$,

$$[\partial_\mu, \partial_\nu]_p f = \frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 f}{\partial x^\nu \partial x^\mu} = 0.$$

6.4 Lie Derivatives

Claim 6.4.1. Suppose X is a smooth vector field and $X_p \neq 0$. Then there is a coordinate chart $\{y^\mu\}$ defined in a neighbourhood U of p such that $X = \frac{\partial}{\partial y^1}$ in U, i.e. $X = (1, 0, \ldots, 0)$.

Proof. (For $n = 4$.) By continuity, there exists a neighbourhood U of p such that $X \neq 0$ in U. Choose a 3-surface Σ in U nowhere tangent to X and arbitrary coordinates (y^2, y^3, y^4) on Σ. There exists at each point of Σ a unique integral curve with parameter t through each point, where $t = 0$ at that point. Define $y^1 = \int \mathrm{d}t$ along such curves and define y^2, y^3, y^4 on these curves to be constant. Then $\{y^\mu\}$ is the required chart. $\square$

6.5 Lie Derivatives of a Function

The Lie derivative of a function with respect to a vector field X at p is

$$(\mathcal{L}_X f)_p = \lim_{\mathrm{d}t \to 0} \left[\frac{f(h_{\mathrm{d}t}(p)) - f(p)}{\mathrm{d}t} \right].$$

To rewrite this in a more useful form, we introduce the chart of claim 6.4.1, $\{y^\mu\}$, such that the vector field $X = (1, 0, \ldots, 0)$. $h_{\mathrm{d}t}(p) = (y^1 + \mathrm{d}t, y^2, \ldots, y^n)$ in coordinates $\{y^\mu\}$ (only moving in y^1). Then the right hand side of our definition evaluates to

$$\left(\frac{\partial F}{\partial y^1} \right)_{\phi(p)} = X_p f, \qquad (\mathcal{L}_X f)_p = X_p f.$$

This is the rate of change of f along the integral curve X.

If two scalars are equal in one chart, they are equal in all charts.

6.6 Lie Derivatives of a Vector Field

We wish to compute $(\mathcal{L}_X Y)_p$ for X, Y vector fields. It is tempting to define

$$\lim_{\mathrm{d}t \to 0} \left[\frac{Y^\mu(h_{\mathrm{d}t}(p)) - Y^\mu(p)}{\mathrm{d}t} \right],$$

but this is not well-defined since vectors at different points live in different tangent spaces.

We use the push-forward $(h_{\mathrm{d}t})_*$ which maps a point $p \in \mathcal{M}$ on γ to a point $h_{\mathrm{d}t}(p) \in \mathcal{M}$ on γ. Therefore if Y_p is a vector field at p, we can define $(h_{\mathrm{d}t})_* Y_{h_{\mathrm{d}t}(p)}$ which is also a vector at p. We define

$$(\mathcal{L}_X Y)_p = \lim_{\mathrm{d}t \to 0} \left[\frac{Y_p - (h_{\mathrm{d}t})_* Y_{h_{\mathrm{d}t}(p)}}{\mathrm{d}t} \right]. \tag{1}$$

In the adapted chart $\{y^\mu\}$, only y^1 changes along the integral curves of X, so

$$\Big((h_{\mathrm{d}t})_* Y_{h_{\mathrm{d}t}(p)}\Big)^\mu = Y^\mu(y^1(p) - \mathrm{d}t, y^2(p), \ldots, y^n(p)).$$

Then the right hand side of (1) is simply

$$\frac{\partial Y^\mu}{\partial y^1} = Y^\mu{}_{,\nu} X^\nu$$

since $X^\nu = (1, 0, \ldots, 0)$ in this chart.

But $Y^\mu{}_{,\nu} X^\nu$ is not a vector field. Note that $X^\mu{}_{,\nu} Y^\nu = 0$ in this chart, so

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

which is a vector field at p.

If two tensors have equal components in one chart, they are equal in all charts.

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

This is the rate of change of Y along integral curves of X.

6.7 Lie Derivatives of Covectors and Tensors

For η a covector and X, Y vector fields, we use the fact that $\eta(Y)$ is a function and the Liebniz rule to calculate the Lie derivative of a covector.

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

so

$$\begin{aligned}(\mathcal{L}_X \eta)(Y) &= \mathcal{L}_X(\eta(Y)) - \eta(\mathcal{L}_X(Y)) \\ &= X(\eta(Y)) - \eta[X, Y] \\ &= X^\mu \partial_\mu(\eta_\nu Y^\nu) - \eta_\nu(Y^\nu{}_{,\mu} X^\mu - X^\nu{}_{,\mu} Y^\mu).\end{aligned}$$

Recall that

$$\begin{aligned}\eta(Y) &= \langle \eta, Y \rangle \\ &= \langle \eta_\nu \, \mathrm{d}x^\nu, Y^\mu \partial_\mu \rangle \\ &= \eta_\nu Y^\mu \langle \mathrm{d}x^\nu, \partial_\mu \rangle \\ &= \eta_\nu Y^\mu \delta^\nu_\mu \\ &= \eta_\nu Y^\nu.\end{aligned}$$

Back to our Lie derivative

$$\begin{aligned}(\mathcal{L}_X\eta)(Y) &= X^\mu\eta_{\nu,\mu}Y^\nu + X^\mu\eta_\nu {Y^\nu}_{,\mu} - \eta_\nu {Y^\nu}_{,\mu}X^\mu + \eta_\nu {X^\nu}_{,\mu}Y^\mu \\ &= X^\mu\eta_{\mu,\nu}Y^\nu + \eta_\nu {X^\nu}_{,\mu}Y^\mu\end{aligned}$$

so $(\mu \leftrightarrow \nu)$

$$(\mathcal{L}_X\eta)_\mu Y^\mu = (X^\nu\eta_{\mu,\nu} + \eta_\nu {X^\nu}_{,\mu})Y^\mu$$

which is true for all Y. Therefore

$$(\mathcal{L}_X\eta)_\mu = X^\nu\eta_{\mu,\nu} + \eta_\nu {X^\nu}_{,\mu}.$$

Example 6.7.1. We will compute the derivative of a $\binom{1}{1}$ tensor. We must use the fact that $T(\eta, Y)$ is a scalar for covector fields η and vector fields Y.

The Leibniz rule implies

$$\begin{aligned}\mathcal{L}_X(T(\eta,Y)) &= (\mathcal{L}_XT)(\eta,Y) + T(\mathcal{L}_X\eta,Y) + T(\eta,\mathcal{L}_XY),\\ (\mathcal{L}_XT)(\eta,Y) &= \mathcal{L}_X(T(\eta,Y)) - T(\mathcal{L}_X\eta,Y) - T(\eta,\mathcal{L}_XY)\\ &= X(T(\eta,Y)) - T(\mathcal{L}_X\eta,Y) - T(\eta,\mathcal{L}_XY)\\ &= X^\mu\partial_\mu({T^\lambda}_\nu\eta_\lambda Y^\nu) - {T^\mu}_\nu(\eta_{\mu,\lambda}X^\lambda + \eta_\lambda {X^\lambda}_{,\mu})Y^\nu\\ &\qquad\qquad - {T^\mu}_\nu\eta_\mu({Y^\nu}_{,\lambda}X^\lambda - {X^\nu}_{,\lambda}Y^\lambda).\end{aligned}$$

Recall that

$$\begin{aligned}T(\eta,Y) &= {T^\mu}_\nu\partial_\mu\otimes \mathrm{d}x^\nu(\eta_\lambda \mathrm{d}x^\lambda, Y^\gamma\partial_\gamma)\\ &= {T^\mu}_\nu\eta_\lambda Y^\gamma\partial_\mu\otimes\mathrm{d}x^\nu(\mathrm{d}x^\lambda,\partial_\gamma)\\ &= {T^\mu}_\nu\eta_\lambda Y^\gamma\delta^\lambda_\mu\delta^\nu_\gamma\\ &= {T^\lambda}_\nu\eta_\lambda Y^\nu.\end{aligned}$$

Returning to our Lie derivative,

$$\begin{aligned}(\mathcal{L}_XT)(\eta,Y) &= X^\mu({T^\lambda}_{\nu,\mu}\eta_\lambda Y^\nu + {T^\lambda}_\nu\eta_{\lambda,\mu}Y^\nu + {T^\lambda}_\nu\eta_\lambda {Y^\nu}_{,\mu})\\ &\quad - {T^\mu}_\nu(\eta_{\mu,\lambda}X^\lambda + \eta_\lambda {X^\lambda}_{,\mu})Y^\nu - {T^\mu}_\nu\eta_\mu({Y^\nu}_{,\lambda}X^\lambda - {X^\nu}_{,\lambda}Y^\lambda)\\ &= {T^\lambda}_{\nu,\mu}X^\mu\eta_\lambda Y^\nu + {T^\mu}_\nu {X^\nu}_{,\lambda}\eta_\mu Y^\lambda - {T^\mu}_\nu {X^\lambda}_{,\mu}\eta_\lambda Y^\nu.\end{aligned}$$

This can be written as

$${(\mathcal{L}_XT)^\mu}_\nu\eta_\mu Y^\nu = {T^\mu}_{\nu,\lambda}X^\lambda\eta_\mu Y^\nu + {T^\mu}_\lambda {X^\lambda}_{,\nu}\eta_\mu Y^\nu - {T^\lambda}_\nu {X^\mu}_{,\lambda}\eta_\mu Y^\nu.$$

This is true for all η, Y, so

$${(\mathcal{L}_XT)^\mu}_\nu = {T^\mu}_{\nu,\lambda}X^\lambda + {T^\mu}_\lambda {X^\lambda}_{,\nu} - {T^\lambda}_\nu {X^\mu}_{,\lambda}.$$

Claim. $\mathcal{L}_X$ and d commute, i.e.

$$\mathcal{L}_X \mathrm{d}f = \mathrm{d}(\mathcal{L}_X f).$$

Proof.

$$\mathcal{L}_X(\mathrm{d}f(Y)) = (\mathcal{L}_X \mathrm{d}f)(Y) + \mathrm{d}f(\mathcal{L}_X Y),$$

therefore

$$\begin{aligned}
(\mathcal{L}_X \mathrm{d}f)(Y) &= \mathcal{L}_X(\mathrm{d}f(Y)) - \mathrm{d}f(\mathcal{L}_X)Y \\
&= X(\mathrm{d}f(Y)) - (\mathcal{L}_X Y)(f) \\
&= X(\mathrm{d}f(Y)) - [X, Y](f) \\
&= X(Y(f)) - X(Y(f)) + Y(X(f)) \\
&= Y(X(f)) \\
&= Y(\mathcal{L}_X(f)) \\
&= \mathrm{d}(\mathcal{L}_X f)(Y) \qquad\qquad \forall Y.
\end{aligned}$$

Since Y is arbitrary,

$$\mathcal{L}_X \mathrm{d}f = \mathrm{d}(\mathcal{L}_X f).$$

□

7 Linear Connections and Covariant Differentiation

We require some concept of differentiation on a manifold that is chart independent (covariant) and we would like a derivative operator whose components transform like a tensor.

Notice, for example, that the partial derivative of a vector field does not transform like a tensor.

$$\begin{aligned}
\frac{\partial V^\mu}{\partial x^\nu} = V^\mu{}_{,\nu} &= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial}{\partial x^{\nu'}} \left\{ \frac{\partial x^\mu}{\partial x^{\mu'}} V^{\mu'} \right\} \\
&= \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial x^\mu}{\partial x^{\mu'}} V^{\mu'}{}_{,\nu'} + \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial^2 x^\mu}{\partial x^{\nu'} \partial x^{\mu'}} V^{\mu'}.
\end{aligned}$$

These are not the components of a $\binom{1}{1}$ tensor.

In order to properly define a derivative operator on a manifold, one needs to be able to compare vectors (or tensors) at different points. The machinery that allows us to do this is the "linear connection".

7.1 Linear Connections

We define a linear connection ∇ to be a map which sends every pair of smooth vector fields X, Y to a smooth vector field,

$$\nabla : X, Y \mapsto \nabla_X Y,$$

satisfying

$$\begin{aligned}\nabla_X(Y+Z) &= \nabla_X(Y) + \nabla_X(Z),\\ \nabla_{fX+Y}(Z) &= f\nabla_X(Z) + \nabla_Y(Z).\end{aligned}$$

It also satisfies

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

and the Liebniz rule

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

$\nabla_X Y$ is the covariant derivative of Y with respect to X.

It is important to note that ∇ is not a tensor since the Liebniz rule implies it is not linear in Y. However, considered as a map from $X \mapsto \nabla_X Y$, i.e.

$$\nabla Y : X \mapsto \nabla_X Y$$

which is a linear map $T_p(\mathcal{M}) \to T_p(\mathcal{M})$ which takes a vector and outputs a vector, ∇Y is a $\binom{1}{1}$ tensor known as the covariant derivative of Y.

7.2 Covariant Derivative of a Vector Field

In an arbitrary basis $\{e_\mu\}$, ∇_{e_μ} is a map taking e_μ to some vector field.

$$\nabla_{e_\mu} e_\nu \equiv \nabla_\mu e_\nu.$$

Since this is a vector field for each e_μ, we may write $\nabla_\mu e_\nu$ as a linear combination of basis vectors

$$\nabla_\mu e_\nu = \Gamma^\lambda_{\nu\mu} e_\lambda$$

where $\Gamma^\lambda_{\nu\mu}$ are known as the connection coefficients and are not the components of a tensor.

Since $X = X^\mu e_\mu$, $Y = Y^\mu e_\mu$, we have

$$\begin{aligned}
\nabla_X Y &= (\nabla_X Y)^\mu e_\mu \equiv (Y^\mu{}_{;\nu} X^\nu) e_\mu \\
&= \nabla_X (Y^\mu e_\mu) \\
&= (\nabla_X Y^\mu) e_\mu + Y^\mu \nabla_X e_\mu \\
&= X(Y^\mu) e_\mu + Y^\mu \nabla_{X^\nu e_\nu} e_\mu \\
&= X^\nu e_\nu Y^\mu e_\mu + Y^\mu X^\nu \nabla_\nu e_\mu \\
&= e_\nu Y^\mu X^\nu e_\mu + Y^\mu X^\nu \Gamma^\lambda_{\mu\nu} e_\lambda \\
&= e_\nu Y^\mu X^\nu e_\mu + Y^\lambda X^\nu \Gamma^\mu_{\lambda\nu} e_\mu .
\end{aligned}$$

Therefore

$$(Y^\mu{}_{;\nu} X^\nu) e_\mu = e_\nu (Y^\mu) X^\nu e_\mu + X^\nu Y^\lambda \Gamma^\mu_{\lambda\nu} e_\mu .$$

This is true for all e_μ and X, so

$$\begin{aligned}
Y^\mu{}_{;\nu} &= e_\nu (Y^\mu) + Y^\lambda \Gamma^\mu_{\lambda\nu} \\
&= Y^\mu{}_{,\nu} + Y^\lambda \Gamma^\mu_{\lambda\nu}
\end{aligned}$$

in a coordinate-induced basis.

To see how the $\Gamma^\lambda_{\mu\nu}$s transform under a coordinate transformation, we will look at

$$\nabla_{e_{\mu'}} e_{\nu'} = \Gamma^{\lambda'}_{\nu'\mu'} e_{\lambda'} = \Gamma^{\lambda'}_{\nu'\mu'} \Lambda^\alpha{}_{\lambda'} e_\alpha .$$

We also have

$$\begin{aligned}
\nabla_{e_{\mu'}} e_{\nu'} &= \nabla_{\Lambda^\alpha{}_{\mu'} e_\alpha} (\Lambda^\beta{}_{\nu'} e_\beta) \\
&= \Lambda^\alpha{}_{\mu'} \nabla_\alpha (\Lambda^\beta{}_{\nu'}) e_\beta + \Lambda^\alpha{}_{\mu'} \Lambda^\beta{}_{\nu'} \nabla_\alpha e_\beta \\
&= \Lambda^\alpha{}_{\mu'} e_\alpha (\Lambda^\beta{}_{\nu'}) e_\beta + \Lambda^\alpha{}_{\mu'} \Lambda^\beta{}_{\nu'} \Gamma^\lambda_{\beta\alpha} e_\lambda . \\
&= \Lambda^\beta{}_{\mu'} e_\beta (\Lambda^\alpha{}_{\nu'}) e_\alpha + \Lambda^\lambda{}_{\mu'} \Lambda^\beta{}_{\nu'} \Gamma^\alpha_{\beta\lambda} e_\alpha .
\end{aligned}$$

Together, this gives us

$$\Gamma^{\lambda'}_{\nu'\mu'} \Lambda^\alpha{}_{\lambda'} e_\alpha = \Lambda^\beta{}_{\mu'} e_\beta (\Lambda^\alpha{}_{\nu'}) e_\alpha + \Lambda^\lambda{}_{\mu'} \Lambda^\beta{}_{\nu'} \Gamma^\alpha_{\beta\lambda} e_\alpha .$$

This is true for all e_α and implies

$$\Gamma^{\lambda'}_{\nu'\mu'} \Lambda^\alpha{}_{\lambda'} = \Lambda^\beta{}_{\mu'} e_\beta (\Lambda^\alpha{}_{\nu'}) + \Lambda^\lambda{}_{\mu'} \Lambda^\beta{}_{\nu'} \Gamma^\alpha_{\beta\lambda} .$$

Then

$$\Gamma^{\lambda'}_{\nu'\mu'} \Lambda^\alpha{}_{\lambda'} \Lambda^{\gamma'}{}_\alpha = \Lambda^{\gamma'}{}_\alpha \Lambda^\beta{}_{\mu'} e_\beta (\Lambda^\alpha{}_{\nu'}) + \Lambda^{\gamma'}{}_\alpha \Lambda^\lambda{}_{\mu'} \Lambda^\beta{}_{\nu'} \Gamma^\alpha_{\beta\lambda}$$

and since $\Lambda^{\alpha}{}_{\lambda'}\Lambda^{\gamma'}{}_{\alpha} = \delta^{\gamma'}_{\lambda'}$, we conclude that

$$\Gamma^{\gamma'}_{\nu'\mu'} = \Lambda^{\gamma'}{}_{\alpha}\Lambda^{\beta}{}_{\mu'}e_{\beta}(\Lambda^{\alpha}{}_{\nu'}) + \Lambda^{\gamma'}{}_{\alpha}\Lambda^{\lambda}{}_{\mu'}\Lambda^{\beta}{}_{\nu'}\Gamma^{\alpha}_{\beta\lambda}.$$

The presence of the first term on the right hand side shows that $\Gamma^{\lambda}_{\mu\nu}$ does not transform like a tensor. In a coordinate basis, we have

$$\Gamma^{\gamma'}_{\nu'\mu'} = \frac{\partial x^{\gamma'}}{\partial x^{\alpha}}\frac{\partial x^{\beta}}{\partial x^{\mu'}}\frac{\partial}{\partial x^{\beta}}\left(\frac{\partial x^{\alpha}}{\partial x^{\nu'}}\right) + \frac{\partial x^{\gamma'}}{\partial x^{\alpha}}\frac{\partial x^{\lambda}}{\partial x^{\mu'}}\frac{\partial x^{\beta}}{\partial x^{\nu'}}\Gamma^{\alpha}_{\beta\lambda}.$$

One can show that $Y^{\mu}{}_{;\nu}$ transforms like a $\binom{1}{1}$ tensor.

$$Y^{\mu}{}_{;\nu} = \underbrace{\underbrace{Y^{\mu}{}_{,\nu}}_{\text{not a tensor}} + \underbrace{\Gamma^{\mu}_{\lambda\nu}Y^{\lambda}}_{\text{not a tensor}}}_{\text{a tensor}}.$$

Lemma 7.2.1. Let $\nabla, \widetilde{\nabla}$ be two connections on $\mathcal{M}$. Then their difference

$$D(X,Y) = \nabla_X Y - \widetilde{\nabla}_X Y$$

is always a tensor.

Proof. We must show that the map D is multilinear. We have already seen that ∇ is linear in X, so we must show linearity in Y.

$$\begin{aligned} D(X,fY) &= \nabla_X(fY) - \widetilde{\nabla}_X(fY) \\ &= f\nabla_X(Y) + X(f)Y - f\widetilde{\nabla}_X(Y) - X(f)Y \\ &= f(\nabla_X Y - \widetilde{\nabla}_X Y) \\ &= fD(X,Y). \end{aligned}$$

It is trivial to show that $D(X, Y+Z) = D(X,Y) + D(X,Z)$. Therefore D is a tensor. □

7.3 Covariant Derivative of Covectors and Tensors

We will use the fact that $\eta(Y) = \eta_\nu Y^\nu$ is a scalar.

$$\nabla_\mu(\eta_\nu Y^\nu) = (\nabla_\mu \eta_\nu)Y^\nu + \eta_\nu \nabla_\mu Y^\nu,$$

so

$$\begin{aligned} (\nabla_\mu\eta_\nu)Y^\nu &= \nabla_\mu(\eta_\nu Y^\nu) - \eta_\nu\nabla_\mu Y^\nu \\ &= \partial_\mu(\eta_\nu Y^\nu) - \eta_\nu(Y^\nu{}_{,\mu} + \Gamma^\nu_{\lambda\mu}Y^\lambda) \\ &= \eta_{\nu,\mu}Y^\nu + \eta_\nu Y^\nu{}_{,\mu} - \eta_\nu Y^\nu{}_{,\mu} - \eta_\nu\Gamma^\nu_{\lambda\mu}Y^\lambda \\ &= \eta_{\nu,\mu}Y^\nu - \eta_\nu\Gamma^\nu_{\lambda\mu}Y^\lambda \\ &= (\eta_{\nu,\mu} - \eta_\lambda\Gamma^\lambda_{\nu\mu})Y^\nu. \end{aligned}$$

This is true for all Y^ν, so

$$\nabla_\mu \eta_\nu = \eta_{\nu,\mu} - \eta_\lambda \Gamma^\lambda_{\nu\mu}.$$

These are the components of the covariant derivative of a covector. Note the negative sign; this was positive for vectors.

Example 7.3.1. We will compute the covariant derivative of a $\binom{1}{1}$ tensor, using the fact that $T^\mu{}_\nu \eta_\mu Y^\nu$ is a scalar.

$$\nabla_\lambda (T^\mu{}_\nu \eta_\mu Y^\nu) = (\nabla_\lambda T^\mu{}_\nu)\eta_\mu Y^\nu + T^\mu{}_\nu (\nabla_\lambda \eta_\mu) Y^\nu + T^\mu{}_\nu \eta_\mu (\nabla_\lambda Y^\nu),$$

so

$$\begin{aligned}
(\nabla_\lambda T^\mu{}_\nu)\eta_\mu Y^\nu &= \nabla_\lambda (T^\mu{}_\nu \eta_\mu Y^\nu) - T^\mu{}_\nu (\nabla_\lambda \eta_\mu) Y^\nu - T^\mu{}_\nu \eta_\nu (\nabla_\lambda Y^\nu) \\
&= \partial_\lambda (T^\mu{}_\nu \eta_\mu Y^\nu) - T^\mu{}_\nu (\eta_{\mu,\lambda} - \eta_\gamma \Gamma^\gamma_{\mu\lambda}) Y^\nu - T^\mu{}_\nu \eta_\mu (Y^\nu{}_{,\lambda} + \Gamma^\nu_{\gamma\lambda} Y^\gamma) \\
&= T^\mu{}_{\nu,\lambda} \eta_\mu Y^\nu + T^\mu{}_\nu \eta_{\mu,\lambda} Y^\nu + T^\mu{}_\nu \eta_\mu Y^\nu{}_{,\lambda} \\
&\quad - T^\mu{}_\nu \eta_{\mu,\lambda} Y^\nu + T^\mu{}_\nu \eta_\gamma \Gamma^\gamma_{\mu\lambda} Y^\nu - T^\mu{}_\nu \eta_\mu Y^\nu{}_{,\lambda} - T^\mu{}_\nu \eta_\mu \Gamma^\nu_{\gamma\lambda} Y^\gamma \\
&= T^\mu{}_{\nu,\lambda} \eta_\mu Y^\nu + T^\mu{}_\nu \eta_\gamma \Gamma^\gamma_{\mu\lambda} T^\nu - T^\mu{}_\nu \eta_\mu \Gamma^\nu_{\gamma\lambda} Y^\gamma \\
&= T^\mu{}_{\nu,\lambda} \eta_\mu Y^\nu + T^\gamma{}_\nu \eta_\mu \Gamma^\mu_{\gamma\lambda} Y^\nu - T^\mu{}_\nu \eta_\mu \Gamma^\gamma_{\nu\lambda} Y^\nu \\
&= (T^\mu{}_{\nu,\lambda} + T^\gamma{}_\nu \Gamma^\mu_{\gamma\lambda} - T^\mu{}_\nu \Gamma^\gamma_{\nu\lambda}) \eta_\mu Y^\nu.
\end{aligned}$$

This is true for all η_μ, Y^ν, so

$$\nabla_\lambda T^\mu{}_\nu \equiv T^\mu{}_{\nu;\lambda} = T^\mu{}_{\nu,\lambda} + T^\gamma{}_\nu \Gamma^\mu_{\gamma\lambda} - T^\mu{}_\gamma \Gamma^\gamma_{\nu\lambda}.$$

These are the components of a $\binom{1}{1}$ tensor.

In general, we have

$$\begin{aligned}
\nabla_\lambda T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\cdots} = T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\cdots,\lambda} &+ \Gamma^{\mu_1}_{\alpha\lambda} T^{\alpha\mu_2\cdots}{}_{\nu_1\nu_2\cdots} + \Gamma^{\mu_2}_{\alpha\lambda} T^{\mu_1\alpha\cdots}{}_{\nu_1\nu_2\cdots} + \cdots \\
&- \Gamma^\alpha_{\nu_1\lambda} T^{\mu_1\mu_2\cdots}{}_{\alpha\nu_2\cdots} - \Gamma^\alpha_{\nu_2\lambda} T^{\mu_1\mu_2\cdots}{}_{\nu_1\alpha\cdots} - \cdots .
\end{aligned}$$

8 Geodesics and Parallel Transport

8.1 Parallel Transport Along a Curve

We need a curved space generalisation of moving a vector along a path while keeping the vector constant. In curved space, the result of parallel transporting a vector (or tensor) from one point to another depends on the path taken.

We say that a tensor T is parallel transported with respect to X a vector if

$$\nabla_X T = 0.$$

Let $\gamma(\tau)$ be some curve and $\phi \circ \gamma = \{x^\mu(\tau)\}$ be the image of γ in $\mathbb{R}^n$. Then T is parallel transported along the curve γ if $\nabla_X T = 0$ holds, with X being the tangent vector to the curve, i.e.

$$X = \left(\frac{\mathrm{d}x^\mu(\tau)}{\mathrm{d}\tau}\right)\left(\frac{\partial}{\partial x^\mu}\right).$$

Then $\nabla_X T = 0$ becomes

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\nabla_\mu T \equiv \frac{DT}{\mathrm{d}\tau} = 0.$$

In particular, for a vector field V, this gives

$$\frac{\mathrm{d}V^\mu}{\mathrm{d}\tau} + \Gamma^\mu_{\beta\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}V^\beta = 0, \qquad \left(\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial V^\mu}{\partial x^\nu} = \frac{\mathrm{d}V^\mu}{\mathrm{d}\tau}\right)$$

the parallel transport equation for a vector field.

8.2 The Geodesic Equation

A geodesic in flat space is a straight line, i.e. the shortest distance between two points. Equivalently, geodesics are paths that parallel transport their own tangent vector.

In curved space, we have not yet introduced the metric and therefore have no well-defined notion of distance. We must use the second definition of a geodesic.

The tangent vector to a path $x^\mu(\tau)$ is simply $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$. These components must satisfy the parallel transport equation, so

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\beta\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = 0.$$

This is the geodesic equation. Note that we also have initial conditions $x^\mu(0) = p$, $\dot{x}^\mu(0) = X^\mu_p$, so that locally there is a unique solution. This is not true globally.

8.3 Normal Coordinates

Given a point $p \in \mathcal{M}$ and a chart $\{x^\mu\}$, we may find a new chart $\{\hat{x}^\mu\}$ such that $\hat{\Gamma}^\mu_{(\nu\lambda)}(p) = 0$, i.e. the geodesics at p satisfy $\ddot{\hat{x}}^\mu = 0$ so that they are locally linear functions of τ. These coordinates are connected by a unique geodesic, which defines the normal neighbourhood of p.

To construct normal coordinates, we take $x^\mu(p) = 0^\mu$ and set

$$\hat{x}^\mu = x^\mu + \frac{1}{2} Q^\mu{}_{\nu\lambda} x^\nu x^\lambda$$

where $Q^\mu{}_{\nu\lambda} = Q^\mu{}_{(\nu\lambda)}$ are constants. Let

$$|x|^2 = |x^1|^2 + |x^2|^2 + \cdots + |x^n|^2$$

and we note that $\hat{x}^\mu = x^\mu + \mathcal{O}(|x|^2)$. Thus

$$x^\mu = \hat{x}^\mu - \frac{1}{2} Q^\mu{}_{\nu\lambda} x^\nu x^\lambda,$$

which may be solved iteratively to obtain

$$x^\mu = \hat{x}^\mu - \frac{1}{2} Q^\mu{}_{\nu\lambda} \hat{x}^\nu \hat{x}^\lambda + \mathcal{O}(|x|^2).$$

Therefore,

$$\begin{aligned} \frac{\partial \hat{x}^\mu}{\partial x^\gamma} &= \delta^\mu_\gamma + Q^\mu{}_{\gamma\lambda} x^\lambda + \cdots , \\ \frac{\partial x^\mu}{\partial \hat{x}^\gamma} &= \delta^\mu_\gamma - Q^\mu{}_{\gamma\lambda} \hat{x}^\lambda + \cdots . \end{aligned}$$

Recall that $\hat{\Gamma}^\mu_{\nu\lambda}$ is related to $\Gamma^\mu_{\nu\lambda}$ by

$$\hat{\Gamma}^\mu_{\nu\lambda} = \frac{\partial \hat{x}^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \hat{x}^\lambda} \frac{\partial^2 x^\alpha}{\partial x^\beta \partial \hat{x}^\nu} + \frac{\partial \hat{X}^\mu}{\partial x^\alpha} \frac{\partial x^\gamma}{\partial \hat{x}^\lambda} \frac{\partial x^\beta}{\partial \hat{x}^\nu} \Gamma^\alpha_{\beta\gamma},$$

implying

$$\begin{aligned} \hat{\Gamma}^\mu_{\nu\lambda}(p) &= \delta^\mu_\alpha \delta^\beta_\lambda (-Q^\alpha{}_{\beta\nu}) + \delta^\mu_\alpha \delta^\gamma_\lambda \delta^\beta_\nu \Gamma^\alpha_{\beta\gamma}\Big|_p \\ &= -Q^\mu{}_{\lambda\nu} + \Gamma^\mu_{\nu\lambda}\big|_p . \end{aligned}$$

Choosing $Q^\mu{}_{\nu\lambda} = \Gamma^\mu_{(\nu\lambda)}(p)$ gives

$$\hat{\Gamma}^\mu_{(\nu\lambda)}(p) = 0$$

as required.

9 Curvature

9.1 Torsion

The torsion tensor is a $\binom{1}{2}$ tensor field defined by

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

It is easy to check multilinearity (i.e. if it is a tensor). To obtain the components of the torsion tensor, we introduce the commutator coefficients

$$[e_\nu, e_\lambda] = \gamma^\mu_{\nu\lambda} e_\mu.$$

Then

$$\begin{aligned} T(e_\nu, e_\lambda) &= T^\mu_{\nu\lambda} e_\mu \\ &= \nabla_\nu e_\lambda - \nabla_\lambda e_\nu - [e_\nu, e_\lambda] \\ &= \Gamma^\mu_{\lambda\nu} e_\mu - \Gamma^\mu_{\nu\lambda} e_\mu - \gamma^\mu_{\nu\lambda} e_\mu, \end{aligned}$$

so

$$\begin{aligned} T^\mu_{\nu\lambda} &= \Gamma^\mu_{\lambda\nu} - \Gamma^\mu_{\nu\lambda} - \gamma^\mu_{\nu\lambda} \\ &= -2\Gamma^\mu_{[\nu\lambda]} - \gamma^\mu_{\nu\lambda}, \end{aligned}$$

completely antisymmetric in lower indices.

Theorem 9.1.1. The torsion on a manifold with a symmetric connection in a coordinate-induced basis is zero.

Proof. A symmetric connection implies

$$\Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{\lambda\nu}, \quad \Gamma^\mu_{[\nu\lambda]} = 0.$$

In a coordinate-induced basis, $\gamma^\mu_{\nu\lambda} = 0$.[2] □

We will usually assume zero torsion.

9.2 The Riemann Curvature Tensor

The Riemann curvature tensor is a $\binom{1}{3}$ tensor field defined by

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

[2]This is obvious, since $\partial_\mu, \partial_\nu$ commute. See also assignment 1 question 9.

where X, Y, Z are vector fields. To see that this is a tensor, we must check multilinearity. Since it is antisymmetric in X, Y, we need only check X, Z.

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

It is trivial to show that $R(X,Y)(Z+W) = R(X,Y)Z + R(X,Y)W$.

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

similarly for $R(X,Y)(fZ)$.

To compute the components of the Riemann curvature tensor, we have

$$\begin{aligned}
R(e_\nu, e_\lambda)e_\rho &= R^\mu{}_{\rho\nu\lambda}e_\mu \\
&= \nabla_\nu\nabla_\lambda e_\rho - \nabla_\lambda\nabla_\nu e_\rho - \nabla_{[e_\nu,e_\lambda]}e_\rho \\
&= \nabla_\nu(\Gamma^\mu_{\rho\lambda}e_\mu) - \nabla_\lambda(\Gamma^\mu_{\rho\nu}e_\mu) - \nabla_{\gamma^\mu_{\nu\lambda}e_\mu}e_\rho \\
&= (\nabla_\nu\Gamma^\mu_{\rho\lambda})e_\mu + \Gamma^\mu_{\rho\lambda}\nabla_\nu e_\mu - (\nabla_\lambda\Gamma^\mu_{\rho\nu})e_\mu - \Gamma^\mu_{\rho\nu}\nabla_\lambda e_\mu - \gamma^\mu_{\nu\lambda}\nabla_\mu e_\rho \\
&= e_\nu(\Gamma^\mu_{\rho\lambda})e_\mu + \Gamma^\mu_{\rho\lambda}\Gamma^\alpha_{\mu\nu}e_\alpha - e_\lambda(\Gamma^\mu_{\rho\nu})e_\mu - \Gamma^\mu_{\rho\nu}\Gamma^\alpha_{\mu\lambda}e_\alpha - \gamma^\mu_{\nu\lambda}\Gamma^\alpha_{\rho\mu}e_\alpha \\
&= e_\nu(\Gamma^\mu_{\rho\lambda})e_\mu + \Gamma^\alpha_{\rho\lambda}\Gamma^\mu_{\alpha\nu}e_\mu - e_\lambda(\Gamma^\mu_{\rho\nu})e_\mu - \Gamma^\alpha_{\rho\nu}\Gamma^\mu_{\alpha\lambda}e_\mu - \gamma^\alpha_{\nu\lambda}\Gamma^\mu_{\rho\alpha}e_\mu.
\end{aligned}$$

This is true for all e_μ, implying

$$R^\mu{}_{\rho\nu\lambda} = e_\nu(\Gamma^\mu_{\rho\lambda}) + \Gamma^\alpha_{\rho\lambda}\Gamma^\mu_{\alpha\nu} - e_\lambda(\Gamma^\mu_{\rho\nu}) - \Gamma^\alpha_{\rho\nu}\Gamma^\mu_{\alpha\lambda} - \gamma^\alpha_{\nu\lambda}\Gamma^\mu_{\rho\alpha},$$

the components of the Riemann curvature tensor.

In a coordinate induced basis, we have

$$\begin{gathered}
(e_\nu \to \partial_\nu), \qquad (\gamma^\alpha_{\nu\lambda} \to 0), \\
R^\mu{}_{\rho\nu\lambda} = \Gamma^\mu_{\rho\lambda,\nu} - \Gamma^\mu_{\rho\nu,\lambda} + \Gamma^\alpha_{\rho\lambda}\Gamma^\mu_{\alpha\nu} - \Gamma^\alpha_{\rho\nu}\Gamma^\mu_{\alpha\lambda},
\end{gathered}$$

the components of the Riemann curvature tensor in a coordinate-induced basis.

This is important in general relativity, as curvature dictates how particles move.

Theorem 9.2.1 (The Ricci Identity)**.** Let X^μ be the components of X with respect to a coordinate-induced basis. We denote

$$X^\mu{}_{;\nu\lambda} = \nabla_\lambda \nabla_\nu X^\mu.$$

The Ricci identity states that, for a symmetric connection,

$$X^\mu{}_{;\nu\lambda} - X^\mu{}_{;\lambda\nu} = -R^\mu{}_{\rho\nu\lambda} X^\rho.$$

Proof.

$$\nabla_\nu X^\mu = X^\mu{}_{;\nu} = X^\mu{}_{,\nu} + \Gamma^\mu_{\rho\nu} X^\rho,$$

$$\begin{aligned}
X^\mu{}_{;\nu\lambda} &= \nabla_\lambda (X^\mu{}_{,\nu} + \Gamma^\mu_{\rho\nu} X^\rho) \\
&= \partial_\lambda (X^\mu{}_{,\nu} + \Gamma^\mu_{\rho\nu} X^\rho) + \Gamma^\mu_{\alpha\lambda}(X^\alpha{}_{,\nu} + \Gamma^\alpha_{\rho\nu} X^\rho) - \Gamma^\beta_{\nu\lambda}(X^\mu{}_{,\beta} + \Gamma^\mu_{\rho\beta} X^\rho) \\
&= X^\mu{}_{,\nu\lambda} + \Gamma^\mu_{\rho\nu,\lambda} X^\rho + \Gamma^\mu_{\rho\nu} X^\rho{}_{,\lambda} + \Gamma^\mu_{\alpha\lambda} X^\alpha{}_{,\nu} + \Gamma^\mu_{\alpha\lambda}\Gamma^\alpha_{\rho\nu} X^\rho \\
&\qquad - \Gamma^\beta_{\nu\lambda} X^\mu{}_{,\beta} - \Gamma^\beta_{\nu\lambda}\Gamma^\mu_{\rho\beta} X^\rho \\
&= X^\mu{}_{,\nu\lambda} + \Gamma^\mu_{\rho\nu,\lambda} X^\rho + \Gamma^\mu_{\alpha\lambda}\Gamma^\alpha_{\rho\nu} X^\rho - \Gamma^\beta_{\nu\lambda}\Gamma^\mu_{\rho\beta} X^\rho \\
&\qquad + \Gamma^\mu_{\rho\nu} X^\rho{}_{,\lambda} + \Gamma^\mu_{\alpha\lambda} X^\alpha{}_{,\nu} - \Gamma^\beta_{\nu\lambda} X^\mu{}_{,\beta},
\end{aligned}$$

$$\begin{aligned}
X^\mu{}_{;\lambda\nu} = X^\mu{}_{,\lambda\nu} &+ \Gamma^\mu_{\rho\lambda,\nu} X^\rho + \Gamma^\mu_{\alpha\nu}\Gamma^\alpha_{\rho\lambda} X^\rho - \Gamma^\beta_{\lambda\nu}\Gamma^\mu_{\rho\beta} X^\rho \\
&+ \Gamma^\mu_{\rho\lambda} X^\rho{}_{,\nu} + \Gamma^\mu_{\alpha\nu} X^\alpha{}_{,\lambda} - \Gamma^\beta_{\lambda\nu} X^\mu{}_{,\beta},
\end{aligned}$$

$$\begin{aligned}
X^\mu{}_{;\nu\lambda} - X^\mu{}_{;\lambda\nu} &= \Gamma^\mu_{\rho\nu,\lambda} X^\rho - \Gamma^\mu_{\rho\lambda,\nu} X^\rho + \Gamma^\mu_{\alpha\lambda}\Gamma^\alpha_{\rho\nu} X^\rho - \Gamma^\mu_{\alpha\nu}\Gamma^\alpha_{\rho\lambda} X^\rho \\
&= -R^\mu{}_{\rho\nu\lambda} X^\rho.
\end{aligned}$$ □

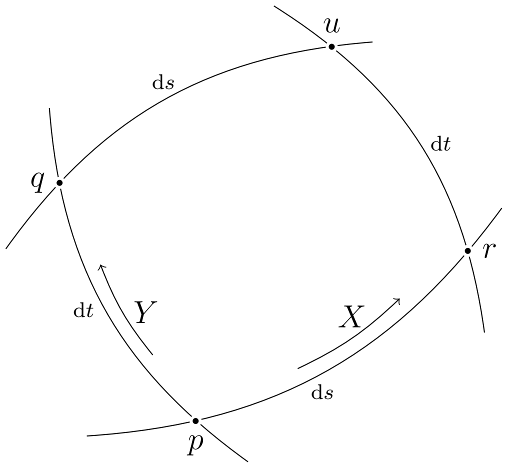
In flat space, the covariant derivative becomes the derivative (; = ,), implying $\nu\lambda$ would commute. Hence $R = 0$ and there is no curvature in flat space.

Recall that for the torsion, T,

$$\left.\begin{array}{r} \Gamma^\mu_{\nu\lambda} = \Gamma^\mu_{\lambda\nu} \\ [e_\nu, e_\lambda] = 0 \end{array}\right\} \quad \Rightarrow \quad T = 0.$$

9.3 Geometric Interpretation of the Riemann Tensor

We set up a similar construction to that of §6.3



We parallel transport a vector Z along integral curves of vector fields X and Y where $[X, Y] = 0$, i.e. the curve is a closed infinitesimal quadrilateral.

We work in normal coordinates so that $\Gamma^\mu_{\nu\lambda}(p) = 0$.

Along pr we have that $\nabla_X Z = 0$, implying

$$\begin{aligned}
\frac{\mathrm{d}Z^\mu}{\mathrm{d}s} + \Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda &= 0, \\
\frac{\mathrm{d}^2 Z^\mu}{\mathrm{d}s^2} &= -\frac{\mathrm{d}}{\mathrm{d}s}(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda) \\
&= -(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda)_{,\rho} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} \\
&= -(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda)_{,\rho} X^\rho.
\end{aligned}$$

Note that

$$\begin{aligned}
(\Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda)_{,\rho} &= \Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda + \Gamma^\mu_{\nu\lambda} Z^\nu{}_{,\rho} X^\lambda + \Gamma^\mu_{\nu\lambda} Z^\nu X^\lambda{}_{,\rho} \\
&= \Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda.
\end{aligned}$$

Taylor expanding Z^μ_r around Z^μ_p,

$$Z^\mu_r = Z^\mu_p + \left(\frac{\mathrm{d}Z^\mu}{\mathrm{d}s}\right)_p \mathrm{d}s + \frac{1}{2}\left(\frac{\mathrm{d}^2 Z^\mu}{\mathrm{d}s^2}\right)_p \mathrm{d}s^2 + \cdots$$

and note that $\frac{\mathrm{d}Z^\mu}{\mathrm{d}s} = 0$ in normal coordinates. Further,

$$\begin{aligned}
Z^\mu_u &= Z^\mu_r + \left(\frac{\mathrm{d}Z^\mu}{\mathrm{d}t}\right)_r \mathrm{d}t + \frac{1}{2}\left(\frac{\mathrm{d}^2 Z^\mu}{\mathrm{d}t^2}\right)_r \mathrm{d}t^2 + \cdots \\
&= Z^\mu_p + \frac{1}{2}\left(\frac{\mathrm{d}^2 Z^\mu}{\mathrm{d}s^2}\right)_p \mathrm{d}s^2 \\
&\quad + \left[\left(\frac{\mathrm{d}Z^\mu}{\mathrm{d}t}\right)_p + \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}Z^\mu}{\mathrm{d}s}\mathrm{d}s\right)_p + \cdots\right]\mathrm{d}t \\
&\quad + \frac{1}{2}\left(\frac{\mathrm{d}Z^\mu}{\mathrm{d}t^2}\right)_p \mathrm{d}t^2 + \cdots \\
&= Z^\mu_p + \frac{1}{2}\left(\frac{\mathrm{d}^2 Z^\mu}{\mathrm{d}s^2}\right)_p \mathrm{d}s^2 + \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}Z^\mu}{\mathrm{d}s}\right)_p \mathrm{d}s\,\mathrm{d}t + \cdots \\
&= Z^\mu_p - \frac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)\mathrm{d}s^2 - (\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda X^\rho)_p \mathrm{d}s\,\mathrm{d}t \\
&\quad - \frac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda Y^\rho)_p \mathrm{d}t^2 + \cdots .
\end{aligned}$$

Parallel transporting Z in the other direction pqu gives

$$\begin{aligned}
Z^\mu_u = Z^\mu_p &- \frac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu Y^\lambda Y^\rho)_p \mathrm{d}t^2 - (\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda Y^\rho)_p \mathrm{d}s\mathrm{d}t \\
&- \frac{1}{2}(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu X^\lambda X^\rho)_p \mathrm{d}s^2 + \cdots .
\end{aligned}$$

The difference represents the change in Z around $pruqp$.

$$\begin{aligned}
\Delta Z^\mu &= -(\Gamma^\mu_{\nu\lambda,\rho} Z^\nu)_p (Y^\lambda X^\rho - X^\lambda Y^\rho)_p \mathrm{d}s\,\mathrm{d}t + \cdots \\
&= (\Gamma^\mu_{\nu\lambda,\rho} - \Gamma^\mu_{\nu\rho,\lambda})_p X^\lambda_p Y^\rho_p Z^\nu_p \mathrm{d}s\,\mathrm{d}t + \cdots \\
&= (R^\mu{}_{\nu\rho\lambda} X^\lambda Y^\rho Z^\nu)_p \mathrm{d}s\,\mathrm{d}t + \cdots
\end{aligned}$$

(since terms like $\Gamma\Gamma$ vanish in normal coordinates). Therefore

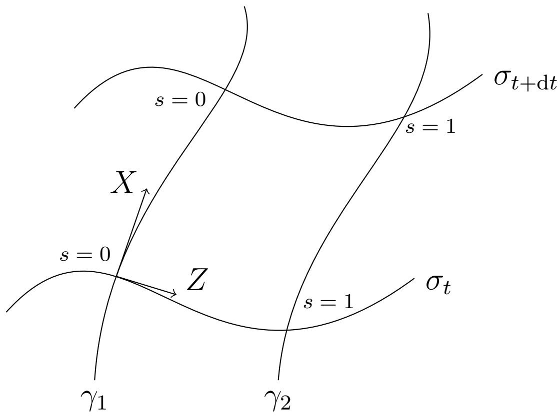
$$R^\mu{}_{\nu\rho\lambda} X^\lambda Y^\rho Z^\nu = \lim_{\substack{\mathrm{d}s\to 0\\ \mathrm{d}t\to 0}} \left(\frac{\Delta Z}{\mathrm{d}s\,\mathrm{d}t}\right).$$

$R(X,Y)Z$ measures the change in Z after parallel transporting around a closed quadrilateral spanned by the vector fields X and Y.

9.4 Geodesic Deviation

We let γ_1, γ_2 be neighbouring integral curves of a vector field X which are parameterised by t. We let $\{\sigma_t\}$ be curves parameterised by s intersecting γ_1

and γ_2 at "time" t. Let Z be the tangent vector to σ_t such that $s = 0$ on γ_1 and $s = 1$ on γ_2.



Z is a vector pointing from one integral curve to another. Since the quadrilateral is closed, we have

$$[X, Z] = 0$$

and Z is known as a connecting vector to the curves.

Theorem 9.4.1. Assuming the torsion vanishes, let X be tangent to a congruence of geodesics and let Z be a connecting vector. Then the acceleration of Z is given by

$$\nabla_X \nabla_X Z = R(X, Z)X.$$

This is the geodesic equation.

Proof. Vanishing torsion gives us

$$T(X, Z) = \nabla_X Z - \nabla_Z X - \underbrace{[X, Z]}_{0} = 0 \quad \Rightarrow \quad \nabla_X Z = \nabla_Z X.$$

Then (as $\nabla_X X = 0$)

$$\begin{aligned}
\nabla_X \nabla_X Z &= \nabla_X \nabla_Z X \\
&= R(X, Z)X + \nabla_Z \nabla_X X + \nabla_{[X,Z]} X \\
&= R(X, Z)X.
\end{aligned}$$

□

The relative acceleration between two neighbouring geodesics is proportional to the curvature.

In component form, we have

$$\begin{aligned}\frac{D^2 Z^\mu}{\mathrm{d}t^2} &= (X^\nu Z^\mu{}_{;\nu})_{;\lambda} X^\lambda \\ &= R^\mu{}_{\nu\lambda\rho} X^\nu X^\lambda Z^\rho.\end{aligned}$$

This equation gives us information about $R^\mu{}_{(\nu\lambda)\rho}$, since $X^\nu X^\lambda$ is symmetric in (ν, λ).

9.5 Symmetries of the Riemann Tensor

Recall that in a coordinate-induced basis,

$$R^\mu{}_{\rho\nu\lambda} = \Gamma^\mu_{\rho\lambda,\nu} - \Gamma^\mu_{\rho\nu,\lambda} + \Gamma^\alpha_{\rho\lambda}\Gamma^\mu_{\alpha\nu} - \Gamma^\alpha_{\rho\nu}\Gamma^\mu_{\alpha\lambda}.$$

Therefore $R^\mu{}_{\rho\nu\lambda} = -R^\mu{}_{\rho\lambda\nu}$, or

$$R^\mu{}_{\rho(\nu\lambda)} = 0$$

(assuming $T = 0$, symmetric connection and coordinate-induced basis).

We also have

$$R^\mu{}_{\nu\lambda\rho} + R^\mu{}_{\rho\nu\lambda} + R^\mu{}_{\lambda\rho\nu} = 0.$$

To see this, we introduce normal coordinates at a point p, where $\Gamma^\mu_{\nu\lambda}(p) = 0$. Then (at p)

$$\begin{aligned}R^\mu{}_{\nu\lambda\rho} &= \Gamma^\mu_{\nu\rho,\lambda} - \Gamma^\mu_{\nu\lambda,\rho}, \\ R^\mu{}_{\rho\nu\lambda} &= \Gamma^\mu_{\rho\lambda,\nu} - \Gamma^\mu_{\rho\nu,\lambda}, \\ R^\mu{}_{\lambda\rho\nu} &= \Gamma^\mu_{\lambda\nu,\rho} - \Gamma^\mu_{\lambda\rho,\nu}.\end{aligned}$$

Adding gives the desired result. This is true in arbitrary coordinates. This symmetry may be written concisely as

$$R^\mu{}_{[\rho\nu\lambda]} = 0.$$

We further have

$$R^\mu{}_{\rho\nu\lambda} = \frac{2}{3}(R^\mu{}_{(\rho\nu)\lambda} - R^\mu{}_{(\rho\lambda)\nu}).$$

To prove this, using normal coordinates the right hand side is

$$\begin{aligned}
&\frac{2}{3}\left(\frac{1}{2}\left[R^{\mu}{}_{\rho\nu\lambda} + R^{\mu}{}_{\nu\rho\lambda}\right] - \frac{1}{2}[R^{\mu}{}_{\rho\lambda\nu} + R^{\mu}{}_{\lambda\rho\nu}]\right) \\
&\quad = \frac{1}{3}\left(\Gamma^{\mu}_{\rho\lambda,\nu} - \Gamma^{\mu}_{\rho\nu,\lambda} + \Gamma^{\mu}_{\nu\lambda,\rho} - \Gamma^{\mu}_{\nu\rho,\lambda} - \Gamma^{\mu}_{\rho\nu,\lambda} + \Gamma^{\mu}_{\rho\lambda,\nu} - \Gamma^{\mu}_{\lambda\nu,\rho} + \Gamma^{\mu}_{\lambda\rho,\nu}\right) \\
&\quad = \frac{1}{3}\left(3\Gamma^{\mu}_{\rho\lambda,\nu} - 3\Gamma^{\mu}_{\rho\nu,\lambda}\right) \\
&\quad = R^{\mu}{}_{\rho\nu\lambda}.
\end{aligned}$$

Additionally, we have the Bianchi identities

$$R^{\mu}{}_{\rho[\nu\lambda;\gamma]} = 0,$$

equivalently

$$R^{\mu}{}_{\rho\nu\lambda;\gamma} + R^{\mu}{}_{\rho\gamma\nu;\lambda} + R^{\mu}{}_{\rho\lambda\gamma;\nu} = 0.$$

Adopting normal coordinates, then schematically we have

$$\begin{aligned}
R &= \partial\Gamma - \partial\Gamma + \Gamma\Gamma - \Gamma\Gamma, \\
\partial R &= \partial\partial\Gamma - \partial\partial\Gamma + \partial\Gamma\cdot\Gamma + \Gamma\cdot\partial\Gamma - \partial\Gamma\cdot\Gamma - \Gamma\cdot\partial\Gamma \\
&= \partial\partial\Gamma - \partial\partial\Gamma,
\end{aligned}$$

implying that

$$\begin{aligned}
R^{\mu}{}_{\rho\nu\lambda;\gamma} &= \Gamma^{\mu}_{\rho\lambda,\nu\gamma} - \Gamma^{\mu}_{\rho\nu,\lambda\gamma}, \\
R^{\mu}{}_{\rho\nu\lambda;\gamma} + R^{\mu}{}_{\rho\gamma\nu;\lambda} + R^{\mu}{}_{\rho\gamma\lambda;\nu} &= \Gamma^{\mu}_{\rho\lambda,\nu\gamma} - \Gamma^{\mu}_{\rho\nu,\lambda\gamma} + \Gamma^{\mu}_{\rho\nu,\gamma\lambda} \\
&\quad - \Gamma^{\mu}_{\rho\gamma,\nu\lambda} + \Gamma^{\mu}_{\rho\gamma,\lambda\nu} - \Gamma^{\mu}_{\rho\lambda,\gamma\nu} \\
&= 0
\end{aligned}$$

and is true for all charts.

10 The Metric

10.1 The Metric Tensor

A metric tensor g is a $\binom{0}{2}$ tensor such that

(i) the magnitude of the vector X is $|g(X,X)|^{\frac{1}{2}}$,

(ii) the angle between two vectors X, Y is

$$\cos^{-1}\left(\frac{g(X,Y)}{|g(X,X)|^{\frac{1}{2}}\,|g(Y,Y)|^{\frac{1}{2}}}\right), \qquad g(X,X) \neq 0, \quad g(Y,Y) \neq 0$$

and if $g(X,Y) = 0$, X, Y are orthogonal,

(iii) the length of a curve whose tangent vector is X between t_1 and t_2 is

$$\int_{t_1}^{t_2} |g(X,X)|^{\frac{1}{2}}\,\mathrm{d}t.$$

The metric gives the equivalent of the dot product, which is needed to define lengths of vectors and angles between them.

In a particular basis $\{e_\mu\}$, the metric tensor is written as

$$g(e_\mu, e_\nu) = g_{\mu\nu} = g_{\nu\mu}.$$

The metric is of fundamental importance in general relativity since it gives the interval between two spacetime points x^μ and $x^\mu + \mathrm{d}x^\mu$, which we call the line element.

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

Example 10.1.1. The line element in Euclidean space in Cartesian coordinates is

$$\begin{aligned}\mathrm{d}s^2 &= \mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z \\ &= \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \\ &= \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j, \qquad i, j = 1, 2, 3.\end{aligned}$$

$\delta_{ij} = \left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right)$ is the 3-dimensional Euclidean metric in Cartesian coordinates. In polar coordinates, which are related to Cartesian coordinates by

$$x = r\cos\phi\sin\theta, \quad y = r\sin\phi\sin\theta, \quad z = r\cos\theta,$$

the 1-forms $\mathrm{d}x^\mu$ transform as

$$\mathrm{d}x^i = \frac{\partial x^i}{\partial x^{i'}}\mathrm{d}x^{i'}.$$

$$\begin{aligned}\mathrm{d}x &= \frac{\partial x}{\partial r}\mathrm{d}r + \frac{\partial x}{\partial \theta}\mathrm{d}\theta + \frac{\partial x}{\partial \phi}\mathrm{d}\phi \\ &= \sin\theta\cos\phi\,\mathrm{d}r + r\cos\theta\cos\phi\,\mathrm{d}\theta - r\sin\theta\sin\phi\,\mathrm{d}\phi, \\ \mathrm{d}y &= \sin\theta\sin\phi\,\mathrm{d}r + r\cos\theta\sin\phi\,\mathrm{d}\theta + r\sin\theta\cos\phi\,\mathrm{d}\phi, \\ \mathrm{d}z &= \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta.\end{aligned}$$

Then

$$\begin{aligned}
\mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \\
&= \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\phi^2 \\
&= g_{i'j'}\mathrm{d}x^{i'}\mathrm{d}x^{j'},
\end{aligned}$$

so the metric in polar coordinates is

$$g_{i'j'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}.$$

The number of positive eigenvalues minus the number of negative eigenvalues gives the signature of the metric. This is basis independent.

If the absolute value of the signature is strictly less than the dimension of the manifold, then the manifold is said to be pseudo-Riemannian. (If they are equal, it is Euclidean or Riemannian.)

For example, for M^4 (4-dimensional Minkowski spacetime) in Cartesian coordinates,

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

The signature is therefore 2 and M^4 is a pseudo-Riemannian manifold. (In general relativity, we will be exclusively concerned with pseudo-Riemannian manifolds.)

10.2 The Inverse Metric

g is a non-degenerate map, since

$$g(X, Y) = 0 \quad \forall Y \quad \Rightarrow \quad X = 0.$$

Therefore g has an inverse whose components are $g^{\mu\nu}$, the components of a $\binom{0}{2}$ tensor satisfying

$$g^{\mu\nu}g_{\nu\lambda} = \delta^\mu_\lambda \qquad (g \cdot g^{-1} = \mathbb{1}).$$

The metric and its inverse define an isomorphism between the tangent space and dual space,

$$\begin{aligned}
T_p(\mathcal{M}) \cong T^*_p(\mathcal{M}), \quad X^\mu &\mapsto g_{\mu\nu}X^\nu \in T^*_p(\mathcal{M}), \\
w_\mu &\mapsto g^{\mu\nu}w_\nu \in T_p(\mathcal{M}),
\end{aligned}$$

i.e.

$$\begin{aligned} X_\mu &= g_{\mu\nu} X^\nu && \text{"lowering the index"}, \\ w^\mu &= g^{\mu\nu} w_\nu && \text{"raising the index"}. \end{aligned}$$

Metric components are used to raise and lower indices of a vector and covector.

This naturally generalises to tensors, e.g.

$$g^{\nu\gamma} T^{\mu}{}_{\gamma\lambda\rho} = T^{\mu\nu}{}_{\lambda\rho}.$$

10.3 The Lorentzian Signature

A Lorentzian signature is of the type $(-,+,+,+)$ (or $(+,+,+,-)$) where the negative eigenvalue is associated with the temporal direction and the positive eigenvalues are associated with the spatial directions.

We can always choose a basis at a point p such that the metric at p looks Minkowskian, i.e. $g_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$ at p. (Gravity doesn't exist locally.)

We say that a vector at p, X_p, is

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

Similarly, we say that the interval is timelike, null or spacelike depending on whether it is negative, zero or positive respectively.

(Using $(+,-,-,-)$, all of these are backwards.)

10.4 Metric Connections & Christoffel Symbols

Up until now the connection has been arbitrary (although some results relied on a symmetric connection), satisfying

$$\nabla_\mu e_\nu = \Gamma^\lambda_{\nu\mu} e_\lambda.$$

However, if the manifold is endowed with a metric there is a unique choice for the connection—the metric or Levi–Civita connection.

Theorem 10.4.1 (The fundamental theorem of Riemannian geometry)**.** If a manifold possesses a metric g, there exists a unique, torsion-free connection such that $\nabla g = 0$.

Proof. Assuming $\nabla g = 0$, we prove existence and uniqueness by explicitly constructing the unique connection. Let X, Y, Z be vector fields. Then $g(Y, Z)$ is a scalar and $X(g(Y, Z)) = \nabla_X g(Y, Z)$

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

Similarly,

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

Combining these,

$$\begin{aligned} &\frac{1}{2}\Big(X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y))\Big) \\ &\qquad = \frac{1}{2}g(\nabla_X Y, Z) + \frac{1}{2}g(Y, \nabla_X Z) \\ &\qquad\quad + \frac{1}{2}g(\nabla_Y Z, X) + \frac{1}{2}g(Z, \nabla_Y X) \\ &\qquad\quad - \frac{1}{2}g(\nabla_Z X, Y) - \frac{1}{2}g(X, \nabla_Z Y) \\ &\qquad = \frac{1}{2}g(Y, \nabla_X Z - \nabla_Z X) + \frac{1}{2}g(X, \nabla_Y Z - \nabla_Z Y) \\ &\qquad\quad + \frac{1}{2}g(Z, \nabla_Y X) + \frac{1}{2}g(\nabla_X Y, Z) \end{aligned}$$

since g is a bilinear form. Assuming vanishing torsion,

$$\begin{aligned} T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] = 0 \\ \Rightarrow \quad \nabla_X Y - \nabla_Y X = [X, Y]. \end{aligned}$$

Then the expression above equals

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

giving the Koszul formula

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

which defines the unique connection (as the metric is non-degenerate). $\square$

To determine the components, we recall that $[e_\mu, e_\nu] = \gamma^\lambda_{\mu\nu} e_\lambda$ and take

$$\begin{aligned} g(\nabla_\mu e_\nu, e_\lambda) &= g(\Gamma^\gamma_{\nu\mu} e_\gamma, e_\lambda) = \Gamma^\gamma_{\nu\mu} g_{\gamma\lambda} \\ &= \frac{1}{2}\Big[e_\mu(g_{\nu\lambda}) + e_\nu(g_{\mu\lambda}) - e_\lambda(g_{\mu\nu}) \\ &\qquad + g(e_\lambda, [e_\mu, e_\nu]) - g(e_\nu, [e_\mu, e_\lambda]) - g(e_\mu, [e_\nu, e_\lambda])\Big] \\ &= \frac{1}{2}\Big[e_\mu(g_{\nu\lambda}) + e_\nu(g_{\mu\lambda}) - e_\lambda(g_{\mu\nu}) + \gamma^\rho_{\mu\nu} g_{\lambda\rho} - \gamma^\rho_{\mu\lambda} g_{\nu\rho} - \gamma^\rho_{\nu\lambda} g_{\mu\rho}\Big]. \end{aligned}$$

In a coordinate-induced basis the commutator vanishes, so

$$\Gamma^\gamma_{\nu\mu} g_{\gamma\lambda} = \frac{1}{2}(g_{\nu\lambda,\mu} + g_{\mu\lambda,\nu} - g_{\mu\nu,\lambda}).$$

Multiplying by the inverse metric,

$$\begin{aligned} \Gamma^\gamma_{\nu\mu} g_{\gamma\lambda} g^{\rho\lambda} &= \Gamma^\gamma_{\nu\mu} \delta^\rho_\gamma \\ &= \frac{1}{2} g^{\rho\lambda}(g_{\nu\lambda,\mu} + g_{\mu\lambda,\nu} - g_{\mu\nu,\lambda}), \end{aligned}$$

so

$$\Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu} = \frac{1}{2} g^{\rho\lambda}(g_{\nu\lambda,\mu} + g_{\mu\lambda,\nu} - g_{\mu\nu,\lambda}).$$

This is the Christoffel symbol. It is sometimes written as

$$\Gamma^\lambda_{\mu\nu} = g^{\lambda\rho}{}_{[\mu\nu,\rho]}.$$

If $\nabla g = 0$, we say that the metric is compatible with the connection.

Example 10.4.1. We will compute the Christoffel symbols for the 3D flat space metric in Cartesian and polar coordinates. In the Cartesian case,

$$g_{ij} = \mathrm{diag}(1,1,1) \quad \Rightarrow \quad \Gamma^i_{jk} = 0$$

since they involve terms like $g_{ij,k}$. In polar coordinates

$$\begin{aligned} g_{ij} = \mathrm{diag}(1, r^2, r^2 \sin^2\theta), \quad \mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \\ &= \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2 + r^2 \sin^2\theta \, \mathrm{d}\phi^2 \end{aligned}$$

(see Example 10.1.1). The Christoffel symbols with an upper r are given by

$$\Gamma^r_{ij} = \frac{1}{2} g^{rr}(g_{ri,j} + g_{jr,i} - g_{ij,r}),$$

so

$$\Gamma^r_{\theta\theta} = -\frac{1}{2} g^{rr} g_{\theta\theta,r} = -r, \quad \Gamma^r_{rr} = 0, \quad \Gamma^r_{\phi\phi} = -r \sin^2\theta$$

and $\Gamma^r_{r\theta} = \Gamma^r_{r\phi} = \Gamma^r_{\theta\phi} = 0$. For $\Gamma^\theta_{ij} = \frac{1}{2} g^{\theta\theta}(g_{\theta i,j} + g_{j\theta,i} - g_{ij,\theta})$,

$$\Gamma^\theta_{rr} = 0, \quad \Gamma^\theta_{\theta\theta} = 0, \quad \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta,$$
$$\Gamma^\theta_{r\theta} = \frac{1}{r}, \quad \Gamma^\theta_{r\phi} = 0, \quad \Gamma^\theta_{\theta\phi} = 0.$$

Similarly, the Γ^ϕ_{ij} terms are

$$\Gamma^\phi_{rr} = 0, \quad \Gamma^\phi_{\theta\theta} = 0, \quad \Gamma^\phi_{\phi\phi} = 0,$$
$$\Gamma^\phi_{r\theta} = 0, \quad \Gamma^\phi_{r\phi} = \frac{1}{r}, \quad \Gamma^\phi_{\theta\phi} = \cot\theta.$$

10.5 Geodesics Revisited

Now that we have a metric, we have an appropriated notion of the length of a curve. If X is the tangent vector to a curve, then the components of X are $\frac{\mathrm{d}x^\mu}{\mathrm{d}s}$ and the length of the curve is

$$\int \mathrm{d}s \sqrt{|g(X,X)|} = \int \mathrm{d}s \sqrt{\left| g_{\mu\nu} \tfrac{\mathrm{d}x^\mu}{\mathrm{d}s} \tfrac{\mathrm{d}x^\nu}{\mathrm{d}s} \right|}.$$

Let us choose our parameterisation to be τ, such that

$$g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\nu}{\mathrm{d}s} = g_{\mu\nu} \dot{X}^\mu \dot{X}^\nu = \begin{cases} -1 & \dot{X}^\mu \text{ timelike} \\ 0 & \dot{X}^\mu \text{ null} \\ 1 & \dot{X}^\mu \text{ spacelike} \end{cases}.$$

In this case, τ is proper time for $\dot{X}^\mu$ timelike and proper length for $\dot{X}^\mu$ spacelike.

$$S = \int \mathrm{d}\tau \sqrt{\varepsilon g_{\mu\nu} \dot{X}^\mu \dot{X}^\nu}, \quad \varepsilon = \begin{cases} -1 & \text{timelike} \\ 1 & \text{spacelike} \end{cases}.$$

We extremise this length via the action principle. For timelike curves, we want to maximise the proper time.[3]

$$S = \int \mathrm{d}\tau \, L.$$

Take $\varepsilon = 1$. Then

$$L = L(X^\mu, \dot{X}^\mu) = \sqrt{g_{\mu\nu} \dot{X}^\mu \dot{X}^\nu}$$

and

$$\delta S = 0 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{x}^\mu} \right) - \frac{\partial L}{\partial x^\mu} = 0.$$

[3]Timelike curves cannot be minima of proper time, since they are infinitesimally close to null curves (of zero proper time).

This will give us the geodesic equation. We can read off the Christoffel symbols by comparing this to the normal geodesic equation.

We could use some other Lagrangian $f(L)$, where f is a continuous non-constant function. Then L and $f(L)$ satisfy the same Euler–Lagrange equation. We will use this to forget about the square root. Instead, take $L = g_{\mu\nu}\dot{X}^\mu \dot{X}^\nu$. $g_{\mu\nu}$ depends on X only.

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}^\lambda} &= g_{\mu\nu}\delta^\mu_\lambda \dot{X}^\nu + g_{\mu\nu}\dot{X}^\mu \delta^\nu_\lambda \\ &= g_{\lambda\nu}\dot{X}^\nu + g_{\mu\lambda}\dot{X}^\mu \\ &= 2g_{\nu\lambda}\dot{X}^\nu\end{aligned}$$

and

$$\frac{\partial L}{\partial x^\lambda} = g_{\mu\nu,\lambda}\dot{X}^\mu \dot{X}^\nu.$$

Therefore,

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\tau}(2g_{\nu\lambda}\dot{X}^\nu) - g_{\mu\nu,\lambda}\dot{X}^\mu \dot{X}^\nu = 0, \\ &2\frac{\mathrm{d}g_{\nu\mu}}{\mathrm{d}\tau}\dot{X}^\nu + 2g_{\nu\lambda}\ddot{X}^\nu - g_{\mu\nu,\lambda}\dot{X}^\mu \dot{X}^\nu = 0, \\ &2\frac{\partial g_{\nu\lambda}}{\partial x^\mu}\dot{X}^\mu \dot{X}^\nu + 2g_{\nu\lambda}\ddot{X}^\nu - g_{\mu\nu,\lambda}\dot{X}^\mu \dot{X}^\nu = 0, \\ &2g_{\nu\lambda}\ddot{X}^\nu + g_{\nu\lambda,\mu}\dot{X}^\mu \dot{X}^\nu + g_{\mu\lambda,\nu}\dot{X}^\mu \dot{X}^\nu - g_{\mu\nu,\lambda}\dot{X}^\mu \dot{X}^\nu = 0, \\ &g_{\nu\lambda}\ddot{X}^\nu + \frac{1}{2}\left(g_{\lambda\mu,\nu} + g_{\nu\lambda,\mu} - g_{\mu\nu,\lambda}\right)\dot{X}^\mu \dot{X}^\nu = 0.\end{aligned}$$

Multiplying by $g^{\lambda\rho}$ gives

$$\begin{aligned}\delta^\rho_\nu \ddot{X}^\nu + \frac{1}{2}g^{\lambda\rho}\left(g_{\lambda\mu,\nu} + g_{\nu\lambda,\mu} - g_{\mu\nu,\lambda}\right)\dot{X}^\mu \dot{X}^\nu = 0, \\ \ddot{X}^\rho + \Gamma^\rho_{\mu\nu}\dot{X}^\mu \dot{X}^\nu = 0,\end{aligned}$$

the geodesic equation. Similarly for timelike curves—both definitions are equivalent provided we parameterise using proper time.

10.6 Affine Parameters

Suppose we change the parameterisation from τ to $s(\tau)$. Then the geodesic equation becomes

$$\left(\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\nu\lambda}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}\right)\left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 + \left(\frac{\mathrm{d}^2 s}{\mathrm{d}\tau^2}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = 0,$$

as

$$\begin{aligned}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}\tau}\right) \\ &= \frac{\mathrm{d}^2x^\mu}{\mathrm{d}s^2}\left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 + \frac{\mathrm{d}^2s}{\mathrm{d}\tau^2}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}.\end{aligned}$$

We retrieve our standard form of the geodesic equation iff

$$\frac{\mathrm{d}s}{\mathrm{d}\tau} \neq 0, \quad \frac{\mathrm{d}^2s}{\mathrm{d}\tau^2} = 0 \quad \Rightarrow \quad s(\tau) = a\tau + b,$$

a, b constants, $a \neq 0$, so we are free to choose origin and scale. Such parameterisations are known as *affine parameters*. They are related to proper time or proper length in a linear way.

10.7 Metric Curvature

Assuming $\mathcal{M}$ has a metric, we can lower the upper index on the Riemann tensor. This gives us

$$R_{\mu\nu\lambda\rho} = g_{\mu\gamma}R^\gamma{}_{\nu\lambda\rho},$$

the metric curvature tensor. The previously mentioned symmetries still hold, i.e.

A. $R_{\mu\nu(\lambda\rho)} = 0$,

B. $R_{\mu[\nu\lambda\rho]} = 0$, etc.

The curvature tensor satisfies some additional symmetries when the manifold has a metric:

C. $R_{(\mu\nu)\lambda\rho} = 0$,

D. $R_{\mu\nu\lambda\rho} = R_{\lambda\rho\mu\nu}$.

How many independent components does $R_{\mu\nu\lambda\rho}$ have? Symmetries A, C and D imply that we can treat the curvature tensor like

$$R_{[\mu\nu][\lambda\rho]},$$

a symmetric $m \times m$ matrix with antisymmetric pairs of indices $\mu\nu$ and $\lambda\rho$ treated as individual indices. A symmetric $m \times m$ matrix has $\frac{1}{2}m(m+1)$ independent components, but each of these components are $n \times n$ antisymmetric matrices with $\frac{1}{2}n(n-1)$ components, so

$$\frac{1}{2}m(m+1) = \frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n).$$

Thinking about symmetry B,

$$\begin{aligned} R_{[\mu\nu\lambda\rho]} &= \frac{1}{4!}\left(3!R_{\mu[\nu\lambda\rho]} + 3!R_{\rho[\mu\nu\lambda]} + 3!R_{\lambda[\rho\mu\nu]} + 3!R_{\nu[\lambda\rho\mu]}\right) \\ &= \frac{3!}{4!}\left(R_{\mu[\nu\lambda\rho]} + R_{\rho[\mu\nu\lambda]} + R_{\lambda[\rho\mu\nu]} + R_{\nu[\lambda\rho\mu]}\right). \end{aligned}$$

Therefore $R_{\mu[\nu\lambda\rho]} = 0$ implies $R_{[\mu\nu\lambda\rho]} = 0$, so imposing $R_{[\mu\nu\lambda\rho]} = 0$ is equivalent to imposing $R_{\mu[\nu\lambda\rho]} = 0$ once the other symmetries have been imposed.

A totally antisymmetric 4-index tensor has $\frac{1}{4!}n(n-1)(n-2)(n-3)$ independent components. The total number of independent components is therefore

$$\frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) - \frac{1}{4!}n(n-1)(n-2)(n-3) = \frac{1}{2}n^2(n^2-1).$$

Gravity will be described by this curvature tensor.

1D	0	independent components	
2D	1	independent component	~ only calculate R_{1212}
3D	6	independent components	
4D	20	independent components	

A manifold $\mathcal{M}$ is locally flat if there exists a chart such that a line element in this chart is

$$\mathrm{d}s^2 = \epsilon_1(\mathrm{d}x^1)^2 + \cdots + \epsilon_n(\mathrm{d}x^n)^2, \qquad \epsilon_i = \pm 1, \quad i = 1, \ldots, n.$$

In this case, $R_{\mu\nu\lambda\rho} = 0$. Conversely, if $R_{\mu\nu\lambda\rho} = 0$ we can show that there exists a chart such that $\mathrm{d}s^2$ is as above and we say the metric is flat.

$$\text{Flat metric} \quad \Leftrightarrow \quad R_{\mu\nu\lambda\rho} = 0.$$

Contracting the first and third indices of the Riemann curvature tensor defines the Ricci curvature tensor

$$R_{\mu\nu} = R^{\gamma}{}_{\mu\gamma\nu}.$$

Contracting with the metric yields the Ricci curvature scalar

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

The Einstein curvature tensor is defined by

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

($G_{\mu\nu} = 0$ is a differential equation for components of the metric. Solving this gives vacuum solutions of Einstein's equations.)

Recall the Bianchi identities, $R^{\mu}{}_{\nu[\lambda\rho;\gamma]} = 0$ or

$$R^{\mu}{}_{\nu\lambda\rho;\gamma} + R^{\mu}{}_{\nu\gamma\lambda;\rho} + R^{\mu}{}_{\nu\rho\gamma;\lambda} = 0.$$

Contracting over μ and ρ,

$$R^{\mu}{}_{\nu\lambda\mu;\gamma} + R^{\mu}{}_{\nu\gamma\lambda;\mu} + R^{\mu}{}_{\nu\mu\gamma;\lambda} = -R_{\nu\lambda;\gamma} + R^{\mu}{}_{\nu\gamma\lambda;\mu} + R_{\nu\gamma;\lambda} = 0,$$

the "once-contracted Bianchi identities". Multiplying by $g^{\nu\lambda}$,

$$-R_{;\gamma} + \left(g^{\nu\lambda} R^{\mu}{}_{\nu\gamma\lambda}\right)_{;\mu} + R^{\lambda}{}_{\gamma;\lambda} = 0,$$

and by $g^{\nu\gamma}$,

$$\begin{aligned} -(g^{\nu\gamma}R)_{;\gamma} + R^{\mu\nu}{}_{;\mu} + R^{\lambda\nu}{}_{;\lambda} &= 0, \\ -(g^{\nu\gamma}R)_{;\gamma} + 2R^{\nu\gamma}{}_{;\gamma} &= 0, \\ G^{\nu\gamma}{}_{;\gamma} &= 0, \end{aligned}$$

the twice-contracted Bianchi identities. The Einstein tensor is conserved.

Finally, we introduce the Weyl tensor defined by

$$C_{\mu\nu\lambda\rho} = R_{\mu\nu\lambda\rho} - \tfrac{2}{n-2}\left(g_{\mu[\lambda}R_{\rho]\nu} - g_{\nu[\lambda}R_{\rho]\mu}\right) + \tfrac{2}{(n-1)(n-2)} R g_{\mu[\lambda}g_{\rho]\nu},$$

also called the "conformal tensor". We can show that the Weyl tensor satisfies all the same symmetries, for example

$$C_{\mu\nu\lambda\rho} = C_{[\mu\nu][\lambda\rho]}, \qquad C_{\mu\nu\lambda\rho} = C_{\lambda\rho\mu\nu}, \qquad C_{\mu[\nu\lambda\rho]} = 0.$$

We can further show that all possible contractions with the metric vanish,

$$g^{\mu\lambda}C_{\mu\nu\lambda\rho} = g^{\mu\nu}C_{\mu\nu\lambda\rho} = g^{\mu\rho}C_{\mu\nu\lambda\rho} = \cdots = 0.$$

Also, the Weyl tensor is invariant under conformal transformations of the metric

$$g_{\mu\nu} \to \Omega^2 g_{\mu\nu}.$$

11 Isometries and Killing Vectors

11.1 Lie Derivatives Revisited

The expression for the Lie derivative of an arbitrary $\binom{r}{l}$ tensor field is

$$\begin{aligned} \mathcal{L}_V T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_l} = {} & V^{\lambda}\partial_{\lambda} T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_l} \\ & - (\partial_{\lambda}V^{\mu_1}) T^{\lambda\mu_2\cdots\mu_r}{}_{\nu_1\cdots\nu_l} - (\partial_{\lambda}V^{\mu_2}) T^{\mu_1\lambda\cdots\mu_r}{}_{\nu_1\cdots\nu_l} - \cdots \\ & + (\partial_{\nu_1}V^{\lambda}) T^{\mu_1\cdots\mu_r}{}_{\lambda\nu_2\cdots\nu_l} + (\partial_{\nu_2}V^{\lambda}) T^{\mu_1\cdots\mu_r}{}_{\nu_1\lambda\cdots\nu_l} + \cdots . \end{aligned}$$

This involves partial derivatives rather than covariant derivatives and is not manifestly tensorial. We can in fact replace the partial derivatives with the covariant derivative (with a symmetric connection) since all connection terms vanish.

$$\begin{aligned}\mathcal{L}_V T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_l} = V^\lambda \nabla_\lambda T^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_l} - (\nabla_\lambda V^{\mu_1}) T^{\lambda\mu_2\cdots\mu_r}{}_{\nu_1\cdots\nu_l} - \cdots \\ + (\nabla_{\nu_1} V^\lambda) T^{\mu_1\cdots\mu_r}{}_{\lambda\nu_2\cdots\nu_l} + \cdots .\end{aligned}$$

Example 11.1.1. We will show that the two representations of a Lie derivative of a vector field are equivalent.

$$\begin{aligned}\mathcal{L}_V X^\mu &= \nabla_\lambda X^\mu V^\lambda - \nabla_\lambda V^\mu X^\lambda \\ &= (X^\mu{}_{,\lambda} + \Gamma^\mu_{\nu\lambda} X^\nu) V^\lambda - (V^\mu{}_{,\lambda} + \Gamma^\mu_{\nu\lambda} V^\nu) X^\lambda \\ &= X^\mu{}_{,\lambda} V^\lambda - V^\mu{}_{,\lambda} X^\lambda + \Gamma^\mu_{\nu\lambda} X^\nu V^\lambda - \Gamma^\mu_{\nu\lambda} V^\nu X^\lambda \\ &= X^\mu{}_{,\lambda} V^\lambda - V^\mu{}_{,\lambda} X^\lambda,\end{aligned}$$

which is the manifestly non-tensorial representation of the Lie derivative. (It's still a tensor, but just involves partial derivatives.)

The manifestly tensorial definition is often more useful.

11.2 Lie Derivative of the Metric and Isometries

The Lie derivative of the metric tensor is

$$\begin{aligned}\mathcal{L}_V g_{\mu\nu} &= V^\lambda (\nabla_\lambda g_{\mu\nu}) + (\nabla_\mu V^\lambda) g_{\lambda\nu} + (\nabla_\nu V^\nu) g_{\mu\lambda} \\ &= (\nabla_\mu V^\lambda) g_{\lambda\nu} + (\nabla_\nu V^\nu) g_{\mu\lambda} \\ &= \nabla_\mu (V^\lambda g_{\lambda\nu}) + \nabla_\nu (V^\nu g_{\mu\lambda}) \\ &= \nabla_\mu V_\nu + \nabla_\nu V_\mu \\ &= 2\nabla_{(\mu} V_{\nu)}.\end{aligned}$$

We say a diffeomorphism h is a symmetry of some tensor T if the tensor is invariant after being pulled back under h,

$$h^* T = T.$$

Suppose we have a continuous one-parameter family of symmetries h_t generated by a vector field V^μ. Then invariance under this symmetry implies that the tensor does not change along integral curves of V^μ, i.e. $\mathcal{L}_V T = 0$.

Symmetric diffeomorphisms of the metric tensor are known as isometries. Suppose that V^μ generates a one-parameter family of isometries. Then

$$\mathcal{L}_V g_{\mu\nu} = 2\nabla_{(\mu} V_{\nu)} = 0 \quad \Leftrightarrow \quad \nabla_{(\mu} V_{\nu)} = 0. \qquad \text{(Killing's equation)}$$

Solutions V^μ are known as Killing (co-)vectors.

12 p-Forms

12.1 Space of p-Forms

A differential p-form is a $\binom{0}{p}$ tensor which is completely antisymmetric. For example, a 2-form is an antisymmetric $\binom{0}{2}$ tensor satisfying

$$T(X,Y) = -T(Y,X)$$

(for X, Y vector fields) or, in a coordinate-induced basis,

$$T = T_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$$

for $T_{\mu\nu} = -T_{\nu\mu}$.

Lemma 12.1.1. If $T(X,X) = 0\ \forall X$ then T is antisymmetric.

Proof. Let $X = Y + Z$. Then

$$\begin{aligned} T(Y+Z, Y+Z) &= T(Y,Y) + T(Y,Z) + T(Z,Y) + T(Z,Z) \\ &= T(Y,Z) + T(Z,Y) = 0 \end{aligned}$$

and hence T is antisymmetric. □

The converse is obviously true: if T is antisymmetric then $T(X,X) = 0\ \forall X$. Repeated indices in a p-form mean certain death.

As an aside, consider the pull-back of a tensor. If T is a $\binom{0}{r}$ tensor, what is h^*T? For a $\binom{0}{1}$ tensor,

$$h^*\eta(X) = \eta(h_*(X)).$$

This has an obvious generalisation to

$$h^*T(X_1, \ldots, X_r) = T(h_*X_1, \ldots, h_*X_r).$$

Consider the number of linearly independent components. For $p = 2$ and dimension n, we have

$$T = T_{\mu\nu}e^\mu \otimes e^\nu, \qquad \mu, \nu = 1, \ldots, n.$$

For $n = 2$,

$$T_{\mu\nu} = -T_{\nu\mu} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} 0 & T_{12} \\ -T_{12} & 0 \end{pmatrix}$$

as antisymmetry implies $T_{11} = T_{22} = 0$. Thus $T_{\mu\nu} = \left(\begin{smallmatrix} 0 & k \\ -k & 0 \end{smallmatrix}\right)$. So for $p = 2$, $n = 2$, there is one independent component. For $p = 2$, $n = 3$, we would have $\frac{n(n-1)}{2} = 3$ independent components.

In general, the number of linearly independent components of a p-form in an n-dimensional space is

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}$$

for $n \geq p$. For $p > n$, all components are zero.

We denote the space of all p-forms by Λ_p and the space of all p-form fields over a manifold $\mathcal{M}$ by $\Lambda_p(\mathcal{M})$.

Note that a 0-form is a function and a 1-form a covector field.

12.2 Wedge Products

Given a p-form P and a q-form Q, we can form a $(p+q)$-form known as the wedge product $P \wedge Q$ by taking the antisymmetric tensor product.

For example, the wedge product of two 1-forms is

$$P \wedge Q = P \otimes Q - Q \otimes P, \qquad P, Q \in T^*(\mathcal{M}).$$

Lemma 12.2.1. The set $\{e^{\mu_1} \wedge e^{\mu_2} \wedge \cdots \wedge e^{\mu_p}\}$ form a basis for Λ_p.

Proof. Clearly, $e^{\mu} \wedge e^{\mu} = 0$ by antisymmetry. Therefore there are $n(n-1)(n-2)\cdots(n-(p-1))$ elements. However, these elements are antisymmetric so that the number of independent components is

$$\frac{n(n-1)\cdots(n-(p-1))}{(n-p)!} = \binom{n}{p} = \dim(\Lambda_p).$$

Therefore $\{e^{\mu_1} \wedge \cdots \wedge e^{\mu_p}\}$ span Λ_p.

For linear independence, we consider

$$\begin{aligned} c_{\mu_1\mu_2\cdots\mu_p} & e^{\mu_1} \wedge e^{\mu_2} \wedge \cdots \wedge e^{\mu_p}(e_{\nu_1}, e_{\nu_2}, \ldots, e_{\nu_p}) \\ &= c_{\mu_1\mu_2\cdots\mu_p}(e^{\mu_1} \otimes e^{\mu_2} \otimes \cdots \otimes e^{\mu_p} - e^{\mu_2} \otimes e^{\mu_1} \otimes \cdots \otimes e^{\mu_p} \\ &\qquad\qquad + \text{all other permutations})(e_{\nu_1}, e_{\nu_2}, \ldots, e_{\nu_p}) \\ &= c_{\mu_1\mu_2\cdots\mu_p}(\delta^{\mu_1}_{\nu_1}\delta^{\mu_2}_{\nu_2}\cdots\delta^{\mu_p}_{\nu_p} - \delta^{\mu_2}_{\nu_1}\delta^{\mu_1}_{\nu_2}\cdots\delta^{\mu_p}_{\nu_p} \\ &\qquad\qquad + \text{all other permutations}) \\ &= c_{\nu_1\nu_2\cdots\nu_p} - c_{\nu_2\nu_1\cdots\nu_p} + \text{all other permutations} \end{aligned}$$

where $\langle e^\mu, e_\nu \rangle = \delta^\mu_\nu$. But $c_{\mu_1\mu_2\cdots\mu_p} = c_{[\mu_1\mu_2\cdots\mu_p]}$, therefore

$$c_{\mu_1\mu_2\cdots\mu_p} e^{\mu_1} \wedge e^{\mu_2} \wedge \cdots \wedge e^{\mu_p}(e_{\nu_1}, e_{\nu_2}, \ldots, e_{\nu_p}) = p!\, c_{\nu_1\nu_2\cdots\nu_p}.$$

If the left hand side is zero, this implies

$$c_{\mu_1\mu_2\cdots\mu_p} = 0.$$

Therefore $\{e^{\mu_1} \wedge \cdots \wedge e^{\mu_p}\}$ are linearly independent and form a basis for Λ_p. □

In a coordinate-induced basis, for example, we would have

$$\{e^{\mu_1} \wedge \cdots \wedge e^{\mu_p}\} = \{\mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_p}\}.$$

If $T \in \Lambda_p$, we write

$$T = \frac{1}{p!} T_{\mu_1\cdots\mu_p} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_p}$$

where the components $T_{\mu_1\cdots\mu_p} = T_{[\mu_1\cdots\mu_p]}$ are completely antisymmetric.

As an example, we will derive an explicit component form for the wedge product of a p-form A and a q-form B.

$$\begin{aligned}
(A \wedge B)_{\nu_1\cdots\nu_{p+q}} &= (A \wedge B)(\partial_{\nu_1}, \ldots, \partial_{\nu_{p+q}}) \\
&= \frac{1}{p!} A_{\mu_1\cdots\mu_p} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_p} \wedge \frac{1}{q!} B_{\mu_{p+1}\cdots\mu_{p+q}} \mathrm{d}x^{\mu_{p+1}} \wedge \cdots \wedge \mathrm{d}x^{\mu_{p+q}} \\
&\qquad\qquad (\partial_{\nu_1}, \ldots, \partial_{\nu_{p+q}}) \\
&= \frac{1}{p!q!} A_{\mu_1\cdots\mu_p} B_{\mu_{p+1}\cdots\mu_{p+q}} dx^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_p} \wedge \mathrm{d}x^{\mu_{p+1}} \wedge \cdots \wedge \mathrm{d}x^{\mu_{p+q}} \\
&\qquad\qquad (\partial_{\nu_1}, \ldots, \partial_{\nu_{p+q}}) \\
&= \frac{1}{p!q!} A_{\mu_1\cdots\mu_p} B_{\mu_{p+1}\cdots\mu_{p+q}} (dx^{\mu_1} \otimes \mathrm{d}x^{\mu_2} \otimes \cdots \otimes \mathrm{d}x^{\mu_{p+q}} \\
&\qquad - \mathrm{d}x^{\mu_2} \otimes \mathrm{d}x^{\mu_1} \otimes \cdots \otimes \mathrm{d}x^{\mu_{p+q}} \\
&\qquad + \text{all other permutations})(\partial_{\nu_1}, \ldots, \partial_{\nu_{p+q}}) \\
&= \frac{1}{p!q!} A_{\mu_1\cdots\mu_p} B_{\mu_{p+1}\cdots\mu_{p+q}} (\delta^{\mu_1}_{\nu_1} \delta^{\mu_2}_{\nu_2} \cdots \delta^{\mu_{p+q}}_{\nu_{p+q}} - \delta^{\mu_2}_{\nu_1} \delta^{\mu_1}_{\nu_2} \cdots \delta^{\mu_{p+1}}_{\nu_{p+q}} \\
&\qquad + \text{all other permutations}) \\
&= \frac{(p+q)!}{p!q!} A_{[\nu_1\cdots\nu_p} B_{\nu_{p+1}\cdots\nu_{p+q}]}.
\end{aligned}$$

So, for example, if A and B are 1-forms, then

$$\begin{aligned}(A\wedge B)_{\mu\nu} &= \frac{3!}{1!1!}A_{[\mu}B_{\nu]}\\ &= 2\tfrac{1}{2}(A_\mu B_\nu - A_\nu B_\mu) = A_\mu B_\nu - A_\nu B_\mu.\end{aligned}$$

Another property of the wedge product is

$$A\wedge B = (-1)^{pq} B\wedge A, \qquad A\in\Lambda_p, \quad B\in\Lambda_q.$$

12.3 Exterior Derivative

We may define a map d known as the exterior derivative such that

$$\mathrm{d}:\Lambda_p\to\Lambda_{p+1}$$

where the components of d are appropriately normalised antisymmetric partial derivatives.

$$(\mathrm{d}A)_{\mu_1\cdots\mu_{p+1}} = (p+1)\partial_{[\mu_1}A_{\mu_2\cdots\mu_{p+1}]}.$$

d satisfies the following properties:

(1) $\mathrm{d}:\Lambda_0\to\Lambda_1$ (recall $\mathrm{d}f = \left(\frac{\partial F}{\partial x^\mu}\right)\mathrm{d}x^\mu$, §3.3);

(2) $\mathrm{d}(A+B) = \mathrm{d}(A)+\mathrm{d}(B) \quad \forall A\in\Lambda_p, B\in\Lambda_q$;

(3) Liebniz rules: $\mathrm{d}(fA) = \mathrm{d}f\wedge A + f\wedge \mathrm{d}A$,
$\mathrm{d}(A\wedge B) = \mathrm{d}A\wedge B + (-1)^p A\wedge \mathrm{d}B$;

(4) $\mathrm{d}^2 A = 0$.

A heuristic explanation of (4) is that d^2 is symmetric, while A is antisymmetric.

As an example, take $\eta = xy\,\mathrm{d}x + y^2\mathrm{d}y$.

$$\begin{aligned}\mathrm{d}\eta &= \mathrm{d}(xy\,\mathrm{d}x) + \mathrm{d}(y^2\,\mathrm{d}y)\\ &= \mathrm{d}(xy)\wedge\mathrm{d}x + xy\,\mathrm{d}^2x + \mathrm{d}(y^2)\wedge\mathrm{d}y + y^2\,\mathrm{d}^2y\\ &= (y\,\mathrm{d}x + x\,\mathrm{d}y)\wedge\mathrm{d}x + 2y\,\mathrm{d}y\wedge\mathrm{d}y\\ &= x\,\mathrm{d}y\wedge\mathrm{d}x\\ &= -x\,\mathrm{d}x\wedge\mathrm{d}y,\end{aligned}$$

$$\begin{aligned}\mathrm{d}^2\eta &= \mathrm{d}(x\,\mathrm{d}y)\wedge\mathrm{d}x - x\,\mathrm{d}y\wedge\mathrm{d}^2x\\ &= \mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}x + x\,\mathrm{d}^2y\wedge\mathrm{d}x\\ &= 0.\end{aligned}$$

12.4 Closed and Exact Forms

Let $A \in \Lambda_p$. If $\mathrm{d}A = 0$, we say that A is closed. We say that A is exact if there exists a $B \in \Lambda_{p-1}$ such that $A = \mathrm{d}B$.

Lemma 12.4.1 (Poincaré)**.** If A is an exact p-form then A is closed.

Proof. Exact implies $A = \mathrm{d}B$.

$$\mathrm{d}A = \mathrm{d}^2 B = 0,$$

therefore A is closed. □

The converse is not necessarily true, although in topologically trivial spacetimes it will be.

12.5 Hodge Duality

We must first introduce the Levi–Civita alternating symbol

$$\tilde{\varepsilon}^{\mu_1 \cdots \mu_p} = \begin{cases} +1 & \text{if } \mu_1 \cdots \mu_p \text{ is an even permutation of } 1 \cdots p \\ -1 & \text{if } \mu_1 \cdots \mu_p \text{ is an odd permutation of } 1 \cdots p \\ 0 & \text{if any index is repeated} \end{cases}.$$

This does not transform like the components of a tensor, but rather like a tensor density of weight -1 (see §13.1).

We can construct a well defined tensor by multiplying by an appropriate scalar density. The Levi–Civita tensor is

$$\varepsilon^{\mu_1 \cdots \mu_p} = \sqrt{-g}\,\tilde{\varepsilon}^{\mu_1 \cdots \mu_p},$$

where $g = \det(g_{\mu\nu})$. We are assuming the manifold is Lorentzian and has a metric.

The Hodge star operator is a map from p-forms to $(n-p)$-forms, defined by

$$(\star A)_{\mu_1 \cdots \mu_{n-p}} = \frac{1}{p!} \varepsilon^{\nu_1 \cdots \nu_p}{}_{\mu_1 \cdots \mu_{n-p}} A_{\nu_1 \cdots \nu_p}.$$

Note that this is metric dependent. Applying the Hodge star twice returns plus or minus the original form,

$$\star \star A = (-1)^{s + p(n-p)} A,$$

where s is the number of minus signs in the eigenvalues of the metric.

Note that in 3D Euclidean space we have, for A, B 1-forms,

$$\star(A \wedge B)_\mu = \varepsilon_\mu{}^{\nu\lambda} A_\nu B_\lambda$$

which is the conventional cross product.

Example 12.5.1 (Electromagnetism)**.** Maxwell's equations in tensor notation are

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{A}$$

$$\partial_{[\mu} F_{\nu\lambda]} = 0, \tag{B}$$

where $F_{\mu\nu} = -F_{\nu\mu}$. Equation B is most succinctly written as the closure of a 2-form,

$$\mathrm{d}F = 0.$$

For topologically trivial spaces (such as Minkowski spacetime), closed implies exact so that there is a 1-form $A = A_\mu \mathrm{d}x^\mu$ with

$$F = \mathrm{d}A.$$

A_μ are the components of the vector potential.

Gauge invariance is now a consequence of the fact that F is invariant under $A \to A + \mathrm{d}\lambda$.

Finally, equation A can be expressed as

$$\mathrm{d}(\star F) = \star J.$$

13 Integration on Manifolds

13.1 Tensor Densities

We denote two charts by x^μ and $x^{\hat{\mu}}$ and we have

$$P^\mu{}_{\hat{\mu}} = \frac{\partial x^\mu}{\partial x^{\hat{\mu}}}, \quad P^{\hat{\mu}}{}_\mu = \frac{\partial x^{\hat{\mu}}}{\partial x^\mu}$$

which are matrices satisfying $P^\mu{}_{\hat{\nu}} P^{\hat{\nu}}{}_\rho = \delta^\mu_\rho$. We denote the determinant of P by $\mathcal{J} = \det(P^\mu{}_{\hat{\mu}}) \neq 0$. Recall the transformation law (§3.5)

$$T^{\hat{\mu}}{}_{\hat{\nu}} = P^{\hat{\mu}}{}_\mu P^\nu{}_{\hat{\nu}} T^\mu{}_\nu.$$

We say that $\mathfrak{T}^\mu{}_\nu$ is a $\binom{1}{1}$ tensor density of weight w if it transforms like

$$\mathfrak{T}^{\hat{\mu}}{}_{\hat{\nu}} = \mathcal{J}^w P^{\hat{\mu}}{}_\mu P^\nu{}_{\hat{\nu}} \mathfrak{T}^\mu{}_\nu.$$

The product of an $\binom{m}{k}$ tensor density of weight w_1 and an $\binom{n}{l}$ tensor density of weight w_2 is an $\binom{n+m}{k+l}$ tensor density of weight $w_1 + w_2$.

Claim 13.1.1. $\sqrt{-g}$ is a scalar density of weight 1.

Proof.

$$\begin{aligned} g_{\hat\mu\hat\nu} &= P^\mu{}_{\hat\mu} P^\nu{}_{\hat\nu} g_{\mu\nu}, \quad g = \det(g_{\mu\nu}), \\ \hat g &= \det(g_{\hat\mu\hat\nu}) \\ &= \det(P^\mu{}_{\hat\mu} P^\nu{}_{\hat\nu} g_{\mu\nu}) \\ &= \det(P^\mu{}_{\hat\mu}) \det(P^\nu{}_{\hat\nu}) \det(g_{\mu\nu}) \\ &= \mathcal{J}^2 g. \end{aligned}$$

Therefore $\hat g$ is a scalar density of weight 2 and $\sqrt{-\hat g} = \mathcal{J}\sqrt{-g}$, so $\sqrt{-g}$ is a scalar density of weight 1. □

One can derive a covariant derivative of a tensor density, for example

$$\nabla_\mu \mathfrak{T}^\nu{}_\lambda = \mathfrak{T}^\nu{}_{\lambda,\mu} + \Gamma^\nu_{\rho\mu}\mathfrak{T}^\rho{}_\lambda - \Gamma^\rho_{\lambda\mu}\mathfrak{T}^\nu{}_\rho - w\Gamma^\rho_{\rho\mu}\mathfrak{T}^\nu{}_\lambda.$$

In general, the covariant derivative of a tensor density need not be a tensor density.

Another important example of a tensor density is the Levi–Civita alternating symbol $\tilde\varepsilon_{\mu_1\cdots\mu_n}$. One can show that[4]

$$\tilde\varepsilon_{\hat\mu_1\cdots\hat\mu_n} = \mathcal{J}^{-1} P^{\mu_1}{}_{\hat\mu_1} \cdots P^{\mu_n}{}_{\hat\mu_n} \tilde\varepsilon_{\mu_1\cdots\mu_n}$$

and therefore is a tensor of weight -1.

13.2 Volume and Surface Elements

Recall that in Euclidean space, the area enclosed by two vectors at p, X_p and Y_p say, is given by

$$\det\begin{pmatrix} X^1 & X^2 \\ Y^1 & Y^2 \end{pmatrix} = X^1Y^2 - Y^1X^2.$$

Consider the wedge product $\mathrm{d}x^1 \wedge \mathrm{d}x^2$ acting on X_p, Y_p.

$$\begin{aligned} (\mathrm{d}x^1 \wedge \mathrm{d}x^2)(X_p, Y_p) &= (\mathrm{d}x^1 \otimes \mathrm{d}x^2)(X_p, Y_p) - (\mathrm{d}x^2 \otimes \mathrm{d}x^1)(X_p, Y_p) \\ &= \mathrm{d}x^1(X_p)\,\mathrm{d}x^2(Y_p) - \mathrm{d}x^2(X_p)\,\mathrm{d}x^1(Y_p) \\ &= X^1Y^2 - X^2Y^1. \end{aligned}$$

Hence, in $\mathbb{R}^2$ $\mathrm{d}x^1 \wedge \mathrm{d}x^2$ is called the area form. We note that we may rewrite this as

$$\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \frac{1}{2!}\tilde\varepsilon_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu.$$

[4] See Carroll's notes (`preposterousuniverse.com/grnotes/grnotes-two.pdf`), with the caveat $\mathcal{J} \to \mathcal{J}^{-1}$.

Moving now to $\mathbb{R}^3$, we recall that the volume V enclosed by vectors X_p, Y_p, Z_p is given by

$$V = (X_p \times Y_p) \cdot Z_p = \tilde{\varepsilon}_{\mu\nu\lambda} X^\mu Y^\nu Z^\lambda.$$

Consider now $\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3(X_p, Y_p, Z_p)$

$$\begin{aligned}
&= (\mathrm{d}x^1 \otimes \mathrm{d}x^2 \otimes \mathrm{d}x^3 - \mathrm{d}x^1 \otimes \mathrm{d}x^3 \otimes \mathrm{d}x^2 + \mathrm{d}x^3 \otimes \mathrm{d}x^1 \otimes \mathrm{d}x^2 \\
&\quad - \mathrm{d}x^3 \otimes \mathrm{d}x^2 \otimes \mathrm{d}x^1 + \mathrm{d}x^2 \otimes \mathrm{d}x^3 \otimes \mathrm{d}x^1 - \mathrm{d}x^2 \otimes \mathrm{d}x^1 \otimes \mathrm{d}x^3)(X_p, Y_p, Z_p) \\
&= X^1Y^2Z^3 - X^1Y^3Z^2 + X^3Y^1Z^2 - X^3Y^2Z^1 + X^2Y^3Z^1 - X^2Y^1Z^3 \\
&= Z^1(X^2Y^3 - X^3Y^2) + Z^2(X^3Y^1 - X^1Y^3) + Z^3(X^1Y^2 - X^2Y^1) \\
&= \tilde{\varepsilon}_{\mu\nu\lambda} X^\mu Y^\nu Z^\lambda.
\end{aligned}$$

$\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$ is the volume form in $\mathbb{R}^3$. Again, we rewrite this as

$$\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 = \frac{1}{3!}\tilde{\varepsilon}_{\mu\nu\lambda} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\lambda = \mathrm{d}\Omega_3.$$

This volume element is not chart independent, however, since the wedge product $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \wedge \mathrm{d}x^\lambda$ transforms like a tensor while $\tilde{\varepsilon}_{\mu\nu\lambda}$ transforms like a tensor density of weight -1.

For a volume integral $\int f \, \mathrm{d}\Omega$ to make sense and be chart independent, f must be a scalar density of weight 1.

For a Lorentzian metric, for example, we can consider integrals of the type

$$\int f\sqrt{-g}\,\mathrm{d}\Omega$$

where f is a scalar and

$$\sqrt{-g}\,\mathrm{d}\Omega = \sqrt{-g}\frac{1}{n!}\tilde{\varepsilon}_{\mu_1\cdots\mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n}$$

is the n-dimensional curved spacetime volume element.

Similarly for surface integrals in $\mathbb{R}^r$, the surface element is

$$\mathrm{d}\Sigma_\mu = \frac{1}{(n-1)!}\tilde{\varepsilon}_{\mu\lambda_1\cdots\lambda_{n-1}} \mathrm{d}x^{\lambda_1} \wedge \cdots \wedge \mathrm{d}x^{\lambda_{n-1}}$$

which transforms as a covector density of weight -1. Therefore, for a general curved space integral of the type $\int X^\mu \, \mathrm{d}\Sigma_\mu$ to make sense, X^μ has to be a vector density of weight $+1$.

Again, for a Lorentzian metric, we can consider integrals of the type

$$\int X^\mu \sqrt{-g}\,\mathrm{d}\Sigma_\mu$$

where X^μ is a vector and $\sqrt{-g}\,\mathrm{d}\Sigma_\mu$ is the curved space surface element.

13.3 Gauss' Divergence Theorem

In $\mathbb{R}^n$, the divergence theorem is

$$\int_V (\nabla \cdot F)\, \mathrm{d}V = \oint_S (F \cdot n)\, \mathrm{d}S.$$

The curved space generalisation is

$$\int_V \nabla_\mu \tilde{\chi}^\mu \, \mathrm{d}\Omega = \int_{\partial V} \tilde{\chi}^\mu \, \mathrm{d}\Sigma_\mu$$

for $\tilde{\chi}^\mu$ a vector density of weight 1.

Then for a Lorentzian metric, we have

$$\int_V \nabla_\mu X^\mu \, \mathrm{d}\Omega = \int_{\partial V} \sqrt{-g} X^\mu \, \mathrm{d}\Sigma_\mu$$

for X^μ a vector.

What next? You can get Debbie Ip's general relativity notes from

`www.maths.tcd.ie/~ipde/GR_Notes.pdf`.