Lecture 8. Connections

This lecture introduces connections, which are the machinery required to allow differentiation of vector fields.

8.1 Differentiating vector fields.

The idea of differentiating vector fields in Euclidean space is familiar: If we choose a fixed orthonormal basis $\{e_1, \dots, e_n\}$ for $\mathbb{R}^n$, then any vector field X can be written in the form $\sum_{i=1}^n X^i e_i$ for some smooth functions $X^1, \dots, X^n$. Then the derivative of X in the direction of a vector v is given by

$$D_v X = \sum_{i=1}^n v(X^i) e_i.$$

In other words, we just differentiate the coefficient functions. The key point is that we think of the vectors $\{e_1, \dots, e_n\}$ as being *constant*.

When we are dealing with vector fields on a manifold, this no longer makes sense: There are no natural vector fields which we can take to be 'constant'.

Example 8.1.1 A failed attempt. One might think that a chart would supply natural vector fields that we can think of as being constant – every vector field can be written as a linear combination of the coordinate tangent vectors $\{\partial_1, \dots, \partial_n\}$. Unfortunately, defining derivatives in terms of these does not really make sense: The difficulty is that changing to a different chart would give a different answer for the derivative of a vector field.

To see this, suppose we have two charts, φ and η, and that a vector field X is given by $\sum_{i=1}^n X^i \partial_i^{(\varphi)}$, and a vector $v = \sum_{i=1}^n v^i \partial_i^{(\varphi)}$. Then we can define a derivative $\nabla_v^{(\varphi)} X$ by

$$\nabla_v^{(\varphi)} X = \sum_{i=1}^n \left(v^j \partial_j^{(\varphi)} \left(X^i\right) \right) \partial_i^{(\varphi)}.$$

Now work in the chart η: We have $\partial_i^{(\varphi)} = \sum_{j=1}^n \Lambda_i^j \partial_j^{(\eta)}$, where Λ is the matrix $D\left(\eta \circ \varphi^{-1}\right)$. Therefore we have

$$\nabla_v^{(\varphi)} X = \sum_{i=1}^{n} v^j \Lambda_j^k \partial_k^{(\eta)} \left(X^i\right) \Lambda_i^l \partial_l^{(\eta)}.$$

However, if we instead worked from the beginning in the chart η, we would have $X = (X^i \Lambda_i^j)\partial_j^{(\eta)}$ and $v = (v^i \Lambda_i^j)\partial_j^{(\eta)}$, and so

$$\nabla_v^{(\eta)} X = v^j \Lambda_j^k \partial_k^{(\eta)} \left(X^i \Lambda_i^l\right) \partial_l^{(\eta)} = \nabla_v^{(\varphi)} X + v^j X^i \Lambda_j^k \partial_k^{(\eta)} \left(\Lambda_i^l\right) \partial_l^{(\eta)}.$$

So these two notions of derivative don't agree if Λ is not constant (i.e. if the second derivatives of the transition map do not vanish).

8.2 Definition of a connection.

Since there does not seem to be any particularly natural way to define derivatives of vector fields on a manifold, we have to introduce extra structure on the manifold to do this. The extra structure required is a *connection* on M.

Definition 8.2.1 Let M be a smooth manifold. A **connection** on M is a map $\nabla : TM \times \mathcal{X}(M) \to TM$ such that $\nabla_v X \in T_x M$ if $v \in T_x M$, and

(1). $\nabla_{c_1 v_1 + c_2 v_2} X = c_1 \nabla_{v_1} X + c_2 \nabla_{v_2} X$ for all v_1 and v_2 in $T_x M$ for any $x \in M$, any real numbers c_1 and c_2, and any $X \in \mathcal{X}(M)$;
(2). $\nabla_v \left(c_1 X_1 + c_2 X_2\right) = c_1 \nabla_v X_1 + c_2 \nabla_v X_2$ for every $v \in TM$, all real numbers c_1 and c_2, and all X_1 and X_2 in $\mathcal{X}(M)$;
(3). $\nabla_v (fX) = v(f)X + f\nabla_v X$ for all $v \in TM$, $f \in C^\infty(M)$, and $X \in \mathcal{X}(M)$.

It is simple to check that the differentiation of vector fields in Eudclidean space satisfies these requirements. Condition (3) is the Leibniz rule for differentiation of vector fields, which requires the differentiation to be consistent with the differentiation of smooth functions.

Example 8.2.2 A natural connection for submanifolds of Euclidean space. Let M be an n-dimensional submanifold of $\mathbb{R}^N$. We will use the inner product structure of $\mathbb{R}^N$ to induce a connection on M: Let $\{e_1, \ldots, e_N\}$ be an orthonormal basis for $\mathbb{R}^N$. Then at a point $x \in M$, $T_x M$ is an n-dimensional subspace of $T_x \mathbb{R}^N \simeq \mathbb{R}^N$. Denote by π the orthogonal projection onto this subspace.

Suppose $v \in T_x M$, and $X \in \mathcal{X}(M)$. Then we can write $X = \sum_{\alpha=1}^N X^\alpha e_\alpha$. In this situation we can regard the vector fields e_α as being constant, so it makes sense to define the derivative $D_v X = \sum_{\alpha=1}^N v(X^\alpha) e_\alpha$. This satisfies all the requirements for a connection, expect that the result may not be a vector in TM, only in $\mathbb{R}^N$. So, to remedy this we can take the orthogonal projection of the result onto TM:

$$\nabla_v X = \pi \left(\sum_{\alpha=1}^{N} v(X^\alpha) e_\alpha \right).$$

Let us check explicitly that this is a connection: The linearity (with constant coefficients) in each argument is immediate. The Leibniz rule also holds:

$$\begin{aligned}
\nabla_v(fX) &= \nabla_v \left(\sum_{\alpha=1}^{N} (fX^\alpha) e_\alpha \right) \\
&= \pi \left(\sum_{\alpha=1}^{N} v(fX^\alpha) e_\alpha \right) \\
&= \pi \left(\sum_{\alpha=1}^{N} \left(v(f) X^\alpha + f v(X^\alpha) \right) e_\alpha \right) \\
&= v(f) \pi \left(\sum_{\alpha=1}^{N} X^\alpha e_\alpha \right) + f \pi \left(\sum_{\alpha=1}^{N} v(X^\alpha) e_\alpha \right) \\
&= v(f) X + f \nabla_v X
\end{aligned}$$

since $\pi X = X$.

Remark. A particle moving subject to a constraint that it lies in the submanifold (with no other external forces) moves as it would in free space, except that any component of its acceleration in direction normal to the surface are automatically cancelled out by the constraint forces. In other words, the motion of the particle is determined by the equation

$$\pi \left(\frac{d^2 x}{dt^2} \right) = 0.$$

Example 8.2.3 The left-invariant connection on a Lie group Let G be a Lie group, and $E_1, \ldots, E_n$ a collection of left-invariant vector fields generated by some basis for $T_e G$. Then any vector field X can be written as a linear combination of these: $X = \sum_{i=1}^{n} X^i E_i$ for some smooth functions $X^1, \ldots, X^n$. We can define a connection on G, called the *left-invariant connection*, by setting the vector fields E_i to be constant:

$$\nabla_v X = \sum_{i=1}^{n} v(X^i) E_i.$$

This connection is independent of the choice of the basis vector fields $E_1, \ldots, E_n$: If $\gamma : I \to G$ is a path with $\gamma'(0) = v \in T_g G$, then the connection can be written in the form

$$\nabla_v X = D_e l_g \left(\frac{d}{dt} \left(D_e l_{\gamma(t)} \right)^{-1} X_{\gamma(t)} \right).$$

This makes sense because $(D_e l_{\gamma(t)})^{-1} X_{\gamma(t)} \in T_eG$ for all t, and we know how to differentiate vectors in a vector space.

Note that we could just as well have considered right-invariant vector fields to define the connection. This gives the *right-invariant connection* on G, which is generally different from the left-invariant connection if G is a non-commutative Lie group.

8.3 Connection coefficients.

Suppose we are working in a chart $\varphi : U \to V \subset \mathbb{R}^n$ for M, with corresponding coordinate tangent vectors $\partial_1, \ldots, \partial_n$. The *connection coefficients* of a connection ∇ with respect to this chart are defined by

$$\nabla_{\partial_i}\partial_j = {\Gamma_{ij}}^k \partial_k.$$

Proposition 8.3.1 *The connection is determined on U by the connection coefficiants.*

Proof. Write $X = X^i\partial_i$. Then we have for any vector $v = v^k\partial_k$

$$\nabla_v X = \nabla_v \left(X^i\partial_i\right) = v(X^i)\partial_i + X^i\nabla_v\partial_i = v(X^i)\partial_i + v^k X^i {\Gamma_{ki}}^j \partial_j.$$

$\square$

Exercise 8.3.2 Suppose we change to a different chart η. Write down the connection coefficients with respect to η, in terms of the connection coefficients for φ, and the map Λ_i^j defined by $D(\eta \circ \varphi^{-1})(e_i) = \Lambda_i^j e_j$.

8.4 Vector fields along curves

Let $\gamma : I \to M$ be a smooth curve. Then by a *vector field along* γ we will mean a smooth map $V : I \to TM$ with $V(t) \in T_{\gamma(t)}M$ for each t. The set of smooth vector fields alonmg γ will be denoted by $\mathcal{X}_\gamma(M)$.

Proposition 8.4.1 *Let M be a smooth manifold and ∇ a connection on M. Then for any smooth curve $\gamma : I \to M$ there is a natural covariant derivative along γ, denoted ∇_t, which takes a smooth vector field along γ to another. This map satisfies:*

(1). $\nabla_t(V + W) = \nabla_t V + \nabla_t W$ *for all V and W in $\mathcal{X}_\gamma(M)$;*
(2). $\nabla_t(fV) = \frac{df}{dt}V + f\nabla_t V$ *for all $f \in C^\infty(I)$ and $V \in \mathcal{X}_\gamma(M)$;*
(3). *If V is given by the restriction of a vector field $\tilde{V} \in \mathcal{X}(M)$ to γ, i.e. $V_t = \tilde{V}_{\gamma(t)}$, then $\nabla_t V = \nabla_{\dot\gamma}\tilde{V}$, where $\dot\gamma$ is the tangent vector to γ.*

The main point of this proposition is that the derivative of a vector field V in the direction of a vector v can be computed if one only knows the values of V along some curve with tangent vector v.

The covariant derivative along γ is defined by

$$\nabla_t \left(V^i(t)\partial_i\right) = \frac{dV^i}{dt}\partial_i + \dot\gamma^j V^i {\Gamma_{ji}}^k(\gamma(t))\partial_k$$

where $\dot\gamma = \dot\gamma^k\partial_k$. The first two statements of the proposition follow immediately, and it remains to check that our definition is consistent with the third.

Let $\tilde V$ be a smooth vector field in M. Then the induced vector field along a curve γ is given by $V(t) = \tilde V_{\gamma(t)}$, or in terms of the coordinate tangent basis, $V(t) = V^i(t)\partial_i$, where $\tilde V_x = \tilde V^i(x)\partial_i$, and $V^i(t) = \tilde V^i(\gamma(t))$.

Then

$$\begin{aligned}
\nabla_{\dot\gamma}\tilde V &= \nabla_{\dot\gamma}\left(\tilde V^i\partial_i\right) \\
&= \left(D\gamma(\partial_t)\right)(\tilde V^i)\partial_i + \dot\gamma^k\tilde V^i{\Gamma_{ki}}^l\partial_l \\
&= \partial_t\left(\tilde V^i\circ\gamma\right)\partial_i + \dot\gamma^k V^i{\Gamma_{ki}}^l(\gamma(t))\partial_i \\
&= \nabla_t V.
\end{aligned}$$

8.5 Parallel transport.

Definition 8.5.1 Let $\gamma : I \to M$ be a smooth curve. A vector field $V \in \mathcal{X}_\gamma(M)$ is said to be **parallel** along γ if $\nabla_t V = 0$.

For example, the constant vector fields in $\mathbb{R}^n$ are parallel with respect to the standard connection, and the left-invariant vector fields on a Lie group are parallel with respect to the left-invariant connection.

Proposition 8.5.2 *Let $\gamma : I \to M$ be a piecewise smooth curve. Then there exists a unique family of linear isomorphisms $P_t : T_{\gamma(0)}M \to T_{\gamma(t)}M$ such that a vector field V along γ is parallel if and only if $V(t) = P_t(V_0)$ for all t.*

The maps P_t are the *parallel transport* operators along γ.

To prove the proposition, consider the differential equation that defines when a vector field along γ is parallel, on an interval I where γ remains inside some chart with corresponding connection coefficients ${\Gamma_{kl}}^m$:

$$0 = \nabla_t V = \left(\frac{dV^i}{dt} + {\Gamma_{kl}}^i\dot\gamma^k V^l\right)\partial_i.$$

This means that for $i = 1, \ldots, n$,

$$0 = \frac{dV^i}{dt} + {\mathit{\Gamma}_{kl}}^i \dot{\gamma}^k V^l. \tag{8.5.1}$$

This is a linear system of n first order ordinary differential equations for the n functions $V^1 \ldots, V^n$ along I, and so (since the coefficient functions $\dot{\gamma}^k {\mathit{\Gamma}_{kl}}^i$ are bounded and piecewise continuous) there exist unique solutions with arbitrary initial values $V^1(0), \ldots, V^n(0)$. We define P_t by $P_t(V^i(0)\partial_i) = V^i(t)\partial_i$, where $V^i(t)$ satisfy the system (8.5.1). We leave it to the reader to check that this defines a linear isomorphism for each t.

Proposition 8.5.3 *The covariant derivative of a vector field $V \in \mathcal{X}_\gamma(M)$ along γ can be written as*

$$\nabla_t V = P_t \left(\frac{d}{dt} \left((P_t)^{-1} V(t) \right) \right).$$

Proof. Choose a basis $\{e_1, \ldots, e_n\}$ for $T_{\gamma(0)}M$, and define $E_i(t) = P_t(e_i) \in T_{\gamma(t)}M$. Then we can write $V(t) = V^i(t)E_i(t) = V^i(t)P_t(e_i) = P_t(V^i(t)e_i)$, or in other words

$$P_t^{-1}(V) = V^i(t)e_i.$$

Therefore we have

$$\frac{d}{dt}\left(P_t^{-1}V(t)\right) = \left(\frac{d}{dt}V^i\right) e_i,$$

and

$$P_t \frac{d}{dt}\left(P_t^{-1}V(t)\right) = \left(\frac{d}{dt}V^i(t)\right) E_i(t).$$

On the other hand we have

$$\nabla_t V = \nabla_t \left(V^i(t)E_i(t)\right) = \left(\frac{d}{dt}V^i(t)\right) E_i(t) + V^i(t)\nabla_t E_i = \left(\frac{d}{dt}V^i(t)\right) E_i(t)$$

since E_i is parallel. □

This tells us that the connection is determined by the parallel transport operators. The parallel transport operators give a convenient way to identify the tangent spaces to M at different points along a smooth curve, and in some ways this is analogous to the left-shift maps $D_e l_g$ on a Lie group. However, it is important to note that the parallel transport operators depend on the curve. In particular the parallel transport operators cannot be extended to give canonical identifications of all the tangent spaces to each other (if we could do this, we could construct non-vanishing vector fields, but we already know this is impossible in some situations). Also, it does not really makes sense to think of a parallel vector field as being "constant", as the following example illustrates:

Example 8.5.4 Parallel transport on the sphere. Let $M = S^2$, and take the submanifold connection given by Example 8.2.2. Then a vector field along a curve in S^2 is parallel if and only if its rate of change (as a vector in $\mathbb{R}^3$) is always normal to the surface of S^2.

Consider the path γ on S^2 which starts at the north pole, follows a line of longitude to the equator, follows the equator for some distance (say, a quarter of the way around) and then follows another line of longitude back to the north pole. Note that each of the three segments of this curve is a geodesic. We compute the vector field given by parallel transport along γ of a vector which is orthogonal to the initial velocity vector at the north pole.

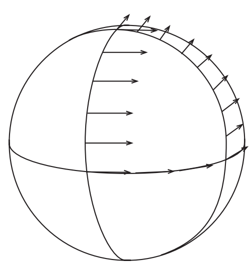
On the first segment, the parallel transport keeps the vector constant as a vector in $\mathbb{R}^3$ (this must be the parallel transport since it remains tangent to S^2, and has zero rate of change, so certainly the tangential component of its rate of change is zero).

On the segment around the equator, we start with a vector tangent to the equator, and parallel transport will give us the tangent vector to the equator of the same length as we move around the equator: Then we have $V = L\frac{dx}{dt}$ for some constant L (the length), and so

$$\dot{V} = L\frac{d^2x}{dt^2} = -Lx$$

since $x(t) = (\cos t, \sin t, 0)$. In particular, $\dot{V}$ is normal to S^2.

On the final segment, the situation is the same as the first segment: We can take V to be constant as a vector in $\mathbb{R}^3$.



So parallel transport around the entire loop has the effect of rotating the vector through $\pi/2$.

8.6 The connection as a projection.

Another way of looking at connections is the following: Let us return to the original problem, of defining a derivative for a vector field. Recall that a vector field $V \in \mathcal{X}(M)$ can be regarded as a smooth map $V : M \to TM$ for which $V_x \in T_xM$ for all $x \in M$. Since this is just a smooth map between manifolds, we can differentiate it!

This gives the derivative map $DV : TM \to TTM$. In other words, we can think of the derivative of a vector field on M in a direction $v \in TM$ as a vector tangent to the $2n$-dimensional manifold $T(TM)$. To get a better understanding of this, look at the situation geometrically: We think of TM as being a union of fibres T_xM. Thus a tangent vector to TM will have some component tangent to the fibre T_xM, and some component transverse to the fibre. In local coordinates $x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n$, the tangent vectors $\partial_1, \ldots, \partial_n$ corresponding to the first n coordinates represent change in position in M, which means that motion in these directions amounts to moving across a family of 'fibres' in TM; the tangent vectors $\dot{\partial}_1, \ldots, \dot{\partial}_n$ corresponding to the last n coordinates (which are just the components of the vector in TM) are tangent to the fibres.

A vector field V can be written in these coordinates as $V(x^1, \ldots, x^n) = (x^1, \ldots, x^n, V^1, \ldots, V^n)$. Then

$$D_vV = v^i \partial_i + v(V^i)\dot{\partial}_i.$$

The idea of the connection is to project this onto the subspace of $T(TM)$ tangent to the fibre, which we can naturally identify with the fibre itself (the fibre is a vector space, so we can identify it with its tangent space at each point). We will denote by $\mathcal{V}$ the subspace of $T(TM)$ spanned by $\dot{\partial}_1, \ldots, \dot{\partial}_n$ (note that this is independent of the choice of local coordinates), and we call this the *vertical subspace* of $T(TM)$. This is naturally identified with TM by the map ι which sends $v^i\dot{\partial}_i$ to $v^i\partial_i$.

Definition 8.6.1 A **vertical projection** on TM is a map $\xi \mapsto \pi(\xi)$, where $\xi = (p, v) \in TM$ and $\pi(\xi)$ is a linear map from $T_\xi(TM) \to T_pM$ such that

(1). $\pi(\xi) = \iota$ on $\mathcal{V}$;

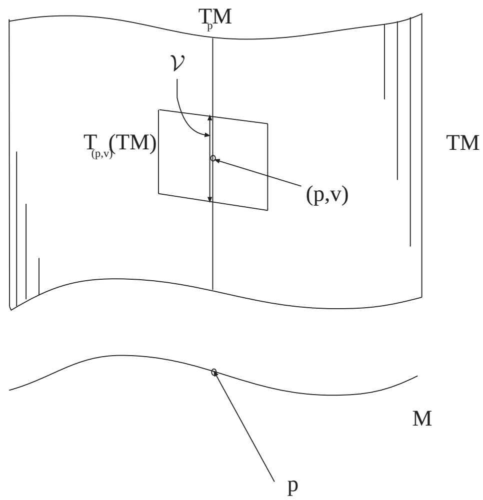
(2). π is consistent with the additive structure on TM: If we take γ_1 and γ_2 to be paths in TM of the form $\gamma_i(t) = (p(t), v_i(t))$, then

$$\pi_{(p,v_1)}(\gamma_1') + \pi_{(p,v_2)}(\gamma_2') = \pi_{(p,v_1+v_2)}(\gamma')$$

where $\gamma(t) = (p(t), v_1(t) + v_2(t))$.

Given a vertical projection π, we can produce a connection as follows: If $X \in \mathcal{X}(M)$ and $v \in T_pM$, define

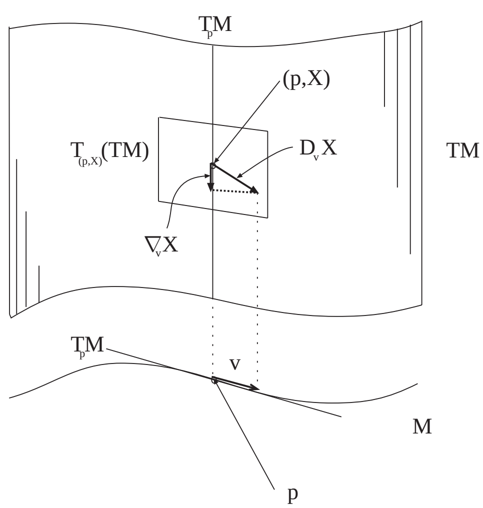
$$\nabla_v X = \pi_{(p,X_p)}(D_vX).$$



Conversely, given a connection, we can produce a corresponding vertical projection by taking

$$\pi_{(p,X)}(v^i\partial_i + \dot{v}^i\dot{\partial}_i) = \nabla_t V,$$

the covariant derivative along the curve $\gamma(t) = \varphi^{-1}\left(\varphi(p) + tv^i e_i\right)$ of the vector field $V(t) = \left(X^i + t\dot{v}^i\right)\partial_i$.



It is interesting to consider the parallel transport operators in terms of the vertical projections: Since π maps a $2n$-dimensional vector space to an n-dimensional vector space, and is non-degenerate on the n-dimensional 'vertical' subspace tangent to the fiber of TM, the kernel of π is an n-dimensional subspace of $T(TM)$ which is complementary to the vertical subspace. We call this the *horizontal subspace* $\mathcal{H}$ of $T(TM)$. A vector field X is parallel along a curve γ if and only if D_tX lies in $\mathcal{H}$ at every point. We will come back to this description when we discuss curvature. A vertical projection is uniquely determined by the choice of a horizontal subspace at each point, complementary to the vertical subspace and consistent with the linear structure.

8.7 Existence of connections.

We will show that every smooth manifold can be equipped with a connection – in fact, there are many connections on any manifold, and no preferred or canonical one (later, when we introduce Riemannian metrics, we will have a way of producing a canonical connection).

Definition 8.7.1 Let M be a smooth manifold. A (smooth) **partition of unity** on M is a collection of smooth functions $\{\rho_\alpha\}_{\alpha\in I}$ which are non-negative, locally finite (i.e. for every $x \in M$ there exists a neighbourhood U of x on which only finitely many of the functions ρ_α are non-zero), and sum to unity: $\sum_\alpha \rho_\alpha(x) = 1$ for all $x \in M$ (note that this is a finite sum for each x). Let $\{U_\beta\}_{\beta\in J}$ be a locally finite covering of M by open sets. Then a partition of unity $\{\rho_\alpha\}$ is **subordinate** to this cover if each of the functions ρ_α has support contained in one of the sets U_β.

We assume that our manifolds are paracompact – that is, every covering of M by open sets has a refinement which is locally finite. In particular, we have an atlas $\mathcal{A}$ for M which is locally finite (i.e. each point of M lies in only finite many of the coordinate charts of the atlas).

Proposition 8.7.2 *For any locally finite cover of M by coordinate charts there exists a subordinate partition of unity.*

Proof. By refining if necessary, assume that the coordinate charts have images in $\mathbb{R}^n$ which are bounded open sets which have smooth boundary. Consider one of these charts, $\varphi_\alpha : U_\alpha \to V_\alpha$. We can choose a number $\epsilon > 0$ such that the distance function $d(y) = d(y, \partial V_\alpha)$ to the boundary of V_α is smooth at points where $d(y) < \varepsilon$. Then define

$$\psi_\alpha = F_\varepsilon \circ d \circ \varphi_\alpha,$$

where $F_\varepsilon : \mathbb{R} \to \mathbb{R}$ is defined by $F_\varepsilon(t) = F_1(\varepsilon t)$, and F_1 is a smooth function which is identically zero for $t \leq 0$, monotone increasing for $t \in (0, 1)$, and identically 1 for $t \geq 1$. It follows that ψ_α is smooth on U_α, extends smoothly to M by taking zero values away from U_α, and is strictly positive on U_α.

Finally, define

$$\rho_\alpha(x) = \frac{\psi_\alpha(x)}{\sum_{\alpha'} \psi_{\alpha'}(x)}.$$

□

Proposition 8.7.3 *For any smooth manifold M there exists a connection on M.*

Proof. Choose a cover of M by coordinate charts $\{U_\alpha\}$, and a subordinate partition of unity $\{\rho_\alpha\}$. We will use this partition of unity to patch together the connections on each chart given by coordinate differentiation: Recall that on each coordinate chart U_α we have a derivative $\nabla^{(\alpha)}$ which is well-defined on vector fields with support inside U_α. Then we define

$$\nabla_v X = \sum_\alpha \nabla_v^{(\alpha)} \left(\rho_\alpha X\right).$$

This makes sense because the sum is actually finite at each point of M. The result is clearly linear (over $\mathbb{R}$) in both v and X, and

$$\begin{aligned}
\nabla_v(fX) &= \sum_\alpha \nabla_v^{(\alpha)} \left(f\rho_\alpha X\right) \\
&= \sum_\alpha \left(v(f)\rho_\alpha X + f\nabla_v^{(\alpha)} \left(\rho_\alpha X\right)\right) \\
&= v(f)X + f\nabla_v X
\end{aligned}$$

where I used the fact that $\sum_\alpha \rho_\alpha = 1$ to get the first term in the last equality.

□

The following exercise is a another application of partitions of unity, to show the existence of F-related vector fields. Recall that if $F : M \to N$ is an smooth map, then $X \in \mathcal{X}(M)$ and $Y \in \mathcal{X}(N)$ are called F-related if $D_x F(X_x) = Y_{F(x)}$ for every $x \in M$.

Exercise 8.7.2 Suppose $F : M \to N$ is an embedding. Show that for any vector field $X \in \mathcal{X}(M)$ there exists a (generally not unique) F-related vector field $Y \in \mathcal{X}(Y)$ [Hint: Cover N by coordinate charts, such that those containing a portion of $F(M)$ are submanifold charts (i.e. $\varphi_\alpha : U_\alpha \to V_\alpha$ such that $F(M) \cap U_\alpha = \varphi_\alpha^{-1}(\mathbb{R}^k \times \{0\})$). Choose a partition of unity subordinate to this cover. Use each submanifold chart to extend X from $F(M)$ to a vector field on the chart in N, and take a zero vector field in all the other charts on N. Patch these together using the partition of unity and check that it gives a vector field with the required properties].

8.9 The difference of two connections

Suppose $\nabla^{(1)}$ and $\nabla^{(2)}$ are two connections on a smooth manifold M. Then we have the following important result:

Proposition 8.9.1 *For any vector $v \in T_xM$ and any vector field $X \in \mathcal{X}(M)$, the difference*

$$\nabla_v^{(2)} X - \nabla_v^{(1)} X$$

depends only on the value of X at the point x.

Proof. In local coordinates, writing $v = v^i \partial_i$ and $X = X^j \partial_j$,

$$\begin{aligned}\nabla_v^{(2)} X - \nabla_v^{(1)} X &= v(X^j)\partial_j + v^j X^i {\Gamma_{ji}^{(1)}}^k \partial_k - v(X^j)\partial_v - v^j X^i {\Gamma_{ji}^{(2)}}^k \partial_k \\ &= v^j X^i \left({\Gamma_{ji}^{(2)}}^k - {\Gamma_{ji}^{(1)}}^k \right) \partial_k.\end{aligned}$$

□

It follows that the difference of two connections defines a map for each $x \in M$ from $T_xM \times T_xM$ to T_xM, which varies smoothly in x and is linear in each argument. We also have a kind of converse to this:

Proposition 8.9.2 *Let M be a smooth manifold, and ∇ a connection on M. Suppose $A_x : T_xM \times T_xM \to T_xM$ is linear in each argument for each $x \in M$, and varies smoothly over M (in the sense that A applied to two smooth vector fields gives a smooth vector field). Then*

$$x, X \mapsto \nabla_v X + A(v, X)$$

is also a connection on M.

The proof is a simple calculation to check that the Leibniz rule still holds. It is easy to show that there are many such suitable maps A (for example, one can write in a local coordinate chart $A(u^i \partial_i, v^j \partial_j) = u^i v^j {A_{ij}}^k \partial_k$ where the coefficient functions ${A_{ij}}^k$ are smooth and have support inside the coordinate chart. Then sums and linear combinations of such things are also suitable. Thus there are many different connections on a manifold.

8.10 Symmetry and torsion

Given a connection ∇, we define its torsion in the following way: If X and Y are two smooth vector fields, then we observe that

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

evaluated at a point $x \in M$, depends only on the values of X and Y at the point x. This follows because

$$\begin{aligned}
\nabla_{fX}(gY) - \nabla_{gY}(fX) - [fX, gY] &= fg\nabla_X Y + fX(g)Y \\
&\quad - fg\nabla_Y X - gY(f)X \\
&\quad - fX(g)Y + gY(f)X - fg[X, Y] \\
&= fg\left(\nabla_X Y - \nabla_Y X - [X, Y]\right).
\end{aligned}$$

This map is called the *torsion* $T(X, Y)$ of ∇ applied to X and Y. This can be written in terms of its components ${T_{ij}}^k$, defined by $\nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = {T_{ij}}^k\partial_k$.

Definition 8.10.1 A connection ∇ is called **symmetric** if the torsion of ∇ vanishes — i.e. $\nabla_X Y - \nabla_Y X = [X, Y]$ for every pair of vector fields X and Y on M.

The standard connection on $\mathbb{R}^n$ is symmetric, as are submanifold connections. However, the left-invariant connection on a Lie group is typically not symmetric: If X and Y are left-invariant vector fields, then $\nabla_X Y = \nabla_Y X = 0$, but if G is non-commutative then usually $[X, Y] \neq 0$.

8.11 Geodesics and the exponential map

Definition 8.11.1 Let M be a smooth manifold, and ∇ a connection. A curve $\gamma : [0, 1] \to M$ is a **geodesic** of ∇ if the tangent vector $\dot\gamma$ is parallel along γ:

$$\nabla_t \dot\gamma = 0.$$

We can think of a path γ as the trajectory of a particle moving in M. Then γ is a geodesic if the acceleration of the particle (measured using ∇) is zero, so γ is in this sense the trajectory of a 'free' particle. This picture is particularly relevant in the case where M is a submanifold of Eulidean space, since then the geodesics are exactly the trajectories of particles which are constrained to lie in M but are otherwise free of external forces.

Proposition 8.11.2 *For any $x \in M$ and any $v \in T_x M$ there exists $\delta > 0$ and a unique geodesic $\gamma : (-\delta, \delta) \to M$ such that $\gamma(0) = x$ and $\dot\gamma(0) = v$.*

Proof. Let $\varphi : U \to V$ be a chart for M about x. Then for any path $\gamma : I \to M$, we have

$$\nabla_t \dot\gamma = \nabla_t \left(\dot\gamma^i \partial_i\right) = \left(\ddot\gamma^i + \Gamma_{kl}{}^i \dot\gamma^k \dot\gamma^l\right) \partial_i,$$

so a geodesic must satisfy the system of second order equations

$$\ddot\gamma^i + \Gamma_{kl}{}^i \dot\gamma^k \dot\gamma^l = 0$$

for $i = 1, \dots, n$. Existence and uniqueness then follow from the standard theory of second order ODE. Alternatively, we can rewrite as a system of $2n$ first-order equations by setting $v^i = \dot\gamma^i$, which gives

$$\begin{aligned}
&\frac{d}{dt}\gamma^i(t) = v^i(t);\\
&\frac{d}{dt}v^i(t) = -\Gamma_{kl}{}^i(\gamma^1(t), \dots, \gamma^n(t)) v^k(t) v^l(t).
\end{aligned}$$

Then apply the ODE existence and uniqueness theorem from Lecture 6 to deduce that there is a unique geodesic for some time starting from any point x and any vector v (i.e. any values of $\gamma^1, \dots, \gamma^n, v^1, \dots, v^n$). □

Using Proposition 8.11.2 we can define the *exponential map* on a smooth manifold with a connection, by analogy with the exponential map on a Lie group: Given $x \in M$ and $v \in T_x M$, we define

$$\exp_p(v) = \exp(p, v) = \gamma(1),$$

where γ is a geodesic with initial position x and initial velocity v. Then we know that the exponential map is defined on a neighbourhood of the zero section $\{(x, 0) : x \in M\}$ within TM. In order to determine the smoothness propoerties of the exponential map, it is useful to interpret it using the flow of a vector field so that we can apply the methods of Lecture 6. As we saw above, the geodesic equation corresponds to a second-order system of equations, and it turns out that this arises not from the flow of a vector field on M, but from the flow of a vector field on the tangent bundle TM: Recall that the tangent bundle can be given local coordinates $x^1, \dots, x^n, \dot x^1, \dots, \dot x^n$. Then let $\partial_1, \dots, \partial_n, \dot\partial_1, \dot\partial_n$ be the corresponding coordinate tangent vectors, and define a vector field $\mathcal{G}$ on TM by

$$\mathcal{G}_{(x,v)} = \dot x^i \partial_i - \Gamma_{kl}{}^i(x) \dot x^k \dot x^l \dot\partial_i.$$

Then our local calculations above show that the geodesic equation is given exactly by the flow for time 1 of the vector field $\mathcal{G}$. More precisely, it is given by the projection onto M of this flow.

Exercise 8.11.3 Show that $\mathcal{G}$, defined as above in local coordinates, is well-defined (i.e. does not depend on the choice of coordinates).

It follows immediately that the exponential map is smooth, and that the map

$$(x, v) \mapsto \left(\exp_x(v), \frac{d}{dt} \exp_x(tv)\Big|_{t=1} \right)$$

from TM to itself is a local diffeomorphism where defined.

Proposition 8.11.4 *For any $x \in M$, the exponential map $\exp_x : T_xM \to M$ from x is a local diffeomorphism near $0 \in T_xM$, onto a neighbourhood of x in M.*

Proof. The smoothness of the map is immediate since it is the composition of the smooth flow of the vector field $\mathcal{G}$ on TM with the smooth projection onto M. Next we calculate the derivative of this map at the origin in the direction of some vector v: We have $\exp_x(v) = \gamma(1)$, where γ is the geodesic through x with velocity v. But then also $\exp_x(tv) = \gamma(t)$, since $s \mapsto \gamma(st)$ is a geodesic through x with velocity tv. Therefore

$$D_0 \exp_x(v) = \frac{d}{dt} \exp_x(tv)\Big|_{t=0} = \gamma'(0) = v$$

Therefore $D_0 \exp_x$ is just the identity map (here we are identifying the tangent space to T_xM at 0 with T_xM itself, in the usual way). In particular, this derivative is non-degenerate, so $\exp_x$ is a local diffeomorphism. □

We can get a slightly stronger statement:

Proposition 8.11.5 *For any $x \in M$ there exists an open neighbourhood U of $x \in M$, and an open neighbourhood $\mathcal{O}$ of the zero section $\{(y, 0) : y \in U\}$ such that for every pair of points p and q in U, there exists a unique geodesic γ_{qr} with $\gamma_{qr}(0) = 1$, $\gamma_{qr}(1) = r$, and $\gamma'(0) \in \mathcal{O}$.*

Roughly speaking, this says that points which are sufficiently close to each other can be joined by a unique "short" geodesic (we cannot expect to drop the shortness condition – consider for example the case where M is a cylinder or a sphere, when there are generally many geodesics joining any pair of points.

Proof. Consider the map $\tilde{\exp} : TM \to M \times M$ defined by

$$\tilde{\exp}(p, v) = (p, \exp_p(v)).$$

Let us compute the derivative of this map at a point $(p, 0)$: Breaking up the $2n \times 2n$ derivative matrix into $n \times n$ blocks, the top left corresponds to the derivative of the map $(p, v) \mapsto p$ with respect to p, which is just the identity map, and the bottom left block is the derivative of the same with respect to v, which is zero. The top right block might be quite complicated, but as we calculated in the proof of proposition 8.11.4, the bottom right block is the derivative of the exponential map from p with respect to v at the origin, which is again just the identity map:

$$D_{(p,0)}\tilde{\exp} = \begin{bmatrix} I & * \\ 0 & I \end{bmatrix}.$$

In particular, this is non-degenerate, so $\tilde{\exp}$ is a local diffeomorphism near $(p, 0)$ for any p. Choose a small neighbourhood $\mathcal{O}$ of $(x, 0)$ in TM, and a neighbourhood U of x in M sufficiently small that $U \times U \subset \tilde{\exp}(\mathcal{O})$. These then satisfy the requirements of the proposition. □

The following exercise gives an interpretation of the torsion of a connection in terms of geodesics and parallel transport:

Exercise 8.11.6 Let M be a smooth manifold with a connection ∇. Fix $p \in M$, and vectors X and Y in T_pM. Let $q(t)$ be the point of M constructed as follows: First, follow the geodesic from p in direction X until time $\sqrt{t}$, to get a point p_1. Then parallel transport Y along that geodesic to get a vector Y_1 at p_1. Also, let X_1 be tangent vector to the geodesic at time t (i.e. at the point p_1). Then follow the geodesic from p_1 in direction Y_1 until time $\sqrt{t}$, to get a point p_2, and parallel transport X_1 and Y_1 along this geodesic to get vectors X_2 and Y_2 at p_2 (Thus, Y_2 is the tangent vector of the geodesic at p_2). Then take the geodesic from p_2 in direction $-X_2$ until time $\sqrt{t}$, parallel transporting X_2 and Y_2 along this to get vectors X_3 and Y_3 at a point p_3. Finally, follow the geodesic from p_3 in direction $-Y_3$ until time $\sqrt{t}$, and call the resulting point $q(t)$.

Show that the tangent vector to the curve $t \mapsto q(t)$ is given by the torsion of ∇ applied to the vectors X and Y. In other words, the torsion measures the extent to which geodesic rectangles fail to close up.