Tangent, Cotangent and Tensor Bundles

In order to systematically keep track of all tangent vectors at all points we introduce the tangent bundle:

Definition 9.10 *The* tangent bundle TM *of a manifold* M *of dimension* n *is the set*

$$TM := \bigcup_{p \in M} \{p\} \times T_pM = \{(p,v)|\ p \in M,\ v \in T_pM\},$$

where the union is disjoint (so $(p,v) = (p',v')$ *if and only if* $p = p'$ *and* $v = v'$*). We view* TM *as a manifold of dimension* $2n$ *with a maximal atlas containing all charts of the form*

$$D\psi : TU \to V \times \mathbb{R}^n : (p,v) \mapsto (\psi(p), D_p\psi(v))$$

such that $v = (D_p\psi(v))^\mu X_\mu$ *in the coordinate basis of* T_pM *determined by* ψ.

To see that the manifold structure of TM is well defined we notice that for any charts $\psi : U \to V$ and $\psi' : U' \to V'$ with $U \cap U' \neq \emptyset$, the change of charts from $D\psi$ to $D\psi'$ is

$$(D\psi \circ (D\psi')^{-1})(\psi'(p), D_p\psi'(v)) = (\psi(p), D_p\psi(v)),$$

which is a diffeomorphism. To see this one may write it in component form and use Equation (15) for the vector components.

If $\chi : M \to M'$ is a C^∞ map, we may define the *tangent map*

$$D\chi : TM \to TM' : D\chi(p,v) := (\chi(p), D_p\chi(v)),$$

where $D_p\chi$ was defined in Equation (16).

Definition 9.11 *A* C^∞ vector field *is a* C^∞ *map* $v : M \to TM$ *such that* $v(p) \in T_p(M)$. *The space of all* C^∞ *vector fields is denoted by* $\Gamma^\infty(M, TM)$.

Because the atlas of TM is closely related to that of M, the smoothness condition can be formulated as follows: for any chart $\psi : U \to V$ in the atlas of M and the corresponding coordinate bases X_μ of T_pM, $p \in U$, the coefficient functions v^μ appearing in

$$v(p) = v^\mu(p) X_\mu,$$

must be C^∞. This local condition is independent of the choice of chart, so it gives rise to a global condition on v. The simplest example of vector fields, at least on the domain U of a chart ψ, are the coordinate basis vector fields X_μ.

For any $p \in M$, T_pM is a vector space, so we may apply the constructions of Section 9.1. We denote the dual space by T_p^*M and for the space of tensors of type (k, l) we introduce the notation

$$T_p^{(k,l)}M := \underbrace{T_pM \otimes \ldots \otimes T_pM}_{k \text{ times}} \otimes \underbrace{T_p^*M \otimes \ldots \otimes T_p^*M}_{l \text{ times}}.$$

We will now show that these linear spaces may be pasted together to form new manifolds, just like the tangent bundle.

Let us first consider the space T_p^*M, which is called the *cotangent space* of M at $p \in M$. The dimension of T_p^*M equals that of T_pM, which is $n = \dim(M)$. We may obtain interesting examples of elements in T_p^*M by choosing a function $f \in C^\infty(M, \mathbb{R})$ and defining $\mathrm{d}_p f \in T_p^*M$ by a funny reversal of perspective in the definition of tangent vectors:

$$\mathrm{d}_p f : T_pM \to \mathbb{R} : v \mapsto v(f). \tag{17}$$

$\mathrm{d}_p f$ is called the *differential* of f at $p \in M$.

Using a chart $\psi : U \to V$ with $p \in U$ we may apply this procedure to the coordinate functions $x^\mu \circ \psi$, which leads to a basis

$$X^{*\mu} := \mathrm{d}\,(x^\mu \circ \psi).$$

At each $p \in U$, this basis is dual to X_μ:

$$X^{*\mu}(X_\nu) = X_\nu(x^\mu \circ \psi) = \partial_{x^\nu}(x^\mu) = \delta^\mu{}_\nu.$$

We therefore call it the *dual coordinate basis* determined by ψ.

Recall that a change of chart at p induces a change of the basis X_μ, which can be written as a matrix multiplication. The change in the dual basis is then given by a multiplication with the inverse matrix, as in the tensor transformation law (13).

Definition 9.12 *The* cotangent bundle T^*M *of a manifold* M *of dimension* n *is the set*

$$T^*M := \bigcup_{p\in M} \{p\} \times T_p^*M = \{(p,\omega)|\ p \in M,\ \omega \in T_P^*M\},$$

where the union is disjoint (so $(p,\omega) = (p',\omega')$ *if and only if* $p = p'$ *and* $\omega = \omega'$*). We view* T^*M *as a manifold of dimension* $2n$ *with a maximal atlas containing all charts of the form*

$$D^*\psi : T^*U \to V \times \mathbb{R}^n : (p,\omega) := (\psi(p), D_p^*\psi(\omega))$$

such that $(D_p^*\psi(\omega))_\mu = \omega(X_\mu)$.

A C^∞ cotangent vector field *(or* dual vector field *or* 1-form field*) is a* C^∞ *map* $\omega : M \to T^*M$ *such that* $v(p) \in T_p^*(M)$*. The space of all* C^∞ *dual vector fields is denoted by* $\Gamma^\infty(M, T^*M)$.

Note that $\omega(X_\mu)$ are the components of ω in the dual coordinate basis, $\omega = \omega(X_\mu)X^{*\mu}$, as may be checked by letting both sides act on the coordinate basis X_ν. The dual coordinate basis $X^{*\mu}$ defines C^∞ cotangent vector fields on the domain U of the given chart and a general C^∞ cotangent vector field can be expressed as

$$\omega(p) = \omega_\mu(p) X^{*\mu}$$

with C^∞ coefficients ω_μ.

General tensors may be treated in an analogous way:

Definition 9.13 *The* tensor bundle $T^{(k,l)}M$ of type (k,l) *on a manifold M of dimension n is the set*

$$T^{(k,l)}M := \bigcup_{p\in M} T_p^{(k,l)}M,$$

where the union is disjoint. We view $T^{(k,l)}M$ as a manifold of dimension n^{k+l+1} with a maximal atlas containing all charts of the form

$$D^{(k,l)}\psi : T^{(k,l)}U \to V \times \mathbb{R}^{n^{k+l}} : (p,S) := (\psi(p), (D_p^{(k,l)}\psi)S)$$

such that the components of $((D_p^{(k,l)}\psi)S)^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l} = S(X^{*\mu_1},\ldots,X^{*\mu_k},X_{\nu_1},\ldots,X_{\nu_l})$.

A C^∞ tensor field of type (k,l) *is a C^∞ map $S : M \to T^{(k,l)}M$ such that $S(p) \in T_p^{(k,l)}(M)$. The space of all C^∞ tensor fields of type (k,l) is denoted by $\Gamma^\infty(M, T^{(k,l)}M)$.*

Again the smoothness of a tensor field S means that the components of S in the basis obtained from X_μ and $X^{*,\nu}$ are smooth functions on M.

A chart $\psi : U \to V$ determines a coordinate basis at each $p \in U$ for the tensor bundle. This basis consists of tensor products of X_μ and their duals $X^{*,\mu}$. We may express the components of a tensor field T in terms of this basis, but we may also use the abstract index notation. Recall that a formula in the abstract index notation is valid when the abstract symbols are replaced by the components of the tensor in any coordinate basis (cf. Convention 9.2).

The outer product and contraction of tensors can also be defined for tensor fields in a point-wise fashion. Given a tensor field S of type (k,l) and a tensor field T of type (k',l'), the outer product is a tensor field of type $(k+k', l+l')$, given by

$$(S\otimes T)^{a_1\cdots a_{k+k'}}{}_{b_1\cdots b_{l+l'}} := S^{a_1\cdots a_k}{}_{b_1\cdots b_l} T^{a_{k+1}\cdots a_{k+k'}}{}_{b_{l+1}\cdots b_{l+l'}},$$

whereas the contraction over the ith upper index and the jth lower index is given by

$$(CT)^{a_1\cdots\widehat{a_i}\cdots a_k}{}_{b_1\cdots\widehat{b_j}\cdots b_l} := T^{a_1\cdots c\cdots a_k}{}_{b_1\cdots c\cdots b_l}.$$

These equations hold point-wise as identities between linear maps. In addition, given any $p \in M$ we may choose a chart $\psi : U \to V$ with $p \in U$ and the equations above (in abstract index notation) imply corresponding equalities for the components in the coordinate basis of ψ. These equations are independent of the choice of the chart ψ, but they only make sense for $p \in U$ and not for the entire manifold M. This is an important reason to use abstract indices: because the equations are independent of the choice of chart and basis X_μ, they make sense on the entire manifold.

Example 9.14 *For a vector field v^a and a dual vector field ω_b we can construct the $(1,1)$-tensor field $(v \otimes \omega)^a{}_b = v^a \omega_b$ and its contraction $(Cv \otimes \omega) = v^a \omega_a$. (In this case there is only one choice of indices that we can contract.) The result is a C^∞ function on M, which at each point $p \in M$ equals the value of $\omega(p)$ when acting on $v(p)$.*

Example 9.15 *For any tensor T of type $(0,l)$ we can define a fully anti-symmetric tensor aT by*

$$(aT)_{b_1\cdots b_l} := \frac{1}{l!} \sum_{\pi \in S_l} \varepsilon(\pi) T_{b_{\pi(1)}\cdots b_{\pi(l)}},$$

where S_l is the group of all permutations π of the numbers $(1,\ldots,l)$ and $\varepsilon(\pi) = \pm 1$ according to whether the permutation is odd or even. Such fully anti-symmetric tensors are called differential forms *in the mathematical literature. Note that aT is indeed anti-symmetric in all its indices and that $a(aT) = aT$. As a matter of notation we will write*

$$T_{[b_1\cdots b_l]} := (aT)_{b_1\cdots b_l},$$

where the square brackets [ ] indicate that the indices should be anti-symmetrised over.

In a similar way we can also define a fully symmetric tensor sT by

$$(sT)_{b_1\cdots b_l} := \frac{1}{l!} \sum_{\pi \in S_l} T_{b_{\pi(1)}\cdots b_{\pi(l)}},$$

which we will write as

$$T_{(b_1 \cdots b_l)} := (sT)_{b_1 \cdots b_l}.$$

where the round brackets ( ) indicate that the indices should be symmetrised over.

With a slight abuse of notation the basis vectors X_μ and their duals $X^{*\mu}$ are often written as

$$\frac{\partial}{\partial x^\mu} := X_\mu, \qquad \mathrm{d}\, x^\mu := X^{*\mu}.$$

We have avoided this notation so far, because it suppresses the choice of the ψ and treats the tangent vectors $\frac{\partial}{\partial x^\mu}$ on V as if they were tangent vectors on U. (The purpose of the charts ψ is of course to translate the usual differential calculus on V into a differential calculus on U, thereby justifying that the abusive notation is indeed valid.)

Covariant Derivatives

We now turn our attention to derivatives of tensor fields on a manifold M. We have already seen that any C^∞ function $f : M \to \mathbb{R}$ gives rise to a one-form $\mathrm{d}\, f$ defined by $\mathrm{d}\, f(v) := v(f)$. In a chart ψ the one-form locally has the expression $(\mathrm{d}\, f)_\mu = \partial_{x^\mu} f$. Under a change of chart, this expression is multiplied by a Jacobi matrix, as usual for a dual vector. For higher order derivatives, however, an expression like $\partial_{x^\mu}\partial_{x^\nu} f$ can be used to define a tensor, but the resulting tensor depends on the choice of coordinates used. In other words, the formula $\partial_{x^\mu}\partial_{x^\nu} f$ is not preserved under changes of coordinates, because we also obtain terms involving derivatives of the Jacobi matrix. We now discuss how to eliminate, or at least keep track of, this dependence on the chart, using the notion of (covariant) derivative operators.

Definition 9.16 *A* (covariant) derivative operator ∇ *on a manifold M is a map (or rather a set of maps, for each (k,l))*

$$\nabla : \Gamma^\infty(M, T^{(k,l)}M) \to \Gamma^\infty(M, T^{(k,l+1)}M) : T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} \to \nabla_{b_0} T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l}$$

with the following properties:

1. linearity*: for all $c_i \in \mathbb{R}$ and tensor fields T_i of type (k,l), $\nabla(c_1 T_1 + c_2 T_2) = c_1 \nabla T_1 + c_2 \nabla T_2$, i.e.*

   $$\nabla_{b_0}(c_1 T_1 + c_2 T_2)^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} = c_1 \nabla_{b_0}(T_1)^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} + c_2 \nabla_{b_0}(T_2)^{a_1 \cdots a_k}{}_{b_1 \cdots b_l},$$

2. Leibniz rule: *for all tensor fields T and S, $\nabla(T\otimes S) = (\nabla T)\otimes S + T\otimes(\nabla S)$, i.e.*

$$\begin{aligned}\nabla_{b_0}(T^{a_1\cdots a_k}{}_{b_1\cdots b_l}S^{a_{k+1}\cdots a_{k+k'}}{}_{b_{l+1}\cdots b_{l+l'}}) &= \nabla_{b_0}T^{a_1\cdots a_k}{}_{b_1\cdots b_l}S^{a_{k+1}\cdots a_{k+k'}}{}_{b_{l+1}\cdots b_{l+l'}} + \\ &\quad T^{a_1\cdots a_k}{}_{b_1\cdots b_l}\nabla_{b_0}S^{a_{k+1}\cdots a_{k+k'}}{}_{b_{l+1}\cdots b_{l+l'}}\end{aligned}$$

   if T is of type (k,l) and T' of type (k',l').

3. commutativity with contractions: *for any tensor field T of type (k,l) with $k,l\geq 1$, $(\nabla CT) = C'(\nabla T)$, where C contracts the ith upper and jth lower index and C' the ith upper and $j+1$st lower index:*

$$\nabla_{b_0}(T^{a_1\cdots c\cdots a_k}{}_{b_1\cdots c\cdots b_l}) = \nabla_{b_0}T^{a_1\cdots c\cdots a_k}{}_{b_1\cdots c\cdots b_l},$$

4. consistency with differentials: *for any C^∞ function f, $\nabla f = d\,f$, i.e.*

$$\nabla_{b_0}f = (d\,f)_{b_0}.$$

5. torsion free: *for any C^∞ function f, $\nabla\nabla f$ is a symmetric tensor, i.e.*

$$\nabla_{b_0}\nabla_{b_1}f = \nabla_{b_1}\nabla_{b_0}f.$$

All these properties are familiar from calculus in $\mathbb{R}^n$ (except perhaps the commutativity with contractions). In fact, using a chart $\psi : U \to V$ and expressing tensors in their coordinate basis, one may choose $\nabla_\mu := \partial_{x^\mu}$ to define a derivative operator on U. This derivative, however, depends on the choice of coordinates. That covariant derivatives exist globally will be seen later, when we prove the existence of a preferred derivative operator.

For a general derivative operator, the action at $p\in M$ on a general tensor field T depends only on the values of T in an infinitesimal neighbourhood of p. This follows form the Leibniz rule and $(\nabla T)(p) = (\nabla(fT))(p)$ when f is a C^∞ function with $f(p) = 1$ and $\mathrm{d}\,f(p) = 0$. Expanding T in a coordinate basis, we may use linearity and the Leibniz rule to express ∇T in terms of derivatives of the basis vector and dual vector fields. Because the derivative operator commutes with contractions and its value on functions is known, we also have

$$\omega_b\nabla_a v^b = \nabla_a(v^b\omega_b) - v^b\nabla_a\omega_b.$$

When ω_b ranges over a basis, this allows us to express the action on any vector field in terms of the action on dual vector fields and functions. Thus, the only freedom we have, is to modify the action on dual vector fields.

Theorem 9.17 *Given two derivative operators on M, ∇ and ∇', there is a unique tensor field C of type $(1,2)$ such that $C^c{}_{ab} = C^c{}_{ba}$ and*

$$\nabla'_a \omega_b - \nabla_a \omega_b = C^c{}_{ab}\omega_c \tag{18}$$

for all dual vector fields ω_a. Conversely, given any derivative operator ∇' and a tensor field $C^c{}_{ab}$ which is symmetric in its lower indices, Equation (18) defines a derivative operator ∇.

Proof: At any point $p \in M$, $\nabla_a \omega_b - \nabla'_a \omega_b$ is a tensor of type $(0,2)$ which depends linearly on ω_a. We will show that it actually only depends on the values $\omega_a(p)$ and not on any derivatives. For this purpose we fix any chart $\psi : U \to V$ with $p \in U$ and we define the coordinate functions $x^\mu \circ \psi$ on U, $\mu = 1, \ldots, n$, where the x^μ are Cartesian coordinates on $\mathbb{R}^n$. The local basis $X^{*\mu}$ can be expressed in an abstract index notation as

$$(X^{*\mu})_a = (\mathrm{d}\,(x^\mu \circ \psi))_a = \nabla_a (x^\mu \circ \psi) = \nabla'_a (x^\mu \circ \psi)$$

and we may write $(X_\mu)^c$ in a similar way. We then view the components ω_μ as functions on U and write $\omega_a = \omega_\mu (X^{*\mu})_a = \omega_\mu \nabla_a (x^\mu \circ \psi)$ and $\omega_\mu = \omega_c (X_\mu)^c$. Then we compute with Leibniz rule and linearity:

$$\begin{aligned}\nabla'_a \omega_b - \nabla_a \omega_b &= (\nabla'_a \omega_\mu - \nabla_a \omega_\mu)(X^{*\mu})_b + \omega_\mu (\nabla'_a \nabla'_b (x^\mu \circ \psi) - \nabla_a \nabla_b (x^\mu \circ \psi)) \\ &= \omega_c (X_\mu)^c (\nabla'_a \nabla'_b (x^\mu \circ \psi) - \nabla_a \nabla_b (x^\mu \circ \psi)).\end{aligned}$$

Note that the first term on the right-hand side of the first line vanishes, because all derivative operators have the same action on the functions ω_μ. We have now shown that $(\nabla'_a \omega_b - \nabla_a \omega_b)(p)$ is a linear map of $\omega_a(p)$ without any derivatives. This entails the existence of the tensor field $C^c{}_{ab}$. We even obtain a formula for this tensor field,

$$C^c{}_{ab} := (X_\mu)^c (\nabla'_a \nabla'_b (x^\mu \circ \psi) - \nabla_a \nabla_b (x^\mu \circ \psi)),$$

(which is actually independent of the choice of chart).

Conversely, given any derivative operator ∇' and a tensor field $C^c{}_{ab}$ which is symmetric in its lower indices, one may define

$$\begin{aligned}\nabla_{b_0} T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} \;&:=\; \nabla'_{b_0} T^{a_1 \cdots a_k}{}_{b_1 \cdots b_l} + \sum_{i=1}^{k} C^{a_i}{}_{b_0 c} T^{a_1 \cdots c \cdots a_k}{}_{b_1 \cdots b_l} \\ &\quad - \sum_{j=1}^{l} C^c{}_{b_0 b_j} T^{a_1 \cdots a_k}{}_{b_1 \cdots c \cdots b_l}.\end{aligned}$$

We will omit the straightforward verification that ∇ is indeed a derivative operator. ☐

In a local chart we may choose ∇' to be a coordinate derivative. Although this derivative is not covariant, an analog of Theorem 9.17 still works, but now the components $C^{\rho}{}_{\mu\nu}$ do not transform like a tensor. In the chosen local chart, these components are written as $\Gamma^{\rho}{}_{\mu\nu}$ and they are called the *Christoffel symbol*.

Given a covariant derivative operator, it is natural to consider tensor fields whose covariant derivative vanishes. A very useful way of doing this is as follows. Let $\xi : I \to M$ be C^{∞} curve with $0 \in I$ and $p := \xi(0)$ and let $v^a(p) \in T_pM$. Given a covariant derivative operator, we may define unique vectors $v^a(q)$ at all points q on ξ such that

$$\dot{\xi}^a \nabla_a v^b = 0$$

on the curve. To see why this is true, we choose a chart and write the equation as

$$0 = \dot{\xi}^{\mu}(\partial_{\mu} v^{\nu} + \Gamma^{\nu}{}_{\mu\rho} v^{\rho}) = \partial_s (v^{\nu} \circ \xi) + \dot{\xi}^{\mu} \Gamma^{\nu}{}_{\mu\rho} v^{\rho}.$$

This is a first order differential equation for the components $v^{\nu} \circ \xi$ as a function of the parameter s, so given the initial values, it admits a unique solution. The vectors $v^a(q)$ are called the *parallel transports* of $v^a(p)$. Of course a similar result holds for any tensor. (The general result may be obtained by linearity and taking tensor products and duals.)

Given a curve $\xi : I \to M$ that goes through $p, p' \in M$, we can use parallel transports to define a map $C_{\xi;p,p'} : T_pM \to T_{p'}M$. Notice that this map is linear, because the parallel transport equation is linear. Moreover, it is invertible, because we can traverse the curve ξ in the opposite direction to find that $C_{\xi;p',p}$ is the inverse. This way of identifying tangent spaces at different points is called a *connection* in the mathematical literature.

As a warning we emphasise that the linear isomorphism $C_{\xi;p,p'} : T_pM \to T_{p'}M$ depends on the choice of the curve ξ.

The fact that connections depend on the choice of a curve ξ is unfamiliar from calculus in $\mathbb{R}^n$. Another marked difference between covariant derivatives and ordinary derivatives is the fact that mixed derivatives do not always commute. We do have $\nabla_a \nabla_b f = \nabla_b \nabla_a f$ for functions f, but for tensors this may no longer be true. This leads to the concept of curvature.

Theorem 9.18 *Given a derivative operator ∇ on a manifold M, there is a tensor field $R_{abc}{}^d$ of type $(1,3)$ such that*

$$\nabla_a\nabla_b\omega_c - \nabla_b\nabla_a\omega_c = R_{abc}{}^d\omega_d$$

for all one-form fields ω_c.

Proof: The left-hand side is a tensor which depends linearly on ω_c. The main point is to show that the value at any $p \in M$ depends only on the value $\omega_c(p)$ and not on any derivatives. For this purpose one may simply work in local coordinates and express ∇ in terms of the coordinate derivatives and a Christoffel symbol. A straightforward computation, which we omit, shows that no derivatives of ω at p appear and

$$R_{\mu\nu\rho}{}^\sigma = -\partial_\mu\Gamma^\sigma{}_{\nu\rho} + \Gamma^\tau{}_{\mu\rho}\Gamma^\sigma{}_{\nu\tau} - (\mu \leftrightarrow \nu).$$

□

Using linearity and the Leibniz rule one may show that for general tensors,

$$\begin{aligned}(\nabla_a\nabla_b - \nabla_b\nabla_a)T^{c_1\cdots c_k}{}_{d_1\cdots d_l} &= -\sum_{i=1}^{k} R_{abe}{}^{c_i}T^{c_1\cdots e\cdots c_k}{}_{d_1\cdots d_l} \\ &\quad + \sum_{j=1}^{l} R_{abd_j}{}^{e}T^{c_1\cdots c_k}{}_{d_1\cdots e\cdots d_l}.\end{aligned}$$

Definition 9.19 *The tensor $R_{abc}{}^d$ is called the* Riemann curvature tensor *of ∇. The tensor $R_{adc}{}^d$ of type $(0,2)$ is called the* Ricci curvature tensor.

Theorem 9.20 *The Riemann curvature tensor has the following properties:*

1. $R_{abc}{}^d = -R_{bac}{}^d$,
2. $R_{[abc]}{}^d = 0$,
3. $\nabla_{[a}R_{bc]d}{}^e = 0$ *(Bianchi identity).*

Proof: The first property of $R_{abc}{}^d$ is obvious. For the second we note that $\nabla_{[\mu}\omega_{\nu]} = \partial_{[\mu}\omega_{\nu]}$, because the Christoffel symbol is symmetric in its lower indices. Similarly,

$$\nabla_{[\mu}\nabla_\nu\omega_{\rho]} = \partial_{[\mu}\partial_\nu\omega_{\rho]} = 0$$

by Equation (19) and the fact that ordinary derivatives commute. Thus, $R_{[\mu\nu\rho]}{}^{\sigma}\omega_\sigma = 2\nabla_{[\mu}\nabla_\nu\omega_{\rho]} = 0$. For the third property we consider

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\nabla_c\omega_d = R_{abc}{}^{e}\nabla_e\omega_d + R_{abd}{}^{e}\nabla_c\omega_e$$

and

$$\nabla_a(\nabla_b\nabla_c - \nabla_c\nabla_b)\omega_d = (\nabla_a R_{bcd}{}^{e})\omega_e + R_{bcd}{}^{e}\nabla_a\omega_e.$$

Antisymmetrising over the indices a, b, c, the left-hand sides become equal. Using the properties of the Riemann curvature that we have already shown on the right-hand side leads to

$$(\nabla_{[a}R_{bc]d}{}^{e})\omega_e = R_{[ab|d|}{}^{e}\nabla_{c]}\omega_e - R_{[bc|d|}{}^{e}\nabla_{a]}\omega_e = 0.$$

As ω_e is arbitrary, the Bianchi identity follows. □

Metrics and Geodesics

So far we do not have sufficient structure on our manifolds to formulate physically relevant questions: What is the distance between two points? What is the shortest path between two points? And is there a preferred choice of covariant derivative?

Let us start with the definition of a (pseudo-) metric:

Definition 9.21 *A (pseudo-)*metric *on a manifold M is a smooth tensor field g_{ab} of type $(0,2)$ such that*

1. *$g_{ab} = g_{ba}$ (symmetry),*
2. *if $v^a \in T_pM$ has $g_{ab}(p)v^a w^b = 0$ for all $w^b \in T_pM$, then $v^a = 0$ (non-degeneracy).*

A pair (M, g_{ab}) consisting of a manifold with a (pseudo-)metric is called a pseudo-Riemannian manifold.

For any $v^a \in TM$, $g_{ab}v^a v^b$ can be viewed as the squared norm of the vector v^a. However, g_{ab} need not be positive definite, i.e. it need not be true that $g_{ab}v^a v^b \geq 0$ for all $v^a \in TM$.

At any point $p \in M$ we can consider the matrix $g_{\mu\nu}(p)$ in a coordinate basis of some chart. This matrix depends on the choice of coordinates, but it has a number of properties which are independent of this choice. It is always symmetric, so it can be diagonalised and the numbers n_+ of positive eigenvalues and $n_- = n - n_+$ of negative eigenvalues are also coordinate independent. In fact, for a given point $p \in M$ we can always choose coordinates such that $g_{\mu\nu}(p)$ is diagonal with eigenvalues $+1$ and -1 only. However, this may hold only hold at the point p! Finally we note that n_+ and n_- do not depend on the choice of p, because the components $g_{\mu\nu}(p)$ vary smoothly with p and $g_{\mu\nu}$ must remain non-degenerate.

Definition 9.22 *The* signature *of a metric is the pair (n_+, n_-), where n_+ is the number of positive eigenvalues of the matrix $g_{\mu\nu}(p)$ (in a coordinate basis at any point $p \in M$), and n_- is the number of negative eigenvalues.*

We call a metric Riemannian *when its signature is $(n, 0)$. In that case we call the pair (M, g_{ab}) a* Riemannian manifold.

We call a metric Lorentzian *when its signature is $(n-1, 1)$ and $n \geq 2$. In that case we call the pair (M, g_{ab}) a* Lorentzian manifold, *or a* spacetime.

For a Riemannian metric, $g_{\mu\nu}$ is positive definite. For a Lorentzian metric, $g_{\mu\nu}$ is similar to the matrix η of Equation (7).

For a Riemannian manifold (M, g_{ab}) and a curve $\xi : (a, b) \to M$ we can define the length of ξ by

$$\int_a^b \sqrt{g_{\mu\nu}(\xi(s))\dot{\xi}^\mu(s)\dot{\xi}^\nu(s)}\,\mathrm{d}\,s,$$

i.e. we integrate the length of the tangent vectors along the curve. Note that this expression is independent of the choice of parameter.

For Lorentzian manifolds (M, g_{ab}) we say that a curve $\gamma : (s_1, s_2) \to M$ is *spacelike* when $g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s) > 0$ for all s. The *length* of γ is then given by

$$l_\gamma := \int_{s_1}^{s_2} \sqrt{g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s)}\mathrm{d}\,s.$$

Similarly, γ is called *timelike* when $g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s) < 0$ for all s. The *proper time* of γ is then given by

$$\tau_\gamma := \int_{s_1}^{s_2} \sqrt{-g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s)}\mathrm{d}\,s.$$

For curves which are neither timelike nor spacelike we do not define a length or proper time.

Because g_{ab} is non-degenerate, we can use it to define a bijection

$$TM \to T^*M : (p, v^b) \mapsto (p, g_{ab}(p)v^b),$$

which is a linear isomorphism from T_pM to T_p^*M for each $p \in M$. The inverse of this map is a tensor of type $(2, 0)$ and will be denoted by g^{ab}, with upper indices, so that

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

is the identity map on TM. As a matter of notation we will define

$$v_a := g_{ab}(p)v^b, \qquad \omega^a := g^{ab}(p)\omega_b$$

for any $(p, v^b) \in TM$ and $(p, \omega_b) \in T^*M$. In this way we can use the metric to raise and lower indices, i.e. to identify vectors and dual vectors. The notation can be extended in an obvious way to all tensors. As a warning, however, we note that in General Relativity, the Lorentzian metric g_{ab} is a dynamical object, so the identification between vectors and dual vectors is not given a priori. Instead, we must solve Einstein's equations to find it.

Exercise 9.23 *Use the inverse metric g^{ab} to show that $g_{ab}g^{ac}g^{bd} = g^{cd}$, so the notation for raising and lowering indices can also be applied in a consistent way to the metric itself.*

Suppose that $\gamma : I \to M$ is a curve and v^a and w^b are vector fields along the curve which are parallel transports for some covariant derivative operator ∇. The inner product $g_{ab}v^aw^b$ along γ changes as follows:

$$\dot{\gamma}^a\nabla_a(g_{bc}v^bw^c) = \dot{\gamma}^a v^b w^c \nabla_a g_{bc}.$$

This means that the inner product remains constant if $\nabla_a g_{bc} = 0$. We now show that for a given (pseudo-)metric g_{ab} there exists a unique covariant derivative operator which satisfies this additional condition:

Theorem 9.24 *Let g_{ab} be any (pseudo-) metric on a manifold M. Then there exists a unique covariant derivative operator ∇ on M which is compatible with g_{ab} in the sense that $\nabla_a g_{bc} = 0$.*

Proof: In any chart $\psi : U \to V$ we can use the coordinate derivative ∂_μ on U. By Theorem 9.17 we can characterise ∇ on U by its Christoffel symbols $\Gamma^c{}_{ab} = C^c{}_{ab}$. We have

$$0 = \nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\sigma{}_{\mu\nu} g_{\sigma\rho} - \Gamma^\sigma{}_{\mu\rho} g_{\nu\sigma}$$

and therefore,

$$\begin{aligned}\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu} &= (\Gamma_{\rho\mu\nu} + \Gamma_{\nu\mu\rho}) + (\Gamma_{\rho\nu\mu} + \Gamma_{\mu\nu\rho}) - (\Gamma_{\nu\rho\mu} + \Gamma_{\mu\rho\nu}) \\ &= 2\Gamma_{\rho\mu\nu},\end{aligned}$$

where we used the symmetry of the Christoffel symbol. This shows that any derivative operator which is compatible with g_{ab} must have a Christoffel symbol which is given in any chart by

$$\Gamma^\sigma{}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}),$$

which proves uniqueness of ∇ on U. Conversely, choosing this Christoffel symbol to define ∇, we also see that ∇ exists on U. We can do this for any chart and on the intersection of any two chart domains, the resulting covariant derivatives must coincide, because both satisfy the compatibility condition. □

The unique covariant derivative operator which is compatible with g_{ab} is called the *Levi-Civita* covariant derivative (or *Levi-Civita connection*). The Levi-Civita derivative operator gives rise to a connection $C_{\xi;p,p'}$ which is not just a linear isomorphism $T_pM \to T_{p'}M$, but it also preserves inner products, i.e. it maps $g_{ab}(p)$ to $g_{ab}(p')$.

We now consider the curvature of the Levi-Civita derivative operator, starting with some additional notations:

Definition 9.25 *The function* $R := g^{ac}R_{ac} = g^{ac}R_{abc}{}^b$ *is called the (Ricci)* scalar curvature *of the Levi-Civita derivative. The tensor* $G_{ab} := R_{ab} - \frac{1}{2}Rg_{ab}$ *is called its* Einstein tensor.

Theorem 9.26 *The curvature of the Levi-Civita derivative operator satisfies*

$$\begin{aligned}R_{abcd} &= -R_{abdc}, \\ G_{ab} &= G_{ba}, \\ \nabla^a G_{ab} &= 0.\end{aligned}$$

Proof: The first property follows directly from $0 = \nabla_{[a}\nabla_{b]}g_{cd}$ and Equation (19). The second property follows from $R_{[abc]}{}^d = 0$, because

$$R_{ba} = R_{bca}{}^c = -R_{abc}{}^c - R_{cab}{}^c = -g^{cd}R_{abcd} + R_{acb}{}^c = R_{ab}.$$

(The first term vanishes by the first property and symmetry reasons.) For the third we consider the Bianchi identity $\nabla_{[a}R_{bc]d}{}^{e} = 0$ (cf. Theorem 9.20) and we use the symmetries of the Riemann curvature to compute:

$$\begin{aligned} 0 &= 3g^{bd}\nabla_{[a}R_{bc]d}{}^{c} \\ &= g^{bd}\nabla_a R_{bcd}{}^{c} + g^{bd}\nabla_b R_{cad}{}^{c} + g^{bd}\nabla_c R_{abd}{}^{c} \\ &= \nabla_a g^{bd}R_{bd} - \nabla^d R_{acd}{}^{c} - \nabla^c g^{bd} R_{abcd} \\ &= \nabla_a R - 2\nabla^d R_{ad} \\ &= -2\nabla^d(R_{ad} - \frac{1}{2}g_{ad}R) = -2\nabla^d G_{ad}. \end{aligned}$$

Because the Einstein tensor is symmetric, the result follows. □

Definition 9.27 *A* geodesic *in a (pseudo-)Riemannian manifold* (M, g_{ab}) *is a curve* $\xi : I \to M$ *whose tangent vectors* $\dot{\xi}^a$ *are parallelly transported along* ξ *with respect to the Levi-Civita covariant derivative operator.*

(Actually, the definition can be made for arbitrary covariant derivative operators on M, but the Levi-Civita derivative operator will be our main interest.)

We think of geodesics as curves which are as straight as possible. The condition on the tangent vectors can be written as the geodesic equation:

$$\dot{\xi}^a \nabla_a \dot{\xi}^b = 0, \tag{19}$$

or, in a chart and using the parameter s along ξ:

$$\ddot{\xi}^\mu + \Gamma^\mu{}_{\nu\rho}(\xi)\dot{\xi}^\nu\dot{\xi}^\rho = 0.$$

This is a second order (non-linear) ordinary differential equation for the components $\xi^\mu(s)$, so its solution is uniquely determined (for s in some maximal interval) by initial values for ξ^μ and $\dot{\xi}^\mu$. In other words, there exists a unique maximal geodesic through any point $(p, v^a) \in TM$.

Exercise 9.28 *Consider* $\mathbb{R}^n$ *as a Riemannian manifold with the constant metric field* $g_{\mu\nu} = \delta_{\mu\nu}$ *in a global Cartesian chart. Given any tangent vector,* $(p, v^\mu) \in T\mathbb{R}^n$*, show that the unique geodesic* ξ *with initial data* (p, v^μ) *is*

$$\xi^\mu(s) := p^\mu + sv^\mu.$$

Exercise 9.29 *Show that for any geodesic* $\xi : I \to M$, $\dot{\xi}^a \dot{\xi}_a$ *is constant along the curve* ξ.

For spacelike or timelike curves the geodesic equation (19) can be obtained from a variational principle. The length l_γ, resp. proper time τ_γ, of a spacelike, respectively timelike, curve γ between two fixed points attains a (local) extremum when γ is a geodesic. To verify this one computes the Euler-Lagrange equations of

$$\int \sqrt{|g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s)|}\,\mathrm{d}\, s$$

to be (in local coordinates (x^μ))

$$0 = \frac{-2}{\sqrt{|g_{\mu\nu}(\xi)\dot{\xi}^\mu\dot{\xi}^\nu|}} \left(-\frac{1}{2}(\partial_\rho g_{\mu\nu})(\xi)\dot{\xi}^\mu\dot{\xi}^\nu + \partial_s(g_{\mu\rho}(\xi)\dot{\xi}^\mu) \right)$$

and the term in brackets can be rewritten, using $\partial_s(g_{\mu\rho}(\xi) = (\partial_\nu g_{\mu\rho})\dot{\xi}^\nu$ and relabeling indices, as

$$g_{\mu\rho}\left(\ddot{\xi}^\mu + \Gamma^\mu{}_{\sigma\nu}\dot{\xi}^\sigma\dot{\xi}^\nu\right).$$

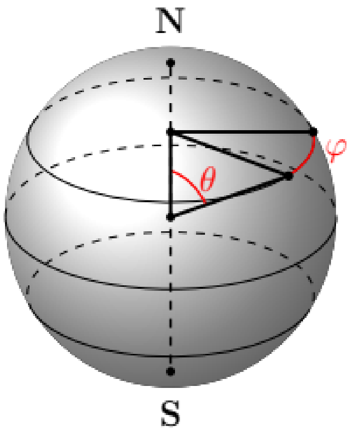
Example 9.30 (Riemann-tensor and geodesics on $\mathbb{S}^2$) *The metric of a 'round' 2-sphere* $\mathbb{S}^2$ *with radius* r *is by definition*

$$ds^2 = r^2\left(d\theta^2 + \sin^2\theta\; d\varphi^2\right)$$

with $\theta = x_1$ *and* $\varphi = x_2$*, or in matrix form*

$$(g_{\mu\nu}) = \begin{pmatrix} r^2 & 0 \\ 0 & r^2\sin^2\theta \end{pmatrix},$$

$$(g^{\mu\nu}) = \begin{pmatrix} \frac{1}{r^2} & 0 \\ 0 & \frac{1}{r^2\sin^2\theta} \end{pmatrix}.$$



One can check using the definition of $\Gamma^\sigma{}_{\mu\nu}$ *that*

$$\begin{aligned} \Gamma^1{}_{22} &= -\sin\theta\cos\theta \\ \Gamma^2{}_{12} = \Gamma^2{}_{21} &= \cot\theta \end{aligned}$$

and that all other components vanish.

$$\begin{aligned} R_{212}{}^1 &= \partial_1\Gamma^1{}_{22} - \partial_2\Gamma^1{}_{12} + \Gamma^1{}_{1\alpha}\Gamma^\alpha{}_{22} - \Gamma^1{}_{2\alpha}\Gamma^\alpha{}_{12} \quad (\textit{sum over } \alpha = 1,2!) \\ &= \left(\sin^2\theta - \cos^2\theta\right) - (0) + (0) - (-\sin\theta\cos\theta)(\cot\theta) \\ &= \sin^2\theta. \end{aligned}$$

To get $R_{\mu\nu\sigma\rho}$ *with lower indices, we use* $R_{\mu\nu\sigma\rho} = g_{\rho\alpha}R_{\mu\nu\sigma}{}^\alpha$. *Then we have*

$$\begin{aligned} \Rightarrow \quad R_{2121} &= g_{1\alpha}R_{212}{}^\alpha \\ &= g_{11}R_{212}{}^1 + \cancel{g_{12}}^{=0}R_{212}{}^2 \\ &= r^2\sin^2\theta. \end{aligned}$$

Using the symmetries

$$R_{11\alpha\beta} = R_{22\alpha\beta} = R_{\alpha\beta11} = R_{\alpha\beta22} = 0$$

we find $R_{2112} = -R_{2121}, \ldots$ *and so on for all 16 components of* $R_{\alpha\beta\gamma\delta}$.

$$\begin{aligned} R_{11} &= g^{\alpha\beta}R_{1\alpha1\beta} \\ &= g^{11}\cancel{R_{1111}}^{=0} + g^{22}R_{1212} + 2g^{12}\cancel{R_{1121}}^{=0} \\ &= 1 \\ R_{22} &= g^{\alpha\beta}R_{2\alpha2\beta} \\ &= g^{11}R_{2121} + g^{22}\cancel{R_{2222}}^{=0} + 2g^{12}\cancel{R_{2122}}^{=0} \\ &= \sin^2\theta \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad R &= g^{11}R_{11} + \cancel{2g^{12}}^{=0}R_{12} + g^{22}R_{22} \\ &= \frac{2}{r^2}. \end{aligned}$$

It shows, that $\mathbb{S}^2$ *has a constant scalar curvature.*

To find the geodesics, let $(\gamma^\mu(t)) = (x_1(t), x_2(t)) = (\theta(t), \varphi(t))$. *Then we have the geodesic equation*

$$\frac{d^2}{dt^2}\gamma^\mu(t) + \Gamma^\mu{}_{\alpha\beta}(\gamma(t))\frac{d}{dt}\gamma^\alpha(t)\frac{d}{dt}\gamma^\beta(t) = 0$$

Written out using $\frac{d}{dt} = \dot{}$ *and* $\mu = 1$ *or* $\mu = 2$*:*

$$\ddot{\theta} - \sin\theta\cos\theta\,(\dot{\varphi})^2 = 0 \quad \textit{and} \quad \ddot{\varphi} + 2\cot\theta\;\dot{\varphi}\dot{\theta} = 0.$$

One solution is obtained for $\varphi = \varphi_0 = const.$ *Then we have:*

$$\dot{\varphi} = \ddot{\varphi} = 0$$

and we find

$$\ddot{\theta} = 0 \quad \textit{which implies} \quad \theta = \theta_0 + \nu t.$$

This describes a segment of a longitudinal great circle of the sphere. By symmetry, one can see that all geodesics are great circles.

Example 9.31 *Another interesting metric is that of* $\mathbb{H}^2$*, the hyperbolic space* $(x_1 = x, x_2 = y > 0)$ *with line element*

$$ds^2 = \frac{dx^2 + dy^2}{y^2},$$

or in matrix form

$$\begin{aligned}(g_{\mu\nu}) &= \begin{pmatrix} \frac{1}{y^2} & 0 \\ 0 & \frac{1}{y^2} \end{pmatrix} \\ (g^{\mu\nu}) &= \begin{pmatrix} y^2 & 0 \\ 0 & y^2 \end{pmatrix} \qquad (20)\end{aligned}$$

One way to obtain geodesics γ *is to go back to the length functional* l_γ *and to remember that geodesics correspond to critical points of this functional. For practical computations, let us parameterize the curve by its* x*-coordinate as*

$$(\gamma^\mu(x)) = (x, y(x)) \;\; \textit{which leads to} \;\; (\dot{\gamma}^\mu(x)) = (1, y'(x))\,.$$

(Note that not all curves in $\mathbb{H}^2$ *can be parameterized that way, but it turns out that all geodesics which are not vertical lines paralell to the* y*-axis can.) Then we have*

$$l_\gamma = \int dx\; \frac{\sqrt{1 + (y')^2}}{y} \qquad \textit{where } y = y(x)$$

and we get the Euler-Lagrange equations

$$\left(\frac{y'}{y\sqrt{1 + y'^2}}\right)' = -\frac{\sqrt{1 + y'^2}}{y^2}.$$

It can be seen that the solutions are arcs of circles in the x*-*y*-plane with the centre on the* x*-axis, or vertical lines (exercise).*