# Riemannian Geometry

Eckhard Meinrenken

Lecture Notes, University of Toronto, Spring 2002

# Contents

## 1. Manifolds

**1.1. Motivation.** One of Riemann's key ideas was to develop the notion of a "manifold" independent from an embedding into an ambient Euclidean space. Roughly, an $n$-dimensional manifold is a space that locally looks like $\mathbb{R}^n$. More precisely, a manifold is a space that can be covered by coordinate charts, in such a way that the change of coordinates between any two charts is a smooth map. The following examples should give an idea what we have in mind.

Example 1.1. Let $S^2 \subset \mathbb{R}^3$ be the unit sphere defined by the equation $(x_0)^2+(x_1)^2+(x_2)^2$. For $j = 0, 1, 2$ let $U_j^+ \subset S^2$ be the subset defined by $x_j > 0$ and $U_j^-$ the subset defined by $x_j < 0$. Let $\phi_j^\pm : U_j^\pm \to \mathbb{R}^2$ be the maps omitting the $j$th coordinate. Then all transition maps $\phi_j^\pm \circ (\phi_k^\pm)^{-1}$ are smooth. For instance,

$$\phi_2^+ \circ \phi_1^- : \phi_1^-(U_1^- \cap U_2^+) \to \phi_2^+(U_1^- \cap U_2^+)$$

is the map $(u, v) \mapsto (u, -\sqrt{1 - u^2 - v^2})$, and this is smooth since $u^2 + v^2 < 1$ on the image of $\phi_1^-$.

Example 1.2. The real projective plane $\mathbb{R}P(2)$ is the set of all lines (=1-dimensional subspaces) in $\mathbb{R}^3$. Any such line is determined by its two points of intersection$\{x, -x\}$ with $S^2$. Thus $\mathbb{R}P(2)$ may be identified with the quotient of $S^2$ by the equivalence relation, $x \sim -x$. Let $\pi : S^2 \to \mathbb{R}P(2)$ be the quotient map. To get a picture of $\mathbb{R}P(2)$, note that for $0 < \epsilon < 1$, the subset $\{(x_0, x_1, x_2) \in S^2 | \ x_2 \geq \epsilon\}$ is a 2-disk, containing at most one element of each equivalence class. Hence its image under $\pi$ is again a 2-disk. On the other hand, the strip $\{(x_0, x_1, x_2) \in S^2 | \ -\epsilon \leq x_2 \geq \epsilon\}$ contains, with any $x$, also the point $-x$. Its image under $\pi$ looks like a Moebius strip. Thus $\mathbb{R}P(2)$ looks like a union of a Moebius strip and a disk, glued along their boundary circles. This is still somewhat hard to imagine, since we cannot perform this gluing in such a way that $\mathbb{R}P(2)$ would become a surface in $\mathbb{R}^3$. Nonetheless, it "should be" a surface: Using the coordinate charts from $S^2$, let $U_j = \pi(U_j^+)$, and let $\phi_j : U_j \to \mathbb{R}^2$ be the unique maps such that $\pi \circ \phi_j = \phi_j^+$. Then the $U_j$ cover $\mathbb{R}P(2)$, and the "change of coordinate" maps are again smooth.

It is indeed possible to embed $\mathbb{R}P(2)$ into $\mathbb{R}^4$: One possibility is the map,

$$[(x_0, x_1, x_2)] \mapsto (x_1x_2, x_0x_2, x_0x_1,\ t_0x_0^2 + t_1x_1^2 + t_2x_2^2) \tag{1}$$

where $t_0, t_1, t_2 \in \mathbb{R}$ are distinct (e.g. $t_0 = 1, t_1 = 2, t_2 = 3$). However, these embedding do not induce the "natural" metric on projective space, i.e. the metric induced from the 2-sphere.

**1.2. Topological spaces.** To develop the concept of a manifold as a "space that locally looks like $\mathbb{R}^n$", our space first of all has to come equipped with some topology (so that the word "local" makes sense). Recall that a *topological space* is a set $M$, together with a collection of subsets of $M$, called *open subsets*, satisfying the following three axioms: (i) the empty set $\emptyset$ and the space $M$ itself are both open, (ii) the intersection of any finite collection of open subsets is open, (iii) the union of *any* collection of open subsets is open. The collection of open subsets of $M$ is also called the topology of $M$. A map $f : M_1 \to M_2$ between topological spaces is called *continuous* if the pre-image of any open subset in $M_2$ is open in $M_1$. A continuous map with a continuous inverse is called a *homeomorphism*.

One basic ingredient in the definition of a manifold is that our topological space comes equipped with a covering by open sets which are homeomorphic to open subsets of $\mathbb{R}^n$.

DEFINITION 1.3. Let $M$ be a topological space. An *$n$-dimensional chart* for $M$ is a pair $(U, \phi)$ consisting of an open subset $U \subset \mathbb{R}^n$ and a continuous map $\phi : U \to \mathbb{R}^n$ such that $\phi$ is a homeomorphism onto its image $\phi(U)$. Two such charts $(U_\alpha, \phi_\alpha)$, $(U_\beta, \phi_\beta)$ are *$C^\infty$-compatible* if the *transition map*

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$$

is a diffeomorphism (a smooth map with smooth inverse). A covering $\mathcal{A} = (U_\alpha)_{\alpha \in A}$ of $M$ by pairwise $C^\infty$-compatible charts is called a *$C^\infty$-atlas*.

EXAMPLE 1.4. Let $X \subset \mathbb{R}^2$ be the union of lines $\mathbb{R} \times \{1\} \cup \mathbb{R} \times \{-1\}$. Let $M = X/\sim$ be its quotient by the equivalence relation, $(u, 1) \sim (u, -1)$ for $u < 0$. Let $\pi : X \to M$ be the quotient map.

Thus $M$ is obtained by gluing to copies of the real line along the negative axis. It is somewhat hard to picture this space, since $\pi(0, \pm 1)$ are *distinct* points in $M$. Nonetheless, $M$ admits a $C^\infty$-atlas: Let $U_+ = \pi(\mathbb{R} \times \{1\})$ and $U_- = \pi(\mathbb{R} \times \{-1\})$, and define $\phi_\pm : U_\pm \to \mathbb{R}$ by $\phi_\pm(u, \pm 1) = u$. Then $(U_\pm, \phi_\pm)$ defines an atlas with two charts (the transition map is just the identity map).

The example just given shows that existence of an atlas does not imply that our space looks "nice". The problem with the example is that the points $\pi(0, \pm 1)$ in $M$ do not admit disjoint open neighborhoods. Recall that a topological space is called *Hausdorff* if any two distinct points in the space admit disjoint open neighborhoods. Thus, we require manifolds to be Hausdorff.

We will impose another restriction on the topology. Recall that A *basis* for a topological space $M$ is a collection $\mathcal{B}$ of open subsets of $M$ such that every open subset of $M$ is a union of open subsets in the collection $\mathcal{B}$. For example, the collection of open balls $B_\epsilon(x)$ in $\mathbb{R}^n$ define a basis. But one already has a basis if one takes only all balls $B_\epsilon(x)$ with $x \in \mathbb{Q}^n$ and $\epsilon \in \mathbb{Q}_{>0}$; this then defines a countable basis. A topological space with countable basis is also called *second countable*. We will require manifolds to admit a countable basis. This will imply, among other things, that the manifold admits a countable atlas, a fact that is useful for certain inductive arguments.

Two atlases on a topological space are called *equivalent* if their union is again an atlas. It is not hard to check that this is indeed an equivalence relation. An equivalence class of atlases is called a $C^\infty$-structure on $M$.

DEFINITION 1.5 (Manifolds). A *$C^\infty$-manifold* is a Hausdorff topological space $M$, with countable basis, together with a $C^\infty$-structure.

It is perhaps somewhat surprising that the two topological restrictions (Hausdorff and countable basis) rule out any further "accidents": The topological properties of manifolds are just as nice as those of Euclidean $\mathbb{R}^n$.

DEFINITION 1.6. A map $F : N \to M$ between manifolds is called smooth (or $C^\infty$) if for all charts $(U, \phi)$ of $N$ and $(V, \psi)$ of $M$, with $F(U) \subset V$, the composite map

$$\psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V)$$

is smooth. The space of smooth maps from $N$ to $M$ is denoted $C^\infty(N, M)$. A smooth map $F : N \to M$ with smooth inverse $F^{-1} : M \to N$ is called a *diffeomorphism*.

PROBLEMS 1.7.

1. Review point set topology: Continuous maps, coverings, neighborhoods, Hausdorff property, compactness, ...
2. Show that equivalence of $C^\infty$-atlases is an equivalence relation. Warning: $C^\infty$-compatibility of charts on a topological space is *not* an equivalence relation. (Why?)
3. Given a manifold $M$ with $C^\infty$-atlas $\mathcal{A}$, let $\mathcal{A}'$ be the collection of all $C^\infty$-charts $(U, \phi)$ on $M$ that are compatible with all charts in $\mathcal{A}$. Show that $\mathcal{A}'$ is again an atlas, and that $\mathcal{A}'$ contains any atlas equivalent to $\mathcal{A}$.
4. Verify that the map (1) is 1-1.

## 2. Examples of manifolds

**Spheres**. The unit sphere $S^n \subset \mathbb{R}^{n+1}$ is a manifold of dimension $n$, with charts $U_j^\pm$ constructed similar to $S^2$. Another choice of atlas, with only two charts, is given by "stereographic projection" from the north and south pole.

**Projective spaces**. Let $\mathbb{R}P(n)$ be the quotient $S^n/\sim$ under the equivalence relation $x \sim -x$. It is easy to check that this is Hausdorff and has countable basis. Let $\pi: S^n \to \mathbb{R}P(n)$ be the quotient map. Just as for $n = 2$, the charts $U_j = \pi(U_j^+)$, with map $\phi_j$ induced from $U_j^+$, form an atlas.

**Products**. If $M_j$ are a finite collection of manifolds of dimensions $n_j$, their direct product is a manifold of dimension $\sum n_j$. For instance, the $n$-torus is defined as the $n$-fold product of $S^1$'s.

**Lens spaces**. Identify $\mathbb{R}^4$ with $\mathbb{C}^2$, thus $S^3 = \{(z, w) : \ |z|^2 + |w|^2 = 1\}$. Given natural numbers $q > p \geq 1$ introduce an equivalence relation, by declaring that $(z, w) \sim (z', w')$ if

$$(z', w') = (e^{2\pi i \frac{k}{q}} z,\ e^{2\pi i \frac{kp}{q}} w)$$

for some $k \in \{0, \ldots, q-1\}$. Let $L(p, q) = S^3/\sim$ be the *lens space*. Note that $L(1, 2) = \mathbb{R}P(3)$. If $p, q$ are relatively prime, $L(p, q)$ is a manifold. Indeed, if $p, q$ are relatively prime then for all $(z, w) \in S^3$, the only solution of

$$(z, w) = \Phi_k(z, w) := (e^{2\pi i \frac{k}{q}} z,\ e^{2\pi i \frac{kp}{q}} w)$$

is $k = 0$. Let $f_k(z, w) = ||(z, w) - \Phi_k(z, w)||$. Then $f_k > 0$ for $k = 1, \ldots, q-1$. Since $S^3$ is compact, each $f_k$ takes on its minimum on $S^3$. Let $\epsilon > 0$ be sufficiently small so that $f_k > \epsilon$ for all $k = 1, \ldots, q-1$.

Then if $U$ is an open subset of $S^3$ that is contained in some open ball of radius $\epsilon$ in $\mathbb{R}^3$, then $U$ contains at most one element of each equivalence class. Let $(U, \phi)$ be a coordinate chart for $S^3$, with $U$ sufficiently small in this sense. Let $V = \pi(U)$, and $\psi : V \to \mathbb{R}^3$ the unique map such that $\psi \circ \pi = \phi$. Then $(V, \psi)$ is a coordinate chart for $L(p, q)$, and the collection of coordinate charts constructed this way defines an atlas.

**Grassmannians.** The set $\mathrm{Gr}(k, n)$ of all $k$-dimensional subspaces of $\mathbb{R}^n$ is called the Grassmannian of $k$-planes in $\mathbb{R}^n$. A $C^\infty$-atlas may be constructed as follows. For any subset

$I \subset \{1, \ldots, n\}$ let $I' = \{1, \ldots, n\} \backslash I$ be its complement. Let $\mathbb{R}^I \subset \mathbb{R}^n$ be the subspace consisting of all $x \in \mathbb{R}^n$ with $x_i = 0$ for $i \notin I$.

If $I$ has cardinality $k$, then $\mathbb{R}^I \in \mathrm{Gr}(k, n)$. Note that $\mathbb{R}^{I'} = (\mathbb{R}^I)^\perp$. Let $U_I = \{E \in \mathrm{Gr}(k, n) | \, E \cap \mathbb{R}^{I'} = 0\}$. Each $E \in U_I$ is described as the graph of a unique linear map $A : \mathbb{R}^I \to \mathbb{R}^{I'}$, that $E = \{x + A(x) | \, x \in \mathbb{R}^I$. This gives a bijection

$$\phi_I : U_I \to L(\mathbb{R}^I, \mathbb{R}^{I'}) \cong \mathbb{R}^{k(n-k)}.$$

We can use this to define the topology on the Grassmannian: It is the smallest topology for which all maps $\phi^I$ are continuous. To check that the charts are compatible, suppose $E \in U_I \cap U_{\tilde{I}}$, and let $A_I$ and $A_{\tilde{I}}$ be the linear maps describing $E$ in the two charts. We have to show that the map taking $A_I$ to $A_{\tilde{I}}$ is smooth. Let $\Pi_I$ denote orthogonal projection $\mathbb{R}^n \to \mathbb{R}^I$. The map $A_I$ is determined by the equations

$$A_I(x_I) = (1 - \Pi_I)x, \quad x_I = \Pi_I x$$

for $x \in E$, and $x = x_I + A_I x_I$. Thus

$$A_{\tilde{I}}(x_{\tilde{I}}) = (I - \Pi_{\tilde{I}})(A_I + 1)x_I, \quad x_{\tilde{I}} = \Pi_{\tilde{I}}(A_I + 1)x_I.$$

The map $S(A_I) : \Pi_{\tilde{I}}(A_I + 1) : \mathbb{R}^I \to \mathbb{R}^{\tilde{I}}$ is an isomorphism, since it is the composition of two isomorphisms $(A_I + 1) : \mathbb{R}^I \to E$ and $\Pi_{\tilde{I}}|_E : E \to \mathbb{R}^{\tilde{I}}$. The above equations show,

$$A_{\tilde{I}} = (I - \Pi_{\tilde{I}})(A_I + 1)S(A_I)^{-1}.$$

The dependence of $S$ on the matrix entries of $A_I$ is smooth, by Cramer's formula for the inverse matrix. It follows that the collection of all $\phi_I : U_I \to \mathbb{R}^{k(n-k)}$ defines on $\mathrm{Gr}(k, n)$ the structure of a manifold of dimension $k(n - k)$.

**Rotation groups.** Let $\mathrm{Mat}_n \cong \mathbb{R}^{n^2}$ be the set of $n \times n$-matrices. The subset $\mathrm{SO}(n) = \{A \in \mathrm{Mat}_n \,|\, A^t A = I, \ \det(A) = 1$ is the group of rotations in $\mathbb{R}^n$. Let $\mathfrak{so}(n) = \{B \in \mathrm{Mat}_n \,|\, B^t + B = 0\} \cong \mathbb{R}^{n(n-1)/2}$. Then $\exp(B) \in \mathrm{SO}(n)$ for all $B \in \mathfrak{so}(n)$. For $\epsilon$ sufficiently small, exp restricts to a bijection from $V = \{B \in \mathfrak{so}(n) | \, ||B|| < \epsilon\}$. For any $A_0$, let $U = \{A \in \mathrm{SO}(n) | \, A = A_0 \exp(B)\}$. Let $\phi$ be the map taking $A$ to $B = \log(A A_0^{-1})$. Then the set of all $(U, \phi)$ constructed this way define an atlas, and give $\mathrm{SO}(n)$ the structure of a manifold of dimension $n(n-1)/2$.

## 3. Submanifolds

Let $M$ be a manifold of dimension $m$.

Definition 3.1. A subset $S \subset M$ is called an *embedded submanifold* of dimension $k \leq m$, if $S$ can be covered by coordinate charts $(U, \phi)$ for $M$ with the property $\phi(U \cap S) = \Phi(U) \cap \mathbb{R}^k$. Charts $(U, \phi)$ of $M$ with this property are called submanifold charts for $S$.

Thus $S$ becomes a $k$-dimensional manifold in its own right, with atlas consisting of charts $(U \cap S, \phi|_{U \cap S})$.

Example 3.2. $S^n$ is a submanifold of $\mathbb{R}^{n+1}$: A typical submanifold chart is

$$V = \{x \in \mathbb{R}^{n+1} | \, x_0 > 0, \textstyle\sum_{i>0} x_i^2 < 1\}, \quad \phi(x) = (x_1, \ldots, x_n, \sqrt{1 - \textstyle\sum_{i>0} x_i^2} - x_0).$$

EXAMPLE 3.3. Similarly, if $f : U \to \mathbb{R}^{n-k}$ is any smooth function on an open subset $U \subset \mathbb{R}^k$, the graph $\Gamma_f = \{(x, f(x)) | \, x \in U\}$ is an embedded submanifold of $U \times \mathbb{R}^k$, with submanifold chart

$$\phi : U \times \mathbb{R}^k \to \mathbb{R}^n, \quad (y', y'') \mapsto (y', f(y') - y'')$$

Recall that for any smooth function $F : V \to \mathbb{R}^m$ on an open subset $V \subset \mathbb{R}^n$, a point $a \in \mathbb{R}^m$ is a *regular value* if for all $x \in F^{-1}(a)$, the Jacobian $DF(x) : \mathbb{R}^n \to \mathbb{R}^m$ is onto. If $m = 1$, this means that the gradient $\nabla F$ does not vanish on $F^{-1}(a)$. (We do not require that $a$ is in the image of $F$ – thus regular *value* is a bit misleading.)

PROPOSITION 3.4. *Let $V \subset \mathbb{R}^n$ open, and $F : V \to \mathbb{R}^m$ smooth. For any regular value $a \in \mathbb{R}^m$ of $F$, the inverse image $F^{-1}(a)$ is an embedded submanifold of dimension $k = n - m$. In fact, there exists a submanifold chart $(U, \phi)$ around any given $x \in F^{-1}(a)$ such that*

$$F(\phi^{-1}(y', y'')) = a + y''$$

*for all $(y', y'') \in \phi(U) \subset \mathbb{R}^k \times \mathbb{R}^m$.*

PROOF. This is really just a version of the implicit function theorem from multivariable calculus. A more familiar version of this theorem states that for all $x \in F^{-1}(a)$, after possibly renumbering the coordinates in $\mathbb{R}^n$, the equation $F(y) = a$ can be "solved" for $y'' = (y_{k+1}, \ldots, y_n)$ as a function of $y' = (y_1, \ldots, y_k)$. That is, there exists a *unique* function $g_a$ from a neighborhood of $x' \in \mathbb{R}^{n-m}$ to $\mathbb{R}^m$, such that on a sufficiently small neighborhood $U$ of $x$

$$F^{-1}(a) \cap U = \{(y', g_a(y'))\} \cap U.$$

This means that on $U$, the level set $F^{-1}(a)$ is the graph of the function $g_a$, and therefore an embedded submanifold.

But in fact, $g_a$ depends smoothly on the value $a = F(x)$. That is, taking $U$ sufficiently small, we have

$$F^{-1}(F(y)) \cap U = \{(y', g_{F(y)}(y'))\} \cap U.$$

for all $y = (y', y'') \in U$. Then $\phi(y) = (y', g_{F(y)}(y') - y'')$ is a submanifold chart with the desired property. □

Manifolds are often described as level sets for regular values:

EXAMPLE 3.5. For $0 < r < R$, the 2-torus can be identified with the embedded submanifold $F^{-1}(r^2)$ where

$$F(x_1, x_2, x_3) = (\sqrt{x_1^2 + x_2^2} - R)^2 + x_3^2$$

is a smooth function on the complement of the $x_3$-axis, $x_1^2 + x_2^2 > 0$. One checks that indeed, $a = r^2$ is a regular value of this function.

The proposition generalizes to maps between manifolds: If $F \in C^\infty(M, N)$, a point $a \in N$ is called a *regular value of $F$* if it is a regular value "in local coordinates": That is, for all $p \in F^{-1}(a)$, and all charts $(U, \phi)$ around $p$ and $(V, \psi)$ around $a$, with $F(U) \subset V$, the Jacobian $D(\psi \circ F \circ \phi^{-1})$ at $\phi(p)$ is onto.

THEOREM 3.6. *If $a \in N$ is a regular value of $F \in C^\infty(M, N)$, then $F^{-1}(a)$ is an embedded submanifold of $M$. In fact, given a coordinate chart $(V, \psi)$ around $a$, with $\psi(a) = 0$, each $p \in F^{-1}(a)$ admits a submanifold chart $(U, \phi)$ with*

$$\psi \circ F \circ \phi^{-1}(y', y'') = y''$$

*for $y = (y', y'') \in \phi(U) \subset \mathbb{R}^n$.*

PROOF. Choose *any* coordinate chart $(U', \phi')$ around $x$. The Proposition, applied to the map $\psi \circ F \circ (\phi')^{-1}$, gives a change of coordinates with the desired properties. □

PROBLEMS 3.7.

1. Show that conversely, every submanifold is locally the graph of a function.
2. Let $S \subset \mathbb{R}^3$ be the 2-torus $(\sqrt{x_1^2 + x_2^2} - R)^2 + x_3^2 = r^2$, and $F : S \to \mathbb{R}$ the function $(x_1, x_2, x_3) \mapsto x_2$. What are the critical points for this function? What is the shape of the level sets $F^{-1}(a)$ for $a$ a singular value?
3. Let $\mathrm{Sym}_n \subset \mathrm{Mat}_n$ be the subspace of symmetric matrices, and $F : \mathrm{Mat}_n \to \mathrm{Sym}_n$ the map $A \mapsto A^t A$. Show that the identity matrix $I$ is a regular value of this map. This proves that the orthogonal group $\mathrm{O}(n)$ is an embedded submanifold of $\mathrm{Mat}_n$, of dimension $n(n-1)/2$. (In fact, by a theorem of E. Cartan, every closed subset $G \subset \mathrm{Mat}_n$, with the property that $G$ is a group under matrix multiplication, is an embedded submanifold of $\mathrm{Mat}_n$.)

## 4. Tangent spaces

For embedded submanifolds $M \subset \mathbb{R}^n$, the tangent space $T_pM$ at $p \in M$ can be defined as the set of all velocity vectors $v = \dot\gamma(0)$, where $\gamma : \mathbb{R} \to M$ is a smooth curve with $\gamma(0) = p$. Thus $T_pM$ becomes a vector subspace of $\mathbb{R}^n$. To extend this idea to general manifolds, note that the vector $v = \dot\gamma(0)$ defines a "directional derivative" $C^\infty(M) \to \mathbb{R}$:

$$v : f \mapsto \tfrac{d}{dt}|_{t=0} f(\gamma(t)).$$

We will define $T_pM$ as a set of directional derivatives.

DEFINITION 4.1. Let $M$ be a manifold, $p \in M$. The tangent space $T_pM$ is the space of all linear maps $v : C^\infty(M) \to \mathbb{R}$ of the form

$$v(f) = \tfrac{d}{dt}|_{t=0} f(\gamma(t))$$

for some smooth curve $\gamma \in C^\infty(\mathbb{R}, M)$ with $\gamma(0) = p$.

The following alternative description of $T_pM$ makes it clear that $T_pM$ is a vector subspace of the space of linear maps $C^\infty(M) \to \mathbb{R}$, of dimension $\dim T_pM = \dim M$.

PROPOSITION 4.2. *Let $(U, \phi)$ be a coordinate chart around $p$, with $\phi(p) = 0$. A linear map $v : C^\infty(M) \to \mathbb{R}$ is in $T_pM$ if and only if it has the form,*

$$v(f) = \sum_{i=1}^m a_i \frac{\partial (f \circ \phi^{-1})}{\partial x_i}|_{x=0}$$

*for some $a = (a_1, \dots, a_m) \in \mathbb{R}^m$.*

Proof. Given a linear map $v$ of this form, let $\gamma(t)$ be any smooth curve with $\phi(\gamma(t)) = ta$ for $|t|$ sufficiently small [1]. Then

$$\tfrac{d}{dt}|_{t=0} f(\gamma(t)) = \tfrac{d}{dt}|_{t=0}(f \circ \phi^{-1})(ta) = \textstyle\sum_{i=1}^m a_i \tfrac{\partial (f\circ\phi^{-1})}{\partial x_i}|_{x=0},$$

by the chain rule. Conversely, given any curve $\gamma$ with $\gamma(0) = p$, let $\tilde{\gamma} = \phi\circ\gamma$ be the corresponding curve in $\phi(U)$ (defined for small $|t|$). Then

$$\tfrac{d}{dt}|_{t=0} f(\gamma(t)) = \tfrac{\partial}{\partial t}(f \circ \phi^{-1})(\tilde{\gamma}(t)) = \textstyle\sum_{i=1}^m a_i \tfrac{\partial (f\circ\phi^{-1})}{\partial x_i}|_{x=0},$$

where $a = \frac{d\tilde{\gamma}}{dt}|_{t=0}$. □

Corollary 4.3. *If $U \subset \mathbb{R}^m$ is an open subset, the tangent space $T_pU$ is canonically identified with $\mathbb{R}^m$.*

We now describe a third definition of $T_pM$ which characterizes "directional derivatives" in a coordinate-free way, without reference to curves $\gamma$. Note first that every tangent vector $v \in T_pM$ satisfies a *product rule*,

$$v(f_1 f_2) = f_1(p)\, v(f_2) + v(f_1)\, f_2(p) \tag{2}$$

for all $f_j \in C^\infty(M)$. Indeed, in local coordinates $(U, \phi)$, this just follows from the product rule from calculus,

$$\frac{\partial}{\partial x_i}(\tilde{f}_1 \tilde{f}_2) = \tilde{f}_1(x)\frac{\partial \tilde{f}_2}{\partial x_i} + \frac{\partial \tilde{f}_1}{\partial x_i}\tilde{f}_2(x)$$

where $\tilde{f}_j = f_j \circ \phi^{-1}$. It turns out that the product rule completely characterizes tangent vectors:

Proposition 4.4. *A linear map $v : C^\infty(M) \to \mathbb{R}$ is a tangent vector if and only if it satisfies the product rule (2).*

Proof. Let $v : C^\infty(M) \to \mathbb{R}$ be a linear map satisfying the product rule (2). To show that $v \in T_pM$, we use the second definition of $T_pM$ in terms of local coordinates.

We first note that by the product rule applied to the constant function $1 = 1 \cdot 1$ we have $v(1) = 0$. Thus $v$ vanishes on constants. Next we show that $v(f_1) = v(f_2)$ if $f_1 = f_2$ near $p$. Equivalently, we show that $v(f) = 0$ if $f = 0$ near $p$. Choose $\chi \in C^\infty(M)$ with $\chi(p) = 1$, zero outside a small neighborhood of $p$ so that $f\chi = 0$. The product rule tells us that

$$0 = v(f\chi) = v(f)\chi(p) + v(\chi)f(p) = v(f).$$

Thus $v(f)$ depends only on the behavior of $f$ in an arbitrarily small neighborhood of $p$. In particular, letting $(U, \phi)$ be a coordinate chart around $p$, with $\phi(p) = 0$, we may assume that $\mathrm{supp}(f) \subset U$.[2] Consider the Taylor expansion of $\tilde{f} = f \circ \phi^{-1}$ near $x = 0$:

$$\tilde{f}(x) = \tilde{f}(0) + \sum_i x_i \frac{\partial}{\partial x_i}|_{x=0}\tilde{f} + r(x)$$

The remainder term $r$ is a smooth function that vanishes at $x = 0$ together with its first derivatives. This means that it can be written (non-uniquely) in the form $r(x) = \sum_i x_i r_i(x)$

[1] More precisely, choose any function $\chi : \mathbb{R} \to \mathbb{R}$ with $\chi(t) = t$ for $|t| < \epsilon/2$ and $\dot{\chi}(t) = 0$ for $|t| \geq \epsilon$. Choose $\epsilon$ sufficiently small, so that the ball of radius $\epsilon||a||$ is contained in $\phi(U)$. Then $\chi(t)a \in \phi(U)$ for all $t$, and $\gamma(t) = \phi^{-1}(\chi(t)a)$ is a well-defined curve with the desired properties.

[2] The support $\mathrm{supp}(f)$ of a function $f$ on $M$ is the closure of the set of all points where it is non-zero.

where $r_i$ are smooth functions that vanish at 0.[3] By the product rule, $v$ vanishes on $r \circ \phi^{-1}$ (since it is a sum of products of functions that vanish at $p$). It also vanishes on the constant $\tilde{f}(0) = f(p)$. Thus

$$v(f) = v(\tilde{f} \circ \phi^{-1}) = \sum_i a_i \frac{\partial}{\partial x_i}|_{x=0} \tilde{f}$$

with $a_i = v(x_i \circ \phi^{-1})$. (Here the coordinates $x_i$ are viewed as functions on $\mathbb{R}^n$, $x \mapsto x_i$.) □

REMARK 4.5. There is a fourth definition of $T_pM$, as follows. Let $C_p^\infty(M)$ denote the subspace of functions vanishing at $p$, and let $C_p^\infty(M)^2$ consist of finite sums $\sum_i f_i\, g_i$ where $f_i, g_i \in C_p^\infty(M)$. Since any tangent vector vanishes $v : C^\infty(M) \to \mathbb{R}$ vanishes on constants, $v$ is effectively a map $v : C_p^\infty(M) \to \mathbb{R}$. Since tangent vectors vanish on products, $v$ vanishes on the subspace $C_p^\infty(M)^2 \subset C_p^\infty(M)$. Thus $v$ descends to a linear map $C_p^\infty(M)/C_p^\infty(M)^2 \to \mathbb{R}$, i.e. an element of the dual space $(C_p^\infty(M)/C_p^\infty(M)^2)^*$. The map

$$T_pM \to (C_p^\infty(M)/C_p^\infty(M)^2)^*$$

just defined is an isomorphism, and can therefore be used as a definition of $T_pM$. This may appear very fancy on first sight, but really just says that a tangent vector is a linear functional on $C^\infty(M)$ that vanishes on constants and depends only on the first order Taylor expansion of the function at $p$.

## 5. Tangent map

DEFINITION 5.1. For any smooth map $F \in C^\infty(M, N)$ and any $p \in M$, the *tangent map* $T_pF : T_pM \to T_{F(p)}N$ is defined by the equation

$$T_pF(v)(f) = v(f \circ F)$$

It is easy to check (using any of the definitions of tangent space) that $T_pF(v)$ is indeed a tangent vector. For example, if $\gamma : \mathbb{R} \to M$ is a curve on $M$ representing $v$, we have

$$T_pF(v)(f) = v(f \circ F) = \frac{d}{dt}|_{t=0} f(F(\gamma(t)))$$

which shows that $T_pF(v)$ is the tangent vector at $F(p)$ represented by the curve $F \circ \gamma : \mathbb{R} \to N$. Similarly, it is easily verified that under composition of functions,

$$T_p(F_2 \circ F_1) = T_{F_1(p)}F_2 \circ T_pF_1.$$

In particular, if $F$ is a diffeomorphism, $T_pF$ is invertible and we have

$$T_{F(p)}F^{-1} = (T_pF)^{-1}.$$

It is instructive to work out the expression for $T_pF$ in local coordinates. We had seen that any chart $(U, \phi)$ around $p$ defines an isomorphism $T_pM \to \mathbb{R}^m$. This is the same as the isomorphism given by the tangent map,

$$T_p\phi : T_pU = T_pM \to T_{\phi(p)}\phi(U) = \mathbb{R}^m.$$

Similarly, a chart $(V, \psi)$ around $F(p)$ gives an identification $T_{F(p)}\psi : T_{F(p)}V \cong \mathbb{R}^n$. Suppose $F(U) \subset V$.

[3] Exercise: Show that if $h$ is any function on $\mathbb{R}^n$ with $h(0) = 0$, then $h$ can be written in the form $h = \sum x_i h_i$ where all $h_i$ are smooth. Show that if the first derivatives of $h$ vanish at 0, then $h_i(0) = 0$.

THEOREM 5.2. *In local charts $(U, \phi)$ and $(V, \psi)$ as above, the map*

$$T_{F(p)}\psi \circ T_pF \circ (T_p\phi)^{-1} : \mathbb{R}^m \to \mathbb{R}^n$$

*is the Jacobian of the map $\tilde{F} = \psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V)$.*

PROOF. Let $a \in \mathbb{R}^m$ represent $v \in T_pM$ in the chart $(U, \phi)$, and let $b \in \mathbb{R}^n$ represent its image under $T_pF$. We denote the coordinates on $\phi(U)$ by $x_1, \ldots, x_m$ and the coordinates on $\psi(V)$ by $y_1, \ldots, y_n$. Let $\tilde{f} = f \circ \psi^{-1} \in C^\infty(\psi(V))$. Then

$$\begin{aligned} v(f \circ F) &= \sum_{i=1}^m a_i \frac{\partial}{\partial x_i}|_{x=\phi(p)}\, f(F(\phi^{-1}(x))) \\ &= \sum_{i=1}^m a_i \frac{\partial}{\partial x_i}|_{x=\phi(p)} \tilde{f}(\tilde{F}(x)) \\ &= \sum_{i=1}^m \sum_{j=1}^n a_i \frac{\partial \tilde{F}_j}{\partial x_i}|_{x=\phi(p)} \frac{\partial \tilde{f}}{\partial y_j}|_{y=\psi(F(p))} \\ &= \sum_{j=1}^n b_j \frac{\partial \tilde{f}}{\partial y_j}|_{y=\psi(F(p))} \end{aligned}$$

where $b_j = \sum_{i=1}^m \frac{\partial \tilde{F}_j}{\partial x_i}|_{x=\phi(p)} a_i = (D\tilde{F})_{ji} a_i$. □

Thus $T_pF$ is just the Jacobian expressed in a coordinate-free way. As an immediate application, we can characterize regular values in a coordinate-free way:

DEFINITION 5.3. A point $q \in N$ is a *regular value* of $F \in C^\infty(M, N)$ if and only if the tangent map $T_pF$ is onto for all $p \in F^{-1}(q)$.

This is clearly equivalent to our earlier definition in local charts.

DEFINITION 5.4. Let $\gamma \in C^\infty(J, M)$ be a smooth curve ($J \subset \mathbb{R}$ an open interval). The tangent (or *velocity*) vector to $\gamma$ at time $t_0 \in J$ is the vector

$$\dot{\gamma}(t_0) := T_{t_0}\gamma(\frac{d}{dt}|_{t=t_0}) \in T_{\gamma(t_0)}M.$$

We will also use the notation $\frac{d\gamma}{dt}(t_0)$ to denote the velocity vector.

PROBLEMS 5.5. 1. Show that if $F \in C^\infty(M, N)$, $T_{\gamma(t)}F(\frac{d\gamma}{dt}) = \frac{d(F\circ\gamma)}{dt}$ for all $t \in J$.

2. Suppose that $S \subset M$ is an embedded submanifold, and let $\iota : S \to M$, $p \mapsto p$ be the inclusion map. Show that $\iota$ is smooth and that the tangent map $T_p\iota$ is 1-1 for all $p \in S$. Show that if $M$ is an open subset of $\mathbb{R}^m$, this becomes the identification of $T_pS$ as a subspace of $\mathbb{R}^m$, as described at the beginning of this section.

3. Suppose $F \in C^\infty(M, N)$ has $q \in N$ as a regular value. Let $S = F^{-1}(q) \hookrightarrow M$ be the level set. For $p \in S$, show that $T_pS$ is the kernel of the tangent map $T_pF$.

## 6. Tangent bundle

Let $M$ be a manifold of dimension $m$. If $M$ is an embedded submanifold of $\mathbb{R}^n$, the *tangent bundle* $TM$ is the subset of $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ given by

$$TM = \{(p, v) \in \mathbb{R}^n \times \mathbb{R}^n |\, p \in M,\, v \in T_pM\}$$

where each $T_pM$ is identified as a vector subspace of $\mathbb{R}^n$. It is not hard to see that $TM$ is, in fact, a smooth embedded submanifold of dimension $2m$. Moreover, the natural map $\pi : TM \to M$, $(p, v) \mapsto p$ is smooth, and its "*fibers*" $\pi^{-1}(p) = T_pM$ carry the structure of vector spaces.

Definition 6.1. A vector bundle of rank $k$ over a manifold $M$ is a manifold $E$, together with a smooth map $\pi : E \to M$, and a structure of a vector space on each fiber $E_p := \pi^{-1}(p)$, satisfying the following *local triviality* condition: Each point in $M$ admits an open neighborhood $U$, and a smooth map

$$\psi : \pi^{-1}(U) \to U \times \mathbb{R}^k,$$

such that $\psi$ restricts to linear isomorphisms $E_p \to \mathbb{R}^k$ for all $p \in U$.

The map $\psi : E_U \equiv \pi^{-1}(U) \to U \times \mathbb{R}^k$ is called a (local) trivialization of $E$ over $U$. In general, there need not be a trivialization over $U = M$.

Definition 6.2. A vector bundle chart for a vector bundle $\pi : E \to M$ is a chart $(U, \phi)$ for $M$, together with a chart $(\pi^{-1}(U), \hat{\phi})$ for $E_U = \pi^{-1}(U)$, such that $\hat{\phi} : \pi^{-1}(U) \to \mathbb{R}^m \times \mathbb{R}^k$ restricts to linear isomorphisms from each fiber $E_p$ onto $\{\phi(p)\} \times \mathbb{R}^k$.

Every vector bundle chart defines a local trivialization. Conversely, if $\psi : E|_U \to U \times \mathbb{R}^k$ is a trivialization of $E_U$, where $U$ is the domain of a chart $(U, \phi)$, one obtains a vector bundle chart $(\pi^{-1}(U), \hat{\phi})$ for $E$.

Example 6.3. (Vector bundles over the Grassmannnian) For any $p \in \mathrm{Gr}(k, n)$, let $E_p \subset \mathbb{R}^n$ be the $k$-plane it represents. Then $E = \cup_{p \in \mathrm{Gr}(k,n)} E_p$ is a vector bundle over $\mathrm{Gr}(k, n)$, called the tautological vector bundle. Recall the definition of charts $\phi_I : U_I \to L(\mathbb{R}^I, \mathbb{R}^{I'})$ for the Grassmannian, where any $p = \{E\} = U_I$ is identified with the linear map $A$ having $E$ as its graph. Let

$$\hat{\phi}_I : \pi^{-1}(U_I) \to L(\mathbb{R}^I, \mathbb{R}^{I'}) \times \mathbb{R}^I$$

be the map $\hat{\phi}_I(v) = (\phi(\pi(v)), \pi_I(v))$ where $\pi_I : \mathbb{R}^n \to \mathbb{R}^I$ is orthogonal projection. The $\hat{\phi}_I$ serve as bundle charts for the tautological vector bundle. There is another natural vector bundle $E'$ over $\mathrm{Gr}(k, n)$, with fiber $E'_p := E_p^\perp$ the orthogonal complement of $E_p$. A special case is $k = 1$, where $\mathrm{Gr}(k, n) = \mathbb{R}P(n-1)$. In this case $E$ is called the tautological line bundle, and $E'$ the hyperplane bundle.

At this stage, we are mainly interested in tangent bundles of manifolds.

Theorem 6.4. *For any manifold $M$, the disjoint union $TM = \cup_{p \in M} T_pM$ carries the structure of a vector bundle over $M$, where $\pi$ takes $v \in T_pM$ to the base point $p$.*

Proof. Recall that any chart $(U, \phi)$ for $M$ gives identifications $T_p\phi : T_pM \to \mathbb{R}^m$ for all $p \in U$. Taking all these maps together, we obtain a bijection,

$$T\phi : \pi^{-1}(U) \to U \times \mathbb{R}^m.$$

We take the collection of $(\pi^{-1}(U), T\phi)$ as vector bundle charts for $TM$. We need to check that the transition maps are smooth. If $(V, \psi)$ is another coordinate chart with $U \cap V \neq \emptyset$, the transition map for $\pi^{-1}(U \cap V)$ is given by,

$$T\psi \circ (T\phi)^{-1} : (U \cap V) \times \mathbb{R}^m \to (U \cap V) \times \mathbb{R}^m.$$

But $T_p\psi \circ (T_p\phi)^{-1} = T_{\phi(p)}(\psi \circ \phi^{-1})$ is just the Jacobian for the change of coordinates $\psi \circ \phi^{-1}$, and as such depends smoothly on $x = \phi(p)$. □

Definition 6.5. A (smooth) section of a vector bundle $\pi : E \to M$ is a smooth map $\sigma : M \to E$ with the property $\pi \circ \sigma = \mathrm{id}_M$. The space of sections of $E$ is denoted $\Gamma^\infty(M, E)$.

Thus, a section is a family of vectors $\sigma_p \in E_p$ depending smoothly on $p$.

Examples 6.6. (a) Every vector bundle has a distinguished section, the *zero section* $p \mapsto \sigma_p = 0$.
(b) A section of the *trivial bundle* $M \times \mathbb{R}^k$ is the same thing as a smooth function from $M$ to $R^k$. In particular, if $\psi : E_U \to U \times \mathbb{R}^k$ is a local trivialization of a vector bundle $E$, the section $\sigma$ (restricted to $U$) becomes a smooth function $\psi \circ \sigma|_U : U \to \mathbb{R}^k$.
(c) Let $\pi : E \to M$ be a rank $k$ vector bundle. A *frame* for $E$ over $U \subset M$ is a collection of sections $\sigma_1, \ldots, \sigma_k$ of $E_U$, such that $(\sigma_j)_p$ are linearly independent at each point $p \in U$. Any frame over $U$ defines a local trivialization $\psi : E_U \to U \times \mathbb{R}^k$, given in terms of its inverse map $\psi^{-1}(p, a) = \sum_j a_j(\sigma_j)_p$. Conversely, each local trivialization gives rise to a frame.

The space $\Gamma^\infty(M, E)$ is a vector space under pointwise addition: $(\sigma_1 + \sigma_2)_p = (\sigma_1)_p + (\sigma_2)_p$. Moreover, it is a $C^\infty(M)$-module under multiplication[4]: $(f\sigma)_p = f_p\sigma_p$.

Definition 6.7. A section of the tangent bundle $TM$ is called a vector field on $M$. The space of vector fields is denoted

$$\mathfrak{X}(M) = \Gamma^\infty(M, TM).$$

Thus, a vector field $X \in \mathfrak{X}(M)$ is a family of tangent vectors $X_p \in T_pM$ depending smoothly on the base point.

In the next section, we will discuss the space of vector fields in more detail.

Problems 6.8. 1. Let $S \subset M$ be an embedded submanifold. Show that for any vector bundle $\pi : E \to M$, the restriction $E|_S \to S$ is a vector bundle over $S$. In particular, $TM|_S$ is defined; its sections are called "vector fields along $S$". The bundle $TM|_S$ contains the tangent bundle $TS$ as a sub-bundle: For all $p \in S$, $T_pS$ is a vector subspace of $T_pM$. The *normal bundle* of $S$ in $M$ is defined as a "quotient bundle" $\nu_S = TM|_S/TS$ with fibers,

$$(\nu_S)_p = T_pM/T_pS$$

Show that this is again a vector bundle.

[4]Here and from now on, we will often write $f_p$ or $f|_p$ for the value $f(p)$.

2. Let $F : M \to N$ be a smooth map, and $\pi : E \to N$ a vector bundle. Show that
$$F^*E := \cup_{p\in M} E_{F(p)}$$
is a vector bundle over $M$. It is called the *pull-back bundle*. Sections of $F^*(TN)$ are called vector fields along (the map) $F$. For instance, if $\gamma : J \to M$ is a smooth curve, the set of velocity vectors $\dot\gamma(t)$ becomes a vector field along $\gamma$.
3. Let $E, E'$ be two vector bundles over $M$. Show that
$$E \oplus E' := \cup_{p\in M} E_p \oplus E'_p$$
is again a vector bundle over $M$. It is called the direct sum (or *Whitney sum*) of $E$ and $E'$. For instance, the direct sum of the two natural bundles $E, E'$ over the Grassmannian has fibers $E_p \oplus E'_p = \mathbb{R}^n$, hence $E \oplus E'$ is the trivial bundle $\mathrm{Gr}(k, n) \times \mathbb{R}^n$.
4. Show that for any vector bundle $E \to M$,
$$E^* = \cup_{p\in M} E_p^*$$
(where $E_p^* = L(E_p, \mathbb{R})$ is the dual space to $E_p$) is again a vector bundle. It is called the dual bundle to $E$. In particular, one defines $T^*M := (TM)^*$, called the cotangent bundle. The sections of $T^*M$ are called covector fields or "1-forms".

## 7. Vector fields as derivations

Let $X \in \mathfrak{X}(M)$ be a vector field on $M$. Each $X_p \in T_pM$ defines a linear map $X_p : C^\infty(M) \to \mathbb{R}$. Letting $p$ vary, this gives a linear map
$$X : C^\infty(M) \to C^\infty(M),\ (X(f))_p = X_p(f).$$
Note that the right hand side really does define a smooth function on $M$. Indeed, this follows from the expression in local coordinates $(U, \phi)$. Let $a \in C^\infty(\phi(U), \mathbb{R}^k)$ be the expression of $X$ in the local trivialization, that is, $(T\phi)(X_p) = (\phi(p), a(\phi(p)))$. Thus
$$a(\phi(p)) = (a_1(\phi(p)), \ldots, a_m(\phi(p)))$$
are simply the components of $X_p$ in the coordinate chart $(U, \phi)$:
$$X_p(f) = \sum_{i=1}^m a_i(\phi(p)) \frac{\partial}{\partial x_i}|_{\phi(p)}(f \circ \phi^{-1}).$$
for $p \in U$. The formula shows that
$$X(f)|_U \circ \phi^{-1} = \sum_{i=1}^m a_i \frac{\partial}{\partial x_i}(f \circ \phi^{-1}).$$
That is, in local coordinates $X$ is represented by the vector field
$$\sum_{i=1}^m a_i \frac{\partial}{\partial x_i} :\ C^\infty(\phi(U)) \to C^\infty(\phi(U)).$$

THEOREM 7.1. *A linear map $X : C^\infty(M) \to C^\infty(M)$ is a vector field if and only if it is a derivation of the algebra $C^\infty(M)$: That is,*
$$X(f_1 f_2) = f_2\, X(f_1) + f_1\, X(f_2)$$
*for all $f_1, f_2 \in C^\infty(M)$.*

Proof. For all $p \in M$, $X$ defines a tangent vector $X_p$ by $X_p(f) = X(f)_p$. We have to show that $p \mapsto X_p$ defines a smooth section of $TM$. Choosing local coordinates $(U, \phi)$ around $p$, taking $p$ to $x = \phi(p)$, the tangent vector $X(f)_p$ is represented by $a(x) = (a_1(x), \dots, a_m(x)) \in \mathbb{R}^m$. That is,

$$X(f)_p(f) = \sum_{j=1}^m a_j(x) \frac{\partial}{\partial x_i}(f \circ \phi^{-1}).$$

Taking for $f$ any function $f_j$ with $f_j \circ \phi^{-1}(x) = x_j$ on some open neighborhood $V \subset \phi(U)$ of the given point $\phi(p)$, we see that

$$a_j = X(f_j) \circ \phi^{-1}$$

on $V$. Since $X(f_j) \in C^\infty(M)$, it follows that $a_j$ is smooth. This proves that $p \mapsto X_p$ is a smooth section of $TM$ over $U$. □

If $X, Y$ are vector fields (viewed as linear maps $C^\infty(M) \to C^\infty(M)$), the composition $X \circ Y$ is *not* a vector field. However, the *Lie bracket* (commutator)

$$[X, Y] := X \circ Y - Y \circ X$$

is a vector field. Indeed, it is easily checked that the right hand side defines a derivation. Alternatively, the calculation can be carried out in local coordinates $(U, \phi)$: One finds that if $X_U$ is represented by $\sum_{i=1}^m a_i \frac{\partial}{\partial x_i}$ and $Y|_U$ by $\sum_{i=1}^m b_i \frac{\partial}{\partial x_i}$, then $[X, Y]|_U$ is represented by

$$\sum_{i=1}^m \sum_{j=1}^m (a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j}) \frac{\partial}{\partial x_i}.$$

Let $F \in C^\infty(M, N)$ be a smooth map. Then the collection of tangent maps $T_pF : T_pM \to T_{F(p)}N$ defines a map $TF : TM \to TN$ which is easily seen to be smooth. The map $TF$ is an example of a vector bundle map: It takes fibers to fibers, and the restriction to each fiber is a linear map. For instance, local trivializations $\psi : E|_U \to U \times \mathbb{R}^k$ are vector bundle maps.

Definition 7.2. Let $F \in C^\infty(M, N)$. Two vector fields $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are called $F$-related if for all $p \in M$, $T_pF(X_p) = Y_{F(p)}$. One writes $X \sim_F Y$.

For example, if $S \subset M$ is an embedded submanifold and $\iota : S \to M$ is the inclusion, vector fields $X$ on $S$ and $Y$ on $M$ are $\iota$-related if and only if $Y$ is tangent to $S$, and $X$ is its restriction.

Theorem 7.3. *a) One has $X \sim_F Y$ if and only if for all $f \in C^\infty(N)$, $X(f \circ F) = Y(f)$.*
*b) If $X_1 \sim_F Y_1$ and $X_2 \sim_F Y_2$ then $[X_1, X_2] \sim_F [Y_1, Y_2]$.*

Proof. At any $p \in M$, the condition $X(f \circ F) = F \circ Y(f)$ says that

$$(T_pF(X_p))(f) = Y(f)_{F(p)} = Y_{F(p)}(f).$$

This proves (a). Part (b) follows from (a):

$$\begin{aligned}
[X_1, X_2](f \circ F) &= X_1(X_2(f \circ F)) - X_2(X_1(f \circ F)) \\
&= X_1(Y_2(f) \circ F) - X_2(Y_1(f) \circ F) \\
&= Y_1(Y_2(f)) \circ F - Y_2(Y_1(f)) \circ F \\
&= [Y_1, Y_2](f) \circ F.
\end{aligned}$$

□

Part (b) shows, for instance, that if two vector fields are tangent to a submanifold $S \subset M$ then their bracket is again tangent to $S$. (Alternatively, one can see this in coordinates, using submanifold charts for $S$.)

PROBLEMS 7.4. 1. Given an example of vector fields $X, Y \in \mathfrak{X}(\mathbb{R}^3)$ such that $X, Y, [X, Y]$ are linearly independent at any point $p \in \mathbb{R}^3$. Thus, there is no 2-dimensional submanifold $S$ with the property that $X, Y$ are tangent to $S$ everywhere.

2. For any $n$, give an example of vector field $X, Y$ on $\mathbb{R}^n$ such that $X, Y$ together with iterated Lie brackets $[X, Y], [[X, Y], Y], \ldots$ span $T_p\mathbb{R}^n = \mathbb{R}^n$ everywhere.

## 8. Flows of vector fields

Suppose $X$ is a vector field on a manifold $M$. A smooth curve $\gamma \in C^\infty(J, M)$, where $J$ is an open neighborhood of $0 \in \mathbb{R}$, is called a *solution curve* to $X$ if for all $t \in J$,

$$\dot{\gamma}(t) = X_{\gamma(t)}.$$

In local coordinates $(U, \phi)$ around a given point $p = \gamma(t_0)$, write

$$\phi \circ \gamma(t) = x(t) = (x_1(t), \ldots, x_n(t))$$

(defined for $t$ sufficiently close to $t_0$). Then

$$\dot{\gamma}(f) = \frac{d}{dt} f(\gamma(t)) = \frac{d}{dt}(f \circ \phi^{-1})(x(t)) = \sum_i \frac{dx_i}{dt} \frac{\partial}{\partial x_i}(f \circ \phi^{-1})\big|_{x(t)}.$$

Let $\sum_i a_i \frac{\partial}{\partial x_i}$ represent $X$ in the chart, that is

$$X_{\gamma(t)}(f) = \sum_i a_i(x(t)) \frac{\partial}{\partial x_i}(f \circ \phi^{-1})\big|_{x(t)}.$$

Hence, the equation for a solution curve corresponds to the following equation in local coordinates:

$$\dot{x}_i = a_i(x(t)),$$

for $i = 1, \ldots, n$. This is a first order system of ordinary differential equations (ODE's). One of the main results from the theory of ODE's reads:

THEOREM 8.1 (Existence and uniqueness theorem for ODE's). *Let $V \subset \mathbb{R}^m$ be an open subset, and $a \in C^\infty(V, \mathbb{R}^m)$. For any given $x_0 \in V$, there exists an open interval $J \subset \mathbb{R}$ around $0$, and a solution $x : J \to V$ of the ODE*

$$\frac{dx_i}{dt} = a_i(x(t)).$$

*with initial condition $x(0) = x_0$. In fact, there is a unique maximal solution of this initial value problem, defined on some interval $J_{x_0}$, such that any other solution is obtained by restriction to some subinterval $\mathcal{J} \subset \mathcal{J}_{x_0}$.*

The solution depends smoothly on initial conditions, in the following sense:

THEOREM 8.2 (Dependence on initial conditions for ODE's). *Let $a \in C^\infty(V, \mathbb{R}^m)$ as above. For any $x_0 \in V$, let $\Phi(t, x_0) := \gamma_{x_0}(t) : \mathcal{J}_{x_0} \to V$ be the maximal solution with initial value $\gamma_{x_0}(0) = x_0$. Let*

$$\mathcal{J} = \bigcup_{x_0 \in V} \mathcal{J}_{x_0} \times \{x_0\} \subset \mathbb{R} \times V.$$

*Then $\mathcal{J}$ is an open neighborhood of $\{0\} \times V$, and the map $\Phi : \mathcal{J} \to V$ is smooth.*

EXAMPLE 8.3. If $V = (0, 1) \subset \mathbb{R}$ and $a(x) = 1$, the solution curves to $\dot{x} = a(x(t)) = 1$ with initial condition $x_0 \in V$ are $x(t) = x_0 + t$, defined for $-x_0 < t < 1 - x_0$. Thus

$$\mathcal{J} = \{(t, x) |\, 0 < t + x < 1\}, \quad \Phi(t, x) = t + x$$

in this case. One can construct a similar example with $V = \mathbb{R}$, where solution curves escape to infinity in finite time: For instance, $a(x) = x^2$ has solution curves, $x(t) = -\frac{1}{t-c}$, these escape to infinity for $t \to c$. Similarly, $a(x) = 1 + x^2$ has solution curves $x(t) = \tan(t - c)$, these reach infinity for $t \to c \pm \frac{\pi}{2}$.

If $a = (a_1, \ldots, a_m) : \phi(U) \to \mathbb{R}^m$ corresponds to $X$ in a local chart $(U, \phi)$, then any solution curve $x : J \to \phi(U)$ for $a$ defines a solution curve $\gamma(t) = \phi^{-1}(x(t))$ for $X$. The existence and uniqueness theorem for ODE's extends to manifolds, as follows:

THEOREM 8.4. *Let $X \in \mathfrak{X}(M)$ be a vector field on a manifold $M$. For any given $p \in M$, there exists a solution curve $\gamma : J \to M$ of with initial condition $\gamma(0) = p$. Any two solutions to the same initial value problem agree on their common domain of definition.*

PROOF. Existence and uniqueness of solutions for small times $t$ follows from the existence and uniqueness theorem for ODE's, by considering the vector field in local charts. To prove uniqueness even for large times $t$, let $\gamma_t : J_1 \to M$ and $\gamma_2 : J_2 \to M$ be two solutions to the IVP. We have to show that $\gamma_1 = \gamma_2$ on $J_1 \cap J_2$. Suppose not. Let $b > 0$ be the infimum of all $t \in J_1 \cap J_2$ with $\gamma_1(t) \neq \gamma_2(t)$. If $\gamma_1(b) = \gamma_2(b)$, the uniqueness part for solutions of ODE's, in a chart around $\gamma_j(b)$, would show that the $\gamma_j(t)$ coincide for $|t - b|$ sufficiently close to $b$. This contradiction shows that $\gamma_1(b) \neq \gamma_2(b)$. But then we can choose disjoint open neighborhoods $U_j$ of $\gamma_j(b)$. For $|t-b|$ sufficiently small, $\gamma_j(t) \in U_j$. In particular, $\gamma_1(t) \neq \gamma_2(t)$ for small $|t-b|$, again in contradiction to the definition of $b$. $\square$

Note that the uniqueness part uses the Hausdorff property in the definition of manifolds. Indeed, the uniqueness part may fail for non-Hausdorff manifolds.

EXAMPLE 8.5. A counter-example is the non-Hausdorff manifold $Y = \mathbb{R} \times \{1\} \cup \mathbb{R} \times \{-1\} / \sim$, where $\sim$ glues two copies of the real line along the strictly negative real axis. Let $U_\pm$ denote the charts obtained as images of $\mathbb{R} \times \{\pm 1\}$. Let $X$ be the vector field on $Y$, given by $\frac{\partial}{\partial x}$ in both charts. It is well-defined, since the transition map is just the identity map. Then $\gamma_+(t) = \pi(t, 1)$ and $\gamma_-(t) = \pi(t, -1)$ are both solution curves, and they agree for negative $t$ but not for positive $t$.

THEOREM 8.6. *Let $X \in \mathfrak{X}(M)$ be a vector field on a manifold $M$. For each $p \in M$, let $\gamma_p : \mathcal{J}_p \to M$ be the maximal solution curve with initial value $\gamma_p(0) = p$. Let $\mathcal{J} = \bigcup_{p \in M} \{p\} \times \mathcal{J}_p$, and let*

$$\Phi : \mathcal{J} \to M, \quad \Phi(t, p) \equiv \Phi_t(p) := \gamma_p(t).$$

*Then $\mathcal{J}$ is an open neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$, and the map $\Phi$ is smooth. If $(t_1, \Phi_{t_2}(p))$, $(t_2, p) \in \mathcal{J}$ then also $(t_1 + t_2, p) \in \mathcal{J}$, and one has*

$$\Phi_{t_1}(\Phi_{t_2}(p)) = \Phi_{t_1+t_2}(p).$$

*The map $\Phi$ is called the flow of the vector field $X$.*

Proof. Define $\mathcal{J} = \cup_{p \in M}\{p\} \times \mathcal{J}_p$ where $\mathcal{J}_p$ is the largest interval around 0 for which there is a solution curve $\gamma_p(t)$ with initial value $\gamma_p(0) = p$. Let $\Phi(t,p) := \gamma_p(t)$. We first establish the property $\Phi_{t_1}(\Phi_{t_2}(p)) = \Phi_{t_1+t_2}(p)$. Given $(t_2, p) \in \mathcal{J}^X$, consider the two curves

$$\gamma(t) = \Phi_t(\Phi_{t_2}(p)), \quad \lambda(t) = \Phi_{t+t_2}(p).$$

By definition of $\Phi$, the curve $\gamma$ is a solution curve with initial value $\gamma(0) = \Phi_{t_2}(p)$, defined for the set of all $t$ with $(t, \Phi_{t_2}(p)) \in \mathcal{J}^X$.

We claim that $\lambda$ is also a solution curve, hence coincides with $\gamma$ by uniqueness of solution curves. We calculate

$$\begin{aligned} \frac{d}{dt}\lambda(t) &= \frac{d}{dt}\Phi_{t+t_2}(p) \\ &= \frac{d}{du}|_{u=t+t_2}\Phi_u(p) \\ &= X_{\Phi_u(p)}|_{u=t+t_2} = X_{\lambda(t)} \end{aligned}$$

It remains to show that $\mathcal{J}$ is open and $\Phi$ is smooth. We will use the property $\Phi_{t_1+t_2} = \Phi_{t_2} \circ \Phi_{t_1}$ of the flow to write the flow for large times $t$ as a composition of flows for small times, where we can use the result for ODE's in local charts.

Let $(t,p) \in \mathcal{J}$, say $t > 0$. Since the interval $[0,t]$ is compact, we can choose $t_i > 0$ with $t = t_1 + t_2 + \ldots + t_N$ such that the curve $\Phi_s(p)$ stays in a fixed coordinate chart $V_0$ for $0 \le s \le t_1$, the curve $\Phi_s(\Phi_{t_1}(p))$ stays in a fixed coordinate chart $V_1$ for $0 \le s \le t_2$, and so on. Also, let $\epsilon > 0$ be sufficiently small, such that $\Phi_s(\Phi_t(p))$ is defined and stays in $V_N := V_{N-1}$ for $-\epsilon \le s \le \epsilon$.

Inductively define $p_0, \ldots, p_N$ by let $p_{k+1} = \Phi_{t_{k+1}}(p_k)$ where $p_0 = p$. Thus $p_N = \Phi_t(p)$. Choose open neighborhoods $U_k$ of $p_k$, with the property that

$$\begin{aligned} &\overline{\Phi_s(U_0)} \subset V_0 \quad \text{for} \quad 0 \le s \le t_1 \\ &\overline{\Phi_s(U_1)} \subset V_1 \quad \text{for} \quad 0 \le s \le t_2 \\ &\qquad \cdots \end{aligned}$$

and $\Phi_s(U_N) \subset V_N$ for $-\epsilon < s < \epsilon$. Let $U$ be the set of all points $q \in M$ such that

$$q \in U_0, \ \Phi_{t_1}(q) \in U_1, \ \Phi_{t_1+t_2}(q) \in U_2, \ldots, \Phi_t(q) \in U_N.$$

Then $U$ is an open neighborhood of $p$. The composition $\Phi_{s+t} = \Phi_s \circ \Phi_{t_N} \circ \cdots \circ \Phi_{t_1}$ is well-defined on $U$, for all $-\epsilon < s < \epsilon$. Thus

$$(t - \epsilon, t + \epsilon) \times U \subset \mathcal{J}.$$

The map $\Phi$, restricted to this set, is smooth, since it is written as a composition of smooth maps:

$$\Phi(t + s, \cdot) = \Phi(s, \Phi_{t_N} \circ \cdots \Phi_{t_1}(\cdot)).$$

$\square$

Let $X$ be a vector field, and $\mathcal{J} = \mathcal{J}^X$ be the domain of definition for the flow $\Phi = \Phi^X$.

DEFINITION 8.7. A vector field $X \in \mathfrak{X}(M)$ is called complete if $\mathcal{J}^X = M \times \mathbb{R}$.

Thus $X$ is complete if and only if all solution curves exist for all time. Above, we had seen some examples of incomplete vector fields on $M = \mathbb{R}$. In these examples, the vector field increases "too fast towards infinity". Conversely, we expect that vector fields $X$ are complete if they vanish outside a compact set. This is indeed the case. The support $\mathrm{supp}(X)$ is defined to be the smallest closed subset outside of which $X$ is zero. That is,

$$\mathrm{supp}(X) = \overline{\{p \in M | X_p \neq 0\}}.$$

PROPOSITION 8.8. *Every vector field of compact support is complete. In particular, this is the case if $M$ is compact.*

PROOF. By compactness, there exists $\epsilon > 0$ such that the flow for any point $p$ exists for times $|t| \leq \epsilon$. But this implies that any integral curve can be extended indefinitely. □

THEOREM 8.9. *If $X$ is a complete vector field, the flow $\Phi_t$ defines a 1-parameter group of diffeomorphisms. That is, each $\Phi_t$ is a diffeomorphism and*

$$\Phi_0 = \mathrm{id}_M, \quad \Phi_{t_1} \circ \Phi_{t_2} = \Phi_{t_1+t_2}.$$

*Conversely, if $\Phi_t$ is a 1-parameter group of diffeomorphisms such that the map $(t,p) \mapsto \Phi_t(p)$ is smooth, the equation*

$$X_p(f) = \frac{d}{dt}|_{t=0} f(\Phi_t(p))$$

*defines a smooth vector field on $M$, with flow $\Phi_t$.*

PROOF. It remains to show the second statement. Clearly, $X_p$ is a tangent vector at $p \in M$. Using local coordinates, one can show that $X_p$ depends smoothly on $p$, hence it defines a vector field. Given $p \in M$ we have to show that $\gamma(t) = \Phi_t(p)$ is an integral curve of $X$. Indeed,

$$\frac{d}{dt}\Phi_t(p) = \frac{d}{ds}|_{s=0}\Phi_{t+s}(p) = \frac{d}{ds}|_{s=0}\Phi_s(\Phi_t(p)) = X_{\Phi_t(p)}.$$

□

By a similar argument, one establishes the identity

$$\frac{d}{dt}\Phi_t^*(f) = \Phi_t^* \frac{d}{ds}|_{s=0}\Phi_s^*(f) = \Phi_t^* X(f)$$

which we will use later on. In fact, this identity may be viewed as a definition of the flow.

EXAMPLE 8.10. Let $X$ be a complete vector field, with flow $\Phi_t$. For each $t \in \mathbb{R}$, the tangent map $T\Phi_t : TM \to TM$ has the flow property,

$$T\Phi_{t_1} \circ T\Phi_{t_2} = T(\Phi_{t_1} \circ \Phi_{t_2}) = T(\Phi_{t_1+t_2}),$$

and the map $\mathbb{R} \times TM \to TM, (t,v) \mapsto \Phi_t(v)$ is smooth (since it is just the restriction of the map $T\Phi : T(\mathbb{R} \times M) \to TM$ to the submanifold $\mathbb{R} \times TM$). Hence, $T\Phi_t$ is a flow on $TM$, and therefore corresponds to a vector field $\widehat{X} \in \mathfrak{X}(TM)$. This is called the *tangent lift* of $X$.

EXAMPLE 8.11. Given $A \in \mathrm{Mat}_m(\mathbb{R})$ let $\Phi_t : \mathbb{R}^m \to \mathbb{R}^m$ be multiplication by the matrix $e^{tA} = \sum_j \frac{t^j}{j!} A^j$ (exponential map of matrices). Since $e^{(t_1+t_2)A} = e^{t_1 A} e^{t_2 A}$, and since $(t,x) \mapsto e^{tA}x$ is a smooth map, $\Phi_t$ defines a flow. What is the corresponding vector field $X$? For any function $f \in C^\infty(\mathbb{R}^m)$ we calculate,

$$\begin{aligned} X_x(f) &= \frac{d}{dt}\Big|_{t=0} f(e^{tA}x) \\ &= \sum_j \frac{\partial f}{\partial x_j}(Ax)_j \\ &= \sum_{jk} A_{jk} x_k \frac{\partial f}{\partial x_j} \end{aligned}$$

showing that $X = \sum_{jk} A_{jk} x_k \frac{\partial}{\partial x_j}$.

PROBLEMS 8.12. 1. Let $X \in \mathfrak{X}(N)$, $Y \in \mathfrak{X}(M)$ be vector fields and $F \in C^\infty(N, M)$ a smooth map. Show that $X \sim_F Y$ if and only if it intertwines the flows $\Phi_t^X, \Phi_t^Y$: That is,

$$F \circ \Phi_t^X = \Phi_t^Y \circ F.$$

2. Let $X$ be a vector field on $U \subset M$, given in local coordinates by $\sum_i a_i \frac{\partial}{\partial x_i}$. Let $(x_1, \ldots, x_m, v_1, \ldots, v_m)$ be the corresponding coordinates on $TU \subset TM$. Show that the tangent lift $\hat{X}$ is given by

$$\sum_i a_i \frac{\partial}{\partial x_i} + \sum_{ij} \frac{\partial a_i}{\partial x_j} v_j \frac{\partial}{\partial v_i}$$

3. Show that for any vector field $X \in \mathfrak{X}(M)$ and any $x \in M$ with $X_x \neq 0$, there exists a local chart around $x$ in which $X$ is given by the constant vector field $\frac{\partial}{\partial x^1}$. Hint: Show that if $S$ is an embedded codimension 1 submanifold, with $x \in S$ and $X_x \notin T_x S$, the map $U \times (-\epsilon, \epsilon) \to M$ is a diffeomorphisms onto its image, for some open neighborhood $U$ of $x$ in $S$. Use the time parameter $t$ and a chart around $x \in U$ to define a chart near $x$.

## 9. Geometric interpretation of the Lie bracket

If $f \in C^\infty(N)$ and $F \in C^\infty(M, N)$ we define the pull-back $F^*(f) = f \circ F \in C^\infty(M)$. Thus pull-back is a linear map,

$$F^* : C^\infty(N) \to C^\infty(M).$$

Using pull-backs, the definition of a tangent map reads

$$T_p F(v) = v \circ F^* : C^\infty(N) \to \mathbb{R}.$$

For instance, the definition of $F$-related vector fields $X \sim_F Y$ can be re-phrased as, $X \circ F^* = F^* \circ Y$. For any vector field $X \in \mathfrak{X}(N)$ and any *diffeomorphism* $F \in C^\infty(M, N)$, we define $F^*X \in \mathfrak{X}(M)$ by

$$F^*X(F^*f) = F^*(X(f)).$$

That is,

$$F^*X = F^* \circ X \circ (F^*)^{-1}.$$

LEMMA 9.1. *If $X, Y$ are vector fields on $N$, $F^*[X,Y] = [F^*X, F^*Y]$.*

Any complete vector field $X \in \mathfrak{X}(M)$ with flow $\Phi_t$ gives rise to a family of maps $\Phi_t^* : \mathfrak{X}(M) \to \mathfrak{X}(M)$. One defines the *Lie derivative* $L_X$ of a vector field $Y \in \mathfrak{X}(M)$ by

$$L_X(Y) = \frac{d}{dt}\Big|_{t=0} \Phi_t^* Y \in \mathfrak{X}(M).$$

The definition of Lie derivative also works for incomplete vector fields, since the definition only involves derivatives at $t = 0$.

THEOREM 9.2. *For any $X, Y \in \mathfrak{X}(M)$, the Lie derivative $L_X Y$ is just the Lie bracket $[X, Y]$. One has the identity*

$$[L_X, L_Y] = L_{[X,Y]}.$$

PROOF. Let $\Phi_t = \Phi_t^X$ be the flow of $X$. For $f \in C^\infty(M)$ we calculate,

$$\begin{aligned} (L_X Y)(f) &= \frac{d}{dt}|_{t=0}(\Phi_t^* Y)(f) \\ &= \frac{d}{dt}|_{t=0}\Phi_t^*(Y(\Phi_{-t}^*(f))) \\ &= \frac{d}{dt}|_{t=0}(\Phi_t^*(Y(f)) - Y(\Phi_t^*(f)) \\ &= X(Y(f)) - Y(X(f)) \\ &= [X, Y](f). \end{aligned}$$

The identity $[L_X, L_Y] = L_{[X,Y]}$ just rephrases the Jacobi identity for the Lie bracket. □

Again, let $X$ be a complete vector field with flow $\Phi$. Let us work out the Taylor expansion of the map $\Phi_t^*$ at $t = 0$. That is, for any function $f \in C^\infty(M)$, consider the Taylor expansion (pointwise, i.e. at any point of $M$) of the function

$$\Phi_t^* f = f \circ \Phi_t \in C^\infty(M)$$

around $t = 0$. We have,

$$\frac{d}{dt}\Phi_t^* f = \frac{d}{ds}|_{s=0}\Phi_{t+s}^* f = \frac{d}{ds}|_{s=0}\Phi_t^*\Phi_s^* f = \Phi_t^* X(f).$$

By induction, this shows

$$\frac{d^k}{dt^k}\Phi_t^* f = \Phi_t^* X^k(f),$$

where $X^k = X \circ \cdots \circ X$ (k times). Hence, the Taylor expansion reads

$$\Phi_t^* f = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k(f).$$

One often writes the right hand side as $\exp(tX)(f)$. Suppose now that $Y$ is another vector field, with flow $\Psi_s$. In general, $\Phi_t \circ \Psi_s$ need not equal $\Psi_s \circ \Phi_t$, that is, the flows need not commute. Let us compare the Taylor expansions of $\Phi_t^*\Psi_s^* f$ and $\Psi_s^*\Phi_t^* f$. We have, in second order,

$$\begin{aligned}\Phi_t^*\Psi_s^* f &= \Phi_t^*(f + sY(f) + \frac{s^2}{2}Y^2(f) + \cdots) \\ &= f + sY(f) + \frac{s^2}{2}Y^2(f) + tX(f) + stX(Y(f)) + \frac{t^2}{2}X^2(f) + \cdots\end{aligned}$$

where the dots indicate cubic or higher terms in the expansion. Interchanging the roles of $X, Y$, and subtracting, we find,

$$(\Phi_t^*\Psi_s^* - \Psi_s^*\Phi_t^*)f = st[X,Y](f) + \ldots$$

This shows that $[X,Y]$ measures the extent to which the flows fail to commute (up to second order to the Taylor expansion). In fact,

THEOREM 9.3. *Let $X, Y$ be complete vector fields. Then $[X,Y] = 0$ if and only if the flows of $X$ and $Y$ commute.*

PROOF. Let $\Phi_t$ be the flow of $X$ and $\Psi_s$ the flow of $Y$. Suppose $[X,Y] = 0$. Then

$$\frac{d}{dt}(\Phi_t)^*Y = (\Phi_t)^* L_X Y = (\Phi_t)^*[X,Y] = 0$$

for all $t$. Integrating from 0 to $t$, this shows $(\Phi_t)^*Y = Y$ for all $t$, which means that $Y$ is $\Phi_t$-related to itself. It follows that $\Phi_t$ takes the flow $\Psi_s$ of $Y$ to itself, which is just the desired equation $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$. Conversely, by differentiating the equation $\Phi_t \circ \Psi_s = \Psi_s \circ \Phi_t$ with respect to $s, t$, we find that $[X,Y] = 0$. □

## 10. Lie groups and Lie algebras

### 10.1. Definition of Lie groups.

DEFINITION 10.1. A *Lie group* is a group $G$, equipped with a manifold structure, such that group multiplication $g_1, g_2 \mapsto g_1g_2$ is a smooth map $G \times G \to G$.

Examples of Lie groups include: The general linear group $\mathrm{GL}(n,\mathbb{R})$ (invertible matrices in $\mathrm{Mat}_n(\mathbb{R})$), the special linear group $\mathrm{SL}(n,\mathbb{R})$ (those with determinant 1), the orthogonal group $\mathrm{O}(n)$ and special orthogonal group $\mathrm{SO}(n)$, the unitary group $\mathrm{U}(n)$ and the special unitary group $\mathrm{SU}(n)$ and the complex general linear or special linear groups $\mathrm{GL}(n,\mathbb{C})$ and $\mathrm{SL}(n,\mathbb{C})$. A important (and not very easy) theorem of E. Cartan says that any subgroup $H$ of a Lie group $G$ that is closed as a subset of $G$, is in fact an embedded submanifold, and hence is a Lie group in its own right. Thanks to Cartan, we don't actually have to check in any of these examples of matrix groups that they are embedded submanifolds: It is automatic from the fact that they are groups and closed subsets. Most examples of Lie groups encountered in practice (for instance, all compact groups) are matrix Lie groups. (An example of a Lie group that is not isomorphic to a matrix Lie group is the double covering of $\mathrm{SL}(2,\mathbb{R})$.) Any $a \in G$ defines two maps $l_a, r_a : G \to G$ with

$$l_a(g) = ag, \quad r_a(g) = ga.$$

The maps $l_a, r_a$ are called left-translation and right-translation, respectively. They are diffeomorphisms of $G$, with inverse maps $l_{a^{-1}}$ and $r_{a^{-1}}$.

Proposition 10.2. *For any Lie group, inversion $g \mapsto g^{-1}$ is a smooth map (hence a diffeomorphism).*

Proof. Consider the map $F : G \times G \to G \times G, \quad (g,h) \mapsto (g, gh)$. We claim that $F$ is a diffeomorphism. Once this is shown, smoothness of the inversion map follows since it can be written as a composition

$$G \longrightarrow G \times G \longrightarrow G \times G \longrightarrow G$$

where the first map is the inclusion $g \mapsto (g,e)$, the second maps is $F^{-1}(g,h) = (g, g^{-1}h)$, and the last map is projection to the second factor. Clearly $F$ is a bijection, with inverse map $F^{-1}(a,b) = (a, a^{-1}b)$. To show that $F$ is a diffeomorphism, it suffices to show that all elements of $G \times G$ are regular values of $F$, i.e. that the tangent map is a bijection everywhere. [5] Let us calculate the tangent map to $F$ at $(g,h) \in G \times G$. Suppose the path $\gamma(t) = (g_t, h_t)$ represents a vector $(v,w)$ in the tangent space, with $g_0 = g$ and $h_0 = h$. To calculate

$$T_{(g,h)}F(v,w) = T_{(g,h)}F(\frac{d\gamma}{dt}|_{t=0}) = \frac{d}{dt}|_{t=0}F(\gamma(t)) = \frac{d}{dt}|_{t=0}(g_t, g_t h_t),$$

we have to calculate the tangent vector to the curve $t \mapsto g_t h_t \in G$. We have

$$\begin{aligned} \frac{d}{dt}|_{t=0}(g_t h_t) &= \frac{d}{dt}|_{t=0}(g h_t) + \frac{d}{dt}|_{t=0}(g_t h) \\ &= T_h l_g(\frac{d}{dt}|_{t=0}(h_t)) + T_g r_h(\frac{d}{dt}|_{t=0}(g_t)) \\ &= T_h l_g(w) + T_g r_h(v). \end{aligned}$$

This shows

$$T_{(g,h)}F(v,w) = (v, T_h l_g(w) + T_g r_h(v))$$

which is 1-1 and therefore a bijection. □

For matrix Lie groups, smoothness of the inversion map also follows from Cramer's rule for the inverse matrix.

### 10.2. Definition of Lie algebras, the Lie algebra of a Lie group.

Definition 10.3. A Lie algebra is a vector space $\mathfrak{g}$, together with a bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying *anti-symmetry*

$$[\xi,\eta] = -[\eta,\xi] \text{ for all } \xi,\eta \in \mathfrak{g},$$

and the *Jacobi identity*,

$$[\xi,[\eta,\zeta]] + [\eta,[\zeta,\xi]] + [\zeta,[\xi,\eta]] = 0 \text{ for all } \xi,\eta,\zeta \in \mathfrak{g}.$$

The map $[\cdot,\cdot]$ is called the Lie bracket.

---

[5] We are using the following corollary of the regular value theorem: If $F \in C^\infty(M,N)$ has bijective tangent map at any point $p \in M$, then $F$ restricts to a diffeomorphism from a neighborhood $U$ of $p$ onto $F(U)$. Thus, if $F$ is a bijection it must be a diffeomorphism. (Smooth bijections need not be diffeomorphisms in general, the map $F : \mathbb{R} \to \mathbb{R}, t \mapsto t^3$ is a counter-example.)

Any associative algebra is a Lie algebra, with bracket the commutator. The space of vector fields $\mathfrak{X}(M)$ on a manifolds is a Lie algebra, with bracket what we've already called the Lie bracket of vector fields.

For any Lie group $G$, one defines a Lie algebra structure on the tangent space to the identity element, $\mathfrak{g} := T_eG$ in the following way. Let $\mathfrak{X}(G)^L$ denote the space of left-invariant vector fields on $G$. Thus $X \in \mathfrak{X}(G)^L$ if and only if $l_a^*(X) = X$ for all $a \in G$. Evaluation at the identity element gives a linear map

$$\mathfrak{X}(G)^L \to \mathfrak{g},\ X \mapsto \xi := X_e.$$

This map is an isomorphism: Given $\xi \in \mathfrak{g}$, one defines a left-invariant vector field $X$ by $X_g = T_eL_g(\xi)$. (Exercise: Check that $X$ is indeed smooth!) The Lie bracket of two vector fields is again left-invariant:

$$l_a^*[X,Y] = [l_a^*X, l_a^*Y] = [X,Y].$$

Thus $\mathfrak{X}(G)^L$ is a Lie subalgebra of the Lie algebra of all vector fields on $G$. Using the isomorphism $\mathfrak{X}(G)^L \cong \mathfrak{g}$, this gives a Lie algebra structure on $\mathfrak{g}$. That is, if we denote by $X = \xi^L$ the left-invariant vector field on $G$ generated by $\xi$, we have,

$$[\xi^L, \eta^L] = [\xi, \eta]^L.$$

Problems 10.4. We defined the Lie bracket on $\mathfrak{g} = T_eG$ by its identification with left-invariant vector fields. A second Lie algebra structure on $\mathfrak{g}$ is defined by identifying $T_eG$ with the space of right-invariant vector fields. How are the two brackets related? (Answer: One has $[\xi^R, \eta^R] = -[\xi, \eta]^R$, so the two brackets differ by sign.)

**10.3. Matrix Lie groups.** Let $G = \mathrm{GL}(n, \mathbb{R})$. Since $\mathrm{GL}(n, \mathbb{R})$ is an open subset of the set $\mathrm{Mat}_{\mathbb{R}}(n)$ of $n \times n$-matrices, all tangent spaces are identified with $\mathrm{Mat}_{\mathbb{R}}(n)$ itself. In particular $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{R}) \cong \mathrm{Mat}_{\mathbb{R}}(n)$. Let us confirm the obvious guess that the Lie bracket on $\mathfrak{g}$ is simply the commutator of matrices. The left-invariant vector field corresponding to $\xi \in \mathfrak{g}$ is

$$\xi^L_g = \frac{d}{dt}|_{t=0}(g\exp(t\xi)) = g\xi$$

(matrix multiplication). Its action on functions $f \in C^\infty(G)$ is,

$$\xi^L(f)_g = \frac{d}{dt}|_{t=0}(g\exp(t\xi)) = \sum_{ij}(g\xi)_{ij}\frac{\partial f}{\partial g_{ij}}|_g.$$

Hence,

$$\begin{aligned}\xi^L\eta^L(f)_g &= \frac{d}{dt}|_{t=0}\sum_{ij}(g\exp(t\xi)\eta)_{ij}\frac{\partial f}{\partial g_{ij}}|_{g\exp(t\xi)}\\ &= \sum_{ij}(g\xi\eta)_{ij}\frac{\partial f}{\partial g_{ij}}|_{g\exp(\xi)} + \ldots,\end{aligned}$$

where $\ldots$ involves second derivatives of the function $f$. (When we calculate Lie brackets, the second derivatives drop out so we need not care about $\ldots$.) We find,

$$(\xi^L\eta^L - \eta^L\xi^L)(f)_g = \sum_{ij}(g(\xi\eta - \eta\xi))_{ij}\frac{\partial f}{\partial g_{ij}}|_{g\exp(\xi)}.$$

Comparing to

$$[\xi,\eta]^L(f)_g = \sum_{ij}(g([\xi,\eta]))_{ij}\frac{\partial f}{\partial g_{ij}}|_{g\exp(\xi)}$$

this confirms that the Lie bracket is indeed just the commutator. [6] We obtain similar results for other matrix Lie groups: For instance, the Lie algebra of $\mathrm{O}(n) = \{A|\, A^t A = I\}$ is the space

$$\mathfrak{o}(n) = \{B|\, B + B^t = 0\},$$

with bracket the commutator, while the Lie algebra of $\mathrm{SL}(n,\mathbb{R})$ is

$$\mathfrak{sl}(n,\mathbb{R}) = \{B|\ \mathrm{tr}(B) = 0\},$$

with bracket the commutator. In all such cases, this follows from the result for the general linear group, once we observe that the exponential map for matrices takes $\mathfrak{g} \subset \mathfrak{gl}(n,\mathbb{R})$ to the corresponding subgroup $G \subset \mathrm{GL}(n,\mathbb{R})$.

**10.4. The exponential map for Lie groups.** There is an alternative characterization of the Lie algebra in terms of 1-parameter subgroups. A *1-parameter subgroup* of a Lie group $G$ is a smooth group homomorphism $\phi : \mathbb{R} \to G$, that is, $\phi(0) = e$ and $\phi(t_1 + t_2) = \phi(t_1)\phi(t_2)$. For any such $\phi$, the velocity vector at $t = 0$ defines an element $\xi \in T_e G = \mathfrak{g}$. Let $\xi^L$ be the corresponding left-invariant vector field. Then $\phi(t)$ is an integral curve for $\xi^L$:

$$\frac{d}{dt}\phi(t) = \frac{d}{ds}|_{s=0}\phi(t+s) = \frac{d}{ds}|_{s=0}\phi(t)\phi(s) = T_e l_{\phi(t)}\frac{d}{ds}|_{s=0}\phi(s) = T_e l_{\phi(t)}\xi = \xi^L_{\phi(t)}.$$

More generally, a similar calculation shows that for all $g \in G$, the curve $\gamma(t) = g\phi(t)$ is an integral curve through $g$. That is, the flow of $\xi^L$ is $\Phi(t,g) = g\phi(t)$.

Suppose conversely that $X$ is a left-invariant vector field. If $\gamma(t)$ is an integral curve, then so is its left translate $g\gamma(t)$ for any $g$. It follows that $X$ is complete and has a left-invariant flow. Let $\phi(t) = \Phi(t,e)$, then $\phi(t)$ is a 1-parameter subgroup, and $X = \xi^L$ for the corresponding $\xi \in \mathfrak{g}$. To summarize, elements of the Lie algebra are in 1-1 correspondence with 1-parameter subgroups. Let $\phi_\xi(t)$ denote the 1-parameter subgroup corresponding to $\xi \in \mathfrak{g}$.

Definition 10.5. For any Lie group $G$, with Lie algebra $\mathfrak{g}$, one defines the *exponential map*

$$\exp : \mathfrak{g} \to G, \quad \exp(\xi) := \phi_\xi(1).$$

Note that this generalizes the exponential map for matrices. Indeed, suppose $G \subseteq \mathrm{GL}(n,\mathbb{R})$ is a matrix Lie group, with Lie algebra $\mathfrak{g} \subseteq \mathfrak{gl}(n,\mathbb{R})$. Then the flow of the left-invariant vector field corresponding to $\xi \in \mathfrak{gl}(n,\mathbb{R})$ is just $\Phi_t(g) = g\exp(t\xi)$ (using the exponential map for matrices).

Theorem 10.6. *The exponential map is smooth, and defines a diffeomorphism from some open neighborhood $U$ of $0$ to $\exp(U)$.*

Proof. We leave smoothness as an exercise. For the second part, it suffices to show that the tangent map at 0 is bijective. Since $\mathfrak{g}$ is a vector space, the tangent space at 0 is identified with $\mathfrak{g}$ itself. Note that

$$\phi_{t\xi}(1) = \phi_\xi(t).$$

[6] This motivates why we used left-invariant vector fields in the definition of Lie bracket: Otherwise we would have found *minus* the commutator at this point.

Hence

$$(T_0 \exp)(\xi) = \frac{d}{dt}|_{t=0} \exp(t\xi) = \frac{d}{dt}|_{t=0}\phi_\xi(t) = \xi$$

thus $T_0 \exp$ is simply the identity map. $\square$

For matrix Lie groups, exp coincides with the exponential map for matrices (hence its name).

**10.5. Group actions.** Lie groups often arise as transformation groups, by some "action" on a manifold $M$.

DEFINITION 10.7. An action of a Lie group $G$ on a manifold $M$ is a group homomorphism $G \to \mathrm{Diff}(M),\ g \mapsto \Phi_g$ such that the action map $\Phi : G \times M \to M,\ (g.p) \mapsto \Phi_g(p)$ is smooth.

Note that an action of $G = \mathbb{R}$ is the same thing as a flow. Every matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$ acts on $\mathbb{R}^n$ in the obvious way. Any Lie group $G$ acts on itself by multiplication from the left $g \mapsto l_g$, multiplication from the right $g \mapsto r_{g^{-1}}$, and also by the *adjoint* (=conjugation) action $g \mapsto l_g r_{g^{-1}}$.

DEFINITION 10.8. An action of a finite dimensional Lie algebra $\mathfrak{g}$ on a manifold $M$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$, $\xi \mapsto \xi_M$ such that the action map $\mathfrak{g} \times M \to TM,\ (\xi, p) \mapsto \xi_M(p)$ is smooth.

THEOREM 10.9. *Given an action of a Lie group $G$ on a manifold $M$, one obtains an action of the corresponding Lie algebra $\mathfrak{g}$, by setting*

$$\xi_M(p) = \frac{d}{dt}|_{t=0}\Phi(\exp(-t\xi)).$$

*The vector field $\xi_M$ is called the generating vector field corresponding to $\xi$.*

EXERCISE 10.10. Prove this theorem. Hints: First verify the theorem for the left-action of a group on itself. (Show that $\xi_M$ equals $-\xi^R$ in this case.) Then, use that the action map $\Phi : G \times M \to M$ is equivariant, i.e. $\Phi \circ (l_a \times \mathrm{id}) = \Phi_a \circ \Phi$. Finally, show that $(-\xi^R, 0) \sim_\Phi \xi_M$. This implies

$$(-[\xi, \eta]^R, 0) = ([\xi^R, \eta^R], 0) \sim_\Phi [\xi_M, \eta_M].$$

Deduce $[\xi_M, \eta_M] = [\xi, \eta]_M$.

Note: Many people omit the minus sign in the definition of the generating vector field $\xi_M$. But then $\xi \mapsto \xi_M$ is not a Lie algebra homomorphism but an "anti-homomorphism". We prefer to avoid "anti" whenever possible.

## 11. Frobenius' theorem

**11.1. Submanifolds.** We defined embedded submanifolds as subsets of manifolds admitting submanifold charts. One often encounters more general submanifolds, in the following sense.

DEFINITION 11.1. Let $S, M$ be manifolds of dimensions $\dim S \leq \dim M$. A smooth map $F \in C^\infty(S, M)$ is an *immersion* if for all $p \in S$, the tangent map $T_pF$ is 1-1. It is called a submanifold if in addition $F$ is 1-1.

Locally, the image of any immersion looks like an embedded submanifold, by the following result:

THEOREM 11.2. *Let $F \in C^\infty(S, M)$ be an immersion. Then every point $p$ in $U$ has an open neighborhood $U \subset S$ such that $F(U)$ is an embedded submanifold.*

PROOF. Using local coordinates, it suffices to prove this for the case that $S$ is an open subset of $\mathbb{R}^s$ and $M$ an open subset of $\mathbb{R}^m$. Given $p$, we may renumber the coordinates such that $T_pF(\mathbb{R}^s) \cap \mathbb{R}^{m-s} = \{0\}$, where we view $\mathbb{R}^{m-s}$ as the subspace where the first $s$ coordinates are 0. Define a map,

$$\widetilde{F}: \ S \times \mathbb{R}^{m-s} \to \mathbb{R}^m, \quad (q, y) \mapsto F(q) + y.$$

It is easily checked that $T_{(p,0)}\widetilde{F}$ is a bijection. Hence, by the regular value theorem, there exists an open neighborhood $U$ of $p$ and an open ball $B_\epsilon(0)$ around $0 \in \mathbb{R}^{m-s}$ such that $\widetilde{F}$ restricts to a diffeomorphism from $U \times B_\epsilon(0)$ onto its image in $M \subset \mathbb{R}^m$. This gives the desired submanifold chart. $\square$

EXAMPLE 11.3. An smooth immersion $\gamma : J \to M$ from an open interval $J \subset \mathbb{R}$ is the same thing as a regular curve: For all $t \in J$, $\dot\gamma(t) \neq 0$.

In general, submanifolds need not be embedded submanifolds: For instance, the integral curves of a complete vector field define submanifolds $\mathbb{R} \to M$, but usually their images are not embedded. (Note that some authors use "submanifold" to denote embedded submanifolds, while others use the same terminology for immersions! We follow the conventions from F. Warner's book.)

**11.2. Integral submanifolds.** Let $X_1, \ldots, X_k$ be a collection of vector fields on a manifold $M$ such that the $X_i$ are pointwise linearly independent. That is, at every $p \in M$ the values $(X_i)_p$ of the vector fields span a $k$-dimensional subspace of the tangent space $T_pM$. A $k$-dimensional submanifold $\iota : S \hookrightarrow M$ is called an *integral submanifold* for $X_1, \ldots, X_k$, if each $X_j$ is tangent to $S$, that is $(X_j)_{\iota(p)} \in T_p\iota(T_pS) \subset T_{\iota(p)}M$ for all $p \in S$. We had seen above that the Lie bracket of any two vector fields tangent to $S$ is again tangent to $S$. Hence, a *necessary* condition for the existence of integral submanifolds through every given point $p \in S$ is that the $X_j$ are *in involution*: That is,

$$[X_i, X_j] = \sum_{l=1}^k c_{ij}^l X_l \tag{3}$$

for some functions $c_{ij}^l$. Frobenius' theorem (see below) asserts that this condition is also sufficient.

EXAMPLE 11.4. On $M = \mathbb{R}^3 \backslash \{x_2 = 0\}$ consider the vector fields,

$$X = x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}, \quad Z = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}$$

We have,

$$[X, Z] = x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1} =: Y.$$

Using $x_1X + x_2Y + x_3Z = 0$, we see that $X, Z$ are in involution.

As stated, the scope of Frobenius' theorem is limited since in general, manifolds need not admit pointwise linearly independent vector fields – often they don't even admit any vector field without zeroes. It is convenient to shift attention to the subbundle of $TM$ spanned by the vector fields, rather than the vector fields themselves:

DEFINITION 11.5. A $k$-dimensional distribution on a manifold $M$ is a rank $k$ vector subbundle $E$ of the tangent bundle $TM$. That is, $M$ can be covered by open subsets $U \subset M$ such that over each $U$, there are $k$ vector fields $X_1, \ldots, X_k$ spanning $E$. The distribution is called *integrable* if any such local basis is in involution. An submanifold $\iota : S \to M$ is called an integral submanifold for a (possibly non-integrable) distribution $E$ if $T_p\iota(T_pS) = E_{\iota(p)}$ for all $p \in S$.

EXERCISE 11.6. Show that the condition of being in involution does not depend on the choice of $X_i$'s: If $X_i' = \sum_j a_{ij} X_j$ and the $X_j$ are in involution, then so are the $X_i'$.

EXAMPLE 11.7. On $M = \mathbb{R}^3 \backslash \{0\}$ consider the three vector fields, $X, Y, Z$ introduced above. They are pointwise linearly dependent: $x_1 X + x_2 Y + x_3 Z = 0$. It follows that the vector bundle $E$ spanned by $X, Y, Z$ has rank 2. The above local calculation shows that $E$ is integrable. The spheres $x_1^2 + x_2^2 + x_3^2 = r^2$ are integral submanifolds.

### 11.3. Frobenius' theorem.

THEOREM 11.8 (Frobenius). *A rank $k$ distribution $E$ on a manifold $M$ is integrable, if and only if there exists an integral submanifold through every point $p \in M$. In this case, every point $p \in M$ admits a coordinate neighborhood $(U, \phi)$ in which $E$ is spanned by the first $k$ coordinate vector fields, $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}$.*

PROOF. We have seen that if there exists an integral submanifold through every point, then $E$ must be integrable. Suppose conversely that $E$ is integrable. It suffices to construct the coordinate charts $(U, \phi)$ described in the theorem: In such coordinates, it is clear that the integral submanifolds are given by setting the coordinates $x_{k+1}, \ldots, x_m$ equal to constants.

Choose an arbitrary chart around $p$, with coordinates $y_1, \ldots, y_m$, where $p$ corresponds to $y = 0$. Using the chart, we may assume that $M$ is an open subset of $\mathbb{R}^m$. Consider the $k$-dimensional subspace $E_p = E_0$. Renumbering the coordinates if necessary, we may assume that $E_0 \cap \mathbb{R}^{m-k} = \{0\}$, where $\mathbb{R}^{m-k}$ is identified with the subspace of $\mathbb{R}^m$ where the first $k$ coordinates are 0. Passing to a small neighborhood of $p = 0$ if necessary, we may assume that $E_q \cap \mathbb{R}^{m-k} = \{0\}$ for all $q$, or equivalently that orthogonal projection from $E_q$ to $\mathbb{R}^k$ is an isomorphism. That is, $E$ is spanned by vector fields of the form,

$$X_i = \frac{\partial}{\partial y_i} + \sum_{r=k+1}^{m} a_{ir} \frac{\partial}{\partial y_r}.$$

It turns out that we got very lucky: the $X_i$ commute! Indeed, by definition of the Lie bracket we have,

$$[X_i, X_j] = \sum_{r=k+1}^{m} (\frac{a_{jr}}{\partial y_i} - \frac{a_{ir}}{\partial y_j}) \frac{\partial}{\partial y_r},$$

but since the $X_i$ are in involution, we also have

$$[X_i, X_j] = \sum_l c_{ij}^l X_l = \sum_{l=1}^k c_{ij}^l \frac{\partial}{\partial y_l} + \sum_{l=k+1}^m \big(\sum_{\nu=1}^k c_{ij}^\nu a_{\nu l}\big) \frac{\partial}{\partial y_l}.$$

Comparing the coefficients of $\frac{\partial}{\partial y_l}$ for $l \le k$, we find that $c_{ij}^l = 0$, showing that the $X_i$ commute. Hence their flow $\Phi_{t_i}^i$ commute (wherever defined). Choose $\epsilon > 0$ sufficiently small, and let $U'$ be a small open neighborhood of $p$ such that for all $t = (t_1, \ldots, t_k) \in B_\epsilon(0)$ and all $q \in U$ the "joint flow"

$$\Phi(t_1, \ldots, t_k, q) = \Phi_{t_1}^1 \circ \cdots \circ \Phi_{t_k}^k(q)$$

is defined. Since the flows commute, we have

$$\frac{\partial}{\partial s}|_{s=0} \Phi(t_1, \ldots, t_j + s, \ldots, t_k, q) = (X_j)_{\Phi(t,q)}.$$

Define a map

$$F : B_\epsilon(0) \times (U' \cap \mathbb{R}^{m-k}) \to U, \ (t, q) \mapsto \Phi_t(q).$$

By construction,

$$TF(\frac{\partial}{\partial t_j}) = X_j, \quad j = 1, \ldots, k$$

and $TF(\frac{\partial}{\partial y_j}) = \frac{\partial}{\partial y_j})$ for $j > k$. In particular, $T_{(0,0)}F$ is invertible, hence $F$ restricts to a diffeomorphism on some open subset $U \subset B_\epsilon(0) \times (U' \cap \mathbb{R}^{m-k})$. The inverse map $F^{-1}$ gives the required change of coordinates. □

One has the following addendum to Frobenius' theorem:

THEOREM 11.9. *Suppose $E \subset TM$ is an integrable distribution. For each $p \in M$, there is a unique maximal connected integral submanifold $\iota : S \to M$ passing through $p$. That is, if $\iota' : S' \to M$ is any other integral submanifold through $p$, then there exists a smooth map $F : S' \to S$ such that $F$ is a diffeomorphism onto its image and $\iota' = \iota \circ F$.*

This is analogous to the fact that every vector field has a unique maximal integral curve through every given point of $M$. The idea of proof is to "patch together" the local solutions. Again, the theorem fails for non-Hausdorff manifolds.

The maximal integral submanifolds are called the *leaves* of the integrable distribution, and the decomposition of $M$ into leaves is called a *foliation*.

**11.4. Applications to Lie groups.** A *homomorphism of Lie groups* is a smooth group homomorphism $F : H \to G$. The tangent map at the identity $T_0F : \mathfrak{h} \to \mathfrak{g}$ is then a homomorphism of Lie algebras, i.e. takes brackets to brackets. (To see this, one proves that the left-invariant vector fields corresponding to $\xi$ and to $T_0F(\xi)$ are $F$-related.) A 1-1 Lie group homomorphism is called a Lie subgroup, in this case $T_0F : \mathfrak{h} \to \mathfrak{g}$ is a Lie subalgebra.

THEOREM 11.10. *Let $G$ be a Lie group, with Lie algebra $\mathfrak{g}$, and $j : \mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Then there exists a unique Lie subgroup $F : H \to G$ having $\mathfrak{h}$ as its Lie algebra: That is, $j = T_0F$.*

Proof. Let $E \subset TG$ be the distribution spanned by the left-invariant vector fields, $\xi^L$ with $\xi \in \mathfrak{h}$. Since $[\xi^L, \eta^L] = [\xi, \eta]^L$, this distribution is integrable. Let $\mathcal{L}_g$ denote the leaf through $g \in G$. The distribution is left-invariant: That is, for all $a \in G$ the tangent map to left translation, $Tl_a : TG \to TG$ takes $E$ to itself. Hence, for any $a \in G$ the left translate $l_a(\mathcal{L}_g) = \mathcal{L}_{ag}$ is again a leaf. Let $H := \mathcal{L}_e$. If $h_1, h_2 \in H$ we have $\mathcal{L}_{h_1} = \mathcal{L}_{h_2} = H$, and acting by $a = h_2^{-1}$ we get

$$\mathcal{L}_{h_2^{-1}h_1} = \mathcal{L}_e = H,$$

proving that $h_2^{-1}h_1 \in H$. This shows that $H$ is a group. Smoothness of the group operations follows from that for $G$. $\square$

We next describe an application of Frobenius' theorem to actions of Lie groups and Lie algebras on manifolds.

Definition 11.11. An action of a Lie group $G$ on a manifold $M$ is a group homomorphism $G \to \mathrm{Diff}(M),\ g \mapsto \Phi_g$ such that the action map

$$\Phi : G \times M \to M,\ (g.p) \mapsto \Phi_g(p)$$

is smooth. An action of a finite dimensional Lie algebra $\mathfrak{g}$ on a manifold $M$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$, $\xi \mapsto \xi_M$ such that the action map $\mathfrak{g} \times M \to TM,\ (\xi, p) \mapsto \xi_M(p)$ is smooth.

Examples 11.12. 1) Note that an action of the (additive) Lie group $G = \mathbb{R}$ is the same thing as a global flow, while an action of the Lie algebra $\mathfrak{g} = \mathbb{R}$ (with zero bracket) is the same thing as a vector field.

2) Every matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$, and every matrix Lie algebra acts on $\mathbb{R}^n$ by multiplication.

3) The rotation action of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ restricts to an action on the sphere, $S^{n-1} \subset \mathbb{R}^n$.

4) Any Lie group $G$ acts on itself by multiplication from the left, $l_a(g) = ag$, multiplication from the right $r_{a^{-1}}(g) = ga^{-1}$, and also by the *adjoint* (=conjugation) action

$$\mathrm{Ad}_a(g) := l_a r_{a^{-1}}(g) = aga^{-1}.$$

The maps

$$\xi \mapsto \xi^L,\ \ \xi \mapsto -\xi^R,\ \ \xi \mapsto \xi^L - \xi^R$$

are all Lie algebra actions of $\mathfrak{g}$ on $G$.

Theorem 11.13. *Given an action of a Lie group $G$ on a manifold $M$, one obtains an action of the corresponding Lie algebra $\mathfrak{g}$, by setting*

$$\xi_M(p) = \frac{d}{dt}|_{t=0}\Phi(\exp(-t\xi), p).$$

*The vector field $\xi_M$ is called the* generating vector field *corresponding to $\xi$.*

Proof. Let us first note that if $G$ acts on manifolds $M_1, M_2$, and if $F : M_1 \to M_2$ is a $G$-equivariant map, i.e.

$$F(g.p) = g.F(p)\ \ \forall p \in M_1$$

then $\xi_{M_1} \sim_F \xi_{M_2}$. This follows because $F$ takes integral curves for $\xi_{M_1}$ to integral curves for $\xi_{M_2}$. Thus, if we can show $[\xi, \eta]_{M_1} = [\xi_{M_1}, \eta_{M_1}]$. then a similar property holds for $M_2$.

We apply this to the special case

$$F : G \times M \to M, \quad (g, p) \mapsto \Phi(g^{-1}, p)$$

$\Phi : G \times M \to M$, with $G$-acting on $G \times M$ by the right-action on $G$ and trivial action on $M$, and on $M$ by the given action. The map is equivariant.

This reduces the problem to the special case $M = G$ with action the right-action $a \mapsto r_{a^{-1}}$ of $G$ on itself. We claim that

$$\xi_M = \xi^L$$

in this case. Indeed, the flow $g \mapsto g \exp(-t\xi)^{-1} = g \exp(t\xi)$ commutes with left translations, hence it is the flow of a left invariant vector field. Taking the derivative at $g = e, t = 0$ we see that this vector field is $\xi^L$, as claimed. But $[\xi^L, \eta^L] = [\xi, \eta]^L$. □

Exercise 11.14. Show that the generating vector field for the left action of $G$ on itself is $-\xi^R$, and the generating vector field for the adjoint action is $\xi^L - \xi^R$.

Note: Many people omit the minus sign in the definition of the generating vector field $\xi_M$. But then $\xi \mapsto \xi_M$ is not a Lie algebra homomorphism but an "anti-homomorphism". We prefer to avoid "anti" whenever possible.

Let us now consider the inverse problem: Try to integrate a given Lie algebra action to an action of the corresponding group!

Suppose $G$ is a connected Lie group, with Lie algebra $\mathfrak{g}$. We assume that $G$ is also *simply connected*: That is, every loop in $G$ can be contracted to a point. For instance, $G = \mathrm{SU}(n)$ is simply connected. If $G$ is a compact Lie group with finite center, one also knows that some finite cover of $G$ is simply connected.

Theorem 11.15. *Every Lie algebra action $\xi \mapsto \xi_M$ of $\mathfrak{g}$ on a compact manifold $M$ "exponentiates" uniquely to a Lie group action of the simply connected Lie group $G$, that is, an action for which $\xi_M$ are the generating vector fields.*

Sketch of proof. Every $G$-action on $M$ decomposes $G \times M$ into submanifolds $\mathcal{L}_p = \{(g^{-1}, g.p) | \, g \in G\}$, and the action may be recovered from this decomposition. The idea of proof, given a $\mathfrak{g}$-action, is to construct $\mathcal{L}_p$ as leafs of a foliation. Let $E \subset T(G \times M)$ be the distribution, of rank equal to $\dim G$, spanned by all vector fields $(\xi^L, \xi_M) \in \mathfrak{X}(G \times M)$ as $\xi$ ranges over the Lie algebra. Since

$$[(\xi^L, \xi_M), (\eta^L, \eta_M)] = ([\xi, \eta]^L, [\xi, \eta]_M),$$

the distribution is involutive. Hence it defines a foliation of $G \times M$ into submanifolds of dimension $\dim G$.

Given $p \in M$, let $\mathcal{L}_p \hookrightarrow G \times M$ be the unique leaf containing the point $(e, p)$. Projection to the first factor induces a smooth map $\mathcal{L}_p \to G$, with tangent map taking $(\xi^L, \xi_M)$ to $\xi^L$. Since the tangent map is an isomorphism, the map $\mathcal{L}_p \to G$ is a local diffeomorphism (that is, every point in $\mathcal{L}_p$ has an open neighborhood over which the map is a diffeomorphism onto its image). We claim that this map is surjective. Proof: By the Lemma given below, and since the exponential map $\exp : \mathfrak{g} \to G$ is a diffeomorphism on some neighborhood of 0, every $g \in G$ can be written as a product $g_1 \ldots g_N$ of elements $g_j = \exp(\xi_j)$ where $\xi_j \in \mathfrak{g}$. The curve $t \mapsto g_1 \ldots g_{j-1} \exp(t\xi_j)$ is an integral curve of the left-invariant vector field $\xi_j^L$. Taking all this curves together defines a piecewise smooth curve $\gamma$ connecting $e$ to $g$. This curve lifts to $\mathcal{L}_p$:

Since $M$ is compact, each $\xi_M$ is complete, hence each smooth segment of $\gamma$ to an integral curve of $(\xi^L, \xi_M)$.

We have shown at this stage that the map $\mathcal{L}_p \to G$ is a local diffeomorphism onto its image. Since $G$ is simply connected by assumption, it follows that the map is in fact a diffeomorphism. Hence, for every $g$, $\mathcal{L}_p$ contains a unique point of the form $(g^{-1}, p')$. Define $g.p = \Phi(g,p) := p'$. We leave it as an exercise to check that this map defines a smooth $G$-action. $\square$

LEMMA 11.16. *Let $G$ be a connected Lie group, and $U \subset G$ an open neighborhood of the group unit $e \in G$. Then every $g \in G$ can be written as a finite product $g = g_1 \cdots g_N$ of elements $g_j \in U$.*

PROOF. We may assume that $g^{-1} \in U$ whenever $g \in U$. For each $N$, let $U^N = \{g_1 \cdots g_N |\, g_j \in U\}$. We have to show $\bigcup_{N=0}^{\infty} U^N = G$. Each $U^N$ is open, hence their union is open as well. If $g \in G \backslash \bigcup_{N=0}^{\infty} U^N$, then $gU \in G \backslash \bigcup_{N=0}^{\infty} U^N$ (for if $gh \in \bigcup_{N=0}^{\infty} U^N$ with $h \in U$ we would have $g = (gh)h^{-1} \in \bigcup_{N=0}^{\infty} U^N$.) This shows that $G \backslash \bigcup_{N=0}^{\infty} U^N$ is also open. Since $G$ is connected, it follows that the open and closed set $\bigcup_{N=0}^{\infty} U^N$ is all of $G$. $\square$

## 12. Riemannian metrics

Let us quickly recall some linear algebra. A *bilinear form* on a vector space $V$ is a bilinear map $g : V \times V \to \mathbb{R}$. Such a bilinear form is called symmetric if $g(v,w) = g(w,v)$ for all $v, w$, and in this case is completely determined by the associated quadratic form $q(v) = g(v,v)$. $g$ is called an inner product if it is *positive definite*, i.e. $g(v,v) > 0$ for all $v \in V$. More generally, a symmetric form $g$ is called non-degenerate if $g(v,w) = 0$ for all $w$ implies $v = 0$. Non-degenerate symmetric bilinear forms are also called *indefinite inner products*.

Given a basis $e_1, \ldots, e_n$ of $V$, one can describe any bilinear form in terms of the matrix $g_{ij} = g(e_i, e_j)$. The bilinear form $g$ is symmetric if and only if the matrix $g_{ij}$ is symmetric, and in this case one can always choose the basis such that $g_{ij}$ is diagonal. In fact, one can choose the basis in such a way that only $+1, 0, -1$ arise as diagonal entries. Let $d_+, d_0, d_-$ the number of $+1, 0, -1$ diagonal entries. Then $g$ is non-degenerate if $d_0 = 0$, and is an inner product if and only if $d_0 = d_- = 0$, i.e. if there exists a basis such that $g_{ij} = \delta_{ij}$.

EXERCISE 12.1. Show that one can split $V = V_+ \oplus V_-$ where $\dim V_\pm = d_\pm$ and $g$ is positive definite on $V_+$, negative definite on $V_-$. However, looking at the case $(d_+, d_-) = (1,1)$, observe that this splitting is *not unique*.

DEFINITION 12.2. A *Riemannian metric* on a manifold $M$ is a family of inner products $g_p : T_pM \times T_pM \to \mathbb{R}$, depending smoothly on $p$ in the sense that the quadratic form

$$q : TM \to \mathbb{R}, \quad v \mapsto g_p(v,v) \text{ for } v \in T_pM$$

is a smooth map $q \in C^\infty(TM)$. More generally, a *pseudo-Riemannian* metric of signature $(d_+, d_-)$ is defined by letting the $g_p$ be indefinite inner products of signature $(d_+, d_-)$.

The case of signature $(3,1)$ is relevant to general relativity, with 3 space dimensions and 1 time dimension. Again, there is no distinguished splitting into “space” and “time” directions.

LEMMA 12.3. *Any pseudo-Riemannian metric defines a symmetric $C^\infty(M)$-bilinear map*

$$g : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M), \quad g(X,Y)_p = g_p(X_p, Y_p).$$

*Conversely, every symmetric $C^\infty(M)$-bilinear map $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$, with the property that $g(X,Y)_p = 0$ for all $Y$ implies $X_p = 0$, defines a pseudo-Riemannian metric.*

Proof. Let $g$ be a pseudo-Riemannian metric, with quadratic form $q$. View vector fields as smooth sections, $X \in \Gamma^\infty(M, TM)$. Then

$$g(X,Y) = \tfrac{1}{2}(g(X+Y, X+Y) - g(X,X) - g(Y,Y)) = \tfrac{1}{2}(q \circ (X+Y) - q \circ X - q \circ Y)$$

is smooth, while $C^\infty(M)$-bilinearity is obvious. Conversely, suppose we are given a $C^\infty(M)$-bilinear map $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ with the property that $g(X,Y)_p = 0$ for all $Y$ implies $X_p = 0$. The following Lemma shows that $g(X,Y)_p$ depends only on $X_p, Y_p$. Hence we can define

$$g_p(X_p, Y_p) := g(X,Y)_p.$$

If $g_p(v,w) = 0$ for all $v$, choose $Y$ with $Y_p = w$. Then $g(X,Y)_p = g(X_p, Y_p) = 0$ for all $X, Y$, which by assumption implies $Y_p = 0$. Hence $g_p$ is non-degenerate. Using the formula $q \circ X = g(X,X)$, and passing to local coordinates, one sees that $g_p$ depends smoothly on $p$, hence it defines a Riemannian metric. □

Lemma 12.4. *If $A : \mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M) \to C^\infty(M)$ is a $C^\infty(M)$-multilinear map, then the value of $A(X_1, \ldots, X_r)$ at $p \in M$ depends only on $(X_1)_p, \ldots, (X_r)_p$. More generally, this Lemma holds true for any $C^\infty(M)$-multilinear map from $\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)$ to a $C^\infty(M)$-module.*

Proof. It suffices to consider the case $r = 1$. We have to show that if $X$ vanishes at $p$, then $A(X)$ vanishes at $p$. But if $X_p = 0$, we can write (using local coordinates, and the Taylor expansion) $X = \sum_i f_i X_i$ where $X_i \in \mathfrak{X}(M)$ and where $f_i \in C^\infty(M)$ vanish at $p$. Hence,

$$A(X)_p = A(\sum_i f_i X_i)_p = \sum_i f_i(p) A(X_i)_p = 0,$$

by $C^\infty(M)$-linearity. □

Definition 12.5. A (pseudo)-*Riemannian* manifold $(M,g)$ is a manifold $M$ together with a (pseudo)-*Riemannian* metric. An isometry between (pseudo)-*Riemannian* manifold $(M_1, g_1)$ and $(M_2, g_2)$ is a diffeomorphism $F : M_1 \to M_2$ such that for all $p \in M_1$, the tangent map $T_pF : T_pM_1 \to T_{F(p)}M_2$ is an isometry, i.e. preserves inner products.

In local coordinates $x_1, \ldots, x_m$ on $U \subset M$, any pseudo-Riemannian metric is determined by smooth functions

$$g_{ij}(x) = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}).$$

Indeed, one recovers $g$ from the $g_{ij}$ by

$$g(\sum_i a_i \frac{\partial}{\partial x_i}, \sum_j b_j \frac{\partial}{\partial x_j}) = \sum_{ij} g_{ij} a_i b_i.$$

Conversely, every collection of smooth functions $g_{ij}$, such that each $(g_{ij}(x))$ is a non-degenerate symmetric bilinear form, defines a Riemannian metric. In particular, to $g_{ij} = \delta_{ij}$ defines the *standard metric* on $\mathbb{R}^n$. How does $g_{ij}$ depend on the choice of coordinates? Let $y = \phi(x)$ be a coordinate change, and let $\tilde{g}_{ij}(y)$ denote the matrix in $y$-coordinates. We have,

$$\frac{\partial}{\partial y_i} = \sum_a \frac{\partial x_a}{\partial y_i} \frac{\partial}{\partial x_a}.$$

Hence,

$$\tilde{g}_{ij}(y) = g(\frac{\partial}{\partial y_i}, \frac{\partial}{\partial y_j}) = \sum_{ab} \frac{\partial x_a}{\partial y_i} \frac{\partial x_b}{\partial y_j} g_{ab}(\phi(y))$$

LEMMA 12.6. *Let $S \subset M$ be an embedded submanifold, and $g$ a Riemannian metric on $M$. Then the restriction of $g$ to the tangent spaces $T_pS \subset T_pM$ defines a Riemannian metric on $M$. More generally, if $\iota : S \to M$ is an immersion, there is a unique Riemannian metric on $S$ such that each tangent map $T_p\iota : T_pS \to T_{\iota(p)}M$ is an isometry onto its image.*

In particular, every embedded submanifold of $\mathbb{R}^m$ inherits a Riemannian metric from the standard Riemannian metric on $\mathbb{R}^m$.

EXAMPLE 12.7. The 2-torus $T^2$ can be defined as a direct product of the circle $S^1 \subset \mathbb{R}^2$ with itself. Correspondingly we have an embedding $T^2 \to R^4$ and the corresponding induced metric $g$ on $T^2$. The resulting metric on $T^2$ is simply the product of the metrics on the $S^1$ factors, and in particular is *flat*: $T^2$ is locally isometric to $\mathbb{R}^2$. It follows that there is no embedding of $T^2$ into $\mathbb{R}^3$, inducing the same metric $g$: We had seen in our curves and surfaces course that any connected surface in $\mathbb{R}^3$ with vanishing first fundamental form is an open subset of a plane, hence cannot be a compact surface.

## 13. Existence of Riemannian metrics

To show that every manifold admits a Riemannian metric, we need an important technical tool called *partitions of unity*.

THEOREM 13.1 (Partitions of unity). *Let $M$ be a manifold.*

*a) Any open cover $\{U_\alpha\}$ of $M$ has a* locally finite refinement *$\{V_\beta\}$: That is, $\{V_\beta\}$ is an open cover, each $V_\beta$ is contained in some $U_\alpha$, and the cover is locally finite in the sense that each point in $M$ has an open neighborhood meeting only finitely many $V_\beta$'s.*

*b) For any locally finite cover $U_\alpha$ of $M$, there exists a* partition of unity*, that is a collection of functions $\chi_\alpha \in C^\infty(M)$ with* $\mathrm{supp}(\chi_\alpha) \subset U_\alpha$*, such that $0 \le \chi_\alpha \le 1$ and*

$$\sum_\alpha \chi_\alpha = 1.$$

Note that the sum $\sum_\alpha \chi_\alpha$ is well-defined, since only finitely many $\chi_\alpha$'s are non-zero near any given point. We will omit the somewhat technical proof of this result. The proof is contained in most books on differential geometry (e.g. Helgason), and can also be found in the lecture notes from my "manifolds" course. The main steps for part (b) are as follows:

(i) One constructs a "shrinking" of the open cover $U_\alpha$ to a new cover $V_\alpha$, such that $\overline{V}_\alpha \subset U_\alpha$. The new cover is still locally finite.
(ii) One constructs functions $f_\alpha \in C^\infty(M)$ supported on $U_\alpha$, such that $f_\alpha > 0$ on $V_\alpha$,
(iii) One defines $f = \sum_\alpha f_\alpha > 0$, and sets $\chi_\alpha = f_\alpha / f$.

THEOREM 13.2. *Every manifold $M$ admits a Riemannian metric.*

PROOF. Choose an atlas $\{(U_\alpha, \phi_\alpha)\}$ of $M$. Passing to a refinement, we may assume that the atlas is locally finite. Choose a partition of unity $\chi_\alpha$ for the cover $\{U_\alpha\}$. Since $\phi_\alpha$ identifies $U_\alpha$ with an open subset of $\mathbb{R}^m$, we obtain Riemannian metrics $g_\alpha$ on $U_\alpha$ from the standard Riemannian metrics on $\mathbb{R}^m$. For all $p \in M$, the sum $g_p = \sum \chi_\alpha(p)(g_\alpha)_p$ is well-defined. Since all $\chi_\alpha(p) \geq 0$, with at least one strictly positive, $g_p$ is an inner product with clearly a smooth dependence on $p$. Thus $g$ is a Riemannian metric on $M$. □

It is *not* true that every manifold admits a *pseudo*-Riemannian metric of given signature $(d_+, d_-)$, where both $d_\pm \neq 0$.

## 14. Length of curves

Suppose $(M, g)$ is a Riemannian manifold (that is, a manifold with a Riemannian metric).[7] For any tangent vector $v \in T_pM$, we define its length as $||v|| = g_p(v, v)^{1/2}$.

DEFINITION 14.1. Let $\gamma : [a, b] \to M$ be a smooth curve in $M$.[8] One defines the *length* of $\gamma$ to be the integral

$$L(\gamma) = \int_a^b ||\dot{\gamma}(t)|| \, \mathrm{d}t.$$

The length functional is invariant under reparametrizations of the curve $\gamma$. Somewhat more generally, we have:

PROPOSITION 14.2. *Let $\sigma : [a, b] \to \mathbb{R}$ be a smooth function, with the property that $\sigma(t_1) \leq \sigma(t_2)$ for $t_1 \leq t_2$. Let $\gamma : [a, b] \to M$ be a smooth curve of the form $\gamma = \tilde{\gamma} \circ \sigma$. Then $L(\gamma) = L(\tilde{\gamma})$.*

PROOF. By substitution of variables $\tilde{t} = \sigma(t)$, [9]

$$\begin{aligned} L(\gamma) &= \int_a^b ||\frac{d}{dt}(\tilde{\gamma} \circ \sigma)|| \mathrm{d}t \\ &= \int_a^b ||\frac{d\tilde{\gamma}}{d\tilde{t}}(\sigma(t))|| \, |\frac{d\sigma}{dt}| \, \mathrm{d}t \\ &= \int_{\sigma(a)}^{\sigma(b)} ||\frac{d\tilde{\gamma}}{d\tilde{t}}|| \mathrm{d}\tilde{t} \\ &= L(\tilde{\gamma}) \end{aligned}$$

□

The definition of $L(\gamma)$ applies to *piecewise smooth curves*: That is, continuous curves $\gamma : [a, b] \to M$ such that there exists a subdivision $a = t_0 \leq \cdots \leq t_N = b$ of the interval, with each $\gamma|_{[t_i, t_{i+1}]}$ a smooth curve.

[7] For the following discussion, see chapter 1.4 in Jost's book.

[8] Here smooth means that $\gamma$ extends to a smooth curve on an open interval $J$ containing $[a, b]$.

[9] If $\sigma : [a, b] \to \mathbb{R}$ is a continuous, piecewise smooth map, which is weakly increasing in the sense that $\sigma(t_1) \leq \sigma(t_2)$ for $t_1 \leq t_2$, then $\int_a^b f(\sigma(t))|\frac{d\sigma}{dt}|\mathrm{d}t = \int_{\sigma(a)}^{\sigma(b)} f(\tilde{t})\mathrm{d}\tilde{t}$.

If the curve $\gamma$ is contained in a fixed coordinate chart $(U, \phi)$, and $(x_1(t), \dots, x_m(t))$ describes the curve in local coordinates, we have

$$L(\gamma) = \int_a^b \sqrt{\sum_{ij} g_{ij}(x(t))\, \dot{x}_i \dot{x}_j}\ \mathrm{d}t.$$

Definition 14.3 (Distance function). Let $(M, g)$ be a connected Riemannian manifold. For $p, q \in M$, the distance $d(p, q)$ between any two points on $M$ is infimum of $L(\gamma)$, as $\gamma$ varies over all piecewise smooth curves $\gamma : [0, 1] \to M$ with $\gamma(0) = p$ and $\gamma(1) = q$. (If no such path exists, we set $d(p, q) = \infty$.)

Problems 14.4. 1. Show that for any manifold $M$, the following are equivalent: (i) $M$ is connected, (ii) any two points $p, q$ can be joined by a continuous path, (iii) any two points $p, q$ can be joined by a piecewise smooth path, (iv) any two points $p, q$ can be joined by a smooth path. Hence $d(p, q) < \infty$ for a connected manifold.

2. Show that in the definition of distance function, one can replace piecewise smooth paths by smooth paths. In fact, any piecewise smooth path is of the form $\gamma = \lambda \circ \sigma$, where $\sigma$ is weakly increasing and piecewise smooth, and $\lambda$ is smooth.

Lemma 14.5. *Let $(U, \phi)$ be a coordinate chart in which $g$ is given by $g_{ij}(x)$, and $K \subset \phi(U)$ a compact subset. Then there exist $\lambda \geq \mu > 0$ with*

$$\mu \sqrt{\sum_i \xi_i \xi_i} \geq \sqrt{\sum_{ij} g_{ij}(x) \xi_i \xi_j} \geq \lambda \sqrt{\sum_i \xi_i \xi_i}. \tag{4}$$

*for $x \in K$, $\xi \in \mathbb{R}^n$.*

Proof. The set of all $(x, \xi) \in \mathbb{R}^{2n}$ with $x \in K$ and $\sum_i \xi_i \xi_i = 1$ is compact. Hence the function $\sum_{ij} g_{ij}(x) \xi_i \xi_j$ takes on its maximum $\lambda$ and minimum $\mu$ on this set. By definition of a Riemannian metric, $\mu > 0$. □

Theorem 14.6. *For any connected manifold $M$, the distance function $d$ defines a metric on $M$. That is, $d(p, q) \geq 0$ with equality if and only if $p = q$, and for any three points $p, q, r$, one has the triangle inequality*

$$d(p, q) + d(q, r) \geq d(p, r).$$

Proof. The triangle inequality is immediate from the definition. Suppose $p \neq q$. We have to show $d(p, q) > 0$. Choose a chart $(U, \phi)$ around $p$, with $\phi(p) = 0$, and let $\epsilon > 0$ be sufficiently small, such that the closed ball $\overline{B_\epsilon}$ is contained in $\phi(U)$ and $\phi^{-1}(\overline{B_\epsilon})$ does not contain $q$. Let $g_{ij}$ represent the metric in the chart $(U, \phi)$.

Given a curve $\gamma$ from $p$ to $q$, let $t_1 < b$ is such that $\gamma(t) \in \phi(U)$ for $a \leq t \leq t_1$ and $\gamma(t_1) \in \overline{B}_\epsilon \backslash B_\epsilon$. Write $\phi(\gamma(t)) = x(t)$ for $a \leq t \leq t_1$. Using the Lemma,

$$L(\gamma) \geq \int_a^{t_1} \sqrt{\sum_{ij} g_{ij}(x(t)) \dot{x}_i \dot{x}_j}\ \mathrm{d}t \geq \lambda \int_a^{t_1} \sqrt{\sum_i \dot{x}_i \dot{x}_i}\ \mathrm{d}t \geq \lambda \epsilon,$$

since the length of the path from $\phi(p) = 0$ to $x(t_1)$ must be at least the Euclidean distance $\epsilon$. Hence also

$$d(p, q) = \inf_\gamma L(\gamma) \geq \lambda \epsilon > 0.$$

□

THEOREM 14.7. *For any manifold $M$, the topology defined by the metric coincides with the manifold topology.*

PROOF. This follows from the Lemma: In local charts, $\epsilon$-balls for the metric $d$ contain sufficiently small Euclidean $\delta$-balls, and vice versa. □

## 15. Connections and parallel transport

In this section, we will define the parallel transport of tangent vectors on any Riemannian manifold $(M, g)$. If $M \subset \mathbb{R}^m$ is an embedded submanifold of $\mathbb{R}^m$, with metric induced from $\mathbb{R}^m$, we can follow the strategy from the "curves and surfaces" course: At any $p \in M$ we have an orthogonal projection $\Pi_p : \mathbb{R}^m \to T_pM$. If $\gamma(t)$ is a curve and $X(t) \in T_{\gamma(t)}M$ a vector field along $\gamma$, we say that $X$ is parallel along $\gamma$ if the *covariant derivative*

$$\frac{\nabla X}{\mathrm{d}t} := \Pi_{\gamma(t)} \frac{\mathrm{d}X}{\mathrm{d}t}$$

vanishes for all $t$. Here we have used that $X(t)$ can be viewed as an $\mathbb{R}^m$-valued function of $t$. Using the existence and uniqueness theorem for ODE's, one finds that any parallel vector field along $\gamma$ is determined by its value $X(t_0)$ at any fixed time $t_0$.

For a general Riemannian manifold $(M, g)$, we don't have "orthogonal projection" at our disposal. It is remarkable that there exists, nevertheless, a well-defined concept of parallel transport on any Riemannian manifold $(M, g)$. That is, parallel transport is really an intrinsic property. Our starting point for defining parallel transport is to define generalized covariant derivatives, called affine connections. We will then show that any Riemannian manifold carries a distinguished affine connection.

**15.1. Affine connections.** Let $M$ be a manifold.

DEFINITION 15.1. An affine connection on $M$ is a bi-linear map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad (X, Y) \to \nabla_X(Y)$$

such that

$$\begin{aligned} \nabla_X(fY) &= f\nabla_X(Y) + X(f)Y \\ \nabla_{fX}(Y) &= f\nabla_X(Y). \end{aligned}$$

for all $f \in C^\infty(M)$, $X, Y \in \mathfrak{X}(M)$.

The second condition says that the operator $\nabla_X$ is $C^\infty$-linear in the $X$ variable. One calls $\nabla_X(Y)$ the covariant derivative of $Y$ in the direction of $X$. $Y$ is called covariant constant in the direction of $X$ if $\nabla_X(Y) = 0$.

If $M = U$ is an open subset of $\mathbb{R}^m$, any affine connection $\nabla$ is determined by its values on coordinate vector fields. The functions $\Gamma^i_{jk} \in C^\infty(U)$ defined by

$$\nabla_{\frac{\partial}{\partial x_j}}\left(\frac{\partial}{\partial x_k}\right) = \sum_i \Gamma^i_{jk} \frac{\partial}{\partial x_i} \tag{5}$$

are called the *Christoffel symbols* of $\nabla$. The full connection is given in terms of Christoffel symbols by the formula,

$$\nabla_{\sum_j a_j \frac{\partial}{\partial x_j}}(\sum_k b_k \frac{\partial}{\partial x_k}) = \sum_j a_j \left( \sum_k \frac{\partial b_k}{\partial x_j} \frac{\partial}{\partial x_k} + \sum_{k,i} \Gamma^i_{jk} b_k \frac{\partial}{\partial x_i} \right).$$

Conversely, it is easily checked that *any* collection of smooth functions $\Gamma^i_{jk}$ defines an affine connection by this formula. In particular, open subsets $U \subset \mathbb{R}^m$ have the *standard* affine connection, given by $\Gamma^i_{jk} = 0$.

More generally, for affine connections on manifolds one defines Christoffel symbols of a connection with respect to a given chart. First we note that if $U \subset M$ is an open subset, affine connections $\nabla$ have a unique restriction $\nabla|_U$ with the property

$$(\nabla|_U)_{X|_U}(Y|_U) = \nabla_X(Y)|_U.$$

Moreover, every connection is determined by its restrictions to elements of an open cover of $M$. Hence we may define:

Definition 15.2. Let $\nabla$ be an affine connection on a manifold $M$. If $(U, \phi)$ is a chart, defining local coordinates $x_1, \ldots, x_m$, one defines the *Christoffel symbols* $\Gamma^i_{jk}$ of $\nabla|_U$ in the given chart to be the functions defined by (5).

Problems 15.3. 1. Calculate the Christoffel symbols of the standard connection on $\mathbb{R}^2$ in polar coordinates. The solution shows that Christoffel symbols may vanish in one coordinate system but be non-zero in another. 2. Work out the transformation property of Christoffel symbols under change of coordinates.

Proposition 15.4. *For any affine connection $\nabla$ on $M$, the map $T : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ given by*

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

*is $C^\infty(M)$-linear in both $X$ and $Y$. It is called the* torsion *of $\nabla$.*

Proof. For all $f \in C^\infty(M)$,

$$\begin{aligned} T(X, fY) - fT(X, Y) &= \nabla_X(fY) - f\nabla_X(Y) - [X, fY] + f[X, Y] \\ &= X(f)\,Y - X(f)Y = 0. \end{aligned}$$

Similarly $T(fX, Y) - fT(X, Y) = 0$. □

In local coordinates we have, in terms of the Christoffel symbols,

$$T(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}) = \sum_{jk} (\Gamma^i_{jk} - \Gamma^i_{kj}) \frac{\partial}{\partial x_i}.$$

Hence, the connection is torsion-free if and only if the Christoffel symbols $\Gamma^i_{jk}$ are symmetric in $j, k$. In particular, if the Christoffel symbols have this symmetry property in one system of coordinates, then also in every other system.

### 15.2. The Levi-Civita connection.

PROPOSITION 15.5. *Let $(M,g)$ be a (pseudo-)Riemannian manifold. For any affine connection $\nabla$ on $M$, and any $Z\in\mathfrak{X}(M)$ the map $\nabla_Z g:\ \mathfrak{X}(M)\times\mathfrak{X}(M)\to C^\infty(M)$ given by*

$$(\nabla_Z g)(X,Y)=Z\,g(X,Y)-g(\nabla_Z(X),Y)-g(X,\nabla_Z(Y))$$

*is $C^\infty(M)$-linear in both $X,Y$. It is called the covariant derivative of $g$ in the direction of $Z$. The connection $\nabla$ is called a* metric connection *if $\nabla_X g=0$.*

The proof is straightforward. In local coordinates and the corresponding Christoffel symbols for $\nabla$, we have

$$(\nabla g)^k_{ij}:=(\nabla_{\frac{\partial}{\partial x_k}}g)(\frac{\partial}{\partial x_i},\frac{\partial}{\partial x_j})=\frac{\partial g_{ij}}{\partial x_k}-\sum_l \Gamma^l_{ki}g_{lj}-\Gamma^l_{kj}g_{li}.$$

and $\nabla$ is a metric connection if and only of the right hand side vanishes.

THEOREM 15.6 (Fundamental Theorem of Riemannian Geometry). *Suppose $(M,g)$ is a pseudo-Riemannian manifold. There exists a unique torsion-free metric connection $\nabla$ on $M$. It is called the* Levi-Civita connection.

PROOF. Suppose $\nabla$ is a torsion-free metric connection. Since $\nabla$ is metric, we have

$$Z\,g(X,Y)=g(\nabla_Z(X),Y)+g(X,\nabla_Z(Y)).$$

Using the torsion free condition $\nabla_Z(X)=\nabla_X(Y)+[Z,X]$ this gives,

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

Permuting letters we also have

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

Use these equations to eliminate $\nabla_X$ and $\nabla_Y$, and obtain

$$Z\,g(X,Y)-X\,g(Y,Z)+Y\,g(Z,X))=2g(X,\nabla_Z(Y))+g([Z,X],Y)-g([X,Y],Z)+g([Y,Z],X),$$

that is,

$$\begin{aligned}&2g(X,\nabla_Z(Y))\\ &=Z\,g(X,Y)-X\,g(Y,Z)+Y\,g(Z,X))-g([Z,X],Y)+g([X,Y],Z)-g([Y,Z],X).\end{aligned}\tag{6}$$

Since $g$ is non-degenerate, any vector field $W$ is completely determined by its parings $g(X,W)$ with all vector fields $X$. In particular, (6) specifies the vector field $W=\nabla_Z(Y)$. This shows that a torsion-free metric connection $\nabla$ is determined by the metric $g$. Conversely, it is straightforward to check that formula (6) defines a torsion-free metric connection. For instance, if we replace $Y$ by $fY$ for some function $f$, we find

$$\begin{aligned}&2(g(X,\nabla_Z(fY))-f\,g(X,\nabla_Z(Y))\\ &=\ Z(f)\,g(X,Y)-X(f)\,g(Y,Z)+g(X(f)Y,Z)+g(Z(f)Y,X)\\ &=\ 2\,Z(f)\,g(X,Y)\end{aligned}$$

which shows that $\nabla_Z(fY)-f\nabla_Z(Y)=Z(f)\,Y$. The other properties are checked similarly. $\square$

EXERCISE 15.7. Try to re-derive the explicit formula (6) for the Levi-Civita connection without looking at the notes. Fill in the details of showing that this formula defines a torsion-free metric connection.

COROLLARY 15.8. *Every manifold admits a torsion-free affine connection $\nabla$.*

PROOF. We have seen that every manifold admits a Riemannian metric $g$. Thus one can take the Levi-Civita connection with respect to $g$. $\square$

Taking $X = \frac{\partial}{\partial x_l}, Y = \frac{\partial}{\partial x_k}, Z = \frac{\partial}{\partial x_j}$ to be coordinate vector fields in (6), we obtain a formula for the Christoffel symbols $\Gamma^i_{jk}$ of the Levi-Civita connection:

$$2\sum_i \Gamma^i_{jk} g_{il} = \frac{\partial g_{kl}}{\partial x_j} + \frac{\partial g_{jl}}{\partial x_k} - \frac{\partial g_{jk}}{\partial x_l}.$$

Letting $(g^{-1})_{ij}$ denote the inverse matrix to $g_{ij}$, this gives:

THEOREM 15.9. *In local coordinates, the Christoffel symbols for the Levi-Civita connection are given by*

$$\Gamma^i_{jk} = \frac{1}{2}\sum_l (g^{-1})_{il}\left(\frac{\partial g_{kl}}{\partial x_j} + \frac{\partial g_{jl}}{\partial x_k} - \frac{\partial g_{jk}}{\partial x_l}\right).$$

We had seen a similar formula in the curves and surfaces course. In fact, we could have used this formula to define a connection in local coordinates, and then check that the local definitions patch together. However, the significance of this rather complicated formula would remain obscure from such an approach.

It is immediate from this formula that $\nabla$ is torsion-free, since the Christoffel symbols are symmetric in $j, k$.

**15.3. Parallel transport.** Let $\nabla_X(Y)$ be an affine connection on a manifold $M$. Since $\nabla_X(Y)$ is $C^\infty$-linear in the $X$-variable, the value of $\nabla_X(Y)$ at $p$ depend only on $X_p$. Thus if $v \in T_pM$ one can define $\nabla_v(Y) \in T_pM$ by $\nabla_v(Y) := \nabla_X(Y)_p$ where $X$ is any vector field with $X_p = v$. If $\gamma\colon J \to M$ is any curve, one can therefore define

$$\nabla_{\dot\gamma(t)} Y \in T_{\gamma(t)}M.$$

If $x(t)$ is the description of the curve $\gamma$ in local coordinates $x_1, \ldots, x_m$, so that $\dot\gamma = \sum_i \dot x_i \frac{\partial}{\partial x_i}$, and $Y = \sum_k b_k \frac{\partial}{\partial x_k}$,

$$\begin{aligned}
\nabla_{\dot\gamma(t)} Y &= \sum_{ij} \dot x_j \left(\frac{\partial b_i}{\partial x_j} + \sum_k \Gamma^i_{jk} b_k\right)\frac{\partial}{\partial x_i}.\\
&= \sum_i \left(\frac{db_i}{dt} + \sum_{jk} \Gamma^i_{jk} \dot x_j b_k\right)\frac{\partial}{\partial x_i}.
\end{aligned}$$

Here $\frac{db_i}{dt} = \frac{d}{dt} b_i(x(t))$. Note that this formula depends only on the "restriction" of $Y$ to $\gamma$, or more precisely on the section of the pull-back bundle $\gamma^*(TM) \to J$ defined by $Y$. In fact, the formula makes sense for any *vector field along* $\gamma$, that is, any section of $\gamma^*(TM) \to J$. In local coordinates, vector fields along $\gamma$ are given by expressions $Y = \sum_i b_k(t) \frac{\partial}{\partial x_k} \in T_{\gamma(t)}M$

depending smoothly on $t$, and the above formula in local coordinates defines a new vector field along $\gamma$,[10]

$$\frac{DY}{dt} \equiv \nabla_{\dot\gamma(t)} Y.$$

DEFINITION 15.10. A vector field $Y$ along a curve $\gamma : J \to M$, is called *parallel* along $\gamma$ if the covariant derivative $\frac{DY}{dt}$ vanishes everywhere.

THEOREM 15.11. *Let $\nabla$ be a metric connection on a manifold $M$. Let $\gamma : J \to M$ be a smooth curve, $X_0 \in T_{\gamma(t_0)}M$ where $t_0 \in J$. Then there is a unique parallel vector field $X(t) \in T_{\gamma(t)}$ along $\gamma$, with the property $X(t_0) = X_0$. The linear map*

$$T_{\gamma(t_0)}M \to T_{\gamma(t)}M, \quad X_0 \mapsto X(t)$$

*is called parallel transport along $\gamma$, with respect to the connection $\nabla$.*

PROOF. In local coordinates as above, parallel vector fields are the solutions of the first order ordinary differential equations,

$$\frac{db_i}{dt} + \sum_{jk} \Gamma^i_{jk} \dot{x_j} b_k = 0.$$

Hence, for "short times" the theorem follows from the existence and uniqueness theorem for ODE's, and for "long times" by patching together local solutions. □

PROPOSITION 15.12. *Let $(M, g)$ be a pseudo-Riemannian manifold, and $\nabla$ an affine connection on $M$. Then*

$$\frac{d}{dt} g(X(t), Y(t)) = (\nabla_{\dot\gamma} g)(X(t), Y(t)) + g(\frac{DX}{dt}, Y) + g(X, \frac{DY}{dt}).$$

PROOF. In local coordinates, write $X(t) = \sum_i a_i(t) \frac{\partial}{\partial x_i}$ and $Y(t) = \sum_j b_j(t) \frac{\partial}{\partial x_j}$, and let $x(t) = (x_1(t), \ldots, x_m(t))$ be the coordinate expression for the curve $\gamma$. Then

$$\begin{aligned} \frac{d}{dt} g(X(t), Y(t)) &= \frac{d}{dt} \sum_{ij} g_{ij} a_i b_j \\ &= \sum_{ijk} \dot{x}_k \frac{\partial g_{ij}}{\partial x_k} + \sum_{ij} g_{ij} \dot{a}_i b_j + \sum_{ij} g_{ij} a_i \dot{b}_j \end{aligned}$$

and

$$\begin{aligned} g(\frac{DX}{dt}, Y) &= \sum_{ij} (\dot{a}_i + \sum_{lm} \Gamma^i_{lm} \dot{x}_l a_m) b_j, \\ g(X, \frac{DY}{dt}) &= \sum_{ij} a_i (\dot{b}_j + \sum_{lm} \Gamma^j_{lm} \dot{x}_l b_m). \end{aligned}$$

[10] Of course, it would be better to give a coordinate free definition. For this, one has to generalize the notion of an affine connection, and introduce connections $\nabla$ on vector bundles $E \to M$. For any $X \in \mathfrak{X}(M)$, $\nabla_X$ is an endomorphism of the space of sections $\Gamma^\infty(E)$. For any smooth map $F : N \to M$ one then obtains a connection $F^*\nabla$ on the pull-back bundle $F^*E$. In our case, we obtain a connection $\gamma^*\nabla$ on $\gamma^*(TM)$. One then defines

$$\frac{DY}{dt} := (\gamma^*\nabla)_{\frac{\partial}{\partial t}} Y(t).$$

Taking this three equations together, and using

$$(\nabla g)^k_{ij} = (\nabla_{\frac{\partial}{\partial x_k}} g)(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}) = \frac{\partial g_{ij}}{\partial x_k} - \sum_l \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{li},$$

the Proposition follows.[11] $\square$

As an immediate consequence, we have:

PROPOSITION 15.13. *An affine connection $\nabla$ on a pseudo-Riemannian manifold $(M, g)$ is a metric connection if and only if parallel transport along curves preserves inner products.*

## 16. Geodesics

Let $\nabla$ be an affine connection on a manifold $M$.

DEFINITION 16.1. A smooth curve $\gamma : J \to M$ is called a geodesic for the connection $\nabla$, if and only if the velocity vector field $\dot{\gamma}$ is parallel along $\gamma$.

EXERCISE 16.2. Show that if $\gamma : J \to M$ is a geodesic, and $\phi : \tilde{J} \to J$ is a diffeomorphism (change of parameters), then

$$\tilde{\gamma}(\tilde{t}) = \gamma(\phi(\tilde{t}))$$

is a geodesic if and only if $\frac{d\phi}{d\tilde{t}} = \text{const}$, i.e. if and only if $\phi(\tilde{t}) = a\tilde{t} + b$ for some $a \neq 0, b$.

As a special case of the differential equation for a parallel vector field $X(t) = \sum_i b_i(t) \frac{\partial}{\partial x_i}$, here $X(t) = \dot{\gamma}$ i.e. $b_i = \dot{x}_i$, we find:

THEOREM 16.3. *In local coordinates, geodesics are the solutions of the second order ordinary differential equation,*

$$\frac{d^2 x^i}{dt^2} + \sum_{jk} \Gamma^i_{jk} \dot{x}_j \dot{x}_k = 0.$$

Notice that only the symmetric part $\Gamma^i_{jk} + \Gamma^i_{kj}$, that is the torsion-free part of $\nabla$, contributes to the geodesic equation. Thus, if one is interested in the geodesic flow of a metric connection $\nabla$, one might as well assume that $\nabla$ is the Levi-Civita connection. On $\mathbb{R}^m$ with the standard Riemannian metric, geodesics are straight lines with constant speed parametrization.

It is a standard trick in ODE theory to reduce higher order ODE's to a system of first order ODE's, by introducing derivatives as parameters. In our case, if we introduce $\dot{x}_i =: \xi_i$, the geodesic equation becomes a system,

$$\begin{aligned} \frac{dx_i}{dt} &= \xi_i \\ \frac{d\xi_i}{dt} &= -\sum_{jk} \Gamma^i_{jk} \xi_j \xi_k. \end{aligned}$$

[11] We had to resort to this terrible proof since we defined the covariant derivative along curves in coordinates only. In the coordinate free definition, the Proposition is almost a triviality because it is essentially just the definition of $\nabla g$!

Notice that $x_i, \xi_i$ are just the standard local coordinates on $TM$ induced by the local coordinates $x_i$ on $M$. Hence, the above first order system defines a vector field $\mathcal{S}$ on $TM$, given in local coordinates by

$$\mathcal{S} = \sum_i \xi_i \frac{\partial}{\partial x_i} - \sum_{ijk} \Gamma^i_{jk} \xi_j \xi_k \frac{\partial}{\partial \xi_i}$$

Definition 16.4. The vector field $\mathcal{S}$ is called the *geodesic spray* of $\nabla$, and its flow is called the geodesic flow.

Theorem 16.5. *For any $p \in M$, $v \in T_pM$ there exists a unique maximal geodesic $\gamma_v : J \to M$, where $\gamma_v(0) = p$, $\dot{\gamma}_v(0) = v$.*

Proof. Let $\Phi_t$ denote the geodesic flow, and $\pi : TM \to M$ the base point projection. The geodesics on $M$ are just the projections of solution curves of the geodesic spray $\mathcal{S}$. In particular, $\gamma_v$ is given by

$$\gamma_v(t) = \pi(\Phi_t(v)).$$

$\square$

Notice that the geodesic flow has the property

$$\frac{\mathrm{d}}{\mathrm{d}t}(\pi(\Phi_t(v))) = \Phi_t(v).$$

This is the coordinate free reformulation of $\dot{x}_i = \xi_i$. Furthermore, it has the property

$$\Phi_t(av) = \Phi_{at}(v)$$

for $a \in \mathbb{R}$; this just says that if $\gamma(t)$ is a geodesic, with $\dot{\gamma}(0) = v$, then $t \mapsto \gamma(at)$ is also a geodesic, but with initial velocity $av$.

Exercise 16.6. Show that every non-constant geodesic is regular, i.e. $\dot{\gamma} \neq 0$ everywhere.

Definition 16.7. The manifold $M$ with affine connection $\nabla$ is called *geodesically complete* if the geodesics spray is a complete vector field. A (pseudo-)Riemannian manifold $(M, g)$ is called geodesically complete if it is geodesically complete for the Levi-Civita connection.

Thus geodesic completeness means that all geodesics exist for all time.

The property $\gamma_{av}(t) = \gamma_v(at)$ for all $a \in \mathbb{R}$ is reminiscent of a property of 1-parameter subgroups of Lie groups. Similar to the Lie groups case we define:

Definition 16.8. Suppose $(M, \nabla)$ is geodesically complete. The map

$$\mathrm{Exp}_p : T_pM \to M, \quad v \mapsto \gamma_v(1)$$

is called the exponential map based at $p$.

Compare with the very similar definition of exponential maps for Lie groups – the curves $\gamma_v$ play the role of 1-parameter subgroups! In terms of the exponential map, we have

$$\gamma_v(t) = \mathrm{Exp}_p(tv).$$

Theorem 16.9. *The exponential map $\mathrm{Exp}_p$ is smooth. It defines a diffeomorphism from a neighborhood of $0 \in T_pM$ onto a neighborhood of $p \in M$.*

PROOF. Let $\Phi : \mathbb{R}\times TM \to TM$ denote the flow of the geodesic spray, $\mathcal{S}$ for the connection $\nabla$, and let $\pi : TM \to M$ be the base point projection. Then $\operatorname{Exp}_p$ is just the restriction of the map $\pi \circ \Phi$ to the submanifold $\{1\} \times T_pM$, and hence is smooth. Compute $T_0 \operatorname{Exp}_p$: For $v \in T_pM$ we have,

$$T_0 \operatorname{Exp}_p(v) = \frac{d}{dt}|_{t=0} \operatorname{Exp}_p(tv) = \frac{d}{dt}|_{t=0}\gamma_v(t) = v,$$

so $T_0 \operatorname{Exp}_p$ is just the identity map $T_pM \to T_pM$. From the inverse function theorem, it then follows that $\operatorname{Exp}_p$ is a diffeomorphism on some small neighborhood of $0 \in T_pM$. $\square$

If one chooses a basis in $T_pM$, thus identifying $T_pM \cong \mathbb{R}^m$, the exponential map gives a system of local coordinates $x_1, \ldots, x_m$ on a neighborhood of $p$. These coordinates are called *normal coordinates* at $p$, and have very nice properties:

THEOREM 16.10. *In normal coordinates $x_1, \ldots, x_m$ based at $p \in M$, the geodesics through $p$ are given by straight lines,*

$$x_i(t) = ta_i \quad a_i \in \mathbb{R}.$$

*Moreover, all Christoffel symbols $\Gamma^i_{jk}$ vanish at 0.*

PROOF. By definition of the exponential map, $\operatorname{Exp}_p(ta)$ for $a \in \mathbb{R}^m \cong T_pM$ is the geodesic with initial velocity $v = a$. Inserting $x_i(t) = ta_i$ into the geodesic equation, we obtain

$$\sum_{jk} \Gamma^i_{jk}(ta)\, a_j a_k = 0$$

for all $a$. Setting $t = 0$, it follows that $\Gamma^i_{jk}(0) = 0$. $\square$

We now specialize to the case that $\nabla$ is the Levi-Civita connection corresponding to a (pseudo-)Riemannian metric $g$ on $M$. Define the *energy function*

$$E \in C^\infty(TM), \quad E(v) = \tfrac{1}{2} g_p(v, v), \quad v \in T_pM.$$

(Thus, the energy function is just the quadratic function associated to $g$, up to the factor $\frac{1}{2}$.) Since parallel transport for a metric connection preserves inner products, the geodesic flow preserves the energy: That is, $\mathcal{S}(E) = 0$. It follows that $\mathcal{S}$ is tangent to the level surfaces of the energy functional.

Geodesics for the Levi-Civita connection have an important alternative characterization, as critical points of the *action functional*.

DEFINITION 16.11. Let $\gamma : [a, b] \to M$ be a smooth curve in $M$.[12] One defines the *action* of $\gamma$ by

$$A(\gamma) = \int_a^b E(\dot\gamma(t))\ \mathrm{d}t = \tfrac{1}{2} \int_a^b ||\dot\gamma(t)||^2\ \mathrm{d}t.$$

In local coordinates, $A(\gamma) = \frac{1}{2} \int_a^b \sum_{ij} g_{ij}(x(t))\, \dot x_i \dot x_j\ \mathrm{d}t$. The action functional is closely related to the length functional (assuming that $g$ is positive definite, so that $L(\gamma)$ is defined):

[12]Here smooth means that $\gamma$ extends to a smooth curve on an open interval $J$ containing $[a, b]$.

LEMMA 16.12. *Let $\gamma : [a,b] \to M$ be a smooth curve. Then*

$$L(\gamma)^2 \leq 2(b-a)A(\gamma).$$

*Equality holds if and only if $\gamma$ has constant speed, that is $||\dot{\gamma}||$ is constant.*

PROOF. The Cauchy-Schwartz inequality[13] implies

$$(\int_a^b f(t)\mathrm{d}t)^2 \leq (b-a)\int_a^b f(t)^2 \ \mathrm{d}t$$

with equality if and only if $f$ is constant. □

Suppose $\gamma : [a,b] \to M$ is a smooth curve. A 1-parameter variation of $\gamma$ is a family of curves $\gamma_s : [a,b] \to M$ defined for $-\epsilon < s < \epsilon$, with $\gamma_s(a) = \gamma(a)$ and $\gamma_s(b) = \gamma(b)$ for all $t$, $\gamma_0 = \gamma$, and such that the map $(s,t) \mapsto \gamma_s(t)$ is smooth.

THEOREM 16.13. *A smooth curve $\gamma : [a,b] \to M$ is a geodesic if and only if for all 1-parameter variations $\gamma_s$ of $\gamma$,*

$$\frac{d}{ds}|_{s=0} A(\gamma_s) = 0.$$

PROOF. Let $\gamma_s(t)$ be a 1-parameter variation. We can view $\gamma_s(t)$ as a curve with parameter $t$, depending on $s$ as a parameter, or vice versa. Let a $'$ indicate $s$-derivatives. Since the Levi-Civita connection $\nabla$ is torsion-free,

$$\frac{D\gamma'}{dt} = \frac{D\dot{\gamma}}{ds}.$$

(In local coordinates, the left hand side is given by $\frac{D\gamma'}{dt} = \sum_i \left(\frac{d^2 x_i}{dsdt} + \sum_{jk} \Gamma^i_{jk} \dot{x}_j x'_k\right) \frac{\partial}{\partial x_i}$, while the right hand side is given by a similar expression with $s,t$-derivatives in opposite order. The two expressions are the same since the Christoffel symbols for a torsion free connection are symmetric in $j,k$.) Since $\nabla$ is a metric connection, we can therefore compute,

$$\begin{aligned}
\frac{d}{ds}A(\gamma_s) &= \tfrac{1}{2}\int_a^b \frac{\partial}{\partial s}\, g(\dot{\gamma},\dot{\gamma})\ dt \\
&= \int_a^b g(\frac{D\dot{\gamma}}{ds},\dot{\gamma})\ dt \\
&= \int_a^b g(\frac{D\gamma'}{dt},\dot{\gamma})\ dt \\
&= \int_a^b \frac{d}{dt} g(\gamma',\dot{\gamma})\ dt - \int_a^b g(\gamma',\frac{D\dot{\gamma}}{dt})\ dt \\
&= -\int_a^b g(\gamma',\frac{D\dot{\gamma}}{dt})\ dt.
\end{aligned}$$

[13] The Cauchy-Schwartz inequality for integrals says that

$$(\int_a^b f(t)g(t)\mathrm{d}t)^2 \leq (\int_a^b f(t)^2 \mathrm{d}t)\ (\int_a^b g(t)^2 \mathrm{d}t)$$

with equality if and only if $f,g$ are linearly dependent (i.e. proportional). The desired inequality follows by setting $g = 1$.

Here we have used that $\gamma'(a) = \gamma'(b) = 0$. The resulting expression vanishes at $s = 0$, for all variations, if and only if $\frac{D\dot\gamma}{dt} = 0$, i.e. if and only if $\gamma$ is a geodesic. $\square$

In particular, if $\gamma:\,[a,b] \to M$ minimizes the action, in the sense that

$$A(\gamma) \le A(\tilde\gamma)$$

for all paths $\tilde\gamma:\,[a,b] \to M$ (defined on the same interval $[a,b]$) with $\tilde\gamma(a) = \gamma(a),\ \ \tilde\gamma(b) = \gamma(b)$, then $\gamma$ is a geodesic. (However, it is not necessary for a geodesic to minimize the action.)

THEOREM 16.14. *A curve $\gamma:\,[a,b] \to M$ with $||\dot\gamma(t)|| = \mathrm{const}$ is a geodesic if and only if, for all 1-parameter variations $\gamma_s$ of $\gamma$,*

$$\frac{\partial}{\partial s}|_{s=0} L(\gamma_s) = 0.$$

We leave the proof as an exercise. We have to put in by hand the assumption that $\gamma$ has constant speed, since the length functional is invariant under reparametrizations. (The 1-parameter variations $\gamma_s$ need not have finite speed.) In particular, length minimizing, constant speed curves are always geodesics.

## 17. The Hopf-Rinow Theorem

The Hopf-Rinow Theorem says that a Riemannian manifold $(M,g)$ is geodesically complete (the geodesic flow is complete, i.e. all geodesics exists for all time) if and only if it is completene as a metric space (every Cauchy sequence converges). To prepare for the proof, we need some more facts on normal coordinates and the exponential map.

DEFINITION 17.1. Let $(M,g)$ be a Riemannian manifold. The *injectivity radius* $i_p(M) > 0$ of $p \in M$ is the supremum of the set of all $r > 0$ such that the exponential map $\mathrm{Exp}_p$ is defined on the open ball $B_r(0)$ and is injective. The injectivity radius $i(M) \ge 0$ of $M$ is the infimum of all $i_p(M)$ with $p \in M$.

EXAMPLE 17.2. For the unit circle $S^1 \subset \mathbb{R}^2$ with the standard Riemannian metric, each point has injectivity radius $\pi$. Similarly, for the sphere $M = S^{m-1} \subset \mathbb{R}^m$, the injectivity radius of any point is $i_p(M) = \pi$. For $M = \mathbb{R}^m$, $i_p(M) = \infty$.

THEOREM 17.3. *For all $0 < r < i_p(M)$, the radial geodesics $\mathrm{Exp}_p(tv)$ intersect the spheres $\mathrm{Exp}_p(S_r(0))$ orthogonally. For any $v \in S_r(0)$, the point $q = \mathrm{Exp}_p(v)$ has distance $d(p,q) = r$ from $p$, and the geodesic $\mathrm{Exp}_p(tv)$ is the unique (up to reparametrization) curve of length $d(p,q)$ connecting $p,q$. In particular,*

$$\mathrm{Exp}_p(S_r(0)) = S_r(p)$$

*for any $0 < r < i_p(M)$.*

We will obtain this result as a consequence of the following Lemma on "geodesic polar coordinates" around $p$. Let $x_1,\ldots,x_m$ denote the normal coordinates on a neighborhood $U$ of $p$, obtained by choosing an orthornormal basis in $T_pM$. In this coordinates,

$$g_{ij}(0) = \delta_{ij}, \quad \Gamma^i_{jk}(0) = 0.$$

Introduce polar coordinates $(\rho, \phi_1, \ldots, \phi_{m-1})$ on $T_pM$, thus $\rho^2 = \sum x_i^2$ and $\phi_1,\ldots,\phi_{m-1}$ are local coordinates on the unit sphere $S^{m-1} \subset T_pM$. (The particular choice of coordinates on

$S^{m-1}$ will be irrelevant.) Using $\mathrm{Exp}_p$, we can view these as coordinates on (suitable open subsets of) $\mathrm{Exp}_p(B_r(0))$ for $r < i_p(M)$. In particular, the coordinate vector field $\frac{\partial}{\partial\rho}$ given as

$$\frac{\partial}{\partial\rho} = \frac{1}{||x||}\sum_{i=1}^{m} x_i \frac{\partial}{\partial x_i}$$

is a well-defined vector fiel on $\mathrm{Exp}_p(B_r(0)\backslash\{0\})$. Note that its integral curves are exactly the unit speed radial geodesics.

Lemma 17.4 (Geodesic polar coordinates). *In geodesic polar coordinates around p,*

$$g_{\rho\rho} \equiv g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\rho}) = 1, \quad g_{\rho\phi_j} \equiv g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j}) = 0.$$

*That is, the radial geodesics* $\mathrm{Exp}_p(tv)$ *are orthogonal to the spheres* $\mathrm{Exp}_p(S_r(0))$*, for all* $0 < r < i_p(M)$*.*

Proof. Thus $\nabla_{\frac{\partial}{\partial\rho}}\frac{\partial}{\partial\rho} = 0$. In particular, the length of $\frac{\partial}{\partial\rho}$ is constant along radial geodesics. But

$$\lim_{t\to 0} g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\rho})|_{tx} = 1,$$

since $g_{ij}(0) = \delta_{ij}$ and since $\frac{\partial}{\partial\rho}$ has length one in the Euclidean metric. It follows that $g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\rho}) = 1$ everywhere. Furthermore, using that the connections is torsion-free,

$$\begin{aligned} \frac{\partial}{\partial\rho} g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j}) &= g(\nabla_{\frac{\partial}{\partial\rho}}\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j}) + g(\frac{\partial}{\partial\rho}, \nabla_{\frac{\partial}{\partial\rho}}\frac{\partial}{\partial\phi_j}) \\ &= g(\frac{\partial}{\partial\rho}, \nabla_{\frac{\partial}{\partial\phi_j}}\frac{\partial}{\partial\rho}) \\ &= \tfrac{1}{2}\frac{\partial}{\partial\phi_j} g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\rho}) \\ &= \tfrac{1}{2}\frac{\partial}{\partial\phi_j}\ 1 = 0. \end{aligned}$$

Thus $g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j})$ is constant in radial directions. But $\lim_{t\to 0} g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j})|_{tx} = 0$, again since $g_{ij}(0) = \delta_{ij}$. Thus $g(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\phi_j}) = 0$ everywhere. □

Proof of Theorem 17.3. Let $\gamma(t)$ $(0 \leq t \leq 1)$ be any curve with $\gamma(0) = p$ and $\gamma(1) = q$. Suppose first that $\gamma(t) \in \mathrm{Exp}_p(B_r(0)\backslash\{0\})$ for $0 < t < 1$. In geodesic polar coordinates

$$\dot\gamma = \dot\rho\frac{\partial}{\partial\rho} + \sum_j \dot\phi_j \frac{\partial}{\partial\phi_j}$$

thus $g(\dot\gamma,\dot\gamma) \geq |\dot\rho|^2$ with equality if and only if $\phi_j = \text{const}$. It follows that

$$\begin{aligned} L(\gamma) &= \int_0^1 g(\dot\gamma,\dot\gamma)^{1/2}\,dt \\ &\geq \int_0^1 |\dot\rho|\,dt \\ &\geq \int_0^1 \dot\rho\,dt \\ &= \rho(1) = r, \end{aligned}$$

with equality if and only if $\phi_j = \text{const}$ and $\dot\rho \geq 0$ for all $t$. Clearly, curves leaving the set $\operatorname{Exp}_p(B_r(0))$ for some time $t \in (0,1)$ will be even longer. □

Corollary 17.5. *Let $p,q \in M$. Suppose there exists a piecewise smooth curve $\gamma : [0,1] \to M$ of length $d(p,q)$ from $p$ to $q$. Then $\gamma$ is a reparametrization of a smooth (!) geodesic of length $d(p,q)$.*

Proof. Since $\gamma([0,1]) \subset M$ is compact, the infimum of the set of all injectivity radii $i_{\gamma(t)}(M)$ is strictly positive. Let $\epsilon > 0$ be smaller than this infimum. Then for any two points on the curve, of distance less than $\epsilon$, the unique shortest curve connecting these points is the geodesic given by the exponential map. In particular, $\gamma$ must coincide with that geodesic up to reparametrization. □

We are now ready to prove the Hopf-Rinow theorem. We recall that a sequence $x_n,\ n = 1,\ldots,\infty$ in a metric space $(X,d)$ (where $d$ is the metric = distance function) is a *Cauchy sequence* if for all $\epsilon > 0$, there exists $N > 0$ such that $d(x_n,x_m) < \epsilon$ for $n,m \geq N$. In particular, every convergent sequence is a Cauchy sequence. A metric space is called *complete* if every Cauchy sequence in $X$ converges. For instance, every compact metric space is complete, while e.g. bounded open subsets of $\mathbb{R}^m$ (with induced metric) are incomplete.

Exercise 17.6. Show that every Cauchy sequence is bounded. That is, there exists $p \in X$ and $R > 0$ such that $x_n \in B_R(p)$ for all $n$.

Theorem 17.7 (Hopf-Rinow). *A Riemannian manifold $(M,g)$ is geodesically complete, if and only if it is complete as a metric space. In this case, any two points $p,q$ may be joined by a smooth geodesic of length $d(p,q)$.*

Proof. We may assume that $M$ is connected.

Suppose $M$ is geodesically incomplete. That is, there exists a *maximal* unit speed geodesic $\gamma : (a,b) \to M$ with $b < \infty$. Since $d(\gamma(t_i),\gamma(t_j)) \leq |t_j - t_i|$, it follows that the sequence $\gamma(t_i)$ for $t_i \to b$ is a Cauchy sequence. On the other hand, this sequence cannot converge since $\gamma(t)$ leaves every given compact set[14] for $t \to b$.

Thus we have found a non-convergent Cauchy sequence, showing that $M$ is incomplete as a metric space.

The other direction is a bit harder: Suppose $M$ is geodesically complete. Pick $p \in M$. We will show that every closed metric ball $\overline{B_r(p)}$ is compact, which implies that $M$ is metrically

[14]For any compact set $K$, there exists $\epsilon > 0$ less than the injectivity radius of any point in $K$. Hence, unit speed geodesics for points starting in $K$ exist at least for time $\epsilon$.

complete (Any Cauchy sequence is bounded, hence is contained in $\overline{B_r(p)}$ for $r$ sufficiently large. Any Cauchy sequence in a compact set converges.) By geodesic completeness, the exponential map

$$\mathrm{Exp}_p : T_pM \to M$$

is defined. It suffices to show that for all $r > 0$,

$$\mathrm{Exp}_p(\overline{B_r(0)}) = \overline{B_r(p)},$$

where $B_r(0) \subset T_pM$ is the ball of radius $r$ for the inner product $g_p$. Indeed, $\overline{B_r(0)}$ is compact, and images of compact sets under continuous maps are again compact. The inclusion $\subseteq$ is clear; the harder part is the opposite inclusion $\overline{B_r(p)} \subseteq \mathrm{Exp}_p(\overline{B_r(0)})$. Let

$$H = \{r > 0 \mid \mathrm{Exp}_p(\overline{B_r(0)}) = \overline{B_r(p)}\}.$$

We have to show $H = [0, \infty)$. We first show that $H$ is closed. Let $r_n \in H$ with $\lim_{n\to\infty} r_n = r$. We have to show $r \in H$. Given $q \in \overline{B_r(p)}$, choose $q_n \in \overline{B_{r_n}(p)}$ with $q_n \to q$. Choose $v_n \in \overline{B_{r_n}(0)}$ with $\mathrm{Exp}_p(v_n) = q_n$. Since $\overline{B_{r_n}(0)} \subset \overline{B_r(0)}$ which is compact, there exists a convergent subsequence. Let $v \in \overline{B_r(0)}$ be the limit point; then $\mathrm{Exp}_p(v) = q$. Since $q$ was arbitrary, this shows $r \in H$.

We next show that if $r \in H$ then $r + \epsilon \in H$ for $\epsilon > 0$ sufficiently small. Since $H$ is closed, this will finish the proof $H = [0, \infty)$. Let $q \in \overline{B_{r+\epsilon}(p)}$.

Choose $0 < \epsilon < i_K(M) := \inf_{p\in K} i_p(M)$, where $K$ is the compact subset $K = \overline{B_r(p)}$. Thus, for any $x \in \overline{B_r(p)}$ with $d(x, q) \leq \epsilon$, there exists a unique geodesic in $M$ joining $x, q$, of length $d(x, q)$. To find such a point $x$, choose a sequence of curves $\gamma_n : [0, 1] \to M$ connecting $p, q$, of length $\leq d(p, q) + \frac{1}{n}$. Let $t_n \in [0, 1]$ be the smallest value such that $x_n := \gamma_n(t_n) \in \partial\overline{B_r(p)}$. We have

$$d(p, q) \leq d(p, x_n) + d(x_n, q) \leq L(\gamma_n) \leq d(p, q) + \frac{1}{n}.$$

Since $\overline{B_r(p)}$ is compact some subsequence of the sequence $x_n$ converges to a limit point $x \in \partial\overline{B_r(p)}$, with

$$d(p, x) + d(x, q) = d(p, q).$$

Since $d(p, q) \leq r + \epsilon$ and $d(p, x) = r$, this implies $d(x, q) \leq \epsilon$. Choose $v \in \overline{B_r(0)}$ with $\mathrm{Exp}_p(v) = p$.

Since $d(x, q) \leq \epsilon$, there exists a unique unit speed geodesic of length $d(x, q)$ from $x$ to $q$. Together with the unit speed geodesic $\mathrm{Exp}_p(tv/r)$, we obtain a piecewise smooth curve of length $d(p, q)$ from $p$ to $q$. As observed above, it is automatic that this curve is smooth, hence a geodesic. It hence coincides with the unique continuation of the geodesic $\mathrm{Exp}_p(tv/r)$. It follows that $\mathrm{Exp}_p(\tilde{v}) = q$ for

$$\tilde{v} = \frac{r + \epsilon}{r} v = (1 + \epsilon/r)v \in \overline{B_{r+\epsilon}(0)}.$$

$\square$

Note that we didn't quite use geodesic completeness in the proof: We only used that $\mathrm{Exp}_p$ is defined on all of $T_pM$. One might call this *geodesic completeness at p*. What we've shown is that geodesic completeness at any point $p$ implies geodesic completeness everywhere.

## 18. The curvature tensor

An affine connection $\nabla$ on a manifold $M$ is called flat if, around any point, there exist local coordinates in which all Christoffel symbols of $\nabla$ vanish. A (pseudo-)Riemannian metric is called flat if the corrsponding Levi-Civita connection is flat.

Flatness of a connection implies that parallel transport along a path does not change under 1-parameter variations of the path. In practice, the definition is not always easy to verify, mainly because Christoffel symbols may vanish in one coordinate system and be non-zero in another. One is therefore interested in *invariants* of a connection: That is, quantities constructed from the connection whose vanishing does not depend on a choice of coordinates. One such example is the *torsion* $T(X,Y) = \nabla_X(Y) - \nabla_Y(X) - [X,Y]$ of a connection: Recall that by $C^\infty$-linearity, it defines a bi-linear map $T : T_pM \times T_pM \to T_pM$, and clearly $T$ has to vanish for any flat connection. A second invariant is the *curvature operator* to be discussed now.

DEFINITION 18.1. For vector fields $X, Y$ one defines the *curvature operator* $R(X,Y) : \mathfrak{X}(M) \to \mathfrak{X}(M)$ by

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

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

THEOREM 18.2. *The map $(X,Y,Z) \mapsto R(X,Y)(Z)$ is $C^\infty(M)$-linear in $X, Y, Z$. It follows that for $u, v \in T_pM$, there is a well-defined linear map $R_p(u,v) : T_pM \to T_pM$ such that*

$$R_p(u,v)(w) = (R(X,Y)(Z))_p$$

*whenever $X_p = u, Y_p = v, Z_p = w$.*

PROOF. $R$ is $C^\infty(M)$-linear in $Z$: For all $f$,

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

Subtracting the last two equations from the first, we find $R(X,Y)(fZ) = fR(X,Y)(Z)$ as desired. Similarly one checks $C^\infty(M)$-linearity in $X, Y$. $\square$

$C^\infty$-linearity of the curvature operator implies that in local charts, $R$ is determined by its values on coordinate vector fields. We can thus introduce components $R^l_{ijk}$ of the curvature tensor, defined by

$$R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})(\frac{\partial}{\partial x_k}) = \sum_l R^l_{ijk} \frac{\partial}{\partial x_l}.$$

These can be expressed in terms of Christoffel symbols: We find, after short calculation,

$$R^l_{ijk} = \frac{\partial \Gamma^l_{jk}}{\partial x_i} - \frac{\partial \Gamma^l_{ik}}{\partial x_j} + \sum_r \Big( \Gamma^r_{jk} \Gamma^l_{ir} - \Gamma^r_{ik} \Gamma^l_{jr} \Big).$$

Recall that this complicated expression appeared in the proof of Gauss' theorem egregium in the curves and surfaces course, but it was somewhat unmotivated back then!

Since $R$ has four indices, the curvature tensor seems to give $(\dim M)^4$ invariants of a connection. In reality, the number is much smaller, due to symmetry properties of the curvature tensor. First of all, it is of course anti-symmetric in $X, Y$. More interesting is:

THEOREM 18.3 (Bianchi identity). *Suppose $\nabla$ has vanishing torsion. Then*

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0.$$

*That is, in local coordinates, $R^l_{ijk} + R^l_{jki} + R^l_{kij} = 0$.*

PROOF. We show that the left hand side vanishes at any given $p \in M$. Let $\mathrm{Exp}_p : B_r(0) \to M$ be the exponential map, where $r < i_p(M)$. Introduce normal coordinates on $U = \mathrm{Exp}_p(B_r(0)) = B_r(p)$. Then all Christoffel symbols $\Gamma^k_{ij}$ vanish at 0, and we have (at 0)

$$R^l_{ijk} + R^l_{jki} + R^l_{kij} = \frac{\partial \Gamma^l_{jk}}{\partial x_i} - \frac{\partial \Gamma^l_{ik}}{\partial x_j} + \frac{\partial \Gamma^l_{ki}}{\partial x_j} - \frac{\partial \Gamma^l_{ji}}{\partial x_k} + \frac{\partial \Gamma^l_{ij}}{\partial x_k} - \frac{\partial \Gamma^l_{kj}}{\partial x_i}.$$

In the torsion-free case, this vanishes since the Christoffel symbols are symmetric in the lower indices. □

EXERCISE 18.4. Give a coordinate-free proof of the Bianchi identity.

Suppose now that $g$ is a (pseudo-)Riemannian metric and $\nabla$ the corresponding Levi-Civita connection. For vector fields $X, Y, Z, W$ define the *curvature tensor* of $g$ by

$$R(X,Y,Z,W) = g(R(X,Y)Z, W).$$

In components $R_{ijkl} = R(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_l})$ we have,

$$R_{ijkl} = \sum_r R^r_{ijk} g_{rl}.$$

THEOREM 18.5. *The curvature tensor has the symmetry properties,*

$$R(X,Y,Z,W) = -R(Y,X,Z,W) = -R(X,Y,W,Z) = R(Z,W,X,Y)$$

*and*

$$R(X,Y,Z,W) + R(Y,Z,X,W) + R(Z,X,Y,W) = 0.$$

PROOF. The last identity is just re-stating the Bianchi identity. Anti-symmetry of $R(X,Y,Z,W)$ in the first two entries $X, Y$ is obvious from the definition. To prove anti-symmetry in the last two entries, it is enough to show that $R(X,Y,Z,Z) = 0$ for all $X, Y, Z$. We have,

$$\begin{aligned}
R(X,Y,Z,Z) &= g(\nabla_X \nabla_Y (Z), Z) - g(\nabla_Y \nabla_X (Z), Z) - g(\nabla_{[X,Y]} Z, Z) \\
&= X\, g(\nabla_Y(Z), Z) - g(\nabla_Y(Z), \nabla_X(Z)) - Y\, g(\nabla_X(Z), Z) + g(\nabla_X(Z), \nabla_Y(Z)) \\
&\quad -\tfrac{1}{2}[X,Y]\, g(Z,Z) \\
&= X\, g(\nabla_Y(Z), Z) - Y\, g(\nabla_X(Z), Z) - \tfrac{1}{2}[X,Y]\, g(Z,Z) \\
&= \tfrac{1}{2}\big(X(Y\, g(Z,Z)) - Y(X\, g(Z,Z)) - [X,Y]\, g(Z,Z)\big) \\
&= 0.
\end{aligned}$$

It remains to prove $R(X,Y,Z,W) = R(Z,W,X,Y)$. In fact, this is a consequence of the other symmetry properties, although in a rather non-obvious way. First, one adds the four equations obtained from the Bianchi identity $R(X,Y,Z,W) + R(Y,Z,X,W) + R(Z,X,Y,W) = 0$ by interchaning $W$ with $X,Y,Z$. This gives

$$\begin{aligned} & R(X,Y,Z,W) + R(Y,Z,X,W) + R(Z,X,Y,W) \\ + \quad & R(W,Y,Z,X) + R(Y,Z,W,X) + R(Z,W,Y,X) \\ + \quad & R(X,W,Z,Y) + R(W,Z,X,Y) + R(Z,X,W,Y) \\ + \quad & R(X,Y,W,Z) + R(Y,W,X,Z) + R(W,X,Y,Z) = 0. \end{aligned}$$

Using anti-symmetry in the first two entries and the last two entries, we obtain some cancellations and find,

$$R(W,X,Y,Z) + R(W,Y,Z,X) + R(W,Z,X,Y) = 0.$$

Using the Bianchi identity again, we have

$$\begin{aligned} R(Z,X,W,Y) &= -R(W,Z,X,Y) - R(X,W,Z,Y) \\ &= -R(W,Z,X,Y) - R(W,X,Y,Z) \\ &= R(W,Y,Z,X). \end{aligned}$$

EXERCISE 18.6. Prove anti-symmetry of $R(X,Y,Z,W)$ in $Z,W$ in local (normal) coordinates, similar to our proof of the Bianchi identity.

$\square$

## 19. Connections on vector bundles

In this section we define connections on vector bundles $E \to M$ be a vector bundle. (We are mainly interested in $E = TM$, but other bundles will appear as well.)

DEFINITION 19.1. A connection (covariant derivative) on $E$ is a bi-linear map

$$\nabla : \mathfrak{X}(M) \times \Gamma^\infty(E) \to \Gamma^\infty(E), (X,\sigma) \mapsto \nabla_X \sigma,$$

such that $\nabla$ is $C^\infty$-linear in the $X$ variable and

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

for all $f \in C^\infty(M)$, $X \in \mathfrak{X}(M)$, $\sigma \in \Gamma^\infty(E)$.

DEFINITION 19.2. The curvature operator corresponding to the connection $\nabla$ is the linear map $R(X,Y) : \Gamma^\infty(E) \to \Gamma^\infty(E)$

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

As for affine connections, the curvature operator $R(X,Y)$ is in fact a $C^\infty$-linear map, and moreover is $C^\infty$-linear in $X,Y$ also.

**19.1. Connections on trivial bundles.** Let us first consider the case of a trivial vector bundle, $E = M \times \mathbb{R}^k$. Let $e_1,\dots,e_k$ be the standard basis of $\mathbb{R}^k$. These define "constant" sections $\epsilon_1,\dots,\epsilon_k$ of $M \times \mathbb{R}^k$, and the most general section has the form,

$$\sigma = \sum_a \sigma_a \epsilon_a.$$

where the $\sigma_a$ are functions. It is immediate that

$$\nabla^0_X \sigma := \sum_a X(\sigma_a)\epsilon_a$$

defines a connection. This is called the *trivial* connection on the trivial bundle $E = M \times \mathbb{R}^k$. Now let $\nabla_X$ be *any* connection. Define a map

$$A : \mathfrak{X}(M) \to C^\infty(M, \mathrm{End}(\mathbb{R}^k)),\ X \mapsto A(X)$$

by

$$\nabla_X \sigma = \nabla^0_X \sigma + A(X)\sigma.$$

Thus $A(X)$ is a matrix-valued function on $M$, measuring the difference from the trivial connection. Letting $A_{ab}(X)$ be its components, we have

$$\nabla_X \sigma = \sum_a X(\sigma_a)\epsilon_a + \sum_{ab} A_{ab}(X)\sigma_b\, \epsilon_a.$$

That is,

$$(\nabla_X \sigma)_a = X(\sigma_a) + \sum_b A_{ab}(X)\sigma_b.$$

Notice that the map $X \mapsto A(X)$ is $C^\infty(M)$-linear. Conversely, every $C^\infty(M)$-linear map of this form defines a connection. That is:

Proposition 19.3. *The space of connections on a* trivial *bundle $E = M \times \mathbb{R}^k$ is in 1-1 correspondence with the space of $C^\infty(M)$-linear maps, $\mathfrak{X}(M) \to C^\infty(M, \mathrm{End}(\mathbb{R}^k)),\ X \mapsto A(X)$. Under this correspondence, the map $A$ defines the connection*

$$\nabla_X = \nabla^0_X + A(X).$$

*One calls $A$ the* connection 1-form *of the connection $\nabla$.*

Suppose now that $\epsilon'_a \in \Gamma^\infty(M, \mathbb{R}^k)$ is a new basis of the space of sections. That is,

$$\epsilon'_a = g_{ba}\epsilon_b$$

where the matrix-valued function $g$ with coefficients $g_{ab} \in C^\infty(M)$ is invertible everywhere.

Let $\sigma'_a$ denote the components of $\sigma$ in the new basis, i.e.

$$\sigma'_a = g_{ab}\sigma_b.$$

Define the connection 1-form $A'$ of $\nabla$ in the new basis by

$$\nabla_X \sigma = \sum_a (X(\sigma'_a) + \sum_b A'(X)_{ab}\sigma'_b)\epsilon'_a.$$

We have $\epsilon'_a = \sum_b (g^{-1})_{ba}\epsilon_b$, therefore

$$\begin{aligned} X(\sigma_c) + \sum_b A(X)_{cb}\sigma_b &= \sum_a (g^{-1})_{ca}(X(\sigma'_a) + \sum_b A'(X)_{ab}\sigma'_b) \\ &= \sum_a g^{-1}_{ca}(\sum_b g_{ab}X(\sigma_b) + \sum_b X(g_{ab})\sigma_b + \sum_b A'(X)_{ab}g_{bc}\sigma_c) \\ &= X(\sigma_c) + \sum_{ab}(g^{-1})_{ca}X(g_{ab}) + \sum_{abd} g^{-1}_{ca}A'(X)_{ab}g_{bd}\sigma_d. \end{aligned}$$

Comparing, we read off, using matric notation,

$$A(X) = g^{-1}A'(X)g + g^{-1}X(g),$$

or equivalently,

$$A'(X) = gA(X)g^{-1} - X(g)g^{-1}.$$

In the theoretical physics literature, connections are called *gauge fields*, and sections of (possibly trivial) bundles $E$ are called *particle fields*. The change of basis using $g$ is called a *gauge transformation*, and the above formula is called the gauge group action of $C^\infty(M, \mathrm{Gl}(k, \mathbb{R}))$.

EXERCISE 19.4. Show that the curvature operator $R(X, Y)$ on $E \times \mathbb{R}^k$ transforms according to

$$R'(X, Y) = g\,R(X, Y)\,g^{-1}.$$

Give a formula for $R$ in terms of connection 1-forms.

**19.2. Connections on non-trivial vector bundles.** This discussion carries over to more general vector bundles, as follows. Let $\mathrm{End}(E) \to M$ be the *endomorphism bundle* of $E$, with fibers $\mathrm{End}(E)_p = \mathrm{End}(E_p)$ the vector space[15] of endomorphisms $E_p \to E_p$. The space of sections of $\mathrm{End}(E)$ is isomorphic to the space of $C^\infty(M)$-*linear* endomorphisms of the vector space $\Gamma^\infty(E)$:

$$\Gamma^\infty(\mathrm{End}(E)) = \mathrm{End}_{C^\infty(M)}(\Gamma^\infty(E)).$$

A connection gives a linear map $\nabla_X \in \mathrm{End}(\Gamma^\infty(E))$, which is not $C^\infty(M)$-linear. However, the difference between any two connections is:

$$A(X) = \nabla'_X - \nabla_X \in \mathrm{End}_{C^\infty(M)}(\Gamma^\infty(E)) = \Gamma^\infty(\mathrm{End}(E)).$$

Conversely, if $\nabla_X$ is any connection, and is

$X \mapsto A(X) \in \Gamma^\infty(\mathrm{End}(E))$ is $C^\infty(M)$-linear, then $\nabla'_X = \nabla_X + A(X)$ defines a new connection. This proves half of:

PROPOSITION 19.5. *Every vector bundle $E$ admits a connection $\nabla$. The most general connection is $\nabla'_X = \nabla_X + A(X)$ for some $C^\infty(M)$-linear map, $A : \mathfrak{X}(M) \to \Gamma^\infty(\mathrm{End}(E))$.*

PROOF. Any local trivializations $E|_U \cong U \times \mathbb{R}^k$ defines a connection $\nabla_U$ on $E|_U$ coming from the trivial connection on $U \times \mathbb{R}^k$. Let $U_\alpha$ be a locally finite open cover of $M$, with local trivializations of $E|_{U_\alpha}$, and let $\nabla^\alpha$ be the corresponding local connections. Let $\chi_\alpha$ be a partition of unity, and define

$$\nabla_X(\sigma) = \sum_\alpha \chi_\alpha \nabla^\alpha_X(\sigma|_{U_\alpha}).$$

This has all the properties of a connection. □

Let $E|_{U_\alpha} \cong U_\alpha \times \mathbb{R}^k$ be a local trivialization. Thus $\nabla$ becomes a connection on $U_\alpha \times \mathbb{R}^k$, described by some $A_\alpha(X) \in C^\infty(U_\alpha, \mathrm{End}(\mathbb{R}^k))$. The maps $X \mapsto A_\alpha$ are called the *local connection 1-forms* for $\nabla$. If $U_\alpha$ is a coordinate chart, with local coordinates $x_1 \ldots, x_m$, $A_\alpha$ is described by $m$ matrix valued functions

$$A_\alpha(\frac{\partial}{\partial x_i}) \in C^\infty(U_\alpha, \mathrm{End}(\mathbb{R}^k)).$$

[15] In fact, each fiber $\mathrm{End}(E_p)$ is an algebra, and accordingly $\mathrm{End}(E)$ is an example of an *algebra bundle*.

The components

$$\Gamma^b_{ia} := (A_\alpha)_{ab}(\frac{\partial}{\partial x_i})$$

are also called the *Christoffel symbols* of the connection with respect to the given local coordinates.

**19.3. Constructions with connections.** Given a vector bundle $E$, let $E^*$ be its dual bundle. There is a natural pairing of the spaces of sections,

$$\langle\cdot,\cdot\rangle : \Gamma^\infty(E^*) \times \Gamma^\infty(E) \to C^\infty(M), \quad \langle\tau,\sigma\rangle_p := \langle\tau_p,\sigma_p\rangle \equiv \tau_p(\sigma_p).$$

In other words, $\Gamma^\infty(E^*)$ is identified with the space of $C^\infty$-linear maps $\Gamma^\infty(E) \to C^\infty(M)$.

Proposition 19.6 (Duals). *For any connection $\nabla$ on $E$, there is a unique connection $\nabla^*$ on $E^*$ with property,*

$$X\langle\tau,\sigma\rangle = \langle\nabla^*_X\tau,\sigma\rangle + \langle\tau,\nabla_X\sigma\rangle.$$

Proof. Try to *define* $\nabla^*$ by this equation:

$$\langle\nabla^*_X\tau,\sigma\rangle = X\langle\tau,\sigma\rangle - \langle\tau,\nabla_X\sigma\rangle.$$

For $f \in C^\infty(M)$ we have,

$$X\langle f\tau,\sigma\rangle - \langle f\tau,\nabla_X\sigma\rangle = \langle X(f)\tau,\sigma\rangle + f(X\langle\tau,\sigma\rangle - \langle\tau,\nabla_X\sigma\rangle)$$

Showing that $\nabla^*_X(f\tau) = X(f)\tau + f\nabla^*_X\tau$ as desired. □

If $E, E'$ are two vector bundles over $M$, we can form the direct sum $E \oplus E'$, with

$$\Gamma^\infty(E \oplus E') = \Gamma^\infty(E) \oplus \Gamma^\infty(E').$$

Proposition 19.7 (Direct sums). *If $\nabla$ is a connection on $E$ and $\nabla'$ a connection on $E'$, there is a unique connection $\nabla \oplus \nabla'$ on $E \oplus E'$ such that*

$$(\nabla \oplus \nabla')_X(\sigma \oplus \sigma') = \nabla_X\sigma \oplus \nabla'_X\sigma'.$$

Finally, recall that if $E$ is a vector bundle over $M$, and $F \in C^\infty(N, M)$ a smooth map from a manifold $N$, we define a pull-back bundle $F^*E$ with fibers $(F^*E)_q = E_{F(q)}$. Its space of sections $\Gamma^\infty(F^*E)$ is generated (as a $C^\infty(N)$-module) by the subspace $F^*\Gamma^\infty(E)$.

Proposition 19.8. *Let $E \to M$ be a vector bundle with connection $\nabla$, and $F \in C^\infty(N, M)$. Then there is a unique connection $F^*\nabla$ such that for all $\sigma \in \Gamma^\infty(E)$, $q \in N$, $w \in T_qN$*

$$(F^*\nabla)_w(F^*\sigma) = \nabla_{T_qF(w)}\sigma.$$

Proof. Exercise. □

The pull-back connection $F^*\nabla$ can be desribed in terms of connection 1-forms: If $E|_U \cong U \times \mathbb{R}^k$ is a local trivialization of $E$, and $X \mapsto A_{ab}(X)$ the connection 1-form of $\nabla$ in terms of this local trivialization. Then we obtain a local trivialization $F^*E|_{F^{-1}(U)} \cong F^{-1}(U) \times \mathbb{R}^k$, with connection 1-forms given by the pull-back forms, $F^*A_{ab}$. [16]

[16] Recall that $C^\infty(M)$- linear maps $\mathfrak{X}(M) \to C^\infty(M)$ are identified with sections of $T^*M$, i.e. 1-forms, and that there is a natural pull-back map $F^*\Gamma^\infty(T^*M) \to \Gamma^\infty(T^*N)$ given by $(F^*\alpha)_q = (T_qF)^*\alpha_{F(q)}$.

**19.4. Parallel transport.** Suppose $E$ is a vector bundle over $M$ with connection $\nabla$, and $\gamma : J \to M$ is any smooth curve. Sections of the pull-back bundle $\gamma^*E$ are called *sections of* $E$ *along* $\gamma$. A connection $\nabla$ on $E$ induces a pull-back connection $\gamma^*\nabla$ on $\gamma^*E$, and one can define a *covariant derivative along* $\gamma$ by

$$\frac{D}{Dt} : \Gamma^\infty(\gamma^*E) \to \Gamma^\infty(\gamma^*E), \quad \frac{D\sigma}{Dt} := (\gamma^*\nabla)_{\frac{\partial}{\partial t}}\sigma.$$

A section $\sigma$ along $\gamma$ is called *parallel* if $\frac{D\sigma}{Dt} = 0$.

Suppose $E|_U = U \times \mathbb{R}^k$ is a local trivialization of $E$ with $\gamma(t) \in U$, given by a basis $\epsilon_1, \ldots, \epsilon_k \in \Gamma^\infty(U, E|_U)$ of the space of sections. Then we can write

$$\sigma(t) = \sum_a \sigma_a(t)(\epsilon_a)_{\gamma(t)} \in E_{\gamma(t)},$$

and the components of the covariant derivative is given by the formula,

$$\left(\frac{D\sigma}{Dt}\right)_a = \frac{d\sigma_a}{\mathrm{d}t} + \sum_b A_{ab}(\dot\gamma)\sigma_b(t).$$

Furthermore, if $U$ is the domain of a coordinate chart, defining local coordinates $x_1, \ldots, x_k$, and

$$\Gamma^b_{ia} = A_{ab}(\frac{\partial}{\partial x_i})$$

are the corresponding Christoffel symbols, the formula becomes,

$$\left(\frac{D\sigma}{Dt}\right)_a = \frac{d\sigma_a}{\mathrm{d}t} + \sum_{ib} \Gamma^b_{ia}\dot x_i \sigma_b(t).$$

As for affine connections, one shows that for any given $\sigma_{t_0} \in E_{\gamma(t_0)}$, there is a unique parallel section $\sigma(t)$ along $\gamma$ with initial value $\sigma(t_0) = \sigma_{t_0}$. In this way, connections $\nabla$ define parallel transport in vector bundles.

There is a more geometric way of understanding paralel transport on a vector bundle $\pi : E \to M$. Consider the tangent map $T\pi : TE \to TM$. Its kernel at $u \in E_p$ is

$$\ker(T_u\pi) = T_u(E_p),$$

the tangent space at $u$ to the fiber $E_p$. It is called the *vertical subspace*

$$V_uE = T_u(E_p) \subset T_uE.$$

Note that since $E_p$ is a vector space, $T_u(E_p) \cong E_p$. This means that we have a natural isomorphism, $VE = \pi^*E$ (the pull-back of $E$ to a vector bundle over $E$.).

Since the map $T_u\pi : T_uE \to T_pM$ is clearly onto, it we have

$$T_uE/V_uE = T_pM.$$

It turns out that every connection $\nabla$ defines a complementary *horizontal subspace* $H_uE \subset T_uE$, with

$$T_uE = V_uE \oplus H_uE.$$

In fact, $HE$ is a vector subbundle of $TE$, called the horizontal bundle, and $TE = VE \oplus HE$.

In a coordinate free way, one may define $H_uE$ as follows:

THEOREM 19.9. *Let $\nabla$ be a connection on $E$. Given $u \in E_p$, there exists a section $\sigma \in \Gamma^\infty(E)$ with $\sigma_p = e$ and $(\nabla\sigma)_p = 0$. The image*

$$H_uE := T_p\sigma(T_pM)$$

*is independent of the choice of $\sigma$.*

EXERCISE 19.10. 1. Proof this theorem.

2. Given an alternative definition of $H_uE$ in local coordinates and using a local trivialization of $E$: Show that the horizontal space in $E = U \times \mathbb{R}^k$ is spanned by all

$$\frac{\partial}{\partial x_i} - \sum_a \Gamma^b_{ia} u_b \epsilon_a(p)$$

Note that it is impossible, in general, to choose $\sigma$ with $\sigma_p = u$ and $\nabla\sigma = 0$ everywhere: This is related to the problem that in general, distributions of rank $\geq 2$ need not be integrable.

One can characterize parallel transport in terms of the horizontal bundle $HE \subset TE$ as follows: For any curve $\gamma(t)$ in $M$, and any given $u \in E_{\gamma(t_0)}$, there is a unique curve $\sigma(t)$ in $E$ such that $\sigma(t_0) = u$ and

$$\pi(\sigma(t)) = \gamma(t), \quad \dot\sigma \in H_{\sigma(t)}E$$

for all $t$. The curve $\sigma(t)$ is called the *horizontal lift* of $\gamma$. Note that a curve $\sigma(t)$ projecting to $\gamma(t)$ is the same thing as a section of $E$ along $\gamma$.

The splitting $TE = HE \oplus VE = \pi^*TM \oplus VE$ given by $\nabla$ defines a horizontal lift of vector fields:

$$\mathrm{Lift}_\nabla : \mathfrak{X}(M) \to \mathfrak{X}(E).$$

Here $\mathrm{Lift}_\nabla(X)_u$ is the unique tangent vector in $H_uE$ projecting to $X_p$. The horizontal lifts of integral curves of $X$ are integrla curves of its horizontal lift $\mathrm{Lift}_\nabla(X)$. Note that by construction,

$$\mathrm{Lift}_\nabla(X) \sim_\pi X.$$

Hence if $X, Y$ are two vector fields,

$$[\mathrm{Lift}_\nabla(X), \mathrm{Lift}_\nabla(Y)] \sim_\pi [X, Y].$$

This shows that the vector field $[\mathrm{Lift}_\nabla(X), \mathrm{Lift}_\nabla(Y)] - \mathrm{Lift}_\nabla([X, Y])$ must be *vertical.* That is, it is a section of $VE = \pi^*E$. What is this section?

THEOREM 19.11. *For any $u \in E_p$, and any $X, Y$, we have*

$$[\mathrm{Lift}_\nabla(X), \mathrm{Lift}_\nabla(Y)]_u - \mathrm{Lift}_\nabla([X, Y])_u = R(X, Y)_p\, u.$$

THEOREM 19.12. *The following are equivalent:*

(a) *The curvature $R$ of $\nabla$ vanishes.*
(b) *Horizontal lift $\mathfrak{X}(M) \to \mathfrak{X}(E)$ is a Lie algebra homomorphism.*
(c) *The horizontal distribution $HE \subset TE$ is integrable.*
(d) *Parallel transport along paths is invariant under homotopies leaving the end points fixed.*

We leave the proofs as exercises, or to be looked up in textbooks.