Differential Geometry

Will J. Merry

ETH Zürich

Lecture notes for the course Differential Geometry given in the academic year 2018-2019. The Problem Sheet solutions were written by Mads Bisgaard, Giada Franz, Francesco Palmurella, Alessio Pellegrini, Alessandro Pigati and Alexandre Puttick. Please let me know by email at merry@math.ethz.ch if you spot any typos!

Smooth manifolds

Let us begin with a short history lesson on how you learned to identify (continuously) differentiable functions.

(i) **(High school)** A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous if its graph doesn't have any jumps. The derivative $f'(x)$ at a point x is the slope of the graph of $f(x) = y$ at the point x.

(ii) **(First class in Analysis)** The (ε, δ) definition of continuity. A function $f\colon \mathbb{R} \to \mathbb{R}$ is differentiable at the point x if the limit

$$\lim_{u \to 0} \frac{f(x+u) - f(x)}{u}$$

exists. This limit is denoted by $f'(x)$. The function f is continuously differentiable if $x \mapsto f'(x)$ is itself a continuous function.

(iii) **(Second class in Analysis)** Now you learned how to handle functions with more than one variable. Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^k$ is a continuous function. Then f is differentiable at $x \in \mathbb{R}^n$ if there exists a *linear* map $T\colon \mathbb{R}^n \to \mathbb{R}^k$ (that is, a $k \times n$ matrix) such that

$$\lim_{|u| \to 0} \frac{|f(x+u) - f(x) - Tu|}{|u|} = 0. \tag{1.1}$$

We denote T by $Df(x)$. It is the matrix of partial derivatives of $f = (f^1, \dots, f^k)$ at the point $x = (x^1, \dots, x^n)$:

$$Df(x) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x) & \cdots & \frac{\partial f^1}{\partial x^n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^k}{\partial x^1}(x) & \cdots & \frac{\partial f^k}{\partial x^n}(x) \end{pmatrix}$$

Of course, this reduces to the same definition as before if $n = k = 1$, since a 1×1 matrix is just a number, and in this case $Df(x)$ is simply multiplication by the number $f'(x)$. As before, the function f is continuously differentiable if $x \mapsto Df(x)$ is a continuous function (this is now a function $\mathbb{R}^n \to \{k \times n \text{ matrices}\} \cong \mathbb{R}^{kn}$).

(iv) **(First class in topology)** Suppose now X and Y are topological spaces and $f\colon X \to Y$ is a function. You learned that f is continuous if $f^{-1}(U)$ is an open set in X for every open set U in Y. If X and Y are metric spaces then this reduces to the old (ε, δ) definition of continuity. But how does one define differentiability in this setting? Equation (1.1) *does not make sense* any more, since in an arbitrary topological space one cannot simply "add" points, and there is no such thing as a "linear" map!

Here endeth the history lesson. The tl;dr version is:

- It's easy to differentiate functions on Euclidean spaces (or more generally, on vector spaces[1]).
- Most topological spaces are not vector spaces.
- Bummer.

Indeed, this is a real shame. Measuring the rate at which things change—that is, differentiating them—is absolutely crucial to all applications of mathematics (and is arguably the single most important concept in theoretical physics). However most "real life" systems are *not* defined on open sets in vector spaces (the whole point of your topology course was to introduce classes of spaces appropriate for such models).

This is where differential geometry comes in. Our first aim is to define a special type of topological space, called a **smooth manifold**, on which it *is* possible to make sense of differentiating a continuous function. The definition of a smooth manifold will:

- Include open sets in vector spaces as a special case.
- Be sufficiently general so that the topological spaces that occur in "real life" systems (in theoretical physics, economics, computer science, robotics, genetics, cooking etc) are smooth manifolds.

So let's get started.

In fact, we will define smooth manifolds in two stages. We will first define a **topological manifold**, which is a topological space that locally resembles Euclidean space. We will then endow a topological manifold with an additional piece of data called a **smooth structure**. The smooth structure is what will allow us to actually go ahead and differentiate things. A topological manifold equipped with a smooth structure is then called a **smooth manifold**.

Let us first recall a few elementary concepts from point-set topology[2].

DEFINITION 1.1. Let X be a topological space. We say that X is **Hausdorff** if for every pair $x \neq y$ of points in X, there are open subsets $U, V \subset X$ such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

Any metric space is Hausdorff. A topological space X is said to be **connected** if it is not the disjoint union of nonempty open sets. A topological space X is said to be **path connected** if for any two points $x, y \in X$ there exists a continuous map $\gamma \colon [0,1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$. A path connected space is connected, but the converse need not hold. In general any topological space can be decomposed into its **connected components** (resp. **path components**), where

[1] In this course, *all* vector spaces are implicitly assumed to be finite-dimensional real vector spaces, unless otherwise specified.

[2] This is hopefully revision for most of you. However it is not the end of the world if you are not that familiar (or have forgotten) most of this material, as we will make very little use of it throughout the course.

the connected component (resp. path component) containing a given point x is the union of all the connected (resp. path connected) sets containing x.

Recall that an **open cover** of a topological space X is a collection $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of open subsets of X, where $\mathtt{A}$ is some index set, such that $X = \bigcup_{\mathtt{a}\in\mathtt{A}} U_{\mathtt{a}}$. If the index set $\mathtt{A}$ is a finite set, we say that the open cover is a **finite cover**. A **subcover** of an open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ consists of a subset $\mathtt{A}' \subset \mathtt{A}$ such that the collection $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}'\}$ is still an open cover.

Definition 1.2. Let X be a topological space. We say that X is **compact** if every open cover has a finite subcover.

As you hopefully remember from your point-set topology course, compact spaces are typically the most "useful" class of topological spaces, in the sense that many powerful theorems only hold for compact spaces. Unfortunately, since we want manifolds to include Euclidean spaces as a special case, we cannot require manifolds to be compact (indeed, a subset $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded—this is the Heine-Borel theorem.)

We will therefore impose a weaker condition, which requires two more preliminary definitions about covers. Suppose $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is an open cover. A **refinement** is another open cover $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ with the property that for every $\mathtt{b} \in \mathtt{B}$ there exists $\mathtt{a} \in \mathtt{A}$ such that $V_{\mathtt{b}} \subset U_{\mathtt{a}}$. Next, an open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of X is said to be **locally finite** if for every $x \in X$ there exists a neighbourhood[3] W of x such that the set $\{\mathtt{a} \in \mathtt{A} \mid U_{\mathtt{a}} \cap W \neq \emptyset\}$ is a finite set.

Definition 1.3. A topological space X is said to be **paracompact** if every open cover has a locally finite refinement.

Thus compact spaces are obviously paracompact, but the latter is more general. For instance, $\mathbb{R}^n$ is paracompact, but as we have just observed, not compact. In fact, the following result holds.

Theorem 1.4. *Every metric space is paracompact.*

Although Theorem 1.4 is not too hard to prove, we will not do so, as it involves ideas from outside the course. Now let us introduce the final (and most important) concept needed to define topological manifolds.

Definition 1.5. A topological space X is said to be **locally Euclidean of dimension** n if for every point $x \in X$, there exists a neighbourhood U of x, an open set $O \subset \mathbb{R}^n$, and a homeomorphism $\sigma \colon U \to O$.

Let

$$B^n := \{x \in \mathbb{R}^n \mid |x| < 1\}$$

denote the *open* unit ball in $\mathbb{R}^n$. If $O \subset \mathbb{R}^n$ is open then there exists a subset $O' \subset O$ such that O' is homeomorphic to B^n. Moreover since B^n is homeomorphic to $\mathbb{R}^n$ itself, we can equivalently define a locally Euclidean space of dimension n to be a topological space X with the property that every $x \in X$ admits a neighbourhood U which is homeomorphic to $\mathbb{R}^n$.

[3] I use the convention that a neighbourhood of a point in a topological space is an *open* set containing that point.

Before stating the next remark, let me introduce a convention that will hold throughout the entire course: *anything marked with a* (♣) *is non-examinable.* There are various reasons for marking something with a (♣):

- it is only tangentially related to the course,
- it is rather technical or difficult,
- it is just a sketch,
- it requires more background knowledge (eg. algebraic topology, functional analysis, etc) than the rest of the course assumes.

In any case, you are welcome to ignore anything marked with a (♣). This holds for both the Lecture Notes and the Problem Sheets.

(♣) Remark 1.6. Suppose $n \neq k$ are two non-negative integers. Is it possible for a topological space to be locally Euclidean of dimension n *and* locally Euclidean of dimension k? Equivalently, is $\mathbb{R}^n$ homeomorphic to $\mathbb{R}^k$ for $n \neq k$? The answer to this is "no", but this is surprisingly difficult to prove. This result is called the *Invariance of Domain Theorem*, and was first proved by Brouwer in 1912. The easiest proof uses tools from algebraic topology. I proved it last year in my course here.

Let us now finally give the first key definition of the course.

Definition 1.7. A topological space M is called a **topological manifold of dimension** n if:

(i) M is locally Euclidean of dimension n,

(ii) M is Hausdorff and has at most countably many connected components,

(iii) M is paracompact.

We refer to n as the **dimension** of M.

As we have already alluded to, the most important part of the definition is the locally Euclidean part. The Hausdorff condition is included to rule out pathologies. The requirement that M only has countably many components is almost never needed—in fact, in this course only one important theorem will use this hypothesis (namely, Sard's Theorem 5.17). Note that this tells us that a countable collection of points is a zero-dimensional manifold (with the discrete topology), but an uncountable collection is not.

The third condition guarantees the existence of *partitions of unity*, which are an important technical tool we will discuss in Lecture 3 (cf. Theorem 3.13). The dimension of M is well-defined due to Remark 1.6. In general the phrase "topological manifold" means a topological manifold of some unspecified dimension n.

Remark 1.8. If X is a topological space then a **basis** for the topology on X is a set $\mathcal{B}$ of open sets of X with the property that every open set in X is a union of sets in $\mathcal{B}$. A topological space is said to be **second countable** if it admits a countable basis. Many authors define topological manifolds as locally Euclidean second countable Hausdorff topological spaces. In fact this is the same as Definition 1.7, due to the following result: *a Hausdorff locally Euclidean topological space is second countable if and only if it is paracompact and has at most countably many connected components*. I prefer mentioning paracompactness explicitly in the definition (over second countability) since paracompactness is what we will actually use when constructing partitions of unity in Lecture 3.

REMARK 1.9. Topological manifolds enjoy many nice point-set topological properties. Let us go through some of them. Do not worry too much about the technical terms—we will never really use them in the course, except in passing.

(i) A topological space X is said to be **locally compact** if for every point $x \in X$ there exists a compact set K and a neighbourhood U of x such that $U \subset K$. If the topological space is Hausdorff, this is equivalent to asking that every point has a neighbourhood with compact closure. Any locally Euclidean Hausdorff space (and hence any topological manifold) obviously has this property.

(ii) A topological space is said to be **Lindelöf** if every open cover has a *countable* subcover. Any locally compact paracompact space with at most countably many components is Lindelöf, and hence the same is true of any topological manifold.

(iii) A topological space X is **locally path connected** if for every point $x \in X$ and every neighbourhood U of x, there exists a path connected neighbourhood V of x with $V \subset U$. A locally Euclidean space is obviously locally path-connected. For a locally path connected space, the path components and the connected components coincide. Thus in particular, a topological manifold M is connected if and only if it is path connected.

(iv) Every paracompact Hausdorff space is a **normal** topological space. This means that given any two closed disjoint subsets K_1, K_2 of M there are open sets U_1, U_2 of M such that $K_i \subset U_i$ for $i = 1, 2$ and $U_1 \cap U_2 = \emptyset$. Thus certainly any topological manifold is normal.

(v) Every topological manifold M is **metrisable**. That is, there exists some metric on M that induces the given topology on M. Thus one can always view a manifold as a metric space—we will carry out this construction explicitly in Lecture 52. In Lecture 6 we will give a proof of this fact for *smooth* manifolds (cf. Definition 1.18 below)—a similar argument works for topological manifolds too, but it is a bit more involved.

EXAMPLE 1.10. $\mathbb{R}^n$ is trivially a topological manifold of dimension n. More generally, any n-dimensional vector space is a topological manifold of dimension n. Similarly any non-empty open subset of a topological manifold of dimension n is also a topological manifold of dimension n.

We will see more interesting examples later in this lecture, but let us briefly note a *non*-example.

Example 1.11. The *closed* unit ball

$$D^n := \{x \in \mathbb{R}^n \mid |x| \leq 1\}$$

is *not* a topological manifold of dimension n. It is an illustrative exercise to try and work out why. In fact, D^n is an example of a more general concept of a **manifold with boundary** that we will come back to later in Lecture 21.

Let us now get back to the point of view discussed at the beginning of the lecture: we are trying to develop a class of topological spaces for which it is possible to differentiate functions on. One might naively believe that the locally Euclidean condition built into the definition of a topological manifold is enough. Indeed, to check whether a function $f\colon \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at a point $x \in \mathbb{R}^n$, we need only examine f in a small neighbourhood of x—this is clear from (1.1). Thus if we are given a continuous map between two topological manifolds, we can locally view it as a continuous map between two Euclidean spaces, and thus we could conceivably say our original map is differentiable if this local map is. But herein lies a problem: a topological manifold is only *homeomorphic* to Euclidean space, and a different choice of homeomorphism might affect whether the local map is differentiable or not.

The solution to this is to introduce more structure. Before doing so, let us recall the chain rule for continuously differentiable functions between Euclidean spaces. We will give two different versions: one for the total differential $Df(x)$ (the matrix) and one for the partial derivatives $\frac{\partial f^i}{\partial x^j}$.

Proposition 1.12 (The Chain Rule). *Let $O \subset \mathbb{R}^n$, $\Omega \subset \mathbb{R}^k$ be open sets. Let $f\colon O \to \mathbb{R}^k$ and $g\colon \Omega \to \mathbb{R}^l$ be continuously differentiable functions satisfying $f(O) \subset \Omega$.*

(i) *The function $g \circ f$ is also continuously differentiable, and its derivative at the point x is given by*

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x).$$

(ii) *Write $x = (x^1, \dots, x^n)$ for the coordinates on $\mathbb{R}^n$ and $y = (y^1, , \dots, y^k)$ for the coordinates on $\mathbb{R}^k$, and write $f = (f^1, \dots, f^k)$ and $g = (g^1, \dots, g^l)$. Then the partial derivatives of $g \circ f$ are given by*

$$\frac{\partial (g^i \circ f)}{\partial x^j}(x) = \sum_{r=1}^{k} \frac{\partial g^i}{\partial y^r}(f(x)) \frac{\partial f^r}{\partial x^j}(x), \qquad \text{for all } 1 \leq i \leq l, 1 \leq j \leq n.$$

We now define higher order derivatives.

Definition 1.13. Let $O \subset \mathbb{R}^n$ and $\Omega \subset \mathbb{R}^k$ be open sets and suppose $f\colon O \to \Omega$ is a differentiable map. We say that f is of class C^r if each partial derivative $\frac{\partial f^i}{\partial x^j}$ is a $(r-1)$-times continuously differentiable function. We say that f is **smooth** or of class C^∞ if f is of class C^r for every $r \geq 1$. If f is both smooth and bijective and the inverse function is also smooth then we say that f is a **diffeomorphism**.

It follows from part (ii) of Proposition 1.12 that the composition of smooth functions defined on open sets in Euclidean spaces is again a smooth function.

Remark 1.14. If f is a diffeomorphism then necessarily $n = k$. This follows immediately from part (i) of Proposition 1.12, which tells us that if f is a diffeomorphism then $Df(x)$ is an invertible matrix. (Its inverse is given by $D(f^{-1})(f(x))$.) A $k \times n$ matrix can only be invertible if $n = k$. Thus in particular $\mathbb{R}^n$ cannot be *diffeomorphic* to $\mathbb{R}^k$ for $n \neq k$ (compare to Remark 1.6).

We these preliminaries in hand, let us get started on the definition of a smooth manifold.

Definition 1.15. Let M be a topological manifold of dimension n. A **smooth atlas** on M is a collection

$$\Sigma = \{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$$

where $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is an open cover of M, each $O_{\mathtt{a}}$ is an open set in $\mathbb{R}^n$, and each $\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}}$ is a homeomorphism such that the following *compatibility condition* is satisfied: Suppose $\mathtt{a}, \mathtt{b} \in \mathtt{A}$ are such that $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$. Then the composition (often called the *transition map*)

$$\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1} \colon \sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}}) \to \sigma_{\mathtt{b}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$$

should be a diffeomorphism. This makes sense, since both $\sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$ and $\sigma_{\mathtt{b}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$ are open subsets of $\mathbb{R}^n$. We call the maps $\sigma_{\mathtt{a}}$ the **charts** of the atlas Σ.

We say that two smooth atlases Σ_1 and Σ_2 are **equivalent** if their union is also a smooth atlas, that is, if given any chart σ of Σ_1 and any chart τ of Σ_2 such that the domains of σ and τ intersect, the composition $\tau \circ \sigma^{-1}$ is also a diffeomorphism. It is immediate that this notion defines an equivalence relation on the set of smooth atlases on a given topological manifold.

Definition 1.16. A **smooth structure** on a topological manifold is an equivalence class of smooth atlases.

Remark 1.17. Given an equivalence class of smooth atlases, there is a unique *maximal* smooth atlas in that class (simply take the union of all the atlases in the given equivalence class). Thus there is a 1-1 correspondence between smooth structures and maximal smooth atlases. Since dealing with equivalence relations can be tedious, it is usually more convenient to regard a smooth structure as a maximal smooth atlas, and we will do so without further comment.

We now finally arrive at the main definition of this first lecture.

Definition 1.18. A **smooth manifold of dimension** n is a pair (M, Σ) where M is a topological manifold of dimension n and Σ is a smooth structure on M.

Since a smooth atlas is contained in a unique maximal smooth atlas, it is sufficient when defining a smooth manifold to specify a smooth atlas on the underlying topological manifold. Whenever possible we will omit the Σ from the notation and just write M. For smooth manifolds the fact that the dimension is well-defined is much easier than for topological manifolds (we only need Remark 1.14, which does not require any algebraic topology).

EXAMPLE 1.19. The standard smooth structure on $\mathbb{R}^n$ is the one containing the smooth atlas consisting of exactly one chart: the identity map $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$. The reason for the word "standard" will become clear by the end of the lecture. More generally, if V is any n-dimensional real vector space, then the standard smooth structure on V is the one induced by the smooth atlas consisting of a single chart $T\colon V \to \mathbb{R}^n$, where T is some linear isomorphism. (*Exercise*: Why is this independent of the choice of T?)

Just as with topological manifolds, an open subset of a smooth manifold is also a smooth manifold:

LEMMA 1.20. *Let M be a smooth manifold of dimension n and let $W \subset M$ be a non-empty open set. Then W naturally inherits the structure of a smooth manifold of dimension n.*

Proof. We have already remarked in Example 1.10 that W is a topological manifold of dimension n. Let $\Sigma = \{\sigma_{\mathtt{a}}\colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be a smooth atlas on M. Then

$$\{\sigma_{\mathtt{a}}|_{W\cap U_{\mathtt{a}}}\colon W \cap U_{\mathtt{a}} \to \sigma_{\mathtt{a}}(W \cap U_{\mathtt{a}}) \subset O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$$

is a smooth atlas for W. ■

Thus any open subset of a vector space is a smooth manifold. Let us now consider a slightly less trivial example. Recall we denote by S^n denote the unit sphere:

$$S^n := \left\{x \in \mathbb{R}^{n+1} \mid |x| = 1\right\}.$$

PROPOSITION 1.21. *The sphere S^n is a compact smooth manifold of dimension n.*

Proof. We give S^n the subspace topology from $\mathbb{R}^{n+1}$. Then S^n is a metric space (being a subset of a metric space), and hence is Hausdorff and paracompact (Theorem 1.4). Moreover S^n is connected. We will directly exhibit a smooth atlas on S^n (thus proving at the same time that S^n is a topological manifold). Let $x_N = (0, \dots, 0, 1)$ denote the "north pole" and let $x_S := (0, \dots, , 0, -1)$ denote the "south pole". Let $U_N = S^n \setminus \{x_N\}$ and $U_S := S^n \setminus \{x_S\}$. Then $\{U_N, U_S\}$ is an open cover of S^n. Define charts

$$\sigma_N\colon U_N \to \mathbb{R}^n, \qquad \sigma_N(x^1, \dots, x^{n+1}) := \frac{1}{1 - x^{n+1}}(x^1, \dots, x^n)$$

and

$$\sigma_S\colon U_S \to \mathbb{R}^n, \qquad \sigma_S(x^1, \dots, x^{n+1}) := \frac{1}{1 + x^{n+1}}(x^1, \dots, x^n).$$

The maps σ_N and σ_S are stereographic projection from the north and south pole respectively. Both the transition maps

$$\sigma_N \circ \sigma_S^{-1}\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\} \qquad \text{and} \qquad \sigma_S \circ \sigma_N^{-1}\colon \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}$$

are given by

$$(y^1, \dots, y^n) \mapsto \frac{1}{\sum_{i=1}^n (y^i)^2}(y^1, \dots, y^n)$$

which is obviously a diffeomorphism. Thus we have defined a smooth atlas on S^n. We refer to this smooth structure as the standard smooth structure on S^n. ■

All we really needed to do in the previous proof was check differentiability of the transition function $\sigma_N \circ \sigma_S^{-1}$. This is because (as a subset of $\mathbb{R}^{n+1}$), S^n already carried a nice topology. Sometimes however we will want to build a smooth manifold "from scratch". For this, the next result is very useful.

PROPOSITION 1.22 (Constructing smooth manifolds). *Let M be a set. Suppose we are given a collection $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of subsets of M together with bijections $\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \subset \mathbb{R}^n$, where $O_{\mathtt{a}}$ is an open subset of $\mathbb{R}^n$. Assume in addition that:*

- *For any $\mathtt{a}, \mathtt{b} \in \mathtt{A}$, $\sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$ is open in $\mathbb{R}^n$.*
- *If $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$, the map $\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1} \colon \sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}}) \to \sigma_{\mathtt{b}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$ is a diffeomorphism.*
- *Countably many of the $U_{\mathtt{a}}$ cover M.*
- *If $x \neq y$ are points in M then either there exists $\mathtt{a}$ such that x and y belong to $U_{\mathtt{a}}$, or there exists $\mathtt{a}, \mathtt{b}$ with $U_{\mathtt{a}} \cap U_{\mathtt{b}} = \emptyset$ such that $x \in U_{\mathtt{a}}$ and $y \in U_{\mathtt{b}}$.*

Then M has a unique smooth manifold structure for which $\{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is a smooth atlas.

The proof is essentially trivial: we simply took the definition of a smooth manifold and inserted it into the hypotheses.

Proof. Define a topology on M by declaring all the $\sigma_{\mathtt{a}}$ to be homeomorphisms. That this is well-defined topology follows from the fact that the $\sigma_{\mathtt{a}}$ are bijections, together with the first two bullet points. The locally Euclidean property is then immediate. The last two bullet points guarantee this topology is Hausdorff, has countably many components, and is paracompact, thus turning M into a topological manifold. Finally the fact that $\{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is a smooth atlas on M is clear from the second bullet point. ■

REMARK 1.23. Historically, a manifold M (smooth or topological) was called *open* if M was non-compact and *closed* if M was compact. This however is bad terminology for two reasons:

(i) Thought of as an abstract topological space, every manifold is both open and closed! (This is true of any topological space.)

(ii) If however our given manifold M is a subspace of a larger space N, then it does make sense to ask whether M is open or closed in the subspace topology of N. For example, the unit ball B^n is open in $\mathbb{R}^n$ and the unit sphere S^n is closed in $\mathbb{R}^{n+1}$. Historically, all manifolds were thought of as subspaces—actually *submanifolds*[4]—of some Euclidean space $\mathbb{R}^k$, and in fact any manifold can be embedded inside Euclidean space[5]. However even then the terminology "open" and "closed" does not make sense! For instance, if we identify $\mathbb{R}^2$ with the set of points in $\mathbb{R}^3$ whose last coordinate is zero then $\mathbb{R}^2$ is closed as a subspace of $\mathbb{R}^3$, but $\mathbb{R}^2$ is not compact as a manifold.

[4]We will define submanifolds precisely in Lecture 5.

[5]In the smooth case, this is known as the *Whitney Embedding Theorem*, which we will prove in Lecture 6.

Thus throughout this course, we will only use the words "open" and "closed" in their topological context (i.e. to speak of open sets and closed sets). If we wish to indicate a given manifold is compact, we will use the rather more logical terminology "compact manifold".

The only caveat to this is that when we define (both smooth and topological) manifolds with boundary later on (Lecture 21), we will need to differentiate between the terms "compact manifold with boundary" and "compact manifold without boundary". Indeed, as we have already mentioned, the closed unit ball D^n is an example of a compact smooth manifold with boundary.

On Problem Sheet A there are many more examples (and non-examples) of smooth manifolds for you to play with. Going back to the general theory, we have now achieved the goal we set out at the beginning of the lecture: to come up with an appropriate class of topological spaces for which it makes sense to say whether a map is differentiable or not.

Definition 1.24. Let $\varphi\colon M \to N$ be a continuous map between two smooth manifolds. We say that φ is of class C^r if for every point $x \in M$, if $\sigma\colon U \to O$ is any chart on M with $x \in U$ and $\tau\colon V \to \Omega$ is any chart on N with $\varphi(x) \in V$, the composition

$$\tau \circ \varphi \circ \sigma^{-1}\colon \sigma(U \cap \varphi^{-1}(V)) \to \tau(\varphi(U) \cap V)$$

is of class C^r. If φ is of class C^r for all r then we say φ is **smooth** (or of class C^∞). If φ is smooth and bijective and the inverse function $N \to M$ is also smooth then φ is said to be a **diffeomorphism**.

It follows from the definition of smooth atlases that it does not matter which charts we use to check differentiability (i.e. we could replace "any chart" with "every chart" above).

Example 1.25. If (M, Σ) is a smooth manifold and $\sigma\colon U \to O$ belongs to Σ, then if we think of U and O as smooth manifolds in their own right (using Lemma 1.20 and Example 1.19) then σ is a diffeomorphism.

Similarly if $W \subset M$ is any open set (endowed with the smooth structure from Lemma 1.20) then the inclusion map $\imath\colon W \hookrightarrow M$ is a smooth map.

The next result also follows immediately from the chain rule in Euclidean spaces (Proposition 1.12).

Proposition 1.26. *Let M, N and L be smooth manifolds, and suppose $\varphi\colon M \to N$ and $\psi\colon N \to L$ are smooth maps. Then $\psi \circ \varphi\colon M \to L$ is smooth.*

Proof. Let $x \in M$. Let σ be a chart on M containing x, let τ be a chart on N containing $\varphi(x)$, and let ρ be a chart on L containing $\psi(\varphi(x))$. We want to show that the composition $\rho \circ (\psi \circ \varphi) \circ \sigma^{-1}$ is smooth where defined. But

$$\rho \circ (\psi \circ \varphi) \circ \sigma^{-1} = \left(\rho \circ \psi \circ \tau^{-1}\right) \circ \left(\tau \circ \varphi \circ \sigma^{-1}\right),$$

and by assumption each of the two bracketed terms on the right-hand side is a smooth map. Since the composition of smooth maps (defined on open sets in Euclidean space) is smooth, the left-hand side is also smooth. ∎

Remark 1.27. Consider the following curiosity. We have defined what it means for a continuous map between two smooth manifolds to be differentiable (Definition 1.24), but we have not defined what the derivative $D\varphi(x)$ is yet! This is somehow backwards—in normal calculus one first defines the derivative $Df(x)$ and then says the map is differentiable if the derivative $Df(x)$ always exists. In fact, the definition of the derivative of a map between two smooth manifolds is a little tricky, and this is what we will do in the next three lectures.

A smooth structure is defined as an equivalence class of smooth atlases. We can take this one step further and look at equivalence classes of smooth structures.

Definition 1.28. We say that two smooth structures Σ_1 and Σ_2 on a given topological manifold M belong to the same **diffeomorphism class** if there exists a diffeomorphism $(M, \Sigma_1) \to (M, \Sigma_2)$. This is clearly another equivalence relation. We write $\mathcal{S}(M)$ for the set of diffeomorphic classes of smooth structures on M.

Example 1.29. As an example to show that smooth structures and diffeomorphism classes really are different concepts, take $M = \mathbb{R}$. Let Σ denote the maximal smooth atlas containing the chart $x \mapsto x^3$. On Problem Sheet A you will check that this is *not* the same smooth structure as the standard one described in Example 1.19. However, there is an obvious diffeomorphism between the two smooth structures (namely, $x \mapsto x^3$). Thus they belong to the same diffeomorphism class.

(♣) Remark 1.30. Does every topological manifold admit a smooth structure (i.e. can every topological manifold be turned into a smooth manifold)? Can a topological manifold admit more than one diffeomorphism class? These questions are typically very hard to solve (and there are many open problems). Here are some interesting facts, all of which are way too hard to prove in this course.

- If M is a topological manifold of dimension 0,1,2 or 3 then $\mathcal{S}(M)$ contains a unique element.
- In higher dimensions, there may be more than one diffeomorphism class. For example, $\mathcal{S}(S^7)$ has exactly 28 elements, and there are more than sixteen million different elements in $\mathcal{S}(S^{31})$!
- For any $n \neq 4$, $\mathbb{R}^n$ admits a unique diffeomorphism class. However $\mathcal{S}(\mathbb{R}^4)$ has infinitely many elements. In general the most "wild" phenomena occur in dimension 4.
- There exist topological manifolds that do not admit *any* smooth structures at all: $\mathcal{S}(M) = \emptyset$.

We conclude this lecture with another esoteric (and non-examinable!) remark.

(♣) Remark 1.31. *Manifolds can be infinite-dimensional too!* Let me briefly outline the relevant definitions. This is for *interest only*—we will not use infinite-dimensional manifolds in this course.

Fix a Banach space E. We say that a topological space X is **locally modelled on** E if every point in X has a neighbourhood which is homeomorphic to an open set in E. A **topological Banach manifold** is a Hausdorff paracompact topological space with at most countably many connected components that is locally modelled on some Banach space E. A smooth Banach manifold is defined similarly—here we use the fact that differentiating functions on Banach spaces works in exactly the same way as differentiating functions on Euclidean spaces.

You should compare this to how you initially learned linear algebra. To begin with all vector spaces were finite-dimensional and linear operators were just matrices. Then two years later they told you that actually things could be infinite-dimensional. All the theorems you knew and loved from linear algebra continued to hold (provided a few more assumptions were made), only the proofs were much harder and it was no longer called "linear algebra," it was called "functional analysis". The same is true in differential geometry—infinite-dimensional differential geometry is sometimes referred to as "global analysis".

As a concrete example of an infinite-dimensional manifold, let M and N be two finite-dimensional manifolds, and let $1 \leq r < \infty$. Then the space $C^r(M, N)$ of maps from M to N of class C^r is an infinite-dimensional Banach manifold.

LECTURE 2

Tangent spaces

The goal of the next few lectures is to associate to an n-dimensional smooth manifold an n-dimensional vector space, denoted by T_xM, to each point $x \in M$. We call T_xM the **tangent space** to M at x. Although it *won't* be immediate from the definition why, the tangent space is what you would naturally "guess" it would be. See Figure 2.1 for the case of S^2 (which should be thought of as sitting inside $\mathbb{R}^3$). We will use this construction to define the *derivative* of a smooth map $\varphi \colon M \to N$:

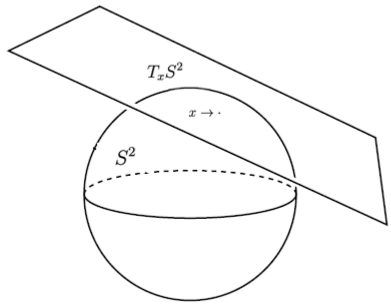

Figure 2.1: The tangent space to S^2

this will be a *linear* map $D\varphi(x) \colon T_xM \to T_{\varphi(x)}N$ for each $x \in M$. In Lecture 4 we will "glue" the vectors spaces together to form one larger space called the **tangent bundle** of M. This will be smooth manifold of twice the dimension of M. A smooth map $\varphi \colon M \to N$ will then induce a smooth map $D\varphi \colon TM \to TN$. In Lecture 5 we will look at submanifolds—it will not be until then that we can rigorously prove that the tangent space we define in this lecture really is the actual "tangent space" as in Figure 2.1 (cf. Example 5.16).

Definition 2.1. A **smooth function** on a manifold is a smooth map[1] $f \colon M \to \mathbb{R}$ in the sense of Definition 1.24, where $\mathbb{R}$ is given the standard smooth structure from Example 1.19. Thus f is a smooth function if for any chart $\sigma \colon U \to O$ on M, the composition $f \circ \sigma^{-1} \colon O \to \mathbb{R}$ is a smooth function (in the normal sense).

We denote by $C^\infty(M)$ the space of smooth functions. If $W \subset M$ is an open set, we define $C^\infty(W)$ to be the space of smooth functions that are only defined on W (where W is thought of a smooth manifold in the sense of Lemma 1.20). The space $C^\infty(M)$ is an **algebra**[2] (and thus in particular a ring and a vector space), under

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] As a general convention, I will use φ, ψ etc for smooth maps from one manifold to another, and f, g for smooth maps from a manifold to a Euclidean space.

[2] Don't worry too much if you're not familiar with the term "algebra"—in this case it just means a vector space where you can also multiply two elements together.

the operations[3]

$$(f+g)(x) := f(x)+g(x), \qquad (fg)(x) := f(x)g(x), \qquad (cf)(x) := cf(x), \qquad c \in \mathbb{R}.$$

(♣) Remark 2.2. Let M be a manifold of dimension $n > 0$, and let $W \subset M$ be a non-empty open set. Then as a vector space, $C^\infty(W)$ is always infinite-dimensional. There are many ways to see this, but here is an easy one. Let $f \in C^\infty(W)$ be any smooth function[4] which is not constant on some connected component of W. Then $f(W)$ is an infinite subset of $\mathbb{R}$ (since it contains an interval). Consider now the vector space $\mathbb{R}[t]$ of all polynomials. This is an infinite-dimensional vector space—a basis is the set of monomials $\{t^n \mid n \geq 0\}$. Any polynomial $p(t)$ is completely determined by its values on an infinite set, and thus in particular if $p \in \mathbb{R}[t]$ then p is completely determined by its values on $f(W)$. Therefore $\{p \circ f \mid p \in \mathbb{R}[t]\}$ is an infinite-dimensional subspace of $C^\infty(W)$, and hence the latter must be infinite-dimensional itself.

Before going any further, let us go back to $\mathbb{R}^n$ and introduce some more notation. To begin with, this will feel somewhat redundant, but we will see next lecture that it makes the various formulae easier to understand. Denote by $u^i \colon \mathbb{R}^n \to \mathbb{R}$ the function

$$u^i(x^1, \dots, x^n) = x^i. \tag{2.1}$$

Let e_i denote the ith standard basis vector in $\mathbb{R}^n$, so that

$$u^i(e_j) = \delta^i_j, \tag{2.2}$$

where δ^j_i is the *Kronecker delta* defined by

$$\delta^j_i = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Now suppose $f \colon O \subset \mathbb{R}^n \to \mathbb{R}^k$ is a smooth map defined on an open subset of $\mathbb{R}^n$. If $x \in O$ and $v \in \mathbb{R}^n$ then the vector[5] $Df(x)[v]$ can be thought of as the partial derivative of f in the direction v:

$$Df(x)[v] = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Definition 2.3. We abbreviate $Df(x)[e_j]$ by $D_j f(x)$:

$$D_j f(x) = Df(x)[e_j] = \lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t}.$$

Let us summarise the various different ways we can write the derivative:

---

[3] We will always write $\circ$ to denote composition, meanwhile juxtaposition indicates the pointwise product.

[4] If you are worried why such a function exists, you could for instance use Lemma 3.2 from the next lecture.

[5] I will typically use square brackets to indicate a matrix eating a vector.

Let $f\colon O \subset \mathbb{R}^n \to \mathbb{R}^k$ be a smooth map, and let $x \in O$. Then:

- $Df(x)$ is a $k \times n$ matrix.
- $D_j f(x)$ is an element of $\mathbb{R}^k$. It is the jth column of the matrix $Df(x)$.
- $D(u^i \circ f)(x)$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$. One can think of it as the ith row of the matrix $Df(x)$.
- $D_j(u^i \circ f)(x)$ is a number. It is the (i, j)th entry of the matrix $Df(x)$.

In more familiar notation

$$D_j(u^i \circ f)(x) = \frac{\partial f^i}{\partial x^j}(x). \tag{2.3}$$

In general I will prefer the slightly more cumbersome expression on the left-hand side of (2.3). This is because next lecture the symbol $\frac{\partial}{\partial x^i}$ will take on a special meaning (cf. Example 3.6), and I do not want to confuse you all. (See also Definition 7.4.)

Remark 2.4. In our new notation, part (ii) of the chain rule in Euclidean spaces (Proposition 1.12) reads:

$$D_j(u^i \circ g \circ f)(x) = \sum_{r=1}^{k} D_r(u^i \circ g)(f(x)) D_j(u^r \circ f)(x), \qquad \text{for all } 1 \le i \le l, 1 \le j \le n.$$

This should already give you a clue as to why the new notation is "better": we did not need to explicitly name our coordinate systems on $\mathbb{R}^n$ and $\mathbb{R}^k$.

Going back to manifolds, we can use the u^i to give an example of a smooth function.

Example 2.5. If $\sigma\colon U \to O$ is a chart on M, for each $i = 1, \dots, n$ the function $u^i \circ \sigma$ is a smooth function on U.

This type of smooth function is especially important, so it gets its own special name.

Definition 2.6. If $x \in M$ and σ is a chart defined on a neighbourhood of x then we will often denote the function $u^i \circ \sigma$ simply by x^i. We call the functions x^i the **coordinates** of the chart σ, and we say that the x^i are **local coordinates about** x.

Remark 2.7. The advantage of this phrasing is that it allows us to omit explicit reference to the chart σ. Thus the sentence

Let $x \in M$ and choose local coordinates $(x^1, \dots, x^n)$ about x...

is shorthand for

Let $x \in M$. Choose a chart σ defined on a neighbourhood of x, and let $x^i = u^i \circ \sigma$ denote the coordinates of σ...

In fact, sometimes we will go one step further and abuse the language even more by writing things like (!):

Let $x = (x^1, \ldots, x^n)$ *be a point in* $M \ldots$

The way to think about this is as follows: Differential Geometry is essentially a way of formalising calculus so that it makes sense on manifolds. The formalism is designed to make things "look" as similar as possible to calculus on Euclidean space. In fact, if you are ever stuck when trying to compute something (for instance, a derivative), you can just "pretend" that everything is actually defined on Euclidean space, and then simply follow the normal rules of multivariable calculus. The formalism is designed so that this will ~~always~~ usually give you the correct answer[6].

DEFINITION 2.8. Let M be a smooth manifold and let $x \in M$. Let U and V be two neighbourhoods of x, and suppose $f \in C^\infty(U)$ and $g \in C^\infty(V)$. We say that f and g have the same **germ** at x if there exists a smaller neighbourhood $W \subset U \cap V$ of x such that

$$f|_W \equiv g|_W.$$

One can think of this as follows: define an equivalence relation on the set of smooth functions defined on a neighbourhood of x by saying that $(U, f) \sim (V, g)$ if there exists a neighbourhood $W \subset U \cap V$ such that $f|_W \equiv g|_W$ (it is immediate that this does indeed define an equivalence relation). A germ is then an equivalence class under this relation. We denote the germ by $\underline{f}$ and we let $\mathcal{F}_x M$ denote the set of germs at x.

In fact, $\mathcal{F}_x M$ is another algebra. We can add germs together: if $\underline{f}$ and $\underline{g}$ are two germs with representatives (U, f) and (V, g) respectively, then $\underline{f} + \underline{g}$ is the germ represented by $(U \cap V, f + g)$. Similarly $\underline{f}\,\underline{g}$ is the germ represented by $(U \cap V, fg)$, and for a real number c, $c\underline{f}$ is the germ represented by (U, cf). We denote by $\underline{c}$ the germ of any function which is constant and equal to c in a neighbourhood of x. The map $\mathbb{R} \to \mathcal{F}_x M$ given by $c \mapsto \underline{c}$ is then an inclusion of algebras.

A germ at x has a well-defined value at x (although nowhere else), and this gives us map

$$\mathrm{eval}_x \colon \mathcal{F}_x M \to \mathbb{R}, \qquad \mathrm{eval}_x(\underline{f}) := f(x),$$

where (U, f) is any representative of $\underline{f}$.

(♣) REMARK 2.9. This remark requires you to know a little bit more ring theory to understand[7]. Since eval_x is clearly a ring homomorphism, the kernel of eval_x is an ideal in the ring $\mathcal{F}_x M$. Since the map eval_x is surjective (as $\mathrm{eval}_x(\underline{c}) = c$), this is actually a *maximal* ideal. In fact, it is the unique maximal ideal, since if $\mathrm{eval}(\underline{f}) \neq 0$ then $\underline{f}$ is invertible in $\mathcal{F}_x M$. Indeed, if (U, f) is a representative of $\underline{f}$ then there exists $V \subset U$ such that f is never zero on V. Thus there is a well-defined function $g := 1/f \colon V \to \mathbb{R}$, and $\underline{g}$ is then an inverse to $\underline{f}$. This shows that $\mathcal{F}_x M$ is a *local ring*.

[6]Of course, this is *not* the correct way to go about learning the subject! To learn the subject well, you should have a deep and profound understanding of every single concept…

[7]If you are not familiar with this material, just ignore this remark—note the (♣).

We now motivate the definition of a tangent vector by looking again at Euclidean spaces.

EXAMPLE 2.10. Let $f\colon O \subset \mathbb{R}^n \to \mathbb{R}$ be smooth, where $O \subset \mathbb{R}^n$ is open. Let $x \in O$ and $v \in \mathbb{R}^n$. The usual interpretation of the derivative is that the matrix $Df(x)$ eats the vector v to produce a real number $Df(x)[v]$. However we could also let f vary and consider the action of differentiation as a map

$$O \times \mathbb{R}^n \times C^\infty(O) \to \mathbb{R}, \qquad (x, v, f) \mapsto Df(x)[v].$$

Instead, let us consider (x, v) as fixed, and just let f vary:

$$(x, v)\colon C^\infty(O) \to \mathbb{R}, \qquad (x, v)(f) := Df(x)[v].$$

It follows from equation (1.1) that differentiability is a local property[8], in the sense that the value of $Df(x)[v]$ depends only on the germ of f at x. Thus we can think of the vector v as defining a *linear* map

$$v\colon \mathcal{F}_x O \to \mathbb{R}, \qquad v(\underline{f}) := Df(x)[v]$$

(here we are thinking of O as a smooth manifold). In fact, the map $v\colon \mathcal{F}_x O \to \mathbb{R}$ is not just any linear map, it is also a **derivation** in the sense that

$$v(\underline{f}\,\underline{g}) = \mathrm{eval}_x(\underline{f})v(\underline{g}) + \mathrm{eval}_x(\underline{g})v(\underline{f}).$$

Indeed, this is just a fancy way of expressing the **Leibniz rule**:

$$D(fg)(x)[v] = f(x)Dg(x)[v] + g(x)Df(x)[v].$$

Motivated by Example 2.10, we will define a tangent vector as a derivation on the space of germs.

DEFINITION 2.11. Let M be a smooth manifold of dimension n and let $x \in M$. A **tangent vector at** x is a linear map

$$v\colon \mathcal{F}_x M \to \mathbb{R}$$

which satisfies the *derivation property*:

$$v(\underline{fg}) = \mathrm{eval}_x(\underline{f})v(\underline{g}) + \mathrm{eval}_x(\underline{g})v(\underline{f}).$$

Since a tangent vector is a linear map from the vector space $\mathcal{F}_x M$ to $\mathbb{R}$, the set of tangent vectors is itself a vector space, and we denote it by $T_x M$.

Here is an easy lemma about derviations.

LEMMA 2.12. *Suppose $v\colon \mathcal{F}_x M \to \mathbb{R}$ is a tangent vector at x and $\underline{c} \in \mathcal{F}_x M$ is a constant germ. Then $v(\underline{c}) = 0$.*

Proof. Since $\underline{c} = c\,\underline{1}$ we have $v(\underline{c}) = cv(\underline{1})$ by linearity. But by the derivation property:

$$v(\underline{1}) = v(\underline{1}\,\underline{1}) = 2\ \mathrm{eval}_x(\underline{1})v(\underline{1}) = 2v(\underline{1})$$

and thus $v(\underline{1}) = 0$. Thus also $v(\underline{c}) = 0$. ∎

[8]If you yearn for a precise definition of the term "local property", see Definition 16.17.

In the special case where $O \subset \mathbb{R}^n$ is an open set, Example 2.10 showed that every vector $v \in \mathbb{R}^n$ defines an element of $T_x O$ (in the sense of Definition 2.11). In fact, these are *all* the elements of $T_x O$, although this requires a bit of work to see. More generally, one has:

THEOREM 2.13. *Let M be a smooth manifold of dimension n and let $x \in M$. Then the vector space $T_x M$ has dimension n.*

Theorem 2.13 is not immediate. Indeed, from Definition 2.11 it is not remotely clear why $T_x M$ should even be finite-dimensional! We will prove Theorem 2.13 in the next lecture by explicitly finding a basis of $T_x M$.

LECTURE 3

Partitions of unity

We begin this lecture by reformulating the definition of a tangent vector in a slightly more convenient way. Since germs are defined via equivalence classes, they are often tedious to work with, and we would like to dispense with them.

Definition 3.1. Let M be a smooth manifold, let $x \in M$, and let W be any neighbourhood of x (for instance W could be all of M). A **derivation of $C^\infty(W)$ at x** is a linear map $w \colon C^\infty(W) \to \mathbb{R}$ which satisfies the *derivation property*

$$w(fg) = f(x)w(g) + g(x)w(f).$$

If $v \in T_xM$ then v naturally defines a derivation w of $C^\infty(W)$ for any open W containing x by setting

$$w(f) := v(\underline{f}). \tag{3.1}$$

In fact, the converse is also true, as we will prove in Proposition 3.3 below. First we need a preliminary lemma. To state it, recall that for a smooth function $\eta \colon M \to \mathbb{R}$, we denote by $\operatorname{supp}(\eta)$ the **support** of η, defined by:

$$\operatorname{supp}(\eta) := \overline{\{x \in M \mid \eta(x) \neq 0\}}.$$

Lemma 3.2 (Cutoff functions). *Let M be a smooth manifold and let $K \subset U$ be subsets, where K is closed and U is open. Then there exists a smooth function $\eta \colon M \to \mathbb{R}$ such that:*

1. $0 \leq \eta(x) \leq 1$ *for all* $x \in M$,
2. $\operatorname{supp}(\eta) \subset U$,
3. $\eta(x) = 1$ *for all* $x \in K$.

The proof of Lemma 3.2 will be carried out at the end of this lecture, when we discuss **partitions of unity**. It is *not* obvious—as we will see this is the main reason we require manifolds to be paracompact.

Proposition 3.3. *Let M be a smooth manifold, let $x \in M$, and let W be any neighbourhood of x. Then there is a linear isomorphism between T_xM and the space of derivations of $C^\infty(W)$ at x.*

Proof. Let W be a neighbourhood of x. We prove the result in three steps.

1. Let $w \colon C^\infty(W) \to \mathbb{R}$ be a derivation at x. Suppose $f \in C^\infty(W)$ is identically zero on a neighbourhood $V \subset W$ of x. We claim that $w(f) = 0$. For this, choose a smooth function $\eta \colon M \to \mathbb{R}$ such that $\eta(x) = 1$ and $\operatorname{supp}(\eta) \subset V$ (i.e. apply Lemma 3.2 with $K = \{x\}$ and "U" equal to V). Let $g = \eta f$, thought of as a function $W \to \mathbb{R}$. Then g is identically zero, and hence $w(g) = 0$ by linearity. But by the derivation property

$$w(g) = w(\eta f) = \eta(x)w(f) + f(x)w(\eta) = w(f)$$

since $\eta(x) = 1$ and $f(x) = 0$. Thus $w(f) = 0$.

2. Suppose now $\underline{f} \in \mathcal{F}_x M$. We claim that we can always find a representative for $\underline{f}$ with domain W, i.e. a smooth function $g \colon W \to \mathbb{R}$ such that $\underline{g} = \underline{f}$. For this, let (V, f) be any representative of $\underline{f}$. By shrinking[1] V if necessary, we may assume that $V \subset W$. Now choose[2] a smaller neighbourhood U of x with $\overline{U} \subset V \subset W$. Our goal now is to *extend* f to a smooth function g defined on W such that $g|_U = f$. For this, we apply Lemma 3.2 again, this time with $K = \overline{U}$ and "U" equal to V. Now consider the smooth function

$$g \colon W \to \mathbb{R}, \qquad g(x) := \begin{cases} \eta(x)f(x), & x \in V, \\ 0, & x \in W \setminus V. \end{cases}$$

Since $g|_U = f$, we certainly have $\underline{g} = \underline{f}$.

3. We now complete the proof. Let $w \colon C^\infty(W) \to \mathbb{R}$ be a derivation at x. We define a a linear map $v \colon \mathcal{F}_x M \to \mathbb{R}$ by setting

$$v(\underline{f}) := w(f), \qquad \text{where } (W, f) \text{ is any representative of } \underline{f}.$$

That such a representative (W, f) exists was the content of Step 2, and the fact that v is well-defined follows from Step 1. Indeed, if (W, h) was another representative of $\underline{f}$ then by assumption there exists a smaller neighbourhood V of x such that $f|_V \equiv h|_V$. Then by linearity $w(f) - w(h) = w(f - h)$ and $w(f - h) = 0$ by Step 1. Finally, it is clear that v is a derivation. This association $w \mapsto v$ obviously inverts (3.1), and thus this completes the proof. ■

Thanks to Proposition 3.3, we will from now always regard a tangent vector v as a derivation of $C^\infty(W)$ at x for any open W containing x. Typically we take W either to be the domain of a chart, or the whole manifold M. We emphasise that Proposition 3.3 implies that it doesn't matter which W we choose. We also obtain immediately:

Corollary 3.4. *Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Regard W as a smooth manifold in its own right, as in Lemma 1.20. Then for any $x \in W$ there is a canonical identification $T_x M = T_x W$.*

Let us also note the following corollary of Lemma 2.12.

[1]The definition of the equivalence relation for a germ always allows us to replace a given representative with a representative defined on a *smaller* set. The aim of this step is replace a given representative with one defined on a *larger* set.

[2]Why does such a neighbourhood U exist? This comes down to point-set topology. If X is any locally compact Hausdorff space, x is any point of X, and W is any neighbourhood of x, then there exists a neighbourhood U of x such that $\overline{U}$ is compact and contained in W. Any manifold is locally compact by part (i) of Remark 1.9. Alternatively one could use the fact that any manifold M is normal, cf. part (iv) of Remark 1.9.

Corollary 3.5. *Let M be a smooth manifold, let $x \in M$, and let $f \in C^\infty(W)$ for some open W containing x. If f is constant in a neighbourhood of x then $v(f) = 0$ for all $v \in T_xM$.*

Let us give now give a concrete example of a tangent vector.

Example 3.6. Let M be a smooth manifold of dimension n, and let $\sigma\colon U \to O$ be a chart on M and write $x^i = u^i \circ \sigma$ for the local coordinates of σ. Let x be any point in U. Define a derivation of $C^\infty(U)$ at x by:

$$\frac{\partial}{\partial x^i}\Big|_x \colon C^\infty(U) \to \mathbb{R}, \qquad \frac{\partial}{\partial x^i}\Big|_x (f) := D_i(f \circ \sigma^{-1})(\sigma(x)),$$

where the right-hand side uses the convention from Definition 2.3. We will shortly prove that the collection $\frac{\partial}{\partial x^i}\big|_x$ for $i = 1, \dots, n$ form a basis of T_xM, thus establishing Theorem 2.13.

(♣) Remark 3.7. Given Proposition 3.3, you may wonder why we didn't immediately define T_xM as the vector space of derivations of $C^\infty(M)$ at x. There are several reasons. Firstly, using germs better encapsulates the fact that differentiation is a local property. Secondly, the advantage of our approach is that Corollary 3.4 is tautological; if we had defined T_xM directly as derivations of $C^\infty(M)$ then this would have required proof (and in fact the proof would essentially be the same argument used in Proposition 3.3).

Thirdly, in certain other contexts, the analogue of Lemma 3.2 is false. For instance, there is an analogous theory of **analytic manifolds**, which are defined in exactly the same way as smooth manifolds, except the word "smooth" should be replaced with "real-analytic" everywhere (thus an analytic manifold has a real-analytic atlas, and maps between real-analytic manifolds are required to be real-analytic, etc). We will not study analytic manifolds in this course, although they are very important in certain fields. In the real-analytic category, Lemma 3.2 is false: there do *not* exist real-analytic cutoff functions. (Exercise: Why?) Thus for analytic manifolds, Proposition 3.3 is false, and one is forced to work with germs to define the tangent space.

Finally, later in the course (Lecture 17) we will discuss **sheaves**, and germs are a motivating example for the construction of the **stalk** of a sheaf.

Let us now get started on the proof of Theorem 2.13. We will need the following easy lemma from multivariable calculus. Recall an open set $O \subset \mathbb{R}^n$ such that $0 \in O$ is said to be **star-shaped** if given any $x \in O$, the line segment from 0 to x is also contained in O.

Lemma 3.8. *Let $O \subset \mathbb{R}^n$ with $0 \in O$ be a star-shaped open set. Suppose $h\colon O \to \mathbb{R}$ is a smooth function. Then there exist n smooth functions $g_i\colon O \to \mathbb{R}$ for $i = 1, \dots, n$ such that $g_i(0) = D_ih(0)$ and such that*

$$h = h(0) + \sum_{i=1}^n u^i\, g_i,$$

where u^i is as in (2.1).

Proof. Fix $x = (x^1, \dots, x^n) \in O$ and consider the line segment $\gamma(t) = tx$. Set $\delta = h \circ \gamma \colon [0,1] \to \mathbb{R}$. Then by the chain rule

$$\delta'(t) = \sum_{i=1}^{n} x^i \, D_i h(tx).$$

Thus

$$h(x) - h(0) = \delta(1) - \delta(0) = \int_0^1 \delta'(t)\, dt = \sum_{i=1}^{n} x^i \int_0^1 D_i h(tx)\, dt.$$

Since $x^i = u^i(x)$ by definition, the claim follows with $g_i(x) := \int_0^1 D_i h(tx)\, dt$. ■

Theorem 2.13 from the last lecture follows immediately from the next statement.

PROPOSITION 3.9. *Let M be a smooth manifold of dimension n. Let $\sigma \colon U \to O$ be a chart on M and let $x \in U$. Then any tangent vector $v \in T_x M$ can be uniquely written as a linear combination*

$$v = \sum_{i=1}^{n} a^i \, \frac{\partial}{\partial x^i}\Big|_x .$$

In fact, $a^i = v(x^i)$. Thus $\left\{ \frac{\partial}{\partial x^i}\big|_x \mid i = 1, \dots, , n \right\}$ is a basis of $T_x M$.

Proof. We may assume without loss of generality that $\sigma(x) = 0$ and that O is star-shaped (here we are using Corollary 3.4). Let $f \in C^\infty(U)$ and apply Lemma 3.8 with $h = f \circ \sigma^{-1}$. We obtain $f = f(x) + \sum_{i=1}^n x^i \,(g_i \circ \sigma)$, where

$$g_i(0) = D_i(f \circ \sigma^{-1})(0) = \frac{\partial}{\partial x^i}\Big|_x (f).$$

Thus for any derivation v, one has

$$v(f) = \underbrace{v(f(x))}_{=0} + \sum_{i=1}^{n} \left(v(x^i) g_i(0) + \underbrace{x^i(x)}_{=0} v(g_i \circ \sigma) \right) = \sum_{i=1}^{n} v(x^i) \, \frac{\partial}{\partial x^i}\Big|_x (f),$$

where we used Corollary 3.5 and the assumption that $\sigma(x) = 0$.

This shows that $\left\{ \frac{\partial}{\partial x^i}\big|_x \mid i = 1, \dots, , n \right\}$ spans $T_x M$. It remains to prove linear independence. For this we note that:

$$\frac{\partial}{\partial x^i}\Big|_x (x^j) = \frac{\partial}{\partial x^i}\Big|_x (u^j \circ \sigma) = D_i(u^j \circ \sigma \circ \sigma^{-1})(\sigma(x)) = D_i u^j(\sigma(x)) = \delta_i^j, \qquad (3.2)$$

where we used the fact that $Du^j = u^j$ as u^j is a linear function, together with (2.2). Thus if $v = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}\big|_x = 0$ then feeding x^j to v gives $a^j = 0$. This shows linear independence, and thus completes the proof. ■

REMARK 3.10. Suppose σ and τ are two charts about x, with corresponding coordinate systems $x^i := u^i \circ \sigma$ and $y^i := u^i \circ \tau$. Taking $v = \frac{\partial}{\partial y^j}\big|_x$ in Proposition 3.9 tells us that

$$\frac{\partial}{\partial y^j}\Big|_x = \sum_{i=1}^{n} \frac{\partial}{\partial y^j}\Big|_x (x^i) \, \frac{\partial}{\partial x^i}\Big|_x .$$

But unravelling the definitions,

$$\frac{\partial}{\partial y^j}\Big|_x (x^i) = D_j(x^i \circ \tau^{-1})(\tau(x)) = D_j(u^i \circ \sigma \circ \tau^{-1})(\tau(x)), \tag{3.3}$$

which is just the (i, j)th entry of the matrix $D(\sigma \circ \tau^{-1})(\tau(x))$. This means that the transition matrix from the basis $\left\{ \frac{\partial}{\partial y^i}\big|_x \mid i = 1, \ldots, n \right\}$ to the basis $\left\{ \frac{\partial}{\partial x^i}\big|_x \mid i = 1, \ldots, n \right\}$ is given by the matrix $D(\sigma \circ \tau^{-1})(\tau(x))$.

We conclude this lecture by proving Lemma 3.2. In fact, we will first establish a special case of Lemma 3.2 where the smaller set K is *compact* (instead of merely closed).

Lemma 3.11 (Cutoff functions, the compact case). *Let M be a manifold and let $K \subset U$ be subsets, where K is compact and U is open. Then there exists a smooth function $\eta\colon M \to \mathbb{R}$ such that:*

1. $0 \le \eta(x) \le 1$ *for all* $x \in M$,
2. $\operatorname{supp}(\eta) \subset U$,
3. $\eta(x) = 1$ *for all* $x \in K$.

The proof of Lemma 3.11 is non-examinable.

(♣) *Proof.* We prove the result in four steps.

1. We first prove that for any pair of real numbers $r < R$ there exists a smooth function $f\colon \mathbb{R} \to [0, 1]$ such that $f(t) = 1$ for $t \le r$, $f(t) = 0$ for all $t \ge R$, and $0 < f(t) < 1$ for all $t \in (r, R)$. For this, consider the function

$$h\colon \mathbb{R} \to \mathbb{R}, \qquad h(t) := \begin{cases} e^{-1/t}, & t > 0, \\ 0, & t \le 0. \end{cases}$$

A somewhat tedious computation[3] shows that h is smooth. Our desired function f is then given by

$$f(t) := \frac{h(R - t)}{h(R - t) + h(t - r)}.$$

One can easily check this function f has the desired properties.

2. Now let us extend this to $\mathbb{R}^n$. Let $B_r \subset \mathbb{R}^n$ denote the open ball of radius r about the origin (so that $B_1 = B^n$). Then for any $0 < r < R$ there exists a smooth function $g\colon \mathbb{R}^n \to \mathbb{R}$ such that $g(x) = 1$ for all $x \in \overline{B}_r$, $g(x) = 0$ on $\mathbb{R}^n \setminus B_R$, and $0 < g(x) < 1$ for all $x \in B_R \setminus \overline{B}_r$. Indeed, the function $g(x) := f(|x|)$, where f is as in the previous step works.

3. Now let M be a smooth manifold, let $x \in M$, and let U be an arbitrary neighbourhood of x. Then we can choose a smaller neighbourhood $V \subset U$ of x with $\overline{V} \subset U$ that has the following property: there exists a smooth function $\eta\colon M \to \mathbb{R}$ such that $\eta(x) = 1$ for all $x \in \overline{V}$, $0 \le \eta(x) \le 1$ for all $x \in M$, and $\eta(x) = 0$ for all $x \in M \setminus U$. Indeed, this follows from the previous step, by choosing an appropriate chart about x.

[3] This would have made a good Analysis I question.

4. We now complete the proof. For each point $x \in K$, choose[4] neighbourhoods $V_x \subset U_x$ such that $\overline{V}_x \subset K$ and $U_x \subset U$. Since K is compact, there are finitely many points $x_1, \dots, x_N$ such that $K \subset \bigcup_{i=1}^N V_{x_i}$. For each i, choose functions $\eta_i \colon M \to \mathbb{R}$ such that $\eta_i(x) = 1$ for all $x \in \overline{V}_i$, $0 \le \eta_i(x) \le 1$ for all $x \in M$, and $\eta_i(x) = 0$ for all $x \in M \setminus U_i$. Now set

$$\eta := 1 - \prod_{i=1}^N (1 - \eta_i(x)) .$$

One easily checks this η does the trick. ∎

The proof of the general version of Lemma 3.2 is considerably harder. We will make use of the following tool, which will be of use throughout the course.

Definition 3.12. Let M be a smooth manifold. A **partition of unity** is a collection $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of smooth functions $\lambda_{\mathtt{a}} \colon M \to \mathbb{R}$ such that:

1. $0 \le \lambda_{\mathtt{a}}(x) \le 1$ for all $x \in M$ and $\mathtt{a} \in \mathtt{A}$.
2. The collection $\{\operatorname{supp}(\lambda_{\mathtt{a}}) \mid \mathtt{a} \in \mathtt{A}\}$ is locally finite, i.e. for any $x \in M$ there are most finitely many $\mathtt{a} \in \mathtt{A}$ such that $x \in \operatorname{supp}(\lambda_{\mathtt{a}})$.
3. For all $x \in M$ one has
$$\sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}}(x) = 1$$
(note by (2) this sum only has finitely many non-zero terms for every x).

We say that a partition of unity $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is **subordinate** to an open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ if $\operatorname{supp}(\lambda_{\mathtt{a}}) \subset U_{\mathtt{a}}$ for each $\mathtt{a} \in \mathtt{A}$. The next result is the reason we require our manifolds to be (Hausdorff and) paracompact.

Theorem 3.13. *Let M be a smooth manifold. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M. There exists a locally finite refinement $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ and a partition of unity $\{\lambda_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ subordinate to $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ with the additional property that $\operatorname{supp}(\lambda_{\mathtt{b}})$ is a compact subset of M for every $\mathtt{b} \in \mathtt{B}$.*

Of course, the main content of the theorem is the existence of the partition of unity $\{\lambda_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$—the existence of the locally finite refinement $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ is just the very definition of paracompactness. This proof is non-examinable.

(♣) *Proof.* Paracompactness guarantees us the existence of a locally finite refinement $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$. In fact, we can do a little better than this: we can find a locally finite refinement $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ together with another open cover $\{W_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ (with the same index set $\mathtt{B}$) such that $\overline{W}_{\mathtt{b}}$ is compact for each $\mathtt{b} \in \mathtt{B}$ and such that $\overline{W}_{\mathtt{b}} \subset V_{\mathtt{b}}$. This argument uses the fact that M is also locally compact (cf. part (i) of Remark 1.9). We won't dwell on the details as they not important to the main theme of the course.

[4]We are using the fact that M is normal again here, cf. part (iv) of Remark 1.9.

We now apply Lemma 3.11 to each pair $\overline{W}_{\mathtt{b}} \subset V_{\mathtt{b}}$ to obtain a smooth function $\eta_{\mathtt{b}} \colon M \to \mathbb{R}$ such that $0 \le \eta_{\mathtt{b}}(x) \le 1$ for all $x \in M$, $\eta_{\mathtt{b}}|_{\overline{W}_{\mathtt{b}}} \equiv 1$, and $\operatorname{supp}(\eta_{\mathtt{b}}) \subset V_{\mathtt{b}}$ is compact. The desired partition of unity is then given by

$$\lambda_{\mathtt{b}} := \frac{\eta_{\mathtt{b}}}{\sum_{\mathtt{b} \in \mathtt{B}} \eta_{\mathtt{b}}}.$$

This completes the proof. ∎

In fact, the following corollary is typically more useful.

Corollary 3.14 (Partitions of unity). *Let M be a smooth manifold. For any open cover of M, there exists a partition of unity subordinate to that cover.*

That is, if we don't need our partition of unity to have compact supports, then we don't even need to refine our original cover.

Proof. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an arbitrary open cover. Let $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ be a locally finite refinement and let $\{\lambda_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ be a partition of unity subordinate to $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$, whose existence are guaranteed by Theorem 3.13. Choose a function $\beta \colon \mathtt{B} \to \mathtt{A}$ such that $V_{\mathtt{b}} \subset U_{\beta(\mathtt{b})}$ for each $\mathtt{b} \in \mathtt{B}$. Now define

$$\lambda_{\mathtt{a}} := \sum_{\mathtt{b} \in \beta^{-1}(\mathtt{a})} \lambda_{\mathtt{b}}.$$

If $\beta^{-1}(\mathtt{a}) = \emptyset$ this should be interpreted as the zero function. Then

$$\operatorname{supp}(\lambda_{\mathtt{a}}) = \overline{\bigcup_{b \in \beta^{-1}(\mathtt{a})} \{x \in M \mid \lambda_{\mathtt{b}}(x) \neq 0\}} = \bigcup_{b \in \beta^{-1}(\mathtt{a})} \operatorname{supp}(\lambda_{\mathtt{b}}) \subset U_{\mathtt{a}},$$

where the second equality used the fact that the collection $\{\operatorname{supp}(\lambda_{\mathtt{b}}) \mid \mathtt{b} \in \mathtt{B}\}$ is locally finite. (Exercise: Justify this!) It is immediate that the collection $\{\operatorname{supp}(\lambda_{\mathtt{a}}) \mid \mathtt{a} \in \mathtt{A}\}$ is locally finite, and thus we conclude that $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is another partition of unity which is subordinate to our original cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$. Note however that $\lambda_{\mathtt{a}}$ need not have compact support. ∎

Finally, our original Lemma 3.2 is an easy consequence of Corollary 3.14:

Proof of Lemma 3.2. Consider the open cover $\{U, M \setminus K\}$ of M. By Corollary 3.14 there exists a partition of unity $\{\lambda_U, \lambda_{M \setminus K}\}$. The function $\eta := \lambda_U$ has the properties we desire. ∎

The derivative and the tangent bundle

Let us now finally define the derivative of a smooth map.

DEFINITION 4.1. Let M and N be smooth manifolds, and let $\varphi \colon M \to N$ be a smooth map. Fix $x \in M$ and $v \in T_x M$. We define a tangent vector $w \in T_{\varphi(x)} N$ by setting

$$w(f) := v(f \circ \varphi), \qquad \forall\, f \in C^\infty(N).$$

It is clear w is a linear derivation of $C^\infty(N)$ at $\varphi(x)$, and hence an element of $T_{\varphi(x)}N$. Moreover if we denote w by $D\varphi(x)[v]$ then it is immediate that the map $v \mapsto D\varphi(x)[v]$ is a *linear* map. We call $D\varphi(x)$ the **derivative of φ at x**.

The chain rule becomes essentially tautologous.

PROPOSITION 4.2 (The chain rule on manifolds). *Let M, N and L be smooth manifolds, and suppose $\varphi \colon M \to N$ and $\psi \colon N \to L$ are smooth maps. Then*

$$D(\psi \circ \varphi)(x) = D\psi(\varphi(x)) \circ D\varphi(x).$$

Proof. Take $v \in T_x M$ and $f \in C^\infty(L)$. Then

$$D(\psi \circ \varphi)(x)[v](f) = v(f \circ \psi \circ \varphi) = D\varphi(x)[v](f \circ \psi) = D\psi(\varphi(x)) \circ D\varphi(x)[v](f).$$

The claim follows. ∎

REMARK 4.3. You may wonder why the chain rule is so (suspiciously) easy to prove. After all, the Euclidean version (Proposition 1.12) is quite tricky. Does Proposition 4.2 give a shortcut to proving the Euclidean version? The answer is sadly no: indeed, we already used the Euclidean version at least twice (in Proposition 1.26 and Lemma 3.8), and hence any attempt to "prove" the Euclidean version via Proposition 4.2 would yield a circular argument.

Let us compute the map $D\varphi(x)$ in coordinates.

LEMMA 4.4. *Let $\varphi \colon M \to N$ be a smooth map between two smooth manifolds, where M has dimension n and N has dimension k. Let $x \in M$, and let $\sigma \colon U \to O$ be a chart on M about x, and let $\tau \colon V \to \Omega$ be a chart on N about $\varphi(x)$. Denote the local coordinates of σ by (x^i) and denote the local coordinates of τ by (y^i). Then the matrix of $D\varphi(x)$ with respect to the bases $\left\{ \frac{\partial}{\partial x^j}\big|_x \mid j = 1, \dots, n \right\}$ of $T_x M$ and $\left\{ \frac{\partial}{\partial y^i}\big|_{\varphi(x)} \mid i = 1, \dots, , k \right\}$ of $T_{\varphi(x)} N$ is given by the matrix $D(\tau \circ \varphi \circ \sigma^{-1})(\sigma(x))$.*

Proof. We compute

$$\begin{aligned} D\varphi(x)\left[\frac{\partial}{\partial x^j}\Big|_x\right] &= \sum_{i=1}^k D\varphi(x)\left[\frac{\partial}{\partial x^j}\Big|_x\right](y^i)\,\frac{\partial}{\partial y^i}\Big|_{\varphi(x)} \\ &= \sum_{i=1}^k \frac{\partial}{\partial x^j}\Big|_x (y^i\circ\varphi)\,\frac{\partial}{\partial y^i}\Big|_{\varphi(x)} \\ &= \sum_{i=1}^k D_j(u^i\circ\tau\circ\varphi\circ\sigma^{-1})(\sigma(x))\,\frac{\partial}{\partial y^i}\Big|_{\varphi(x)} \end{aligned}$$

The number $D_j(u^i\circ\tau\circ\varphi\circ\sigma^{-1})(\sigma(x))$ is the (i,j)th entry of the matrix $D(\tau\circ\varphi\circ\sigma^{-1})(\sigma(x))$, and thus the proof is complete. ∎

We shall see shortly that this implies we are not overreaching ourselves—this new definition coincides with the old one when $M=\mathbb{R}^n$ and $N=\mathbb{R}^k$. Indeed, in this case take σ to be the identity chart on $\mathbb{R}^n$ and take τ to be the identity chart on $\mathbb{R}^k$. Then the coordinates $(x^1,\dots,x^n)$ of σ are just the usual coordinates on $\mathbb{R}^n$, and the coordinates $(y^1,\dots,y^k)$ are just the usual coordinates of $\mathbb{R}^k$. Define a linear isomorphism

$$T_x\colon \mathbb{R}^n\to T_x\mathbb{R}^n, \qquad T_x e_i = \frac{\partial}{\partial x^i}\Big|_x. \tag{4.1}$$

Similarly if we let e'_j denote the standard basis of $\mathbb{R}^k$, we have a linear isomorphism:

$$T_y\colon \mathbb{R}^k\to T_y\mathbb{R}^k, \qquad T_y e'_j = \frac{\partial}{\partial y^j}\Big|_y.$$

Now suppose $f\colon O\subset\mathbb{R}^n\to\mathbb{R}^k$ is smooth. We now have two (!) definitions of the map $Df(x)$. We temporarily write the two maps as $Df(x)^{\mathrm{calc}}$ and $Df(x)^{\mathrm{man}}$. Thus $Df(x)^{\mathrm{calc}}$ is a linear map $\mathbb{R}^n\to\mathbb{R}^k$; it is the matrix of partial derivatives, as at the beginning of Lecture 1. Meanwhile $Df(x)^{\mathrm{man}}$ is a linear map $T_x\mathbb{R}^n\to T_{f(x)}\mathbb{R}^k$. Lemma 4.4 tells us that these are the same map, or equivalently, that the following diagram commutes[1]:

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "Df(x)^{\mathrm{calc}}"] \arrow[d, "T_x"'] & \mathbb{R}^k \arrow[d, "T_{f(x)}"] \\ T_x\mathbb{R}^n \arrow[r, "Df(x)^{\mathrm{man}}"'] & T_{f(x)}\mathbb{R}^k \end{tikzcd} \tag{4.2}$$

From now on we will drop the "calc" and "man" superscripts, and just call both maps $Df(x)$. It should be clear from the context which is meant.

Remark 4.5. Actually this is still not completely satisfactory, since the isomorphism T_x required us to fix bases e_i and $\frac{\partial}{\partial x^i}\big|_x$ of $\mathbb{R}^n$ and $T_x\mathbb{R}^n$ respectively. On Problem Sheet B you will give an alternative proof of (4.1) and (4.2) that does *not* require fixing a basis first.

[1] This means that going clockwise is the same as going anticlockwise.

More generally, the same computation shows us that if $\varphi\colon M \to N$ is an arbitrary smooth map between smooth manifolds of dimension n and k respectively, and σ is a chart about $x \in M$ with coordinates (x^i), and τ is a chart about $y \in N$ with coordinates (y^j), then with T_x and $T_{\varphi(x)}$ defined as above, we have a commutative diagram:

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "D(\tau\circ\varphi\circ\sigma^{-1})(\sigma(x))"] \arrow[d, "T_x"'] & \mathbb{R}^k \arrow[d, "T_{\varphi(x)}"] \\ T_xM \arrow[r, "D\varphi(x)"'] & T_{\varphi(x)}N \end{tikzcd} \tag{4.3}$$

We now give an entirely different way of defining tangent vectors. This approach is not quite as aesthetically pleasing as using derivations, but it has the advantage that it is easier to compute.

Suppose $\gamma\colon (a,b) \to \mathbb{R}^n$ is a smooth map. We will usually write the coordinate on $\mathbb{R} = \mathbb{R}^1$ as t instead of x^1, and we will usually denote the derivative of γ at a point t by $\gamma'(t)$. Writing $\gamma = (\gamma^1, \dots, , \gamma^n)$, the vector $\gamma'(t)$ is just the row vector $((\gamma^1)'(t), \dots, (\gamma^n)'(t))$. Our aim now is to extend this to manifolds.

Definition 4.6. A **curve** in a smooth manifold M is a smooth map $\gamma\colon (a,b) \to M$, where we think of (a,b) as a 1-dimensional smooth manifold. Fix $t \in (a,b)$. There are, a priori, two different ways we could define an element $\gamma'(t)$ of $T_{\gamma(t)}M$, which we will call the **velocity vector of γ at time t**.

1. Firstly, we can define a derivation on $C^\infty(M)$ at $\gamma(t)$ by setting
$$\gamma'(t)(f) := (f \circ \gamma)'(t), \qquad f \in C^\infty(M). \tag{4.4}$$
(Exercise: Why is this a derivation?)

2. Secondly, if we think of γ as a smooth map between manifolds then we can define a tangent vector $\gamma'(t)$ at $\gamma(t)$ via the derivative $D\gamma(t)$:
$$\gamma'(t) := D\gamma(t)\left[\frac{\partial}{\partial t}\bigg|_t\right] \in T_{\gamma(t)}M. \tag{4.5}$$

To see that these two definitions agree, let $\sigma\colon U \to O$ be a chart defined on a neighbourhood of $\gamma(t)$. Let (x^i) denote the coordinates of σ. Let $\gamma^i := x^i \circ \gamma$ so that γ^i is a curve in $\mathbb{R}$. Applying Proposition 3.9 to (4.4), we see that

$$\gamma'(t) = \sum_{i=1}^n (\gamma^i)'(t)\,\frac{\partial}{\partial x^i}\bigg|_{\gamma(t)}, \tag{4.6}$$

since $\gamma'(t)(x^i) = (x^i \circ \gamma)'(t) = (\gamma^i)'(t)$. But similarly by applying Lemma 4.4 to (4.5) we see that this definition also gives the same formula (4.6) for $\gamma'(t)$.

Lemma 4.7. *Let M be a smooth manifold and let $\gamma, \delta\colon (-\varepsilon, \varepsilon) \to M$ be two smooth curves such that $\gamma(0) = \delta(0)$. Then $\gamma'(0) = \delta'(0)$ as elements of $T_{\gamma(0)}M$ if and only if for some (and hence any) chart $\sigma\colon U \to O$ defined on a neighbourhood of $\gamma(0)$, we have*
$$(\sigma \circ \gamma)'(0) = (\sigma \circ \delta)'(0). \tag{4.7}$$

Note both sides of (4.7) are the derivatives of smooth maps $(-\varepsilon, \varepsilon) \to O$, and hence these derivatives are the usual derivatives in the sense of multivariate calculus.

Proof. Let (x^i) denotes the coordinates of σ. The stated condition is equivalent to requiring that $(\gamma^i)'(0) = (\delta^i)'(0)$ for each i, where $\gamma^i = x^i \circ \gamma$ and $\delta^i = x^i \circ \delta$. The claim follows from (4.6), since $\left\{ \frac{\partial}{\partial x^i}\big|_{\gamma(0)} \right\}$ is a basis of $T_{\gamma(0)}M$. ∎

What is less clear is that *every* tangent vector can be written as the velocity vector of a curve.

PROPOSITION 4.8. *Let M be a smooth manifold of dimension n, let $x \in M$ and let $v \in T_xM$. There exists a smooth curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ such that $\gamma'(0) = v$.*

Proof. Choose a chart $\sigma\colon U \to O \subset \mathbb{R}^n$, where O is an open set containing 0 such that $\sigma(x) = 0$. Let the coordinates be denoted by (x^i) as usual, and suppose that $v = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}\big|_x$, where the a^i are real numbers. For sufficiently small $\varepsilon > 0$, the vector $(ta^1, \dots, ta^n)$ belongs to O for all $|t| < \varepsilon$. Then if we define

$$\gamma\colon (-\varepsilon, \varepsilon) \to M, \qquad \gamma(t) := \sigma^{-1}(ta^1, ta^2, \dots, ta^n),$$

then γ is well-defined, smooth, and satisfies $\gamma(0) = x$, and finally (4.6) shows us that $\gamma'(0) = v$. ∎

REMARK 4.9. This tells us that we can make the following alternative definition of T_xM: a tangent vector at $x \in M$ is an equivalence class of smooth curves $\gamma\colon (-\varepsilon, \varepsilon) \to M$ such that $\gamma(0) = x$, where $\gamma \sim \delta$ if and only if for some chart σ centred about x, (4.7) holds.

We note however that this only works because we *already* established that T_xM was a vector space with basis $\{ \frac{\partial}{\partial x^i}\big|_x \}$. If one wanted to *start* with this definition of T_xM, one would need to use Problem B.1 to endow T_xM with a vector space structure.

Let us examine how velocity vectors behave with respect to smooth maps.

PROPOSITION 4.10. *Let $\varphi\colon M \to N$ be a smooth map between two smooth manifolds, and let $\gamma\colon (a, b) \to M$ be a smooth curve in M. Then*

$$D\varphi(\gamma(t))[\gamma'(t)] = (\varphi \circ \gamma)'(t).$$

Proof. We will give two proofs, one for each of the two (equivalent) definitions (4.4) and (4.5) of $\gamma'(t)$. Of course these are really the same proof.

- *Proof using* (4.4) *as the definition of $\gamma'(t)$:* Take $f \in C^\infty(N)$. Then by the definition of $D\varphi(x)$ and (4.4)

$$D\varphi(\gamma(t))[\gamma'(t)](f) = \gamma'(t)(f \circ \varphi) = (f \circ \varphi \circ \gamma)'(t) = (\varphi \circ \gamma)'(t)(f).$$

- *Proof using* (4.5) *as the definition of $\gamma'(t)$:* For this we simply use the chain rule (Proposition 4.2):

$$\begin{aligned} D\varphi(\gamma(t))[\gamma'(t)] &= D\varphi(\gamma(t)) \circ D\gamma(t) \left[\frac{\partial}{\partial t}\Big|_t \right] \\ &= D(\varphi \circ \gamma)(t) \left[\frac{\partial}{\partial t}\Big|_t \right] \\ &= (\varphi \circ \gamma)'(t). \end{aligned}$$

This completes the proof (twice). ∎

Let us now look at the dual space to T_xM.

DEFINITION 4.11. Let M be a smooth manifold of dimension n and let $x \in M$. We denote the dual vector space $\mathrm{L}(T_xM, \mathbb{R})$ by T_x^*M and call it the **cotangent space** of M at x.

Thus T_x^*M is another vector space of dimension n. Since elements of T_xM are linear derivations eating functions, the standard duality construction tells us that we can interpret elements of T_x^*M as functions eating linear derivations.

EXAMPLE 4.12. Let M be a smooth manifold of dimension n and let $x \in M$. Let U be a neighbourhood of x and let $f \in C^\infty(U)$. Then f defines an element $df|_x \in T_x^*M$ by

$$df|_x(v) := v(f), \qquad v \in T_xM.$$

One calls $df|_x$ the **differential** of f at x.

REMARK 4.13. Thus $df|_x$ is a linear function $T_xM \to \mathbb{R}$. In contrast, the derivative[2] $Df(x)$ is a linear function $T_xM \to T_{f(x)}\mathbb{R}$. Under the identification $T_{f(x)}\mathbb{R} \cong \mathbb{R}$ given by (4.1) these become the same map:

$$Df(x)[v] = df|_x(v)\, \frac{\partial}{\partial t}\bigg|_{f(x)}.$$

PROPOSITION 4.14. *Let M be a smooth manifold of dimension n and let $x \in M$. Let $\sigma\colon U \to O$ be a chart about x with local coordinates $x^i = u^i \circ \sigma \in C^\infty(U)$. Then $\{dx^i|_x\}$ is a basis of T_x^*M.*

Proof. We need only note that $\{dx^i|_x\}$ is the dual basis to $\left\{\frac{\partial}{\partial x^i}\big|_x\right\}$ since

$$dx^j|_x\left(\frac{\partial}{\partial x^i}\bigg|_x\right) = \frac{\partial}{\partial x^i}\bigg|_x (x^j) = \delta_i^j,$$

by (3.2) from the last lecture. ∎

We now aim to "glue" the vector spaces T_xM together into one big manifold TM.

DEFINITION 4.15. Let M be a smooth manifold. The **tangent bundle** of M is the disjoint union of the tangent spaces:

$$TM = \bigsqcup_{x \in M} T_xM.$$

We denote an element of TM as a pair (x, v) to indicate that $v \in T_xM$. There is a map $\pi\colon TM \to M$ given by $\pi(x, v) = x$. We call π the **footpoint** map.

As it stands TM is only a set. Let us now prove it is actually a manifold.

[2] Note: "derivative" and "differential" are two different words!

THEOREM 4.16. *Let M be a smooth manifold of dimension n. The smooth structure on M naturally induces a smooth structure on TM, making TM into a smooth manifold of dimension $2n$. Moreover the map $\pi\colon TM \to M$ is smooth.*

Proof. Let $\Sigma = \{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be our smooth atlas on M. Write $x^i_{\mathtt{a}} = u^i \circ \sigma_{\mathtt{a}}$ for the local coordinates of $\sigma_{\mathtt{a}}$. We build a chart $\tilde{\sigma}_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to O_{\mathtt{a}} \times \mathbb{R}^n$ by setting

$$\tilde{\sigma}_{\mathtt{a}}(x, v) = \left(\sigma_{\mathtt{a}}(x), \sum_{i=1}^n dx^i_{\mathtt{a}}|_x(v)\, e_i \right), \qquad x \in U_{\mathtt{a}},\ v \in T_x M.$$

We will prove that if $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$ then for all $z \in \sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$ and $w \in \mathbb{R}^n$, one has:

$$\tilde{\sigma}_{\mathtt{b}} \circ \tilde{\sigma}_{\mathtt{a}}^{-1}(z, w) = \left(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1}(z), D(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1})(z)[w] \right), \tag{4.8}$$

From this it follows from Proposition 1.22 that TM is a smooth manifold. To prove (4.8), write $w = \sum_{j=1}^n c^j\, e_j$ and set $x := \sigma_{\mathtt{a}}^{-1}(z) \in U_{\mathtt{a}} \cap U_{\mathtt{b}}$. Then

$$\tilde{\sigma}_{\mathtt{a}}^{-1}(z, w) = \left(x, \sum_{j=1}^n c^j \left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x \right).$$

Fix $1 \leq i \leq n$. We compute

$$dx^i_{\mathtt{b}}|_x \left(\sum_{j=1}^n c^j \left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x \right) = \sum_{j=1}^n c^j\, dx^i_{\mathtt{b}}|_x \left(\left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x \right) = \sum_{j=1}^n c^j \left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x (x^i_{\mathtt{b}})$$

By (3.3) in Remark 3.10, the number $\left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x (x^i_{\mathtt{b}})$ is the (i,j)th entry of the matrix $D(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1})(z)$. Thus

$$\sum_{i=1}^n dx^i_{\mathtt{b}}|_x \left(\sum_{j=1}^n c^j \left.\frac{\partial}{\partial x^j_{\mathtt{a}}}\right|_x \right) e_i = D(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1})(z)[w],$$

and (4.8) is proved.

The right-hand side of (4.8) is a diffeomorphism by assumption. Thus

$$\tilde{\Sigma} = \left\{ \tilde{\sigma}_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to O_{\mathtt{a}} \times \mathbb{R}^n \mid \mathtt{a} \in \mathtt{A} \right\}$$

is a smooth atlas on TM. This proves that TM is a smooth manifold of dimension $2n$. To check that π is smooth, we simply observe that if $z \in O_{\mathtt{a}}$ and $w \in \mathbb{R}^n$ then

$$\sigma_{\mathtt{a}} \circ \pi \circ \tilde{\sigma}_{\mathtt{a}}^{-1}(z, w) = z,$$

which is obviously smooth. ■

REMARK 4.17. The tangent bundle is the prototypical example of a more general construction called a **vector bundle** over a smooth manifold. We will take up their study in Lecture 13.

We can use the tangent bundle to unify the derivatives $D\varphi(x)$ from Definition 4.1 into a single map.

DEFINITION 4.18. Let $\varphi\colon M \to N$ be a smooth map between two smooth manifolds. Define the **derivative of** φ to be the map

$$D\varphi\colon TM \to TN, \qquad D\varphi(x,v) := (\varphi(x), D\varphi(x)[v]) .$$

On Problem Sheet C you will prove this map is smooth. We conclude this lecture by defining the dual version of the tangent bundle:

DEFINITION 4.19. Let M be a smooth manifold. The **cotangent bundle** of M is the disjoint union of the cotangent spaces:

$$T^*M = \bigsqcup_{x\in M} T_x^*M.$$

We denote an element of T^*M as a pair (x,p) to indicate that $p \in T_x^*M$. We denote again by $\pi\colon T^*M \to M$ the **footpoint** map $\pi(x,p) = x$.

On Problem Sheet C you will show that T^*M is also naturally a smooth manifold of twice the dimension of M.

Submanifolds and the Implicit Function Theorem

In this lecture we define *submanifolds*, which are smaller manifolds sitting inside larger ones. Let us first recall the Inverse Function Theorem for maps defined on Euclidean space. We say that a smooth map $f\colon O \subset \mathbb{R}^n \to \mathbb{R}^k$ has **rank** r at $x \in O$ if the $k \times n$ matrix $Df(x)$ has rank r. We say that f has **maximal rank at** x if the rank of f at x is as large as it can be (which is thus equal to the minimum of n and k). If $n = k$ then f has maximal rank at x if and only if $Df(x)$ is invertible.

THEOREM 5.1 (The Inverse Function Theorem). *Let $f\colon O \subset \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map, where O is open. Let $x \in O$ and assume the matrix $Df(x)$ has maximal rank $(= n)$. Then there exists a neighbourhood $\Omega \subset O$ of x such that the restriction $f\colon \Omega \to f(\Omega)$ is a diffeomorphism.*

The theorem immediately carries over to manifolds. We say that a smooth map $\varphi\colon M \to N$ has **rank** r at a point x if the linear subspace $D\varphi(x)[T_xM]$ has dimension r inside of $T_{\varphi(x)}N$.

THEOREM 5.2 (The Inverse Function Theorem for manifolds). *Let M and N be smooth manifolds of the same dimension n and suppose $\varphi\colon M \to N$ is a smooth map. Let $x \in M$ and assume that φ has maximal rank $(= n)$ at x. Then there exists a neighbourhood W of x such that the restriction $\varphi\colon W \to \varphi(W)$ is a diffeomorphism.*

Proof. The assertion is purely local. Choose a chart $\sigma\colon U \to O$ on M at x and a chart $\tau\colon V \to \Omega$ on N at $\varphi(x)$. Since σ and τ are diffeomorphisms (cf. Example 1.25), the derivative of the map

$$\tau \circ \varphi \circ \sigma^{-1}\colon \sigma(U \cap \varphi^{-1}(V)) \to \tau(\varphi(U) \cap V)$$

has rank n at $\sigma(x)$. Thus by Theorem 5.1 there exists $O' \subset \sigma(U \cap \varphi^{-1}(V))$ such that $\tau \circ \varphi \circ \sigma^{-1}|_{O'}$ is a diffeomorphism. Then using once more that σ and τ are diffeomorphisms, if $W \coloneqq \sigma^{-1}(O')$ then $\varphi|_W\colon W \to \varphi(W)$ is also a diffeomorphism. ■

We now move onto the Implicit Function Theorem. We shall give a quick proof using the Inverse Function Theorem.

THEOREM 5.3 (The Implicit Function Theorem). *Let O be a neighbourhood of 0 in $\mathbb{R}^n$ and suppose $f\colon O \to \mathbb{R}^k$ is a smooth map such that $f(0) = 0$.*

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

(i) *Assume $n \leq k$ and that the matrix $Df(0)$ has maximal rank ($= n$) at 0. Let $\imath \colon \mathbb{R}^n \to \mathbb{R}^k$ denote the inclusion*

$$\imath(x^1, \dots, x^n) := (x^1, \dots, x^n, 0, \dots, 0).$$

Then there exists a chart g about 0 on $\mathbb{R}^k$ such that $g \circ f = \imath$ on a neighbourhood of 0 in $\mathbb{R}^n$.

(ii) *Assume $n \geq k$ and that the matrix $Df(0)$ has maximal rank ($= k$) at 0. Let $\pi \colon \mathbb{R}^n \to \mathbb{R}^k$ denote the projection*

$$\pi(x^1, \dots, x^n) := (x^1, \dots, x^k).$$

Then there exists a chart h about 0 in $\mathbb{R}^n$ such that $f \circ h = \pi$ on a neighbourhood of 0 in $\mathbb{R}^n$.

Proof. We start with (i). The matrix $Df(0)$ has rank n. By rearranging the coordinate functions $f^i = u^i \circ f$ if necessary (this corresponds to composing f with a linear isomorphism $\mathbb{R}^k \to \mathbb{R}^k$, which is a diffeomorphism), we may assume that the $n \times n$ submatrix $\left(\frac{\partial f^i}{\partial x^j}(0)\right)_{1 \leq i,j \leq n}$ is invertible. Now define a map $\tilde{f} \colon O \times \mathbb{R}^{k-n} \to \mathbb{R}^k$ by

$$\tilde{f}(x^1, \dots, x^k) = f(x^1, \dots, x^n) + (0, \dots, 0, x^{n+1}, \dots, x^k).$$

Then $\tilde{f} \circ \imath = f$ and the derivative $D\tilde{f}(0)$ takes the following form:

$$D\tilde{f}(0) = \begin{pmatrix} \left(\frac{\partial f^i}{\partial x^j}(0)\right)_{1 \leq i,j \leq n} & 0 \\ * & \mathrm{id}_{\mathbb{R}^{k-n}} \end{pmatrix}$$

where $*$ denotes the other entries of $Df(0)$. Thus $\det D\tilde{f}(0) \neq 0$, and consequently $D\tilde{f}(0)$ has rank k. Thus by Theorem 5.1, there exists a neighbourhood $O' \subset O \times \mathbb{R}^{k-n}$ of the origin $0 \in \mathbb{R}^k$ such that $\tilde{f} \colon O' \to \tilde{f}(O')$ is a diffeomorphism. If g denotes the inverse to $\tilde{f}|_{O'}$ then $g \circ f = g \circ \tilde{f} \circ \imath = \imath$. This proves (1).

The proof of (ii) is very similar. This time we may assume that the submatrix $\left(\frac{\partial f^i}{\partial x^j}(0)\right)_{1 \leq i,j \leq k}$ is invertible, and we define $\tilde{f} \colon O \to \mathbb{R}^n$ by

$$\tilde{f}(x^1, \dots, x^n) := (f(x^1, \dots, x^n), x^{k+1}, \dots, x^n).$$

Then $f = \pi \circ \tilde{f}$ and the derivative $D\tilde{f}(0)$ takes the following form:

$$D\tilde{f}(0) = \begin{pmatrix} \left(\frac{\partial f^i}{\partial x^j}(0)\right)_{1 \leq i,j \leq k} & * \\ 0 & \mathrm{id}_{\mathbb{R}^{n-k}} \end{pmatrix}$$

This is invertible, whence $\tilde{f}$ has a local inverse h, and $f \circ h = \pi \circ \tilde{f} \circ h = \pi$. ∎

We will shortly prove a version of the Implicit Function Theorem for manifolds. First, some definitions. In order to make the statements more succinct, we will start adopting the convention that writing $M = M^n$ means that M has dimension n.

Definition 5.4. Let $\varphi \colon M^n \to N^k$ be a smooth map.

- We say that φ is an **immersion** if the linear map $D\varphi(x) \colon T_x M \to T_{\varphi(x)} N$ is injective for every $x \in M$. (Note this implies $n \le k$).
- If in addition φ itself is injective then we say that φ is an **injective immersion**.
- If in addition φ maps M homeomorphically onto $\varphi(M)$ (where $\varphi(M)$ is endowed with the subspace topology in N) we say that φ is an **embedding**.

Remark 5.5. If M is compact, then an injective immersion $\varphi \colon M \to N$ is automatically an embedding, as you will prove on Problem Sheet C. However in the non-compact case, this need not be the case (see again Problem Sheet C). However an (injective or not) immersion is always *locally* an embedding, as we will now prove.

The next result is a manifold version of part (i) of the Implicit Function Theorem 5.1.

Proposition 5.6. *Suppose $\varphi \colon M^n \to N^k$ is an immersion. Then for any $x \in M$, there exists a neighbourhood U of x and a chart $\tau \colon V \to \Omega$ on N, where V is some neighbourhood of $\varphi(x)$ such that:*

(i) *If $y^i = u^i \circ \tau$ denotes the local coordinates of τ then*

$$\varphi(U) \cap V = \left\{ y \in V \mid y^{n+1}(y) = \cdots = y^k(y) = 0 \right\}, \tag{5.1}$$

that is,

$$\tau(\varphi(U) \cap V) = (\mathbb{R}^n \times \{0\}) \cap \Omega.$$

(ii) *$\varphi|_U$ is an embedding.*

Proof. The assertion is again local. Let $\imath \colon \mathbb{R}^n \to \mathbb{R}^k$ denote the inclusion, as in part (i) of the Implicit Function Theorem 5.3. Let σ denote a chart on M with $\sigma(x) = 0$ and let $\tilde{\tau}$ denote a chart on N with $\tilde{\tau}(\varphi(x)) = 0$. Then $\tilde{\tau} \circ \varphi \circ \sigma^{-1}$ has maximal rank at 0, and hence by part (i) of the Implicit Function Theorem there exists a chart g on $\mathbb{R}^k$ about 0 and a neighbourhood O of 0 in $\mathbb{R}^n$ such that

$$g \circ \tilde{\tau} \circ \varphi \circ \sigma^{-1}|_O = \imath|_O.$$

Set $U := \sigma^{-1}(O)$ and set $\tau := g \circ \tilde{\tau}$. Then after restricting the domain if necessary, (5.1) holds. To prove the second statement, simply note that $\varphi|_U = \tau^{-1} \circ \imath \circ \sigma|_U$ is the composition of embeddings, and thus is an embedding. ■

Remark 5.7. If φ is an embedding then the set $\varphi(U)$ from Proposition 5.6 can be written as $\varphi(U) = \varphi(M) \cap W$ for some open set $W \subset N$. (This is just the definition of the subspace topology). Replacing V with $W \cap V$, (5.1) becomes

$$\varphi(M) \cap V = \left\{ y \in V \mid y^{n+1}(y) = \cdots = y^k(y) = 0 \right\}. \tag{5.2}$$

Definition 5.8. Let M and N be manifolds with $M \subset N$ (as sets). We say that M is a **embedded submanifold** of N if the inclusion $M \hookrightarrow N$ is an embedding. If the inclusion is merely an immersion (note the inclusion is always injective!), we say that M is an **immersed submanifold**.

If M is an embedded submanifold of N then Remark 5.7 tells us we can always choose charts on N that satisfy (5.2). We give such a chart a special name:

Definition 5.9. Let M^n be an embedded submanifold on N^k. A **slice chart for M in N** is a chart $\tau \colon V \to \Omega$ on N such that, writing $y^i = u^i \circ \tau$ for the local coordinates of τ, one has

$$M \cap V = \left\{ y \in V \mid y^{n+1}(y) = \cdots = y^k(y) = 0 \right\}.$$

In fact, the existence of slice charts is and "if and only if" condition, in the sense that we can use slice charts to endow a subset with a smooth structure. The next result makes this more precise.

Proposition 5.10. *Let N^k be a smooth manifold and suppose $M \subset N$ is a subset is with the property that around every point $x \in M$ there exists a slice chart for M in N, that is, a chart $\tau \colon V \to \Omega$ on N with $x \in V$ such that, writing $y^i = u^i \circ \tau$ for the local coordinates of τ, one has*

$$M \cap V = \left\{ y \in V \mid y^{n+1}(y) = \cdots = y^k(y) = 0 \right\}.$$

Then if we endow M with the subspace topology on N, M is a topological manifold of dimension n, and moreover it has a smooth structure that makes it into an embedded submanifold of N.

Proof. Let $\pi \colon \mathbb{R}^k \to \mathbb{R}^n$ denote projection as above. Fix $x \in M$ and let $\tau \colon V \to \Omega$ be such a slice chart for M in N with $x \in V$. Let $U := M \cap V$ and let $O := \pi(\Omega)$. Let $\sigma := \pi \circ \tau|_U$. If M is given the subspace topology then σ is a homeomorphism. If we do this at every point $x \in M$, we end up with a collection of maps for which the hypotheses of Proposition 1.22 are satisfied. Thus M is a smooth manifold of dimension n. Moreover the topology on M that Proposition 1.22 provides coincides with the subspace topology, since the maps σ were already homeomorphisms in the subspace topology. Finally if $\imath \colon M \hookrightarrow N$ denotes the inclusion that with τ, σ as above, one has $\tau \circ \imath \circ \sigma^{-1}$ equal to the map $(x^1, \dots, x^n) \mapsto (x^1, \dots, x^n, 0, \dots, 0)$ which is smooth. ■

If $\varphi \colon M \to N$ is an injective immersion (resp. an embedding) then M is diffeomorphic to an immersed submanifold (resp. an embedded submanifold) of N: namely $\varphi(M)$, where $\varphi(M)$ is endowed with the smooth structure such that $\varphi \colon M \to \varphi(M)$ is a diffeomorphism.

(♣) Remark 5.11. Suppose $\varphi_1 \colon M_1 \to N$ and $\varphi_2 \colon M_2 \to N$ are two injective immersions. We say that φ_1 is **equivalent** to φ_2 if there exists a diffeomorphism $\psi \colon M_1 \to M_2$ such that $\varphi_2 \circ \psi = \varphi_1$. It is clear that this defines an equivalence relation on the set of injective immersions into N. Each equivalence class contains a unique immersed submanifold.

We now move onto the case where the first manifold M is the "larger" one.

DEFINITION 5.12. Let $\varphi\colon M^n \to N^k$ be smooth. A point $x \in M$ is said to be a **regular point** of φ if φ has rank k at x (note this implies $n \geq k$). A point $x \in M$ is called a **critical point** if it is not a regular point. Similarly a point $y \in N$ is called a **regular value** if every point in $\varphi^{-1}(y)$ is a regular point. A point $y \in N$ is called a **critical value** if it is not a regular value. Thus if $y \notin \varphi(M)$ then y is vacuously a regular value.

Here is the main result of today's lecture. One can think of it as a manifold version of part (ii) of the Implicit Function Theorem 5.1, although unlike Proposition 5.6 this is a much deeper result, as the assertion is *not* local.

THEOREM 5.13 (The Implicit Function Theorem for manifolds). *Let $\varphi\colon M^n \to N^k$ be a smooth map with $n \geq k$. Suppose $y \in N$ is a regular value of φ and $L := \varphi^{-1}(y)$ is not empty. Then L is a topological manifold of dimension $n-k$. Moreover there exists a smooth structure on L which makes L into a smooth embedded submanifold of M.*

This proof is non-examinable, since it rather fiddly.

(♣) *Proof.* We prove the result in four steps.

1. Let us first fix some notation. Write $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$. Let π_1 and π_2 denote the two projections $\mathbb{R}^n \to \mathbb{R}^k$ and $\mathbb{R}^{n-k}$ respectively:

$$\pi_1(x^1, \dots, x^n) = (x^1, \dots, x^k), \qquad \pi_2(x^1, \dots, x^n) = (x^{k+1}, \dots, x^n),$$

and let $\jmath\colon \mathbb{R}^{n-k} \to \mathbb{R}^n$ denote the inclusion onto the *last* $n-k$ coordinates:

$$\jmath(x^1, \dots, x^{n-k}) = (0, \dots, 0, x^1, \dots, x^{n-k}).$$

Now let $\tau\colon V \to \tau(V) \subset \mathbb{R}^k$ denote a chart on N such that $\tau(y) = 0$. Fix a point $x \in L$ and let $\sigma\colon U \to \sigma(U) \subset \mathbb{R}^n$ denote a chart on M such that $\sigma(x) = 0$. Then $\tau \circ \varphi \circ \sigma^{-1}$ has maximal rank k at $0 \in \mathbb{R}^n$, and hence by part (ii) of the Implicit Function Theorem 5.3 there exists a chart h on $\mathbb{R}^n$, defined on an open ball O containing the origin such that

$$\tau \circ \varphi \circ \sigma^{-1} \circ h|_O = \pi_1|_O.$$

Shrinking O if necessary, we may assume $O = \pi_1(O) \times \pi_2(O) \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$. Set $\Omega := \pi_2(O)$. Then

$$\tau \circ \varphi \circ \sigma^{-1} \circ h \circ \jmath|_\Omega = \pi_1 \circ \jmath|_\Omega \equiv 0.$$

Thus if $\zeta := \sigma^{-1} \circ h \circ \jmath|_\Omega$ then $\zeta(\Omega) \subset L$.

2. We now prove that

$$\zeta(\Omega) = L \cap (\sigma^{-1} \circ h)(O). \tag{5.3}$$

Indeed,

$$\zeta(\Omega) = (\sigma^{-1} \circ h \circ \jmath)(\Omega) = (\sigma^{-1} \circ h)\big(O \cap \big(0 \times \mathbb{R}^{n-k}\big)\big) \subset L \cap (\sigma^{-1} \circ h)(O).$$

To see the other direction, if $z \in L \cap (\sigma^{-1} \circ h)(O)$ then $z = (\sigma^{-1} \circ h)(w)$ for a unique $w \in O$, and since

$$\pi_1(w) = \left(\tau \circ \varphi \circ \sigma^{-1} \circ h\right)(w) = \tau \circ \varphi(z) = 0,$$

we can write $w = (0, u)$ for a unique $u \in \Omega$. Then $z = \zeta(u)$. This proves the other inclusion, and hence establishes (5.3).

3. We now show that L is a smooth manifold. The equation (5.3) tells us that ζ maps Ω homeomorphically onto a neighbourhood of x in L in the subspace topology. Thus the inclusion $L \hookrightarrow M$ is a topological embedding. Set $W := \zeta(\Omega)$ and set $\rho := \zeta^{-1}$. Then $\rho \colon W \to \Omega$ is a chart on L. We claim that the collection of all such charts, as x ranges over L, determines a smooth structure on L. Indeed, suppose x_1 was another point in L with corresponding chart $\sigma_1 \colon U_1 \to \sigma_1(U_1) \subset \mathbb{R}^n$. Assume that $U \cap U_1 \neq \emptyset$. Let h_1 denote the corresponding diffeomorphism of $\mathbb{R}^n$, and define ζ_1 and ρ_1 similarly. Then by assumption $\sigma_1 \circ \sigma^{-1}$ is a diffeomorphism where defined, and hence so is $l := h_1^{-1} \circ \sigma_1 \circ \sigma^{-1} \circ h$. Moreover from (5.3) we can write $l(0, u) = (0, l_1(u))$ for l_1 a diffeomorphism defined on a neighbourhood of 0 in $\mathbb{R}^{n-k}$. Thus

$$\rho_1 \circ \rho^{-1} = \jmath^{-1} \circ l \circ \jmath = l_1$$

is a diffeomorphism where defined. Thus we have built a smooth structure on L.

4. To complete the proof, we show that the inclusion $\imath \colon L \hookrightarrow M$ is smooth. For this we note that with σ, ρ and h as above,

$$\sigma \circ \imath \circ \rho^{-1} = \sigma \circ \imath \circ \zeta = h \circ \jmath,$$

which is smooth. This completes the proof. ∎

Definition 5.14. A smooth map $\varphi \colon M^n \to N^k$ is called a **submersion** if every point of M is a regular point of φ, i.e. if $D\varphi(x)$ is surjective for every $x \in M$. Thus if φ is a submersion then by the Implicit Function Theorem 5.2, every point $x \in M$ belongs to the $(n-k)$-dimensional embedded submanifold $\varphi^{-1}(\varphi(x))$. A surjective submersion is necessarily a quotient map (see Lemma 24.7 for a proof of this fact).

Proposition 5.15. *Let $\varphi \colon M^n \to N^k$ be a smooth map and let $y \in N$ be a regular value of φ such that $L := \varphi^{-1}(y) \neq \emptyset$. Let $\imath \colon L \hookrightarrow M$ denote the inclusion. Then for all $x \in L$, one has*

$$D\imath(x)[T_x L] = \ker D\varphi(x).$$

Proof. By assumption both sides are linear subspaces of $T_x M$ of dimension $n - k$, so it suffices to show that $D\imath(x)[T_x L] \subset \ker D\varphi(x)$. For this take $f \in C^\infty(N)$ and $v \in T_x L$. Then by the chain rule (Proposition 4.2), one has

$$D\varphi(x) \circ D\imath(x)[v](f) = D(\varphi \circ \imath)(x)[v](f) = v(f \circ \varphi \circ \imath).$$

But $f \circ \varphi \circ \imath \in C^\infty(L)$ is the constant function $x \mapsto f(y)$ and hence by Corollary 3.5 one has $v(f \circ \varphi \circ \imath) = 0$. The result follows. ∎

Proposition 5.15 finally allows us to recover the "intuitive" definition of the tangent space for S^2 given at the beginning of Lecture 2.

Example 5.16. Let $f \colon \mathbb{R}^{n+1} \to \mathbb{R}$ be the map $f(x) = |x|^2 - 1$. It is straightforward to check that f is smooth and that the only critical point of f is $0 \in \mathbb{R}^{n+1}$. Thus by the Implicit Function Theorem 5.13, $S^n = f^{-1}(0)$ is a smooth manifold of dimension n. I leave it to you to check that this is the same smooth structure as the one given in Proposition 1.21. If we denote by $\imath \colon S^n \to \mathbb{R}^{n+1}$ the inclusion then (as you will check on Problem Sheet C), one has

$$D\imath(x)[T_x S^n] = \mathcal{J}_x(x^\perp), \tag{5.4}$$

where $\mathcal{J}_x \colon \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ was defined in Problem B.3 and

$$x^\perp := \left\{ y \in \mathbb{R}^{n+1} \mid \langle x, y \rangle = 0 \right\},$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean dot product. Now a moment's thought shows that (5.4) implies that the tangent space to S^n at a point x is the hyperplane tangent to S^n at x, as Figure 2.1 claimed.

We now state a version of Sard's Theorem valid for manifolds. We will only give a brief sketch of the proof, which is non-examinable. This theorem is the main reason we assume that manifolds have at most countably many components (the theorem is false if this condition is not imposed).

Theorem 5.17 (Sard's Theorem for Manifolds). *Let $\varphi \colon M^n \to N^k$ be a smooth map. The set of critical values of φ has measure zero and is nowhere dense. The set of regular values of φ is residual and thus dense in N. In particular, if $n < k$ then every point of M is necessarily a critical point of φ, and hence $N \setminus \varphi(M)$ is dense in N.*

(♣) *Proof.* The classical version of Sard's Theorem says that[1] if $O \subset \mathbb{R}^n$ is an open set and $f \colon O \to \mathbb{R}^k$ is a smooth map, then the set of critical values of f has measure zero in $\mathbb{R}^k$. Since manifolds have only countably many components, they can covered by countably many open sets that are diffeomorphic to balls in Euclidean spaces, cf. part (1) of Remark 1.9. Since the countable union of measure zero sets is also of measure zero, the result follows. ■

We conclude this lecture by briefly returning to the setting of the Implicit Function Theorem in Euclidean spaces (Theorem 5.3). Thus suppose O is a neighbourhood of 0 in $\mathbb{R}^n$ and $f \colon O \to \mathbb{R}^k$ is a smooth map. As Theorem 5.3 showed, if we assumed that the rank of f at 0 was maximal (and thus either equal to n or k, depending which was larger), then the rank of f was also maximal for all x near 0 too. Thus having maximal rank is an *open* condition.

If the rank is not maximal, then it can "jump", i.e. if f has (non-maximal) rank r at 0 then for x arbitrarily close to 0 the rank of f at x can be different to r. However if one adds as a hypothesis that the rank of f does *not* jump, then an analogous result to Theorem 5.3 holds. Here is a precise statement:

[1] A nice proof can be found in Chapter 3 of Milnor's classic textbook "Topology from a Differentiable Viewpoint".

THEOREM 5.18 (The Constant Rank Theorem). *Let O be a neighbourhood of* 0 *in* $\mathbb{R}^n$ *and suppose* $f\colon O \to \mathbb{R}^k$ *is a smooth map such that* $f(0) = 0$. *Assume that* f *has constant rank* r *for all* $x \in O$, *and let* $\vartheta\colon \mathbb{R}^n \to \mathbb{R}^k$ *denote the map*

$$\vartheta(x^1, \dots, x^n) = (x^1, \dots, x^r, 0, \dots, 0). \tag{5.5}$$

There exists chart g *about* 0 *on* $\mathbb{R}^k$ *and a chart* h *about* 0 *on* $\mathbb{R}^n$ *such that* $g \circ f \circ h = \vartheta$ *on a neighbourhood of* 0 *on* $\mathbb{R}^n$.

The proof is similar (albeit slightly messier) than Theorem 5.3, and we omit the details. Just as in Proposition 5.6, one can immediately translate this to a local statement about smooth maps between manifolds:

COROLLARY 5.19. *Let* $\varphi\colon M^n \to N^k$ *be a smooth map. Let* $x \in M$ *and assume there exists a neighbourhood of* x *such that* φ *has constant rank* r *on that neighbourhood. Then there exists a chart* σ *on* M *about* x *and a chart* τ *on* N *about* $\varphi(x)$ *such that* $\tau \circ \varphi \circ \sigma^{-1} = \vartheta$, *where* ϑ *is as in* (5.5).

Corollary 5.19 has the following global consequence, which will be useful in Lecture 9.

COROLLARY 5.20. *Let* $\varphi\colon M^n \to N^k$ *be a smooth map. Assume that the rank of* φ *is constant on all of* M. *If* φ *is either injective or surjective, then the rank is automatically maximal (and thus* φ *is either an immersion or a submersion respectively). In particular, a bijective constant rank map is necessarily a diffeomorphism.*

Proof. If φ has rank $r < n$ on a neighbourhood of x then it is clear that φ is not injective near x, since with charts σ and τ as in Corollary 5.19, the map $\vartheta = \tau \circ \varphi \circ \sigma^{-1}$ is not injective (since $(0, \dots, 0, \varepsilon)$ is mapped to $(0, \dots, 0)$ for all ε small). Meanwhile if φ has rank $r < k$ then[2] by Sard's Theorem 5.17 $\varphi(M)$ is nowhere dense in N, and thus certainly φ is not surjective. Finally, if φ is bijective, then φ must have rank $r = n = k$, and thus by the Inverse Function Theorem 5.2 φ is locally a diffeomorphism in a neighbourhood of every point, and hence also a global diffeomorphism. ∎

[2] Actually using Sard's Theorem here is overkill. A more elementary argument goes as follows: For any point $x \in M$, it follows from Corollary 5.19 that there is a neighbourhood U_x of x such that $\varphi(\overline{U}_x)$ is nowhere dense in N. Since M is Lindelöf (cf. part (ii) of Remark 1.9), we can cover M with countably many such sets U_x. Thus $\varphi(M)$ is the countable union of nowhere dense sets, and hence by the Baire Category Theorem (which is valid as any manifold is locally compact and Hausdorff, cf. part (i) of Remark 1.9), $\varphi(M)$ is itself nowhere dense in N, and thus in particular φ is not surjective.

The Whitney Theorems

In this lecture we will prove two famous theorems of Whitney. The first states that *every* smooth manifold can be embedded inside Euclidean space. Recall a continuous function $f\colon X \to Y$ between two topological spaces is **proper** if the preimage of any compact set in Y is compact in X. If X is compact and Y is Hausdorff then every continuous function is proper.

THEOREM 6.1 (The Whitney Embedding Theorem). *Let M be a smooth manifold of dimension n. Then there exists a proper embedding $\varphi\colon M \to \mathbb{R}^{2n}$.*

Theorem 6.1 is a genuinely difficult result. It is much easier to prove that M^n always embeds in $\mathbb{R}^{2n+1}$ (this is sometimes called the "Weak Whitney Embedding Theorem"). This is still too hard for us, however, so we will prove this only for the special case of compact manifolds. We call this the "Baby Whitney Embedding Theorem".

THEOREM 6.2 (The Baby Whitney Embedding Theorem). *Let M be a compact smooth manifold of dimension n. Then there exists a (proper) embedding $\varphi\colon M \to \mathbb{R}^{2n+1}$.*

The "proper" is in parentheses, as this is automatic when M is compact.

Proof. We prove the result in four steps.

1. We begin by showing that M admits an embedding into *some* Euclidean space $\mathbb{R}^k$ (this method will typically produce a very large k). In the next step we will reduce k down to $2n+1$. Since M is compact we can find a finite cover $\{V_1, \dots, V_r\}$ of open sets, with the property that there exist charts $\sigma_i\colon U_i \to O_i \subset \mathbb{R}^n$ for $i = 1, \dots, r$ with $\overline{V}_i \subset U_i$. Now let $\eta_i\colon M \to \mathbb{R}$ denote a smooth cutoff function (whose existence is guaranteed by Lemma 3.2) such that $\eta_i(\overline{V}_i) \equiv 1$, $0 \le \eta_i(x) \le 1$ for all $x \in M$ and $\operatorname{supp}(\eta_i) \subset U_i$. Set $f_i = \eta_i\,\sigma_i$, which we think of as a function from $M \to \mathbb{R}^n$ be extending by zero outside of U_i. Then define

$$\varphi\colon M \to \mathbb{R}^{nr+r}, \qquad \varphi(x) = \left(f_1(x), \dots, f_r(x), \eta_1(x), \dots, \eta_r(x)\right).$$

We claim that φ is an injective immersion. Since M is compact, it then follows (Problem C.4) that φ is an embedding. To see that φ is injective, suppose $\varphi(x) = \varphi(y)$. Since the sets $\{V_i\}$ cover M, there is some i such that $x \in V_i$, and hence $\eta_i(x) = 1$. Since $\varphi(x) = \varphi(y)$, we also have $\eta_i(y) = 1$, and thus $y \in \operatorname{supp}(\eta_i) \subset U_i$. Then also

$$\sigma_i(x) = \eta_i(x)\sigma_i(x) = f_i(x) = f_i(y) = \eta_i(y)\sigma_i(y) = \sigma_i(y).$$

But σ_i is a diffeomorphism, and hence in particular injective. Thus $x = y$.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Finally, to check φ is an immersion, pick an arbitrary $x \in M$. Then $x \in V_i$ for some i. Since $\eta_i \equiv 1$ on a neighbourhood of x, we have $Df_i(x) = D\sigma_i(x)$, which is injective. Thus also $D\varphi(x)$ is injective. This completes the proof of the weak version we wished to prove, where we took $k = nr + r$.

2. Replacing M by $\varphi(M)$, we now have $M \subset \mathbb{R}^k$. Assume that $k > 2n+1$, otherwise there is nothing to prove. Think of $\mathbb{R}^{k-1}$ as sitting inside $\mathbb{R}^k$ as the hyperplane $\{(x^1, \dots, x^k) \mid x^k = 0\}$. Given $v \in \mathbb{R}^k \setminus \mathbb{R}^{k-1}$, let $p_v \colon \mathbb{R}^k \to \mathbb{R}^{k-1}$ denote the projection parallel to v (i.e. the linear map with kernel equal to $\mathbb{R} \cdot v$). We will look for unit vectors v with the property that

$$p_v|_M \colon M \to \mathbb{R}^{k-1}$$

is an embedding. Using Problem C.4 again, it suffices to show that $p_v|_M$ is an injective immersion. But what that mean in this context? Saying that $p_v|_M$ is injective is saying that v is not parallel to any secant of M, that is,

$$v \neq \frac{x-y}{|x-y|}, \qquad \forall\, x, y \in M. \tag{6.1}$$

Here and elsewhere in this lecture, the norm $|\cdot|$ is taken with respect to the standard Euclidean norm. This is true both for points in M and for points in $TM \subset T\mathbb{R}^k = \mathbb{R}^{2k}$. The kernel of the linear map p_v is the line through v. Since p_v is linear, its derivative is the same linear map. Thus a tangent vector $w \in T_xM$ lies in the kernel of $Dp_v(x)$ if and only if w is parallel to v. We therefore see that p_v is an immersion if

$$v \neq \frac{w}{|w|}, \qquad \forall\, w \in T_xM,\ \forall\, x \in M. \tag{6.2}$$

3. We will use Sard's Theorem 5.17 to prove a v exists such that both (6.1) and (6.2) hold. For (6.1), consider the map

$$\psi \colon (M \times M) \setminus \Delta \to S^{k-1}, \qquad (x, y) \mapsto \frac{x-y}{|x-y|}.$$

Here Δ is the *diagonal* inside $M \times M$:

$$\Delta := \{(x, x) \mid x \in M\}.$$

Clearly v satisfies (6.1) if and only if v is *not* in the image of ψ. Note that $(M \times M) \setminus \Delta$ is an open set of $M \times M$, and thus $(M \times M) \setminus \Delta$ is a manifold of dimension $2n$ by Lemma 1.20 and Problem A.2. The map ψ is visibly smooth. Since $2n < k - 1 = \dim S^{k-1}$, by Sard's Theorem 5.17 the image of ψ is nowhere dense in S^{k-1}. Thus in particular, any non-empty open set of S^{k-1} contains a point v satisfying (6.1).

Now we consider (6.2). It suffices to check that it holds for all vectors w of norm 1. To this end we focus on the *unit tangent bundle*

$$SM := \{(x, w) \in TM \mid |w| = 1\}.$$

We will come back to unit tangent bundles next semester when we discuss Riemannian geometry. To see this is a manifold, consider the map $h \colon T\mathbb{R}^k \to \mathbb{R}$ given by $h(x,w) = |w|^2$. It is easy to see that 1 is a regular value of $h|_{TM}$ and that $SM = h|_{TM}^{-1}(1)$. By the Implicit Function Theorem SM is a manifold of dimension $2n-1$. It is easy to see that SM is compact (since M is). Now identify TM with a subset of $M \times \mathbb{R}^k$; then SM is identified with a subset of $M \times S^{k-1}$. Projecting onto the second factor, this gives us a map

$$SM \to M \times S^{k-1} \to S^{k-1}$$

which geometrically takes a unit vector based at a point in M and translates it to a unit vector based at the origin in $\mathbb{R}^k$. Using Sard's Theorem 5.17 again, the image of the composite map $SM \to S^{k-1}$ is nowhere dense. Since SM is compact, it follows that the complement—let us call it W—of the image is a dense open set in S^{k-1}. Thus W meets $S^{k-1} \cap (\mathbb{R}^k \setminus \mathbb{R}^{k-1})$ in an non-empty open set W_0. From what we already proved, such a non-empty open set W_0 contains a vector v which is not in the image of ψ.

4. We now complete the proof. The choice of v found above gives us an embedding $p_v \colon M \to \mathbb{R}^{k-1}$. If $k-1 = 2n+1$ we are done, if not then $2n+1 < k-1$, and the same argument again works to provide a new embedding in $\mathbb{R}^{k-2}$. By induction, we eventually obtain our desired embedding $M \to \mathbb{R}^{2n+1}$. ■

Remark 6.3. Extending Theorem 6.2 to cover all smooth manifolds (not just compact ones) is not that much more work. However it requires the concept of a *manifold with boundary* that we won't define until later on in the course, and hence we shall content ourselves with the compact case only. We emphasise though that the stronger result (Theorem 6.1, where $2n+1$ is reduced down to $2n$) is *much* harder.

Theorem 6.1 implies one could equivalently *define* a manifold as an embedded submanifold of Euclidean space (this is how manifolds are defined in most "baby" courses on differential geometry).

Definition 6.4 (Alternative definition of a manifold). Let $n \leq k$. A subset $M \subset \mathbb{R}^k$ is called a **smooth manifold of dimension** n if each point x in M has a neighbourhood V in $\mathbb{R}^k$ such that $M \cap V$ is diffeomorphic to an open set in $\mathbb{R}^n$.

In more detail, this means: for each point $x \in M$ there exists an open set $O \subset \mathbb{R}^n$ and a neighbourhood $V \subset \mathbb{R}^k$ of x, together with an injective smooth map $\zeta \colon O \to \mathbb{R}^k$ of maximal rank ($= n$) everywhere such that $\zeta(O) = M \cap V$ and $\sigma := \zeta^{-1}|_{M \cap V} \colon M \cap V \to O$ is continuous (where M is given the subspace topology of $\mathbb{R}^k$). One usually calls ζ a **parametrisation** of M. The inverse σ of ζ is then a chart on M in the normal sense. Note that if $n = k$ then this forces M to be an open subset of $\mathbb{R}^k$, and hence if M is compact then one necessarily has $n < k$.

Remark 6.5. Definition 6.4 is superficially much simpler than our original definition (Definition 1.18)—there is no need to first define topological manifolds, or even mention Hausdorff, paracompactness, etc. The equivalence of the definitions follows from Theorem 6.1 and the existence of slice charts (Definition 5.9). Moreover it is immediate from Definition 6.4 that manifolds are metrisable, since any subset of a metric space inherits a metric that determines its subspace topology.

You might therefore reasonably ask: was there any point in the abstract definition? The answer is of course "yes", as I will now try to explain.

Indeed, an embedded submanifold of Euclidean space should really be thought of as a pair (M, φ), where M is an (abstract) smooth manifold and φ is a choice of embedding. It is possible to embed a given manifold in many different ways (cf. Remark 5.11, also, if you can embed M in $\mathbb{R}^k$ then you can also embed M in $\mathbb{R}^l$ for any $l \geq k$), and a different choice of embedding can lead to dramatically different geometry (this will be particularly evident when we study *Riemannian geometry* next semester). Thus when proving results about embedded submanifolds, one always needs to ask the question: is this proof really a statement about the manifold itself, or does it depend on the embedding? This can often vastly complicate the proofs. The upshot is that having a more complicated definition leads to simpler proofs, and hence in the long run—since you only need to define things once but there are many theorems to prove!—it is better to work with the abstract definition whenever possible.

Still another reason to prefer the abstract definition is the following: One of the key applications of differential geometry in theoretical physics is Einstein's theory of *General Relativity*. Here one views the universe as 4-dimensional (curved) space-time. In the finite universe model, the spacial part of space-time is taken to be compact 3-dimensional hyperbolic manifold. Since (by definition) the universe is "everything", it doesn't make any sense *at all* to require the theory to begin by embedding the universe in a larger Euclidean space...

(♣) Remark 6.6. Here are some additional remarks about the (strong) Whitney Embedding Theorem 6.1:

(i) The Whitney Embedding Theorem is sharp in the sense that if $n = 2^r$ then $\mathbb{R}P^n$ cannot be embedded in $\mathbb{R}^{2n-1}$. This can be proved using characteristic classes (see Proposition 37.22).

(ii) There are various other versions of the Whitney Embedding Theorem. For instance, if M is a compact *orientable* smooth manifold of dimension n (we will define orientability in Lecture 20) then M embeds inside $\mathbb{R}^{2n-1}$. This does not contradict the previous statement, since for n even $\mathbb{R}P^n$ is not orientable.

(iii) In many cases the upper bound can be improved—for instance, we in Lecture 1 we saw that S^n embeds into $\mathbb{R}^{n+1}$. Another result (due to Haefliger) is that if M is a compact smooth manifold of dimension n whose homotopy groups $\pi_i(M)$ vanish for $i \leq k$ then if $2k + 3 \leq n$ one can embed M in $\mathbb{R}^{2n-k}$. In general, if $e(M)$ denotes the optimal k such that M embeds inside $\mathbb{R}^k$ then computing $e(M)$ is an open problem for many manifolds M.

We now aim to prove another theorem, also due to Whitney (Theorem 6.14), that allows us replace a continuous map with a smooth one. We begin with the following statement, which says a continuous function from a manifold to a Euclidean space can be approximated arbitrarily well by a smooth one.

Proposition 6.7. *Let M be a smooth manifold and let $h \colon M \to \mathbb{R}^k$ be a continuous function. Given any positive continuous function $\delta \colon M \to \mathbb{R}$, there exists a smooth function $f \colon M \to \mathbb{R}^k$ such that*

$$|f(x) - h(x)| < \delta(x), \qquad \forall\, x \in M.$$

Proof. Fix $x \in M$ and let U_x be a neighbourhood of x such that for all $y \in U_x$, one has

$$\delta(y) > \frac{1}{2}\delta(x), \qquad |h(y) - h(x)| < \frac{1}{2}\delta(x).$$

Such a neighbourhood exists as h and δ are assumed to be continuous. Then in particular we have that

$$|h(y) - h(x)| < \delta(y), \qquad \forall\, y \in U_x.$$

The collection $\{U_x \mid x \in M\}$ is an open cover of M. Let $\{\lambda_x \mid x \in M\}$ be a partition of unity subordinate to this open cover and define

$$f\colon M \to \mathbb{R}^k, \qquad f(y) := \sum_{x \in M} \lambda_x(y) h(x).$$

Recall that the right-hand side is actually a finite sum at every point, since $\{\mathrm{supp}(\lambda_x)\}$ is locally finite, and hence f is smooth. Moreover since $\sum_x \lambda_x \equiv 1$ and $\mathrm{supp}(\lambda_x) \subset U_x$, one has for any $y \in M$ that

$$\begin{aligned} |f(y) - h(y)| &= \left| \sum_{x \in M} \lambda_x(y) h(x) - h(y) \right| \\ &= \left| \sum_{x \in M} \lambda_x(y) h(x) - \sum_{x \in M} \lambda_x(y) h(y) \right| \\ &\leq \sum_{x \in M} \lambda_x(y) |h(y) - h(x)| \\ &< \sum_{x \in M} \lambda_x(y) \delta(y) = \delta(y). \end{aligned}$$

This completes the proof. ∎

Our aim now is to improve Proposition 6.7 to the case where the target space is another manifold, not a Euclidean space. The "obvious" tactic (given that we just proved the Whitney Embedding Theorem) is to embed the target manifold in a Euclidean space, and then approximate via the result we just proved. Unfortunately this doesn't quite work, as even though the function f can be chosen to be very close to h, it may still be the case that f "misses" our newly embedded manifold (remember an embedded manifold is not an open subset unless it is of full dimension). Thus we need a way to correct this. We will do this my making use of *tubular neighbourhoods*, which will be defined shortly.

DEFINITION 6.8. Let M^n be an embedded submanifold of $\mathbb{R}^k$. We define the **normal space to M at x** to be the $(k-n)$-dimensional subspace $\mathrm{Norm}_x M \subset T_x\mathbb{R}^k$ consisting of all vectors that are orthogonal to $T_x M$ with respect to the Euclidean dot product. We define the **normal bundle** of M as the set

$$\mathrm{Norm}(M) := \left\{ (x, v) \in T\mathbb{R}^k = \mathbb{R}^k \times \mathbb{R}^k \mid x \in M, v \in \mathrm{Norm}_x M \right\}.$$

On Problem Sheet C you proved that $\mathrm{Norm}(M)$ is an embedded k-dimensional submanifold of $T\mathbb{R}^k = \mathbb{R}^k \times \mathbb{R}^k$. We define a map

$$T\colon \mathrm{Norm}(M) \to \mathbb{R}^k, \qquad T(x,v) := x + v.$$

We emphasise that this only makes sense as M is embedded in $\mathbb{R}^k$. In general one cannot add points together on a manifold! The map T is smooth, since it is the restriction to $\mathrm{Norm}(M)$ of the addition map $\mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^k$. If o_M denotes the *zero-section*:

$$o_M := \{(x,0) \mid x \in M\}.$$

(the explanation of the name "zero section" will come in Lecture 10, when we discuss vector bundles) then one has

$$T(o_M) = M.$$

Thus it is reasonable to hope that a small neighbourhood of o_M in $\mathrm{Norm}(M)$ gets mapped under T to a small neighbourhood of M in $\mathbb{R}^k$. This motivates the following definition.

DEFINITION 6.9. A **tubular neighbourhood of** M is a neighbourhood U of M in $\mathbb{R}^k$ which is the diffeomorphic image under T of an open subset $V \subset \mathrm{Norm}(M)$ of the of form

$$V = \{(x,v) \in \mathrm{Norm}(M) \mid |v| < \varepsilon(x)\}, \tag{6.3}$$

where $\varepsilon\colon M \to \mathbb{R}$ is a strictly positive continuous function.

It is a non-trivial fact that such neighbourhoods always exist:

THEOREM 6.10 (The Tubular Neighbourhood Theorem). *Every embedded submanifold $M \subset \mathbb{R}^k$ admits a tubular neighbourhood.*

This proof is non-examinable, since I skipped it in class.

(♣) *Proof.* We prove the result in four steps.

1. We will prove that $DT(x,0)$ is invertible at every point $(x,0) \in o_M$. Since $T|_{o_M}\colon o_M \to M$ is obviously a diffeomorphism, one sees that $DT(x,0)$ maps $T_{(x,0)}o_M \subset T_{(x,0)}\mathrm{Norm}(M)$ isomorphically onto T_xM. Secondly, if we restrict T to the fibre Norm_xM, T just becomes the affine map $v \mapsto x + v$, and thus $DT(x,0)$ maps $T_{(x,0)}\mathrm{Norm}_xM$ isomorphically onto Norm_xM (cf Problem B.4).

Thus by the Inverse Function Theorem 5.2 we see that for each $x \in M$ there exists an $\varepsilon_x > 0$ such that if

$$U(x,\varepsilon_x) := \{(y,v) \in \mathrm{Norm}(M) \mid |x-y| < \varepsilon_x,\ |v| < \varepsilon_x\}$$

then $T|_{U(x,\varepsilon_x)}$ is a diffeomorphism. To complete the proof we need to show that there is open set of the form (6.3) on which T is a global diffeomorphism.

2. Let $\varepsilon\colon M \to \mathbb{R}$ be the function that assigns to a point $x \in M$ the supremun of all $\varepsilon \le 1$ such that T is a diffeomorphism on $U(x,\varepsilon)$. Then ε is strictly positive, as $\varepsilon(x) \ge \varepsilon_x$. We now claim that ε is actually a continuous function. Indeed, suppose $x, y \in M$ and suppose that $|x-y| < \varepsilon(x)$. Then for $\delta := \varepsilon(x) - |x-y|$, one has by the triangle inequality that $U(y,\delta) \subset U(x,\varepsilon(x))$, and hence $\varepsilon(y) \geq \varepsilon(x) - |x-y|$. Thus if $|x-y| < \varepsilon(x)$ then

$$\varepsilon(x) - \varepsilon(y) \leq |x-y|.$$

On the other hand, if $\varepsilon(x) \leq |x-y|$ then since $\varepsilon(y) \geq 0$ by definition, one trivially also has

$$\varepsilon(x) - \varepsilon(y) \leq |x-y|.$$

Reversing the roles of x and y shows that

$$|\varepsilon(x) - \varepsilon(y)| \leq |x-y|,$$

which proves ε is continuous.

3. Set

$$V := \left\{ (x,v) \in \mathrm{Norm}(M) \mid |v| < \frac{1}{2}\varepsilon(x) \right\}.$$

We claim that T is injective on V. Indeed, suppose (x,v) and (y,w) both belong to V and satisfy $x + v = T(x,v) = T(y,w) = y + w$. Without loss of generality, assume $\varepsilon(y) \leq \varepsilon(x)$. Then

$$|x-y| = |v-w| \leq |v| + |w| \leq \frac{1}{2}\varepsilon(x) + \frac{1}{2}\varepsilon(x) = \varepsilon(x),$$

where the first equality used $x + v = y + w$. Thus both (x,v) and (y,w) belong to $U(x,\varepsilon(x))$. But on this set T is injective by construction. Thus $(x,v) = (y,w)$ as required.

4. We complete the proof. Set $U := T(V)$. Then U is open as T is a local diffeomorphism. Since $T|_V$ is injective, we see that $T\colon V \to U$ is smooth bijection, and hence (as T is a local diffeomorphism), also a diffeomorphism. This completes the proof. ■

Remark 6.11. Next semester we will define another "tubular neighbourhood" associated to compact submanifold M of any Riemannian manifold (N,m). This is more general than the construction discussed here, since N does not have to be equal to a Euclidean space.

Definition 6.12. Let $Y \subset X$ be a subspace of a topological space. A **retraction** of X onto Y is a continuous map $r\colon X \to Y$ such that $r|_Y$ is the identity map on Y.

Corollary 6.13. *Let $M \subset \mathbb{R}^k$ be an embedded submanifold, and let U be a tubular neighbourhood of M. There exists a smooth map $r\colon U \to M$ which is both a retraction and a submersion.*

Proof. Let $T\colon V \subset \mathrm{Norm}(M) \to U$ be our tubular neighbourhood, and write $\pi\colon \mathrm{Norm}(M) \to M$ for the footpoint map that sends a pair (x,v) to x. Define $r\colon U \to M$ by $r := \pi \circ T^{-1}|_U$. Since $T|_V$ is a diffeomorphism and π is clearly a submersion, it follows that r is a submersion. Finally since $T(x,0) = x$, we see that $r(x) = \pi \circ T^{-1}(x) = x$, and hence r is a retraction. ■

The next result (also due to Whitney) allows us make contact with the topological world. Recall that if $h_0, h_1 \colon X \to Y$ are two continuous maps, we say they are **homotopic** if there exists a continuous map $H \colon X \times [0,1] \to Y$ such that $H(\cdot, 0) = h_0$ and $H(\cdot, 1) = h_1$. We can now state and prove another result due to Whitney. We will use this result later on in the course when we discuss the *homotopy invariance of de Rham cohomology* in Lecture 23.

THEOREM 6.14 (The Whitney Approximation Theorem). *Let $h \colon M \to N$ be a continuous map between two smooth manifolds. Then h is homotopic to a smooth map $\varphi \colon M \to N$.*

Proof. By the Whitney Embedding Theorem 6.1, we may assume that N is a properly embedded submanifold of some Euclidean space $\mathbb{R}^k$. Let U be a tubular neighbourhood of N, and let $r \colon U \to N$ be a smooth submersive retraction (whose existence is guaranteed by Corollary 6.13). Given $y \in N$, let

$$0 < \varepsilon(y) := \sup \{ \varepsilon \le 1 \mid B_\varepsilon(y) \subset U \},$$

where $B_\varepsilon(y)$ denotes the ball of radius ε about y (in the Euclidean norm). We claim that ε is actually a continuous function. This argument is essentially identical to the proof of Step 2 of Theorem 6.10, but we give it again anyway. So let $y, z \in N$ and first suppose that $|y - z| < \varepsilon(y)$. Then for $\delta := \varepsilon(y) - |y - z|$, one has by the triangle inequality that $B_\delta(z) \subset B_{\varepsilon(y)}(y)$, and hence $\varepsilon(z) \ge \varepsilon(y) - |y - z|$. Thus if $|y - z| < \varepsilon(y)$ then

$$\varepsilon(y) - \varepsilon(z) \le |y - z|.$$

On the other hand, if $\varepsilon(y) \le |y - z|$ then since $\varepsilon(z) > 0$ by definition, one trivially also has

$$\varepsilon(y) - \varepsilon(z) \le |y - z|.$$

Reversing the roles of y and z shows that

$$|\varepsilon(y) - \varepsilon(z)| \le |y - z|,$$

which proves ε is continuous. Now define $\delta := \varepsilon \circ h \colon M \to \mathbb{R}$. Then δ is continuous (as h and ε are) and positive (as ε is). By Proposition 6.7, there exists a smooth function $f \colon M \to \mathbb{R}^k$ such that

$$|f(x) - h(x)| < \delta(x), \qquad \forall\, x \in M.$$

Define

$$H \colon M \times [0,1] \to N, \qquad H(x,t) := r((1-t)h(x) + tf(x)).$$

This is well-defined due to our choice of function δ, which implies that $(1-t)h(x) + tf(x) \in U$ for all $t \in [0,1]$. Since r is the identity on $N \subset U$ and h takes values in N, we see that $H(\cdot, 0) = h$. Moreover if $\varphi := r \circ f$ then φ is smooth and $H(\cdot, 1) = \varphi$. This completes the proof. ■

We conclude this lecture with a a couple of non-examinable remarks. First, a definition.

Definition 6.15. Suppose M and N are smooth manifolds and $A \subset M$ is an arbitrary set. We say a map $\varphi\colon A \to N$ is **smooth on** A if it can be locally smoothly extended, i.e. if for every $x \in A$ there exists a neighbourhood U of x in M and a smooth map $\tilde{\varphi}\colon U \to N$ such that $\tilde{\varphi}|_{U \cap A} = \varphi$.

(♣) Remark 6.16. With a little bit more work, Theorem 6.14 can be improved to give the following statement: Suppose $h\colon M \to N$ is a continuous map between two smooth manifolds. Suppose $A \subset M$ is a closed set and $h|_A$ is already smooth (in the sense of Definition 6.15). Then the homotopy H can be chosen such that if $x \in A$ then $H(x,t) = h(x)$ for all t. In particular, the final smooth map $\varphi\colon M \to N$ satisfies $\varphi|_A = h|_A$.

(♣) Remark 6.17. One can also play the same game with smooth homotopies. Two smooth maps $\varphi, \psi\colon M \to N$ are **smoothly homotopic** if there exists a smooth map $M \times [0,1] \to N$ (note we are using Definition 6.15 again here to make sense of this) such that $H(\cdot, 0) = \varphi$ and $H(\cdot, 1) = \psi$. The homotopy version of the Whitney Approximation Theorem says that if two smooth maps are homotopic (in the normal sense) then they are also smoothly homotopic. Similarly if the given normal homotopy H from φ to ψ is stationary on some *closed* set A (i.e. $H(x,t) = \varphi(x)$ for all $x \in A$—note this implies $\varphi|_A \equiv \psi|_A$) then the approximating smooth homotopy can also be chosen to be stationary on A.

LECTURE 7

Vector fields

In this lecture we will define **vector fields**, which are smooth sections of the tangent bundle. We first introduce[1] the following standard notational convention, which will hold for the remainder of the course.

The Einstein Summation Convention. If the same index appears *exactly* twice in any monomial, written once as an *upper* index and once as a *lower* index, then that term is understood to be summed over all possible values of that index. Here are two examples:

1. If e_i denotes the standard ith basis vector in $\mathbb{R}^n$, then we write

$$v = a^i e_i \qquad \text{as an abbreviation for} \qquad v = \sum_{i=1}^n a^i e_i$$

2. If M is an n-dimensional smooth manifold, $x \in M$, (x^i) are local coordinates about x, and $v \in T_xM$, we write

$$v = a^i \frac{\partial}{\partial x^i}\Big|_x \qquad \text{as an abbreviation for} \qquad v = \sum_{i=1}^n a^i \frac{\partial}{\partial x^i}\Big|_x$$

Here $\frac{\partial}{\partial x^i}$ is understood to have i as a lower index, despite the fact that x^i has i as an upper index, because it is on the bottom of a fraction.

This convention will vastly simplify equations throughout the course. For instance, when we start to talk about tensors, we will have cause to consider quantities which have local expressions such as

$$A = A^{ij}_{kl} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \otimes dx^k \otimes dx^l,$$

which is much simpler than writing this abomination

$$A = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n A^{ij}_{kl} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \otimes dx^k \otimes dx^l.$$

The caveat is that in order for the convention to "work", the choice of whether to write a given quantity as an upper index or a lower index is *not* arbitrary.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] I delayed introducing this convention immediately, so that you could all see how cumbersome proofs with multiple summation signs could be (eg. Theorem 4.16) and thus fully appreciate the new convention!

Definition 7.1. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set (possibly equal to all of M). A **vector field** X on W is a smooth map $X \colon W \to TM$ (where we regard W as a smooth manifold in its own right) that satisfies the **section property**:

$$\pi(X(x)) = x, \qquad \forall\, x \in W, \tag{7.1}$$

where $\pi \colon TM \to M$ is the footpoint map.

In order to keep the notation under control, we will start to be a little sloppy when referring to points in the tangent bundle. If $v \in T_xM$, instead of writing (x, v) for the corresponding point in TM, we will sometimes continue to just write v. With this convention, one can think of a vector field as a smooth map $X \colon W \to TM$ such that $X(x) \in T_xM$ for each $x \in W$. We denote by $\mathfrak{X}(W)$ the set of all vector fields on W.

Let us give various equivalent ways of expressing what smooth means in this context. Suppose M has dimension n and let $\sigma \colon U \to O$ be a chart on M. Suppose $X \colon U \to TM$ is *any* function satisfying the section property (7.1) (not necessarily smooth). Let $x \in U$. Since $\left\{ \frac{\partial}{\partial x^i}\big|_x \mid i = 1, \dots, n \right\}$ is a basis of T_xM, we can write

$$X(x) = X^i(x) \frac{\partial}{\partial x^i}\bigg|_x, \tag{7.2}$$

(note here we are using the Einstein Summation Convention to omit the $\sum_{i=1}^n$) for some real numbers $X^i(x)$. If we do this for every point $x \in U$, we can think of the X^i as defining functions $X^i \colon U \to \mathbb{R}$. In general these functions X^i need not even be continuous, but as will shortly see, if X is smooth (i.e. a vector field on U) then the X^i are actually smooth functions.

Here is yet another way to think about it. Suppose $f \in C^\infty(U)$, and suppose as above X is any map $U \to TM$ satisfying the section property. Then for any $x \in U$, thinking of $X(x)$ as a derivation of $C^\infty(U)$ at x, we can feed f to $X(x)$ to get a number $X(x)(f)$. This gives us a function $X(f) \colon U \to \mathbb{R}$:

$$X(f)(x) := X(x)(f), \qquad \forall\, x \in U. \tag{7.3}$$

Once again, if X is just any map satisfying the section property then $X(f)$ will not in general even be continuous. However if X is smooth (i.e. a vector field) then $X(f)$ is smooth.

Proposition 7.2. *Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Let $X \colon W \to TM$ be any function satisfying the section property (7.1). Then the following are equivalent.*

(i) *X is a vector field on W.*

(ii) *If $\sigma \colon U \to O$ is any chart on M with $U \subset W$ then the functions X^i defined in (7.2) belong to $C^\infty(U)$.*

(iii) *If $V \subset W$ is any open set (possibly equal to all of W) and $f \in C^\infty(V)$ then the function $X(f)$ defined by (7.3) also belongs to $C^\infty(V)$.*

Proof. We begin with proving that (i) $\Leftrightarrow$ (ii). Let $x \in W$, and let $\sigma\colon U \to O$ be a chart about x. By definition, the function X^i defined in (7.2) is smooth if and only if $X^i \circ \sigma^{-1}$ is smooth in the normal sense. Note that by Proposition 3.9 and the definition of dx^i the function $X^i \circ \sigma^{-1}$ is the function

$$z \mapsto dx^i|_{\sigma^{-1}(z)}(X(\sigma^{-1}(z))), \qquad z \in O. \tag{7.4}$$

Now let us recall from the proof of Theorem 4.16 that a chart $\sigma\colon U \to O$ on M defines a chart $\tilde{\sigma}\colon \pi^{-1}(U) \to O \times \mathbb{R}^n$ on TM by[2]

$$\tilde{\sigma}(x, v) = \left(\sigma(x), dx^i|_x(v)\, e_i\right), \qquad x \in U,\ v \in T_xM.$$

By definition, X is smooth at x if and only if the composition

$$\tilde{\sigma} \circ X \circ \sigma^{-1}\colon O \to O \times \mathbb{R}^n$$

is smooth at $\sigma(x)$. Explicitly this is the map

$$O \to O \times \mathbb{R}^n, \qquad z \mapsto \left(z, dx^i|_{\sigma^{-1}(z)}(X(\sigma^{-1}(z)))\, e_i\right). \tag{7.5}$$

We see immediately that this map is smooth if and only if (7.4) is smooth for each $i = 1, \dots, n$. This proves (1) $\Leftrightarrow$ (2).

Now let us prove (2) $\Rightarrow$ (3). Let $V \subset W$ and let $f \in C^\infty(V)$. Choose a chart $\sigma\colon U \to O$ with $U \subset V$. Then for $x \in U$, we have that

$$X(f)(x) = X^i(x) \frac{\partial}{\partial x^i}\Big|_x (f)$$

The function $x \mapsto \frac{\partial}{\partial x^i}\big|_x(f)$ is smooth (this is just the function $x \mapsto D_i(f \circ \sigma^{-1})(\sigma(x))$. By (2) the X^i are also smooth functions, and hence $X(f)$ is a finite sum of the pointwise product of smooth functions and hence is smooth. We have proved $X(f)$ is smooth on U. But since U was arbitrary and smoothness is a local property, it follows that $X(f)$ is smooth on all of V. This proves (3).

Finally we prove (3) $\Rightarrow$ (2). Indeed, (2) is a special case of (3), since if $\sigma\colon U \to O$ is a chart about x with local coordinates x^i, then the function X^i defined in (7.2) is simply $X(x^i)$, where we think of the x^i as elements of $C^\infty(U)$. This completes the proof. ∎

Example 7.3. Suppose $\sigma\colon U \to O$ is a chart on M with local coordinates $x^i = u^i \circ \sigma$. Then we can think of $\frac{\partial}{\partial x^i}$ as defining a vector field on U via:

$$\frac{\partial}{\partial x^i}(x) := \frac{\partial}{\partial x^i}\Big|_x.$$

It is immediate from Proposition 7.2 that $\frac{\partial}{\partial x^i}$ is smooth.

We now introduce a notational convention that is both totally logical and somewhat confusing at the same time:

[2]Note how prettier this formula is with the Einstein Summation Convention in effect.

Definition 7.4. If $f \in C^\infty(U)$ then we denote the function $\frac{\partial}{\partial x^i}(f)$ from (7.3) with $X = \frac{\partial}{\partial x^i}$ by $\frac{\partial f}{\partial x^i}$. Thus $\frac{\partial f}{\partial x^i}$ is the function

$$\frac{\partial f}{\partial x^i}(x) := \frac{\partial}{\partial x^i}(x)(f) = \frac{\partial}{\partial x^i}\Big|_x (f) = D_i(f \circ \sigma^{-1})(\sigma(x)).$$

If our given manifold is an open subset of $\mathbb{R}^n$ and σ is the identity chart then the notation $\frac{\partial f}{\partial x^i}$ is consistent with the "usual" definition of partial derivative.

Let us now continue with the general case, where $W \subset M$ is any non-empty open subset. The space $\mathfrak{X}(W)$ is a real vector space under pointwise addition:

$$(X+Y)(x) := X(x) + Y(x), \qquad (cX)(x) := cX(x), \qquad X, Y \in \mathfrak{X}(W),\ c \in \mathbb{R}.$$

In fact, $\mathfrak{X}(W)$ forms a module over the ring $C^\infty(W)$ by defining

$$(fX)(x) := f(x)X(x), \qquad X \in \mathfrak{X}(W),\ f \in C^\infty(W).$$

In order for this to be well-defined, one needs to know that eg. $X+Y$ is smooth and fX is smooth. This however is immediate from Proposition 7.2.

Remark 7.5. Pay attention to the ordering. If $X \in \mathfrak{X}(W)$ and $f \in C^\infty(W)$ then $X(f)$ belongs to $C^\infty(W)$ whereas fX belongs to $\mathfrak{X}(W)$!

We now extend Definition 3.1 to derivations that are not based at a point.

Definition 7.6. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. A **derivation on** $C^\infty(W)$ is a linear map

$$\mathcal{X} \colon C^\infty(W) \to C^\infty(W)$$

satisfying the *derivation property*

$$\mathcal{X}(fg) = f\mathcal{X}(g) + g\mathcal{X}(f), \qquad \forall\, f, g \in C^\infty(W).$$

Let us temporarily denote by $\mathfrak{X}^{\mathrm{deriv}}(W)$ the set of derivations on W. Observe that $\mathfrak{X}^{\mathrm{deriv}}(W)$ is another module over $C^\infty(W)$. It follows from Proposition 3.3 that any vector field $X \in \mathfrak{X}(W)$ defines a derivation $\mathcal{X} \in \mathfrak{X}^{\mathrm{deriv}}(W)$ via

$$\mathcal{X}(f) := X(f)$$

In fact, the converse is true.

Proposition 7.7. *Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Then $\mathfrak{X}^{\mathrm{deriv}}(W)$ and $\mathfrak{X}(W)$ are isomorphic as modules over $C^\infty(W)$.*

Proof. Suppose $\mathcal{X}$ is a derivation on $C^\infty(W)$. Fix $x \in W$. Then $\mathcal{X}$ defines a derivation on $C^\infty(W)$ at x, which we suggestively write as $X(x)$, via the formula

$$X(x)(f) := \mathcal{X}(f)(x), \qquad \forall\, f \in C^\infty(W).$$

We can then think of X as defining a map $W \to TM$ by via $x \mapsto X(x)$. (Here we are using Proposition 3.3 repeatedly). We claim that X is smooth, and hence defines a vector field on W. For this by part (iii) of Proposition 7.2 we need only check that $X(f)$ is smooth for any $f \in C^\infty(W)$. But by construction $X(f) = \mathcal{X}(f)$, which is then smooth by assumption. ■

From now on we will identify a vector fields $X \in \mathfrak{X}(W)$ with the corresponding derivation $\mathcal{X}$ of $C^\infty(W)$ and write both with a Latin letters X. We will also abandon the notation $\mathfrak{X}^{\mathrm{deriv}}(W)$ and just write $\mathfrak{X}(W)$. Our next goal is to turn $\mathfrak{X}(W)$ into an algebra, that is, to have a bilinear operation

$$\mathfrak{X}(W) \times \mathfrak{X}(W) \to \mathfrak{X}(W).$$

The naive guess would be to try composition of derivations:

$$X \circ Y : C^\infty(W) \to C^\infty(W), \qquad (X \circ Y)(f) := X(Y(f)).$$

Unfortunately, this is *not* a derivation. Indeed, if we take $f, g \in C^\infty(W)$ and compute:

$$\begin{aligned}(X \circ Y)(fg) &= X\Big(fY(g) + gY(f)\Big) \\ &= \Big(f(X \circ Y)(g) + g(X \circ Y)(f)\Big) + \Big(X(f)Y(g) + X(g)Y(f)\Big)\end{aligned}$$

However, observe that the "error" term $X(f)Y(g) + X(g)Y(f)$ is symmetric in X and Y. This means that if we consider the **commutator**

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

then the error term cancels, and thus $[X, Y]$ is a derivation. We have thus justified the following definition.

DEFINITION 7.8. Let $X, Y \in \mathfrak{X}(W)$. Then the commutator $[X, Y] := X \circ Y - Y \circ X$ is another derivation. We call $[X, Y]$ the **Lie bracket** of X and Y.

REMARK 7.9. **Warning:** A few authors[3] define the Lie bracket with the *opposite* sign: $[X, Y] := Y \circ X - X \circ Y$. From a 'high-level" point of view, this other sign convention is actually the "correct" one, but this requires a little bit of infinite-dimensional Lie group theory to understand, as we will explain in Remark 10.24. The convention I am using, namely $[X, Y] := X \circ Y - Y \circ X$, is consistent with the majority of the literature.

The next proposition gives a formula for $[X, Y]$ in coordinates. The proof is deferred to Problem Sheet D.

PROPOSITION 7.10. *Let $\sigma : U \to O$ be a chart on M with local coordinates x^i, and let $X, Y \in \mathfrak{X}(U)$. Write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$. Then*

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j},$$

where $\frac{\partial Y^j}{\partial x^i}$ and $\frac{\partial X^j}{\partial x^i}$ are the functions from Definition 7.4.

In order to explain the name, we need an algebraic definition.

[3] Notably, Joel Robbin and Dietmar Salamon use the other sign convention in their wonderful lecture notes.

Definition 7.11. A (real) **Lie algebra** is a vector space $\mathfrak{g}$ endowed with a bilinear operation called the **Lie bracket**

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \qquad (v, w) \mapsto [v, w]$$

which in addition is antisymmetric, $[v, w] = -[w, v]$ and satisfies the **Jacobi identity**

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0, \qquad \forall\, u, v, w \in \mathfrak{g}.$$

Thus a Lie algebra is a non-associative algebra. The name "Lie" comes from the Norwegian mathematician Sophus Lie. It is traditional to write Lie algebras using fraktur symbols $\mathfrak{g}$ and $\mathfrak{h}$. The **dimension** of a Lie algebra $\mathfrak{g}$ is simply the dimension of $\mathfrak{g}$ as a vector space. If $\mathfrak{g}$ is a Lie algebra then a linear subspace $\mathfrak{h} \subset \mathfrak{g}$ is called a **Lie subalgebra** if $[v, w] \in \mathfrak{h}$ for all $v, w \in \mathfrak{h}$.

Example 7.12. Here are some examples of Lie algebras:

(i) The cross product $[x, y] := x \times y$ makes $\mathbb{R}^3$ into a 3-dimensional Lie algebra.

(ii) The set $\mathrm{Mat}(n)$ of $n \times n$ matrices is an n^2-dimensional Lie algebra under the normal commutator $[A, B] := AB - BA$.

(iii) If V is any vector space then we can turn V into a (rather boring) Lie algebra by defining $[v, w] := 0$. Such an Lie algebra is called **abelian**.

You will probably not be surprised to learn we have just constructed another example:

Theorem 7.13. *Let M be a smooth manifold and let $W \subset M$ be an open set. Then $\mathfrak{X}(W)$ is a Lie algebra.*

Proof. The only thing left to check is the Jacobi identity. This is Problem D.2 on Problem Sheet D. ∎

(♣) Remark 7.14. As long as $\dim M > 0$ then for any non-empty open subset W, $\mathfrak{X}(W)$ is an infinite-dimensional Lie algebra. To see this, we need only note that $\mathfrak{X}(W)$ is a module over $C^\infty(W)$, and $C^\infty(W)$ is an infinite-dimensional vector space (cf. Remark 2.2).

Let us now investigate how functions and vector fields can be "pushed forward" with a diffeomorphism.

Definition 7.15. Let $\varphi\colon M \to N$ be a diffeomorphism. We define an algebra homomorphism

$$\varphi_\star\colon C^\infty(M) \to C^\infty(N), \qquad f \mapsto \varphi_\star(f)$$

where

$$\varphi_\star(f) := f \circ \varphi^{-1}.$$

The claim that $\varphi_\star$ is an algebra homomorphism is just the assertion that

$$\varphi_\star(f + g) = \varphi_\star(f) + \varphi_\star(g), \qquad \varphi_\star(fg) = \varphi_\star(f)\varphi_\star(g), \qquad \varphi_\star(cf) = c\varphi_\star(f)$$

for all $f, g \in C^\infty(M)$ and $c \in \mathbb{R}$, which is immediate from the definitions.

Remark 7.16. **Warning!** Many authors use the notation $\varphi_\star$ for the derivative $D\varphi$.

Definition 7.17. Suppose $\varphi\colon M \to N$ is a diffeomorphism and $X \in \mathfrak{X}(M)$. We define the **pushforward** vector field $\varphi_\star(X) \in \mathfrak{X}(N)$ by defining

$$\varphi_\star(X)(y) := D\varphi(\varphi^{-1}(y))[X(\varphi^{-1}(y))].$$

To check this is well defined, we need to know that (a) $\varphi_\star(X)(y) \in T_yN$ for each $y \in N$, which is obvious, and (b) that $\varphi_\star(X)\colon N \to TN$ is smooth. The latter holds because it is simply the composition

$$N \xrightarrow{\varphi^{-1}} M \xrightarrow{X} TM \xrightarrow{D\varphi} TN$$

of smooth maps, and hence is smooth.

The map $\varphi_\star$ is again linear:

$$\varphi_\star(X+Y) = \varphi_\star(X) + \varphi_\star(Y), \qquad \forall\, X, Y \in \mathfrak{X}(M).$$

Moreover one has

$$\varphi_\star(fX) = \varphi_\star(f)\varphi_\star(X), \qquad \forall\, X \in \mathfrak{X}(M),\ \forall\, f \in C^\infty(M).$$

Remark 7.18. It may at first seem confusing that we have defined two different maps (one from functions to functions and one from vector fields to vector fields) and called them *both* $\varphi_\star$. The reason for this will become clear when we discuss the *tensor algebra* $\mathcal{T}(M)$ of a manifold. Roughly speaking, the tensor algebra is a big direct sum:

$$\mathcal{T}(M) = \bigoplus_{r,s\geq 0} \mathcal{T}^{r,s}(M)$$

where $\mathcal{T}^{r,s}(M)$ denotes the *tensors of type* (r,s). As we will eventually see, a tensor of type $(0,0)$ is simply a function (so $\mathcal{T}^{0,0}(M) = C^\infty(M)$) and a tensor of type $(1,0)$ is simply a vector field (so $\mathcal{T}^{1,0}(M) = \mathfrak{X}(M)$). Given a diffeomorphism $\varphi\colon M \to N$, in Lecture 18 we will construct a "master" morphism (see Definition 18.11)

$$\varphi_\star\colon \mathcal{T}(M) \to \mathcal{T}(N) \tag{7.6}$$

that preserves type, i.e.

$$\varphi_\star(\mathcal{T}^{r,s}(M)) \subset \mathcal{T}^{r,s}(N).$$

The map $\varphi_\star$ from Definition 7.15 is the restriction of the master $\varphi_\star$ from (7.6) to $\mathcal{T}^{0,0}(M) \subset \mathcal{T}(M)$ and the map $\varphi_\star$ from Definition 7.17 is the restriction of the master $\varphi_\star$ from (7.6) to $\mathcal{T}^{1,0}(M) \subset \mathcal{T}(M)$. Thus it makes sense to denote them both by $\varphi_\star$.

Definition 7.19. Let $\mathfrak{g}$ and $\mathfrak{h}$ be two Lie algebras. A **Lie algebra homomorphism** is a linear map $T\colon \mathfrak{g} \to \mathfrak{h}$ which respects the Lie brackets, i.e.

$$[Tv, Tw] = T[v,w], \qquad \forall\, v, w \in \mathfrak{g},$$

where the left-hand side is the Lie bracket in $\mathfrak{h}$ and the right-hand side is the Lie bracket in $\mathfrak{g}$. A **Lie algebra isomorphism** is a bijective Lie algebra homomorphism whose inverse is also a Lie algebra homomorphism.

Proposition 7.20. *Let $\varphi\colon M \to N$ be a diffeomorphism. Then $\varphi_\star\colon \mathfrak{X}(M) \to \mathfrak{X}(N)$ is a Lie algebra isomorphism.*

Proposition 7.20 is a special case of part (ii) of Problem D.5.

LECTURE 8

Flows and the Lie derivative

Let us begin this lecture by recalling two theorems from the theory of ordinary differential equations.

THEOREM 8.1 (Existence of solutions). *Let $O \subset \mathbb{R}^n$ be open and let $f\colon O \to \mathbb{R}^n$ be smooth. For any $z \in O$ there exists a neighbourhood V of z and an open interval (a, b) with $a < 0 < b$, together with a smooth map $h\colon (a, b) \times V \to O$ such that:*

(i) *$h(0, y) = y$, for all $y \in V$,*

(ii) *If we write*
$$\frac{d}{dt}h(t, y) := \lim_{s \to 0} \frac{h(t + s, y) - h(t, y)}{s}$$
then
$$\frac{d}{dt}h(t, y) = f(h(t, y)), \qquad \forall\, (t, y) \in (a, b) \times V.$$

Theorem 8.1 can be interpreted as follows. Suppose $\gamma = (\gamma^1, \dots, \gamma^n)\colon (a, b) \to O$ is a smooth curve. One calls γ an **integral curve** of $f = (f^1, \dots, f^n)$ if

$$(\gamma^i)'(t) = f^i \circ \gamma(t), \qquad \forall\, 1 \le i \le n. \tag{8.1}$$

Thus Theorem 8.1 tells us that integral curves $\gamma(t) = h(t, y)$ exist for arbitrary initial conditions $\gamma(0) = y$, and depend smoothly on their initial conditions. Moreover, they all locally exist for a common time (i.e. for every y in V, the integral curve with initial condition lasts for all $t \in (a, b)$. Next, we address uniqueness of solutions.

THEOREM 8.2 (Uniqueness of solutions). *Let $O \subset \mathbb{R}^n$ be open and let $f\colon O \to \mathbb{R}^n$ be smooth. If $\gamma, \delta\colon (a, b) \to O$ are two integral curves of f with $\gamma(t) = \delta(t)$ for some $t \in (a, b)$ then $\gamma \equiv \delta$.*

We will not prove either Theorem 8.1 or Theorem 8.2. They are both hopefully familiar to you from previous courses you took on ordinary differential equations. Instead, we will generalise them to manifolds.

DEFINITION 8.3. Let M be a manifold and let X be a vector field on M. Let $(a, b) \subset \mathbb{R}$ be an interval, and suppose $\gamma\colon (a, b) \to M$ is a smooth map. We say that γ is an **integral curve** of X if

$$\gamma'(t) = X(\gamma(t)), \qquad \forall\, t \in (a, b). \tag{8.2}$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Remark 8.4. This definition is consistent with the usual one (8.1) in the special case where $M = O$ is an open subset of $\mathbb{R}^n$. A vector field X on O canonically determines (and is determined by) a smooth function $f\colon O \to \mathbb{R}^n$ via $X(x) = \mathcal{J}_x(f(x))$, where $\mathcal{J}_x\colon \mathbb{R}^n \to T_x\mathbb{R}^n$ is the isomorphism from Problem B.3, and I leave it up to you to check that under the map $\mathcal{J}_{\gamma(t)}$, the condition (8.2) (which is an equality of tangent vectors in $T_{\gamma(t)}O$) is transformed into (8.1) (which is a series of equality of real numbers).

Before stating the next result, let us introduce a convention. If M is a manifold and (a,b) is an interval then $(a,b)\times M$ is also a manifold. Given $x \in M$ we denote by $\imath_x\colon (a,b) \to (a,b)\times M$ the map $\imath_x(t) := (t,x)$. We denote by

$$\frac{\partial}{\partial t}\bigg|_{(t,x)} := D\imath_x(t)\left[\frac{\partial}{\partial t}\bigg|_t\right] = \imath_x'(t) \tag{8.3}$$

the tangent vector in $T_{(t,x)}((a,b)\times M)$ obtained from the canonical generator $\frac{\partial}{\partial t}\big|_t \in T_t\mathbb{R}$. One can think of $(t,x) \mapsto \frac{\partial}{\partial t}\big|_{(t,x)}$ as defining a vector field on $(a,b)\times M$. (Exercise: Why?)

Theorem 8.5 (Local flow). *Let M be a smooth manifold and let $X \in \mathfrak{X}(M)$. For any $x \in M$ there exists a neighbourhood W of x and an interval (a,b) with $a < 0 < b$, together with a smooth map*

$$\theta^{\mathrm{loc}}\colon (a,b)\times W \to M.$$

such that

(i) $\theta^{\mathrm{loc}}(0,y) = y$, *for all* $y \in W$.

(ii) *For all* $(t,y) \in (a,b)\times W$ *one has*

$$D\theta^{\mathrm{loc}}(t,y)\left[\frac{\partial}{\partial t}\bigg|_{(t,y)}\right] = X(\theta^{\mathrm{loc}}(t,y)). \tag{8.4}$$

We call θ^{loc} a **local flow** of X. We will shortly get rid of the “loc”.

Proof. Let $\sigma\colon U \to O$ be a chart around x with local coordinates $(x^1,\dots,x^n)$. Let $\tilde{\sigma}\colon \pi^{-1}(U) \to O\times\mathbb{R}^n$ denote the corresponding chart on TM. Then we can write (cf. (7.5))

$$\tilde{\sigma}\circ X\circ\sigma^{-1} = (\mathrm{id}, f)$$

where $f\colon O \to \mathbb{R}^n$ is smooth. In fact[1], writing $f = (f^1,\dots,f^n)$ one has

$$f^i(z) := dx^i|_{\sigma^{-1}(z)}(X(\sigma^{-1}(z))).$$

Theorem 8.1 gives us a neighbourhood V of $\sigma(x)$, an interval (a,b), and a smooth map $h\colon (a,b)\times V \to O$ such that the two stated conditions holds. To complete the proof, set $W := \sigma^{-1}(V)$ and define

$$\theta^{\mathrm{loc}}(t,y) := \sigma^{-1}\circ h(t,\sigma(y)), \qquad (t,y) \in (a,b)\times W.$$

That θ^{loc} satisfies the two required conditions is immediate from the fact that h did. ∎

[1] One should be aware of the slightly confusing fact that when we think of $X(x)$ as a tangent vector, we usually write it as a column vector. When we think of X (or in this case, its local representative f) as a function, we normally write it as a row vector!

Remark 8.6. The condition (8.4) is simpler that it looks. Given $y \in W$, set $\gamma_y(t) := \theta^{\mathrm{loc}}(t, y)$, so that $\gamma_y \colon (a, b) \to U$ is a curve in our manifold. Then from how velocity vectors are defined (cf. Definition 4.6), one has

$$D\theta^{\mathrm{loc}}(t, y) \left[\frac{\partial}{\partial t}\bigg|_{(t,y)} \right] = (\theta^{\mathrm{loc}} \circ \imath_y)'(t) = \gamma_y'(t).$$

Thus (8.4) asserts that γ_y is an integral curve of X.

A similar argument also proves the manifold version of Theorem 8.2:

Theorem 8.7. *Let M be a smooth manifold and let $X \in \mathfrak{X}(M)$. If $\gamma, \delta \colon (a, b) \to M$ are two integral curves of X with $\gamma(t) = \delta(t)$ for some $t \in (a, b)$ then $\gamma \equiv \delta$.*

Thanks to Theorem 8.7, it makes sense to talk about the maximal integral curve through a given point.

Definition 8.8. Let X be a vector field on M. Given a point $x \in M$, we denote by $\big(t^-(x), t^+(x)\big)$ the maximal interval around 0 on which the (unique by Theorem 8.7) integral curve $\gamma_x \colon \big(t^-(x), t^+(x)\big) \to M$ of X whose initial condition is $\gamma_x(0) = x$ is defined. We call γ_x the **maximal integral curve through** x.

Remark 8.9. It readily follows from maximality that for any point $x \in M$ one has

$$t^{\pm}(\gamma_x(s)) = t^{\pm}(x) - s, \qquad \forall\, s \in \big(t^-(x), t^+(x)\big). \tag{8.5}$$

This will useful later in establishing the group property for flows, see Remark 8.11.

We emphasise that $\big(t^-(x), t^+(x)\big)$ typically will be larger than the interval (a, b) given by Theorem 8.5—indeed, by construction $\theta^{\mathrm{loc}}(t, x)$ never leaves the open set U that the chart σ was defined on. Thus whilst $\gamma_x(t) = \theta^{\mathrm{loc}}(t, x)$ for small enough t, in general the curve γ_x could wander all over the manifold. Our aim now is to globalise Theorem 8.5 so that the equality $\gamma_x(t) = \theta(t, x)$ holds whenever the former is defined.

Theorem 8.10 (Maximal flow). *Let M be a smooth manifold and let $X \in \mathfrak{X}(M)$. There exists a unique open set $\mathcal{D} \subset \mathbb{R} \times M$ and a unique smooth map $\theta \colon \mathcal{D} \to M$ such that*

(i) *For all $x \in M$ one has*

$$\mathcal{D} \cap \big(\mathbb{R} \times \{x\}\big) = \big(t^-(x), t^+(x)\big) \times \{x\}.$$

(ii) *$\theta(t, x) = \gamma_x(t)$ for all $(t, x) \in \mathcal{D}$.*

We call θ the **flow** of X. If there is more than one vector field under consideration we write θ^X.

Proof. Note that (i) determines $\mathcal{D}$ uniquely, and (ii) does the same for θ. It remains therefore to show that $\mathcal{D}$ is open and θ is smooth.

Fix $x \in M$ and let I denote the set of all $t \in \big(t^-(x), t^+(x)\big)$ for which there exists some neighbourhood of (t, x) contained in $\mathcal{D}$ on which θ is smooth. Note that

I is clearly open by definition. (Its defining property is an open condition). We will prove that I is nonempty and closed, whence it follows that $I = \big(t^-(x), t^+(x)\big)$.

Firstly, I is non-empty, since $0 \in I$ by Theorem 8.5. Now suppose $t_0 \in \overline{I}$. Set $x_0 := \gamma_x(t_0)$. We apply Theorem 8.5 at the point x_0 to obtain a local flow $\theta^{\mathrm{loc}} \colon (a,b) \times U_0 \to M$ about x_0.

Since t_0 belongs to the closure of I, we may choose $t_1 \in I$ close enough to t_0 such that $t_0 - t_1$ belongs to (a,b) and such that $\gamma_x(t_1)$ belongs to U_0 (here we are using the fact that γ_x is continuous at t_0 and that U_0 is a neighbourhood of x_0).

Since (a,b) is an interval, we can do a little better: we can choose an interval I_0 about t_0 such that $t - t_1 \in (a,b)$ for all $t \in I_0$. Finally, by continuity of θ at (t_1, x), there exists a neighbourhood V of x such that $\theta(\{t_1\} \times V) \subset U_0$.

We now claim that our original θ is defined and smooth on all of $I_0 \times V$, so that in particular $t_0 \in I$. Indeed, if $t \in I_0$ and $y \in V$ then $t - t_1 \in (a,b)$ and $\theta(t_1, y) \in U_0$. Thus $\theta^{\mathrm{loc}}(t - t_1, \theta(t_1, y))$ is defined and smooth. But the curve $s \mapsto \theta^{\mathrm{loc}}(s - t_1, \theta(t_1, y))$ is an integral curve of X which passes through $\theta(t_1, y)$ at t_1. By uniqueness, this curve is $\theta(t, y)$. We thus see that $\theta(t,y) = \theta^{\mathrm{loc}}(t - t_1, \theta(t_1, y))$ is defined and smooth at (t, y).

We have thus shown that for all $x \in M$ and for all $t \in \big(t^-(x), t^+(x)\big)$, there exists a neighbourhood of (t,x) in $\mathcal{D}$ on which θ is smooth. Thus $\mathcal{D}$ is open and $\theta \colon \mathcal{D} \to M$ is smooth. This completes the proof. ■

Remark 8.11. We can play the same game and partition M up. Given $t \in \mathbb{R}$, let

$$M_t := \{x \in M \mid (t, x) \in \mathcal{D}\}.$$

Then M_t is open in M and there is a well-defined smooth map $\theta_t \colon M_t \to M_{-t}$ given by

$$\theta_t(x) := \theta(t, x), \qquad x \in M_t$$

(the fact that θ_t takes values in M_{-t} follows from (8.5)). This map θ_t is a diffeomorphism, since $\theta_{-t} \colon M_{-t} \to M_t$ is an inverse. More generally, if $s, t \in \mathbb{R}$ then the domain of $\theta_s \circ \theta_t$ is contained in (though not necessarily equal to) M_{s+t}. If s and t have the same sign then we have equality. In any case, one has $\theta_s \circ \theta_t = \theta_{s+t}$ on the domain of $\theta_s \circ \theta_t$.

Restricting the maps θ_t to open subsets of M is annoying. The next condition rules this out.

Definition 8.12. A vector field X is **complete** if the set $\mathcal{D}$ from Theorem 8.10 is all of $\mathbb{R} \times M$. Equivalently, a vector field is complete if either (a) its integral curves exist for all time or (b) the maps $\theta_t(x) := \theta(t, x)$ are all diffeomorphisms of the entire manifold M.

Definition 8.13. We write $\mathrm{Diff}(M)$ for the set of diffeomorphisms $\varphi \colon M \to M$. Note that $\mathrm{Diff}(M)$ is actually a group under composition, where the identity element is just the identity map.

(♣) Remark 8.14. Assume that M is compact. Then more is true: the group $\mathrm{Diff}(M)$ can itself be given a (Fréchet) manifold structure. We will say more about this Remark 10.24.

Definition 8.15. A **one-parameter group of diffeomorphisms** is a smooth[2] group homeomorphism $\mathbb{R} \to \mathrm{Diff}(M)$. Writing this as $t \mapsto \theta_t$, the group property tells us that

$$\theta_0 = \mathrm{id}, \qquad \theta_{s+t} = \theta_s \circ \theta_t, \qquad \forall\, s,t \in \mathbb{R}.$$

If $\{\theta_t\}$ is a one-parameter group of diffeomorphisms then we define its **infinitesimal generator** as the (necessarily complete) vector field

$$X(x) := D\theta(0,x)\left[\frac{\partial}{\partial t}\Big|_{(0,x)}\right], \tag{8.6}$$

where we wrote $\theta(t,x) := \theta_t(x)$ and used the convention from (8.3). Then the flow of X is simply the one-parameter group θ_t. We have thus proved:

Proposition 8.16. *Let M be a smooth manifold. Then there is a bijective correspondence between one-parameter subgroups of diffeomorphisms and complete vector fields.*

Example 8.17. Perhaps the easiest example of a *non-complete* vector field is given by taking $M = \mathbb{R}^2 \setminus 0$ and taking $X = \frac{\partial}{\partial x^1}$. If $(x^1, x^2) \in \mathbb{R}^2 \setminus 0$ then the flow line passing through (x^1, x^2) takes the form: $(x^1, x^2) \mapsto (t + x^1, x^2)$. It is then obvious that something must go wrong if you take $(x^1, x^2) = (-1, 0)$ and try and flow forwards—indeed, if the flow existed for all time then at time $t = 1$ you would fall out the manifold through the hole... (Exercise: Make this rigorous.)

For the remainder of this lecture (and indeed, the course), we will switch between the notations $\theta(t,x)$, $\theta_t(x)$ and $\gamma_x(t)$ whenever convenient. Here is an easy way to guarantee completeness.

Lemma 8.18. *Let X be a vector field on M. Assume there exists $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \subset \big(t^-(x), t^+(x)\big)$ for all $x \in M$. Then X is complete.*

Proof. If not, there exists some $x \in M$ such that either $t^+(x) < \infty$ or $t^-(x) > -\infty$. Assume the former (the proof in the other case is almost identical). Choose a number t_0 such that $t^+(x) - \varepsilon < t_0 < t^+(x)$. Set $x_0 := \gamma_x(t_0)$. By assumption $\gamma_{x_0}(t)$ is defined for all $t \in (-\varepsilon, \varepsilon)$. Now consider the curve

$$\gamma(t) := \begin{cases} \gamma_x(t), & t^-(x) < t < t^+(x), \\ \gamma_{x_0}(t - t_0), & t_0 - \varepsilon < t < t_0 + \varepsilon. \end{cases}$$

These two definitions agree on the overlap, since

$$\gamma_{x_0}(t - t_0) = \theta_{t-t_0}(x_0) = \theta_{t-t_0} \circ \theta_{t_0}(x) = \theta_t(x) = \gamma_x(t).$$

But then γ is an integral curve for X with initial condition x which is defined on $\big(t^-(x), t_0 + \varepsilon\big)$. Since $t_0 + \varepsilon > t^+(x)$, this contradicts the maximality of $t^+(x)$. ∎

[2] Here "smooth" should be interpreted as saying that $t \mapsto \theta_t$ is a smooth map from the manifold $\mathbb{R}$ to the (infinite-dimensional) manifold $\mathrm{Diff}(M)$. In more down-to-earth language, this just means that $(t,x) \mapsto \theta_t(x)$ is a smooth function $\mathbb{R} \times M \to M$. See also Remark 10.24.

We define the **support** of a vector field X in exactly the same way as we define the support of a function:

$$\operatorname{supp}(X) := \overline{\{x \in M \mid X(x) \neq 0 \in T_xM\}}.$$

COROLLARY 8.19. *Let X be a vector field with compact support. Then X is complete.*

Proof. By Theorem 8.5 for each $x \in \operatorname{supp}(X)$ there exists a neighbourhood U_x of x and an interval $(-\varepsilon_x, \varepsilon_x)$ such that the flow is defined on $(-\varepsilon_x, \varepsilon_x) \times U_x$. Since $\operatorname{supp}(X)$ is compact, we may select finitely many points $x_1, \dots, x_N$ such that $\operatorname{supp}(X) \subset \bigcup_{i=1}^N U_{x_i}$. Now set $\varepsilon := \min_{i=1,\dots,N} \varepsilon_{x_i}$. Then for every $x \in \operatorname{supp}(X)$ one has $(-\varepsilon, \varepsilon) \subset \big(t^-(x), t^+(x)\big)$. Since X is identically zero on $M \setminus \operatorname{supp}(X)$, every integral curve of X starting at some point in $M \setminus \operatorname{supp}(X)$ is trivially defined for all $t \in \mathbb{R}$ (and is constant). Thus the hypotheses of Lemma 8.18 are satisfied, and the proof is complete. ∎

COROLLARY 8.20. *If M is compact then every vector field on M is complete.*

Proof. If M is compact then certainly every vector field has compact support. ∎

Here is another variant on Lemma 8.18 which is sometimes more useful.

LEMMA 8.21. *Let X be a vector field on M. If the maximal domain of an integral curve γ_x is not all of $\mathbb{R}$, then the image of that curve cannot be contained in any compact subset of M.*

Proof. Assume for instance that $t^+(x) < \infty$ and that γ_x is contained in a compact set K. Choose a sequence t_n such that $t_i \to t^+(x)$ from below. By compactness, $\gamma_x(t_i)$ converges to some point x_0. By Theorem 8.5, a local flow θ^{loc} of X is defined on $(-\varepsilon, \varepsilon) \times U$ for some $\varepsilon > 0$ and some neighbourhood U of x_0. Choose i large enough so that $\gamma_x(t_i) \in U$ and $t_i + \varepsilon > t^+(x)$. Then arguing just as in the proof of Lemma 8.18, the curve

$$\gamma(t) := \begin{cases} \gamma_x(t), & t^-(x) < t < t^+(x), \\ \theta^{\mathrm{loc}}(t - t_i, \gamma_x(t_i)), & t_i - \varepsilon < t < t_i + \varepsilon. \end{cases}$$

is a well-defined integral curve of X starting at x, and thus contradicts the maximality of $t^+(x)$. ∎

We now move onto defining the **Lie derivative** associated to a vector field X. As with maps $\varphi_\star$ from the last lecture, we will actually give two definitions, one for the Lie derivative eating a function, and one for the Lie derivative eating a vector field. When we discuss tensors (cf. Remark 7.18), these two definitions will eventually be unified to give one "master" Lie derivative that eats any tensor and spits out another tensor of the same type.

DEFINITION 8.22. Let $X \in \mathfrak{X}(M)$ with flow θ_t. We define the **Lie derivative** of X to be the map

$$\mathcal{L}_X \colon C^\infty(M) \to C^\infty(M)$$

given by

$$\mathcal{L}_X(f)(x) := \lim_{t \to 0} \frac{f \circ \theta_t(x) - f(x)}{t}.$$

To see that this is well-defined (i.e. why the limit exists and defines a smooth function), we prove:

LEMMA 8.23. $\mathcal{L}_X(f) = X(f)$.

Proof. From the definitions one has

$$X(f)(x) = X(x)(f) = \gamma_x'(0)(f) = (f \circ \gamma_x)'(0).$$

But then clearly

$$(f \circ \gamma_x)'(0) = \lim_{t \to 0} \frac{f \circ \gamma_x(t) - f(x)}{t} = \lim_{t \to 0} \frac{f \circ \theta_t(x) - f(x)}{t}.$$

■

Now we define the Lie derivative on vector fields.

DEFINITION 8.24. Let $X \in \mathfrak{X}(M)$ have flow θ_t. We define the **Lie derivative** of X to be the map

$$\mathcal{L}_X : \mathfrak{X}(M) \to \mathfrak{X}(M)$$

by

$$\mathcal{L}_X(Y)(x) := \lim_{t \to 0} \frac{D\theta_{-t}(\theta_t(x))[Y(\theta_t(x))] - Y(x)}{t}. \tag{8.7}$$

To see that this is well-defined (i.e. why the limit exists and defines a vector field) we prove:

THEOREM 8.25. *For any two vector fields on M, one has $\mathcal{L}_X(Y) = [X, Y]$.*

Theorem 8.25 also explains the name "Lie derivative". Before going any further, let us emphasise once more: the main point of the Lie derivative is that we will eventually extend this to an operator $\mathcal{L}_X : \mathcal{T}(M) \to \mathcal{T}(M)$ on the tensor algebra of M (see Theorem 18.18 in Lecture 18). For now though, we can simply think of the Lie derivative of giving more insight into the Lie bracket; an example of this is Proposition 8.27 below. The proof of Theorem 8.25 requires a preliminary lemma, which can be thought of as a manifold version of Lemma 3.8.

LEMMA 8.26. *Let $U \subset M$ be open and let $a < 0 < b$. Let $f : (a, b) \times U \to \mathbb{R}$ be a smooth function such that $f(0, x) = 0$ for all $x \in U$. Then there exists another smooth function $h : (a, b) \times U \to \mathbb{R}$ such that*

$$f(t, x) = th(t, x), \qquad \frac{\partial}{\partial t}\Big|_{(0,x)}(f) = h(0, x), \qquad \forall\, (t, x) \in (a, b) \times U.$$

Here $(t, x) \mapsto \frac{\partial}{\partial t}\big|_{(t,x)}$ is the vector field on $(a, b) \times U$ defined in (8.3).

Proof. Simply define

$$h(t, x) := \int_0^1 \frac{\partial}{\partial t}\Big|_{(st,x)}(f)\, ds.$$

Then h is smooth. To see that $f(t,x) = th(t,x)$ one considers the curve $\gamma(s) \coloneqq f \circ \imath_x(st)$. Then

$$f(t,x) = f(t,x) - f(0,x) = \gamma(1) - \gamma(0) = \int_0^1 \gamma'(s)\,ds.$$

But by definition

$$\gamma'(s) = t\frac{\partial}{\partial t}\Big|_{(st,x)}(f).$$

This completes the proof. ■

We now prove Theorem 8.25.

Proof of Theorem 8.25. Fix $x \in M$. By Theorem 8.5, there exists $a < 0 < b$ and a neighbourhood U of x such that $(a,b) \times U \subset \mathcal{D}$, the domain of θ. Now fix $g \in C^\infty(M)$. We apply Lemma 8.26 to the function $f(t,y) \coloneqq g(\theta_t(y)) - g(y)$ to obtain a function h, which, writing $h_t(y) \coloneqq h(t,y)$, we have:

$$g \circ \theta_t = g + th_t, \qquad h_0 = X(g),$$

where we used Lemma 8.23. Thus for another vector field Y we have

$$\begin{aligned} D\theta_{-t}(\theta_t(x))[Y(\theta_t(x))](g) &= Y(\theta_t(x))(g \circ \theta_{-t}) \\ &= Y(\theta_t(x))(g - th_{-t}) \\ &= Y(g) \circ \theta_t(x) - tY(h_{-t}) \circ \theta_t(x). \end{aligned}$$

We therefore have

$$\begin{aligned} \mathcal{L}_X(Y)(g)(x) &= \lim_{t\to 0} \frac{Y(g) \circ \theta_t(x) - (Y(g))(x)}{t} - \lim_{t \to 0} Y(h_{-t}) \circ \theta_t(x) \\ &= \mathcal{L}_X(Y(g))(x) - Y(h_0)(x) \\ &= X(Y(g))(x) - Y(X(g))(x) \\ &= [X,Y](g)(x). \end{aligned}$$

Since x and g were arbitrary, this completes the proof. ■

An application of Theorem 8.25 is the following result, whose proof is deferred to Problem Sheet E.

Proposition 8.27. *Let X and Y be vector fields on M with flows θ_t^X and θ_t^Y respectively. Then $[X,Y] \equiv 0$ if and only if the two flows commute, i.e. $\theta_t^X \circ \theta_s^Y = \theta_s^Y \circ \theta_t^X$ for all s,t small.*

LECTURE 9

Lie groups

In the next four lectures we will cover the basic theory of *Lie groups*. Lie groups are important in many areas of mathematics (not just geometry!)—including representation theory, harmonic analysis, differential equations and more. Lie groups also crop up naturally in physics—both classically (eg. Noether's theorem that every smooth symmetry of a physical system has a corresponding conservation law), and in high-energy particle physics, via *gauge theory*. We will come back to gauge theory in Differential Geometry II when we study connections on principal bundles.

DEFINITION 9.1. A **Lie group** G is a smooth manifold that is also a group in the algebraic sense, with the property that the group multiplication

$$m\colon G \times G \to G, \qquad m(a,b) = ab,$$

and group inversion

$$i\colon G \to G, \qquad i(a) = a^{-1},$$

are both smooth maps.

REMARK 9.2. In contrast to manifolds, where we normally use the letters x and y to denote points, for Lie groups we will use a and b. This distinction will become important next lecture when we discuss Lie groups *acting* on manifolds.

DEFINITION 9.3. A **Lie group homomorphism** $\varphi\colon G \to H$ is a smooth map $G \to H$ which is also a group homomorphism. A **Lie group isomorphism** is a Lie group homomorphism which is also a diffeomorphism (and thus the inverse is automatically a Lie group homomorphism.)

EXAMPLE 9.4. Here are some examples of Lie groups.

(i) $\mathbb{R}^n$ is a Lie group under addition.

(ii) $\mathbb{R} \setminus \{0\}$ is a Lie group under multiplication.

(iii) The set $\mathrm{GL}(n)$ of invertible $n \times n$ matrices is a Lie group under matrix multiplication. Indeed, it is a manifold of dimension n^2 (cf. Problem A.1). Multiplication is smooth because the matrix entries of a product AB are given by polynomials in the entries of A and B, and inversion is smooth by Cramer's rule.

(iv) If G is a Lie group and $H \subset G$ is an **open subgroup** (that is, a subgroup which is also an open set in G) then H naturally inherits a Lie group structure (cf. Proposition 1.20). Thus the set $\mathrm{GL}^+(n)$ of invertible matrices with positive determinant is a Lie group.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

(v) The n-torus $T^n = \mathbb{R}^n/\mathbb{Z}^n$ is an abelian Lie group (the group structure is induced by addition on $\mathbb{R}^n$). In fact, one can show that any compact abelian Lie group is (isomorphic to) a torus.

(vi) The same underlying smooth manifold can carry multiple Lie group structures. For instance, a different Lie group structure on $\mathbb{R}^3$ is given by

$$m(x,y) := \big(x^1+y^1, x^2+y^2, x^3+y^3+x^1y^2\big), \qquad x = (x^1,x^2,x^3),\ y = (y^1,y^2,y^3).$$

This is the **Heisenberg group**. In order to see that this does indeed define a group structure, we identify $\mathbb{R}^3$ with upper triangular 3×3 matrices:

$$(x^1,x^2,x^3) \qquad \longleftrightarrow \qquad \begin{pmatrix} 1 & x^1 & x^3 \\ 0 & 1 & x^2 \\ 0 & 0 & 1 \end{pmatrix}$$

The group multiplication m corresponds to normal matrix multiplication.

(vii) Not all smooth manifolds can be made into Lie groups. For instance, S^n admits a Lie group structure only for $n = 0, 1$ and 3. The reason for this is briefly discussed in Remark 12.14.

Going back to the general theory, we will usually denote group composition simply by juxtaposition, and write e for the identity element. (Exception: if G is a group of matrices then we denote the identity element by I).

Definition 9.5. Let G be a Lie group and let $a \in G$. We let $l_a\colon G \to G$ and $r_a\colon G \to G$ denote the **left translation** and **right translation** by a respectively

$$l_a(b) := ab, \qquad r_a(b) = ba.$$

These maps are both diffeomorphisms. For instance, $l_a = m \circ \imath_a$, where $\imath_a\colon G \to G\times G$ is the map $b \mapsto (a,b)$, and hence is the composition of smooth maps. Moreover $l_{a^{-1}}$ is the inverse of l_a. A similar argument applies for r_a.

Remark 9.6. Throughout this lecture (and indeed, this course), we will almost exclusively work with left translations. This is purely a convention—everything we do could be reformulated (with appropriate modifications) to work with right translations instead. See also Remark 10.7 in the next lecture.

Proposition 9.7. *Every Lie group homomorphism has constant rank.*

Proof. Let $\varphi\colon G \to H$ be a Lie group homomorphism. Fix $a \in G$. We show that φ has the same rank at a as it does at e. Indeed, since φ is a homomorphism, for all $b \in G$ one has

$$\varphi(l_a(b)) = \varphi(ab) = \varphi(a)\varphi(b) = l_{\varphi(a)}(\varphi(b)),$$

that is,

$$\varphi \circ l_a = l_{\varphi(a)} \circ \varphi.$$

Now differentiate both sides at e and use the chain rule for manifolds (Proposition 4.2) to obtain

$$D\varphi(a) \circ Dl_a(e) = Dl_{\varphi(a)}(e) \circ D\varphi(e). \tag{9.1}$$

Since l_a and $l_{\varphi(a)}$ are diffeomorphisms, both $Dl_a(e)$ and $Dl_{\varphi(a)}$ are linear isomorphisms. The claim follows. ∎

Corollary 9.8. *A Lie group homomorphism is a Lie group isomorphism if and only if it is bijective.*

Proof. This follows immediately from Proposition 9.7 and Corollary 5.20. ∎

Definition 9.9. Let G be a Lie group. A **Lie subgroup** of G is a subgroup H endowed with a topology and a smooth structure that simultaneously makes H into a Lie group and into an *immersed* submanifold of G.

In fact, *embedded* submanifolds are automatically Lie subgroups.

Proposition 9.10. *Let G be a Lie group, and let $H \subset G$ be a subgroup which is an embedded submanifold. Then H is a Lie subgroup.*

Proof. We need to check that H is a Lie group in its own right. Thus for instance we must show that the group multiplication $m \colon H \times H \to H$ is smooth. For this we need the following two facts:

- If $M \subset N$ is an immersed submanifold and $\varphi \colon N \to L$ is smooth then $\varphi|_M \colon M \to L$ is also smooth. (*Proof*: the inclusion map $\imath \colon M \to N$ is smooth by definition of an immersed submanifold, and $\varphi|_M = \varphi \circ \imath$.)
- If $M \subset N$ ia an embedded submanifold and $\varphi \colon L \to N$ is a smooth map with $\varphi(L) \subset M$ then $\varphi \colon L \to M$ is also smooth. (*Proof*: This is immediate from the definition of the subspace topology.)

Going back to the proof, from the first bullet point, $m|_{H\times H} \colon H \times H \to G$ is smooth. Since H is a subgroup, $m(H \times H) \subset H$. By the second bullet point, $m|_{H\times H} \colon H \times H \to H$ is smooth. A similar argument applies for inversion. ∎

The following result is much deeper. It is not that difficult to prove, but it would take the entire lecture (and then some), so we will skip it.

Theorem 9.11 (The Closed Subgroup Theorem). *Let G be a Lie group and suppose H is any subgroup of G. The following are equivalent:*

(i) *H is a closed subgroup (i.e. H is a closed set in G).*

(ii) *H is an embedded submanifold of G.*

(iii) *H is an embedded Lie subgroup of G.*

Clearly (iii) implies (ii). Proposition 9.10 proved that (ii) implies (iii). On Problem Sheet E you are asked to prove that (ii) implies (i). The trickier bit is to show that (i) implies (ii), and this is what we will skip.

Example 9.12. Let $\mathrm{O}(n) \subset \mathrm{GL}(n)$ denote the orthogonal matrices, i.e. those matrices A with $AA^T = I$. Then $\mathrm{O}(n)$ is closed in $\mathrm{GL}(n)$, and hence by the Closed Subgroup Theorem 9.11, it is a Lie subgroup.

Due to its importance, let us give a direct proof of this example.

Proposition 9.13. *The set of orthogonal matrices $\mathrm{O}(n)$ is a Lie subgroup of $\mathrm{GL}(n)$ of dimension $\frac{1}{2}n(n-1)$.*

Proof. By Proposition 9.10 we need only show that $\mathrm{O}(n)$ is an embedded submanifold. For this first consider the set $\mathrm{Sym}(n)$ of *symmetric* matrices. Clearly we can identify $\mathrm{Sym}(n) \cong \mathbb{R}^{\frac{n(n+1)}{2}}$, and thus $\mathrm{Sym}(n)$ is naturally a smooth manifold. Consider the (obviously smooth) map $\varphi \colon \mathrm{GL}(n) \to \mathrm{Sym}(n)$ given by

$$\varphi(A) := AA^T,$$

where A^T denotes the transpose of A. Then $\mathrm{O}(n) = \varphi^{-1}(I)$, where I is the $n \times n$ identity matrix. Thus by the Implicit Function Theorem 5.13, we need only show that I is regular value of φ, whence $\mathrm{O}(n)$ is an embedded submanifold of $\mathrm{GL}(n)$ of dimension $n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$. For $A \in \mathrm{O}(n)$ one has $\varphi \circ r_A = \varphi$, and thus $D\varphi(A) \circ Dr_A(I) = D\varphi(I)$. Since r_A is a diffeomorphism, it follows that the rank of φ at A is the same as the rank of φ at I. Thus we need only show that φ has maximal rank at I, i.e. that $D\varphi(I)$ is surjective.

Since $\mathrm{GL}(n)$ is an open subset of the vector space $\mathrm{Mat}(n)$, its tangent space is canonically identified with $\mathrm{Mat}(n)$, and similarly $T_I \mathrm{Sym}(n) \cong \mathrm{Sym}(n)$. Using the isomorphism $\mathcal{J}_I$ from Problem B.3, if $A \in \mathrm{Mat}(n)$ then[1]

$$\begin{aligned} D\varphi(I)(\mathcal{J}_I(A)) &= \frac{d}{dt}\Big|_{t=0} \varphi(I + tA) \\ &= \frac{d}{dt}\Big|_{t=0} \left(I + t(A + A^T) + t^2 AA^T\right) \\ &= A + A^T \end{aligned}$$

Now fix an arbitrary $S \in \mathrm{Sym}(n) \cong T_I \mathrm{Sym}(n)$. To complete the proof we need to find $A \in \mathrm{Mat}(n)$ such that $A + A^T = S$. But this is easy: take $A := \frac{1}{2}S$. ∎

Remark 9.14. A **matrix Lie group** is a closed subgroup of $\mathrm{GL}(n)$. Thus a matrix Lie group is necessarily a Lie group in its own right, by the Closed Subgroup Theorem 9.11. As in Proposition 9.13, it is typically possible to prove this directly. Another example of this is on Problem Sheet E.

Definition 9.15. Let G be a Lie group. We define the **Lie algebra** of G, which we will usually write[2] as $\mathfrak{g}$, as the tangent space to G at the identity element e:

$$\mathfrak{g} := T_e G.$$

Of course, for this definition not to be completely insane, the Lie algebra of a Lie group better be a Lie algebra (in the sense of Definition 7.11). Luckily this is indeed the case, as we will shortly prove in Corollary 9.20.

To make sure we are all on the same page, let us fix once and for all our notation for matrix Lie groups.

[1]This makes sense as $I + tA$ necessarily belongs to $\mathrm{GL}(n)$ for t small enough (Exercise: Why?)

[2]The convention is that the Lie algebra of a given Lie group is written with the same letter, only with a lower case Fraktur letter. Thus the Lie algebra of H is $\mathfrak{h}$, and the Lie algebra of K is $\mathfrak{k}$, etc.

Notation: Let V and W be vector spaces.

(i) We write $\mathrm{L}(V, W)$ for the vector space of all linear maps from V to W.

(ii) We write $\mathrm{GL}(V) \subset \mathrm{L}(V, V)$ for the open set of all invertible linear maps.

(iii) We write $\mathfrak{gl}(V) := T_I \,\mathrm{GL}(V)$ for its Lie algebra. Since $\mathrm{GL}(V)$ is an open subset of $\mathrm{L}(V, V)$, the map $\mathcal{J}_I$ defines a canonical isomorphism

$$\mathcal{J}_I \colon \mathrm{L}(V, V) \to \mathfrak{gl}(V).$$

Often the isomorphism $\mathcal{J}_I$ will be suppressed from the notation and we write $\mathrm{L}(V, V) \cong \mathfrak{gl}(V)$ (or even $\mathrm{L}(V, V) = \mathfrak{gl}(V)$!)

(iv) Other matrix Lie subgroups of $\mathrm{GL}(V)$ are written as you would guess. For instance, if V is endowed with an inner product $\langle \cdot, \cdot \rangle$ then $\mathrm{O}(V) \subset \mathrm{GL}(V)$ denotes the linear transformations that preserve $\langle \cdot, \cdot \rangle$, and $\mathfrak{o}(V) := T_I \,\mathrm{O}(V)$.

Finally, in the special case $V = \mathbb{R}^n$, we write $\mathrm{Mat}(n) := \mathrm{L}(\mathbb{R}^n, \mathbb{R}^n)$, $\mathrm{GL}(n) := \mathrm{GL}(V)$ and $\mathfrak{gl}(n) := \mathfrak{gl}(V)$.

Here are some examples.

Example 9.16. Here are some examples of Lie algebras of Lie groups.

(i) The Lie algebra of $\mathrm{GL}(n)$ is $\mathfrak{gl}(n) \cong \mathrm{Mat}(n)$.

(ii) The Lie algebra of $\mathrm{O}(n)$ is

$$\mathfrak{o}(n) := \left\{ A \in \mathfrak{gl}(n) \mid A + A^T = 0 \right\}.$$

This follows from Proposition 9.13 together with Proposition 5.15.

(iii) The Lie algebra of the T^n is $\mathbb{R}^n$. Indeed, for $n = 1$ this is clear, and for $n > 1$ this follows from Problem C.5. More generally, the Lie algebra of any abelian Lie group is an abelian (and the converse holds if the Lie group is connected), as you will prove on Problem Sheets E and G.

The key to proving that the Lie algebra of a Lie group is indeed a Lie algebra is the following concept.

Definition 9.17. Let G be a Lie group. A vector field $X \in \mathfrak{X}(G)$ is said to be **left-invariant** if $(l_a)_\star(X) = X$ for all $a \in G$. Equivalently, this means that

$$Dl_a(b)[X(b)] = X(ab), \qquad \forall\, a, b \in G.$$

We denote by $\mathfrak{X}_l(G) \subset \mathfrak{X}(G)$ the set of left-invariant vector fields.

It is immediate that $\mathfrak{X}_l(G)$ is a linear subspace of $\mathfrak{X}(G)$. In fact, much more is true: the Lie bracket of two left-invariant vector fields is again left-invariant:

Proposition 9.18. *Let G be a Lie group and let $X, Y \in \mathfrak{X}_l(G)$. Then $[X, Y]$ also belongs to $\mathfrak{X}_l(G)$. Consequently $\mathfrak{X}_l(G)$ is a Lie subalgebra of $\mathfrak{X}(G)$.*

Proof. Fix $a \in G$. Then by Proposition 7.20 one has

$$(l_a)_\star[X, Y] = [(l_a)_\star(X), (l_a)_\star(Y)] = [X, Y].$$

Since a was arbitrary, the result follows. ■

The next result is the main step needed to show that $\mathfrak{g}$ is a Lie algebra.

Theorem 9.19. *Let G be a Lie group with Lie algebra $\mathfrak{g} = T_eG$. The evaluation map*

$$\mathrm{eval}_e \colon \mathfrak{X}_l(G) \to \mathfrak{g}, \qquad \mathrm{eval}_e(X) = X(e)$$

is a vector space isomorphism. Thus $\mathfrak{X}_l(G)$ is a vector space of the same dimension as G.

Proof. The map eval_e is clearly linear. If $\mathrm{eval}_e(X) = 0$ then X is identically zero, since for any $a \in G$ one has by left-invariance.

$$X(a) = Dl_a(e)[X(e)] = 0.$$

Thus we need only show that eval_e is surjective. For this, fix an arbitrary $v \in \mathfrak{g} = T_eG$. We define a map $X_v \colon G \to TG$ by

$$X_v(a) := Dl_a(e)[v]. \tag{9.2}$$

Then X_v certainly satisfies the section property (7.1), since $Dl_a(e) \colon T_eG \to T_aG$. To show that X_v is a vector field, it suffices to show that $X_v(f)$ is smooth for any $f \in C^\infty(G)$. For this, choose a smooth curve $\gamma \colon (-\varepsilon, \varepsilon) \to G$ such that $\gamma(0) = e$ and $\gamma'(0) = v$. Then following through the definitions, for any $a \in G$ one has

$$X_v(f)(a) = X_v(a)(f) = Dl_a(e)[v](f) = v(f \circ l_a) = (f \circ l_a \circ \gamma)'(0).$$

The curve $f \circ l_a \circ \gamma$ is given by $t \mapsto f(m(a, \gamma(t)))$. Since f, γ and m are all smooth, this is smooth.

Next, we claim X_v is left-invariant. Indeed, if $a, b \in G$ then

$$Dl_a(b)[X_v(b)] = Dl_a(b) \circ Dl_b(e)[v] = D(l_a \circ l_b)(e)[v] = Dl_{ab}(e)[v] = X_v(ab).$$

Thus $X_v \in \mathfrak{X}_l(G)$. Since $\mathrm{eval}_e(X_v) = X_v(e) = v$, this shows eval_e is surjective, and thus completes the proof. ■

Corollary 9.20. *Let G be a Lie group of dimension n. Then its Lie algebra is a Lie algebra (!) of dimension n.*

Proof. We need only define a Lie bracket on $\mathfrak{g}$. For this, using the notation from Theorem 9.19, we simply set

$$[v, w] := \mathrm{eval}_e([X_v, X_w]), \qquad v, w \in \mathfrak{g}.$$

This works by Theorem 9.19 and Proposition 9.18. ■

For the next result, let us recall from Problem D.5 that if $\varphi\colon M \to N$ is a smooth map between manifolds, and $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ then we say that X and Y are **φ-related** if $D\varphi(x)[X(x)] = Y(\varphi(x))$ for all $x \in M$.

Proposition 9.21. *Let $\varphi\colon G \to H$ be a Lie group homomorphism between two Lie groups. Then $D\varphi(e)\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.*

Proof. Let $v \in \mathfrak{g}$ and let $X_v \in \mathfrak{X}_l(G)$ denote the unique left-invariant vector field such that $X_v(e) = v$. Let $w := D\varphi(e)[v]$ and let $Y_w \in \mathfrak{X}_l(H)$ denote the unique left-invariant vector field such that $Y_w(e) = w$. We claim that X_v and Y_w are φ-related. Indeed, by (9.1) one has

$$D\varphi(a)[X_v(a)] = D\varphi(a)\circ Dl_a(e)[v] = Dl_{\varphi(a)}(e)\circ D\varphi(e)[v] = Dl_{\varphi(a)}(e)[w] = Y_w(\varphi(a)).$$

Now by part (ii) of Problem D.5, if $v_1, v_2 \in \mathfrak{g}$ and $w_i := D\varphi(e)[v_i]$ then $[X_{v_1}, X_{v_2}]$ is φ-related to $[Y_{w_1}, Y_{w_2}]$, and evaluating both sides at e gives

$$D\varphi(e)[v_1, v_2] = [w_1, w_2].$$

This completes the proof. ■

Suppose now H is a Lie subgroup of G. Let $\imath\colon H \hookrightarrow G$ denote the inclusion. Then since $D\imath(e)\colon \mathfrak{h} = T_eH \to \mathfrak{g} = T_eG$ is injective, we can regard $\mathfrak{h}$ as a linear subspace of $\mathfrak{g}$. A priori however, this identification might not respect the Lie brackets. Thanks to Proposition 9.21, however, it does:

Corollary 9.22. *Let $H \subset G$ be a Lie subgroup, and identify $\mathfrak{h}$ with its image in $\mathfrak{g}$. The Lie bracket on $\mathfrak{h}$ is simply the restriction of the Lie bracket on $\mathfrak{g}$ to $\mathfrak{h}$. Thus $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.*

Proof. Apply Proposition 9.21 with φ the inclusion (note that the roles of H and G have been reversed!) ■

Let us go back to $\mathrm{GL}(n)$. We now have potentially two different Lie brackets on $\mathfrak{gl}(n) \cong \mathrm{Mat}(n)$: the one coming from Corollary 9.20, and the commutator bracket (cf. part (ii) of Example 7.12). The next result (whose proof is deferred to Problem Sheet E) tells us that these coincide.

Proposition 9.23. *The Lie bracket on $\mathfrak{gl}(n)$ is given by matrix commutation, i.e.*

$$[A, B] = AB - BA, \qquad \forall\, A, B \in \mathfrak{gl}(n) \cong \mathrm{Mat}(n).$$

Combining Corollary 9.22 and Proposition 9.23, we end up with:

Corollary 9.24. *Let G be a matrix Lie group. Then the Lie bracket on $\mathfrak{g}$ is given by matrix commutation.*

We conclude this lecture by stating the following result, which tells us that the Lie group-Lie algebra correspondence goes both ways. The proof will be given in Lecture 12 after we have proved the Frobenius Theorem in Lecture 11.

Theorem 9.25. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$ then there is a unique connected Lie subgroup H of G whose Lie algebra is $\mathfrak{h}$.*

The exponential map

In this lecture we define the **exponential map** of a Lie group G, which will be a map $\exp\colon \mathfrak{g} \to G$. The reason for the name will become apparent in Proposition 10.14—namely, the exponential map of a matrix Lie group is given by matrix exponentiation.

PROPOSITION 10.1. *Let G be a Lie group and let $X \in \mathfrak{X}_l(G)$. Then X is complete.*

Proof. By Theorem 8.5 there exists some $\varepsilon > 0$ such that the integral curve $\gamma_e(t)$ of X with initial condition e is defined on $(-\varepsilon, \varepsilon)$. Now observe that $l_a \circ \gamma_e$ is an integral curve of X starting at a, and hence is equal to γ_a. Thus γ_a is also defined on $(-\varepsilon, \varepsilon)$. The claim now follows from Lemma 8.18. ∎

DEFINITION 10.2. Let G be a Lie group with Lie algebra $\mathfrak{g}$. A **one-parameter subgroup** of G is a Lie group homomorphism $\mathbb{R} \to G$.

REMARK 10.3. In Definition 8.15 we defined a *one-parameter subgroup of diffeomorphisms*. The two concepts are related, as we will discuss in Remark 10.24 at the end of the lecture.

PROPOSITION 10.4. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $v \in \mathfrak{g}$, and let $X_v \in \mathfrak{X}_l(G)$ denote the unique left-invariant vector field with $X_v(e) = v$ (defined as in (9.2)). Let $\gamma^v = \gamma^v_e \colon \mathbb{R} \to G$ denote the integral curve of X_v with $\gamma^v(0) = e$ (this is defined on all of $\mathbb{R}$ by Proposition 10.1). Then γ^v is a one-parameter subgroup. Moreover if $\gamma\colon \mathbb{R} \to G$ is any one-parameter subgroup then $\gamma = \gamma^v$ for some $v \in \mathfrak{g}$.*

Proof. To show that γ^v is a one-parameter subgroup we must show that

$$\gamma^v(s+t) = \gamma^v(s)\gamma^v(t)$$

for all $s, t \in \mathbb{R}$, where on the right-hand side we use multiplication in G For this, consider the curve

$$\delta(t) := \gamma^v(s)^{-1}\gamma^v(s+t).$$

Then $\delta(0) = e$, and by the chain rule

$$\begin{aligned}
\delta'(t) &= Dl_{\gamma^v(s)^{-1}}(\gamma^v(s+t))[(\gamma^v)'(s+t)] \\
&= Dl_{\gamma^v(s)^{-1}}(\gamma^v(s+t))[X_v(\gamma^v(s+t))] \\
&\overset{(\dagger)}{=} X_v\big(\gamma^v(s)^{-1}\gamma^v(s+t)\big) \\
&= X_v(\delta(t)).
\end{aligned}$$

where $(\dagger)$ used left-invariance. Thus by uniqueness of integral curves, one must have $\delta(t) = \gamma^v(t)$.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Conversely, suppose γ is a one-parameter subgroup. Let $v := \gamma'(0) \in \mathfrak{g}$. We claim that $\gamma'(t) = X_v(\gamma(t))$. Since $\gamma(t+s) = \gamma(t)\gamma(s) = l_{\gamma(t)}\gamma(s)$, we have

$$\begin{aligned}\gamma'(t) &= \frac{d}{ds}\Big|_{s=0}\gamma(t+s) \\ &= \frac{d}{ds}\Big|_{s=0} l_{\gamma(t)}\gamma(s) \\ &= Dl_{\gamma(t)}(\gamma(0))[\gamma'(0)] \\ &= Dl_{\gamma(t)}(e)[v] \\ &= X_v(\gamma(t)).\end{aligned}$$

where the last line used the definition of X_v. Thus again by uniqueness of integral curves we have $\gamma \equiv \gamma^v$. This completes the proof. ∎

We can play a similar game by replacing $v \in \mathfrak{g}$ with a scalar multiple sv.

Lemma 10.5. *For any $s, t \in \mathbb{R}$ one has*

$$\gamma^v(st) = \gamma^{sv}(t).$$

Proof. First note as $Dl_a(e)$ is a linear map one has for any $a \in G$ that

$$X_{sv}(a) = Dl_a(e)[sv] = sDl_a(e)[v] = sX_v(a).$$

Thus $X_{sv} = sX_v$. Now by the chain rule

$$\frac{d}{dt}\Big|_{t=t_0}\gamma^v(st) = s(\gamma^v)'(st_0) = sX_v(\gamma^v(st_0)) = X_{sv}(\gamma^v(st_0)).$$

Thus $t \mapsto \gamma^v(st)$ is an integral curve of X_{sv} with initial condition e, and hence by uniqueness of integral curves once more, one has $\gamma^v(st) \equiv \gamma^{sv}(t)$. ∎

Let us now write $\theta^v_t \colon G \to G$ for the flow of X_v. Thus by definition $\gamma^v(t) = \theta^v_t(e)$.

Proposition 10.6. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $\gamma \colon \mathbb{R} \to G$ be a smooth curve with $\gamma(0) = e$ and $\gamma'(0) = v \in \mathfrak{g}$. The following are equivalent:*

(i) *γ is a one-parameter subgroup,*

(ii) *$\gamma(t) = \gamma^v(t)$,*

(iii) *$\theta^v_t = r_{\gamma(t)}$.*

Proof. We already know that condition (i) is equivalent to condition (ii). To see that (iii) implies (ii) observe that if (iii) holds then

$$\gamma^v(t) = \theta^v_t(e) = r_{\gamma(t)}e = \gamma(t).$$

so that condition (ii) holds. Now assume that (ii) holds, and fix $a \in G$. Then since X_v is left-invariant, as in the proof of Proposition 10.1 $a\gamma^v = l_a \circ \gamma^v$ is another integral curve of X_v with initial condition a, and thus by uniqueness of integral curves one last time we have $r_{\gamma(t)}(a) = a\gamma^v(t) = \theta^v_t(a)$. Since a was arbitrary, (iii) holds. ∎

Remark 10.7. In the following it will occasionally be useful for us also to consider *right-invariant* vector fields. To this end we denote by $\tilde{X}_v$ the vector field on G defined by

$$\tilde{X}_v(a) = Dr_a(e)[v].$$

I will leave it to you to check that everything we proved for left-invariant vector fields works in exactly the same way for right-invariant vector fields. In particular, $v \mapsto \tilde{X}_v$ is a isomorphism from $\mathfrak{g}$ to the space $\mathfrak{X}_r(G)$ of right-invariant vector fields.

In particular, if $\tilde{\gamma}^v$ is the integral curve of $\tilde{X}_v$ with initial condition e and $\tilde{\theta}^v_t$ denotes the flow of $\tilde{X}_v$ then in Proposition 10.6 we could replace condition (ii) with $\gamma(t) = \tilde{\gamma}^v(t)$ and we could replace condition (iii) with $\tilde{\theta}^v_t = l_{\gamma(t)}$.

We can now finally define the exponential map.

Definition 10.8. Let G be a Lie group with Lie algebra $\mathfrak{g}$. The **exponential map** is the map

$$\exp\colon \mathfrak{g} \to G, \qquad v \mapsto \gamma^v(1).$$

The following result is an immediate corollary of Proposition 10.4, Lemma 10.5, and Proposition 10.6.

Proposition 10.9 (Properties of the exponential map). *The exponential map* $\exp\colon \mathfrak{g} \to G$ *satisfies:*

(i) $\exp((s+t)v) = \exp(sv)\exp(tv)$ *for all* $v \in \mathfrak{g}$ *and* $s, t \in \mathbb{R}$,

(ii) $\exp(-tv) = (\exp(tv))^{-1}$ *for all* $v \in \mathfrak{g}$ *and* $t \in \mathbb{R}$,

(iii) *The map* $t \mapsto \exp(tv)$ *is precisely the one-parameter subgroup* $\gamma^v(t)$.

(iv) *The flow* θ^v_t *of* X_v *is given by* $\theta^v_t = r_{\exp(tv)}$.

The following property of the exponential map is less obvious.

Theorem 10.10. *The exponential map* $\exp\colon \mathfrak{g} \to G$ *is smooth. Moreover up to the canonical isomorphism* $T_0\mathfrak{g} \cong \mathfrak{g}$, *the derivative of the exponential map at* $0 \in \mathfrak{g}$ *is the identity.*

Proof. We prove the result in two steps.

1. Consider the map $\widehat{X}$ on the product manifold $G \times \mathfrak{g}$ given by

$$\widehat{X}(a, v) := (X_v(a), 0) \in T_aG \times T_v\mathfrak{g} \cong T_{(a,v)}(G \times \mathfrak{g}),$$

where we are using the result from Problem C.5. We claim that $\widehat{X}$ is a vector field. It clearly satisfies the section property 7.1, and thus we need only check that $\widehat{X}$ is smooth. For this, suppose $f \in C^\infty(G \times \mathfrak{g})$ is smooth. Given $v \in \mathfrak{g}$, let $f_v := f(\cdot, v)\colon G \to \mathbb{R}$ denote the smooth function given by regarding v as fixed. Then

$$\widehat{X}(f)(a, v) = X_v(f_v)(a).$$

The vector field X_v depends linearly on v (and thus also smoothly). The function f_v depends smoothly on v as f is smooth as a function of both a and v. Thus the expression $(a, v) \mapsto X_v(f_v)(a)$ depends smoothly on both a and v. Thus by

Proposition 7.2, $\widehat{X}$ is indeed a vector field. Thus its flow $\widehat{\theta}$ of $\widehat{X}$ is also smooth. By Proposition 10.6 this flow is given by

$$\widehat{\theta}(t,a,v) := \big(a\exp(tv), v\big), \qquad (t,a,v) \in \mathbb{R}\times G\times\mathfrak{g}.$$

In particular, $\widehat{\theta}(1,e,\cdot)\colon \mathfrak{g}\to G\times\mathfrak{g}$ is smooth. This is the map $v\mapsto(\exp(v),v)$. Thus exp is smooth.

2. We now compute the derivative of exp. Our claim is that if $\mathcal{J}_0\colon\mathfrak{g}\to T_0\mathfrak{g}$ is the map from Problem B.3 then the following diagram commutes:

$$\begin{tikzcd} T_0\mathfrak{g} \arrow[rr, "D\exp(0)"] & & \mathfrak{g} \\ & \mathfrak{g} \arrow[ul, "\mathcal{J}_0"] \arrow[ur, "\mathrm{id}"'] & \end{tikzcd}$$

So take $v\in\mathfrak{g}$. Then $\mathcal{J}_0(v)=\delta'(0)$ where $\delta(t)=tv$.

$$D\exp(0)[\mathcal{J}_0(v)] = (\exp\circ\delta)'(0) = \frac{d}{dt}\Big|_{t=0}\exp(tv) = \frac{d}{dt}\Big|_{t=0}\gamma^v(t) = v.$$

This completes the proof. ∎

Corollary 10.11. *The exponential map is a diffeomorphism of some neighbourhood of the origin in $\mathfrak{g}$ onto its image in G.*

Proof. Since exp has maximal rank at 0 by Theorem 10.10, this follows immediately from the Inverse Function Theorem. ∎

Now let us investigate how the exponential map behaves with respect to Lie group homomorphisms.

Proposition 10.12. *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\varphi\colon G\to H$ be a Lie group homomorphism. Then the following diagram commutes:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "D\varphi(e)"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "\varphi"'] & H \end{tikzcd}$$

Proof. If $\gamma\colon\mathbb{R}\to G$ is a homomorphism then since φ is a homomorphism so is $\varphi\circ\gamma\colon\mathbb{R}\to H$. Applying this with $\gamma(t)=\exp(tv)$ shows that $t\mapsto\varphi(\exp(tv))$ is a one-parameter subgroup of H. Since

$$(\varphi\circ\gamma)'(0) = \frac{d}{dt}\Big|_{t=0}\varphi(\exp(tv)) = D\varphi(e)\circ D\exp(0)[v] = D\varphi(e)[v],$$

where we used the chain rule and Theorem 10.10, we see by uniqueness of integral curves that $\varphi(\exp(tv))=\exp(tD\varphi(e)[v])$, which is what we wanted to prove. ∎

Applying Proposition 10.12 to an inclusion of a subgroup, as in Corollary 9.22 tells us:

Corollary 10.13. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $H \subset G$ a Lie subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then the exponential map $\exp\colon \mathfrak{h} \to H$ is the restriction of $\exp\colon \mathfrak{g} \to G$ to $\mathfrak{h}$.*

We now identify the exponential map for $G = \mathrm{GL}(n)$.

Proposition 10.14. *Let $A \in \mathfrak{gl}(n) = \mathrm{Mat}(n)$. Then the matrix exponential*

$$\exp(A) := \sum_{h=0}^{\infty} \frac{1}{h!} A^h$$

converges and defines an element of $\mathrm{GL}(n)$. Moreover $A \mapsto \exp(A)$ is the exponential map of $\mathrm{GL}(n)$.

The proof is deferred to Problem Sheet F. As with Corollary 9.24, this also allows us to characterise the matrix exponential for matrix Lie groups.

Corollary 10.15. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then the exponential map $\exp\colon \mathfrak{g} \to G$ is given by matrix exponentiation: $\exp(A) = e^A$.*

Proof. Apply Corollary 10.13 and Proposition 10.14. ∎

Remark 10.16. **Warning:** Next semester we will define another map called the "exponential map". This is defined for any *spray* $\mathbb{S}$ on any manifold M (not necessarily a Lie group G) (see Definition 43.2), and behaves similarly to the exponential map defined here. In general there is no relation between the two exponential maps (apart from sharing similar properties), and thus the terminology is a bit unfortunate. However if G is a Lie group and m is a *bi-invariant* metric (meaning that $l_a^\star m = r_a^\star m = m$ for all $a \in G$; then the exponential map corresponding to the *geodesic spray* of the Levi-Civita connection (see Lecture 45) of m agrees with the exponential map defined here.

We now look at a Lie group *acting* on manifold.

Definition 10.17. Let G be a Lie group and let M be a manifold. A smooth map $\mu\colon G \times M \to M$ satisfying

$$\mu(ab, x) = \mu(a, \mu(b, x)), \qquad \mu(e, x) = x$$

for all $a, b \in G$ and $x \in M$ is called **a left action of G on M**. For fixed $a \in G$, this implies that $x \mapsto \mu(a, x)$ is a diffeomorphism of M, which we denote by μ_a.

Remark 10.18. One can analogously define a *right* action of a Lie group on a manifold. We will come back to this in Lecture 24 when we define principal bundles.

Definition 10.19. A **linear** action of a Lie group G on a vector space V is a smooth[1] left action $\mu\colon G \times V \to V$ such that μ_a is a linear map for each $a \in G$. Thus one can think of $a \mapsto \mu_a$ as Lie group homomorphism $G \to \mathrm{GL}(V)$. In algebra it is common to call this a **representation** of G.

[1] In fact, if μ is continuous then μ is automatically smooth—this can be proved using Problem F.6.

A point $x \in M$ is called a **fixed point** if $\mu(a, x) = x$ for all $a \in G$.

PROPOSITION 10.20. *Let $\mu\colon G \times M \to M$ be a left action of a Lie group G on a manifold M. Assume that x is a fixed point of μ. Then the map*

$$\varphi\colon G \to \mathrm{GL}(T_x M), \qquad \varphi(a) := D\mu_a(x)$$

is a Lie group homomorphism (i.e. a representation).

Proof. Let us first check φ is a homomorphism. For this, observe

$$\varphi(ab) = D\mu_{ab}(x) = D(\mu_a \circ \mu_b)(x) = \varphi(a)\varphi(b).$$

The smoothness issue is a little more delicate. Suppose $\dim M = n$, so that $T_x M \cong \mathbb{R}^n$ and $\mathrm{GL}(T_x M) \cong \mathrm{GL}(n)$. Under such an identification, $\varphi(a)$ is an $n \times n$ matrix with (i,j)th entry $\varphi(a)^i_j$. We need to prove that $a \mapsto \varphi(a)^i_j$ is smooth for each (i,j). If $e_1, \dots, e_n$ is our given basis of $T_x M$ with dual basis $e^1, \dots, e^n$ of $T^*_x M$, then the (i,j)th entry of $\varphi(a)$ is $e^i(\varphi(a)[e_j])$. Thus we need to prove that $a \mapsto e^i(\varphi(a)[e_j])$ is smooth. In fact we will show something more general that does not require any choice of basis: if $v \in T_x M$ and $p \in T^*_x M$ are *any* two fixed elements then

$$a \mapsto p\big(D\mu_a(x)[v]\big)$$

is smooth. More generally still, it suffices to show that $a \mapsto D\mu_a(x)[v]$ is smooth as a map $G \to TM$. But this is simply the composition

$$G \to TG \times TM \cong T(G \times M) \to TM$$

where the first map is $a \mapsto \big((a, 0), (x, v)\big)$, the second map is the canonical identification coming[2] from Problem C.5, and the last map is $D\mu$. ■

There are two main type of Lie group actions that we are interested in the next few lectures. The first occurs when $M = G$ and the action is given by conjugation. The second occurs when the action of G on M is *transitive*. In this case M becomes a **homogeneous space** and is necessarily diffeomorphic to G/H for some Lie subgroup $H \subset G$. We will study these in Lecture 12. In Lecture 24 we will come back to Lie group actions in the context of principal bundles.

DEFINITION 10.21. A Lie group G acts smoothly on itself on the left via inner automorphisms:

$$\mu_a\colon G \to G, \qquad \mu_a(b) = aba^{-1} = l_a(r_{a^{-1}}(b)) = r_{a^{-1}}(l_a(b)).$$

The identity e is a fixed point of this action, and hence by Proposition 10.20 we obtain a Lie group homomorphism $G \to \mathrm{GL}(\mathfrak{g})$. This is called the **adjoint representation** and is denoted by

$$\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g}).$$

We usually write $\mathrm{Ad}(a) = \mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$.

[2] It follows from Problem C.5 that for any two manifolds M, N, the tangent bundles $TM \times TN$ and $T(M \times N)$ are canonically diffeomorphic. Exercise: Why?

We can then go one step further and differentiate Ad. This requires us to look at the Lie algebra of $\mathrm{GL}(\mathfrak{g})$, which we write as

$$\mathfrak{gl}(\mathfrak{g}) = \{\text{all linear maps } \mathfrak{g} \to \mathfrak{g}\}\,.$$

Definition 10.22. The derivative of the adjoint representation is denoted by

$$\mathrm{ad} \coloneqq D(\mathrm{Ad})(e) \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}).$$

We usually write $\mathrm{ad}(v) = \mathrm{ad}_v \colon \mathfrak{g} \to \mathfrak{g}$.

Then Proposition 10.12 gives us a commutative diagram:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\mathrm{ad}"] \arrow[d, "\exp"'] & \mathfrak{gl}(\mathfrak{g}) \arrow[d, "\exp"] \\ G \arrow[r, "\mathrm{Ad}"'] & \mathrm{GL}(\mathfrak{g}) \end{tikzcd}$$

The map ad has a pleasing description. The proof of the next result is deferred to Problem Sheet F.

Proposition 10.23. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then for $v, w \in \mathfrak{g}$ one has* $\mathrm{ad}_v(w) = [v, w]$.

We conclude this lecture by briefly discussing the most important infinite-dimensional Lie group. This will also explain the discrepancy from Remark 7.9 on the sign of the Lie bracket. This material is strictly non-examinable.

(♣) Remark 10.24. Let M be a compact manifold. The group $\mathrm{Diff}(M)$ can itself be given a (Fréchet) manifold structure, and thus $\mathrm{Diff}(M)$ is an infinite-dimensional Lie group. A Fréchet manifold is a weaker and less useful concept than that of a Banach manifold (cf. Remark 1.31). The difference is that a Fréchet manifold is locally modelled on a Fréchet space rather than a Banach space. The reason they are less useful is that the Inverse and Implicit Function Theorems are valid for Banach manifolds, but not for Fréchet manifolds.

Sadly though we have no choice in the matter. Even if we wanted to work with lower regularity, whilst the space of C^r-diffeomorphisms $C^r(M, M)$ does have a Banach manifold structure, it is *not* a Lie group. Indeed, although right multiplication is smooth as a map from $C^r(M, M)$ to itself, left multiplication is not even continuous! (Exercise: Why?)

In any case, if we give $\mathrm{Diff}(M)$ its Fréchet smooth structure, then one can show that

$$T_{\mathrm{id}}\mathrm{Diff}(M) = \mathfrak{X}(M),$$

(as one would expect, the tangent space to an infinite-dimensional manifold is itself infinite-dimensional).

A one-parameter group $t \mapsto \theta_t$ (in the sense of Definition 8.15) simply as a curve in the manifold $\mathrm{Diff}(M)$. Thus its velocity vector $\frac{d}{dt}\big|_{t=0}\theta_t$ should be a vector field on M, and indeed, one can check that that the formalism gives the infinitesimal generator from in (8.6).

Proposition 10.1 and Proposition 10.4 can thus be thought of as justifying Definition 8.15. The exponential map $\exp\colon \mathfrak{X}(M) \to \mathrm{Diff}(M)$ assigns to a vector field X the flow θ_t^X—this is well-defined by Corollary 8.20. The inner automorphism action $\mu_\varphi(\psi) := \varphi \circ \psi \circ \varphi^{-1}$ gives rise to the adjoint map

$$\mathrm{Ad}_\varphi\colon \mathfrak{X}(M) \to \mathfrak{X}(M)$$

which one easily sees is given by

$$\mathrm{Ad}_\varphi(X) = \varphi_\star(X).$$

But now the "bug" in the definition becomes apparent: if we differentiate this to get $\mathrm{ad}\colon \mathfrak{X}(M) \to \mathfrak{gl}(\mathfrak{X}(M))$, we find that

$$\mathrm{ad}_X(Y) = \mathcal{L}_Y X = [Y, X], \qquad \forall\, X, Y \in \mathfrak{X}(M),$$

which contradicts (!) Proposition 10.23.

Of course there is no actual contradiction, since this is all a matter of conventions. What we have learnt is that: *if we want to think of* $\mathfrak{X}(M)$ *as the Lie algebra of the infinite-dimensional Lie group* $\mathrm{Diff}(M)$ *then the Lie bracket "should" be given by* $[X, Y] = \mathcal{L}_Y X$.

This is the reason why some authors define the Lie bracket the other way round. Nevertheless, as mentioned in Remark 7.9, I have chosen the "incorrect" sign convention so as to be consistent with the vast majority of the literature.

Distributions and the Frobenius Theorem

In this lecture we introduce *distributions* and prove the *Frobenius Theorem*. This theorem will be used repeatedly throughout the course (and is the cornerstone of an area of differential geometry called *foliation theory*). We will see some immediate applications next lecture to Lie groups. We begin with the following preliminary result.

PROPOSITION 11.1. *Let M be a smooth manifold of dimension n and let $W \subset M$ be a non-empty open set. Suppose $X_1, \dots, X_k \in \mathfrak{X}(W)$ are vector fields such that*

(i) *There exists $x_0 \in W$ such that the vectors $X_i(x_0)$ are all linearly independent in $T_{x_0}M$ (and thus necessarily $k \leq n$)*

(ii) *For all i, j one has $[X_i, X_j] \equiv 0$.*

Then there exists a neighbourhood $U \subset W$ of x_0 and a chart $\sigma \colon U \to O$ with local coordinates $x^i = u^i \circ \sigma$ such that $\frac{\partial}{\partial x^i} = X_i|_U$ for all $1 \leq i \leq k$.

An immediate corollary is the following extension of Problem D.1.

COROLLARY 11.2. *Let M be a smooth manifold of dimension n and let $W \subset M$ be a non-empty open set. Let $X \in \mathfrak{X}(W)$ and suppose $X(x_0) \neq 0$ for some $x_0 \in W$. Then there exists a neighbourhood $U \subset W$ of x_0 and a chart $\sigma \colon U \to O$ such that $X|_U = \frac{\partial}{\partial x^1}$.*

Proof of Proposition 11.1. We prove the result in two steps. The first step reduces the problem to $\mathbb{R}^n$. That this is possible should be clear from the statement, since the assertion is visibly local.

1. Suppose $\sigma \colon U \to O$ is a chart on M. Let us write $D_i \in \mathfrak{X}(O)$ for the vector field given by[1]

$$D_i(y)(f) := D_i f(y) = Df(y)[e_i], \qquad y \in O, \ f \in C^\infty(O). \tag{11.1}$$

Let $x^i = u^i \circ \sigma$ denote the local coordinates of σ. Then by definition one has

$$\sigma_\star \left(\frac{\partial}{\partial x^i} \right) = D_i.$$

Since $\sigma_\star \left(\frac{\partial}{\partial x^i} \right)$ is uniquely determined by σ, it is sufficient to find a chart $\sigma \colon U \to O$ such that

$$\sigma_\star (X_i|_U) = D_i, \qquad \forall\, 1 \leq i \leq k.$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] We often denote this vector field by $\frac{\partial}{\partial x^i}$, but for obvious reasons that would be too confusing in this proof...

Now suppose $\tau \colon U \to O$ is any chart on M about x_0. Assume that $\tau(x_0) = 0 \in O$ for simplicity, and let $Y_i \in \mathfrak{X}(O)$ denote the unique vector field such that

$$\tau_\star(X_i|_U) = Y_i.$$

Since the X_i are linearly independent at x_0, the Y_i are linearly independent at 0. Thus there exists a linear isomorphism $T \colon \mathbb{R}^n \to \mathbb{R}^n$ that maps $\mathcal{J}_0^{-1}(Y_i(0))$ to the standard basis vector e_i for each $1 \le i \le k$. Then if we replace τ by $T \circ \tau$ we may assume that the vector fields Y_i satisfy

$$Y_i(0) = D_i(0), \qquad \forall\, 1 \le i \le k. \tag{11.2}$$

We emphasise this identity only holds at the point 0. We wish to find a local diffeomorphism h defined on a neighbourhood $V \subset \mathbb{R}^n$ about 0 such that $h(0) = 0$ and such that for all $1 \le i \le k$ one has

$$h_\star(Y_i) = D_i \qquad \text{on } V.$$

Then setting $\sigma = h \circ \tau$ one has

$$\sigma_\star(X_i) = h_\star \circ \tau_\star(X_i) = h_\star(Y_i) = D_i.$$

2. It thus suffices to find such an h. Note by Proposition 7.20 that the vector fields Y_i satisfy $[Y_i, Y_j] \equiv 0$. Let θ^i_t denote the flow of Y_i. For a sufficiently small neighbourhood Ω of 0 in $\mathbb{R}^n$ there is a well defined smooth function

$$f \colon \Omega \to \mathbb{R}^n, \qquad f(y^1, \dots, y^n) = \big(\theta^1_{y^1} \circ \cdots \circ \theta^k_{y^k}\big)(0, \dots, 0, y^{k+1}, \dots, y^n).$$

Let $g \in C^\infty(\Omega)$. We compute

$$\begin{aligned}
Df(y)[D_1(y)](g) &= D_1(y)(g \circ f) \\
&= \lim_{t \to 0} \frac{\big(g \circ \theta^1_{y^1+t} \circ \cdots \circ \theta^k_{y^k}\big)(0, \dots, 0, y^{k+1}, \dots, y^n) - (g \circ f)(y)}{t} \\
&= \lim_{t \to 0} \frac{g \circ \theta^1_t(f(y)) - g(f(y))}{t} \\
&= Y_1(f(y))(g).
\end{aligned}$$

Thus $f_\star(D_1) = Y_1$. Since the Lie brackets vanish, using induction and Proposition 8.27 we have for any $1 \le i \le k$ that

$$\theta^1_{y^1} \circ \cdots \circ \theta^i_{y^i} \circ \cdots \circ \theta^k_{y^k} = \theta^i_{y^i} \circ \cdots \circ \theta^1_{y^1} \circ \cdots \circ \theta^k_{y^k},$$

and thus exactly the same argument shows that $f_\star(D_i) = Y_i$ for all $1 \le i \le k$. Moreover for $k < i \le n$ if we take $y = 0$ we have

$$\begin{aligned}
Df(0)[D_i(0)](g) &= D_i(0)(g \circ f) \\
&= \lim_{t \to 0} \frac{(g \circ f)(0, \dots, 0, t, 0 \dots, 0) - (g \circ f)(0)}{t} \\
&= \lim_{t \to 0} \frac{g(0, \dots, 0, t, 0 \dots, 0) - g(0)}{t} \\
&= D_i(0)(g).
\end{aligned}$$

Since $Y_i(0) = D_i(0)$ by (11.2), we have $Df(0)[D_i(0)] = D_i(0)$ for all $1 \le i \le n$, which tells us that $Df(0)$ is the identity. Thus by the Inverse Function Theorem there exists $V \subset \Omega$ containing 0 such that $f|_V$ is a diffeomorphism. Set $h := f|_V^{-1}$. Then $h_\star(Y_i) = D_i$ and the proof is complete. ■

We now introduce the notion of a distribution.

Definition 11.3. Let M be a smooth manifold of dimension n, and let $k \le n$. A **distribution** Δ **on** M **of dimension** k is a choice of k-dimensional linear subspace $\Delta_x \subset T_xM$ for each $x \in M$ that varies smoothly with x in the following sense: For each point $x_0 \in M$ there exists a neighbourhood U of x_0 and k vector fields $X_1, \dots, X_k \in \mathfrak{X}(U)$ such that

$$\Delta_x = \operatorname{span}_{\mathbb{R}}\{X_1(x), \dots, X_k(x)\}, \qquad \forall\, x \in U.$$

The simplest example is $k = 1$.

Example 11.4. Suppose X is a **nowhere-vanishing** vector field on M (this means that $X(x) \neq 0$ for all $x \in M$). Then X defines a one-dimensional distribution by setting $\Delta_x := \operatorname{span}_{\mathbb{R}}\{X(x)\}$ for each $x \in M$.

Remark 11.5. Not every manifold admits such a vector field. Indeed, if n is even then every vector field on S^n vanishes in at least one point. This is the so-called "Hairy Ball Theorem". See Problem L.8 for a proof of this fact.

In fact, the Hairy Ball Theorem is a purely topological result, and thus the smoothness assumption is not necessary: if n is even then any continuous map $S^n \to TS^n$ satisfying the section property 7.1 must vanish somewhere. This can be proved by applying the Whitney Approximation Theorem 6.14 to the smooth case, but it is also easy to show using some basic algebraic topology. I proved it about half-way through Algebraic Topology I last year here.

Definition 11.6. Let Δ be a k-dimensional distribution on M, and suppose $L \subset M$ is a k-dimensional immersed submanifold. We say that L is an **integral manifold** of Δ if

$$D\imath(x)[T_xL] = \Delta_x, \qquad \forall\, x \in L,$$

where $\imath\colon L \hookrightarrow M$ is the inclusion.

Example 11.7. Let Δ be a one-dimensional distribution on M. We claim through any point $x \in M$ there exist (many) integral manifolds L of Δ. Indeed, by definition of a distribution, given any $x_0 \in M$ there exists a neighbourhood W of x_0 and a vector field $X \in \mathfrak{X}(W)$ such that $\Delta_x = \operatorname{span}_{\mathbb{R}}\{X(x)\}$ for all $x \in W$. Since in particular $X(x_0) \neq 0$, by Corollary 11.2 there exists a neighbourhood $U \subset W$ of x_0 and a chart $\sigma\colon U \to O$ such that $\sigma(x_0) = 0$ and such that if x^i are the local coordinates of σ then $X|_U = \frac{\partial}{\partial x^1}$. Then the set

$$L := \big\{x \in U \mid x^2(x) = \dots = x^n(x) = 0\big\}$$

is an embedded one-dimensional submanifold of M by Proposition 5.10. Moreover the proof of Proposition 5.10 shows that $D\imath(x)[T_xL] = \frac{\partial}{\partial x^1}\big|_x$ for all $x \in L$, where $\imath\colon L \hookrightarrow M$ is the inclusion. Thus L is an integral manifold of Δ containing x_0. In fact, any connected integral manifold of Δ which is contained in U is of this form.

For $k > 1$ however this is not necessarily the case—there exist distributions that not admit integral manifolds through every point. Here is the standard example of such a distribution.

Example 11.8. Consider the distribution on $\mathbb{R}^3$ spanned by the vector field

$$X := \frac{\partial}{\partial x^1} + x^2 \frac{\partial}{\partial x^3}, \qquad Y := \frac{\partial}{\partial x^2}.$$

See Figure 11.1. Suppose an integral manifold L existed through the origin. Then (as the picture indicates), one would have $T_{(0,0,0)}L$ equal to the (x^1, x^2)-plane. But now suppose $\gamma \colon S^1 \to L$ is a closed curve in L that circles round the x^3-axis. Since γ is tangent to Δ, one readily sees that the x^3-component of γ is an increasing function. But then γ endlessly spirals upwards, and hence cannot close up—contradiction.

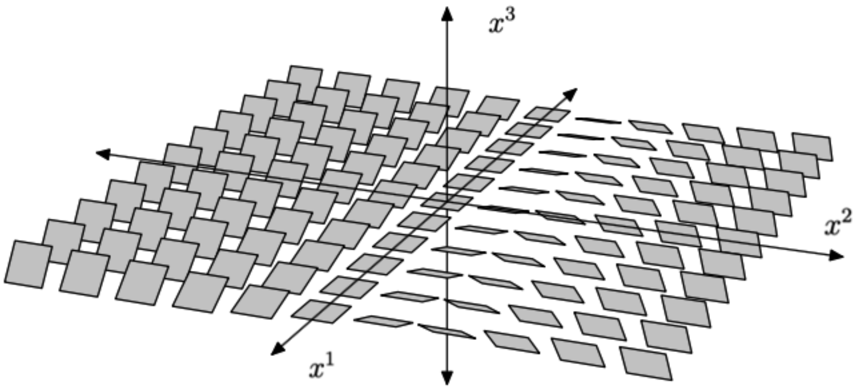

Figure 11.1: The standard contact distribution on $\mathbb{R}^3$. (Taken from Wikipedia.)

(♣) Remark 11.9. Example 11.8 is the starting point for the field of geometry called **contact geometry**. In general a **contact distribution** on a manifold is a distribution which is "maximally" non-integrable. Such a manifold is necessarily odd-dimensional. Contact manifolds are the odd-dimensional cousins of **symplectic manifolds**. We won't study either contact or symplectic manifolds in this course. If you are interested (and you should be interested!) then I recommend attending Jagna Wiśniewska's course next semester.

We now formulate a condition that implies integral manifolds always exist.

Definition 11.10. Let Δ be a distribution on M and let X be a vector field on M. We say that X **belongs to** Δ if $X(x) \in \Delta_x$ for each $x \in M$.

Definition 11.11. A distribution Δ is said to be **integrable** if $[X, Y]$ belongs to Δ whenever X and Y belong to Δ. Thus an integrable distribution is one for which the space of vector fields belonging to it forms a Lie subalgebra of $\mathfrak{X}(M)$.

Remark 11.12. Some authors use the word *involutive* instead of integrable to describe a distribution satisfying the conditions of Definition 11.11.

The following condition is occasionally useful for checking integrability.

LEMMA 11.13. *Let Δ be a distribution on M of dimension k. If for every $x \in M$ there exists a neighbourhood U of x and k vector fields $X_1, \dots, X_k \in \mathfrak{X}(U)$ such that Δ is spanned by the X_i over U and such that $[X_i, X_j](y) \in \Delta_y$ for all $y \in U$ and for all $1 \le i, j \le k$ then Δ is integrable.*

Proof. Let $x \in M$ and let X and Y be vector fields that belong to Δ. Choose a neighbourhood U of x for which there exist vector fields spanning Δ as in the hypotheses of the Lemma. Then on U we can write

$$X|_U = a^i X_i, \qquad Y|_U = b^i X_i$$

for some smooth[2] functions $a^i, b^i \colon U \to \mathbb{R}$. By Problem D.4 one has on U that

$$[X, Y]|_U = [a^i X_i, b^j X_j] = a^i b^j [X_i, X_j] + a^i X_i(b^j) X_j - b^j X_j(a^i) X_i.$$

Since $[X_i, X_j](y) \in \Delta_y$ for all $y \in U$, this shows that $[X, Y]$ belongs to Δ for every point in U. Since x was arbitrary, if follows that $[X, Y]$ belongs to Δ. ■

Every one-dimensional distribution is integrable. (Exercise: Why?) More generally, we have:

PROPOSITION 11.14. *Let Δ be a distribution on M. Assume that for every $x \in M$ there exists an integral manifold L_x of Δ with $x \in L_x$. Then Δ is integrable.*

Proof. Let X and Y belong to Δ. Fix an arbitrary point $x \in M$, and let $\imath \colon L_x \hookrightarrow M$ denote the inclusion. In the language of Problem D.6, X and Y are tangent to L_x. By part (iii) of Problem D.6, $[X, Y]$ is also tangent to L_x, or equivalently, $[X, Y](x) \in D\imath_x(x)[T_x L_x] = \Delta_x$. Since x was arbitrary, we conclude $[X, Y]$ belongs to Δ. ■

A more difficult result states that the converse holds.

REMARK 11.15. In this lecture it will be convenient to be slightly more flexible in the definition of a slice chart (cf. Definition 5.9). Let $\mathbb{I}^r := (-1, 1)^r$ denote the r-dimensional open unit cube, and write an element of $\mathbb{I}^r$ as a tuple $c = (c^1, \dots, c^r)$. Suppose M^n is a manifold and $\sigma \colon U \to \mathbb{I}^n$ is a chart with corresponding local coordinates x^i. A **slice** in M of dimension k is an embedded submanifold of the form

$$L(c) := \left\{ x \in U \mid x^{k+1}(x) = c^1, \dots, x^n(x) = c^{n-k} \right\}, \qquad c = (c^1, \dots, c^{n-k}) \in \mathbb{I}^{n-k}.$$

Thus the difference (compared to Definition 5.9 and Proposition 5.10) is that instead of requiring the last $n - k$ coordinates to all be zero, instead we are requiring them to be some fixed element in $\mathbb{I}^{n-k}$. Of course, this makes no difference in the grand scheme of things.

THEOREM 11.16 (The Local Frobenius Theorem). *Let M be a smooth manifold of dimension n, and let Δ be an integrable k-dimensional distribution on M. Then*

[2] If you are worried why these functions are smooth, see Remark 16.9.

for every $x \in M$ there exists a chart $\sigma\colon U \to \mathbb{I}^n$ with $\sigma(x_0) = 0$ and such that for any $c \in \mathbb{I}^{n-k}$, the slice

$$L(c) := \left\{x \in U \mid x^{k+1}(x) = c^1, \dots, x^n(x) = c^{n-k}\right\}$$

is an integral manifold of Δ. Moreover every connected integral manifold of Δ contained in U is of this form.

Proof. Once again, the statement is purely local, so by arguing as in Step 1 of Proposition 11.1, we may assume that $M = \mathbb{R}^n$, $x_0 = 0$, and Δ_0 is spanned by the vector fields $D_i(0)$ for $i = 1, \dots, k$. (Exercise: Fill in these details!)

As in the proof of the Implicit Function Theorem 5.13, Let π_1 and π_2 denote the two projections $\mathbb{R}^n \to \mathbb{R}^k$ and $\mathbb{R}^{n-k}$ respectively:

$$\pi_1(x^1, \dots, x^n) = (x^1, \dots, x^k), \qquad \pi_2(x^1, \dots, x^n) = (x^{k+1}, \dots, x^n).$$

Let $P_x := D\pi_1(x)|_{\Delta_x}\colon \Delta_x \to T_x\mathbb{R}^k$. Then $x \mapsto P_x$ is a smooth family of linear maps (whose domain also ranges smoothly with x). By assumption P_0 is an isomorphism. Since being invertible is an open condition[3], it follows that there is a neighbourhood W of 0 in $\mathbb{R}^n$ such that P_x is an isomorphism for all $x \in W$. Thus there are unique vector fields $X_i \in \mathfrak{X}(W)$ belonging to Δ that are π_1-related to D_i for $i = 1, \dots, k$. By part (ii) of Problem D.5 one has that $[X_i, X_j]$ is π_1-related to $[D_i, D_j]$. By Proposition 7.10, $[D_i, D_j] = 0$, and thus $[X_i, X_j]$ is π_1-related to the zero vector field. Now since Δ is integrable, $[X_i, X_j]$ belongs to Δ, and since $D\pi_1(x)|_{\Delta_x} = P_x$ is injective for $x \in W$, it follows that $[X_i, X_j] = 0$. (This is the only place where we use integrability of Δ!)

Thus by Proposition 11.1 there is a chart $\sigma\colon U \to \mathbb{I}^n$ defined on $U \subset W$ such that $X_i|_U = \frac{\partial}{\partial x^i}$. Now let $\varphi := \pi_2 \circ \sigma\colon U \to \mathbb{I}^{n-k}$. Then φ is a smooth surjective submersion, and thus by the Implicit Function Theorem 5.13, for any $c \in \mathbb{I}^{n-k}$, the set $L(c) := \varphi^{-1}(c)$ is an embedded submanifold of M, and any $x \in U$ belongs to a unique $L(c)$ (namely, $c = \varphi(x)$). Moreover by Proposition 5.15, if we denote by $\imath\colon L(c) \hookrightarrow U$ the inclusion then for any $x \in L(c)$ one has

$$\begin{aligned}
D\imath(x)[T_xL(c)] &= \ker D\varphi(x) \\
&= \left\{v \in T_xU \mid v(x^i) = 0,\ \forall\, i = k+1, \dots, n\right\} \\
&= \operatorname{span}_{\mathbb{R}} \left\{ \frac{\partial}{\partial x^i}\bigg|_x \mid 1 \le i \le k \right\} \\
&= \Delta_x.
\end{aligned}$$

Finally, if L is any integral manifold of Δ contained in U then for any $x \in L$ and $v \in T_xL$, one has $D\imath(x)[v](x^i) = 0$ for $i = k+1, \dots, n$. Thus $D(x^i \circ \imath)(x)$ is the zero map for each $i = k+1, \dots, n$, and hence $x \mapsto x^i(\imath(x))$ is locally constant. If L is connected, then it is constant, and thus $L = L(c)$ is a slice as described above. This completes the proof. ∎

We now globalise Theorem 11.16. Note that so far all our integral manifolds have actually been embedded, despite the fact that the definition only required them to

[3] This is a very special case of the Inverse Function Theorem 5.1 for linear maps—alternatively one could simply use that $x \mapsto \det P_x$ is a continuous function.

be immersed. The reason for this is that "glue" together the integral manifolds (this will be made precise below) we typically lose the embedded property. Let us formalise this with a definition.

Definition 11.17. Let M be a smooth manifold of dimension n. A **k-dimensional foliation** $\mathcal{F}$ of M is a partition of M into k-dimensional connected immersed submanifolds, called the **leaves** of the foliation, such that:

(i) The collection of tangent spaces to the leaves defines a distribution Δ on M.

(ii) Any connected integral manifold of Δ is contained in some leaf of $\mathcal{F}$.

Each leaf L of $\mathcal{F}$ is called a **maximal integral manifold** of Δ. One says that the distribution Δ is **induced by** $\mathcal{F}$.

Here is the global version of Theorem 11.16.

Theorem 11.18 (The Global Frobenius Theorem). *Let Δ be an integrable distribution on M. Then Δ is induced by a foliation.*

The proof is rather tricky, and we will be a little sketchy. This is non-examinable.

(♣) *Proof.* Let Δ be an integrable distribution. By Theorem 11.16 for any point $x_0 \in M$ there is a chart $\sigma \colon U \to \mathbb{I}^n$ such that the slices

$$L(c) := \left\{ x \in U \mid x^{k+1}(x) = c^1, \dots, x^n(x) = c^{n-k} \right\} \tag{11.3}$$

for $c \in \mathbb{I}^{n-k}$ are integral manifolds of Δ. Since M is Lindelöf (cf. part (ii) of Remark 1.9), there is a countable set $\{x_h\}_{h=1}^{\infty}$ of points such that the corresponding charts $\sigma_h \colon U_h \to \mathbb{I}^n$ about x_h cover M. Now let $\mathcal{L}$ denote the collection of all slices $L(c)$ of the form (11.3) for the charts σ_h. Define an equivalence relation on $\mathcal{L}$ by declaring that $L \sim L'$ if there exists a finite sequence $L = L_0, L_1, \dots, L_r = L'$ such that $L_i \cap L_{i+1} \neq \emptyset$ for $i = 0, \dots, r-1$. Each equivalence class can only contain countably many slices $L \in \mathcal{L}$, since if $L \subset U_h$ is a slice then L can intersect another U_j in at most countably many components (since U_j is connected and L has at most countably many components). Now consider the union of all slices in a given equivalence class. This is a connected immersed integral manifold of Δ. Moreover any two such unions are either equal or disjoint, and by definition any connected integral manifold of Δ is contained in such a union. ■

Although the leaves of a foliation are typically not embedded submanifolds, they do retain some properties of embedded submanifolds. Here is one:

Definition 11.19. Let $L \subset M$ be an immersed submanifold. We say that L is **weakly embedded** if for every smooth manifold N and every smooth map $\varphi \colon N \to M$ such that $\varphi(N) \subset L$, the map φ is also smooth as a map $N \to L$.

Embedded submanifolds are automatically weakly embedded (we used this in the proof of Proposition 9.10).

Proposition 11.20. *Let Δ be an integrable distribution on a smooth manifold M. Every integral manifold L of Δ is a weakly embedded submanifold of M.*

Proof. Assume that $\varphi\colon N \to M$ is a smooth map such that $\varphi(N) \subset L$. Let $y_0 \in N$ and let $x_0 \coloneqq \varphi(y_0) \in L$. Choose a chart $\sigma\colon U \to \mathbb{I}^n$ with $\sigma(x_0) = 0$ such that all connected integral submanifolds of Δ contained in U are of the form

$$L(c) = \left\{x \in U \mid x^{k+1}(x) = c^1, \ldots, x^n(x) = c^{n-k}\right\}$$

for $c \in \mathbb{I}^{n-k}$. Thus there are countably many $c_h \in \mathbb{I}^{n-k}$ such that $L \cap U$ is contained in the union of the slices $L(c_h)$. Choose a connected neighbourhood $V \subset N$ of y_0 and a chart τ defined on V. Let $f \coloneqq \sigma \circ \varphi \circ \tau^{-1}$. Then if $f^i = u^i \circ f$ the last $(n-k)$ functions $(f^{k+1}, \ldots, f^n)$ can only take values in the countable set $\{c_h\}_{h=1}^{\infty}$, and thus they are locally constant. Since V is connected, it follows they are actually constant. Thus $\varphi(V)$ is contained in a single slice $L(c_{h_0})$. Since $L \cap L(c_{h_0})$ is an open subset of L that is embedded in M, it follows that $\varphi|_V$ is a smooth map from V to $L \cap L(c_{h_0})$, and thus also the composition $\varphi|_V\colon V \to L \cap L(c_{h_0}) \hookrightarrow L$. Since y_0 was an arbitrary point of N the claim follows. ∎

Homogeneous spaces

In this lecture we study *homogeneous spaces*, which are manifolds that admit a *transitive* Lie group action of diffeomorphisms. We begin however by proving Theorem 9.25 from Lecture 9. This is our first application of the Frobenius Theorem.

THEOREM 12.1. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$ then there is a unique connected Lie subgroup H of G whose Lie algebra is $\mathfrak{h}$.*

Proof. Given $a \in G$, let Δ_a denote the subspace of T_aG given by the set of all $X_v(a)$, where $X_v \in \mathfrak{X}_l(G)$ is a left-invariant vector field such that $v = X_v(e) \in \mathfrak{h} \subset \mathfrak{g}$. Thus

$$\Delta_a := \{Dl_a(e)[v] \mid v \in \mathfrak{h}\}.$$

To see that Δ really is a distribution, note that if $\{v_i\}$ is a basis of $\mathfrak{h}$ then the left-invariant vector fields $\{X_{v_i}(a)\}$ span Δ_a at every point $a \in G$. Moreover since $\mathfrak{h}$ is a Lie subalgebra, $[v_i, v_j] \in \mathfrak{h}$ for each i, j and thus $[X_{v_i}, X_{v_j}] = X_{[v_i,v_j]}$ belongs to Δ for every i, j. Thus by Lemma 11.13 it follows that Δ is integrable. Let H be the leaf of the foliation induced by Δ containing e. For any $b \in G$ be have $Dl_b(a)[\Delta_a] = \Delta_{ba}$ by construction, and hence Dl_b leaves the distribution invariant. Thus l_b permutes the leaves of the foliation, i.e. it maps the leaf passing through a diffeomorphically onto the leaf passing through ba. In particular, if $b \in H$ then $l_{b^{-1}}$ maps H to the leaf containing e, which is just H again. Thus $l_{b^{-1}}(H) = H$, which proves that H is a subgroup. It remains to prove that the multiplication map $m \colon H \times H \to H$ is smooth. We know that the multiplication $m \colon H \times H \to G$ is smooth and $m(H \times H) \subset H$. Thus by Proposition 11.20 from the last lecture, m is also smooth as a map $H \times H \to H$. This complete the proof. ■

COROLLARY 12.2. *Every Lie subgroup H of a Lie group G is weakly embedded.*

Proof. Immediate from the proof of Theorem 12.1 and Proposition 11.20. ■

COROLLARY 12.3. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let $H \subset G$ be a Lie subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Let $v \in \mathfrak{g}$. Then*

$$v \in \mathfrak{h} \qquad \Leftrightarrow \qquad \exp(tv) \in H, \ \forall\, t \in \mathbb{R}.$$

Proof. If $v \in \mathfrak{h}$ then $\exp(tv) \in H$ for all $t \in \mathbb{R}$ by Corollary 10.13. Conversely if $\gamma(t) := \exp(tv) \in H$ for all $t \in \mathbb{R}$, then $\gamma \colon \mathbb{R} \to G$ is a smooth map with image in H. Since H is weakly embedded by Corollary 12.2, γ is smooth as a map from $\mathbb{R}$ into H. Thus γ is a 1-parameter subgroup of H, and so in particular $v = \gamma'(0)$ belongs to $\mathfrak{h}$ by Proposition 10.4. ■

Note that the foliation of G defined in the proof of Theorem 12.1 has leaves the left cosets of H in G, i.e. the leaf through a is the coset $aH = l_a(H)$. We will exploit this fact again in the proof of the next theorem.

Theorem 12.4. *Let G be a Lie group and let $H \subset G$ be a closed subgroup. Let G/H denote the set of left cosets of H. Then G/H is a topological manifold with the quotient topology of dimension $\dim(G/H) = \dim G - \dim H$. Moreover there exists a smooth structure on G/H such that the projection map $\pi \colon G \to G/H$ is a smooth submersion.*

Note that $\dim H$ makes sense, for H is automatically a Lie group in its own right (and thus in particular a manifold), thanks to the Closed Subgroup Theorem 9.11. Note also that Theorem 12.4 is *not* asserting that G/H is a Lie group. Indeed, G/H is not even a group! If H is a *normal* subgroup however then G/H is a group, and we will see in Proposition 12.8 that in this case G/H is also a Lie group. The proof of Theorem 12.4 is non-examinable, since it is a bit technical.

(♣) *Proof.* We prove the result in four steps.

1. Let us first show that G/H is a Hausdorff paracompact space with at most countably many components. These are all standard point-set topological arguments; we will cover the Hausdorff one in detail and leave the others as exercises.

Observe that the quotient map π is an open map for the quotient topology, as if $U \subset G$ is open then

$$\pi^{-1}(\pi(U)) = \bigcup_{b \in H} r_b(U)$$

is open in G. To see that G/H is Hausdorff, first consider the set $C \subset G \times G$ of all pairs (a, b) with the property that there exists $a_1 \in H$ such that $a = a_1 b$. Then C is closed in $G \times G$ as H is closed in G (it is the preimage of H under the continuous map $(a, b) \mapsto b^{-1}a$). Now suppose $\pi(a) \neq \pi(b)$ as elements of G/H. This means that $(a, b) \notin C$. Since C is closed, there exist neighbourhoods U and V of a and b respectively such that $(U \times V) \cap C = \emptyset$. Then $\pi(U)$ and $\pi(V)$ are open neighbourhoods (since π is an open map) of $\pi(a)$ and $\pi(b)$ in G/H such that $\pi(U) \cap \pi(V) = \emptyset$. Thus G/H is Hausdorff[1].

Similarly the fact that G/H is paracompact is again pure point-set topology, and I will omit the details[2]. Finally the claim that G/H has at most countably many components is clear, since the connected components of G/H are the images of the connected components of G under H (this is true for all quotient spaces).

2. We now start the construction of a smooth atlas on G/H. In this step we find a set $V \subset G$ containing e such that $\pi \colon V \to \pi(V) \subset G/H$ is bijective. For this, let π_1 and π_2 denote (as usual) the two projections $\mathbb{R}^n \to \mathbb{R}^k$ and $\mathbb{R}^n \to \mathbb{R}^{n-k}$ respectively:

$$\pi_1(x^1, \dots, x^n) = (x^1, \dots, x^k), \qquad \pi_2(x^1, \dots, x^n) = (x^{k+1}, \dots, x^n).$$

As in the proof of Theorem 12.1, we apply the Frobenius Theorem to the distribution Δ given by $\Delta_a \coloneqq Dl_a(e)[T_e H]$. As mentioned before the statement of Theorem

[1]This is a special case of the following more general point-set topological fact: if X is any Hausdorff space and $\sim$ is an equivalence relation on $X \times X$ such that $\pi \colon X \to X/\sim$ is an open map then $\sim$ is closed in $X \times X$ if and only if $X/\sim$ is Hausdorff.

[2]A stronger result, which is much harder to prove and is due to Antonyan, states that if X is paracompact Hausdorff topological group and $Y \subset X$ a locally compact subgroup then the orbit space X/Y is paracompact (which in particular applies to our situation).

12.4, the corresponding foliation of G has leaves the left cosets of the connected component of H containing e in G. Theorem 11.16 gives us a chart $\sigma \colon U \to \mathbb{I}^n$ about e, where $\sigma(e) = 0$, such that each slice

$$\left\{a \in U \mid x^{k+1}(a) = c^1, \dots x^n(a) = c^{n-k}\right\}$$

for $c = (c^1, \dots, c^{n-k}) \in \mathbb{I}^{n-k}$ is contained in a left coset of H (here $x^i = u^i \circ \sigma$ are the local coordinates of σ as usual). Let L denote the slice containing e itself (i.e. where $c^i = 0$ for each i). Note L is connected by Theorem 11.16.

Since H is an embedded submanifold, there exists a connected neighbourhood $U_1 \subset U$ of e such that $U_1 \cap L = U_1 \cap H$. Choose a smaller neighbourhood $U_2 \subset U_1$ of e such that $U_2 = U_2^{-1}$ and that $U_2 \cdot U_2 \subset U_1$ (i.e. $a \in U_2$ if and only if $a^{-1} \in U_2$, and $a, b \in U_2$ implies $ab \in U_1$). Now let

$$V \coloneqq \left\{a \in U_2 \mid \pi_1(\sigma(a)) = 0 \in \mathbb{R}^k\right\}. \tag{12.1}$$

We claim that π is injective on V. Indeed, if $a, b \in V$ and $\pi(a) = \pi(b)$ then $a^{-1}b \in U_1 \cap H = U_1 \cap L$, and thus $b \in l_a(U_1 \cap L)$. Now $l_a(U_1 \cap L)$ is a connected integral manifold of Δ which lies in U, and hence by Theorem 11.16 it lies in a single slice. Since this set also contains a, we have that a and b lie in the same slice. Thus $\sigma(a) = \sigma(b)$. Since σ is a diffeomorphism (and thus in particular bijective), we have $a = b$.

3. We are now ready to construct our smooth atlas, and thus prove that G/H is a smooth manifold. With V as in (12.1), the map $\pi|_V$ is open and bijective, and hence is a homeomorphism. Set $W \coloneqq \pi(V)$ and define

$$\tau \coloneqq \pi_2 \circ \sigma \circ \pi|_V^{-1} \colon W \to \tau(W) \subset \mathbb{R}^{n-k}.$$

Then τ is also a homeomorphism, and it is our desired chart on G/H around $\pi(e)$. We now use left translations to produce charts around any other point $\pi(a)$: let $\tilde{l}_a \colon G/H \to G/H$ denote the homeomorphism

$$\tilde{l}_a(\pi(b)) \coloneqq \pi(ab).$$

Then $\tau_a \coloneqq \tau \circ \tilde{l}_{a^{-1}} \colon \tilde{l}_a(W) \to \tau(W)$ is a homeomorphism about $\pi(a)$. The transition function $\tau_a \circ \tau_b^{-1}$ is given by

$$\tau \circ \tilde{l}_{a^{-1}} \circ \left(\tau \circ \tilde{l}_{b^{-1}}\right)^{-1} = \pi_2 \circ \sigma|_V \circ l_{a^{-1}b} \circ (\pi_2 \circ \sigma|_V)^{-1},$$

which is the composition of smooth maps and hence is smooth.

4. It remains to show that π is a smooth map. Keeping with the notation from above, observe that if $a \in G$ then $\sigma \circ l_{a^{-1}}$ is a chart of G about a and $\tau \circ \tilde{l}_{a^{-1}}$ is a chart on G/H about $\pi(a)$, and

$$\left(\tau \circ \tilde{l}_{a^{-1}}\right) \circ \pi \circ \left(\sigma \circ l_{a^{-1}}\right)^{-1} = \tau \circ \pi \circ \sigma^{-1} = \pi_2,$$

which is smooth. This completes the proof. ∎

In fact, the proof given showed slightly more.

Corollary 12.5. *Let G be a Lie group and let $H \subset G$ be a closed subgroup. Endow G/H with the smooth structure constructed in Theorem 12.4. Then for any $a \in G$ there exists a neighbourhood W_a of $\pi(a)$ and a smooth map $\psi\colon W_a \to G$ such that $\pi \circ \psi = \mathrm{id}$.*

(♣) *Proof.* Using the notation from the proof of Theorem 12.4, set $W_a = \tilde{l}_a(W)$ and set $\psi := \pi|_V^{-1} \circ \tilde{l}_{a^{-1}}$. ■

We call the map ψ a **local smooth section** of π.

Corollary 12.6. *Let G be a Lie group and let $H \subset G$ be a closed subgroup. The topological manifold G/H of left cosets has a unique smooth structure for which $\pi\colon G \to G/H$ is a smooth map that admits local smooth sections.*

Proof. Suppose $(G/H)'$ is the same topological manifold, but endowed with a different smooth atlas for which π is smooth and admits local smooth structures. We claim that $\mathrm{id}\colon G/H \to (G/H)'$ is a diffeomorphism:

$$\begin{tikzcd} & G \arrow[dl, "\pi"'] \arrow[dr, "\pi"] & \\ G/H \arrow[rr, "\mathrm{id}"'] & & (G/H)' \end{tikzcd}$$

Indeed, locally the identity map and its inverse can be expressed as the composition of local smooth sections into G followed by π. Thus the smooth atlases on G/H and $(G/H)'$ both define the same smooth structure. ■

This allows us to make the following definition.

Definition 12.7. A **homogeneous space** is a smooth manifold M which is diffeomorphic to a smooth manifold of the form G/H, where G is a Lie group, H is a closed subgroup, and G/H is given the smooth structure from Theorem 12.4.

Proposition 12.8. *Let G be a Lie group and let H be a closed normal subgroup. Then the homogeneous space G/H with its natural group structure is a Lie group.*

Proof. The multiplication on G/H is given by

$$m(\pi(a), \pi(b)) := \pi(ab).$$

To check this is smooth, fix $a, b \in G$ and let ψ_a and ψ_b be smooth local sections of G/H near a and b respectively. Then near the point $(\pi(a), \pi(b)) \in G/H \times G/H$, one has

$$m = \pi \circ m_G \circ (\psi_a, \psi_b),$$

where $m_G\colon G \times G \to G$ is the multiplication on G. Thus near $(\pi(a), \pi(b))$, m is the composition of smooth maps, and hence is smooth. Since $\pi(a)$ and $\pi(b)$ were arbitrary, m is smooth everywhere. A similar argument works for the inversion map. ■

Many manifolds are homogeneous spaces (we will shortly see some examples). The key tool used to prove a given manifold is a homogeneous space is Theorem 12.11 below, which needs a few preliminary definitions.

DEFINITION 12.9. Let $\mu \colon G \times M \to M$ be a left action of a Lie group G on a smooth manifold M, as in the sense of Definition 10.17. The action μ is said to be **transitive** if for any two points $x, y \in M$ there exists $a \in G$ such that $\mu_a(x) = y$.

Suppose now $\mu \colon G \times M \to M$ is a left action. Fix a point $x \in M$, and set

$$H := \{a \in G \mid \mu_a(x) = x\}.$$

Then H is a closed subgroup of G, and the action of G restricted to H gives an action of H on M for which x is a fixed point. We call H the **isotropy group** at x.

DEFINITION 12.10. Since x is a fixed point for the action of H, by Proposition 10.20 there is a Lie group homomorphism (a representation) $H \to \mathrm{GL}(T_xM)$. We call the image of H inside $\mathrm{GL}(T_xM)$ the **linear isotropy group** at x. We will come back to this in Lecture 25.

THEOREM 12.11. *Let $\mu \colon G \times M \to M$ be a transitive left action of a Lie group G on a smooth manifold M. Fix $x \in M$ and let H denote the isotropy group at x. Let $\pi \colon G \to G/H$ denote the quotient map, and endow G/H with the smooth structure from Theorem 12.4. Define*

$$\varphi \colon G/H \to M, \qquad \varphi(\pi(a)) := \mu_a(x).$$

Then φ is a diffeomorphism, and hence M is a homogeneous space.

Proof. First observe that φ is well defined, since if $b \in H$ then $\mu_{ab}(x) = \mu_a(\mu_b(x)) = \mu_a(x)$. Next note that φ is surjective as μ is a transitive action. Moreover φ is injective since if $\varphi(\pi(a)) = \varphi(\pi(b))$ then $\mu_{a^{-1}b}(x) = x$, whence $a^{-1}b \in H$ and thus $\pi(a) = \pi(b)$.

It remains to show that φ is a diffeomorphism. By the Inverse Function Theorem 5.2, it suffices to show that φ is smooth and that $D\varphi$ has maximal rank at every point of G/H. To show that φ is smooth in a neighbourhood of a point $\pi(a)$, it suffices to show that $\varphi \circ \pi$ is smooth near a. Indeed, if $\varphi \circ \pi$ is smooth at a and $\psi \colon G/H \to G$ is a smooth local section of π at $\pi(a)$ then $\varphi = (\varphi \circ \pi) \circ \psi$ is the composition of smooth maps. Now observe that $\varphi \circ \pi = \mu \circ \imath_x$, where $\imath_x \colon G \to G \times M$ is the smooth map $a \mapsto (a, x)$, and thus $\varphi \circ \pi$ is the composition of smooth maps and hence is smooth.

Since π is a submersion, by Proposition 5.15 we have $\ker D\pi(a) = T_a(l_a(H))$, where we are suppressing the inclusion map $l_a(H) \hookrightarrow G$, and thus to prove that φ has maximal rank at $\pi(a)$ it suffices to show that

$$\ker D(\varphi \circ \pi)(a) = \ker D\pi(a).$$

Since

$$\varphi \circ \pi = \mu_a \circ (\varphi \circ \pi) \circ l_{a^{-1}} \tag{12.2}$$

for any $a \in G$, and μ_a and $l_{a^{-1}}$ are diffeomorphisms, it suffices to work at e. Let $\mathfrak{g}$ denote the Lie algebra of G and $\mathfrak{h}$ denote the Lie algebra of H. We need to prove that

$$\ker D(\varphi \circ \pi)(e) = \ker D\pi(e) = \mathfrak{h}. \tag{12.3}$$

It is obvious that $\mathfrak{h} \subset \ker D(\varphi \circ \pi)(e)$, so we must verify the opposite inclusion. Let $v \in \mathfrak{g}$, and suppose

$$v \in \ker D(\varphi \circ \pi)(e). \tag{12.4}$$

To show that $v \in \mathfrak{h}$, it suffices by Corollary 12.3 to show that $\exp(tv) \in H$ for all $t \in \mathbb{R}$. Let $\gamma(t) := \varphi(\pi(\exp(tv)))$. We will show that $\gamma'(t) \equiv 0$, which implies that γ is the constant curve $\gamma(t) \equiv x$, and then by definition of H it follows that $\exp(tv) \in H$ for all t. We now compute using (12.2) with $a = \exp(tv)$ that

$$\begin{aligned} \gamma'(t) &= D\mu_{\exp(tv)} \circ D(\varphi \circ \pi) \circ \underbrace{Dl_{\exp(-tv)}(\exp(tv)) \left[\frac{d}{dt}\bigg|_t \exp(tv) \right]}_{=v \text{ by Proposition 10.9}} \\ &= D\mu_{\exp(tv)} \circ \underbrace{D(\varphi \circ \pi)(e)[v]}_{=0 \text{ by (12.4)}} = 0. \end{aligned}$$

This completes the proof. ■

Thus we can equivalently define a homogeneous space as a smooth manifold that admits a transitive Lie group action. We emphasise a given smooth manifold can sometimes be made into a homogeneous space in multiple ways.

Example 12.12. The Lie group $\mathrm{GL}(n)$ acts on $\mathbb{R}^n$. This in itself is not very interesting, but observe the action of $\mathrm{O}(n) \subset \mathrm{GL}(n)$ restricts to a *transitive* action on $S^{n-1} \subset \mathbb{R}^n$ by elementary linear algebra. Moreover the isotropy subgroup of $e_n = (0, 0, \dots, 0, 1) \in S^{n-1}$ is given by those matrices $A \in \mathrm{O}(n)$ of the form

$$A = \begin{pmatrix} \begin{pmatrix} & & \\ & B & \\ & & \end{pmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 \end{matrix} & 1 \end{pmatrix}$$

where $B \in \mathrm{O}(n-1)$. We conclude that S^{n-1} is the homogeneous space

$$S^{n-1} \cong \mathrm{O}(n) / \mathrm{O}(n-1).$$

The same argument works to show that

$$S^{n-1} \cong \mathrm{SO}(n) / \mathrm{SO}(n-1).$$

Example 12.13. Let $\mathrm{U}(n) \subset \mathrm{GL}(n; \mathbb{C})$ denote the unitary group and $\mathrm{SU}(n) \subset \mathrm{U}(n)$ the special unitary group. If we regard S^{2n-1} as the unit sphere in $\mathbb{C}^n$ then a similar argument shows that

$$S^{2n-1} \cong \mathrm{U}(n) / \mathrm{U}(n-1) \qquad \text{and} \qquad S^{2n-1} \cong \mathrm{SU}(n) / \mathrm{SU}(n-1).$$

For $n = 1$, $\mathrm{SU}(1)$ is just the 1×1 identity matrix. Thus S^3 is diffeomorphic to $\mathrm{SU}(2)$, and hence S^3 can be given a Lie group structure.

(♣) Remark 12.14. Not all smooth manifolds admit the structure of a Lie group. For instance, S^n admits a Lie group structure only for $n = 0, 1$ or $n = 3$. For $n = 0$ this is trivial. For $n = 1$, this was part (vi) from Example 9.4 above, and we just did the case of S^3 in Example 12.13. The proof that no other sphere admits a Lie group structure is quite tricky, but roughly speaking proceeds as follows: suppose S^n admits a Lie group structure for $n > 1$. Since S^n is simply connected for $n > 1$, the Lie group structure is necessarily non-abelian. Next, one can show[3] that any compact non-abelian Lie group G carries a natural closed but not exact bi-invariant differential 3-form. Thus $H^3(G; \mathbb{R}) \neq 0$. For S^n this forces $n = 3$.

[3]This will be an exercise on one of the Problem Sheets next semester.

LECTURE 13

Fibre bundles and vector bundles

In this lecture we define the general notion of a **fibre bundle**. We then quickly specialise to the case of **vector bundle**. In Lecture 24 we will focus on another type of fibre bundle called a **principal bundle**.

DEFINITION 13.1. Let E, M and F be smooth manifolds, and suppose $\pi\colon E \to M$ is a smooth surjective map. We say that $\pi\colon E \to M$ is a **fibre bundle over** M **with fibre** F if for every point $x \in M$ there exists a neighbourhood U of x and a smooth map

$$\alpha\colon \pi^{-1}(U) \to F$$

such that

$$(\pi, \alpha)\colon \pi^{-1}(U) \to U \times F$$

is a diffeomorphism. We call $\alpha\colon \pi^{-1}(U) \to F$ a **bundle chart**. We call E the **total space** of the bundle, M the **base space**, and F the **fibre**.

We should really say "smooth fibre bundle", but since we won't ever have cause in this course to look at non-smooth fibre bundles, we omit the adjective smooth. Note that unlike the definition of a smooth atlas on a manifold, there are no compatibility conditions on the bundle charts α. This is because the spaces E, M and F are all already manifolds. In particular if $\alpha\colon \pi^{-1}(U) \to F$ and $\beta\colon \pi^{-1}(V) \to F$ are two bundle charts with $U \cap V \neq \emptyset$ then $\alpha|_{\pi^{-1}(U\cap V)}$ and $\beta|_{\pi^{-1}(U\cap V)}$ are two more bundle charts. Any collection of bundle charts with the property that the corresponding sets $U \subset M$ form an open cover of M is called a **bundle atlas**.

EXAMPLE 13.2. The simplest example of a fibre bundle is the product manifold $E = M \times F$ with $\pi\colon M \times F \to M$ the first projection. In this case we can take U to be all of M and define $\alpha\colon M \times F \to F$ to be the second projection. More generally, any fibre bundle E which is globally diffeomorphic to $M \times F$ is called a **trivial bundle**.

We will usually abuse notation and write just E for the fibre bundle over M (or sometimes just π), rather than $\pi\colon E \to M$. A typical point in E will be written with the letter p or q (in contrast to points in M that are written with x and y).

DEFINITION 13.3. Given a fibre bundle E over M with fibre F, we set $E_x := \pi^{-1}(x)$ for $x \in M$ and call E_x the **fibre** over x.

Each fibre E_x is diffeomorphic to F, as the next lemma shows.

LEMMA 13.4. *Let $\pi\colon E \to M$ be a fibre bundle with fibre F. Then π is a submersion, and moreover each fibre E_x is an embedded submanifold of E which is diffeomorphic to F.*

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Proof. Let $\alpha\colon \pi^{-1}(U) \to F$ be a bundle chart, and let $\mathrm{pr}_1\colon U\times F \to U$ and $\mathrm{pr}_2\colon U\times F\to F$ denote the two projections. Then $D\mathrm{pr}_1(x,z)$ maps $T_{(x,z)}(M\times F)$ onto T_xM, and hence pr_1 is a submersion. Since (π,α) is a diffeomorphism, its differential at a point p is a bijection $T_pE \to T_{(\pi(p),\alpha(p))}(M\times F)$. Then $D\pi(p)$ is the composition

$$D\pi(p) = D\,\mathrm{pr}_1(\pi(p),\alpha(p))\circ D(\pi,\alpha)(p)\colon T_pE\to T_{\pi(p)}M$$

and hence is surjective. Thus π is submersion. The Implicit Function Theorem 5.13 then tells us that each fibre E_x is naturally an embedded submanifold of E. Finally, for each $x\in U$, (π,α) maps E_x diffeomorphically onto the embedded submanifold $\{x\}\times F$ of $U\times F$, and pr_2 is a diffeomorphism from this onto F. ■

Definition 13.5. Let $M\subset N$ be an embedded submanifold. Suppose $\pi\colon E\to M$ is a fibre bundle with fibre F and $\pi_1\colon E_1\to N$ is another fibre bundle with fibre F_1. Assume that $F\subset F_1$ and $E\subset E_1$ are also both embedded submanifolds. We say that E is a **subbundle** of E_1 if for every bundle chart $\alpha\colon\pi^{-1}(U)\to F$ of E and any $x\in U$ there exists a bundle chart $\beta\colon \pi_1^{-1}(V)\to F_1$, where V is an open neighbourhood of x in N such that

$$\big(\pi,\alpha|_{\pi^{-1}(U\cap V)}\big) = (\pi_1,\beta)|_{\pi_1^{-1}(U\cap V)}.$$

In short, this means that the bundle chart β on E_1 restricts to a bundle chart of E. Not all bundle charts have this property if $F\neq F_1$, as there are diffeomorphisms of F_1 that do not map F to itself.

Example 13.6. If E is any fibre bundle and $\alpha\colon\pi^{-1}(U)\to F$ is a bundle chart, then we can consider $\pi|_{\pi^{-1}(U)}\colon\pi^{-1}(U)\to U$ as a fibre bundle in its own right. This fibre bundle is trivial and is a subbundle of E. As a result we often say that E is **trivial over** U if there exists a bundle chart with domain $\pi^{-1}(U)$.

Suppose now that $\alpha\colon\pi^{-1}(U)\to F$ and $\beta\colon\pi^{-1}(V)\to F$ are two bundle charts with $U\cap V\neq\emptyset$. Then $\alpha|_{\pi^{-1}(U\cap V)}$ and $\beta|_{\pi^{-1}(U\cap V)}$ are two more bundle charts. Since for each $x\in U\cap V$, both α and β map the fibre E_x diffeomorphically onto F, we have a well defined map:

$$\rho_{\alpha\beta}\colon U\cap V\to \mathrm{Diff}(F),\qquad \rho_{\alpha\beta}(x) := \alpha|_{E_x}\circ\beta|_{E_x}^{-1}.$$

We usually call $\rho_{\alpha\beta}$ the **transition function**[1] from the bundle chart α to the bundle chart β, and refer to the collection $\{\rho_{\alpha\beta}\}$ of all transitions functions arising from the bundle atlas as the **transition functions of the bundle atlas**. Thus if $p\in\pi^{-1}(U\cap V)$ one has

$$\alpha(p) = \rho_{\alpha\beta}(\pi(p))(\beta(p)).$$

If $\gamma\colon\pi^{-1}(W)\to F$ is another bundle chart with $U\cap V\cap W\neq\emptyset$ then the following **cocycle condition** is automatically satisfied:

$$\rho_{\alpha\gamma}(x) = \rho_{\alpha\beta}(x)\circ\rho_{\beta\gamma}(x),\qquad \forall\, x\in U\cap V\cap W. \tag{13.1}$$

The composition on the right-hand side occurs in $\mathrm{Diff}(F)$. The meaning of the word "cocycle" will be explained in Remark 14.12 next lecture. In particular,

$$\rho_{\alpha\beta}(x)^{-1} = \rho_{\beta\alpha}(x). \tag{13.2}$$

[1] *Warning*: This is a slightly different meaning of the word "transition function" than was used in Definition 1.15.

Remark 13.7. Suppose $\alpha\colon \pi^{-1}(W) \to F$ is a bundle chart on E. Let $\sigma\colon U \to O$ and $\tau\colon V \to \Omega$ be (manifold) charts on M and F respectively with $U \subset W$. Then $(\sigma \circ \pi, \tau \circ \alpha)$ is a manifold chart on an open set in E which is compatible with the given smooth structure on E.

It is often useful to work backwards. Suppose we begin with a *set* E and a surjective map $\pi\colon E \to M$, where M is a smooth manifold. Suppose in addition we are given another smooth manifold F and an open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of M, together with a collection of functions

$$\alpha_{\mathtt{a}}\colon \pi^{-1}(U_{\mathtt{a}}) \to F$$

such that $(\pi, \alpha_{\mathtt{a}})$ is a bijection for each $\mathtt{a} \in \mathtt{A}$. We can then attempt to *define* a smooth structure by declaring that charts on E are of the form $(\sigma \circ \pi, \tau \circ \alpha_{\mathtt{a}})$, where σ is a chart on M defined on an open subset of $U_{\mathtt{a}}$, and τ is some chart on F. Of course, now there is something to check. By Proposition 1.22, if one can verify that the transition functions are diffeomorphisms, this will endow E with a smooth manifold structure in such a way that the $\{\alpha_{\mathtt{a}}\}$ become a bundle atlas.

Definition 13.8. Let F be a smooth manifold and suppose we are given a left action $\mu\colon G \times F \to F$ of a Lie group G on F. We say that μ is an **effective** action if the only element $a \in G$ for which $\mu_a = \mathrm{id}$ is $a = e$.

If $\mu\colon G \times F \to F$ is an effective action then the map $a \mapsto \mu_a$ defines an injective group homomorphism $G \to \mathrm{Diff}(F)$. Thus via μ we can view G as a subgroup of $\mathrm{Diff}(F)$.

Definition 13.9. Suppose $\pi\colon E \to M$ is a fibre bundle with fibre F, and suppose G is a Lie group acting effectively on F via $\mu\colon G \times F \to F$. We say that two bundle charts $\alpha\colon \pi^{-1}(U) \to F$ and $\beta\colon \pi^{-1}(V) \to F$ with $U \cap V \neq \emptyset$ are **(G, μ)-compatible** if there exists a smooth map $\tilde{\rho}_{\alpha\beta}\colon U \cap V \to G$ such that

$$\rho_{\alpha\beta}(x)(z) = \mu(\tilde{\rho}_{\alpha\beta}(x), z), \qquad \forall\, x \in U \cap V,\ \forall\, z \in F. \tag{13.3}$$

We will usually omit explicit reference to the action μ and just say that the two bundle charts are **G-compatible**. Similarly wherever possible we will suppress the difference between $\rho_{\alpha\beta}$ and $\tilde{\rho}_{\alpha\beta}$ and write them both as $\rho_{\alpha\beta}$. This is usually harmless, since most of the time the action μ is considered to be fixed, and is thus suppressed from the notation. In particular, if $F = \mathbb{R}^k$ is a vector space and $G \subset \mathrm{GL}(k)$ is a matrix Lie group then the action μ is *always* understood to be the standard one. If however the particular choice of μ is important (or non-standard) then we will continue to include it in our notation and terminology.

Definition 13.10. Let $\pi\colon E \to M$ be a fibre bundle with fibre F, and let $\mu\colon G \times F \to F$ be an effective Lie group action. A **G-bundle atlas** is a bundle atlas for E with the property that if $\alpha\colon \pi^{-1}(U) \to F$ and $\beta\colon \pi^{-1}(V) \to F$ are any two charts in this atlas whose domains intersect, then α and β are G-compatible. If such an atlas exists, we say that E is a fibre bundle with **structure group** G.

Remark 13.11. Just as with smooth atlases on manifolds, since G-bundle atlases come with compatibility conditions, the union of two G-bundle atlases may not be still be a G-bundle atlas. However we can define an equivalence relation on the set of G-bundle atlases by declaring two atlases to be equivalent if their union is another G-bundle atlas. We then define a **G-bundle structure** to be an equivalence class. Alternatively, a G-bundle structure can be thought of as a maximal G-bundle atlas. (Compare Remark 1.17). In practice however, just as with smooth atlases versus smooth structures on manifolds, the distinction is usually unimportant.

EXAMPLE 13.12. Let G be a Lie group and $H \subset G$ a closed subgroup. Then $\pi\colon G \to G/H$ is a fibre bundle with fibre H and structure group H (where H acts on itself via left translations). *Exercise:* Prove this using the proof of Theorem 12.4.

DEFINITION 13.13. A **vector bundle of rank** k is a fibre bundle $\pi\colon E \to M$ whose fibre is $F = \mathbb{R}^k$ and whose structure group is $G = \mathrm{GL}(k)$, or some matrix Lie subgroup thereof.

REMARK 13.14. It is important to realise that E can have structure group G for many different Lie groups G (and thus we should really say "a structure group" rather than "the structure group"). Indeed, it is often advantageous to make the structure group as small as possible: if E has structure group G and $H \subset G$ is a subgroup, then sometimes it is possible to find a new G-bundle atlas such that each transition function $\rho_{\alpha\beta}$ takes image in $H \subset G$. Then this G-bundle atlas is actually an H-bundle atlas, and we say that we have **reduced the structure group to** H.

For instance, if E is a vector bundle of rank k then it is always possible to reduce the structure group to $\mathrm{O}(k) \subset \mathrm{GL}(k)$ (we will prove this next semester when we discuss metrics, see Corollary 36.13). On the other hand, only sometimes is it possible to reduce the structure group to $\mathrm{GL}^+(k)$ (as we will see, this is equivalent to the bundle being *orientable*, cf Proposition 20.16).

EXAMPLE 13.15. Let M be a smooth manifold. Then the tangent bundle $\pi\colon TM \to M$ is a vector bundle of rank n over M. Indeed, if $\sigma\colon U \to O$ is a chart on M with local coordinates x^i then if we set

$$\alpha\colon \pi^{-1}(U) \to \mathbb{R}^n, \qquad \alpha(x,v) = dx^i|_x(v)e_i$$

as in Theorem 4.16, then α is a bundle chart. Moreover if $\tau\colon V \to \Omega$ is another chart on M with overlapping domain and β the corresponding map $\beta\colon \pi^{-1}(V) \to \mathbb{R}^n$ then

$$\rho_{\alpha\beta}(x) = D(\sigma\circ\tau^{-1})(\tau(x)) \in \mathrm{GL}(n) \subset \mathrm{Diff}(\mathbb{R}^n).$$

by (4.8). A similar argument shows that the cotangent bundle $T^\star M$ is another vector bundle of rank n over M.

Thus a vector bundle of rank k is a fibre bundle whose fibre is $\mathbb{R}^k$, but not every fibre bundle whose fibre is $\mathbb{R}^k$ is a vector bundle (due to the additional requirement that the structure group is $\mathrm{GL}(k)$). The next proposition clarifies the difference.

PROPOSITION 13.16. *Let $\pi\colon E \to M$ be a fibre bundle with fibre $F = \mathbb{R}^k$. Then E is a vector bundle if and only if it is possible to endow each fibre E_x with a vector space structure and find a bundle atlas with the property that if $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ is any bundle chart belonging to the atlas then $\alpha|_{E_x}\colon E_x \to \mathbb{R}^k$ is a vector space isomorphism for each $x \in U$.*

Proof. If $\pi\colon E \to M$ is a vector bundle of rank k then the fact that each fibre E_x admits the structure of a vector space of dimension k follows from Problem B.1, since each transition function $\rho_{\alpha\beta}(x)$ is a linear isomorphism. By construction, this vector space structure has the property that each $\alpha|_{E_x}\colon E_x \to \mathbb{R}^k$ is a vector space isomorphism.

For the converse, we simply note that if $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ and $\beta\colon \pi^{-1}(V) \to \mathbb{R}^k$ are two overlapping bundle charts as in the statement then for $x \in U \cap V$, $c \in \mathbb{R}$ and $v, w \in \mathbb{R}^k$, one has

$$\begin{aligned}\rho_{\alpha\beta}(x)[cv + w] &= \alpha|_{E_x} \circ \beta|_{E_x}^{-1}(cv + w) \\ &= \alpha|_{E_x}\big(c\beta|_{E_x}^{-1}(v) + \beta|_{E_x}^{-1}(w)\big) \\ &= c\rho_{\alpha\beta}(x)[v] + \rho_{\alpha\beta}(x)[w]\end{aligned}$$

Thus each $\rho_{\alpha\beta}$ is linear, as required. ■

Proposition 13.16 allows us to make the following alternative definition.

Definition 13.17. Let $\pi\colon E \to M$ be a surjective map between two smooth manifolds, and set $E_x := \pi^{-1}(x)$. We say that E is a **vector bundle of rank** k if each E_x admits the structure of a k-dimensional vector space, and any $x \in M$ has a neighbourhood U together with a smooth map $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ such that

(i) $(\pi, \alpha)\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$ is a diffeomorphism,

(ii) $\alpha|_{E_x}\colon E_x \to \mathbb{R}^k$ is a vector space isomorphism.

We will call such a map α a **vector bundle chart** (instead of a $\mathrm{GL}(k)$-bundle chart).

Let us explore some more constructions with fibre bundles.

Example 13.18. Let $\pi_i\colon E_i \to M_i$ be fibre bundles with fibres F_i and structure groups G_i for $i = 1, 2$. Then $(\pi_1, \pi_2)\colon E_1 \times E_2 \to M_1 \times M_2$ is another fibre bundle with fibre $F_1 \times F_2$ and structure group $G_1 \times G_2$. In particular, the product of two vector bundles is another vector bundle. An example of this is the tangent bundle $T(M_1 \times M_2)$, which is diffeomorphic to $TM_1 \times TM_2$.

Here is a generalisation of the preceding example, which is typically much more interesting.

Example 13.19. Let $\varphi\colon M \to N$ be a smooth map, and suppose $\pi\colon E \to N$ is a fibre bundle with fibre F. Then $M \times E$ is a trivial fibre bundle over M with fibre E. Whilst this is not a particularly interesting bundle, we can use φ to define a much more interesting subbundle. The **pullback bundle** $\varphi^\star E$ is defined as follows: Set

$$\varphi^\star E := \{(x, p) \in M \times E \mid \varphi(x) = \pi(p)\},$$

with projection $\mathrm{pr}_1\colon \varphi^\star E \to M$. The fibre of $\varphi^\star E$ over x in M is $\{x\} \times E_{\varphi(x)}$, which is diffeomorphic to $E_{\varphi(x)}$ under pr_2. If $\alpha\colon \pi^{-1}(U) \to F$ is a bundle chart for E then $\alpha \circ \mathrm{pr}_2\colon \mathrm{pr}_1^{-1}(\varphi^{-1}(U)) \to F$ is a bundle chart for $\varphi^\star E$. Thus $\varphi^\star E$ is a fibre bundle over M with fibre F, and we can summarise this with the following commuting picture:

$$\begin{tikzcd} \varphi^\star E \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\mathrm{pr}_1"'] & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

Note that $\varphi^\star E$ is a subbundle of $M \times E$. On Problem Sheet G you are asked to verify the following two additional properties of $\varphi^\star E$:

(i) If E has structure group G then $\varphi^\star E$ has structure group a Lie subgroup of G. In particular, if E is a vector bundle then so is $\varphi^\star E$.

(ii) The tangent bundle of $\varphi^\star E$ (this makes sense, since $\varphi^\star E$ is a manifold) is given by

$$T_{(x,p)}(\varphi^\star E) = \{(v,\zeta) \in T_xM \times T_pE \mid D\varphi(x)[v] = D\pi(p)[\zeta]\}\,.$$

Another example is given by composition.

EXAMPLE 13.20. Let $\pi\colon E \to M$ be a fibre bundle with fibre F. Assume in addition that E is itself the base space of another fibre bundle $\pi_1\colon E_1 \to E$ with fibre F_1. Then $\pi \circ \pi_1\colon E_1 \to M$ is a fibre bundle with fibre $F \times F_1$ called the **composite bundle**. Indeed, if $\alpha\colon \pi^{-1}(U) \to F$ is a bundle chart for E over $U \subset M$ and $\alpha_1\colon \pi_1^{-1}(U_1) \to F_1$ is a bundle chart for E_1 over $U_1 \subset E$ such that $W := \pi(U_1) \cap U \neq \emptyset$, then

$$\alpha_2 := (\alpha \circ \pi_1, \alpha_1)\colon (\pi \circ \pi_1)^{-1}(W) \to F \times F_1$$

is a bundle chart for $\pi \circ \pi_1$ over W. If $\beta\colon \pi^{-1}(V) \to F$ and $\beta_1\colon \pi_1^{-1}(V_1) \to F_1$ are two more choices of bundle charts on E and E_1 respectively such that $W_1 := \pi(V_1) \cap V$ has non-empty intersection with W, then if β_2 is the corresponding bundle chart for $\pi \circ \pi_1$ over W_1 then I invite you to check that the transition function $\rho_{\alpha_2\beta_2}$ is given by

$$\rho_{\alpha_2\beta_2}(x)(z,z_1) := \big(\rho_{\alpha\beta}(x)(z), \rho_{\alpha_1\beta_1} \circ (\pi,\beta)^{-1}(x,z)(z_1)\big)$$

for $(z,z_1) \in F \times F_1$ and $x \in W \cap W_1$.

As a concrete example of this, if M^n is a manifold then the tangent bundle TM is another manifold, and its tangent bundle $T(TM)$ is then a vector bundle over M of rank $3n$.

LECTURE 14

Constructing new vector bundles

We begin this lecture with a recipe for constructing fibre bundles starting from the transition functions. We state the general version for fibre bundles, but we will mainly use this for vector bundles.

THEOREM 14.1. *Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open covering of a manifold M. Let G be a Lie group. Suppose for each $\mathtt{a}, \mathtt{b} \in \mathtt{A}$ such that $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$, we are given a smooth map $\rho_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G$ such that the following cocycle condition is satisfied:*

$$\begin{cases} \rho_{\mathtt{ac}}(x) = \rho_{\mathtt{ab}}(x)\rho_{\mathtt{bc}}(x), & \forall x \in U_{\mathtt{a}} \cap U_{\mathtt{b}} \cap U_{\mathtt{c}}, \text{ if } U_{\mathtt{a}} \cap U_{\mathtt{b}} \cap U_{\mathtt{c}} \neq \emptyset, \\ \rho_{\mathtt{aa}}(x) = e, & \forall\, x \in U_{\mathtt{a}}, \ \forall\, \mathtt{a} \in \mathtt{A}. \end{cases} \tag{14.1}$$

Suppose in addition we are given an effective action $\mu \colon G \times F \to F$ of G on a manifold F. Then there exists a fibre bundle $\pi \colon E \to M$ with fibre F and structure group G. Moreover there is a bundle atlas $\{\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to F \mid \mathtt{a} \in \mathtt{A}\}$ such that the transition function $\rho_{\alpha_{\mathtt{a}}\alpha_{\mathtt{b}}}$ are given by $\rho_{\mathtt{ab}}$.

As you might expect from a theorem with such complicated hypotheses (compare the Proposition 1.22), the proof is basically trivial—most of the work is in formulating the hypotheses correctly!

Proof. Let

$$E := \left(\bigsqcup_{\mathtt{a} \in \mathtt{A}} U_{\mathtt{a}} \times F \right) \Big/ \sim,$$

where we identify $(x, p) \in U_{\mathtt{a}} \times F$ with $(y, q) \in U_{\mathtt{b}} \times F$ if and only if $x = y$ and

$$p = \mu(\rho_{\mathtt{ab}}(x), q).$$

Let $\pi \colon E \to M$ denote the map induced by first projection. Let $\Pi \colon \bigsqcup_{\mathtt{a} \in \mathtt{A}} U_{\mathtt{a}} \times F \to E$ denote the projection. Then for each $\mathtt{a} \in \mathtt{A}$, the restriction of Π to $U_{\mathtt{a}} \times F$ onto its image in E is a homeomorphism. Its inverse is of the form $(\pi, \alpha_{\mathtt{a}})$, where $\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to F$. This is our desired bundle atlas on E: we first make E into a smooth manifold using the procedure outlined in the second half of Remark 13.7—the fact that this gives a well-defined smooth structure follow from (14.1). It is then immediate from the definition that the transition functions of this bundle atlas are given by the maps $\rho_{\mathtt{ab}}$. This completes the proof. ∎

EXAMPLE 14.2. Take $G = \mathbb{Z}_2 = \{\pm 1\}$ and take $F = \mathbb{R}$, with G acting by multiplication. We take $M = S^1 \subset \mathbb{C}$. Let $U_1 = S^1 \setminus \{i\}$ and $U_2 := S^1 \setminus \{-i\}$. By Theorem 14.1 a smooth map $\rho_{12} \colon U_1 \cap U_2 \to G$ determines a vector bundle of rank 1 over M. If we set

$$\rho_{21}(z) := \begin{cases} 1, & \Re(z) > 0, \\ -1, & \Re(z) < 0, \end{cases}$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

then the vector bundle so obtained is called the **Möbius band**[1].

For the rest of this lecture we will work exclusively with vector bundles, since this allows for slightly simpler definitions. We now formulate the correct notion of a map from one vector bundle to another.

DEFINITION 14.3. Let $\pi_i\colon E_i \to M_i$ denote two vector bundles. Suppose we are given two smooth maps $\Phi\colon E_1 \to E_2$ and $\varphi\colon M_1 \to M_2$. We say that Φ is a **vector bundle morphism along** φ if the restriction of Φ to each fibre $E_1|_x$ is a *linear* map from the vector space $E_1|_x$ to the vector space $E_2|_{\varphi(x)}$. Thus the following commutes:

$$\begin{tikzcd} E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\ M_1 \arrow[r, "\varphi"'] & M_2 \end{tikzcd}$$

If Φ maps each fibre $E_1|_x$ isomorphically onto $E_2|_{\varphi(x)}$ then Φ is called a **vector bundle isomorphism along** φ, and we say that E_1 and E_2 are **isomorphic vector bundles along** φ.

EXAMPLE 14.4. Let $\varphi\colon M \to N$ be a smooth map between two smooth manifolds. Then $D\varphi\colon TM \to TN$ is a vector bundle morphism along φ.

Let us spell out a particular case in detail, where we take φ to be the identity.

DEFINITION 14.5. Let $\pi_i\colon E_i \to M$ be two vector bundles over the same base space M. A **vector bundle homomorphism** $\Phi\colon E_1 \to E_2$ is simply a vector bundle morphism along the identity map $\mathrm{id}\colon M \to M$. That is, Φ is a smooth map that maps each fibre linearly to itself. If Φ maps each fibre $E_1|_x$ isomorphically onto $E_2|_x$ (i.e. so that Φ is a vector bundle isomorphism along the identity map) then Φ is called a **vector bundle isomorphism**, and E_1 and E_2 are said to be **isomorphic vector bundles**.

REMARK 14.6. A vector bundle homomorphism is a vector bundle isomorphism if and only if it is a diffeomorphism. This is not true for vector bundle morphisms along a map. For instance, if M is a manifold and $x \in M$ then (thinking of $\{x\}$ as a zero-dimensional manifold) we have a smooth map $\imath_x\colon \{x\} \hookrightarrow M$ given by inclusion. If E is any vector bundle over M then the inclusion map $E_x \hookrightarrow E$ is a vector bundle isomorphism along $\imath_x$, but of course it is not a diffeomorphism.

EXAMPLE 14.7. Let us revisit the pullback bundle construction from Example 13.19. Let $\pi\colon E \to N$ be a vector bundle and let $\varphi\colon M \to N$ be a smooth map. Then $\varphi^\star E$ is isomorphic to E along φ via $\mathrm{pr}_2\colon \varphi^\star E \to E$. Moreover $\varphi^\star E$ is unique up to isomorphism in the following sense: if $E_1 \to M$ is another vector bundle on M then E_1 is isomorphic to $\varphi^\star E$ if and only if E_1 is isomorphic to E along φ. Moreover pullbacks are functorial[2]: if $\psi\colon L \to M$ is another smooth map then $\psi^\star(\varphi^\star E)$ is isomorphic to $(\varphi \circ \psi)^\star E$ (see Problem Sheets G and H).

[1]More accurately: a Möbius band of infinite width.

[2]We are about to define this word precisely!

EXAMPLE 14.8. If $M \subset N$ is an embedded submanifold then a vector bundle E over M is a **vector subbundle** (i.e. a subbundle in the sense of Definition 13.5, where all bundle charts are required to be vector bundle charts) of a vector bundle $E_1 \to N$ if and only the inclusion $E \hookrightarrow E_1$ is a vector bundle morphism along the inclusion $M \hookrightarrow N$.

EXAMPLE 14.9. If Δ is a distribution on M then one can think of Δ as a vector subbundle of TM.

The next result clarifies the relation between the isomorphism class of a vector bundle and its transition functions. The proof is deferred to Problem Sheet H.

PROPOSITION 14.10. *Let M be a smooth manifold and suppose $\pi_i \colon E_i \to M$ are two vector bundles over M of the same rank k. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both*[3] *E_1 and E_2 admit $\mathrm{GL}(k)$-bundle atlases over the $U_{\mathtt{a}}$. Let*

$$\rho^1_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k), \qquad \text{and} \qquad \rho^2_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k)$$

denote the transition functions of E_1 and E_2 with respect to these bundle atlases. Then E_1 and E_2 are isomorphic if and only if there exists a smooth family $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to \mathrm{GL}(k)$ of functions such that

$$\nu_{\mathtt{a}}(x) \circ \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \circ \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}.$$

COROLLARY 14.11. *The vector bundle constructed in Theorem 14.1 is unique up to isomorphism.*

(♣) REMARK 14.12. Proposition 14.10 tells us that the isomorphism classes of rank k vector bundles over M are in bijective correspondence with $\check{H}^1(M; \underline{\mathrm{GL}}(k))$, i.e. the first *Čech cohomology* group of M with coefficients in the *sheaf* $C^\infty(\square, \mathrm{GL}(k)) =: \underline{\mathrm{GL}}(k)$.

We now move onto to a general method to create new vector bundles from old ones. We will first state and imprecise "metatheorem" and explain how to use it. A formal proof is rather more involved, and requires ideas from a field of mathematics called *category theory*.

METATHEOREM. *Anything you can do with vector spaces, you can also do with vector bundles.*

What does this mean? Let us look at some examples:

(i) If V is a vector space then its dual space $V^* = \mathrm{L}(V, \mathbb{R})$ is another vector space of the same dimension. Thus if $E \to M$ is a vector bundle, we can form a new vector bundle $E^* \to M$ called the **dual bundle** by setting

$$E^* := \bigsqcup_{x \in M} E^*_x$$

where $E^*_x = \mathrm{L}(E_x, \mathbb{R})$ is the dual vector space to E_x. The cotangent bundle is an example of this construction: it is the dual bundle to the tangent bundle.

[3] This can always be achieved by taking intersections.

(ii) If V and W are vector spaces then their **direct sum** $V \oplus W$ is another vector space of dimension $\dim V + \dim W$. Thus if $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ are two vector bundles over the same manifold M, we can form the **direct sum bundle** $E_1 \oplus E_2 \to M$ by setting

$$E_1 \oplus E_2 := \bigsqcup_{x \in M} E_1|_x \oplus E_2|_x$$

Warning: As vector spaces, the direct sum $V \oplus W$ is the same thing as the product $V \times W$, and we often use the notation interchangeably (for instance, we usually write $\mathbb{R}^n \times \mathbb{R}^k$, not $\mathbb{R}^n \oplus \mathbb{R}^k$, and next lecture we will typically write $\times$ instead of $\oplus$ when discussing multilinear maps). However as vector bundles, we will exclusively use the direct sum notation, because the product $E_1 \times E_2$ refers to the product bundle over $M \times M$ as defined in Example 13.18. The two concepts are related however: if $\delta \colon M \to M \times M$ is the *diagonal map* $x \mapsto (x, x)$ then

$$\delta^\star(E_1 \times E_2) \cong E_1 \oplus E_2,$$

as I invite you to verify.

(iii) If V and W are vector spaces then $\mathrm{L}(V, W)$, the space of linear maps from V to W is another vector space of dimension $\dim V \cdot \dim W$. (The dual space is the special case where $W = \mathbb{R}$). Thus if $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ are two vector bundles, the **homomorphism bundle** $\mathrm{Hom}(E_1, E_2)$ is defined by

$$\mathrm{Hom}(E_1, E_2) := \bigsqcup_{x \in M} \mathrm{L}(E_1|_x, E_2|_x) \tag{14.2}$$

The fact that the left-hand side is written $\mathrm{Hom}(E_1, E_2)$ rather than the (slightly more logical name) $\mathrm{L}(E_1, E_2)$ will be explained in Lecture 16 (see part (iv) of Example 16.2).

(iv) One can iterate these constructions: for instance, if E_1, E_2, E_3, E_4, E_5 are vector bundles of rank k_i then

$$\mathrm{L}(E_1 \oplus E_2^*, \mathrm{L}(E_3^*, E_4 \oplus E_5))$$

is another vector bundle of rank $(k_1 + k_2)k_3(k_4 + k_5)$.

In the next lecture, we will see two more important cases of this: the **tensor bundle** $E_1 \otimes E_2$ and the **exterior algebra bundle** $\bigwedge(E)$.

Of course, in all of these examples we have not proved that these constructions give rise to vector bundles: we have simply stuck vector spaces over each point in M and claimed that the resulting object is a vector bundle. In all of these cases, it is not too hard to prove this directly (i.e. that the total space is a smooth manifold, that the relevant bundle atlas exists, etc). However doing so would be very "ad hoc"—one would need a separate proof for every example, and, as part (iv) above showed, since we can iterate we would swiftly need infinitely many proofs. This is not ideal. Thus we will search for a way to prove everything in one fell swoop.

Enter category theory...

Definition 14.13. A **category** C consists of three ingredients. The first is a *class* $\mathrm{obj}(\mathsf{C})$ of **objects**. Secondly, for each ordered pair of objects (A, B) there is a *set* $\mathrm{Hom}(A, B)$ of **morphisms** from A to B. Sometimes instead of $f \in \mathrm{Hom}(A, B)$ we write $f : A \to B$ or $A \xrightarrow{f} B$. Finally, there is a rule, called **composition**, which associates to every ordered triple (A, B, C) of objects a map

$$\mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C),$$

written

$$(f, g) \mapsto g \circ f,$$

which satisfies the following three axioms:

1. The Hom sets are pairwise disjoint; that is, each $f \in \mathrm{Hom}(A, B)$ has a unique **domain** A and a unique **target** B.

2. Composition is associative whenever defined, i.e. given

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$$

one has

$$(h \circ g) \circ f = h \circ (g \circ f).$$

3. For each $A \in \mathrm{obj}(\mathsf{C})$ there is a unique morphism $\mathrm{id}_A \in \mathrm{Hom}(A, A)$ called the *identity* which has the property that $f \circ \mathrm{id}_A = f$ and $\mathrm{id}_B \circ f = f$ for every $f : A \to B$.

(♣) Remark 14.14. Note that we said that $\mathrm{obj}(\mathsf{C})$ was a *class* and $\mathrm{Hom}(A, B)$ was a *set*. There is (an important, but technical) difference between a class and a set. If you've ever taken a class on logic/set theory, you'll know that not every "collection" of objects is formally a set. For instance, the collection of all sets is itself not a set. A class is a more general concept (the collection of all sets is a class). Nevertheless, as far as this course is concerned, the distinction is essentially irrelevant.

Remark 14.15. A word of warning: category theory is often (lovingly) referred to as abstract nonsense. But fear not: nothing we will do will ever be *that* abstract!

Here are six examples of categories. The first three are algebraic in nature.

Example 14.16. The category Sets of **sets**. The objects of Sets are all the sets, and $\mathrm{Hom}(A, B)$ is just the set $\mathrm{Maps}(A, B)$ of all functions from A to B, and composition is just the usual composition of functions.

Example 14.17. The category Groups of **groups**. The objects of Groups are just groups, and $\mathrm{Hom}(G, H)$ is the set of all *group homomorphisms* from G to H, and composition is just the usual composition of homomorphisms.

Example 14.18. The category $\mathsf{Vect} = \mathsf{Vect}_{\mathbb{R}}$ of **finite-dimensional real vector spaces**. The objects of Vect are finite-dimensional real vector spaces, and $\mathrm{Hom}(V, W)$ is the set $\mathrm{L}(V, W)$ of all *linear maps* from V to W.

Here are three more examples more pertinent to this course.

Example 14.19. The category Top of **topological spaces**. The objects of Top are all the topological spaces, and $\mathrm{Hom}(X, Y)$ is just the set $C(X, Y)$ of all *continuous* functions from X to Y, and composition is just the usual composition of functions.

Example 14.20. The category Man of **smooth manifolds**. The objects of Man are smooth manifolds, and $\mathrm{Hom}(M, N)$ is the set $C^\infty(M, N)$ of all *smooth* maps $\varphi\colon M \to N$. Composition is given by normal composition of maps; this is well defined by Proposition 1.26.

Example 14.21. The category $\mathsf{VectBundles}$ of **vector bundles**. The objects of $\mathsf{VectBundles}$ are vector bundles $\pi\colon E \to M$, and morphism from $\pi_1\colon E_1 \to M_1$ to $\pi_2\colon E_2 \to M_2$ is a pair (Φ, φ), where $\varphi\colon M_1 \to M_2$ is a smooth map and $\Phi\colon E_1 \to E_2$ is a vector bundle morphism from E_1 to E_2 along φ.

Remark 14.22. The category Vect is rather special: its morphism sets are themselves objects of the category. That is, if V and W are vector spaces then $\mathrm{L}(V, W)$ is itself naturally a vector space. This is not true in the category of Groups—the set of all group homomorphisms from one group to another typically does not have a group structure. Similarly the set $C^\infty(M, N)$ of smooth maps between two smooth manifolds is never itself a (finite-dimensional) manifold when $\dim M > 0$.

Remark 14.23. The fact that we require the morphism sets to be pairwise disjoint has several pedantic consequences. For example, suppose $A \subsetneq B$ are two sets. Then the inclusion $\imath\colon A \hookrightarrow B$ and the identity map $\mathrm{id}_A\colon A \to A$ are different morphisms, since they have different targets. One should be aware that we only allow the composition $g \circ f$ when the range of f is *exactly* the same as the domain of g. Suppose L, M, N and P are manifolds, and suppose M is an embedded submanifold of N. Let $\varphi\colon L \to M$ be smooth and let $\psi\colon N \to P$ be smooth. Then as we have seen, the composition $\psi \circ \varphi\colon L \to P$ is also smooth (since M is embedded). Nevertheless, from the point of view of category theory, the composition $\psi \circ \varphi$ *does not exist!* Rather, one must first take the inclusion $\imath\colon M \hookrightarrow N$ and then consider the composition $\psi \circ \imath \circ \varphi$, which is a well-defined element of the morphism space $C^\infty(L, P)$.

Definition 14.24. Suppose C and D are two categories. We say that C is a **subcategory** of D if:

1. $\mathrm{obj}(\mathsf{C}) \subseteq \mathrm{obj}(\mathsf{D})$,
2. $\mathrm{Hom}_{\mathsf{C}}(A, B) \subseteq \mathrm{Hom}_{\mathsf{D}}(A, B)$ for all $A, B \in \mathrm{obj}(\mathsf{C})$, where we denote Hom sets in C by $\mathrm{Hom}_{\mathsf{C}}(\square, \square)$,
3. if $f \in \mathrm{Hom}_{\mathsf{C}}(A, B)$ and $g \in \mathrm{Hom}_{\mathsf{C}}(B, C)$ then the composite $g \circ f \in \mathrm{Hom}_{\mathsf{C}}(A, C)$ is equal to the composite $g \circ f \in \mathrm{Hom}_{\mathsf{D}}(A, C)$,
4. if $C \in \mathrm{obj}(\mathsf{C})$ then $\mathrm{id}_C \in \mathrm{Hom}_{\mathsf{C}}(C, C)$ is equal to $\mathrm{id}_C \in \mathrm{Hom}_{\mathsf{D}}(C, C)$.

If for every pair $A, B \in \mathrm{obj}(\mathsf{C})$ one always has $\mathrm{Hom}_{\mathsf{C}}(A, B) = \mathrm{Hom}_{\mathsf{D}}(A, B)$ then we say that C is a **full subcategory** of D.

EXAMPLE 14.25. Here are two examples of subcategories:

(i) The category Ab of *abelian* groups is a full subcategory of the category Groups.

(ii) Let $\mathsf{Vect}^{\leq\infty}$ denote the category of *all* real vector spaces (finite-dimensional or infinite-dimensional). Then Vect is a full subcategory of $\mathsf{Vect}^{\leq\infty}$.

A **functor** is a map from one category to another. These come in two flavours: *covariant* and *contravariant*. We discuss the former first.

DEFINITION 14.26. Suppose C and D are two categories. A **covariant functor** $\mathsf{F}\colon \mathsf{C} \to \mathsf{D}$ associates to each $A \in \mathrm{obj}(\mathsf{C})$ an object $\mathsf{F}(A) \in \mathrm{obj}(\mathsf{D})$, and to each morphism $A \xrightarrow{f} B$ in C a morphism $\mathsf{F}(A) \xrightarrow{\mathsf{F}(f)} \mathsf{F}(B)$ in D which satisfies the following two axioms:

1. If $A \xrightarrow{f} B \xrightarrow{g} C$ in C then $\mathsf{F}(A) \xrightarrow{\mathsf{F}(f)} \mathsf{F}(B) \xrightarrow{\mathsf{F}(g)} \mathsf{F}(C)$ in D and
$$\mathsf{F}(g \circ f) = \mathsf{F}(g) \circ \mathsf{F}(f).$$

2. $\mathsf{F}(\mathrm{id}_A) = \mathrm{id}_{\mathsf{F}(A)}$ for every $A \in \mathrm{obj}(\mathsf{C})$.

The easiest example of a functor is a **forgetful functor**:

EXAMPLE 14.27. The forgetful functor $\mathsf{Top} \to \mathsf{Sets}$ simply “forgets” the topological structure. Thus it assigns to each topological space its underlying set, and to each continuous function it assigns the same function, considered now simply as a map between two *sets* (i.e. it “forgets” the function is continuous). The same thing works as a functor $\mathsf{Man} \to \mathsf{Top}$, where one “forgets” the smooth manifold structure.

EXAMPLE 14.28. There is slightly more interesting forgetful functor $\mathsf{VectBundles} \to \mathsf{Man}$ that sends a vector bundle $\pi\colon E \to M$ to its base space M (i.e. it “forgets” the vector bundle sitting over the base). On morphisms, this functor just “forgets” the vector bundle morphism: $(\Phi, \varphi) \mapsto \varphi$.

Here is a pertinent example of a functor from the category Vect to itself:

EXAMPLE 14.29. Let V be a fixed vector space. Then there is a covariant functor
$$\mathrm{L}(V, \square)\colon \mathsf{Vect} \to \mathsf{Vect}$$
that assigns to a vector space W the vector space $\mathrm{L}(V, W)$. If $T\colon W_1 \to W_2$ is a linear map then
$$\mathrm{L}(V, \square)(T)\colon \mathrm{L}(V, W_1) \to \mathrm{L}(V, W_2)$$
is given by $S \mapsto T \circ S$.

(♣) REMARK 14.30. Algebraic topology is an excellent source of functors. For instance, the **fundamental group** π_1 is a covariant functor from the **pointed homotopy category** hTop_* to Groups, and the **higher homotopy groups** are covariant functors $\pi_n\colon \mathsf{hTop}_* \to \mathsf{Ab}$. **Singular homology** (or indeed, any homology theory) is a covariant functor $\mathsf{hTop}^2 \to \mathsf{Ab}$, where hTop^2 is the homotopy category of pairs.

One can also formulate the definition of a functor of more than one variable. This requires us to define the notion of a product category.

Definition 14.31. Let C and D be two categories. The **product category** (C, D) is the category whose objects are ordered pairs (C, D) where $C \in \operatorname{obj}(\mathsf{C})$ and $D \in \operatorname{obj}(\mathsf{D})$, and

$$\operatorname{Hom}_{(\mathsf{C},\mathsf{D})}((C, D), (C', D')) = \{(f, g) \mid f \in \operatorname{Hom}_{\mathsf{C}}(C, C')\, g \in \operatorname{Hom}_{\mathsf{D}}(D, D')\}\,.$$

The composition $(f, g) \circ_{\mathsf{C}\times\mathsf{D}} (f', g')$ is defined as you expect:

$$(f, g) \circ_{(\mathsf{C},\mathsf{D})} (f', g') \coloneqq ((f \circ_{\mathsf{C}} f'), (g \circ_{\mathsf{D}} g'))\,.$$

The identity element $\mathrm{id}_{(C,D)}$ is simply the pair $(\mathrm{id}_C, \mathrm{id}_D)$.

Example 14.32. The category $(\mathsf{Vect}, \mathsf{Vect})$ has objects ordered pairs (V, W) of vector spaces, and morphisms pairs of linear maps.

Definition 14.33. A **covariant functor of two variables** is a covariant functor defined on a product category: $\mathsf{F}\colon (\mathsf{C}, \mathsf{D}) \to \mathsf{E}$.

Example 14.34. Let V and W be vector spaces. Then the[4] **direct sum** $V \oplus W$ of V and W is another vector space. Thus we get a functor $\oplus\colon (\mathsf{Vect}, \mathsf{Vect}) \to \mathsf{Vect}$ that assigns to (V, W) the vector space $V \oplus W$, and assigns to a pair (S, T) of linear maps $S\colon V_1 \to V_2$ and $T\colon W_1 \to W_2$ the linear map $S \oplus T\colon V_1 \oplus W_1 \to V_2 \oplus W_2$.

In the same way, one can form a **r-fold product category** $(\mathsf{C}_1, \dots, \mathsf{C}_r)$ of categories C_i, and a **covariant functor of r variables** is a covariant functor of the form $\mathsf{F}\colon (\mathsf{C}_1, \dots, \mathsf{C}_r) \to \mathsf{D}$. For example, there is a functor

$$(\mathsf{Vect}, \dots, \mathsf{Vect}) \to \mathsf{Vect}, \qquad (V_1, \dots, V_r) \to V_1 \oplus \cdots \oplus V_r.$$

A contravariant functor is defined in almost the same way, but it reverses the arrows.

Definition 14.35. Suppose C and D are two categories. A **contravariant functor** $\mathsf{G}\colon \mathsf{C} \to \mathsf{D}$ associates to each $A \in \operatorname{obj}(\mathsf{C})$ an object $\mathsf{G}(A) \in \operatorname{obj}(\mathsf{D})$, and to each morphism $A \xrightarrow{f} B$ in C a morphism $\mathsf{G}(B) \xrightarrow{\mathsf{G}(f)} \mathsf{G}(A)$ in D which satisfies the following two axioms:

1. If $A \xrightarrow{f} B \xrightarrow{g} C$ in C then $\mathsf{G}(C) \xrightarrow{\mathsf{G}(g)} \mathsf{G}(B) \xrightarrow{\mathsf{G}(f)} \mathsf{G}(A)$ in D and
$$\mathsf{G}(g \circ f) = \mathsf{G}(f) \circ \mathsf{G}(g).$$
2. $\mathsf{G}(\mathrm{id}_A) = \mathrm{id}_{\mathsf{G}(A)}$ for every $A \in \operatorname{obj}(\mathsf{C})$.

Here is a simple example of a contravariant functor on the category of vector spaces.

[4] I will use the direct sume notation here as well (instead of the product) so as not to confuse you with the product category.

Example 14.36. Let W be a fixed vector space. Then there is a contravariant functor

$$\mathrm{L}(\Box, W) \colon \mathsf{Vect} \to \mathsf{Vect}$$

that assigns to a vector space V the vector space $\mathrm{L}(V, W)$. If $T \colon V_1 \to V_2$ is a linear map then

$$\mathrm{L}(\Box, W)(T) \colon \mathrm{L}(V_2, W) \to \mathrm{L}(V_1, W)$$

is given by $S \mapsto S \circ T$ (note the order of V_1 and V_2!) It is important for you to understand why $\mathrm{L}(V, \Box)$ is covariant but $\mathrm{L}(\Box, W)$ is contravariant.

Taking $W = \mathbb{R}$ shows that $V \mapsto V^*$ is a contravariant functor.

(♣) Remark 14.37. Going back to algebraic topology, **singular cohomology** is a contravariant functor $\mathsf{hTop}^2 \to \mathsf{Ab}$. Later in this course we will look at **de Rham cohomology**.

Similarly one can consider contravariant functors of more than one variable. In fact, one can even consider functors that are covariant in some variables and contravariant in others. This is easiest to see with an example.

Example 14.38. Let $\mathrm{L}(\Box, \Box) \colon (\mathsf{Vect}, \mathsf{Vect}) \to \mathsf{Vect}$ denote the functor that sends a pair (V, W) to the vector space $\mathrm{L}(V, W)$. As Example 14.29 and 14.36 showed, this is contravariant in the first variable and covariant in the second variable. If $S \colon V_1 \to V_2$ and $T \colon W_1 \to W_2$ then

$$\mathrm{L}(\Box, \Box)(S, T) \colon \mathrm{L}(V_2, W_1) \to \mathrm{L}(V_1, W_2)$$

sends a linear map $A \colon V_2 \to W_1$ to the linear map $T \circ A \circ S \colon V_1 \to W_2$.

We have now almost arrived at the correct setting for which to prove the Metatheorem. The only thing that is left is to take into the account that we require our functors to be smooth.

Definition 14.39. Let $\mathsf{F} \colon \mathsf{Vect} \to \mathsf{Vect}$ be a covariant functor. We say that F is **smooth** if for any two vector spaces V, W, the map

$$\mathrm{L}(V, W) \to \mathrm{L}(\mathsf{F}(V), \mathsf{F}(W)), \qquad T \mapsto \mathsf{F}(T)$$

is itself smooth in the normal sense.

A similar definition makes sense for functors of r variables which are covariant in some variables and contravariant in others, provided one remembers to flip the domain and target in each contravariant variable:

Definition 14.40. Let $\mathsf{F} \colon (\mathsf{Vect}, \dots, \mathsf{Vect}) \to \mathsf{Vect}$ be a functor of r variables of either (or mixed) variance. We say that F is a **smooth functor** if for any vector spaces $V_1, \dots V_r$ and $W_1, \dots, W_r$, the induced map

$$\bigoplus_{i=1}^{r} \widetilde{\mathrm{L}}(V_i, W_i) \to \mathrm{L}(\mathsf{F}(V_1, \dots, V_r), \mathsf{F}(W_1, \dots W_r)), \qquad (T_1, \dots, T_r) \mapsto \mathsf{F}(T_1, \dots, T_r) \tag{14.3}$$

where

$$\widetilde{\mathrm{L}}(V_i, W_i) := \begin{cases} \mathrm{L}(V_i, W_i), & \text{if } \mathsf{F} \text{ is covariant in the } i\text{th variable,} \\ \mathrm{L}(W_i, V_i), & \text{if } \mathsf{F} \text{ is contravariant in the } i\text{th variable,} \end{cases}$$

is a smooth map in the usual sense (note again each side is simply a vector space).

In fact, in all the examples we have seen, the map (14.3) is actually a linear map (and so is certainly smooth). We emphasise though that for a general functor this may not be the case. Here now is a precise statement of the Metatheorem.

THEOREM 14.41. *Let* $\mathsf{F}\colon (\mathsf{Vect}, \dots, \mathsf{Vect}) \to \mathsf{Vect}$ *be a smooth functor of* r *variables of either variance in each variable. Let* $\pi_i\colon E_i \to M$ *be* r *vector bundles. Define*

$$\mathsf{F}(E_1, \dots, E_r) := \bigsqcup_{x \in M} \mathsf{F}(E_1|_x, \dots, E_r|_x),$$

with associated projection $\pi\colon \mathsf{F}(E_1, \dots, E_r) \to M$. *Then* $\mathsf{F}(E_1, \dots, E_r)$ *is a vector bundle.*

Warning: This proof is very easy, but it is notationally quite challenging. I recommend you write out for yourself the case $r = 2$ where F is say, contravariant in the first variable and covariant in the second (the $\mathrm{L}(\square, \square)$ functor from Example 14.38 is such an example). Once you understand this, the general case is just messier. In any case, the proof is non-examinable.

(♣) *Proof.* Choose an open set $U \subset M$ over which all the E_i are trivial, i.e. so that there exist vector bundle charts $\alpha_i\colon \pi_i^{-1}(U) \to \mathbb{R}^{k_i}$, where E_i has rank k_i. Then for each $x \in U$ and each i, we have a linear isomorphism $\alpha_i|_{E_i|_x}\colon E_i|_x \to \mathbb{R}^{k_i}$. Set

$$\tilde{\alpha}_{i,x} := \begin{cases} \alpha_i|_{E_i|_x}\colon E_i|_x \to \mathbb{R}^{k_i}, & \text{if } \mathsf{F} \text{ is covariant in the } i\text{th variable,} \\ \alpha_i|_{E_i|_x}^{-1}\colon \mathbb{R}^{k_i} \to E_i|_x, & \text{if } \mathsf{F} \text{ is contravariant in the } i\text{th variable.} \end{cases}$$

Since F is a functor, we can feed it the morphisms $\tilde{\alpha}_{i,x}$ to get a map

$$\tilde{\alpha}_x = \mathsf{F}(\tilde{\alpha}_{1,x}, \dots, \tilde{\alpha}_{r,x}) \in \mathrm{L}\left(\mathsf{F}(E_1|_x, \dots, E_r|_x), \mathsf{F}(\mathbb{R}^{k_1}, \dots, \mathbb{R}^{k_r})\right)$$

By functoriality, $\tilde{\alpha}_x$ is linear isomorphism. Define $\tilde{\alpha}\colon \pi^{-1}(U) \to \mathsf{F}(\mathbb{R}^{k_1}, \dots, \mathbb{R}^{k_r})$ by letting $\tilde{\alpha}$ be equal to $\tilde{\alpha}_x$ on $\mathsf{F}(E_1|_x, \dots, E_r|_x)$. We now declare that $(\pi, \tilde{\alpha})$ is a bundle chart for $\mathsf{F}(E_1, \dots, E_r)$ over U.

To complete the proof, we need to show that the transition functions are smooth linear isomorphisms. For this, suppose $\beta_i\colon \pi_i^{-1}(U) \to \mathbb{R}^{k_i}$ were different choices of vector bundle chart on each E_i, with corresponding chart $\tilde{\beta}$ on $\mathsf{F}(E_1, \dots, E_r)$. We must show that the transition function $\rho_{\tilde{\alpha}\tilde{\beta}}$ is smooth and linear. But this again follows almost immediately from functorality. If $x \in U$ then

$$\begin{aligned} \rho_{\tilde{\alpha}\tilde{\beta}}(x) &= \tilde{\alpha}|_{\mathsf{F}(E_1|_x, \dots, E_r|_x)} \circ \tilde{\beta}|^{-1}_{\mathsf{F}(E_1|_x, \dots, E_r|_x)} \\ &= \mathsf{F}(\tilde{\alpha}_{1,x}, \dots, \tilde{\alpha}_{r,x}) \circ \mathsf{F}(\tilde{\beta}_{1,x}, \dots, \tilde{\beta}_{r,x})^{-1} \\ &= \mathsf{F}(\tilde{\alpha}_{1,x}, \dots, \tilde{\alpha}_{r,x}) \circ \mathsf{F}(\tilde{\beta}^{-1}_{1,x}, \dots, \tilde{\beta}^{-1}_{r,x}) \\ &= \mathsf{F}\left(\tilde{\alpha}_{1,x} \circ \beta^{-1}_{1,x}, \dots \tilde{\alpha}_{r,x} \circ \beta^{-1}_{r,x}\right) \\ &= \mathsf{F}(\tilde{\rho}_{\alpha_1\beta_1}(x), \dots \tilde{\rho}_{\alpha_r\beta_r}(x)), \end{aligned}$$

where

$$\tilde{\rho}_{\alpha_i\beta_i}(x) \coloneqq \begin{cases} \rho_{\alpha_i\beta_i}(x) & \text{if } \mathsf{F} \text{ is covariant in the } i\text{th variable,} \\ \rho_{\alpha_i\beta_i}(x)^{-1} & \text{if } \mathsf{F} \text{ is contravariant in the } i\text{th variable.} \end{cases}$$

Thus since F is a functor, $\mathsf{F}(\tilde{\rho}_{\alpha_1\beta_1}(x), \dots \tilde{\rho}_{\alpha_r\beta_r}(x))$ is a linear isomorphism. Moreover since F is a smooth functor, $x \mapsto \mathsf{F}(\tilde{\rho}_{\alpha_1\beta_1}(x), \dots \tilde{\rho}_{\alpha_r\beta_r}(x))$ depends smoothly on x. This completes the proof. ■

LECTURE 15

Tensor and exterior algebras

In this lecture we continue our theme of constructing new vector bundles from old, but this time we focus on two constructions you may be less familiar with on the linear algebra level.

DEFINITION 15.1. Let V and W be two vector spaces. Their **tensor product** is the vector space $V \otimes W$ which is defined as follows. First, let $\mathrm{Free}(V \times W)$ denote (infinite-dimensional) vector space which has as basis all pairs (v, w) where $v \in V$ and $w \in W$. Thus an element of $\mathrm{Free}(V \times W)$ consists of a finite linear combination of pairs (v, w) with $v \in V$ and $w \in W$. Now let $R(V, W)$ denote the linear subspace of $\mathrm{Free}(V \times W)$ generated by the set of all elements of the form

$$\begin{cases} (v_1 + v_2, w) - (v_1, w) - (v_2, w), & v_1, v_2 \in V,\ w \in W, \\ (v, w_1 + w_2) - (v, w_1) - (v, w_2), & v \in V,\ w_1, w_2 \in W, \\ c(v, w) - (cv, w), & v \in V,\ w \in W,\ c \in \mathbb{R}, \\ c(v, w) - (v, cw), & v \in V,\ w \in W,\ c \in \mathbb{R}. \end{cases}$$

Let $V \otimes W$ denote the quotient vector space $\mathrm{Free}(V, W)/R(V, W)$. The coset in $V \otimes W$ containing (v, w) is denoted by $v \otimes w$. By construction one has

$$\begin{cases} (v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w, & v_1, v_2 \in V,\ w \in W, \\ v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2, & v \in V,\ w_1, w_2 \in W, \\ c(v \otimes w) = (cv) \otimes w, & v \in V,\ w \in W,\ c \in \mathbb{R}, \\ c(v \otimes w) = v \otimes (cw), & v \in V,\ w \in W,\ c \in \mathbb{R}. \end{cases}$$

A typical element in $V \otimes W$ is a finite sum $\sum_i a_i\, v_i \otimes w_i$ where the a_i are real numbers. An element of the form $v \otimes w$ is called **decomposable**.

There is a natural bilinear map $\otimes \colon V \times W \to V \otimes W$ that sends $(v, w) \mapsto v \otimes w$. Here is a useful property of the tensor product.

LEMMA 15.2. *Let V, W and U be vector spaces and suppose $B \colon V \times W \to U$ is a bilinear map. Then there exists a unique linear map $T \colon V \otimes W \to U$ such that the following diagram commutes:*

$$\begin{tikzcd} V \times W \arrow[rr, "\otimes"] \arrow[dr, "B"'] & & V \otimes W \arrow[dl, dashed, "T"] \\ & U & \end{tikzcd}$$

Moreover this property uniquely characterises $V \otimes W$.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Proof. Let $B\colon V\times W\to U$ be a bilinear function. Recall $V\otimes W = \operatorname{Free}(V,W)\big/R(V,W)$. We first extend B by linearity to a map $\tilde{B}\colon \operatorname{Free}(V,W)\to U$. Bilinearity then tells us that $R(V,W)\subset \ker\tilde{B}$, and hence $\tilde{B}$ factors to define a homomorphism $T\colon V\otimes W\to U$ such that $T(v\otimes w) = B(v,w)$ for all $(v,w)\in V\times W$. Moreover the map T is unique, since the decomposable elements generate $V\otimes W$.

Finally, to see why this property uniquely determines $\otimes$, suppose that X is another vector space equipped with a bilinear map $b\colon V\times W\to X$ with the property that if $B\colon V\times W\to U$ is bilinear then there exists a *unique* linear map $S\colon X\to U$ such that the diagram commutes:

$$\begin{tikzcd} V\times W \arrow[rr, "b"] \arrow[dr, "B"'] & & X \arrow[dl, dashed, "S"] \\ & U & \end{tikzcd}$$

We apply this with $U = V\otimes W$ and $B = \otimes$. This gives us a unique linear map $S\colon X\to V\otimes W$ such that the diagram commutes. Now we go back to our original diagram and chose $U = X$ and $B = b$. Thus we get a unique linear map $T\colon V\otimes W\to X$ such that the diagram commutes. The composition $T\circ S$ makes the following diagram commute:

$$\begin{tikzcd} V\times W \arrow[r, "b"] \arrow[d, "\otimes"'] & X \arrow[dl, dashed, "S"] \\ V\otimes W \arrow[dr, dashed, "T"] & \\ V\times W \arrow[u, "\otimes"] \arrow[r, "b"'] & X \end{tikzcd}$$

Thus the composition makes this diagram commute:

$$\begin{tikzcd} V\times W \arrow[rr, "b"] \arrow[dr, "b"'] & & X \arrow[dl, dashed, "T\circ S"] \\ & X & \end{tikzcd}$$

But there is meant to only be one map that makes this diagram commute, and another choice is the identity map $X\to X$. Thus $T\circ S = \mathrm{id}_X$. Similarly $S\circ T = \mathrm{id}_{V\otimes W}$, and we conclude that X and $V\otimes W$ are isomorphic, as claimed. ∎

(♣) Remark 15.3. In category-theoretic language, we just established that the tensor product could be defined by a *universal property*. Thus the last half of the proof (proving uniqueness) was formally unnecessary, since solutions to universal properties are *always* unique up to isomorphism (if they exist).

Corollary 15.4. *Let V and W denote vector spaces, and let $V^* = \mathrm{L}(V, \mathbb{R})$ denote the dual space. Then there is a natural isomorphism $\mathrm{L}(V, W) \cong V^* \otimes W$.*

Proof. Define $B\colon V^* \times W \to \mathrm{L}(V, W)$ by $B(\lambda, w)(v) := \lambda(v) \cdot w$. This gives us a linear map $T\colon V^* \otimes W \to \mathrm{L}(V, W)$ by Lemma 15.2. This map is an isomorphism, as an inverse $S\colon \mathrm{L}(V, W) \to V^* \otimes W$ is given by $S(L) := e^i \otimes Le_i$, where e_i is any basis of V and e^i is the dual basis of V^*. ■

Corollary 15.5. *If (e_i) is a basis for V and (e'_j) is a basis for W then $e_i \otimes e'_j$ is a basis for $V \otimes W$. Thus $\dim(V \otimes W) = \dim V \cdot \dim W$.*

Corollary 15.6. *If V, W and U are vector spaces then there are natural isomorphisms $V \otimes W \cong W \otimes V$ and $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$. (And thus we can unambiguously write $U \otimes V \otimes W$.)*

The proof of Corollary 15.6 is on Problem Sheet H. Going back to our category-theoretic point of view, we obtain another functor:

Example 15.7. There is a covariant functor $\otimes\colon (\mathsf{Vect}, \mathsf{Vect}) \to \mathsf{Vect}$ given by $(V, W) \mapsto V \otimes W$.

Definition 15.8. Let r and s be non-negative integers. Define a functor $T^{r,s}\colon \mathsf{Vect} \to \mathsf{Vect}$ by setting

$$T^{r,s}(V) := \underbrace{V \otimes \cdots \otimes V}_{r} \otimes \overbrace{V^* \otimes \cdots \otimes V^*}^{s}.$$

One calls an element of $T^{r,s}(V)$ a **tensor of type** (r, s). The vector space $T^{r,s}(V)$ has dimension $(\dim V)^{r+s}$.

Note we are using Corollary 15.6 to write the right-hand side without brackets. Corollary 15.6 also shows us that it is unimportant in which order we present the factors: for convenience we write the V factors first and the V^* factors afterwards.

Let $\mathrm{Mult}_{r,s}(V)$ denote the space of multilinear maps

$$\underbrace{V \times \cdots \times V}_{r} \times \overbrace{V^* \times \cdots \times V^*}^{s} \to \mathbb{R}.$$

Proposition 15.9. *There is a canonical isomorphism between the vector space $T^{r,s}(V)$ and the vector space $\mathrm{Mult}_{s,r}(V)$.*

Note the r and the s swapped round—this is *not* a typo! Recall that a **perfect pairing** of a vector space V with another W is a bilinear map $\beta\colon V \times W \to \mathbb{R}$ such that $\beta(v, \cdot)$ is identically zero if and only if $v = 0$, and $\beta(\cdot, w)$ is identically zero if and only if w is zero. If V and W are finite-dimensional then such a pairing induces an isomorphism $T\colon V \to W^*$ given by $(Tv)(w) := \beta(v, w)$.

Example 15.10. The natural isomorphism[1] $V \cong V^{**}$ arises from the perfect pairing $V \times V^* \to \mathbb{R}$ given by $(v, p) \mapsto p(v)$.

[1]Remember we assume all vector spaces are finite-dimensional!

Proof of Proposition 15.9. We define a perfect pairing of $T^{r,s}(V)$ with $T^{r,s}(V^*)$. Namely, if

$$v = v_1 \otimes \cdots \otimes v_r \otimes p^1 \otimes \cdots \otimes p^s \in T^{r,s}(V)$$

and

$$w = q^1 \otimes \cdots \otimes q^r \otimes w_1 \otimes \cdots \otimes w_s \in T^{r,s}(V^*)$$

then we can naturally feed them each other[2]

$$\beta(v, w) := \prod_{i=1}^{r} q^i(v_i) \cdot \prod_{j=1}^{s} p^j(w_j) \tag{15.1}$$

Now extend this bilinearly to all elements. Thus $T^{r,s}(V)$ is isomorphic to $\big(T^{r,s}(V^*)\big)^*$. Next, a trivial extension of Lemma 15.2 shows that

$$\mathrm{Mult}_{r,s}(V^*) \cong \big(T^{r,s}(V^*)\big)^* \tag{15.2}$$

Indeed, for each $B \in \mathrm{Mult}_{r,s}(V^*)$ there exists a unique $T \colon T^{r,s}(V^*) \to \mathbb{R}$, i.e. $T \in \big(T^{r,s}(V^*)\big)^*$ such that the diagram commutes.

$$\begin{tikzcd}
\underbrace{V^* \times \cdots \times V^*}_{r} \times \overbrace{V \times \cdots \times V}^{s} \arrow[rr, "\otimes"] \arrow[dr, "B"'] & & \underbrace{V^* \otimes \cdots \otimes V^*}_{r} \otimes \overbrace{V \otimes \cdots \otimes V}^{s} \arrow[dl, dashed, "T"] \\
& \mathbb{R} &
\end{tikzcd}$$

Since B is uniquely determined by T, this sets up the desired isomorphism (15.2). Finally we clearly have $\mathrm{Mult}_{r,s}(V^*) \cong \mathrm{Mult}_{s,r}(V)$, and thus the proof is complete. ■

Remark 15.11. On decomposable elements the isomorphism $T^{r,s}(V) \cong \mathrm{Mult}_{s,r}(V)$ is easier to describe. Suppose for simplicity $A \in T^{2,3}(V)$ is the decomposable element

$$A = v_1 \otimes v_2 \otimes p^1 \otimes p^2 \otimes p^3.$$

Then if we use Proposition 15.9 to regard A as an element of $\mathrm{Mult}_{3,2}(V)$ then A is given explicitly as

$$A(w_1, w_2, w_3, q^1, q^2) = q^1(v_1)\, q^2(v_2)\, p^1(w_1)\, p^2(w_2)\, p^3(w_3).$$

We can use Theorem 14.41 ("the Metatheorem") to transfer these linear algebra constructions to vector bundles:

Corollary 15.12. *Let $\pi_i \colon E_i \to M$ be vector bundles for $i = 1, 2$ of rank k_i. Then there is a vector bundle $E_1 \otimes E_2 \to M$ of rank $k_1 k_2$ whose fibre over x is $E_1|_x \otimes E_2|_x$.*

Recall the Hom-bundle from (14.2) in the last lecture.

---

[2] We are *not* using the Einstein Summation Convention in this formula—this is a product not a sum!

Corollary 15.13. *Let $\pi_i \colon E_i \to M$ be vector bundles for $i = 1, 2$. Then there is a natural vector bundle isomorphism*

$$\mathrm{Hom}(E_1, E_2) \cong E_1^* \otimes E_2.$$

Proof. Define $\Psi \colon \mathrm{Hom}(E_1, E_2) \to E_1^* \otimes E_2$ fibrewise by declaring that the map $\Psi_x \colon \mathrm{L}(E_1|_x, E_2|_x) \to E_1^*|_x \otimes E_2|_x$ is the isomorphism from Corollary 15.4. This assignment $x \mapsto \Psi_x$ is smooth and thus Ψ is a vector bundle isomorphism. (If you are worried about why $x \mapsto \Psi_x$ is smooth, you could use Proposition 16.26 from the next lecture). ∎

Similarly Corollary 15.6 tells us that the tensor product of vector bundles is commutative and associative:

Corollary 15.14. *Let $\pi_i \colon E_i \to M$ be vector bundles for $i = 1, 2, 3$. Then the bundles $E_1 \otimes E_2$ and $E_2 \otimes E_1$ are isomorphic, and the bundles $E_1 \otimes (E_2 \otimes E_3)$ and $(E_1 \otimes E_2) \otimes E_3$ are isomorphic.*

(♣) Remark 15.15. Corollary 15.13 and Corollary 15.14 are special cases of a more general result, which goes as follows: Let F_1 and F_2 be two functors as in the statement of Theorem 14.41. Assume there exists a *smooth natural isomorphism* $\tau \colon \mathsf{F}_1 \to \mathsf{F}_2$. Then the vector bundles obtained by applying Theorem 14.41 to F_1 and F_2 are naturally isomorphic. The proof is very similar to that of Theorem 14.41 and is not any harder. However since I have not defined precisely what a natural transformation is, I will not go into the details. If you are interested, here is where I defined natural transformations last year in Algebraic Topology I.

Corollary 15.16. *Let $\pi \colon E \to M$ be a vector bundle. Then there is a vector bundle $T^{r,s}(E) \to M$ whose fibre over $x \in M$ is the vector space $T^{r,s}(E_x)$.*

Let us recall the formal definition of an algebra.

Definition 15.17. A vector space V is said to be an **algebra** if there exists a bilinear map $V \times V \to V$ (or equivalently, a linear map $V \otimes V \to V$), which we call **multiplication**.

Whilst each $T^{r,s}(V)$ is not an algebra, if we sum them all together we obtain one.

Definition 15.18. The **tensor algebra** of V is defined to be

$$T(V) := \bigoplus_{r,s \geq 0} T^{r,s}(V),$$

where $T^{0,0}(V) := \mathbb{R}$. This is a *graded algebra*, in the sense that $\otimes$ gives a natural map

$$\otimes \colon T^{r,s}(V) \times T^{r_1,s_1}(V) \to T^{r+r_1,s+s_1}(V). \tag{15.3}$$

The natural map is defined as one would guess: on decomposable elements it simply tensors everything together and then rearranges the factors so the V elements come first, so as to fit with our convention. We illustrate this with $(r, s) = (1, 2)$ and $(r_1, s_1) = (2, 1)$:

$$\Big((v_1 \otimes p^1 \otimes p^2), (w_1 \otimes w_2 \otimes q^1)\Big) \mapsto v_1 \otimes w_1 \otimes w_2 \otimes p^1 \otimes p^2 \otimes q^1. \tag{15.4}$$

If $(r, s) = (0, 0)$ then tensor multiplication with a scalar is defined to be normal scalar multiplication, i.e.

$$r \otimes v := rv, \qquad r \in \mathbb{R},\ v \in V. \tag{15.5}$$

Remark 15.19. **Warning:** The space $T(V)$ is an infinite-dimensional vector space. This means that T is *not* a functor $\mathsf{Vect} \to \mathsf{Vect}$! Thus if $\pi \colon E \to M$ is a vector bundle, whilst for any finite r, s we can speak of the bundle $T^{r,s}(E) \to M$, we cannot apply[3] Theorem 14.41 to obtain a vector bundle $T(E) \to M$.

We now define another linear algebra construction, called the **exterior algebra**. This will associate to a vector space V another (finite-dimensional) vector space $\bigwedge(V)$ which, like the tensor algebra $T(V)$, admits an algebra structure. This defines a functor $\bigwedge \colon \mathsf{Vect} \to \mathsf{Vect}$, and thus by Theorem 14.41 we can apply it to vector bundles.

Let V be a vector space. Let $T^+(V)$ denote the subalgebra given by $T^+(V) := \bigoplus_{r \geq 0} T^{r,0}(V)$. Let $I(V)$ denote the two-sided ideal in $T^+(V)$ generated by all elements of the form $v \otimes v$ for $v \in V$. Thus for instance $u \otimes v \otimes v \otimes w$ belongs to $I(V)$.

Definition 15.20. The **exterior algebra** is defined to be the quotient algebra $\bigwedge(V) := T^+(V)/I(V)$. We denote the image of $v_1 \otimes \cdots \otimes v_r$ in $\bigwedge(V)$ by $v_1 \wedge \cdots \wedge v_r$ and call $\wedge$ the **wedge product**.

Such an element $v_1 \wedge \cdots \wedge v_r$ is called **decomposable**. If we set $\bigwedge^r(V)$ to be the image of $T^{r,0}(V)$ in $\bigwedge(V)$ under the projection $T^+(V) \to \bigwedge(V)$ there is a canonical vector space isomorphism

$$\bigwedge^r(V) \cong T^{r,0}(V) \Big/ I^r(V),$$

where $I^r(V) := T^{r,0}(V) \cap I(V)$. Note that $\bigwedge^1(V) = V$ and $\bigwedge^0(V) = \mathbb{R}$. This definition may seem a little abstract, so let us unpack things a bit.

Proposition 15.21 (Properties of the wedge product). *Let V be a vector space. Then*

(i) *For all $v, w \in V$, $v \wedge w = -w \wedge v$.*

(ii) *Assume $r, s > 0$. If $v \in \bigwedge^r(V)$ and $w \in \bigwedge^s(V)$ then $v \wedge w \in \bigwedge^{r+s}(V)$ and*

$$v \wedge w = (-1)^{rs} w \wedge v, \tag{15.6}$$

This continues to hold if either $r = 0$ or $s = 0$ if we use the convention that for a real number a and a vector v, one has $a \wedge v := av$.

[3] Actually, there is an infinite-dimensional analogue of Theorem 14.41, and using this it *is* possible to define $T(E)$, but this goes beyond the scope of the course and we won't need it.

(iii) *If $v_1 \wedge \cdots \wedge v_r \in \bigwedge^r(V)$ is a decomposable element then transposing v_i with v_j acts as multiplication by -1:*

$$v_1 \wedge \cdots \wedge v_i \wedge \cdots \wedge v_j \wedge \cdots \wedge v_r = -v_1 \wedge \cdots \wedge v_j \wedge \cdots \wedge v_i \wedge \cdots \wedge v_r$$

(iv) *If $\varrho \in \mathfrak{S}(r)$ is a permutation on r letters and $v_i \in V$ then*

$$v_{\varrho(1)} \wedge \cdots \wedge v_{\varrho(r)} = \operatorname{sgn}(\varrho) v_1 \wedge \cdots \wedge v_r.$$

Proof. To prove part (i), we note that for any $u \in V$, $u \otimes u$ belongs to $I(V)$, and thus in $\bigwedge(V)$, $u \wedge u = 0$. Applying this with $u = v + w$ we have

$$\begin{aligned} 0 &= (v+w) \wedge (v+w) \\ &= v \wedge v + v \wedge w + w \wedge v + w \wedge w \\ &= v \wedge w + w \wedge v. \end{aligned}$$

To prove part (ii), as both sides are linear in v and w, it suffices to verify it for decomposable elements, and for such, the conclusion follows by repeated applications of part (i). Next, to prove part (iii), we may assume $i < j$. Set $u \coloneqq v_{i+1} \wedge \cdots \wedge v_{j-1}$. Then by part (ii) one has

$$v_i \wedge u \wedge v_j = -v_j \wedge u \wedge v_i,$$

and thus part (iii) follows. Finally, part (iv) is immediate from the fact that any permutation may be written as a product of transpositions. ■

There is an analogous universal mapping property for the exterior algebra.

Definition 15.22. Let V and W be vector spaces. Let $\mathrm{Alt}_r(V, W)$ denote the space of **alternating r-linear maps**, i.e. multilinear maps $A \colon V \times \cdots \times V \to W$ (r times) that vanish whenever any two of the arguments are equal:

$$A(\cdots, v, \cdots, v, \cdots) = 0.$$

We abbreviate $\mathrm{Alt}_r(V) = \mathrm{Alt}_r(V, \mathbb{R})$.

The map $\wedge \colon V \times \cdots \times V \to \bigwedge^r(V)$ given by sending $(v_1, \dots, v_r) \mapsto v_1 \wedge \cdots \wedge v_r$ is an example of such a map. We aim to prove the following alternating version of Proposition 15.9:

Proposition 15.23. *There is a canonical isomorphism between $\bigwedge^r(V^*)$ and $\mathrm{Alt}_r(V)$.*

The proof strategy is similar to that of Proposition 15.9, and we will be brief. First, we need an analogue of the universal mapping property (Lemma 15.2).

Lemma 15.24. *Let V and W be vector spaces. For any $A \in \mathrm{Alt}_r(V, W)$ there is a unique linear map $T \colon \bigwedge^r(V) \to W$ such that the following diagram commutes:*

$$\begin{tikzcd} \overbrace{V \times \cdots \times V}^{r} \arrow[rr, "\wedge"] \arrow[dr, "A"'] & & \bigwedge^r(V) \arrow[dl, dashed, "T"] \\ & W & \end{tikzcd}$$

Moreover $\bigwedge^r(V)$ is uniquely characterised by this property.

Corollary 15.25. *Let V be a vector space of dimension k with basis $\{e_1, \dots, e_k\}$. Then*
$$\{e_{i_1} \wedge \cdots \wedge e_{i_r} \mid 1 \le i_1 < \cdots < i_r \le k\}$$
is a basis of $\bigwedge^r(V)$ and $\bigwedge^r(V) = 0$ for $r > k$. Thus $\dim \bigwedge^r(V) = \binom{k}{r}$ and $\dim \bigwedge(V) = 2^k$.

The proofs of Lemma 15.24 and Corollary 15.25 are on Problem Sheet H. Just as with (15.2), it follows that we can identify

$$\mathrm{Alt}_r(V) \cong \left(\bigwedge^r(V)\right)^*. \tag{15.7}$$

The next step is to exhibit a perfect pairing of $\bigwedge^r(V^*)$ with $\bigwedge^r(V)$. This formula is a little harder to guess than in (15.1), but once you know the formula it is easy to check. Namely, we define

$$\alpha\colon \bigwedge^r(V^*) \times \bigwedge^r(V) \to \mathbb{R}$$

by declaring on decomposable elements that:

$$\alpha\left((p^1 \wedge \cdots \wedge p^r), (v_1 \wedge \cdots \wedge v_r)\right) := \det A \tag{15.8}$$

where A is the $r \times r$ matrix whose (i,j)th entry is $p^i(v_j)$. Then extend this by bilinearity to all elements. I will leave it to you to verify this is indeed a perfect pairing.

We end today's lecture by applying Theorem 14.41 to the functors $\bigwedge^r\colon \mathsf{Vect} \to \mathsf{Vect}$ and $\bigwedge\colon \mathsf{Vect} \to \mathsf{Vect}$.

Corollary 15.26. *Let $\pi\colon E \to M$ be a vector bundle of rank k. Then for any $0 \le r \le k$, there is a vector bundle $\bigwedge^r(E) \to M$ whose fibre over $x \in M$ is given by $\bigwedge^r(E_x)$. It has rank $\binom{k}{r}$. Similarly there is a vector bundle $\bigwedge(E) \to M$ of rank 2^k whose fibre over $x \in M$ is given by $\bigwedge(E_x)$. It is the direct sum of the vector bundles $\bigwedge^r(E)$.*

Remark 15.27. Let V be k-dimensional vector space which is also an algebra in the sense of Definition 15.17. Let $\pi\colon E \to M$ be a vector bundle of rank k. We say that E is an **algebra bundle** if each fibre E_x admits the structure of an algebra, and there exists a bundle atlas of charts $\alpha\colon \pi^{-1}(U) \to V$ such that for each $x \in U$ the map $\alpha|_{E_x}\colon E_x \to V$ is not only a linear isomorphism but also an algebra isomorphism. Thus the exterior algebra bundle $\bigwedge(E) \to M$ is an example of an algebra bundle.

Sections of vector bundles

A fibre bundle $\pi\colon E \to M$ is a surjective submersion between manifolds with the property that the domain E has extra structure. Smooth maps that go in the opposite direction are—from the point of view of fibre bundles—uninteresting unless they respect this extra structure.

DEFINITION 16.1. Let $\pi\colon E \to M$ be a fibre bundle. A **section** of E is a smooth map $s\colon M \to E$ such that $\pi \circ s = \mathrm{id}$, that is, a smooth map $s\colon M \to E$ such that

$$s(x) \in E_x, \qquad \forall x \in M. \tag{16.1}$$

The set of all sections is denoted by $\Gamma(E)$. A **local section** of E is a section of the bundle $\pi^{-1}(U) \to U$ of E over an open set $U \subset M$. We denote by $\Gamma(U, E)$ the set of all local sections with domain U.

EXAMPLE 16.2. Here are some examples of sections:

(i) Let M be a manifold. A vector field X on M is a section of the tangent bundle. Thus
$$\mathfrak{X}(M) = \Gamma(TM).$$
Similarly a vector field X defined on an open subset of M is a local section:
$$\mathfrak{X}(U) = \Gamma(U, TM).$$
In particular, if $\sigma\colon U \to O$ is a chart on M with local coordinates x^i then $\frac{\partial}{\partial x^i}$ is an element of $\Gamma(U, TM)$.

(ii) In a similar vein, if $f \in C^\infty(M)$ then in Example 4.12 we defined a section df of T^*M. If $f \in C^\infty(U)$ then $df \in \Gamma(U, T^*M)$.

(iii) A section of the trivial fibre bundle $M \times F \to M$ is the same thing as a smooth map $M \to F$. Thus for instance, a section of $M \times \mathbb{R} \to M$ is just a smooth function on M.

(iv) Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ denote two vector bundles, and let $\mathrm{Hom}(E_1, E_2) \cong E_1^* \otimes E_2 \to M$ denote the bundle obtained by applying Theorem 14.41 to the functor $\mathrm{L}(\Box, \Box)$ (cf. (14.2)). A section $\Phi \in \Gamma(\mathrm{Hom}(E_1, E_2))$ is a smooth map $x \mapsto \Phi_x$ where $\Phi_x\colon E_1|_x \to E_2|_x$ is a linear map. Thus:
$$\Gamma(\mathrm{Hom}(E_1, E_2)) = \{\text{vector bundle homomorphisms } \Phi\colon E_1 \to E_2\}.$$

(v) If G is a Lie group and $H \subset G$ is a closed subgroup then the projection map $\pi\colon G \to G/H$ is a fibre bundle with fibre H (see Example 13.12). We proved in Corollary 12.5 that π admitted local sections around every point, and exploited this fact repeatedly in Theorem 12.11.

Lemma 16.3. *Let $\pi\colon E \to M$ be a fibre bundle and let $s \in \Gamma(U, E)$. Then $s(U)$ is an embedded submanifold of E of dimension equal to the dimension of M.*

Proof. If σ is a chart on $V \subset U$ then $\sigma \circ \pi$ is a chart on $s(V)$. ∎

Example 16.4. Let $\pi\colon E \to M$ be a vector bundle. The **zero section** $o\colon M \to E$ assigns to each $x \in M$ the zero vector in E_x. This allows us to see $M \cong o(M) \subset E$ as an embedded submanifold of E.

The space of sections of a vector bundle has extra structure not present in normal fibre bundles. We already saw this for vector fields in Lecture 7, but let us go over it again here[1].

Lemma 16.5. *Let $\pi\colon E \to M$ be a vector bundle. Then for any non-empty open set $U \subset M$, the set $\Gamma(U, E)$ is a vector space and a module over the ring $C^\infty(U)$.*

Proof. Suppose $s \in \Gamma(U, E)$. Let $\sigma\colon V \to O$ be a chart on $V \subset U$ and let α be a vector bundle chart on E defined on $\pi^{-1}(V)$. Then as in Remark 13.7, we may take $(\sigma \circ \pi, \alpha)$ as a chart on E. The assumption that s is smooth means that the composition

$$(\sigma \circ \pi, \alpha) \circ s \circ \sigma^{-1}\colon O \to O \times \mathbb{R}^k$$

is smooth (here k is the rank of E as a bundle). Moreover the section property tells us that this local map is of the form

$$(\sigma \circ \pi, \alpha) \circ s \circ \sigma^{-1} = (\mathrm{id}, \tilde{s}) \tag{16.2}$$

where $\tilde{s}\colon O \to \mathbb{R}^k$ is some smooth map. Just as in the proof of Proposition 7.2, this argument reverses, and we see that a map s satisfying the section property is smooth if and only if each local map $\tilde{s}$ is smooth.

With this in hand the lemma is trivial: if s and t are two sections and $a \in \mathbb{R}$ then $x \mapsto a\, s(x) + t(x)$ certainly satisfies the section property (16.1), and its local expression is given by $a\tilde{s} + \tilde{t}$ which is smooth if $\tilde{s}$ and $\tilde{t}$ are. Moreover if $f \in C^\infty(U)$ then we define

$$(fs)(x) := f(x)\, s(x), \qquad x \in U.$$

The local expression of fs is $\tilde{f}\tilde{s}$ where $\tilde{f} = f \circ \sigma^{-1}$. ∎

Definition 16.6. Let $\pi\colon E \to M$ be a vector bundle of rank k and let $U \subset M$ be open. A **local frame** for E over U is a collection $\{e_1, \dots, e_k\}$ of sections $e_i \in \Gamma(U, E)$ such that $\{e_1(x), \dots e_k(x)\}$ form a basis of the vector space E_x for each $x \in U$.

Local frames always exist for vector bundles: if $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ is a vector bundle chart then if we define

$$e_i(x) := \alpha|_{E_x}^{-1}(e_i), \tag{16.3}$$

where e_i is the standard basis vector in $\mathbb{R}^k$, then e_i is smooth (use the argument from the proof of Lemma 16.5) and the collection $\{e_i(x)\}$ is a basis of E_x since

[1] In fact, many of the arguments in this lecture will feel reminiscent of arguments from Lecture 3 and 7—the reason for this will become clear next lecture, when we unify things using sheaves.

$\alpha|_{E_x}$ is a linear isomorphism. Conversely, the existence of a local frame over U determines a vector bundle chart $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$. Indeed, if such a frame exists then every point $p \in \pi^{-1}(U)$ can be written as a linear combination $p = a^i\, e_i(x)$. We define

$$\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k, \qquad p \mapsto (a^1, \dots, a^k). \tag{16.4}$$

A **global frame** of a vector bundle is a frame defined on $U = M$. The next lemma is the vector bundle version of Problem F.4.

Corollary 16.7. *A vector bundle $\pi\colon E \to M$ admits a global frame if and only if it is trivial.*

Remark 16.8. On Euclidean spaces, we typically used $x = (x^1, \dots, x^n)$ for a point. Then on manifolds we continued to use the notation x^i, only instead of points these were now local functions (see Remark 2.7). The same is now true on vector bundles: e_i no longer denotes a single vector in $\mathbb{R}^k$, it is now a local section of a vector bundle. Once again, the idea is to make the notation "look" as similar as possible to standard differential calculus on Euclidean spaces.

Remark 16.9. If $\{e_1, \dots, e_k\}$ is a local frame for E over U then any map $s\colon U \to E$ satisfying the section property (16.1) can be written as

$$s = a^i\, e_i, \qquad \text{for some functions } a^i\colon U \to \mathbb{R}.$$

If we take the vector bundle chart α on E from (16.4) associated to the local frame $\{e_1, \dots, e_k\}$ then for any chart σ on M with appropriate domain, the function $\tilde{s}$ associated to s from (16.2) is given by

$$\tilde{s}(z) = \big(a^1(\sigma^{-1}(z)), \dots, a^k(\sigma^{-1}(z))\big).$$

This tells us that s is smooth (and hence belongs to $\Gamma(U, E)$) if and only if the functions a^i are smooth functions on U.

Definition 16.10. A local frame $\{e_1, \dots, e_k\}$ of E over U determines a local frame $\{\varepsilon^1, \dots, \varepsilon^k\}$ of the dual bundle E^* over U by requiring that

$$\varepsilon^i(x)(e_j(x)) = \delta^i_j, \qquad \text{for all } x \in U.$$

(Exercise: Why is this smooth?) One calls $\{\varepsilon^1, \dots, \varepsilon^k\}$ the **dual frame**.

Remark 16.11. If $s \in \Gamma(U, E)$ then if we write $s = a^i\, e_i$ for smooth functions a^i as per Remark 16.9 then observe that

$$\varepsilon^i(x)(s(x)) = a^i(x).$$

Similarly if $\omega \in \Gamma(U, E^*)$ is any section of the dual bundle then we can write $\omega = b_i\, \varepsilon^i$ where the function $b_i \in C^\infty(U)$ is given by

$$b_i(x) = \omega(x)(e_i(x)).$$

EXAMPLE 16.12. Let M be a smooth manifold of dimension n and let $\sigma \colon U \to O$ be a chart on M with local coordinates x^i. Then

$$\left\{ \frac{\partial}{\partial x^i} \mid i = 1, \dots, n \right\}$$

is a local frame for TM over U. Similarly

$$\{ dx^i \mid i = 1, \dots, n \}$$

is a local frame for T^*M over U. This is the dual frame. Taking this one step further,

$$\left\{ \frac{\partial}{\partial x^i} \otimes dx^j \mid 1 \le i, j \le n \right\}$$

is a local frame for $TM \otimes T^*M$ over U.

In general, a section of a vector bundle is often (although not always) called a **field** on M, where the "type" of field depends on the vector bundle. Thus for instance, the tangent bundle consists of vectors; thus a section of the tangent bundle is called a vector field.

DEFINITION 16.13. A **tensor field** of type (r, s) on M is a section of $T^{r,s}(TM)$. We normally use the special notation $\mathcal{T}^{r,s}(M)$ for tensor fields. The space of sections $\mathcal{T}^{r,s}(U) := \Gamma(U, T^{r,s}(TM))$ is defined similarly; these are the tensor fields of type (r, s) over U. We already briefly met these in Remark 7.18, and we will study these in more depth in Lecture 18. Let us unpack this a bit. The bundle $T^{r,s}(TM)$ is the bundle whose fibre over $x \in M$ is

$$T^{r,s}(T_xM) := \underbrace{T_xM \otimes \cdots \otimes T_xM}_{r} \otimes \overbrace{T_x^*M \otimes \cdots \otimes T_x^*M}^{s}.$$

If $A \in \mathcal{T}^{r,s}(M)$ then we can think of the value of A at x, which we write either as $A(x)$ or A_x (the latter is preferred if there are many variables) as a multilinear map

$$A_x \colon \underbrace{T_x^*M \times \cdots \times T_x^*M}_{r} \times \overbrace{T_xM \times \cdots \times T_xM}^{s} \to \mathbb{R}$$

thanks to Proposition 15.9. A tensor field of type $(1, 0)$ is just a vector field: in this case we think of $X(x) \colon T_x^*M \to \mathbb{R}$ as the linear map given by $X(x)(p) := p(X(x))$. If $A \in \mathcal{T}^{r,s}(M)$ then and $\sigma \colon U \to O$ is a chart on M then locally we can write[2]

$$A = A^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s} \tag{16.5}$$

where the function $A^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(U)$ is defined by

$$A^{i_1 \dots i_r}_{j_1 \dots j_s}(x) = A_x \left(dx^{i_1}|_x, \dots, dx^{i_r}|_x, \left.\frac{\partial}{\partial x^{j_1}}\right|_x, \dots, \left.\frac{\partial}{\partial x^{j_s}}\right|_x \right).$$

[2]Expressions of this form are the main reason we introduced the Einstein Summation Convention!

An even more important example is the notion of a *differential form*, which is a section of the exterior algebra bundle of the cotangent bundle.

Definition 16.14. A **differential r-form** (often simply called "an r-form") on M is a section of $\bigwedge^r(T^*M)$. We use the special notation $\Omega^r(M)$ for the space of differential r-forms. If $\omega \in \Omega^r(M)$ and $x \in M$ then we can think of $\omega(x) = \omega_x$ as an alternating map

$$\omega_x \colon \underbrace{T_xM \times \cdots \times T_xM}_{r} \to \mathbb{R},$$

thanks to Proposition 15.23.

We will come back to tensor fields and differential forms in Lectures 18 and 19. They are particularly important since they are the objects that can be *integrated* on manifolds.

We now investigate how vector bundle homomorphisms act on sections. The main result (Theorem 16.30), which we call the "Hom Γ Theorem[3]", gives an alternative way to define a vector bundle homomorphism. In the next lecture we will study this more abstractly; this will lead us to the relation between vector bundles and *locally free sheaves*.

Definition 16.15. Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ denote two vector bundles over the same manifold M. Let $\Phi \colon E_1 \to E_2$ denote a vector bundle homomorphism. We define a map

$$\Phi_\star \colon \Gamma(E_1) \to \Gamma(E_2), \qquad s \mapsto \Phi \circ s$$

The map $\Phi_\star$ is clearly a linear map between the two vector spaces $\Gamma(E_1)$ and $\Gamma(E_2)$. Moreover a moment's thought shows that $\Phi_\star$ is actually a *module homomorphism*, i.e. it is linear over $C^\infty(M)$. Indeed, if $f \in C^\infty(M)$, $s \in \Gamma(E_1)$, and $x \in M$ then

$$\begin{aligned}
\Phi_\star(fs)(x) &= \Phi \circ (fs)(x) \\
&= \Phi|_{E_x}(f(x)s(x)) \\
&\overset{(\dagger)}{=} f(x)\Phi|_{E_x}(s(x)) \\
&= (f\Phi_\star(s))(x)
\end{aligned}$$

where $(\dagger)$ used the fact that $\Phi|_{E_x}$ is a linear map. The Hom-Γ Theorem tells us that the converse holds: any $C^\infty(M)$-linear map $\Gamma(E_1) \to \Gamma(E_2)$ is induced by a vector bundle homomorphism. This will require some preparation though, and we begin with the following easy application of the Cutoff Functions Lemma 3.2.

Lemma 16.16. *Let $\pi \colon E \to M$ be a vector bundle and let $s \in \Gamma(U, E)$. Fix $x \in U$. Then there exists a global section $\hat{s} \in \Gamma(E)$ such that $\hat{s}$ agrees with s on a neighbourhood of x.*

[3] Pronounced: the Hom-Gamma Theorem. NB: I made this name up...

Proof. Choose a neighbourhood V of x with $\overline{V} \subset U$. Choose a cutoff function $\eta\colon M \to \mathbb{R}$ such that $\eta(y) = 1$ for all $y \in V$ and such that $\operatorname{supp}(\eta) \subset U$. Define $\hat{s}\colon M \to E$ by

$$\hat{s}(y) := \begin{cases} \eta(y)s(y), & y \in U, \\ 0, & y \in M \setminus U. \end{cases}$$

Then $\hat{s}$ is smooth[4] and agrees with s on the neighbourhood V of x. ■

Definition 16.17. Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ be two vector bundles over the same manifold M. Suppose $\chi\colon \Gamma(E_1) \to \Gamma(E_2)$ is an $\mathbb{R}$-linear operator. We say that χ is a **local operator** if whenever a section $s \in \Gamma(E_1)$ vanishes on an open set $U \subset M$, $\chi(s) \in \Gamma(E_2)$ also vanishes on U. We call χ a **point operator** if whenever a section $s \in \Gamma(E_1)$ vanishes at a point x, $\chi(s)$ also vanishes at x.

Any point operator is clearly a local operator, but the converse is not true.

Example 16.18. By part (iii) of Example 16.2, the space $C^\infty(\mathbb{R})$ can be identified with the space of all sections of the trivial bundle $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. The operator

$$C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \qquad f \mapsto f'$$

is a local operator (since if f is constant on an open set its derivative is also constant on that open set) but it is not a point operator.

More generally:

Example 16.19. Let M be a smooth manifold, and let $X \in \mathfrak{X}(M)$ denote a vector field. Regard X as a derivation of $C^\infty(M)$ as in Proposition 7.7, or equivalently, as a linear operator $\Gamma(M \times \mathbb{R}) \to \Gamma(M \times \mathbb{R})$ (as in part (iii) of Example 16.1). Then $f \mapsto X(f)$ is a local operator by Corollary 3.5, but not a point operator.

Definition 16.20. Let $\pi\colon E \to M$ be a vector bundle. An operator $\chi\colon \Gamma(E) \to \Gamma(E)$ is said to satisfy the **Leibniz rule** if there exists a vector field X on M such that for any $f \in C^\infty(M)$ and $s \in \Gamma(E)$ one has

$$\chi(fs) = (Xf)s + f\chi(s).$$

More generally still, we have the following result, whose proof is on Problem Sheet I.

Proposition 16.21. *If χ satisfies the Leibniz rule then χ is a local operator. If $X \not\equiv 0$ then χ is not a point operator.*

We will eventually show that every $C^\infty(M)$-linear map $\chi\colon \Gamma(E_1) \to \Gamma(E_2)$ is a point operator.

Proposition 16.22. *Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ be two vector bundles and suppose $\chi\colon \Gamma(E_1) \to \Gamma(E_2)$ is a local operator. Then for each open set $U \subset M$, there is a unique linear map $\chi_U\colon \Gamma(U, E_1) \to \Gamma(U, E_2)$, called the* **restriction** *of χ to U, such that for any global section s, one has*

$$\chi_U(s|_U) = \chi(s)|_U. \tag{16.6}$$

[4]Exercise: Why?

Anticipating language that will be introduced next lecture, Proposition 16.20 tells us that local operators define *presheaf morphisms*.

Proof. Let $s \in \Gamma(U, E)$ and fix $x \in U$. By Lemma 16.16 there exists a global section $\hat{s}$ of E that agrees with s in some neighbourhood V of x. We define

$$\chi_U(s)(x) \coloneqq \chi(\hat{s})(x).$$

This is well-defined, i.e. independent of the choice of $\hat{s}$ since χ is a local operator. Since $\chi(\hat{s})$ is smooth by assumption, it follows $\chi_U(s)$ is smooth at x, and since x was arbitrary, $\chi_U(s)$ is smooth. Finally, if s is a global section then s is an extension of $s|_U$ for any open U, and thus (16.6) follows. ∎

As a first step to proving that every $C^\infty(M)$-linear operator is a point operator, let us prove the weaker statement that every such operator is a local operator.

PROPOSITION 16.23. *Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ be two vector bundles and let $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ be a $C^\infty(M)$-linear map. Then χ is a local operator.*

Proof. Suppose $s \in \Gamma(E_1)$ vanishes on an open set U. Let $x \in U$, and choose a cutoff function $\eta \colon M \to \mathbb{R}$ such that $\eta(x) = 1$ and $\operatorname{supp}(\eta) \subset U$ (using Lemma 3.2 again). Then $\eta\, s$ is identically zero on M, and so $\chi(\eta\, s)$ is identically zero. However evaluating at x and using $C^\infty(M)$-linearity,

$$0 = \chi(\eta\, s)(x) = \eta(x)\chi(s)(x) = \chi(s)(x).$$

Since x was an arbitrary point of U, we have $\chi(s)|_U \equiv 0$ as required. ∎

PROPOSITION 16.24. *Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ be two vector bundles and let $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ be a $C^\infty(M)$-linear map. Let $U \subset M$ be open. Then the restriction operator $\chi_U \colon \Gamma(U, E_1) \to \Gamma(U, E_2)$ is a $C^\infty(U)$-linear map.*

Proof. Let $s \in \Gamma(U, E_1)$ and let $f \in C^\infty(U)$. Fix $x \in U$ and let $\hat{s} \in \Gamma(E_1)$ denote a global section that agrees with s on a neighbourhood of x, and let $\hat{f}$ be a global smooth function that agrees with f on a neighbourhood[5] of x. Then

$$\chi_U(fs)(x) = \chi(\hat{f}\hat{s})(x) = \hat{f}(x)\chi(\hat{s})(x) = f(x)\chi_U(s)(x).$$

Since x was arbitrary, we see that $\chi_U(fs) = f\chi_U(s)$, as required. ∎

We can now prove:

PROPOSITION 16.25. *Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ be two vector bundles. Let $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ be a $C^\infty(M)$-linear map. Then χ is a point operator. Conversely, if χ is an $\mathbb{R}$-linear point operator then χ is $C^\infty(M)$-linear.*

[5] The existence of $\hat{f}$ is a special case of Lemma 16.16, cf. part (iii) of Example 16.2, but it was also proved directly in Step 2 or Proposition 3.3.

Proof. Let $s \in \Gamma(E_1)$. Suppose $s(x) = 0$. Choose an neighbourhood U of x admitting a local frame $\{e_i\}$. Then we can write

$$s|_U = a^i\, e_i, \qquad a^i \in C^\infty(U).$$

Since $s(x) = 0$ we have $a^i(x) = 0$ for each i. We now compute:

$$\begin{aligned} \chi(s)(x) &\overset{(\dagger)}{=} \chi_U(s|_U)(x) \\ &= \chi_U(a^i\, e_i)(x) \\ &\overset{(\ddagger)}{=} a^i(x)\chi_U(e_i)(x) \\ &= 0 \end{aligned}$$

where (†) used Proposition 16.22 and (‡) used Proposition 16.24.

The converse is easier: fix $f \in C^\infty(M)$, $s \in \Gamma(E_1)$ and $x \in M$. Let $c := f(x)$. Then $fs - cs$ vanishes at x, and thus $\chi(fs - cs)(x) = 0$ as χ is a point operator. Since χ is $\mathbb{R}$-linear,

$$\chi(fs)(x) = \chi(cs)(x) = c\chi(s)(x) = f(x)\chi(s)(x).$$

Since x was arbitrary, $\chi(fs) = f\chi(s)$. ■

As we have seen in Example 16.19, a vector field on a manifold can be thought of an operator on the space of sections of the trivial bundle $M \times \mathbb{R}$ via $f \mapsto X(f)$. In Proposition 7.2 we proved that a map $X\colon M \to TM$ satisfying the section property (7.1) was smooth if and only if $X(f)$ was a smooth function for every smooth function f. The next result generalises this to arbitrary vector bundles.

PROPOSITION 16.26. *Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ be two vector bundles. Suppose $\Phi\colon E_1 \to E_2$ is a fibre-preserving map such that $\Phi|_{E_x}$ is linear for every $x \in M$. Then Φ is smooth (and hence a vector bundle homomorphism) if and only if $\Phi_\star(s) := \Phi \circ s$ belongs to $\Gamma(U, E_2)$ for every $s \in \Gamma(U, E_1)$.*

Proof. If Φ is smooth then certainly $\Phi \circ s$ is smooth. For the converse, let $x \in M$ and suppose $\sigma\colon U \to O$ is a chart on M with local coordinates x^i. We may assume that both E_1 and E_2 admit local frames over U; call them $\{e_j \mid j = 1, \dots, k\}$ and $\{e'_i \mid i = 1, \dots, l\}$ respectively. Since $\Phi_\star$ maps smooth sections to smooth sections, there are functions $f^i_j \in C^\infty(U)$ such that

$$\Phi_\star(e_j) = f^i_j\, e'_i.$$

Now suppose $p \in \pi_1^{-1}(U)$. We can write $p = a^j\, e_j(x)$ for real numbers a^j. Let α_i denote the corresponding vector bundle chart on E_i as in (16.4). Then $(\sigma \circ \pi_i, \alpha_i)$ is a manifold chart on E_i on $\pi_i^{-1}(U)$, and the local expression of Φ is of the form:

$$(\sigma \circ \pi_2, \alpha_2) \circ \Phi \circ (\sigma \circ \pi_1, \alpha_1)^{-1} = \big(\mathrm{id}, (a^j f^1_j, \dots, a^j f^l_j) \circ \sigma^{-1}\big),$$

which is smooth. ■

We need one more lemma, whose proof is analogous to Problem D.1 and is thus left as an exercise.

Lemma 16.27. *Let $\pi\colon E \to M$ be a vector bundle. Let $x \in M$ and let $p \in E_x$. Then there exists a section $s \in \Gamma(E)$ with $s(x) = p$.*

The main step in the proof of the Hom-Γ Theorem is the next claim.

Proposition 16.28. *Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ be two vector bundles. Suppose $\chi\colon \Gamma(E_1) \to \Gamma(E_2)$ is a $C^\infty(M)$-linear map. Then for each $x \in M$ there is a unique linear map $\Phi_x\colon E_1|_x \to E_2|_x$ such that for all $s \in \Gamma(E_1)$, one has*

$$\Phi_x(s(x)) = \chi(s)(x).$$

Proof. Fix $x \in M$ and $p \in E_1|_x$. By Lemma 16.27 there exists a section s such that $s(x) = p$. Define $\Phi_x(p) := \chi(s)(x)$. This definition is independent of the choice of s, since if s_1 was another such section then $(s - s_1)(x) = 0$, and thus $\chi(s)(x) - \chi(s_1)(x) = \chi(s - s_1)(x) = 0$ since χ is a point operator by Proposition 16.25.

We claim that Φ_x is a linear map. Indeed, if $p_1, p_2 \in E_1|_x$ and $a^1, a^2 \in \mathbb{R}$, then if s_1 and s_2 are sections such that $s_i(x) = p_i$ then $a^1 s_1 + a^2 s_2$ is a section whose value at x is $a^1 p_1 + a^2 p_2$ and

$$\begin{aligned}
\Phi_x(a^1 p_1 + a^2 p_2) &= \chi(a^1 s_1 + a^2 s_2)(x) \\
&= a^1\chi(s_1)(x) + a^2\chi(s_2)(x) \\
&= a^1\Phi_x(p_1) + a^2\Phi_x(p_2).
\end{aligned}$$

■

Here is a corollary of Proposition 16.28 that will be useful in Lecture 18

Corollary 16.29. *Let M be a smooth manifold. A $C^\infty(M)$-linear map $\omega\colon \mathfrak{X}(M) \to C^\infty(M)$ is a differential 1-form on M, i.e. an element of $\Omega^1(M)$.*

Proof. Apply Proposition 16.28 with $E_1 = TM$ and E_2 equal to the trivial bundle $M \times \mathbb{R}$. Thus if we are given a $C^\infty(M)$-linear map $\omega\colon \mathfrak{X}(M) \to C^\infty(M)$ then we get for each $x \in M$ a linear map

$$\omega_x\colon T_xM \to \mathbb{R}.$$

By assumption $\omega_x(X(x))$ is a smooth function on M for every $x \in M$, which tells us that $x \mapsto \omega_x$ is a smooth section of T^*M (use Remark 16.9 and Remark 16.11). Since $\bigwedge^1(V) = V$ for any vector space V, this tells us that ω is a smooth section of $\bigwedge^1(T^*M)$, which is the same thing as a differential 1-form. ■

We now move onto our main result.

Theorem 16.30 (The Hom-Γ Theorem). *Let $\pi_1\colon E_1 \to M$ and $\pi_2\colon E_2 \to M$ be two vector bundles. Then there is a one-to-one correspondence between*

$$\{\textit{vector bundle homomorphisms } \Phi\colon E_1 \to E_2\}$$

and

$$\{C^\infty(M)\textit{-linear maps } \chi\colon \Gamma(E_1) \to \Gamma(E_2)\}$$

given by

$$\Phi \mapsto \Phi_\star.$$

(♣) Remark 16.31. Let us explain the name "Hom-Γ Theorem". This needs a little more algebra. Let R be a commutative ring. The category Mod_R of R-modules has objects the R-modules, and morphisms

$$\mathrm{Hom}(N_1, N_2) = \{f \colon N_1 \to N_2 \text{ is an } R\text{-module homomorphism.}\}$$

If you are not familiar with modules, just think of the case $R = \mathbb{R}$. Then an $\mathbb{R}$-module is a vector space, and

$$\mathrm{Hom}(V, W) = \mathrm{L}(V, W).$$

Alternatively, take $R = \mathbb{Z}$: a $\mathbb{Z}$-module is just an abelian group. Since a $C^\infty(M)$-linear map $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ is the same thing as a $C^\infty(M)$-module homomorphism, Theorem 16.30 tells us that

$$\Gamma(\mathrm{Hom}(E_1, E_2)) \cong \mathrm{Hom}(\Gamma(E_1), \Gamma(E_2)),$$

which can be loosely interpreted as

The functors Hom *and* Γ *commute.*

Proof of Theorem 16.30. We first prove surjectivity. If $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ is a $C^\infty(M)$-linear map then by Proposition 16.28 there exists a linear map $\Phi_x \colon E_1|_x \to E_2|_x$ such that for any $s \in \Gamma(E_1)$, $\Phi_x(s(x)) = \chi(s)(x)$. Define $\Phi \colon E_1 \to E_2$ by declaring that $\Phi|_{E_1|_x} = \Phi_x$. Then by Proposition 16.26, the map Φ is a vector bundle homomorphism, and clearly $\Phi_\star = \chi$.

To prove injectivity, suppose $\Phi_\star = \Psi_\star$. Let $x \in M$ and $p \in E_1|_x$ and let $s \in \Gamma(E_1)$ be a section such that $s(x) = p$ (using Lemma 16.27). Then

$$\Phi(p) = \Phi(s(x)) = \Phi_\star(s)(x) = \Psi_\star(s)(x) = \Psi(s(x)) = \Psi(p).$$

Since x and p were arbitrary we conclude $\Phi = \Psi$ as required. ■

LECTURE 17

Sheaves and manifolds

Today's lecture is another algebraic interlude (this will be the last such interlude of the semester). We introduce the notion of a **sheaf**, and use this to unify many of the concepts we've looked at during so far. Just as with the category theory from Lecture 14, we will never actually use any genuine theorems in sheaf theory—for us it will merely be a convenient way to concisely formulate other concepts.

Roughly speaking, a presheaf is a way to assign data locally to open subsets of a topological space in such a way that it is compatible with restrictions. A sheaf is a presheaf for which it is possible to go backwards and reassemble global data from local data.

Definition 17.1. Let X denote a topological space. A **presheaf $\mathcal{F}$ of sets on X** consists of:

(i) A set $\mathcal{F}(U)$ for every open set $U \subset X$.

(ii) For every pair $U \subset V$ of open sets a map $\operatorname{res}^V_U \colon \mathcal{F}(V) \to \mathcal{F}(U)$ called the **restriction map** such that $\operatorname{res}^U_U = \operatorname{id}_{\mathcal{F}(U)}$ for every U and such that

$$\operatorname{res}^W_U = \operatorname{res}^V_U \circ \operatorname{res}^W_V, \qquad \text{whenever } U \subset V \subset W.$$

Definition 17.2. Let $\mathcal{F}$ and $\mathcal{G}$ be two presheaves on X. A **morphism of presheaves** $\chi \colon \mathcal{F} \to \mathcal{G}$ is a family of maps $\chi_U \colon \mathcal{F}(U) \to \mathcal{G}(U)$ such that for every pair of open sets $U \subset V$ the following diagram commutes:

$$\begin{tikzcd}
\mathcal{F}(V) \arrow[r, "\chi_V"] \arrow[d, "\operatorname{res}^V_U"'] & \mathcal{G}(V) \arrow[d, "\operatorname{res}^V_U"] \\
\mathcal{F}(U) \arrow[r, "\chi_U"'] & \mathcal{G}(U)
\end{tikzcd}$$

If $\chi \colon \mathcal{F} \to \mathcal{G}$ and $\xi \colon \mathcal{G} \to \mathcal{H}$ are two morphisms of presheaves over X then their composition $\xi \circ \chi \colon \mathcal{F} \to \mathcal{H}$ is defined as one would guess:

$$(\xi \circ \chi)_U := \xi_U \circ \chi_U.$$

An **isomorphism** is a presheaf morphism such that each χ_U is a bijection.

This gives us the category $\mathsf{PSh}(X; \mathsf{Sets})$ of presheaves on X whose objects are the presheaves on X and whose morphisms are presheaf morphisms.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Definition 17.3. Let C be an arbitrary category. A **presheaf $\mathcal{F}$ on X with values in C** is defined in almost the same way, only now each $\mathcal{F}(U)$ must be an object of C, each restriction map res_U^V must be a morphism in C, and morphisms between two presheaves must also be morphisms in C.

To give a concrete example, let's take $\mathsf{C} = \mathsf{Vect}$. A presheaf of vector spaces is thus an assignment of a vector space $\mathcal{F}(U)$ for every open set $U \subset X$, and the restriction maps res_U^V must be linear transformations $\mathcal{F}(V) \to \mathcal{F}(U)$. Finally if $\chi \colon \mathcal{F} \to \mathcal{G}$ is a morphism of presheaves of vector spaces then each χ_U must be a linear transformation $\mathcal{F}(U) \to \mathcal{G}(U)$. In particular, an isomorphism of presheaves of vector spaces requires each χ_U to be a linear isomorphism.

(♣) Remark 17.4. Here[1] is an alternative more categorical definition of a presheaf. Let $\mathsf{Open}(X)$ denote the category whose objects are the open sets of X and, for two open sets U, V, the morphism space $\operatorname{Hom}(U, V)$ consists of the inclusion map $U \hookrightarrow V$ if $U \subset V$ and is empty otherwise. Then a presheaf on X with values in C is simply a contravariant functor $\mathsf{Open}(X) \to \mathsf{C}$. A morphism $\chi \colon \mathcal{F} \to \mathcal{G}$ is a *natural transformation* χ between the two functors.

If $\mathcal{F}$ is a presheaf on X and $s \in \mathcal{F}(V)$ then for $U \subset V$ we normally abbreviate

$$s|_U \coloneqq \operatorname{res}_U^V(s)$$

This fits in with the idea that we are "restricting" s to U. In fact, every single presheaf we will care about in the course will be a presheaf of *functions*, which we now define, and in this case restriction really *is* restriction.

Definition 17.5. Let X be a topological space and let S be a fixed set. A **presheaf of S-valued functions** is a presheaf with the property that $\mathcal{F}(U) \subset \operatorname{Maps}(U, S)$ for all open sets $U \subset X$, where $\operatorname{Maps}(U, S)$ denotes the set of all functions from U to S (i.e. the morphism set in category Sets).

Definition 17.6. Let $\mathcal{F}$ and $\mathcal{G}$ be two presheaves on X. We say that $\mathcal{F}$ is a **subpresheaf** of $\mathcal{G}$ if for every open set $U \subset X$, $\mathcal{F}(U) \subset \mathcal{G}(U)$, and for all $U \subset V$ open sets the restriction maps $\mathcal{F}(V) \to \mathcal{F}(U)$ are induced by the restriction maps $\mathcal{G}(V) \to \mathcal{G}(U)$.

Thus if $\mathcal{F}$ is any presheaf of S-valued functions on X then $\mathcal{F}$ is a subpresheaf of the presheaf of *all* S-valued functions on X.

Example 17.7. Let us see some standard examples of presheaves that will be relevant to this course.

(i) Let X be a topological space and take $S = \mathbb{R}$. Let $\mathcal{C}_X$ denote the presheaf that assigns to an open set $U \subset X$ the set of continuous real-valued functions on X:

$$\mathcal{C}_X(U) \coloneqq C(U, \mathbb{R}) = \{ f \colon U \to \mathbb{R} \text{ continuous} \} .$$

$\mathcal{C}_X$ is not just a presheaf of sets, but a presheaf of $\mathbb{R}$-algebras (and thus also a presheaf of rings and (infinite-dimensional) vector spaces).

[1] This lecture has quite a lot of (♣) content—if you are not familiar with categories and colimits, just ignore them!

(ii) We can also consider differentiable functions. Take $X = \mathbb{R}$ and let $\mathcal{F}(U) = C^\infty(U)$ denote the set of all smooth functions $U \to \mathbb{R}$. This is a subpresheaf of $\mathcal{C}_{\mathbb{R}}$. We can think of differentiation as a morphism $D\colon \mathcal{F} \to \mathcal{F}$. This is a morphism of presheaves of vector spaces, since

$$D(af + bg) = af' + bg' = aD(f) + bD(g)$$

but it is *not* a morphism of presheaves of algebras, since

$$D(fg) = fg' + f'g \neq D(f)D(g).$$

(iii) More generally, let M be a smooth manifold. Then the assignment $U \mapsto C^\infty(U)$ is a presheaf of $\mathbb{R}$-algebras on M. As before, differentiation is a morphism of presheaves of vector spaces, but not of algebras. We normally denote this presheaf by $\mathcal{C}^\infty_M$.

(iv) Let $\pi\colon E \to M$ be a vector bundle. Then $U \mapsto \Gamma(U, E)$ is a presheaf of (infinite-dimensional) vector spaces on M. It is not a presheaf of algebras (unless E is an algebra bundle, cf. Remark 15.27), since in general there is no way to multiply two sections together. We usually denote this presheaf by $\mathcal{E}_E$.

(v) Let X be any topological space and let S be any set. Let $\mathcal{F}(U)$ denote the set of all *constant* functions $U \to S$. Since a constant function $f\colon U \to S$ can be identified with its image $s := f(U)$, one can simply think of $\mathcal{F}(U)$ as being equal to S itself. In this case, all restriction maps are the identity map id_S. We call this the **constant presheaf on X with values in S**.

Let us now introduce a sheaf, which is a presheaf with an additional property.

Definition 17.8. Let $\mathcal{F}$ be a presheaf on X (of sets, rings, groups, etc.). We say that $\mathcal{F}$ is a **sheaf** if the following condition is satisfied: for any open set $U \subset X$ and any open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of U, if we are given a collection $s_{\mathtt{a}} \in \mathcal{F}(U_{\mathtt{a}})$ such that

$$s_{\mathtt{a}}|_{U_{\mathtt{a}} \cap U_{\mathtt{b}}} = s_{\mathtt{b}}|_{U_{\mathtt{a}} \cap U_{\mathtt{b}}}, \qquad \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A} \text{ such that } U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset, \tag{17.1}$$

then there exists a *unique* $s \in \mathcal{F}(U)$ such that $s|_{U_{\mathtt{a}}} = s_{\mathtt{a}}$ for all $\mathtt{a} \in \mathtt{A}$.

Remark 17.9. Taking $U = \emptyset$ and choosing the covering with empty index set $\mathtt{A} = \emptyset$ shows that if $\mathcal{F}$ is a sheaf then $\mathcal{F}(\emptyset)$ is a set consisting of one element.

A morphism $\chi\colon \mathcal{F} \to \mathcal{G}$ of sheaves is simply a morphism of the underlying presheaves, and we denote by $\mathsf{Sh}(X; \mathsf{C})$ the category of sheaves on X with values in C.

Remark 17.10. If we start with a presheaf of functions, as in Definition 17.5, the condition (17.1) can be phrased in a slightly simpler fashion: if $\mathcal{F}$ is a presheaf of S-valued functions on X then $\mathcal{F}$ is a sheaf if and only if for any open set $U \subset X$ and any open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ of U, if $f\colon U \to S$ is any function such that $f|_{U_{\mathtt{a}}} \in \mathcal{F}(U_{\mathtt{a}})$ for each $\mathtt{a} \in \mathtt{A}$, then $f \in \mathcal{F}(U)$.

This reformulation makes it clear that the presheaf $\mathcal{C}_X$ of continuous functions on a topological space is actually a sheaf. The proof of the following result is on Problem Sheet I.

Proposition 17.11. *Let M be a smooth manifold. Then $\mathcal{C}^\infty_M$ is a sheaf. More generally, if $\pi\colon E \to M$ is any vector bundle over M then $\mathcal{E}_E$ is a sheaf.*

Not everything is a sheaf however: the presheaf of constant functions from part (v) of Example 17.7 is not a sheaf if X contains two disjoint non-empty open subsets and S has more than one element[2].

There is a natural way to turn a presheaf into a sheaf. This procedure is called the **sheafification** of a presheaf. The definition is rather complicated, and for our purposes unimportant (since the relevant presheaves in this course are already sheaves thanks to Proposition 17.11). Thus we will content ourselves with giving the definition only in the special case of a presheaf of functions.

Proposition 17.12. *Let X be a topological space and let S be a set. Suppose $\mathcal{F}$ is a presheaf of S-valued functions on X. Let*

$$\tilde{\mathcal{F}}(U) := \Big\{ f\colon U \to S \mid \text{there exists an open covering } \{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\} \\ \text{of } U \text{ such that } f|_{U_{\mathtt{a}}} \in \mathcal{F}(U_{\mathtt{a}}) \text{ for all } \mathtt{a} \in \mathtt{A}. \Big\}$$

Then $\tilde{\mathcal{F}}$ is a sheaf and the inclusion $\mathcal{F}(U) \hookrightarrow \tilde{\mathcal{F}}(U)$ induces a morphism of presheaves $\imath\colon \mathcal{F} \to \tilde{\mathcal{F}}$.

Proof. This is clear from the reformulation of the sheaf condition given in Remark 17.10: we simply added in all the functions that were missing in order for $\mathcal{F}$ into a sheaf. ∎

Remark 17.13. If $\mathcal{F}$ already was a sheaf, then clearly $\mathcal{F} = \tilde{\mathcal{F}}$.

Example 17.14. Let $\mathcal{F}$ be the presheaf of constant S-valued functions on X. As we have remarked before, this is typically not a sheaf. However it is very easy to describe the sheaf obtained from $\mathcal{F}$ via Proposition 17.12. Indeed, a little thought shows that the sheaf $\tilde{\mathcal{F}}$ is exactly the *locally constant* functions on S:

$$\tilde{\mathcal{F}}(U) = \{ f\colon U \to S \mid f \text{ is locally constant} \}.$$

(♣) Remark 17.15. The sheafification can be defined via a universal property (compare Lemma 15.2): Let $\mathcal{F}$ be a presheaf on X. The sheafification $\tilde{\mathcal{F}}$ and the morphism $\imath\colon \mathcal{F} \to \tilde{\mathcal{F}}$ of presheaves has the property that if $\mathcal{G}$ is any sheaf on X and $\chi\colon \mathcal{F} \to \mathcal{G}$ is any morphism of presheaves, then there exist a *unique* morphism of sheaves $\chi\colon \tilde{\mathcal{F}} \to \mathcal{G}$ such that the following diagram commutes:

$$\begin{tikzcd} \mathcal{F} \arrow[rr, "\imath"] \arrow[dr, "\chi"'] & & \tilde{\mathcal{F}} \arrow[dl, dashed, "\tilde{\chi}"] \\ & \mathcal{G} & \end{tikzcd}$$

[2] Exercise: Why?

As such, via abstract nonsense[3], the sheafification is unique up to isomorphism.

We now move onto discussing the **stalk** of a presheaf. This generalises the notation of a germ of a function that we discussed in Lecture 2.

DEFINITION 17.16. Let $\mathcal{F}$ be a presheaf on X, and let $x \in X$. We define the **stalk** of $\mathcal{F}$ at x to be:

$$\mathcal{F}_x := \{(U, s) \mid U \text{ is a neighbourhood of } x,\ s \in \mathcal{F}(U)\} \Big/ \sim$$

where $(U, s) \sim (V, t)$ if there exists a neighbourhood $W \subset U \cap V$ such that $s|_W \equiv t|_W$.

Thus for any neighbourhood U of x there exists a canonical map $\mathcal{F}(U) \to \mathcal{F}_x$ that sends s to the equivalence class of (U, s) in $\mathcal{F}_x$, which we denote by $\underline{s}$.

LEMMA 17.17. *Let $\chi\colon \mathcal{F} \to \mathcal{G}$ be a morphism of presheaves. Then for each $x \in X$ there is a well-defined map $\chi_x\colon \mathcal{F}_x \to \mathcal{G}_x$ defined as follows: if $\underline{s} \in \mathcal{F}_x$ is represented by (U, s), then we declare that $(U, \chi_U(s))$ is a representative of $\chi_x(\underline{s})$. Thus the following diagram commutes:*

$$\begin{tikzcd}
\mathcal{F}(U) \arrow[r, "s \mapsto \underline{s}"] \arrow[d, "\chi_U"'] & \mathcal{F}_x \arrow[d, "\chi_x"] \\
\mathcal{G}(U) \arrow[r, "t \mapsto \underline{t}"'] & \mathcal{G}_x
\end{tikzcd}$$

Proof. We need only check this is well-defined. Suppose $(U, s) \sim (V, t)$. Then there exists $W \subset U \cap V$ such that $s|_W \equiv t|_W$. Since χ is a presheaf morphism, one has that

$$\chi_U(s)|_W = \chi_W(s|_W) = \chi_W(t|_W) = \chi_V(t)|_W.$$

Thus $(U, \chi_U(s)) \sim (V, \chi_V(t))$. ∎

(♣) REMARK 17.18. A more categorical way to define stalks is the following: given $x \in X$, let $\mathsf{Open}_x(X)$ denote the full subcategory of $\mathsf{Open}(X)$ (cf. Remark 17.4) consisting of neighbourhoods of x. Then if $\mathcal{F}$ is a presheaf on X, one has

$$\mathcal{F}_x = \varinjlim \mathcal{F}(U)$$

where the filtered colimit runs over $\mathsf{Open}_x(X)$. Similarly if $\chi\colon \mathcal{F} \to \mathcal{G}$ is a morphism of presheaves then

$$\chi_x = \varinjlim \chi_U.$$

If $\mathcal{F}$ is a presheaf of groups, or rings, or modules, etc, then the stalks also inherit that structure. We saw this in the concrete example where $\mathcal{F} = \mathcal{C}^\infty_M$ just after Definition 2.8. As another example, suppose $\mathcal{F}$ is a sheaf of groups. Then $\mathcal{F}_x$ is also a group, where we define the group law as follows: if $\underline{s}$ is represented by (U, s) and $\underline{t}$ is represented by (V, t), then we declare $\underline{s} \cdot \underline{t}$ to be the element represented by $(U \cap V, s|_{U\cap V} \cdot t|_{U \cap V})$.

[3] i.e. category theory.

(♣) REMARK 17.19. More generally, if C is a category in which filtered colimits exist then for any $x \in X$ there is a functor $\mathsf{PSh}(X; \mathsf{C}) \to \mathsf{C}$ given by $\mathcal{F} \mapsto \mathcal{F}_x$.

Let us now look at some operations on sheaves.

DEFINITION 17.20. Let U be an open set of X. Then if $\mathcal{F}$ is any presheaf on X then we can define a presheaf $\mathcal{F}|_U$ on U by setting $\mathcal{F}|_U(V) := \mathcal{F}(V)$ for $V \subset U$ open. If $\mathcal{F}$ is a sheaf then so is $\mathcal{F}|_U$.

DEFINITION 17.21. Let $\varphi \colon X \to Y$ be a continuous map from one topological space to another. Suppose $\mathcal{F}$ is a presheaf on X. We define a presheaf $\varphi_\star(\mathcal{F})$ on Y by declaring that

$$\varphi_\star(\mathcal{F})(U) := \mathcal{F}(\varphi^{-1}(U)), \qquad U \subset Y \text{ open.}$$

We call $\varphi_\star(\mathcal{F})$ the **direct image** of $\mathcal{F}$ under φ. If $\chi \colon \mathcal{F} \to \mathcal{G}$ is a morphism of presheaves on X then $\varphi_\star(\chi) \colon \varphi_\star(\mathcal{F}) \to \varphi_\star(\mathcal{G})$ is a morphism on presheaves on Y, where

$$\varphi_\star(\chi)_U := \chi_{\varphi^{-1}(U)} \colon \varphi_\star(\mathcal{F})(U) = \mathcal{F}(\varphi^{-1}(U)) \to \mathcal{G}(\varphi^{-1}(U)) = \varphi_\star(\mathcal{G})(U).$$

In this way we get a functor from presheaves on X to presheaves on Y. If $\mathcal{F}$ is a sheaf on X then it is clear that $\varphi_\star(\mathcal{F})$ is a sheaf on Y.

DEFINITION 17.22. A **continuous ringed space** consists of a pair $(X, \mathcal{F})$ where X is a topological space and $\mathcal{F}$ is a sheaf of subalgebras of the sheaf of continuous functions on X. Explicitly, this means:

- $\mathcal{F}$ is a sheaf and $\mathcal{F}(U) \subset C(U, \mathbb{R})$ for each open set $U \subset X$.
- If $f, g \in \mathcal{F}(U)$ and $a, b \in \mathbb{R}$ then $af + bg$ and fg both belong to $\mathcal{F}(U)$.

(♣) REMARK 17.23. The name "continuous ringed space" is not quite standard[4]. In algebraic geometry, given a commutative ring R, one studies the more general notion of a **ringed space**, which is defined to be a pair $(X, \mathcal{F})$, where X is a topological space and $\mathcal{F}$ is a sheaf of commutative, associative and unital R-algebras on X. Thus what I call a "continuous ringed space" is the special case where $R = \mathbb{R}$ and $\mathcal{F}$ is a subalgebra of the sheaf of continuous functions on X.

Algebraic geometers often restrict to a special class of ringed spaces, called **locally ringed spaces**, which are ringed spaces $(X, \mathcal{F})$ with the additional property that the stalk $\mathcal{F}_x$ is a *local ring* for every point $x \in X$ (i.e. it has a unique maximal ideal). All continuous ringed spaces in the sense of Definition 17.22 are locally ringed spaces; see Remark 2.9.

DEFINITION 17.24. Let $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ be two continuous ringed spaces. A **morphism of continuous ringed spaces** is a continuous map $\varphi \colon X \to Y$ with the following property:

$$f \in \mathcal{G}(U) \quad \Rightarrow \quad f \circ \varphi \in \mathcal{F}(\varphi^{-1}(U)), \qquad \text{for all open } U \subset Y. \tag{17.2}$$

[4] Actually, not at all standard, since I just made it up...

Property (17.2) implies there is a well-defined sheaf morphism $\mathcal{G} \to \varphi_\star(\mathcal{F})$ given by

$$f \in \mathcal{G}(U) \mapsto f \circ \varphi \in \varphi_\star(\mathcal{F})(U).$$

An **isomorphism of continuous ringed spaces** is a homeomorphism φ such that both φ and φ^{-1} are morphisms of continuous ringed spaces.

We will now use the notion of a continuous ringed space to give an equivalent definition of a manifold. This definition is more in the spirit of algebraic geometry, and it has several advantages over the standard one, as we will shortly explain.

DEFINITION 17.25. Let $(M, \mathcal{F})$ be a continuous ringed space. We say $(M, \mathcal{F})$ is a **smooth ringed space of dimension** n if for every point $x \in M$ there exists a neighbourhood U of x and a homeomorphism $\sigma \colon U \to O$, where O is some open subset of $\mathbb{R}^n$, such that σ defines an isomorphism of continuous ringed spaces

$$(U, \mathcal{F}|_U) \cong (O, \mathcal{C}^\infty_O).$$

The next theorem tells us that this really is an alternative way to define a manifold.

THEOREM 17.26. *Let M be a smooth manifold of dimension n. Then $(M, \mathcal{C}^\infty_M)$ is a smooth ringed space of dimension n. Conversely, assume that $(M, \mathcal{F})$ is a smooth ringed space, and assume in addition that M is Hausdorff, paracompact, and has at most countably many components. Then there exists a unique smooth structure on M such that $\mathcal{F}$ becomes the sheaf $\mathcal{C}^\infty_M$.*

The proof is easy: one direction is clear from the definition of a smooth function on a manifold (Definition 2.1), and for the other direction we (work a bit and then) apply Proposition 1.22. I invite you to fill in the details (this is **not** examinable though!)

REMARK 17.27. In many ways, starting Lecture 1 by defining a manifold via Definition 17.25 would have been more efficient. Here are some reasons why:

(i) There is no need to worry about equivalence classes of smooth atlases (cf. Remark 1.17).

(ii) The definition of what it means for a continuous map $\varphi \colon (M, \mathcal{C}^\infty_M) \to (N, \mathcal{C}^\infty_N)$ between two smooth manifolds to be smooth is much cleaner: it simply has to be a morphism of continuous ringed spaces[5].

(iii) The definition of a tangent vector as a derivation on the space of germs (i.e. the stalks of the sheaf $\mathcal{C}^\infty_M$) is far more natural.

(iv) This algebraic approach dramatically reduces the need to use local coordinates, which are messy and irritating.

Nevertheless, the best part[6] of differential geometry is the "geometry", and this algebraic approach deletes most of said geometry. So we will not pursue it beyond this lecture.

[5]Exercise: Check this!

[6]In my opinion, at least...

We conclude this lecture by giving a sheaf-theoretic definition of a vector bundle. This will also allow us to reinterpret the Hom Γ Theorem 16.30 from the last lecture in a more algebraic way. First, some preliminary definitions.

DEFINITION 17.28. Let $(X, \mathcal{F})$ be a continuous ringed space. Let $\mathcal{M}$ be a sheaf of abelian groups on X, and assume in addition that for every open set $U \subset X$, the abelian group $\mathcal{M}(U)$ has the structure of an $\mathcal{F}(U)$-module, and moreover the restriction morphisms respect this structure, ie.

$$\mathrm{res}^V_U(fs) = \mathrm{res}^V_U(f)\,\mathrm{res}^V_U(s), \qquad \forall\, f \in \mathcal{F}(V),\ s \in \mathcal{M}(V).$$

Then we say that $\mathcal{M}$ is a **sheaf of $\mathcal{F}$-modules**. A morphism χ from one sheaf $\mathcal{M}$ of $\mathcal{F}$-modules to another sheaf $\mathcal{N}$ of $\mathcal{F}$-modules is one such that each map $\chi_U \colon \mathcal{M}(U) \to \mathcal{N}(U)$ is $\mathcal{F}(U)$-linear. We call such a χ an **$\mathcal{F}$-morphism** of sheaves.

Here is an example.

EXAMPLE 17.29. Let $\pi \colon E \to M$ be a vector bundle. Then the sheaf $\mathcal{E}_E$ of sections of E is a sheaf of $\mathcal{C}^\infty_M$-modules. Indeed, this is just a fancy way of rephrasing Lemma 16.5.

We can also rephrase some of the results from the previous lecture.

COROLLARY 17.30. *Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$, and suppose $\chi \colon \Gamma(E_1) \to \Gamma(E_2)$ is an $\mathbb{R}$-linear operator. Then χ is a local operator in the sense of Definition 16.17 if and only if $\chi = \chi_M$ for a morphism of sheaves $\chi \colon \mathcal{E}_{E_1} \to \mathcal{E}_{E_2}$.*

Proof. This is Proposition 16.22. ∎

COROLLARY 17.31. *Let $\pi_1 \colon E_1 \to M$ and $\pi_2 \colon E_2 \to M$ be two vector bundles over M. Suppose $\chi \colon \mathcal{E}_{E_1} \to \mathcal{E}_{E_2}$ is a C^∞_M-morphism of sheaves. Then $\chi_M \colon \Gamma(E_1) \to \Gamma(E_2)$ is a point operator in the sense of Definition 16.17.*

Proof. This is Proposition 16.25. ∎

Here is another more abstract example of an $\mathcal{F}$-module.

EXAMPLE 17.32. Let $(X, \mathcal{F})$ be a continuous ringed space. Let $n \in \mathbb{N}$. Then the sum

$$\mathcal{F}^k(U) := \underbrace{\mathcal{F}(U) \oplus \cdots \oplus \mathcal{F}(U)}_{k}$$

is a *free* $\mathcal{F}$-module of rank k.

More generally, if $\mathcal{M}$ is any $\mathcal{F}$-module over X then we say that $\mathcal{M}$ is **locally free of rank** k if for any $x \in X$ there exists a neighbourhood U of x and an $\mathcal{F}|_U$-isomorphism of sheaves $\mathcal{M}|_U \cong \mathcal{F}^k$. If $\mathcal{M}$ is locally free of rank k then with a little work one can show that the stalk $\mathcal{M}_x$ is a free $\mathcal{F}_x$-module of rank k.

EXAMPLE 17.33. Let $\pi \colon E \to M$ be a vector bundle of rank k. Then the sheaf $\mathcal{E}_E$ is locally free of rank k. Indeed, this follows from the fact that for any $x \in M$, there exists a neighbourhood U of x such that E admits a local frame $\{e_1, \dots, e_k\}$. Then any $s \in \Gamma(U, E)$ can be written as a

$$s = a^i e_i, \qquad a^i \in C^\infty(U),$$

and the correspondence $s \mapsto (a^1, \dots, a^k)$ sets up an isomorphism $\mathcal{E}_E|_U$ with $(\mathcal{C}^\infty_M|_U)^k$.

Just as in Theorem 17.26, it is actually possible to work backwards and *define* a vector bundle this way.

THEOREM 17.34. *Let M be a smooth manifold and let $\mathcal{M}$ be a sheaf of locally free $\mathcal{C}^\infty_M$-modules of rank k. Then there exists a vector bundle $\pi\colon E \to M$ and a $\mathcal{C}^\infty_M$-isomorphism of sheaves from $\mathcal{M}$ to $\mathcal{E}_E$. Moreover E is unique up to vector-bundle isomorphism.*

This proof is non-examinable and is rather sketchy.

(♣) *Proof.* The stalk $\mathcal{F}_x M$ of $\mathcal{C}^\infty_M$ is a local ring with maximal ideal $\mathfrak{m}_x$ equal to the kernel of the evaluation map (see Remark 2.9). The stalk $\mathcal{M}_x$ is a free $\mathcal{F}_x$-module of rank k. Thus if we set

$$E_x := \mathcal{M}_x / \mathfrak{m}_x \mathcal{M}_x$$

then E_x is a vector space of dimension k. Now set $E = \bigsqcup_{x \in M} E_x$. If $x \in M$ and $U \subset M$ is a neighbourhood such that $\mathcal{M}|_U \cong (\mathcal{C}^\infty_M|_U)^k$ then this gives us a basis $\{e_1(x), \dots, e_k(x)\}$ of E_x for every $x \in U$, and thus a local frame for E. This gives us a bundle chart via (16.4). We use this to define a fibre bundle structure on E via Remark 13.7. The transition functions arising from a different choice of local frame near x are linear by assumption, and thus we have built a vector bundle. ■

(♣) REMARK 17.35. Theorem 17.34 tells us that there is a one-to-one correspondence (up to isomorphism) between vector bundles and locally free sheaves. From the point of view of categories, this gives us a way to go from an object of the category of vector bundles to an object of the category of finite rank locally free sheaves. A souped-up version of $\operatorname{Hom}\Gamma$ Theorem from the previous lecture allows us to extend this to morphisms too: i.e. a vector bundle homomorphism $E_1 \to E_2$ is equivalent to an $\mathcal{C}^\infty_M$-morphism of sheaves. This allows us to conclude the following result: *there is an equivalence of categories between the category of vector bundles over M and the category of finite rank locally free $\mathcal{C}^\infty_M$-modules.*

LECTURE 18

Tensor fields

In this lecture we return to tensor fields, as introduced in Definition 16.13 and study them in more detail. Our eventual aim is to extend the Lie derivative $\mathcal{L}_X$ as a sheaf morphism on the tensor algebra sheaf $\mathcal{T}_M$, thus fulfilling the claims made in Lecture 7 and Lecture 8.

Let M be a smooth manifold. Using Proposition 17.11, for each $r, s \geq 0$ there is a sheaf $\mathcal{T}_M^{r,s}$ over M which associates to an open set $U \subset M$ the infinite-dimensional vector space of tensor fields of type (r, s) on U:

$$\mathcal{T}_M^{r,s}(U) := \mathcal{T}^{r,s}(U) = \Gamma(U, T^{r,s}(TM)).$$

The subscript M is added as a notational hint that we are thinking of $U \mapsto \mathcal{T}^{r,s}(U)$ as a sheaf, and it will sometimes be omitted, particularly when we take $U = M$. Thus $\mathcal{T}_M^{0,0} = \mathcal{C}_M^\infty$ is the sheaf of smooth functions on M and $\mathcal{T}_M^{1,0}$ is the sheaf of vector fields on M.

If we set

$$\mathcal{T}_M(U) := \bigoplus_{r,s\geq 0} \mathcal{T}_M^{r,s}(U) = \bigoplus_{r,s\geq 0} \Gamma(U, T^{r,s}(TM))$$

then $\mathcal{T}$ is a sheaf of graded $\mathbb{R}$-algebras over M, cf. Definition 15.18, where for $A \in \mathcal{T}_M^{r,s}(U)$ and $B \in \mathcal{T}_M^{r_1,s_1}(U)$, the product $A \otimes B$ belongs to $\mathcal{T}_M^{r+r_1,s+s_1}(U)$ and is defined pointwise via Definition 15.18, see (15.3) and (15.4). We call $\mathcal{T}_M$ the **tensor algebra sheaf** on M.

Remark 18.1. As in (15.4), strictly speaking the tensor $A \otimes B$ needs its factors rearranging. If for instance $A = X_1 \otimes \omega_1$ and $B = X_2 \otimes \omega_2$ for vector fields X_i and 1-forms ω_i, then $A \otimes B$ should really be written as $X_1 \otimes X_2 \otimes \omega_1 \otimes \omega_2$ (so that the vector field factors come first). In practice, this is inconvenient, and so we will often not bother and just keep the factors unchanged, thus writing $A \otimes B = X_1 \otimes \omega_1 \otimes X_2 \otimes \omega_2$. This is harmless, since it was merely a convention to put the vector fields first (cf. Corollary 15.14).

Remark 18.2. **Warning**: As explained in Remark 15.19, we cannot form a "vector bundle" $\bigoplus_{r,s\geq 0} T^{r,s}(TM)$, since we have not defined infinite-dimensional vector bundles. Thus one *cannot* interpret $\mathcal{T}$ as a sheaf of sections of a vector bundle over M.

The next result gives an alternative way to interpret tensors.

Theorem 18.3 (The Tensor Criterion). *Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Then there is a canonical identification between $\mathcal{T}^{r,s}(W)$ and $C^\infty(W)$-multilinear functions*

$$\underbrace{\Omega^1(W) \times \cdots \times \Omega^1(W)}_{r} \times \overbrace{\mathfrak{X}(W) \times \cdots \times \mathfrak{X}(W)}^{s} \to C^\infty(W).$$

The proof of Theorem 18.3 is on Problem Sheet I. We will typically suppress this isomorphism from our notation, and thus interchangeably regard a tensor field A over U either as an element of $\mathcal{T}^{r,s}(U)$, or as an appropriate multilinear map.

(♣) REMARK 18.4. In Lecture 15 we defined tensor products for finite-dimensional real vector spaces. However everything would have worked (without any changes at all) if we worked with finite rank modules over a fixed commutative ring R. A more interesting question is to what extent the finite rank hypothesis was needed. Indeed, suppose V is a module over a commutative ring R. The dual module is defined $V^* = \mathrm{Hom}_R(V, R)$, and the space $\mathrm{Mult}_{r,s}(V)$ is then defined be the set of multilinear maps

$$\underbrace{V \times \cdots \times V}_{r} \times \overbrace{V^* \times \cdots \times V^*}^{s} \to R.$$

One can then ask the question: *is it true that* $T^{r,s}(V)$ *and* $\mathrm{Mult}_{s,r}(V)$ *are isomorphic modules*?

$$T^{r,s}(V) \overset{?}{\cong} \mathrm{Mult}_{s,r}(V) \tag{18.1}$$

The answer in general is **no**[1]. Nevertheless (18.1) it is true for *some* infinite-rank modules. Rather than giving a precise theorem, let us just state the special case we care about:

THEOREM 18.5. *Let $\pi\colon E \to M$ be a vector bundle. For any open set $U \subset M$, the space $\Gamma(U, E)$ is a module over the ring $C^\infty(U)$ by Lemma 16.5, and in fact for this module (18.1) is true:*

$$T^{r,s}(\Gamma(U, E)) \cong \mathrm{Mult}_{s,r}(\Gamma(U, E)).$$

With Theorem 18.5 we can state the Tensor Criterion (Theorem 18.3) more succinctly as:

$$\mathcal{T}^{r,s}(W) \cong T^{r,s}(\mathfrak{X}(W)), \tag{18.2}$$

where we are using Corollary 16.29 to identify the dual module of $\mathfrak{X}(W)$ with $\Omega^1(W)$. This could also be called the "Tensor-Γ Theorem" in analogy with the Hom-Γ Theorem 16.30, since (18.2) is just compact notation for

$$\Gamma(W, T^{r,s}(TM)) \cong T^{r,s}(\Gamma(W, TM)),$$

which is (almost) saying that the functors $T^{r,s}$ and Γ "commute".

Note also that $W \mapsto T^{r,s}(\mathfrak{X}(W))$ does *not* naturally form a sheaf (in fact, not even a presheaf). Indeed, if $U \subset W$ then there is no obvious way to "restrict" a multilinear map with arguments in the module $\mathfrak{X}(W)$ (and its dual) to a multilinear map with arguments in $\mathfrak{X}(U)$.

REMARK 18.6. Every differential geometer should at one point in their life compute tensor fields in coordinates. This is ~~soul-destroyingly boring and I really don't want to type it out~~ left as a wholesome exercise for you to enjoy on Problem Sheet J.

[1]Exercise: Spot all the places in the proof of Proposition 15.9 that go wrong when V is allowed to be infinite-dimensional.

In Lecture 7 we showed how a *diffeomorphism* $\varphi\colon M \to N$ can "push forward" a function f on M to a function on N, written $\varphi_\star(f)$ and similarly a vector field X on M to a vector field on N, written $\varphi_\star(X)$. We will shortly generalise this to all tensors, but first let us show how *any* smooth map $\varphi\colon M \to N$ (not necessarily a diffeomorphism) can "pull back" a tensor field of type $(0, s)$ on N to a tensor field of type $(0, s)$ on M.

Definition 18.7. Suppose now that $\varphi\colon M \to N$ is a smooth map between two smooth manifolds, and suppose $A \in \mathcal{T}^{0,s}(N)$ is a $(0, s)$ tensor field on N. We define the **pullback** of A by φ, written[2] $\varphi^\star(A)$ to be the tensor field on $\mathcal{T}^{0,s}(M)$ defined pointwise as follows: If $s > 0$ then:

$$\varphi^\star(A)_x(v_1, \dots, v_s) := A_{\varphi(x)}\big(D\varphi(x)[v_1], \dots, D\varphi(x)[v_s]\big), \qquad \forall\, x \in M, v_i \in T_xM.$$

Meanwhile for $s = 0$ we simply set $\varphi^\star(f) := f \circ \varphi$.

It is immediate that $\varphi^\star$ is $\mathbb{R}$-linear. Moreover if $f \in C^\infty(N)$ then

$$\varphi^\star(fA) = (f \circ \varphi)\varphi^\star(A).$$

Note also that if $\psi\colon L \to M$ then $(\varphi \circ \psi)^\star = \psi^\star \circ \varphi^\star$. This definition immediately extends to define

$$\varphi^\star\colon \bigoplus_{s\geq 0} \mathcal{T}^{0,s}(N) \to \bigoplus_{s\geq 0} \mathcal{T}^{0,s}(M),$$

and this is an algebra morphism, i.e.

$$\varphi^\star(A \otimes B) = \varphi^\star(A) \otimes \varphi^\star(B).$$

Unravelling the definitions also shows that $\varphi^\star$ defines a sheaf morphism

$$\varphi^\star\colon \mathcal{T}^{0,s}_N \to \varphi_\star(\mathcal{T}^{0,s}_M)$$

of sheaves over[3] N.

To extend this definition to tensors of arbitrary type, we need to assume that φ is a diffeomorphism. First let us introduce the notion of a **cotangent lift**.

Definition 18.8. Let $\varphi\colon M \to N$ be a diffeomorphism. Then $D\varphi(x)\colon T_xM \to T_{\varphi(x)}N$ is a linear isomorphism for each x, and thus we can speak of its inverse $(D\varphi(x))^{-1}\colon T_{\varphi(x)}N \to T_xM$ (this coincides with the differential of φ^{-1} at the point $\varphi(x)$). Thus there is a well-defined map

$$(D\varphi)^\dagger\colon T^*M \to T^*N,$$

called the **cotangent lift** defined for $x \in M$, $p \in T^*_xM$ and $v \in T_{\varphi(x)}N$ by

$$(D\varphi)^\dagger(x)[p][v] := p\left((D\varphi(x))^{-1}[v]\right)$$

[2]The upper star indicates that $\varphi \mapsto \varphi^\star$ is contravariant, i.e. it reverses the direction of the arrows.

[3]Note one has to use the direct image sheaf here in order to get a sheaf morphism, since (by definition) sheaf morphisms are only defined over the same base space!

The cotangent lift makes the following diagram commute:

$$\begin{tikzcd} T^*M \arrow[r, "(D\varphi)^\dagger"] \arrow[d, "\pi_M"'] & T^*N \arrow[d, "\pi_N"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

where $\pi_M\colon T^*M \to M$ and $\pi_N\colon T^*N \to N$ are the projections. Thus $(D\varphi)^\dagger$ is a vector bundle morphism along φ.

DEFINITION 18.9. Now suppose $\varphi\colon M \to N$ is a diffeomorphism. Then we can define the pullback tensor $\varphi^\star(A) \in \mathcal{T}^{r,s}(M)$ for a tensor $A \in \mathcal{T}^{r,s}(N)$ of arbitrary type (r, s) by setting

$$\begin{aligned}\varphi^\star(A)_x(p_1, \ldots, p_r, &v_1, \ldots, v_s) \\ &:= A_{\varphi(x)}\Big(D\varphi^\dagger(x)[p_1], \ldots, D\varphi^\dagger(x)[p_r], D\varphi(x)[v_1], \ldots, D\varphi(x)[v_s]\Big)\end{aligned}$$

for $x \in M$, $p_1, \ldots, p_r \in T_x^*M$ and $v_1, \ldots, v_s \in T_xM$.

REMARK 18.10. More generally, if $\varphi\colon U \subset M \to N$ is a locally defined map which is a diffeomorphism onto its image $V = \varphi(U)$ then we can still use φ to pull back tensor fields of arbitrary type from V to U.

Thus when φ is a diffeomorphism we can construct a well-defined map $\varphi^\star\colon \mathcal{T}(N) \to \mathcal{T}(M)$ on the entire tensor algebra, and thus a sheaf morphism $\mathcal{T}_N \to \varphi_\star(\mathcal{T}_M)$ of graded $\mathbb{R}$-algebras, i.e.

$$\varphi^\star(A \otimes B) = \varphi^\star(A) \otimes \varphi^\star(B). \tag{18.3}$$

Finally we can invert this to obtain the desired map from Remark 7.18.

DEFINITION 18.11. Let $\varphi\colon M \to N$ be a diffeomorphism. We define $\varphi_\star := (\varphi^{-1})^\star$, which is therefore a map from $\mathcal{T}(M) \to \mathcal{T}(N)$ (or a sheaf morphism $\mathcal{T}_M \to (\varphi^{-1})_\star(\mathcal{T}_N)$).

In the special case $r = s = 0$, the map $\varphi_\star$ sends a function f to $f \circ \varphi^{-1}$. In the special case $r = 1$ and $s = 0$, one has

$$\varphi_\star(X)(x) = D\varphi(\varphi^{-1}(x))[X(\varphi^{-1}(x))],$$

and thus in both cases these extend the definitions from Lecture 7.

We now work towards extending the Lie derivative $\mathcal{L}_X$ to a *tensor derivation* on $\mathcal{T}(M)$. On $\mathcal{T}^{0,0}(M)$ we already defined $\mathcal{L}_X(f) = X(f)$, and on $\mathcal{T}^{1,0}(M)$ we already defined $\mathcal{L}_X(Y) = [X, Y]$. To extend the definition to higher tensors we first need a little bit of linear algebra.

DEFINITION 18.12. Let V be a vector space and fix $r, s \geq 0$. Choose $h \leq r$ and $k \leq s$. The (h, k)**th contraction**, written $C^{h,k}$ is the linear operator

$$C^{h,k}\colon T^{r,s}(V) \to T^{r-1,s-1}(V)$$

defined on decomposable elements by feeding the hth V-factor to the kth V^* factor:

$$C^{h,k}(v_1 \otimes \cdots \otimes v_h \otimes \cdots \otimes v_r \otimes p^1 \otimes \cdots \otimes p^k \otimes \cdots \otimes p^s) := \\ p^k(v_h) \cdot v_1 \otimes \cdots \otimes v_{h-1} \otimes v_{h+1} \otimes \cdots \otimes v_r \otimes p^1 \otimes \cdots \otimes p^{k-1} \otimes p^{k+1} \otimes \cdots \otimes p^s$$

and then extending by linearity.

LEMMA 18.13. *Assume* $\dim V = n$. *Let* $A \in T^{r,s}(V)$, *and regard* A *as defining an element of* $\mathrm{Mult}_{s,r}(V)$ *as in Proposition 15.9. Let* $\{e_1, \dots, e_n\}$ *be a basis of* V *with dual basis* $\{e^1, \dots, e^n\}$ *of* V^*. *Then if we regard* $C^{h,k}(A)$ *also as an element of* $\mathrm{Mult}_{s-1,r-1}(V)$, *one has*

$$\begin{aligned} &C^{h,k}(A)(w_1, \dots, w_{s-1}, q^1, \dots, q^{r-1}) \\ &\qquad = \sum_{i=1}^n A(w_1, \dots, \underset{k\text{th position}}{e_i}, \dots, w_{s-1}, q^1, \dots, \underset{h\text{th position}}{e^i}, \dots q^{r-1}). \end{aligned}$$

Proof. It suffices to prove the equality for decomposable elements. Let us temporarily write add a tilde to denote the multilinear map corresponding to a given tensor. For simplicity assume $(r, s) = (2, 3)$, and assume $A = v_1 \otimes v_2 \otimes p^1 \otimes p^2 \otimes p^3$. Then as in Remark 15.11, the corresponding map $\tilde{A} \in \mathrm{Mult}_{3,2}(V)$ is given by

$$\tilde{A}(w_1, w_2, w_3, q^1, q^2) = q^1(v_1)\, q^2(v_2)\, p^1(w_1)\, p^2(w_2)\, p^3(w_3).$$

Take $(h, k) = (1, 2)$. Then $C^{1,2}(A) = p^2(v_1) \cdot v_2 \otimes p^1 \otimes p^3$. We compute[4]

$$\begin{aligned} C^{1,2}(\tilde{A})(w_1, w_2, q^1) &\overset{\text{def}}{=} \sum_{i=1}^n \tilde{A}(w_1, e_i, w_2, e^i, q^1) \\ &= \sum_{i=1}^n e^i(v_1)\, q^1(v_2)\, p^1(w_1)\, p^2(e_i)\, p^3(w_2) \\ &= \left(\sum_{i=1}^n e^i(v_1)\, p^2(e_i) \right) q^1(v_2)\, p^1(w_1)\, p^3(w_2). \end{aligned}$$

But

$$\sum_{i=1}^n e^i(v_1)\, p^2(e_i) = p^2(v_1),$$

and thus

$$C^{1,2}(\tilde{A}) = \widetilde{C^{1,2}(A)}$$

which is what we wanted to prove. ∎

A contraction $C^{h,k}$ extends to define an operator on tensor fields in an obvious fashion. For instance, if $A \in \mathcal{T}^{2,1}(M)$ is a the decomposable tensor $X \otimes Y \otimes \omega$ for $X, Y \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$ then $C^{1,1}(A) = \omega(X)Y$. Thus we can think of $C^{h,k}$ as sheaf morphism $\mathcal{T}_M^{r,s} \to \mathcal{T}_M^{r-1,s-1}$. If φ is a (possibly local) diffeomorphism then for any contraction $C^{h,k}$ one has

$$\varphi^\star\big(C^{h,k}(A)\big) = C^{h,k}\big(\varphi^\star(A)\big), \tag{18.4}$$

as can be easily checked.

[4]We include the summation signs to minimise the risk of confusion.

Definition 18.14. A **tensor derivation** is a sheaf morphism $\mathcal{D}\colon \mathcal{T}_M \to \mathcal{T}_M$ that preserves type and which in addition satisfies the following two properties:

(i) For any open set $U \subset M$, $\mathcal{D}_U$ commutes with all contractions of $\mathcal{T}(U)$.

(ii) For any open set $U \subset M$, $\mathcal{D}_U$ is a derivation in the sense that

$$\mathcal{D}_U(A \otimes B) = \mathcal{D}_U(A) \otimes B + A \otimes \mathcal{D}_U(B)$$

for all $A, B \in \mathcal{T}(U)$.

Lemma 18.15. *Suppose $\mathcal{D}$ is a tensor derivation on M, suppose $A \in \mathcal{T}^{r,s}(M)$, and suppose $X_1, \ldots, X_s \in \mathfrak{X}(U)$, and $\omega_1, \ldots, \omega_r \in \Omega^1(U)$. Then:*

$$\begin{aligned}\mathcal{D}_U(A)(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) =&\mathcal{D}_U\big(A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)\big) \\ &- \sum_{i=1}^{r} A(\omega_1, \ldots, \mathcal{D}_U(\omega_i), \ldots, \omega_r, X_1, \ldots, X_s) \\ &- \sum_{i=1}^{s} A(\omega_1, \ldots, \omega_r, X_1, \ldots \mathcal{D}_U(X_i), \ldots X_s)\end{aligned} \tag{18.5}$$

Proof. The $(0,0)$-tensor $A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)$ can be thought of as being obtained from the $(r+s, r+s)$ tensor[5] $A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s$ by repeated contractions. We write this symbolically as

$$A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) = C(A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s),$$

where C stands for repeated contractions. The claim now follows by repeatedly using (i) and (ii) from the definition of a tensor derivation. ■

Corollary 18.16. *Suppose $\mathcal{D}$ and $\mathcal{D}'$ are two tensor derivations that agree on functions and vector fields. Then they are identical.*

Proof. Let ω be a 1-form. Then by Lemma 18.15 with $A = \omega$ we see for an arbitrary vector field X that

$$\begin{aligned}\mathcal{D}(\omega)(X) &= \mathcal{D}(\omega(X)) - \omega(\mathcal{D}(X)) \\ &= \mathcal{D}'(\omega(X)) - \omega(\mathcal{D}'(X)) \\ &= \mathcal{D}'(\omega)(X).\end{aligned}$$

Since X was arbitrary, this shows that $\mathcal{D}(\omega) = \mathcal{D}'(\omega)$, and since ω was arbitrary this shows that $\mathcal{D}$ and $\mathcal{D}'$ coincide on tensors of type $(0,1)$. Now for an arbitrary A, observe that (18.5) expands $\mathcal{D}(A)$ in such a way that all the other terms are of the form $\mathcal{D}$ eating a function, a vector field, or a 1-form. Thus $\mathcal{D}(A) = \mathcal{D}'(A)$ for arbitrary A. ■

The next result shows how one can work backwards and build a tensor derivation if we have something defined on functions and vector fields with the appropriate property.

[5] As in Remark 18.1, we don't bother to reorder the factors in this expression.

Proposition 18.17. *Suppose we have a sheaf morphism $\mathcal{D}$ on smooth functions and vector fields which satisfies*

$$\mathcal{D}_U(fg) = \mathcal{D}_U(f)g + f\mathcal{D}_U(g), \qquad \mathcal{D}_U(fX) = \mathcal{D}_U(f)X + f\mathcal{D}_U(X) \tag{18.6}$$

for all $f, g \in C^\infty(U)$ and $X \in \mathfrak{X}(U)$. Then $\mathcal{D}$ extends uniquely to a tensor derivation.

Note the two conditions in (18.6) are forced if we want $\mathcal{D}$ to be a tensor derivation, since $f \otimes g = fg$ and $f \otimes X = fX$ (cf. (15.5)).

Proof. Uniqueness is immediate from the previous corollary, since we have prescribed what $\mathcal{D}$ must do to vector fields and functions. The proof proceeds very similarly to that of Corollary 18.16: the derivation property coupled with our requirement that $\mathcal{D}$ commutes with contractions means that at every stage we have no choice how to proceed.

Namely, we define $\mathcal{D}$ on 1-forms by setting

$$\mathcal{D}_U(\omega)(X) = \mathcal{D}_U(\omega(X)) - \omega(\mathcal{D}_U(X)). \tag{18.7}$$

The hypotheses imply that $\mathcal{D}\colon \mathcal{T}_M^{0,1} \to \mathcal{T}_M^{0,1}$ is a sheaf morphism. Next to define $\mathcal{D}_U$ on $\mathcal{T}^{1,1}(U)$ we start with a tensor of the form $X \otimes \omega$. The derivation property requires us define

$$\mathcal{D}_U(X \otimes \omega) := X \otimes \mathcal{D}_U(\omega) + \mathcal{D}_U(X) \otimes \omega.$$

If $\mathcal{D}_U$ commutes with the contraction $C^{1,1}\colon \mathcal{T}_M^{1,1} \to \mathcal{C}_M^\infty$ then we need

$$\begin{aligned}\mathcal{D}_U(\omega(X)) &= C(\mathcal{D}_U(X \otimes \omega)) \\ &= C(X \otimes \mathcal{D}_U(\omega) + \mathcal{D}_U(X) \otimes \omega) \\ &= \mathcal{D}_U(\omega)(X) + \omega(\mathcal{D}_U(X)),\end{aligned}$$

and this is true by (18.7). This also shows that (18.7) was forced—no other choice would have worked. Now we use the formula from Lemma 18.15 to define $\mathcal{D}$ on all higher tensors. A check similar to the one we just did shows that the resulting object is a derivation that commutes with all contractions. ∎

We now obtain our promised extension of the Lie derivative.

Theorem 18.18. *Let $X \in \mathfrak{X}(M)$. There exists a unique tensor derivation $\mathcal{L}_X\colon \mathcal{T}_M \to \mathcal{T}_M$ that extends the Lie derivative defined on functions and vector fields from Lecture 8.*

Lemma 18.15 tells us how to compute $\mathcal{L}_X(A)$. For instance, if $A \in \mathcal{T}^{0,s}(M)$ then

$$\mathcal{L}_X(A)(Y_1, \ldots, Y_s) = X\big(A(Y_1, \ldots, Y_s)\big) - \sum_{i=1}^{s} A\big(Y_1, \ldots, Y_{i-1}, \mathcal{L}_X(Y_i), Y_{i+1}, \ldots, Y_s\big).$$

Of course this isn't very satisfactory, since we haven't actually given an explicit formula for $\mathcal{L}_X$. We conclude this lecture by remedying this:

Definition 18.19. Let $X \in \mathfrak{X}(M)$ with flow θ_t. Then for any tensor field A, define a new tensor

$$\tilde{\mathcal{L}}_X(A) := \frac{d}{dt}\Big|_{t=0} \theta_t^\star(A) = \lim_{t\to 0} \frac{\theta_t^\star(A) - A}{t} \tag{18.8}$$

The expression is well-defined by Remark 18.10. We now prove:

Proposition 18.20. *Let $X \in \mathfrak{X}(M)$ with flow θ_t. Then for any tensor A, one has:*

$$\mathcal{L}_X(A) = \tilde{\mathcal{L}}_X(A)$$

To prove Proposition 18.20 we use the following result, which will be useful elsewhere.

Proposition 18.21. *Let (r,s), (r',s') and (r'',s'') be three pairs of non-negative integers. Suppose we are given a $\mathcal{C}^\infty_M$-bilinear sheaf homomorphism*

$$\mathcal{A} \colon \mathcal{T}_M^{r,s} \times \mathcal{T}_M^{r',s'} \to \mathcal{T}_M^{r'',s''}.$$

Assume in addition that $\mathcal{A}$ has the property that if $\varphi \colon U \to V$ is a local diffeomorphism between open sets of M then

$$\varphi^\star(\mathcal{A}_V(A,B)) = \mathcal{A}_U(\varphi^\star(A), \varphi^\star(B)). \tag{18.9}$$

Then for every vector field X on M, one has

$$\tilde{\mathcal{L}}_X(\mathcal{A}(A,B)) = \mathcal{A}(\tilde{\mathcal{L}}_X(A), B) + \mathcal{A}(A, \tilde{\mathcal{L}}_X(B)).$$

The proof of Proposition 18.21 is on Problem Sheet J.

Proof of Proposition 18.20. Since $\theta_t^\star = ((\theta_t)^{-1})_\star = (\theta_{-t})_\star$ it follows from the definitions that $\tilde{\mathcal{L}}_X = \mathcal{L}_X$ on functions and vector fields. Thus if we can show that $\tilde{\mathcal{L}}_X$ is a tensor derivation, it will follow from the uniqueness part of of Theorem 18.18 that $\mathcal{L}_X = \tilde{\mathcal{L}}_X$.

Thus we must show that $\tilde{\mathcal{L}}_X$ is a derivation that commutes with contractions. For this we use Proposition 18.21. Taking $\mathcal{A}(A,B) := A \otimes B$ shows that $\tilde{\mathcal{L}}_X$ is a derivation (note (18.9) is satisfied by (18.3)). Similarly taking for instance $\mathcal{A}(A,B) = C^{1,1}(A \otimes B)$ shows that

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

again, (18.9) is satisfied by (18.4). More generally, taking $\mathcal{A}(A,B) = C^{h,k}(A \otimes B)$ shows that $\tilde{\mathcal{L}}_X$ commutes with $C^{h,k}$. This completes the proof. ■

From now on, we will just write $\mathcal{L}_X$ for both the operator $\mathcal{L}_X$ from Theorem 18.18 and the operator defined in (18.8).

Differential forms

In this lecture we study differential forms in depth. Let M be a smooth manifold of dimension n, and let $0 \leq r \leq n$. We denote by $\Omega^r(M) = \Gamma(\bigwedge^r(T^*M))$ the space of sections of the bundle $\bigwedge^r(T^*M) \to M$. An element of $\Omega^r(M)$ is called a **differential r-form** or just a r-form. Similarly if $U \subset M$ is an open set then we denote by $\Omega^r(U) = \Gamma(U, \bigwedge^r(T^*M))$ the sections defined only on U. The space $\Omega^r(U)$ is a module over the ring $C^\infty(U)$.

The assignment $U \mapsto \Omega^r(U)$ is a sheaf of vector spaces on M by Proposition 17.11. We write the sheaf as Ω^r_M. (As with tensors, the subscript M is simply there as a notational hint that we are thinking of this as a sheaf).

Thus for $r = 0$, a differential 0-form is simply a function[1], and for $r = 1$ a differential 1-form is the same thing as a tensor of type $(0, 1)$. The next result is proved in the same way as the Tensor Criterion (Theorem 18.3), and I leave the details to you.

Theorem 19.1 (The Differential Form Criterion). *Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Then there is a canonical identification between $\Omega^r(W)$ and* ***alternating*** *$C^\infty(W)$-multilinear functions*

$$\underbrace{\mathfrak{X}(W) \times \cdots \times \mathfrak{X}(W)}_{r} \to C^\infty(W).$$

Since an alternating multilinear map is (in particular) a multilinear map, we see that any differential r-form may be regarded as a tensor of type $(0, r)$. But for $r \geq 2$, there are (many) multilinear maps that are not alternating, and thus not every tensor of type $(0, r)$ is a differential form. We define

$$\Omega(M) = \bigoplus_{0 \leq r \leq n} \Omega^r(M),$$

with $\Omega(U)$ defined similarly. The sheaf $U \mapsto \Omega(U)$ is denoted by Ω_M. Thus an element of $\Omega(M)$ is a sum $\sum_{i=0}^n \omega_i$ where $\omega_i \in \Omega^i(M)$.

Definition 19.2. If $\omega \in \Omega^r(M)$ and $\vartheta \in \Omega^s(M)$ then the **wedge product** is the differential form $\omega \wedge \vartheta \in \Omega^{r+s}(M)$ defined pointwise by

$$(\omega \wedge \vartheta)(x) = \omega(x) \wedge \vartheta(x)$$

Since $\bigwedge^r(V) = 0$ if $r > \dim V$, the wedge product $\omega \wedge \vartheta$ is zero if $r + s > n$. Note that by part *(ii)* of Proposition 15.21, one has

$$\omega \wedge \vartheta = (-1)^{rs} \vartheta \wedge \omega, \qquad \omega \in \Omega^r(M), \vartheta \in \Omega^s(M)$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] This is because (by definition) $\bigwedge^0(V) = \mathbb{R}$ for any vector space V.

The wedge product gives $\Omega(M)$ the structure of graded ring, and in fact, also a $C^\infty(M)$-graded skew-commutative[2] algebra. Thus Ω_M is a sheaf of graded skew-commutative algebras.

It follows from Corollary 15.25 that if $\sigma\colon U \to O$ is a chart on M then a local frame from $\bigwedge^r(T^*M) \to M$ over U is given by the collection

$$\{dx^{i_1} \wedge \cdots \wedge dx^{i_r} \mid i_1 < \cdots < i_r\}.$$

Thus by Remarks 16.9 and 16.11, we can locally write a differential r form as

$$\omega = \omega_{i_1 \cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r} \tag{19.1}$$

where $\omega_{i_1 \dots i_r} \in C^\infty(U)$. Here is a useful piece of linear algebra, whose proof is on Problem Sheet J. We let $\mathfrak{S}_r$ denote the group of all permutations on r letters.

Definition 19.3. Let $r, s \geq 0$. A **(r,s)-shuffle** is a permutation $\varrho \in \mathfrak{S}_{r+s}$ such that

$$\varrho(1) < \cdots < \varrho(r) \qquad \text{and} \qquad \varrho(r+1) < \cdots < \varrho(r+s).$$

We let $\mathrm{Shuffle}(r,s) \subset \mathfrak{S}(r+s)$ denote the set of all (r,s)-shuffles.

Lemma 19.4. *Let V be a vector space and suppose $\omega \in \bigwedge^r(V^*)$ and $\vartheta \in \bigwedge^s(V^*)$. Let $v_i \in V$ for $i = 1, \dots, r+s$. Then if we identify ω with an element of $\mathrm{Alt}_r(V)$, ϑ with an element of $\mathrm{Alt}_s(V)$ and $\omega \wedge \vartheta$ with an element of $\mathrm{Alt}_{r+s}(V)$, one has:*

$$(\omega \wedge \vartheta)(v_1, \dots, v_{r+s}) = \frac{1}{r!s!} \sum_{\varrho \in \mathfrak{S}_{r+s}} \mathrm{sgn}(\varrho)\omega\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big)\,\vartheta\big(v_{\varrho(r+1)}, \dots, v_{\varrho(r+s)}\big)$$

or equivalently

$$(\omega \wedge \vartheta)(v_1, \dots, v_{r+s}) = \sum_{\varrho \in \mathrm{Shuffle}(r,s)} \mathrm{sgn}(\varrho)\omega\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big)\,\vartheta\big(v_{\varrho(r+1)}, \dots v_{\varrho(r+s)}\big).$$

For low values of r and s this gives an easy method to compute the wedge product of two differential forms. For instance, we have:

Corollary 19.5. *Let $\omega, \vartheta \in \Omega^1(M)$ denotes two 1-forms. Then*

$$(\omega \wedge \vartheta)_x(v, w) = \omega_x(v)\vartheta_x(w) - \omega_x(w)\vartheta_x(v), \qquad \forall\, x \in M,\ \forall\, v, w \in T_xM.$$

A smooth map $\varphi\colon M \to N$ can pull back differential forms in the same way that it pulls back tensors of type $(0, r)$.

Definition 19.6. Let $\varphi\colon M \to N$ denote a smooth map. Given $\omega \in \Omega^r(N)$, we define the **pullback** form $\varphi^\star(\omega) \in \Omega^r(M)$ by

$$\varphi^\star(\omega)_x(v_1, \dots, v_r) := \omega_{\varphi(x)}(D\varphi(x)[v_1], \dots D\varphi(x)[v_r]).$$

This defines a map $\varphi^\star\colon \Omega^r(N) \to \Omega^r(M)$, and thus also a map $\varphi^\star\colon \Omega(N) \to \Omega(M)$. The next lemma tells us that $\varphi^\star$ is an algebra homomorphism.

[2] The "skew-commutative" refers to the sign $(-1)^{rs}$.

Lemma 19.7. *If* $\varphi\colon M \to N$ *is a smooth map and* $\omega, \vartheta \in \Omega(N)$ *then*

$$\varphi^\star(\omega \wedge \vartheta) = \varphi^\star(\omega) \wedge \varphi^\star(\vartheta)$$

Proof. Immediate from Lemma 19.4 and the definition. ∎

Note also that (just as with tensors), the pullback operation is functorial: if $\varphi\colon M \to N$ and $\psi\colon L \to M$ then

$$(\varphi \circ \psi)^\star = \psi^\star \circ \varphi^\star \tag{19.2}$$

as maps $\Omega(N) \to \Omega(L)$.

(♣) Remark 19.8. We can phrase this abstractly as the following statement: *There is a contravariant functor* Man $\to$ Algebras *that assigns to a manifold* M *the algebra* $\Omega(M)$ *of differential forms, and assigns to a smooth map* $\varphi\colon M \to N$ *the algebra homomorphism* $\varphi^\star\colon \Omega(N) \to \Omega(M)$.

Definition 19.9. Since any differential form $\omega \in \Omega^r(M)$ can be thought of a tensor of type $(0, r)$, we can apply the Lie derivative $\mathcal{L}_X$ to it to obtain another tensor of type $(0, r)$, denoted by $\mathcal{L}_X(\omega)$. In fact, from Lemma 18.15 the tensor $\mathcal{L}_X(\omega)$ is easily seen to be alternating, and hence $\mathcal{L}_X(\omega)$ is another differential r-form. Explicitly, by (18.5) one has

$$\mathcal{L}_X(\omega)(Y) = X(\omega(Y)) - \omega([X, Y]), \qquad \omega \in \Omega^1(M),\ X, Y \in \mathfrak{X}(M), \tag{19.3}$$

and more generally

$$\mathcal{L}_X(\omega)(X_1, \dots, X_r) = X(\omega(X_1, \dots, X_r)) - \sum_{i=1}^r \omega(X_1, \dots, [X, X_i], \dots X_r) \tag{19.4}$$

for $\omega \in \Omega^r(M)$ and $X, X_1, \dots, X_r \in \mathfrak{X}(M)$. Thus the Lie derivative defines an operator $\mathcal{L}_X\colon \Omega^r(M) \to \Omega^r(M)$, and hence also an operator $\mathcal{L}_X\colon \Omega(M) \to \Omega(M)$.

Here is how the Lie derivative behaves with respect to the wedge product.

Lemma 19.10. *Let* $\omega, \vartheta \in \Omega(M)$. *Then*

$$\mathcal{L}_X(\omega \wedge \vartheta) = \mathcal{L}_X(\omega) \wedge \vartheta + \omega \wedge \mathcal{L}_X(\vartheta).$$

Proof. Apply[3] Proposition 18.21 with $\mathcal{A}(\omega, \vartheta) = \omega \wedge \vartheta$. ∎

Remark 19.11. The Lie derivative $\mathcal{L}_X$ gives us a way to "differentiate" a tensor field (or a differential form) with respect to a vector field, but it does *not* allows us differentiate a tensor field (or differential form) with respect to a single tangent vector. Indeed, the value of $\mathcal{L}_X(A)$ at a point x depends on the values of X on a whole neighbourhood of x, not just on $X(x)$. This is because $X \mapsto \mathcal{L}_X$ is

[3]Strictly speaking, this is a slight modification of Proposition 18.21 for differential forms instead of tensors, but the proof is exactly the same.

not $C^\infty(M)$-linear. For instance, if A is a 1-form ω then for $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$,

$$\begin{aligned}\mathcal{L}_{fX}(\omega)(Y) &= (fX)\omega(Y) - \omega([fX, Y]) \\ &\overset{(\dagger)}{=} fX(\omega(Y)) - \omega(f[X,Y] - Y(f)X) \\ &= f\mathcal{L}_X(\omega)(Y) + Y(f)\omega(X),\end{aligned}$$

where ($\dagger$) used Problem D.4. The presence of the "error term" $Y(f)\omega(X)$ shows that $\mathcal{L}_{fX} \neq f\mathcal{L}_X$.

The first topic we will cover next semester will be *connections* in vector bundles and principal bundles (the latter will be defined in Lecture 24). A connection on the tangent bundle TM induces a *covariant derivative* ∇_X on tensor fields associated to every vector field X. As we will see, a covariant derivative ∇_X will have the nice property that $\nabla_{fX} = f\nabla_X$, and thus this will give us a notion of differentiation that works pointwise. The downside it that it requires a choice of extra structure (namely, a connection). Meanwhile the Lie derivative is canonical.

Differential forms are typically more important than tensors in geometry for two key reasons:

- We can differentiate them.
- We can integrate them.

We will discuss differentiation in this lecture. Integration will be covered next week. Let us motivate this by considering the special case of a 0-form, i.e. a smooth function. If $f \in C^\infty(M)$ and X is a vector field, then the Lie derivative $\mathcal{L}_X(f) = X(f) = df(X)$ can be thought of a "directional derivative" of f. However it is not what we would actually call the "derivative" of f: that would be the map $Df\colon TM \to T\mathbb{R}$, or equivalently, the differential[4] of f, which is the 1-form $df \in \Omega^1(M)$. This tells us that the differential of a 0-form is a 1-form. Generalising this, we will define the differential of a r-form ω to be a $(r+1)$-form $d\omega$.

As with tensor derivations, we will think of the differential as a sheaf morphism. Let us start with the following general definition.

Definition 19.12. Let M be a smooth manifold and let $h \in \mathbb{Z}$. A **graded derivation of degree** h on M is an $\mathbb{R}$-linear sheaf morphism $\mathcal{D}\colon \Omega_M \to \Omega_M$ which satisfies:

- If $\omega \in \Omega^r(U)$ then $\mathcal{D}_U(\omega) \in \Omega^{r+h}(U)$.
- If $\omega \in \Omega^r(U)$ and $\vartheta \in \Omega(U)$ then

$$\mathcal{D}_U(\omega \wedge \vartheta) = \mathcal{D}_U(\omega) \wedge \vartheta + (-1)^{hr}\omega \wedge \mathcal{D}_U(\vartheta). \tag{19.5}$$

The first condition tells us that $\mathcal{D}$ restricts to define sheaf morphisms

$$\mathcal{D}\colon \Omega^r_M \mapsto \Omega^{r+h}_M.$$

The second condition can be phrased as saying that $\mathcal{D}$ should be a sheaf morphism of graded skew-commutative algebras.

[4]Recall from Remark 4.13 that the difference between the derivative Df and the differential df is essentially just notation, and indeed many authors denote them both by the same letter.

EXAMPLE 19.13. The Lie derivative $\mathcal{L}_X$ is a graded derivation of degree 0 by Lemma 19.10.

DEFINITION 19.14. A 1-form $\omega \in \Omega^1(U)$ is called **exact** if $\omega = df$ for some $f \in C^\infty(U)$.

Just as a tensor derivation is entirely determined by what it does to functions and vector fields, a graded derivation is entirely determined by what it does to functions and exact 1-forms.

PROPOSITION 19.15. *Suppose $\mathcal{D}$ and $\mathcal{D}'$ are two graded derivations of the same degree h. If $\mathcal{D}$ and $\mathcal{D}'$ agree on functions and exact 1-forms then $\mathcal{D} = \mathcal{D}'$.*

Proof. Since a graded derivation is a sheaf morphism, it is entirely determined by all its restrictions $\mathcal{D}_U$ where $U \subset M$ is the domain of a chart $\sigma : U \to O$. If x^i are the local coordinates of σ then by (19.1), any $\omega \in \Omega^r(U)$ can we written as a sum of elements of the form

$$f\, dx^{i_1} \wedge \cdots \wedge dx^{i_r}.$$

Since $\mathcal{D}$ is $\mathbb{R}$-linear, $\mathcal{D}_U$ is determined by what it does to such a term. But by repeatedly applying (19.5), we see that $\mathcal{D}_U(f dx^{i_1} \wedge \cdots \wedge dx^{i_r})$ is determined by $\mathcal{D}_U(f)$ and $\mathcal{D}_U(dx^{i_j})$. Thus if two graded derivations agree on functions and exact 1-forms then they are identical. ■

REMARK 19.16. Suppose $\mathcal{D}$ and $\mathcal{D}'$ are two graded derivations of degrees h and k respectively. Then

$$\mathcal{D} \circ \mathcal{D}' - (-1)^{hk} \mathcal{D}' \circ \mathcal{D}$$

is another graded derivation of degree $h + k$.

Here is the main result of today's lecture.

THEOREM 19.17 (The exterior differential). *Let M be a smooth manifold of dimension n. There is a unique graded derivation $d : \Omega_M \to \Omega_M$ of degree 1 such that:*

- *If $f \in C^\infty(U)$ then $d_U(f) = df \in \Omega^1(U)$,*
- *$d \circ d = 0$, i.e. $d_U(d_U \omega) = 0$ for any $\omega \in \Omega^r(U)$.*

We call d the **exterior differential operator** and refer to $d\omega$ as the **exterior differential** of ω (often shortened to the just "the differential of ω"). Our proof of Theorem 19.17 will construct d in coordinates. We will give a coordinate-free expression for d in the next lecture (see Theorem 20.7).

Proof. We prove the result in three steps.

1. We first deal with uniqueness. This is immediate from Proposition 19.15, since the first bullet point defines d on functions, and the second bullet point defines d on exact 1-forms (namely: it's identically zero on exact 1-forms).

2. To construct d it suffices to define $d_U\colon \Omega^r(U) \to \Omega^{r+1}(U)$ for any open set U which is the domain of a chart $\sigma\colon U \to O$. In this step we will define an operator $d_{U,\sigma}\colon \Omega^r(U) \to \Omega^{r+1}(U)$ which satisfies the two bullet points but depends on the choice of chart σ. In the last step we will show that in fact $d_{U,\sigma}$ does not depend on σ, which thus completes the proof.

Let x^i denote the local coordinates of σ. To ease the notation we adopt the following shorthand: if $I = (i_1, \dots, i_r)$ is a subset of $\{1, \dots, n\}$ with $i_j < i_{j+1}$ for each $j = 1, \dots, r-1$ then we set:

$$dx^I := dx^{i_1} \wedge \cdots \wedge dx^{i_r}.$$

We also define $dx^I := 1$ if $I = \emptyset$. Thus any $\omega \in \Omega(U)$ can be written as a sum

$$\omega = \sum_I f_I \, dx^I.$$

We define

$$d_{U,\sigma}\omega := \sum_I df_I \wedge dx^I.$$

If $\omega \in \Omega^r(U)$ then $d_{U,\sigma}\omega \in \Omega^{r+1}(U)$. Moreover $d_{U,\sigma}$ is obviously $\mathbb{R}$-linear and satisfies the first bullet point by definition. Thus we need only check that $d_{U,\sigma}(d_{U,\sigma}\omega) = 0$ and that (19.5) holds.

To establish (19.5), we may assume $\omega = f\, dx^I$ and $\vartheta = g\, dx^J$. If I and/or J are empty then (19.5) follows from the Leibniz rule $d(fg) = fdg + gdf$. In the general case we argue as follows. Assume ω has degree r. Then:

$$\begin{aligned}
d_{U,\sigma}(\omega \wedge \vartheta) &= d(fg\, dx^I \wedge dx^J) \\
&= d(fg) \wedge dx^I \wedge dx^J) \\
&= (fdg + gdf) \wedge dx^I \wedge dx^J \\
&= \big(df \wedge dx^I\big) \wedge (g\, dx^J) + (-1)^r (f\, dx^I) \wedge \big(dg \wedge dx^J\big) \\
&= d_{U,\sigma}\omega \wedge \vartheta + (-1)^r \omega \wedge d_{U,\sigma}\vartheta.
\end{aligned}$$

To see that $d_{U,\sigma}(d_{U,\sigma}\omega) = 0$ we first show that $d_{U,\sigma}(df) = 0$ for any function $f \in C^\infty(U)$. For this write $df = \frac{\partial f}{\partial x^i}\, dx^i$ (cf. Definition 7.4). Then

$$d_{U,\sigma}(df) = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \wedge dx^j,$$

where we abbreviate $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^i}\right)$. But by elementary calculus, $\frac{\partial^2 f}{\partial x^i \partial x^j}$ is symmetric in i and j, whereas $dx^i \wedge dx^j$ is anti-symmetric. Thus the sum cancels.

Next, if $f, g \in C^\infty(U)$ then using (19.5) we have

$$d_{U,\sigma}(df \wedge dg) = d_{U,\sigma}(df) \wedge dg - df \wedge d_{U,\sigma}(dg) = 0,$$

and more generally if $f, f_1, \dots f_r \in C^\infty(U)$ are any functions then

$$d_{U,\sigma}(df \wedge df_1 \wedge \cdots \wedge df_r) = 0.$$

Applying this with $f_j = x^{i_j}$ and using $\mathbb{R}$-linearity shows that $d_{U,\sigma}(d_{U,\sigma}(\omega)) = 0$ for any $\omega \in \Omega^r(U)$.

3. We now address the dependence on σ. Suppose τ is another chart defined on an open set V such that $U \cap V \neq \emptyset$. We must show that

$$d_{U\cap V, \sigma|_{U\cap V}} \equiv d_{U\cap V, \tau|_{U\cap V}}.$$

By Proposition 19.15 applied to the graded derivations $d_{U\cap V, \sigma|_{U\cap V}}$ and $d_{U\cap V, \tau|_{U\cap V}}$ on $\Omega(U\cap V)$, it suffices to show they agree on functions and on exact 1-forms. But this is clear: if $f \in C^\infty(U \cap V)$ then one has

$$d_{U\cap V, \sigma|_{U\cap V}}(f) = d_{U\cap V, \tau|_{U\cap V}}(f) = df,$$

and the argument we just gave above showed that

$$d_{U\cap V, \sigma|_{U\cap V}}(df) = d_{U\cap V, \tau|_{U\cap V}}(df) = 0.$$

This completes the proof. ∎

DEFINITION 19.18. A differential form ω is said to be **closed** if $d\omega = 0$. A differential form ω is said to be **exact** if $\omega = d\vartheta$ for some ϑ (this extends Definition 19.14 to r-forms for $r > 1$. Since $d \circ d = 0$, any exact form is closed, but the converse is typically false. One denotes the quotient vector space by

$$H^r_{\mathrm{dR}}(M) := \frac{\{\text{closed } r\text{-forms}\}}{\{\text{exact } r\text{-forms}\}},$$

where the "dR" stands for "de Rham". An element of $H^r_{\mathrm{dR}}(M)$ is written as $[\omega]$, where ω is a closed r-form. Thus by definition

$$[\omega] = [\omega + d\vartheta].$$

We call $H^r_{\mathrm{dR}}(M)$ the **rth de Rham cohomology**[5] **group**[6] of M. In Lecture 23 we will see that the de Rham groups are a topological invariant (cf. Remark 23.19). In Lecture 27 we will strengthen this and prove that the de Rham cohomology groups actually agree with the singular cohomology groups:

$$H^r_{\mathrm{dR}}(M) \cong H^r(M; \mathbb{R}) \qquad \text{(singular cohomology with coefficients in } \mathbb{R})$$

and thus in particular are a topological invariant of M. (*Remark:* All of Lecture 27 is non-examinable!)

LEMMA 19.19. *Let $\varphi \colon M \to N$ be a smooth map and let $\omega \in \Omega(N)$. Then*

$$\varphi^\star(d\omega) = d(\varphi^\star(\omega)).$$

that is, $\varphi^\star$ commutes with the exterior differentials.

---

[5]Do not be scared by the word "cohomology" if you are not familiar with algebraic topology. As far as this course is concerned, all that is important is that $H^r_{\mathrm{dR}}(M)$ is a quotient vector space.

[6]Even though the de Rham groups are actually vector spaces (and not just abelian groups), it is still common to refer to them as the "de Rham cohomology groups".

Proof. We first prove the lemma for the special case of 0-forms, i.e functions. Let $f \in C^\infty(N)$ and $X \in \mathfrak{X}(M)$. Then

$$\varphi^\star(df)(X) = df(D\varphi[X]) = D\varphi[X](f) = X(f \circ \varphi) = d(f \circ \varphi)(X) = d(\varphi^\star f)(X).$$

For the general case, since both sides are $\mathbb{R}$-linear sheaf morphisms, it suffices to work in a chart domain and assume ω is of the form $f\, dx^{i_1} \wedge \cdots \wedge dx^{i_r}$. Then since we already know that $\varphi^\star$ is an algebra homomorphism (Lemma 19.7) and we already proved the result for functions:

$$\begin{aligned}
\varphi^\star(d\omega) &= \varphi^\star\big(df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_r}\big) \\
&= \varphi^\star(df) \wedge \varphi^\star(dx^{i_1}) \wedge \cdots \wedge \varphi^\star(dx^{i_r}) \\
&= d(\varphi^\star(f)) \wedge d(\varphi^\star(x^{i_1})) \wedge \cdots \wedge d(\varphi^\star(x^{i_r})) \\
&= d\big(\varphi^\star(f) \wedge d(\varphi^\star(x^{i_1})) \wedge \cdots \wedge d(\varphi^\star(x^{i_r}))\big) \\
&= d(\varphi^\star(\omega)).
\end{aligned}$$

■

Corollary 19.20. *If $\varphi \colon M \to N$ is a smooth map then $\varphi^\star$ induces a well-defined map (also denoted by) $\varphi^\star \colon H^r_{\mathrm{dR}}(N) \to H^r_{\mathrm{dR}}(M)$ via:*

$$[\omega] \mapsto [\varphi^\star(\omega)].$$

Proof. By Lemma 19.19, $\varphi^\star$ maps closed forms to closed forms and exact forms to exact forms. ■

(♣) Remark 19.21. We have thus created a contravariant functor $\mathsf{Man} \to \mathsf{Vect}$ that assigns to a manifold M its rth de Rham cohomology group $H^r_{\mathrm{dR}}(M)$ and assigns to smooth map $\varphi \colon M \to N$ the induced map $\varphi^\star \colon H^r_{\mathrm{dR}}(N) \to H^r_{\mathrm{dR}}(M)$.

We conclude this lecture by relating the Lie derivative to the exterior differential.

Proposition 19.22. *Let M be a smooth manifold and fix $X \in \mathfrak{X}(M)$. Then $d \circ \mathcal{L}_X = \mathcal{L}_X \circ d$.*

Proof. We first prove the result for functions. For a function f and a vector field Y, one has by (19.3) that

$$\mathcal{L}_X(df)(Y) = X(Y(f)) - [X, Y](f) = Y(X(f)) = d(X(f))(Y) = d(\mathcal{L}_X(f))(Y).$$

Since Y was arbitrary, this shows that $\mathcal{L}_X(df) = d(\mathcal{L}_X(f))$. For the general case, by Remark 19.16, $d \circ \mathcal{L}_X - \mathcal{L}_X \circ d$ is a graded derivation of degree $+1$. By Proposition 19.15, if we can show it vanishes on functions and exact 1-forms then it is identically zero. We just did the case for functions, and for an exact 1-form we have

$$d(\mathcal{L}_X(df)) - \mathcal{L}_X(d(df)) = d(\mathcal{L}_X(df)) - 0 = d(d(\mathcal{L}_X(f)) = 0.$$

This completes the proof. ■

LECTURE 20

Cartan's Magic Formula and orientability

We begin this lecture by stating and proving **Cartan's Magic Formula**[1]. This formula relates the Lie derivative $\mathcal{L}_X$, the exterior differential, and a third operation, the (as yet undefined) **interior product** i_X. We thus begin by defining the interior product, and, as usual, we start at the level of linear algebra.

DEFINITION 20.1. Let V be a vector space, and fix $v \in V$. Define $i_v \colon \bigwedge^r(V^*) \to \bigwedge^{r-1}(V^*)$ by declaring that on decomposable elements $p^1 \wedge \cdots \wedge p^r$ that

$$i_v(p^1 \wedge \cdots \wedge p^r) = \sum_{i=1}^{r} (-1)^{i+1} p^i(v) \cdot p^1 \wedge \cdots \wedge p^{i-1} \wedge p^{i+1} \wedge \cdots \wedge p^r,$$

and then extending by linearity.

Straight from the definition, we see that:

LEMMA 20.2. *Let $\omega \in \bigwedge^r(V^*)$ and $\vartheta \in \bigwedge^s(V^*)$. Then*

$$i_v(\omega \wedge \vartheta) = i_v(\omega) \wedge \vartheta + (-1)^r \omega \wedge i_v(\vartheta).$$

The following result is slightly less clear, and is left as an enjoyable exercise.

LEMMA 20.3. *Let $v \in V$ and let $\omega \in \bigwedge^r(V^*)$. If we regard both ω and $i_v(\omega)$ as elements of $\mathrm{Alt}_r(V)$ and $\mathrm{Alt}_{r-1}(V)$ respectively (via Proposition 15.23), then*

$$i_v(\omega)(v_1, \dots, v_{r-1}) = \omega(v, v_1, \dots, v_{r-1}).$$

Note this shows that $i_v \circ i_v = 0$. We now transfer this to manifolds.

PROPOSITION 20.4. *Let M be a smooth manifold and let $X \in \mathfrak{X}(M)$. There is a graded derivation $i_X \colon \Omega_M \to \Omega_M$ of degree -1 defined by*

$$i_X(\omega)(X_1, \dots, X_{r-1}) := \omega(X, X_1, \dots, X_{r-1}), \qquad \omega \in \Omega^r(M),\ X_i \in \mathfrak{X}(M)$$

if $r \geq 1$ and $i_X(f) := 0$. One has $i_X \circ i_X = 0$. This is the unique graded derivation of degree -1 such that $i_X(\omega) = \omega(X)$ for ω a 1-form and $i_X(f) = 0$ for f a function.

The proof is immediate; uniqueness follows from Proposition 19.15.

COROLLARY 20.5. *Let $X, Y \in \mathfrak{X}(M)$. Then $i_{[X,Y]} = \mathcal{L}_X \circ i_Y - i_Y \circ \mathcal{L}_X$ as morphisms on Ω_M.*

Proof. Both sides are graded derivations of degree -1 by Remark 19.16. Thus it suffices to check on functions and exact 1-forms. For functions both sides are zero. For an exact 1-form df this follows from (19.3) applied with $\omega = df$. ■

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] I didn't make this name up. Seriously.

Here then is the promised magical forumla:

THEOREM 20.6 (Cartan's Magic Formula). *Let $X \in \mathfrak{X}(M)$. Then*

$$\mathcal{L}_X = d \circ i_X + i_X \circ d.$$

Proof. Again, both sides are graded derivations of degree 0 by Remark 19.16. Thus by Proposition 19.15 we need only check they agree on functions and exact 1-forms. On functions this follows from Lemma 8.23 and on exact 1-forms this was Proposition 19.22. ■

Let us use Cartan's Magic Formula to give a coordinate free definition of the exterior differential d.

THEOREM 20.7. *Let M be a smooth manifold, $\omega \in \Omega^r(M)$ and $X_0, \dots X_r \in \mathfrak{X}(M)$. Then:*

$$\begin{aligned} d\omega(X_0, \dots, X_r) = &\sum_{i=0}^{r} (-1)^i X_i\big(\omega(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \sum_{0 \le i < j \le r} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_r). \end{aligned}$$

Here and elsewhere, the caret $\widehat{X}_i$ means that the X_i term should be omitted.

Proof. One has $d\omega(X_0, \dots, X_r) = i_{X_0}(d\omega)(X_1, \dots, X_r)$, which by Cartan's Magic Formula is equal to

$$\mathcal{L}_{X_0}(\omega)(X_1, \dots, X_r) - d(i_{X_0}(\omega))(X_1, \dots, X_r). \tag{20.1}$$

We now argue by induction on r. If $r = 1$ then by (19.3) this becomes

$$\mathcal{L}_{X_0}(\omega(X_1)) - \omega([X_0, X_1]) - d(\omega(X_0))(X_1) = X_0(\omega(X_1)) - X_1(\omega(X_0)) - \omega([X_0, X_1]),$$

which is what we want. Now assume $r \ge 2$ and that the result is true for all forms of degree $r - 1$. By (19.4) the first term of (20.1) is equal to

$$X_0(\omega(X_1, \dots, X_r)) - \sum_{i=1}^{r} \omega(X_1, \dots, [X_0, X_i], \dots, X_r).$$

By induction, we have that $d(i_{X_0}(\omega))(X_1, \dots, X_r)$ is equal to

$$\begin{aligned} &\sum_{i=1}^{r} (-1)^{i-1} X_i\big((i_{X_0}(\omega)(X_1, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &\qquad + \sum_{1 \le i < j \le r} (-1)^{i+j} \big(i_{X_0}(\omega)([X_i, X_j], X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_r)\big). \end{aligned}$$

Putting this into (20.1) and checking the signs carefully gives the desired result. ■

We now move onto a somewhat different topic and discuss *orientations* of vector bundles. This is the first of two preliminary topics we need to cover (the second is manifolds with boundary) before we can state and prove *Stokes' Theorem*, which is about integrating differential forms on oriented manifolds with boundary.

As usual, let us start at the level of linear algebra. Of course, you have all known since kindergarten what an orientation of a vector space is, but perhaps you haven't seen it in the "right" language yet.

Definition 20.8. Let V be a one-dimensional vector space. Then $V \setminus \{0\}$ has two components. An **orientation** of V is a choice of one of these components, which one then labels as "positive". The other component is then labelled "negative". A **positive basis** of V is a choice of any non-zero vector belonging to the positive component. A **negative basis** of V is a choice of any non-zero vector belonging to the negative component.

Example 20.9. The **standard orientation** of $\mathbb{R}$ is given by declaring that the positive numbers are (surprise!) the positive component of $\mathbb{R} \setminus \{0\}$.

Definition 20.10. Let V be a vector space. We will use the notation $\det V$ to mean $\bigwedge^n(V)$ where $n = \dim V$. One calls $\det V$ the **determinant** of V. From Proposition 15.25, the space $\det V$ is a one-dimensional vector space. Moreover if $\{e_1, \dots, e_n\}$ is a basis for V then $e_1 \wedge \cdots \wedge e_n$ is a basis of $\det V$.

Definition 20.11. Let V be a vector space of positive dimension. An **orientation** on V is a choice of orientation on $\det V$. Thus there are exactly two orientations. A basis $\{e_1, \dots, e_n\}$ of V is said to be **positive** if $e_1 \wedge \cdots \wedge e_n$ is a positive basis of $\det V$. If instead $e_1 \wedge \cdots \wedge e_n$ is a negative basis of $\det V$ then $\{e_1, \dots, e_n\}$ is a negative basis of V.

Example 20.12. If e_i denotes the standard ith basis vector in $\mathbb{R}^n$ then the **standard orientation** of $\mathbb{R}^n$ is given by declaring the $e_1 \wedge \cdots \wedge e_n$ is a positive basis of $\det \mathbb{R}^n$ (and hence $\{e_1, \dots, e_n\}$ is a positive basis of $\mathbb{R}^n$.)

Remark 20.13. You are probably more used to thinking of the determinant of a linear transformation, rather than the determinant of a vector space itself. The motivation for this terminology goes as follows. Suppose that V and W are vector spaces of the same dimension n. Since $\bigwedge^n$ is a functor, if $T \colon V \to W$ is a linear map then we get a induced linear map $\bigwedge^n(T) \colon \det V \to \det W$, defined explicitly by

$$\bigwedge^n(T)(v_1 \wedge \cdots \wedge v_n) := Tv_1 \wedge \cdots \wedge Tv_n.$$

This is a linear map between two one-dimensional vector spaces, and hence is multiplication by a scalar. This scalar is non-zero if and only if T is an isomorphism. In general the precise value of this scalar depends on a choice of basis of V and W, but the linear map $\bigwedge^n(T)$ itself clearly does not. If T is an isomorphism and V and W are oriented, then we say that T is **orientation-preserving** if $\bigwedge^n(T)$ maps the positive component of $\bigwedge^n(V)$ to the positive component of $\bigwedge^n(W)$. Otherwise T is **orientation-reversing**.

If $V = W$ and we choose the same basis for both the domain V and the target V then the value of the scalar is independent of the basis. In this case, it is common to call the scalar the **determinant of** T. Explicitly, if $\{e_1, \dots, e_n\}$ is our chosen basis then

$$Te_1 \wedge \cdots \wedge Te_n = (\det T) \cdot e_1 \wedge \cdots \wedge e_n. \tag{20.2}$$

In any case, we can think of the determinant as a functor $\mathsf{det}\colon \mathsf{Vect}^{>0} \to \mathsf{Vect}^1$ from positive-dimensional vector spaces to one-dimensional vector spaces that assigns to a vector space its determinant space $\det V$, and on morphisms if $T\colon V \to W$ is a linear map then $\mathsf{det}(T)\colon \det V \to \det W$ is the linear map given by:

$$\mathsf{det}(T) = \begin{cases} \bigwedge^n(T), & \text{if } \dim V = \dim W = \text{n}, \\ \text{the zero map}, & \text{if } \dim V \neq \dim W. \end{cases}$$

Exercise: Check this new definition of determinant coincides with the one you are used to from linear algebra. Use this to give slicker proofs of everything you learnt in your linear algebra course. (For example: the fact that $\det(S \circ T) = \det S \cdot \det T$ is immediate from the fact that det is a functor.) Conclude that you should have been taught Category Theory in the Basisjahr.

Remark 20.14. If V is a vector space then an orientation on V canonically determines an orientation on the dual space V^* by declaring that the dual basis to a positive basis is positive.

Now we move to vector bundles. In general a vector bundle of rank one is often called[2] a **line bundle**.

Definition 20.15. Let E be a vector bundle over M. The **determinant line bundle** associated to E is the vector bundle $\det E \to M$ of rank one whose fibre over $x \in M$ is $\det E_x$.

Roughly speaking, a vector bundle $\pi\colon E \to M$ is oriented if each fibre E_x is given an orientation which depends smoothly on x. To make this precise, we prove the following result. Recall the notion of the structure group of a vector bundle, cf. Remark 13.14.

Proposition 20.16. *Let $\pi\colon E \to M$ be a vector bundle of rank k over M. The following are equivalent:*

(i) *There is a smooth nowhere vanishing section $\mu \in \Gamma(\det E^*)$.*

(ii) *It is possible to reduce the structure group of E from $\mathrm{GL}(k)$ to $\mathrm{GL}^+(k)$.*

(iii) *The bundle $\det E^* \to M$ is a trivial bundle.*

Proof. We prove the result in three steps.

[2] **Warning:** This terminology is very popular in complex geometry and algebraic geometry too. But typically there people are working with *complex* vector bundles, not real vector bundles. A complex line bundle is (in particular) a two-dimensional real vector bundle. So when taken out of context, beware of the phrase "line bundle" since it may either be referring to a one-dimensional real bundle or a one-dimensional complex bundle.

1. We first prove (i) implies (ii). Let $\{\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to \mathbb{R}^k \mid \mathtt{a} \in \mathtt{A}\}$ be a vector bundle atlas for E. We may assume each $U_{\mathtt{a}}$ is connected. For each $\mathtt{a} \in \mathtt{A}$, we obtain a local frame $\{e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}}\}$ for E over $U_{\mathtt{a}}$ via (16.3). Since μ is non-vanishing, for each $\mathtt{a} \in \mathtt{A}$ the function[3]

$$\mu(e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}}) \colon U_{\mathtt{a}} \to \mathbb{R} \tag{20.3}$$

is either everywhere positive or everywhere negative. If for a given $\mathtt{a}$ one has that (20.3) is positive, we do nothing. If instead for a given $\mathtt{a}$ one has that (20.3) is negative, we replace the local frame $\{e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}}\}$ with the new one $\{-e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}}\}$, and then replace $\alpha_{\mathtt{a}}$ with the vector bundle chart corresponding to this new frame, cf (16.4). Having done this, we may assume that (20.3) is a positive function for every $\mathtt{a} \in \mathtt{A}$.

We claim that our new bundle atlas (still denoted by) $\{\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to \mathbb{R}^k \mid \mathtt{a} \in \mathtt{A}\}$ has its structure group contained in $\mathrm{GL}^+(k)$. Indeed, if $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$ then we can write

$$e_j^{\mathtt{b}}(x) = A_j^i(x) e_i^{\mathtt{a}}(x), \qquad x \in U_{\mathtt{a}} \cap U_{\mathtt{b}}$$

for $A_j^i \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathbb{R}$ a smooth function. In fact, unravelling the definition shows that $A_j^i(x)$ is the (i,j)th entry of the transition matrix $\rho_{\alpha_{\mathtt{b}}\alpha_{\mathtt{a}}}(x)$. Thus for any $x \in U_{\mathtt{a}} \cap U_{\mathtt{b}}$, we have by (15.8) that:

$$\mu(e_1^{\mathtt{b}}, \dots, e_k^{\mathtt{b}})(x) = \Big(\det \rho_{\alpha_{\mathtt{b}}\alpha_{\mathtt{a}}}(x)\Big) \cdot \mu(e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}})(x).$$

Thus $\det \rho_{\alpha_{\mathtt{a}}\alpha_{\mathtt{b}}}(x) > 0$ for all $x \in U_{\mathtt{a}} \cap U_{\mathtt{b}}$, which is what we wanted to prove.

2. Now we show that (ii) implies (i). For this we start with a vector bundle atlas $\{\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to \mathbb{R}^k \mid \mathtt{a} \in \mathtt{A}\}$ with structure group in $\mathrm{GL}^+(k)$ and we have to build a section μ. Let $\{\varepsilon_{\mathtt{a}}^1, \dots, \varepsilon_{\mathtt{a}}^k\}$ denote the dual frame to the local frame $\{e_1^{\mathtt{a}}, \dots, e_k^{\mathtt{a}}\}$ associated to $\alpha_{\mathtt{a}}$, and let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote a partition of unity subordinate to the open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$. We now define

$$\mu \colon M \to \det E^*, \qquad \mu := \sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}} \, \varepsilon_{\mathtt{a}}^1 \wedge \cdots \wedge \varepsilon_{\mathtt{a}}^k.$$

We need only check that μ is nowhere vanishing. Fix $x \in M$. Then there exists $\mathtt{b} \in \mathtt{A}$ such that $x \in U_{\mathtt{b}}$. We evaluate μ on $e_1^{\mathtt{b}} \wedge \cdots \wedge e_k^{\mathtt{b}}$ at x to discover

$$\mu_x(e_1^{\mathtt{b}}(x), \dots, e_k^{\mathtt{b}}(x)) = \sum_{\mathtt{a} \in \mathtt{A}} \Big(\det \rho_{\alpha_{\mathtt{b}}\alpha_{\mathtt{a}}}(x)\Big) \cdot \lambda_{\mathtt{a}}(x),$$

which is positive as desired.

3. Finally, since $\det E^*$ is a one-dimensional vector bundle, it is trivial if and only if it admits a nowhere vanishing section (this is a special case of Corollary 16.7). Thus (i) and (iii) are obviously equivalent. This completes the proof. ∎

We now use Proposition 20.16 to define precisely what it means for a vector bundle to be orientable.

[3]Explicitly, this is the function given by $x \mapsto \mu_x(e_1(x), \dots, e_k(x))$, where we view $\mu_x \in \bigwedge^k(E_x^*)$ as an element of $\mathrm{Alt}_k(E_x)$ via Proposition 15.23.

Definition 20.17. Let $\pi\colon E \to M$ be a vector bundle. We say that E is **orientable** if either of the three conditions from Proposition 20.16 is satisfied. If we make a specific choice of nowhere vanishing section of $\det E^*$ then we say that E is **oriented**.

Thus as with vector spaces, if a vector bundle is orientable then if M is connected there are exactly two orientations. If μ is a non-vanishing section of $\det E^*$ then $f\,\mu$ is another non-vanishing section for any strictly positive function $f\colon M \to (0,\infty)$, and it defines the same orientation as μ. Meanwhile if $h\colon M \to (-\infty, 0)$ is any strictly negative smooth function then $h\,\mu$ defines the other orientation. We denote by $[\mu]$ the equivalence class and often refer to $[\mu]$ as the **orientation**.

Definition 20.18. Suppose now $\pi\colon E \to M$ is an oriented vector bundle of rank k, with orientation $[\mu]$. This allows us to assign an orientation to each vector space E_x as follows: a basis $\{v_1, \dots, v_k\}$ of E_x is positive if and only if $\mu_x(v_1, \dots, v_k) > 0$. This clarifies the intuitive idea that an orientation of a vector bundle is an orientation of each fibre that depends smoothly on x. Similarly we say a local frame $\{e_1, \dots, e_k\}$ is **positively oriented** if the function $\mu(e_1, \dots, e_k)$ is positive.

Specialising this to our favourite type of vector bundle tells what it means for a manifold to be orientable.

Definition 20.19. A manifold M of dimension n is said to be **orientable** if $TM \to M$ is an orientable vector bundle.

In this case since $\Gamma(\det T^*M) = \Omega^n(M)$ is just the top-dimensional differential forms, an orientation μ is a nowhere vanishing differential n-form. This has its own special name:

Definition 20.20. A **volume form** on an n-dimensional smooth manifold M is a nowhere-vanishing differential n-form.

A manifold together with a choice of orientation (i.e. volume form) $[\mu]$ is called an **oriented manifold**. By a slight abuse of notation we often refer to $[\mu]$ as an orientation of M itself (rather than TM).

Definition 20.21. Let $(M, [\mu])$ and $(N, [\nu])$ be two oriented manifolds of the same dimension n. Suppose $\varphi\colon M \to N$ is a diffeomorphism. Then $\varphi^\star(\nu) = f\,\mu$ for a smooth nowhere vanishing function $f \in C^\infty(M)$ (this is because $\Omega^n(M)$ is the space of sections of a one-dimensional bundle). We say that φ is **orientation preserving** if f is everywhere positive and **orientation reversing** if f is everywhere negative. Note that if M and N are not connected, it may be the case that φ is neither orientation preserving or reversing.

Definition 20.22. As a special case of this, a chart $\sigma\colon U \to O$ on an oriented manifold M^n is said to be **positively oriented** if σ is an orientation preserving diffeomorphism between manifolds U and O (here U inherits the orientation from M and O inherits the standard orientation from $\mathbb{R}^n$).

We conclude this lecture by restating Proposition 20.16 in the special case of a tangent bundle, since this will more convenient to refer back to in the future.

Corollary 20.23 (Orientability of manifolds). *Let M be a smooth manifold of dimension n. The following are equivalent:*

(i) *M admits a volume form.*

(ii) *There exists a smooth atlas $\Sigma := \{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ for M such that whenever $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$,*

$$\det D(\sigma_{\mathtt{a}} \circ \sigma_{\mathtt{b}}^{-1})(\sigma_{\mathtt{b}}(x)) > 0, \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}}. \tag{20.4}$$

We call such an atlas a **positively oriented** *smooth atlas. Note that every chart $\sigma_{\mathtt{a}}$ is then positively oriented.*

(iii) *The determinant line bundle of the cotangent bundle T^*M is a trivial bundle.*

If either of these hold, we say M is orientable, and a specific choice thereof (i.e. a choice of volume form, a choice of atlas, a choice of trivialisation) is called an orientation of M.

On Problem Sheet K you will see some examples of orientable and non-orientable manifolds.

Manifolds with boundary

Let us now move on to defining manifolds with *boundary*. A serious defect of differential geometry so far (at least as we have defined it) is that many natural and interesting compact subsets of Euclidean space fail to be manifolds, and thus none of our results are applicable to them.

Two key examples are the *closed* unit ball D^n, or a closed interval $[a, b] \subset \mathbb{R}$. Neither of these are locally Euclidean spaces (of dimension n and 1 respectively), since points on their boundary do not have neighbourhoods that are homeomorphic to open subsets of $\mathbb{R}^n$ (or $\mathbb{R}$). But note in both cases their interior *is* a smooth manifold of the desired dimension. For the closed ball D^n, the interior is B^n which is an n-dimensional manifold, and for the interval $[a, b]$, the interior (a, b) is a one-dimensional manifold. Moreover the boundary in both cases is an $(n-1)$-dimensional manifold: for the closed ball, $\partial D^n = S^{n-1}$, and $\partial[a, b] = \{a, b\}$.

Remark 21.1. **Warning:** In Lecture 1 (cf. Remark 1.23) we noted that manifold theory had re-purposed the words "open" and "closed" and given them their own meanings, which in many cases were *not* the same as the topological definitions of open and closed. In these notes we elected not to use the "manifold" meanings, and thus for us the words "open" and "closed" should always be taken to have their standard topological meaning.

Unfortunately the same is true of the word "boundary". As we will shortly see, the "boundary" of a manifold does *not* necessarily coincide with the topological definition of the word boundary. This time we will favour the manifold definition of the word, and thus when we write ∂M this is always taken to mean the "manifold" definition of the boundary (which we will shortly introduce). We will use the phrase *topological boundary* to denote the boundary in the sense of point-set topology, and use the notation ∂^{top}. Thus for any subset Y of a topological space X,

$$\partial^{\mathrm{top}} Y = \overline{Y} \setminus \mathrm{int}(Y).$$

We will see several examples below where $\partial M \neq \partial^{\mathrm{top}} M$ for a M a manifold with boundary.

Definition 21.2. A pair of **half-spaces** of $\mathbb{R}^n$ is specified by two things: a linear functional $p \in (\mathbb{R}^n)^*$, and a real number a, which gives us

$$\mathbb{R}^n_{p \geq a} := \{x \in \mathbb{R}^n \mid p(x) \geq a\}, \qquad \text{and} \qquad \mathbb{R}^n_{p \leq a} := \{x \in \mathbb{R}^n \mid p(x) \leq a\}.$$

In a similar way we have **open half-spaces**

$$\mathbb{R}^n_{p > a} := \{x \in \mathbb{R}^n \mid p(x) > a\}, \qquad \text{and} \qquad \mathbb{R}^n_{p < a} := \{x \in \mathbb{R}^n \mid p(x) < a\}.$$

EXAMPLE 21.3. Let $p\colon \mathbb{R}^n \to \mathbb{R}$ denote the linear functional $p = u^1$, i.e.

$$p(x^1, \dots, x^n) = u^1(x^1, \dots, x^n) = x^1.$$

We define the **standard half-spaces** to be

$$\mathbb{R}^n_{u^1 \geq 0} := \left\{ (x^1, \dots, x^n) \in \mathbb{R}^n \mid x^1 \geq 0 \right\}, \qquad \mathbb{R}^n_{u^1 \leq 0} := \left\{ (x^1, \dots, x^n) \in \mathbb{R}^n \mid x^1 \leq 0 \right\},$$

which we will typically abbreviate by $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$ respectively.

REMARK 21.4. **Warning:** It is more common in the literature to define the "standard" half-spaces using $p = u^n$ instead. For instance, $\mathbb{R}^n_{u^n \geq 0}$ is the "upper half-plane" $\mathbb{H}^n$ usually used in hyperbolic geometry (which we will introduce in Lecture 49 next semester). I prefer to use the standard half-spaces from Example 21.3 for two reasons:

(i) As we will see next lecture, using $\mathbb{R}^n_-$ as our "model" half-space leads to simpler formulae when discussing integration. The reason for this is explained in Problem K.2.

(ii) The symbol $\mathbb{H}^n$ is usually understood to denote the half-space $\mathbb{R}^n_{u^n \geq 0}$ which *in addition* has been endowed with its standard *hyperbolic metric* (a topic we will come back to extensively in Differential Geometry II). Since we are not making any statements about metrics here, to avoid confusion I prefer not to use the symbol $\mathbb{H}^n$.

Of course, at the end of the day it is essentially irrelevant which half-space we choose as our "standard" one; they all give rise to the same notion. We could equally as well set the entire theory up with our "standard" half space being $\mathbb{R}^n_{p \geq \pi}$, where

$$p(x^1, \dots, x^n) := \sum_{i=1}^{n} (-1)^i x^i - \log \Gamma(n).$$

(This choice would be somewhat inconvenient when it came to computations though!)

With these considerations in mind, let us now define a topological manifold with boundary.

DEFINITION 21.5. A topological space M is called a **topological manifold with boundary of dimension** n if:

(i) Every point $x \in M$ has a neighbourhood homeomorphic to an open subset of the standard half-space $\mathbb{R}^n_-$,

(ii) M is Hausdorff and has at most countably many connected components,

(iii) M is paracompact.

We refer to n as the **dimension** of M.

Any topological manifold of dimension n is also a topological manifold with boundary of dimension n. Indeed, if a space is locally Euclidean of dimension n, then it also satisfies condition (i) above, since any open subset of $\mathbb{R}^n_-$ that does not intersect the hyperplane $\mathbb{R}^n_{u^1=0}$ is also an open subset of $\mathbb{R}^n$. But the new condition (i) is more general, since an open subset of $\mathbb{R}^n_-$ that intersects the hyperplane $\mathbb{R}^n_{u^1=0}$ is *not* an open subset of $\mathbb{R}^n$.

In general, if M is an n-dimensional topological manifold with boundary then we say a point $x \in M$ is an[1] **interior point** if x admits a neighbourhood that is homeomorphic to an open subset of $\mathbb{R}^n$. We denote by $\operatorname{int}(M)$ the set of interior points. If x is not an interior point then we say x is a **boundary point**. We denote by ∂M the collection of boundary points.

The fact that the dimension is well-defined again requires us to invoke Brouwer's Invariance of Domain Theorem (cf. Remark 1.6). In the smooth case however this will be much easier.

EXAMPLE 21.6. Here are some examples of topological manifolds with boundary:

(i) Any topological space M is a topological manifold of dimension n if and only if it is a topological manifold with boundary of dimension n such that $\partial M = \emptyset$.

(ii) Any half space $\mathbb{R}^n_{p\geq a}$ is a topological manifold with boundary of dimension n. The boundary $\partial \mathbb{R}^n_{p\geq a}$ is $\mathbb{R}^n_{p=a}$. More generally any open subset Q of $\mathbb{R}^n_{p\geq a}$ is a topological manifold with boundary, with $\partial Q = Q \cap \mathbb{R}^n_{p=a}$.

(iii) The closed unit ball D^n is a topological manifold with boundary of dimension n. One has $\partial D^n = S^{n-1}$. (Exercise: Prove this!)

(iv) The closed n-dimensional cube $\overline{\mathbb{I}}^n = [-1,1]^n$ that we used in Lecture 11 is a topological manifold with boundary of dimension n. In this case $\partial \overline{\mathbb{I}}^n$ is homeomorphic to S^{n-1}.

(v) The punctured closed unit ball $D^n \setminus \{0\}$ is a topological manifold with boundary, since it is an open subset of the topological manifold with boundary D^n. This is an example where the manifold boundary is *not* the same as the topological boundary, since:

$$\partial(D^n \setminus \{0\}) = S^{n-1}, \qquad \partial^{\mathrm{top}}(D^n \setminus \{0\}) = S^{n-1} \cup \{0\}.$$

(vi) More generally, any annulus which is half-open and half-closed, eg.

$$A^{\leq R}_{>r} := \{x \in \mathbb{R}^n \mid r < |x| \leq R\}, \qquad \text{or} \qquad A^{<R}_{\geq r} := \{x \in \mathbb{R}^n \mid r \leq |x| < R\}$$

is a topological manifold with boundary whose boundary consists of the boundary circle for which one has the non-strict equality:

$$\partial A^{\leq R}_{>r} = \{|x| = R\}, \qquad \partial A^{<R}_{\geq r} = \{|x| = r\},$$

meanwhile

$$\partial^{\mathrm{top}} A^{\leq R}_{>r} = \partial^{\mathrm{top}} A^{<R}_{\geq r} = \{|x| = r\} \cup \{|x| = R\}.$$

[1] This notion coincides with the topological one, see Proposition 21.7 below and Problem Sheet K.

Proposition 21.7. *Let M be a topological manifold with boundary of dimension n. Then $\mathrm{int}(M) \cap \partial M = \emptyset$. Moreover $\mathrm{int}(M)$ is a topological manifold without boundary of dimension n and ∂M is a topological manifold without boundary of dimension $n-1$.*

Proof. The fact that $\mathrm{int}(M) \cap \partial M = \emptyset$ uses Brouwer's Theorem as mentioned above (since $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}^{n-1}$). The rest is clear, since an open subset Q of $\mathbb{R}^n_{p \geq a}$ that does not intersect $\mathbb{R}^n_{p=a}$ is also open in $\mathbb{R}^n$, and if Q is open in $\mathbb{R}^n_{p \geq a}$ then $Q \cap \mathbb{R}^n_{p=a}$ is open in $\mathbb{R}^n_{p=a} \cong \mathbb{R}^{n-1}$. ■

Corollary 21.8. *If M is a topological manifold with boundary and $W \subset M$ is an open set then W is a topological manifold with boundary, and $\partial W = W \cap \partial M$.*

We now define smooth manifolds with boundary. We begin by extending by the definition of a diffeomorphism between open subsets of half-spaces. We already know (Definition 6.15) how to define what it means for a map to be smooth whose domain is not open, so it remains to extend this to the case when the range is also not open.

Definition 21.9. Let $Q \subset \mathbb{R}^n_{p \geq a}$ denote an open set and $f \colon Q \to \mathbb{R}^k_{q \geq b}$ a continuous map. We say that f is **smooth** if the composition $\imath \circ f \colon Q \to \mathbb{R}^k$ is smooth in the sense of Definition 6.15, where $\imath \colon \mathbb{R}^k_{q \geq b} \hookrightarrow \mathbb{R}^k$ is the inclusion. If both $f \colon Q \to f(Q)$ and $f^{-1} \colon f(Q) \to Q$ are homeomorphisms between open sets of half-spaces that are smooth in this sense then we say that f is a **diffeomorphism**.

The next result is standard calculus, and I will leave the proof to you.

Proposition 21.10. *Here are some properties of smooth maps between open sets of half-spaces:*

(i) *Let O be an open subset of $\mathbb{R}^n$ with non-empty intersection with $\mathbb{R}^n_{p \geq a}$. Suppose $f, g \colon O \to \mathbb{R}^k$ are smooth maps in the usual sense (Definition 1.13). If $f = g$ on $O \cap \mathbb{R}^n_{p \geq a}$ then $Df(x) = Dg(x)$ for all $x \in O \cap \mathbb{R}^n_{p \geq a}$.*

(ii) *Let $O \subset \mathbb{R}^n$ be open and $f \colon O \to \mathbb{R}^n_{q \geq b}$ be smooth in the sense of Definition 21.9. If $f(x) \in \mathbb{R}^n_{q=b}$ for all $x \in O$ then $Df(x)$ has image in $\mathbb{R}^n_{q=0}$ for all $x \in O$.*

(iii) *Suppose $Q_1 \subset \mathbb{R}^n_{p \geq a}$ and $Q_2 \subset \mathbb{R}^k_{q \geq b}$ are open sets, and suppose $f \colon Q_1 \to Q_2$ is a diffeomorphism in the sense of Definition 21.9. Assume $\partial Q_1 = Q_1 \cap \mathbb{R}^n_{p=a}$ and $\partial Q_2 = Q_2 \cap \mathbb{R}^k_{q=b}$ are both non-empty. Then f induces diffeomorphisms $\partial Q_1 \to \partial Q_2$ and $\mathrm{int}(Q_1) \to \mathrm{int}(Q_2)$ in the sense of Definition 1.13, where we think of ∂Q_1 and ∂Q_2 as open subsets of $\mathbb{R}^{n-1}$ and $\mathbb{R}^{k-1}$ respectively.*

We then have:

Definition 21.11. Let M be a topological manifold with boundary of dimension n. A **smooth atlas** on M is a collection $\Sigma = \{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to Q_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ where $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is an open cover of M, each $Q_{\mathtt{a}}$ is an open subset of some n-dimensional half-space $\mathbb{R}^n_{p_{\mathtt{a}} \geq a_{\mathtt{a}}}$ (the precise half-space may depend on $\mathtt{a}$), and each $\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to Q_{\mathtt{a}}$ is a homeomorphism such that the usual compatibility condition is satisfied: if $\mathtt{a}, \mathtt{b} \in \mathtt{A}$ are such that $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$ then the composition

$$\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1} \colon \sigma_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}}) \to \sigma_{\mathtt{b}}(U_{\mathtt{a}} \cap U_{\mathtt{b}})$$

should be a diffeomorphism in the sense of Definition 21.9.

We call each such $\sigma_{\mathtt{a}}$ a **half-space chart**. One then defines a smooth structure in exactly the same way as in Definition 1.16, and this gives us the definition of a smooth manifold with boundary.

DEFINITION 21.12. A **smooth manifold with boundary of dimension** n is a pair (M, Σ) where M is a topological manifold with boundary of dimension n, and Σ is a smooth structure on M in the sense of Definition 21.11.

Just as with Proposition 21.7 we have:

PROPOSITION 21.13. *Let M be a smooth manifold with boundary of dimension n. Then $\operatorname{int}(M) \cap \partial M = \emptyset$. Moreover $\operatorname{int}(M)$ naturally inherits the structure of a smooth manifold without boundary of dimension n, and ∂M naturally inherits the structure of a smooth manifold without boundary of dimension $n-1$.*

Proof. This follows from part (iii) of Proposition 21.10. ∎

EXAMPLE 21.14. All the examples from Example 21.6 are naturally smooth manifolds with boundary, except for the unit cube $\mathbb{I}^n$, which is *not* a smooth manifold with boundary when $n \geq 2$. (See Problem Sheet K.)

Although the definition of a smooth atlas does not require all the half-space charts to take values in the same half-space, it is often convenient to assume they do.

DEFINITION 21.15. A **good smooth atlas** on a smooth manifold with boundary M is a smooth atlas as in Definition 21.11 with the additional property that $Q_{\mathtt{a}}$ is an open subset of our preferred standard half-space $\mathbb{R}^n_-$ from Example 21.3.

It is easy to see that we may always assume this:

LEMMA 21.16. *Every smooth manifold with boundary admits a good smooth atlas.*

REMARK 21.17. You might therefore ask what the point was in the more general definition. This is two-fold: firstly it is convenient when proving certain standard spaces are topological (resp. smooth) manifolds with boundary to be allowed more flexibility. Secondly, the distinction between good smooth atlases and normal smooth atlases is meaningful in dimension $n = 1$ when one in addition insists on orientability, as we will see in Proposition 21.23 below.

Many of the concepts we have covered so far in this course make sense for manifolds with boundary, and we don't have the time (or energy) to fill in the details, so let us just briefly summarise some of the important points:

- Partitions of unity still make sense for smooth manifolds with boundary, and they always exist.

- If M is a smooth manifold with boundary of dimension n then T_xM is still an n-dimensional vector space for all $x \in M$. This is clear for $x \in \operatorname{int}(M)$, so suppose $x \in \partial M$. Let $\sigma \colon U \to Q$ denote a half-space chart about x, where Q is an open set in some half-space $\mathbb{R}^n_{p \geq a}$ and $\sigma(x)$ lies in the hyperplane $\mathbb{R}^n_{p=a}$. As before, a function f defined near x on M is smooth at x if and only if

$f\circ\sigma^{-1}$ is smooth near $z := \sigma(x)$. Now recall by definition a function is smooth if and only if it admits a smooth extension to some open neighbourhood of z in $\mathbb{R}^n$. If g and h are any two such extensions of $f\circ\sigma^{-1}$ then by part *(i)* of Proposition 21.10 the derivatives of g and h coincide on $\mathbb{R}^n_{p=a}$. It follows that a derivation on the space of germs of smooth functions at x can be defined in exactly the same way as before, and thus the arguments from Lectures 2 and 3 go through without change to show that the tangent space T_xM at x is again an n-dimensional vector space.

- On the other hand, the tangent space to ∂M at $x \in \partial M$ can be identified with an $(n-1)$-dimensional subspace of T_xM. Indeed, if we let $\imath\colon \partial M \hookrightarrow M$ denote the inclusion then with the notation as above, $\sigma\circ\imath|_{U\cap\partial M}$ is a chart on ∂M and thus
$$D\imath(x)[T_x(\partial M)] = D\sigma(x)^{-1}[T_z\mathbb{R}^n_{p=a}], \tag{21.1}$$
(note that $T_z\mathbb{R}^n_{p=a} \cong \mathbb{R}^n_{p=a} \cong \mathbb{R}^{n-1}$ via Problem B.3.) We usually suppress the $D\imath(x)$ map and thus think of $T_x(\partial M)$ as an actual subspace of T_xM. *Exercise:* Prove that the right-hand side of (21.1) does not depend on the choice of half-space chart σ.

- If N is a smooth manifold (with or without boundary) and $M \subset N$ is a subset endowed with a topology and a smooth structure making it into a smooth manifold with boundary such that the inclusion $M \hookrightarrow N$ is an embedding then M is said to be an **embedded submanifold with boundary**. Immersed submanifolds with boundary are defined similarly. If M is a smooth manifold with boundary then ∂M is an embedded submanifold of M—this follows immediately from the definition. *Exercise:* Investigate how the Implicit Function Theorem 5.13 behaves with respect to manifolds with boundary. What is the correct notion of a slice chart in this setting?

- Both the Whitney Embedding Theorem 6.1 and the Whitney Approximation Theorem 6.14 still work for manifolds with boundary.

- A vector field X on a smooth manifold with boundary M is said to be **tangent to** ∂M if $X(x) \in T_x(\partial M)$ for each $x \in \partial M$. For vector fields that are tangent to M, Theorem 8.10 still works.

- The notion of a fibre bundle still makes sense if the base space is allowed to have boundary. In particular, vector bundles over manifolds with boundary are defined entirely analogously. Things go wrong however if the fibre is allowed to have boundary. *Exercise:* Why?

- Tensors and differential forms are defined in exactly the same way.

We will however go through one aspect in detail, since this will be important in our treatment of the global Stokes' Theorem in Lecture 23. Suppose M is a manifold with boundary and $\pi\colon E \to M$ is a vector bundle over M. If $[\mu]$ is an orientation for E (so $\mu \in \Gamma(\det E^*)$) then $[\mu|_{\partial M}]$ is an orientation on the bundle $E|_{\partial M} \to \partial M$, where $E|_{\partial M} = \pi^{-1}(\partial M)$ (this is a subbundle of E since ∂M is an embedded submanifold of M). For the special case $E = TM$, this gives us an orientation of the bundle $TM|_{\partial M} \to \partial M$. This however is *not* the same thing as an orientation of $T(\partial M) \to \partial M$ (i.e. an orientation of ∂M as a manifold).

DEFINITION 21.18. Let M be a smooth manifold with boundary of dimension n, and let $x \in \partial M$. A tangent vector $v \in T_x M$ is said to be **outward pointing** if for some half-space chart $\sigma \colon U \to Q$ about x, with $Q \subset \mathbb{R}^n_{p \geq a}$ an open set and $z := \sigma(x) \in \mathbb{R}^n_{p=a}$, one has

$$p\big(\mathcal{J}_z^{-1}(D\sigma(x)[v])\big) < 0.$$

To unwrap this: $D\sigma(x)$ is a linear map $T_x M \to T_z \mathbb{R}^n_{p \geq a} = T_z \mathbb{R}^n$. Applying the map $\mathcal{J}_z \colon \mathbb{R}^n \to T_z \mathbb{R}^n$ from Problem B.3, we obtain a vector $\mathcal{J}_z^{-1}(D\sigma(x)[v]) \in \mathbb{R}^n$, which p can then eat to produce a real number. It follows from part (iii) of Proposition 21.10 that the property of being outward pointing is independent of the choice of half-space chart σ.

The definition is rather clearer if we take our preferred half-space $\mathbb{R}^n_-$. Then the condition that $v \in T_x M$ is outward pointing is simply that

$$dx^1|_x(v) > 0,$$

where (x^i) are the local coordinates of σ. See Figure 21.1.

Similarly an **inward-pointing** vector is one for which $dx^1|_x(v) < 0$. This allows us to decompose $T_x M$ as:

$$T_x M = \{\text{outward pointing vectors}\} \sqcup \{\text{inward pointing vectors}\} \sqcup T_x(\partial M),$$

since in such a chart σ, $T_x(\partial M) = \{v \in T_x M \mid dx^1|_x(v) = 0\}$.

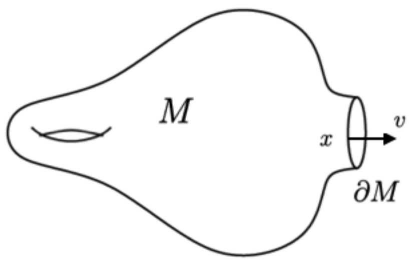

Figure 21.1: An outward pointing vector v

Similarly a section X of $TM|_{\partial M}$ is said to be **outward pointing** if $X(x)$ is outward pointing for every x.

EXAMPLE 21.19. Let M be a manifold with boundary and let $x \in M$. Let $x \in \partial M$ and let $\sigma \colon U \to Q$ denote a half space chart where $Q \subset \mathbb{R}^n_-$ is open. Then $\frac{\partial}{\partial x^1}$ is an outward pointing section of $TM|_{\partial M}$ over $U \cap \partial M$.

In fact, via a standard partition of unity argument, one can produce outward pointing sections defined on the entire boundary.

LEMMA 21.20. *Let M be a smooth manifold with boundary. Then there exists a section $X \in \Gamma(TM|_{\partial M})$ of the bundle $TM|_{\partial M} \to \partial M$ which is outward pointing at every $x \in \partial M$.*

Proof. We may assume M has a good smooth atlas $\{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to Q_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$. Let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote a partition of unity subordinate to $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$. Let $x^i_{\mathtt{a}}$ denote the local coordinates of $\sigma_{\mathtt{a}}$ and set

$$X := \sum_{\mathtt{a}} \lambda_{\mathtt{a}} \frac{\partial}{\partial x^1_{\mathtt{a}}}.$$

This is outward pointing by Example 21.19. ■

Let us now use this to define the induced orientation.

Definition 21.21. Let M be an smooth manifold with boundary of dimension n, and let $\mu \in \Omega^n(M)$ be a volume form. Let X be an outward pointing section. Then we can view $i_X(\mu)$ as an element of $\Omega^{n-1}(\partial M)$. Since X is outward pointing, $i_X(\mu)$ is nowhere vanishing on ∂M, and hence this gives an orientation of ∂M. We call the orientation $[i_X(\mu)]$ of ∂M the **induced orientation** from the orientation $[\mu]$ of M. *Exercise:* Check this is well-defined, i.e. independent of X and of the representative μ of $[\mu]$.

Thus if $(X_1, \dots, X_n)$ is a positively oriented frame for TM (i.e. $\mu(X_1, \dots, X_n) > 0$ over U) such that $U \cap \partial M$ is non-empty, then $(X_2, \dots, X_n)$ is a positively oriented local frame for $TM|_{\partial M}$ over $U \cap \partial M$ with respect to the induced orientation if $X_1|_{U \cap \partial M}$ is outward pointing.

Remark 21.22. In the case $n = 1$, the boundary ∂M is a discrete set of points. We only defined orientations for vector spaces of positive dimension, but this can still be made sense of if we simply think of a boundary point x being positively oriented if $\mu(X)(x) > 0$ (note in this case $\mu(X)$ is simply a function) and negatively oriented otherwise.

Here is an extension of Corollary 20.23 for manifolds with boundary. This is where it is important to make the distinction between a good atlas and a normal one.

Proposition 21.23. *Let M be an oriented smooth manifold with boundary of dimension n. Then M admits a positively oriented smooth atlas (that is, one such that (20.4) holds). If $n \geq 2$ then M admits a positively oriented good smooth atlas.*

Proof. If σ is a chart with local coordinates $(x^1, \dots, x^n)$ that is not positively oriented then we replace it with a new chart $(x^1, -x^2, \dots, x^n)$. For $n = 1$ this changes a chart from being a $\mathbb{R}^n_-$ half-space chart to a $\mathbb{R}^n_+$ half-space chart. ■

For the rest of these notes, **all** manifolds (topological or smooth) should be assumed **not** to have boundary, unless it is explicitly said that they do.

LECTURE 22

Singular cubes and Stokes' Theorem

In this lecture we discuss *integration* on manifolds, and prove a local version of one of the fundamental theorems in differential calculus, known as *Stokes' Theorem*. Next lecture we will globalise this. There are several ways to approach integration; we will choose one that makes the proof of the de Rham Theorem (coming in the non-examinable Lecture 27) relatively painless.

DEFINITION 22.1. Let us abbreviate by C^k the closed cube[1] $[0,1]^k$, thought of as sitting inside $\mathbb{R}^k$. For $k = 0$, $C^0 = \{0\}$ is a point. A **smooth singular k-cube** (often shorted to: a "singular k-cube" or just a "k-cube") in a smooth manifold M is a smooth map $c\colon C^k \to M$. Thus a singular 0-cube is simply a point $c(0)$ in M.

Recall that (by Definition 6.15) a map $c\colon C^k \to M$ is smooth if there exists a neighbourhood U of C^k in $\mathbb{R}^k$ and a smooth map $\tilde{c}\colon U \to M$ such that $\tilde{c}|_{C^k} = c$. Of course the extension $\tilde{c}$ is not unique. (For $k = 0$, we declare that any map $c\colon C^0 \to M$ is smooth).

REMARK 22.2. The adjective "singular" is meant to draw your attention to the fact that c need not be injective or an immersion. Indeed, a valid smooth singular k-cube would be a constant map! Moreover if $k > \dim M$ then no singular k-cube can be an immersion.

The next example is more important than you would first guess.

EXAMPLE 22.3. We let $I^k\colon C^k \to \mathbb{R}^k$ denote the inclusion and call I^k the **standard smooth singular k-cube**.

REMARK 22.4. We will often regard I^k as taking values in C^k. This is harmless, since it does—it simply means we are viewing I^k as the identity map on C^k. Strictly speaking however in this case I^k is not a smooth singular cube, since the range space C^k is not a smooth manifold. For the most part we shall ignore this pedantry.

(♣) REMARK 22.5. You might hope that the machinery of smooth manifolds with boundary that we developed last lecture would allow us to forego the tedious extension business.

This works fine for $k = 1$: $C^1 = [0,1]$ is a smooth manifold with boundary, and a singular 1-cube is simply a smooth map $C^1 \to M$ between manifolds. However for $k = 2$ this goes wrong: C^2 is not a manifold with boundary (see Problem K.4). It is however a **smooth manifold with corners**, which is defined as you might expect: instead of a half space atlas one works with a *quarter space atlas*. If M is a smooth manifold with corners then its boundary ∂M is a smooth manifold with

[1]In the past we used $\mathbb{I}^k$ for the open cube $(-1,1)^k$; here it is more convenient to work on $[0,1]$ itself, so we choose different notation.

boundary, and the boundary of the boundary is then a smooth manifold without boundary. (For C^2, one has ∂C^2 equal to the union of the edges, and $\partial(\partial C^2)$ equal to the four vertices).

Sadly however this *still* isn't enough, since for $n \geq 3$ the space C^k is not a smooth manifold with corners either. The correct notion is that of a **stratified manifold**, which, roughly speaking is a manifold which is allowed to "boundary-like" pieces of arbitrarily high codimension. A manifold with boundary is a stratified manifold with only codimension one strata, and a manifold with corners is a stratified manifold with only codimension one and two strata. In general, C^k is a stratified manifold with k different stratas.

That said, developing the entire theory of stratified manifolds just to dispense with the need to talk about extensions is somewhat inefficient, even by my standards, so we will stick with the extensions. This will therefore be a minor annoyance throughout the lecture.

Definition 22.6. Let $k > 0$, and let $\omega \in \Omega^k(C^k)$ denote a k-form on C^k (you can think of this as meaning: ω is a k-form on some neighbourhood U of C^k in $\mathbb{R}^k$). We can write $\omega = h\, dx^1 \wedge \cdots \wedge dx^k$ for some $h \in C^\infty(C^k)$. We define the **integral of** ω to be the Riemann integral of h:

$$\int_{C^k} \omega := \int_{C^k} h$$

We emphasise the right-hand side is the *normal Riemann integral* of the function h.

We now transfer this to manifolds:

Definition 22.7. Let $k > 0$ and let c be a smooth singular k-cube in M and let $\omega \in \Omega^k(M)$ denote a k-form. Then $c^\star(\omega)$ is a k-form on C^k. We define the **integral of** ω **over** c to be the real number

$$\int_c \omega := \int_{C^k} c^\star(\omega).$$

It would be sufficient if ω was only defined on some neighbourhood of the image of c for this to make sense. For $k = 0$, the definition is simpler; in this case ω is just a function f, and

$$\int_c f := f(c(0)).$$

Remark 22.8. If we write $c^\star(\omega) = h\, dx^1 \wedge \cdots \wedge dx^k$ then the function h is given explicitly by

$$h = c^\star(\omega)\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\right)$$

(this expression is well-defined since $c^\star(\omega)$ is really defined on some open neighbourhood of C^k). Thus an alternative formula is

$$\int_c \omega = \int_{C^k} c^\star(\omega)\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\right),$$

where again the right-hand side is just a normal Riemann integral.

Remark 22.9. Since for any singular k-cube c we have $c = c \circ I^k$, we have using (19.2) that

$$\int_c \omega = \int_{c \circ I^k} \omega = \int_{C^k} (c \circ I^k)^\star(\omega) = \int_{C^k} (I^k)^\star(c^\star(\omega)) = \int_{I^k} c^\star(\omega).$$

Definition 22.10. A singular k-cube $c \colon C^k \to M$ is said to be **degenerate** if there exists $1 \le i \le k$ such that c does not depend on x^i. Otherwise c is said to be **non-degenerate**. Thus a 0-cube is never degenerate, and a 1-cube is degenerate if and only if it is a constant map.

On Problem Sheet L you will prove.

Lemma 22.11. *If $c \colon C^k \to M$ is a degenerate singular k-cube then $\int_c \omega = 0$ for any $\omega \in \Omega^k(M)$.*

The next result is also on Problem Sheet L. You should think of it as a version of the usual change of variables formula from multivariable calculus:

Proposition 22.12 (Change of Variables). *Let $c \colon C^k \to M$ be a smooth singular k-cube in M and let $\varphi \colon C^k \to C^k$ be an orientation preserving diffeomorphism*[2]. *Let $\tilde{c} := c \circ \varphi$. Then*

$$\int_c \omega = \int_{\tilde{c}} \omega.$$

Let us now consider formal sums of singular cubes.

Definition 22.13. Let $Q_k(M)$ denote[3] the (infinite-dimensional) free vector space generated by the collection of all the smooth singular k-cubes in M. Thus an element of $Q_k(M)$ is a formal finite sum $q = \sum_i a_i c_i$ where $a_i \in \mathbb{R}$ and the c_i are smooth singular k-cubes. We call an element $q \in Q_k(M)$ a **smooth singular k-chain**, or (sometimes just a k-chain). A k-chain $q = \sum_i a_i c_i$ is said to be **non-degenerate** if each cube c_i is non-degenerate.

Example 22.14. Since a 0-cube in M is just a point in M, the space $Q_0(M)$ can be thought as the infinite-dimensional vector space with basis the points of M. In particular, if $x, y \in M$ then the expression $x - y$ makes sense in $Q_0(M)$, even though it does not in M.

Remark 22.15. **Warning:** The space $Q_0(M)$ is a vector space with basis the points in M. Thus (by definition) there are no relations between different elements. This can be confusing, particularly if the manifold M happens to be a submanifold of Euclidean space where it *does* make sense to add points together. As an example, let us take $M = \mathbb{R}^n$. Let $v, w \in \mathbb{R}^n$ be two vectors. Then in $\mathbb{R}^n$, we can add v and w together to get a new vector $v + w$. However in $Q_0(\mathbb{R}^n)$, the three elements v, w and $v + w$ are *linearly independent* and thus it is *not* true that $v + w = (v + w)$! A similar issue occurs with scalar multiplication. If this confuses you, consider

[2] As usual, think of this as meaning that φ is the restriction to C^k of an orientation preserving diffeomorphism of some neighbourhood.

[3] To explain the notation: "cube" sounds like it begins with a "Q".

writing the addition and multiplication operations in Q_0 with a different colour, for instance red. Thus if $v, w \in \mathbb{R}^n$ and $a \in \mathbb{R}$ then

$$v + w \neq v + w, \qquad a\, v \neq 1(av).$$

Luckily most of the time this shouldn't be confusing, since typically on manifolds one cannot add points together, and thus the notation is unambiguous already.

Definition 22.16. We define the integral of a k-form over a k-chain in M by linearity: if $q = \sum_i a_i\, c_i$ then

$$\int_q \omega := \sum_i a_i \int_{c_i} \omega.$$

We will also need the concept of the *boundary* of a chain.

Definition 22.17. Fix $1 \leq i \leq k$ and let $c \colon C^k \to M$ denote a singular k-cube. The ith **front face** of c, written $F_i c$, is the singular $(k-1)$-cube defined by

$$F_i c(x^1, \dots, x^{k-1}) := c(x^1, \dots, x^{i-1}, 0, x^i, \dots, x^{k-1}).$$

Similarly the ith **back face** is the singular $(k-1)$-cube defined by

$$B_i c(x^1, \dots, x^{k-1}) := c(x^1, \dots, x^{i-1}, 1, x^i, \dots, x^{k-1}).$$

Let us note the following result

Lemma 22.18. *Let $c \colon C^k \to M$ be a smooth singular k-cube. Let $1 \leq i < j \leq k$. Then:*

$$\begin{aligned} F_i(F_j c) &= F_{j-1}(F_i c), \\ B_i(B_j c) &= B_{j-1}(B_i c), \\ F_i(B_j c) &= B_{j-1}(F_i c), \\ B_i(F_j c) &= F_{j-1}(B_i c). \end{aligned} \tag{22.1}$$

Moreover one has

$$F_i c = c \circ F_i I^k, \qquad B_i c = c \circ B_i I^k. \tag{22.2}$$

The proof is a trivial computation which I leave to you.

Definition 22.19. Let $c \colon C^k \to M$ be a smooth singular k-cube for $k > 0$. We define the **boundary** of c, written ∂c, to be the element of $Q_{k-1}(M)$ given by

$$\partial c := \sum_{i=1}^{k} (-1)^i (F_i c - B_i c).$$

We define the boundary of a 0-cube to be the real number 1. Note that if a cube c is non-degenerate then so is ∂c. We then extend ∂ to arbitrary k-chains by linearity. Thus we may think of ∂ as a linear map $Q_k(M) \to Q_{k-1}(M)$ for all $k \geq 1$ (this works for $k = 0$ too if we define $Q_{-1}(M) := \mathbb{R}$).

REMARK 22.20. Thus this is yet another meaning of the symbol ∂. This one is not as confusing as the topological boundary and the manifold boundary (cf. Remark 21.1), since c is a function, and thus there can be no ambiguity about what is meant.

EXAMPLE 22.21. Let $c\colon [0,1] \to M$ be a 1-cube. Then $F_1 c$ is the 0-cube $c(0)$ and $B_1 c$ is the 0-cube $c(1)$. Thus $\partial c = c(1) - c(0)$. Remember the subtraction is taking place in $Q_0(M)$, not M itself!

Let us now state and prove the one-dimensional version of Stokes' Theorem, also known as the Fundamental Theorem of Calculus, in the language of chains.

THEOREM 22.22 (The Fundamental Theorem of Calculus for Singular 1-Chains). *Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable and let q be a singular 1-chain in $\mathbb{R}$. Then*

$$\int_q df = \int_{\partial q} f$$

Proof. By linearity of the integral, we may assume that q is a single singular 1-cube c. Then we compute

$$\begin{aligned}
\int_c df &= \int_0^1 c^\star(df)\left[\frac{\partial}{\partial t}\right] dt \\
&= \int_0^1 (f \circ c)'(t) dt \\
&\overset{(\dagger)}{=} f(c(1)) - f(c(0)) \\
&= \int_{c(1)} f - \int_{c(0)} f \\
&= \int_{\partial c} f,
\end{aligned}$$

where $(\dagger)$ used the usual Fundamental Theorem of Calculus that you have known since kindergarten. ■

PROPOSITION 22.23. *The boundary operator squares to zero: $\partial^2 = 0$.*

Proof. Since ∂ is linear, it suffices to show that $\partial(\partial c) = 0$ for any cube c.

$$\begin{aligned}
\partial(\partial c) &= \partial\left(\sum_{i=1}^{k}(-1)^i(F_i c - B_i c)\right) \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k-1}(-1)^{i+j}\Big(F_j(F_i c) - F_j(B_i c) - B_j(F_i c) + B_j(B_i c)\Big)
\end{aligned}$$

Using the face relations (22.1), we see that the first and fourth terms cancel in pairs, and the second and third terms cancel each other. ■

Definition 22.24. It follows that we can play a similar game to the definition (Definition 19.18) of the de Rham cohomology groups. Let us say a chain q is **closed** if $\partial q = 0$ and a chain q is **exact** if $q = \partial p$ for some $(k+1)$-chain p. Then every exact chain is also closed (as $\partial^2 = 0$), and thus we can form the quotient vector space:

$$H_k^{\mathrm{cube}}(M;\mathbb{R}) := \frac{\{\text{closed non-degenerate } k\text{-chains}\}}{\{\text{exact non-degenerate } k\text{-chains}\}}.$$

If q is a closed k-chain, we denote by $[q]$ its equivalence class in $H_k^{\mathrm{cube}}(M;\mathbb{R})$. The reason for insisting on non-degeneracy will not be important until Lecture 27. Unlike the groups $H^k_{\mathrm{dR}}(M)$, which are certainly zero for $k > \dim M$, a priori the groups $H_k^{\mathrm{cube}}(M;\mathbb{R})$ could be non-zero for arbitrarily high k. However this is not the case. In fact, as we will explain in Lecture 27, there is an isomorphism

$$H_k^{\mathrm{cube}}(M;\mathbb{R}) \cong H^{n-k}_{\mathrm{dR}}(M), \qquad \forall k \geq 0. \tag{22.3}$$

We can now state and prove the local version of Stokes' Theorem.

Theorem 22.25 (The Local Stokes' Theorem). *Let M be a smooth manifold. Let $q \in Q_k(M)$ and $\omega \in \Omega^{k-1}(M)$. Then*

$$\int_q d\omega = \int_{\partial q} \omega.$$

Note Theorem 22.22 is the special case $M = \mathbb{R}$ and $k = 1$. This proof is non-examinable.

(♣) *Proof.* We prove the result in three steps.

1. Let us first consider the case where $M = \mathbb{R}^k$ and $c = I^k$ is the standard cube from Example 22.3. This actually represents most of the work. By linearity we may assume that ω is of the form

$$\omega = f\, dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^k,$$

where the carat indicates we skip the term dx^j. In this first step, we come up with a nice formula for the right-hand side $\int_{\partial I^k} \omega$.

We have by definition that

$$\int_{\partial I^k} \omega = \sum_{i=1}^k (-1)^i \left(\int_{F_i I^k} \omega - \int_{B_i I^k} \omega \right). \tag{22.4}$$

We now claim that:

$$\int_{F_i I^k} \omega = \begin{cases} \int_{C^{k-1}} f \circ F_i I^k, & i = j, \\ 0, & i \neq j. \end{cases} \tag{22.5}$$

The proof of (22.5) is a little fiddly. One way to argue this is as follows: from Remark 22.8 and (15.8) we have that

$$\int_{F_i I^k} \omega = \int_{C^{k-1}} (f \circ F_i I^k) \cdot \det A,$$

where $A = (A_l^p)$ is the $(k-1) \times (k-1)$ matrix whose entries are given by

$$A_l^p = D_l(u^p \circ F_i I^k), \qquad \text{for } 1 \leq l \leq k-1 \text{ and } 1 \leq p \leq k,\ p \neq j.$$

The function $u^p \circ F_i I^k$ is given by

$$u^p \circ F_i I^k(x^1, \dots, x^{k-1}) = u^p(x^1, \dots, x^{i-1}, 0, x^k, \dots x^{k-1}).$$

Thus if $i = j$ then $A_l^p = \delta_l^p$ and thus $\det A = 1$. However if $i \neq j$ then the entire ith row (A_l^i) is zero (since $u^i \circ F_i I^k$ is the zero function), and thus $\det A = 0$. This proves (22.5). Together with a similar formula for the back face, we see that (22.4) reduces to

$$\int_{\partial I^k} \omega = (-1)^j \int_{C^{k-1}} f \circ F_j I^k - f \circ B_j I^k. \tag{22.6}$$

By Fubini's Theorem and the Fundamental Theorem of Calculus:

$$\begin{aligned}
&\int_{C^{k-1}} f \circ B_j I^k - f \circ F_j I^k \\
&\quad = \int_0^1 \cdots \int_0^1 \Big(f(x^1, \dots, 1, \dots, x^k) - f(x^1, \dots, 0, \dots, x^k) \Big) dx^1 \cdots \widehat{dx^j} \cdots dx^k \\
&\quad = \int_{C^k} \frac{\partial f}{\partial x^j}.
\end{aligned}$$

Thus we conclude from (22.6) that

$$\int_{\partial I^k} \omega = (-1)^{j-1} \int_{C^k} \frac{\partial f}{\partial x^j}. \tag{22.7}$$

2. We now consider the term $\int_{I^k} d\omega$. Since $df = \frac{\partial f}{\partial x^j}\, dx^j$ we have (writing the summation signs for clarity)

$$\begin{aligned}
\int_{I^k} d\omega &= \int_{I^k} df \wedge dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^k \\
&= \int_{I^k} \sum_{i=1}^{k} \frac{\partial f}{\partial x^i}\, dx^i \wedge dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^k) \\
&= (-1)^{j-1} \int_{I^k} \frac{\partial f}{\partial x^j}\, dx^1 \wedge \cdots \wedge dx^k \\
&= (-1)^{j-1} \int_{C^k} \frac{\partial f}{\partial x^j}
\end{aligned}$$

This completes the proof for $M = \mathbb{R}^k$ and $c = I^k$.

3. In the general case, again by linearity we may assume $q = c$ is a singular

k-cube. Then using (22.2) one has

$$\begin{aligned}
\int_{\partial c} \omega &= \sum_{i=1}^{k} (-1)^i \left(\int_{F_i c} \omega - \int_{B_i c} \omega \right) \\
&= \sum_{i=1}^{k} (-1)^i \left(\int_{c \circ F_i I^k} \omega - \int_{c \circ B_i I^k} \omega \right) \\
&= \sum_{i=1}^{k} \left(\int_{C^{k-1}} (c \circ F_i I^k)^\star(\omega) - \int_{C^{k-1}} (c \circ B_i I^k)^\star(\omega) \right) \\
&= \sum_{i=1}^{k} \left(\int_{C^{k-1}} (F_i I^k)^\star(c^\star(\omega)) - \int_{C^{k-1}} (F_i I^k)^\star(c^\star(\omega)) \right) \\
&= \sum_{i=1}^{k} \left(\int_{F_i I^k} c^\star(\omega) - \int_{B_i I^k} c^\star(\omega) \right) \\
&= \int_{\partial I^k} c^\star(\omega) \\
&= \int_{I^k} d(c^\star(\omega))
\end{aligned}$$

by the previous step. But since $c^\star \circ d = d \circ c^\star$ by Lemma 19.19, we have

$$\int_{I^k} d(c^\star(\omega)) = \int_{I^k} c^\star(d\omega) = \int_c d\omega,$$

where we used Remark 22.9 at the end. This completes the proof. ■

Next lecture we shall globalise Theorem 22.25. For now, let us note the following consequence, which will be useful in the last lecture of the semester.

Definition 22.26. Let M be a smooth manifold of dimension n. Then for $0 \leq k \leq n$ we can think of integration as defining a bilinear map

$$\int : Q_k(M) \times \Omega^k(M) \to \mathbb{R}, \qquad (q, \omega) \mapsto \int_q \omega.$$

Corollary 22.27. *The bilinear form $\int$ is also well-defined on the (co)homology level, that is, the map*

$$\int : H_k^{\mathrm{cube}}(M; \mathbb{R}) \times H^k_{\mathrm{dR}}(M) \to \mathbb{R}, \qquad ([q], [\omega]) \mapsto \int_q \omega$$

is well-defined.

Proof. We already know that $\int_q \omega$ vanishes whenever q is degenerate (Lemma 22.11). Thus we need only show that if q is a closed non-degenerate k-chain and ω is a closed k-form, then for any non-degenerate $(k+1)$-chain p and any $(k-1)$-form ϑ, one has

$$\int_{q+\partial p} (\omega + d\vartheta) = \int_q \omega.$$

For this we expand by linearity on both sides:

$$\begin{aligned}
\int_{q+\partial p} (\omega + d\vartheta) &= \int_q \omega + \int_{\partial p} \omega + \int_q d\vartheta + \int_{\partial p} d\vartheta \\
&= \int_q \omega + \int_p d\omega + \int_{\partial q} \vartheta + \int_p (d^2(\vartheta)) \\
&= \int_q \omega + 0,
\end{aligned}$$

where the second equality used Stokes' Theorem and the last used the assumption that q and ω are closed. ∎

Do not be fooled by the apparent similarity of (22.3) and Corollary 22.27. The former is *much* deeper. We shall come back to this at the end of the course.

LECTURE 23

The Poincaré Lemma

In this lecture we give a global version of Stokes' Theorem. We begin by explaining how to make sense of the expression $\int_M \omega$. Unlike the local version in the last lecture, this will only work when M is oriented, and ω is a compactly supported differential form of top degree (i.e. of degree $n = \dim M$).

Remark 23.1. In general we will start omitting explicit mention of the orientation in our notation for an oriented manifold, and thus just write M instead of $(M, [\mu])$. We will also adopt the shorthand notation that if M is an oriented manifold then $-M$ denotes the same manifold, but with the opposite orientation.

Definition 23.2. Let M^n be an oriented manifold. A singular cube $c \colon C^n \to M$ is said to be an **orientation preserving** if there exists a neighbourhood U of C^n in $\mathbb{R}^n$ and an orientation preserving[1] embedding $\tilde{c} \colon U \to M$ such that $\tilde{c}|_{C^n} = c$. Note that $\tilde{c}$ is thus a diffeomorphism onto its image.

Remark 23.3. Note that if M^n is an oriented manifold, we can always find an open cover of M such that each open set U in that cover is contained in the interior of the image of an orientation preserving singular cube $c \colon C^n \to M$. Indeed if $\sigma \colon U \to O$ is a positively oriented chart (cf. Definition 20.22) then one can apply an affine transformation of $\mathbb{R}^n$ so that $C^n \subset O$, and then $c := \sigma^{-1}|_{C^n}$ works.

Definition 23.4. Let M be a smooth manifold, and let $\omega \in \Omega(M)$. The **support** of ω is defined in the same way as normal:

$$\operatorname{supp}(\omega) := \overline{\{x \in M \mid \omega_x \neq 0\}}$$

A differential form ω is said to have **compact support** if $\operatorname{supp}(\omega)$ is compact. We denote by $\Omega_c(M) \subset \Omega(M)$ the subset of differential forms with compact support, and $\Omega_c^r(M)$ the differential r-forms with compact support. Note that by definition of the exterior differential, we have

$$\operatorname{supp}(d\omega) \subseteq \operatorname{supp}(\omega). \tag{23.1}$$

We begin with the following lemma, which explains why for global integration we need our manifold to be oriented. In the following, we will always assume the cube C^k carries the standard orientation inherited from $\mathbb{R}^k$.

Lemma 23.5. *Let M^n be an orientated manifold and $\omega \in \Omega^n(M)$. Let $c_1, c_2 \colon C^n \to M$ be two orientation preserving singular cubes, and assume that*

$$\operatorname{supp}(\omega) \subset \operatorname{int}(\operatorname{im} c_1) \cap \operatorname{int}(\operatorname{im} c_2)$$

Then

$$\int_{c_1} \omega = \int_{c_2} \omega.$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1]Here we always explicitly assume $\mathbb{R}^n$ carries its standard orientation.

Proof. This almost follows from Proposition 22.12, since $c_2^{-1} \circ c_1$ is almost an orientation preserving diffeomorphism of C^n. The only issue is that $c_2^{-1} \circ c_1$ may not be defined on all of C^n. However, since $\operatorname{supp}(\omega) \subset \operatorname{int}(\operatorname{im} c_1) \cap \operatorname{int}(\operatorname{im} c_2)$, the proof of Proposition 22.12 goes through without change to give

$$\int_{c_2} \omega = \int_{c_2 \circ c_2^{-1} \circ c_1} \omega = \int_{c_1} \omega.$$

■

Thus we can unambiguously make the following definition.

DEFINITION 23.6. Let M^n be an oriented manifold and $\omega \in \Omega^n(M)$. Assume that ω has support in the interior of the image of some orientation preserving singular cube c. Then we define

$$\int_M \omega := \int_c \omega.$$

The following lemma is immediate from the definitions.

LEMMA 23.7. *If c is an orientation reversing singular cube and ω has support in* $\operatorname{im} c$ *then*

$$\int_M \omega = -\int_c \omega.$$

Thus (using the convention from Remark 23.1) one has

$$\int_M \omega = -\int_{-M} \omega.$$

We can use a partition of unity to extend this to an arbitrary $\omega \in \Omega_c^n(M)$. We first give the definition, and then prove it is well defined.

DEFINITION 23.8. Let M^n be an oriented manifold and let $\omega \in \Omega_c^n(M)$. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover with the property that each $U_{\mathtt{a}}$ is contained in the interior of the image of some orientation preserving singular cube (cf. Remark 23.3). Let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be a partition of unity subordinate to this cover. We define

$$\int_M \omega := \sum_{\mathtt{a} \in \mathtt{A}} \int_M \lambda_{\mathtt{a}}\, \omega.$$

Note this is a finite sum since ω has compact support and $\operatorname{supp}(\lambda_{\mathtt{a}})$ is locally infinite.

LEMMA 23.9. *The sum in Definition 23.8 is well defined: if $\{V_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ is another open cover with the property that each $V_{\mathtt{b}}$ is contained in the interior of the image of some orientation preserving singular cube and $\{\nu_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ is a partition of unity subordinate to that cover then for any $\omega \in \Omega_c^n(M)$ one has:*

$$\sum_{\mathtt{a} \in \mathtt{A}} \int_M \lambda_{\mathtt{a}}\, \omega = \sum_{\mathtt{b} \in \mathtt{B}} \int_M \nu_{\mathtt{b}}\, \omega$$

Proof. Since

$$\sum_{\mathtt{a}\in\mathtt{A}} \lambda_{\mathtt{a}}(x) = \sum_{\mathtt{b}\in\mathtt{B}} \nu_{\mathtt{b}}(x) = 1, \qquad \forall\, x \in M,$$

we have using linearity of the integral that

$$\begin{aligned}
\sum_{\mathtt{a}\in\mathtt{A}} \int_M \lambda_{\mathtt{a}}\,\omega &= \sum_{\mathtt{a}\in\mathtt{A}} \int_M \left(\sum_{\mathtt{b}\in\mathtt{B}} \nu_{\mathtt{b}} \right) \lambda_a\,\omega \\
&= \sum_{\mathtt{a}\in\mathtt{A}} \sum_{\mathtt{b}\in\mathtt{B}} \int_M \nu_{\mathtt{b}}\,\lambda_{\mathtt{a}}\,\omega \\
&= \sum_{\mathtt{b}\in\mathtt{B}} \int_M \left(\sum_{\mathtt{a}\in\mathtt{A}} \lambda_{\mathtt{a}} \right) \nu_{\mathtt{b}}\,\omega \\
&= \sum_{\mathtt{b}\in\mathtt{B}} \int_M \nu_{\mathtt{b}}\,\omega
\end{aligned}$$

where the rearrangement of the sums is justified as everything is a finite sum. ∎

We now know how to integrate a top-dimensional differential form with compact support on an oriented manifold. Let us extend this to oriented manifolds with boundary. For this we use the following trick:

Definition 23.10. Let M^n be an oriented smooth manifold with boundary. An orientation preserving singular cube $c\colon C^n \to M$ is called **special**[2] if either $\operatorname{im} c \subset \operatorname{int}(M)$ or

$$\partial M \cap \operatorname{im} c = \operatorname{im}(F_1 c),$$

where as usual $F_1 c\colon C^{n-1} \to M$ is the first front face.

Lemma 23.11. *Let M^n be an oriented smooth manifold with boundary, and endow ∂M with the induced orientation. If $c\colon C^n \to M$ is a special singular cube such that $\operatorname{im} c \cap \partial M \neq \emptyset$ then $F_1 c$ is an orientation reversing singular cube for ∂M.*

This is *not* a typo—we really do want $F_1 c$ to reverse orientation! As we shall see, the minus sign will eventually cancel, since the coefficient of $F_1 c$ in ∂c is -1.

Proof. We need only check that $F_1 c$ is orientation reversing with respect to the induced orientation. Let x^i denote the standard coordinates on C^n. Since c is a diffeomorphism, we may take c^{-1} as a $\mathbb{R}^n_+$ half-space chart on M (remember c is really defined on an open neighbourhood of C^n). Let $y^i := x^i \circ c^{-1}$ denote the local coordinates of this chart. Since c is orientation preserving, $\left\{ \frac{\partial}{\partial y^1}, \dots, \frac{\partial}{\partial y^n} \right\}$ is a positive oriented local frame of TM. Note that $\frac{\partial}{\partial y^1}$ is an *inward* pointing

[2]It would be slightly more logical to call such a cube "bad" rather than "special", as I will now explain. Indeed, we defined a "good" half-space chart to be one where the half-space in question was $\mathbb{R}^n_-$. Thus one could quite logically declare a "bad" half-space chart to be one where the half-space was the opposite half-space $\mathbb{R}^n_+$. As we will shortly see, a "special" orientation preserving cube can be thought of as the inverse of an (appropriately restricted) bad chart in this sense. Thus the adjective "bad" makes perfect sense for such a cube. However, I prefer the label "special", since otherwise it would seem rather strange that we defined integration on manifolds with boundary using "bad" cubes...

section (compare Example 21.19—the reason we get inward pointing not outward pointing is that this is a $\mathbb{R}^n_+$ chart not a $\mathbb{R}^n_-$ chart!), and thus (cf. the paragraph after Definition 21.21) the frame $\left\{\frac{\partial}{\partial y^2}, \dots, \frac{\partial}{\partial y^n}\right\}$ is a negatively oriented frame for $T(\partial M)$. Thus $F_1 c$ is orientation reversing as required. ■

Just as in Remark 23.3, if M is a smooth manifold with boundary then we can always find an open cover of M with the property that each open set is contained in the interior of the image of a special orientating preserving singular cube. We use this to extend the definition of integration to manifolds with boundary.

Definition 23.12. Let M^n be an oriented smooth manifold with boundary, and let $\omega \in \Omega^n_c(M)$. Let $\{U_\mathtt{a} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover with the property that each $U_\mathtt{a}$ is contained in the image of some special orientation preserving singular cube (cf. Remark 23.3). Let $\{\lambda_\mathtt{a} \mid \mathtt{a} \in \mathtt{A}\}$ be a partition of unity subordinate to this cover. We define

$$\int_M \omega := \sum_{\mathtt{a} \in \mathtt{A}} \int_M \lambda_\mathtt{a} \, \omega.$$

The same proof as Lemma 23.9 shows this is well-defined. We now state and prove the Global Stokes' Theorem.

Theorem 23.13 (The Global Stokes' Theorem). *Let M^n be an oriented smooth manifold with boundary, and endow ∂M with the induced orientation. Let $\omega \in \Omega^{n-1}_c(M)$. Then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. We prove the result in two steps. Let $\{U_\mathtt{a} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover with the property that each $U_\mathtt{a}$ is contained in the image of some special cube.

1. First assume that $\operatorname{supp}(\omega)$ is contained in one of the sets $U_\mathtt{a}$, which itself is contained in the image of some special cube c. If $\operatorname{im} c \cap \partial M = \emptyset$ then the result is immediate from the Local Stokes' Theorem 22.25, since

$$\int_M d\omega = \int_c d\omega = \int_{\partial c} \omega = 0,$$

since $\operatorname{supp}(\omega)$ does not intersect the image of ∂c. But also clearly $\int_{\partial M} \omega = 0$ since $\operatorname{supp}(\omega)$ does not intersect ∂M.

Now assume that $\operatorname{im} c \cap \partial M \neq \emptyset$. Then we have:

$$\begin{aligned}
\int_M d\omega &= \int_c d\omega \\
&= \int_{\partial c} \omega \\
&= \sum_{i=1}^n (-1)^i \left(\int_{F_i c} \omega - \int_{B_i c} \omega \right) \\
&= -\int_{F_1 c} \omega,
\end{aligned}$$

since $\mathrm{supp}(\omega)$ misses all faces apart from $F_1 c$ by definition of a special cube. Thus by Lemma 23.7 and Lemma 23.11 we have

$$\int_M d\omega = -\int_{F_1 c} \omega = (-1)^2 \int_{\partial M} \omega = \int_{\partial M} \omega.$$

2. Now we prove the general case. Let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be a partition of unity subordinate to the open cover $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$. Then by definition,

$$\begin{aligned}
\int_{\partial M} \omega &= \sum_{\mathtt{a} \in \mathtt{A}} \int_{\partial M} \lambda_{\mathtt{a}}\, \omega \\
&\overset{(\dagger)}{=} \sum_{\mathtt{a} \in \mathtt{A}} \int_M d(\lambda_{\mathtt{a}}\, \omega) \\
&= \sum_{\mathtt{a} \in \mathtt{A}} \int_M d\lambda_{\mathtt{a}} \wedge \omega + \lambda_{\mathtt{a}}\, d\omega \\
&\overset{(\ddagger)}{=} \int_M d\omega + \sum_{\mathtt{a} \in \mathtt{A}} \int_M d\lambda_{\mathtt{a}} \wedge \omega \\
&= \int_M d\omega + \int_M d\left(\sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}}\right) \wedge \omega \\
&\overset{(\ddagger)}{=} \int_M d\omega + 0,
\end{aligned}$$

where $(\dagger)$ used Step 1, both $(\ddagger)$ used the fact that $\sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}} \equiv 1$, and the interchange of summation and integral is always justified as these are always finite sums as $\mathrm{supp}(\omega)$ is compact. This completes the proof.

■

Corollary 23.14. *Let M be a smooth manifold*[3]*, and let $\omega \in \Omega_c^{n-1}(M)$. Then $\int_M d\omega = 0$.*

Proof. M is also a smooth manifold with boundary whose boundary is empty. ■

Corollary 23.15. *Let M^n be an oriented connected compact smooth manifold. Then $H^n_{\mathrm{dR}}(M) \neq 0$.*

Proof. Let μ be a volume form. Then for any orientation preserving cube, we have $\int_c \mu > 0$. Thus $\int_M \mu > 0$. The form μ is closed (as $d\mu = 0$ for dimension reasons). If μ was exact then $\int_M \mu = 0$ by Corollary 23.14. Thus μ is a closed non-exact form, and hence defines a non-zero element in $H^n_{\mathrm{dR}}(M)$. ■

We will eventually improve Corollary 23.15 and show that for an oriented connected compact smooth manifold M^n, one has

$$H^n_{\mathrm{dR}}(M) \cong \mathbb{R}.$$

To prove this we will first establish the **homotopy invariance** property of de Rham cohomology. The key to this is the following innocuous looking statement.

[3]Remember, all manifolds are assumed not to have boundary unless explicitly stated!

PROPOSITION 23.16. *Let M be a smooth manifold. Define for $t \in [0,1]$ a smooth map*

$$\jmath_t \colon M \to M \times [0,1], \qquad \jmath_t(x) := (x,t),$$

(here we view $M \times [0,1]$ as a smooth manifold with boundary). There is a map

$$h \colon \Omega^r(M \times [0,1]) \to \Omega^{r-1}(M)$$

such that for every differential r-form $\omega \in \Omega^r(M \times [0,1])$, one has

$$h(d\omega) + d(h(\omega)) = \jmath_1^\star(\omega) - \jmath_0^\star(\omega)$$

as elements of $\Omega^r(M)$. Thus the induced maps on de Rham cohomology

$$\jmath_0^\star,\ \jmath_1^\star \colon H^r_{\mathrm{dR}}(M \times [0,1]) \to H^r_{\mathrm{dR}}(M)$$

coincide.

Proof. Let Y denote the vector field on $M \times [0,1]$ whose value at (x,t) is

$$Y(x,t) = \left(0, \frac{\partial}{\partial t}\Big|_t\right). \tag{23.2}$$

(Compare (8.3) in Lecture 8—we are using slightly different notation to simplify the formulae to come). The desired map h is then given by

$$h(\omega) := \int_0^1 \jmath_t^\star(i_Y(\omega))\, dt,$$

for $\omega \in \Omega^r(M \times I)$. That is, for any $x \in M$,

$$h(\omega)_x = \int_0^1 \jmath_t^\star(i_Y(\omega)_{(x,t)})\, dt,$$

where the integrand is thought of as a function of t on the vector space $\bigwedge^{r-1}(T_x^* M)$. We emphasise that this is just a normal integral, not an integral on a manifold! By choosing local coordinates, we see that the integral defines a smooth $(r-1)$-form on M. To compute $d(h(\omega))$ it suffices to work locally. In local coordinates (x^i) we can't express $h(\omega)$ as a sum of terms of the form

$$\left(\int_0^1 f(x,t)\, dt\right) dx^I,$$

using the notation introduced in the proof of Theorem 19.17. Applying d to such a term and differentiating under the integral sign[4] gives

$$\sum_j \frac{\partial}{\partial x^j}\left(\int_0^1 f(x,t)\, dt\right) dx^j \wedge dx^I = \left(\int_0^1 \sum_j \frac{\partial f}{\partial x^j}(x,t)\, dt\right) dx^j \wedge dx^I$$

[4]This is sometimes referred to as the Leibniz integral rule.

Putting this together shows us that

$$d(h(\omega)) = \int_0^1 d\left(\jmath_t^\star(i_Y(\omega))\right) dt.$$

Thus using Lemma 19.19 and Cartan's Magic Formula (Theorem 20.6) we see that

$$\begin{aligned} h(d\omega) + d(h(\omega)) &= \int_0^1 \left(\jmath_t^\star(i_Y(d\omega)) + d\left(\jmath_t^\star(i_Y(\omega))\right)\right) dt \\ &= \int_0^1 \left(\jmath_t^\star(i_Y(d\omega)) + \jmath_t^\star(d(i_Y(\omega)))\right) dt \\ &= \int_0^1 \jmath_t^\star(\mathcal{L}_Y(\omega))\, dt. \end{aligned}$$

Let θ_t denote the flow of Y. Then $\theta_t(x,s) = (x, t+s)$, and thus $\jmath_t = \theta_t \circ \jmath_0$ and we can compute the Lie derivative as

$$\begin{aligned} \jmath_t^\star(\mathcal{L}_Y(\omega)) &= \jmath_0^\star(\theta_t^\star(\mathcal{L}_Y(\omega))) \\ &\overset{(\dagger)}{=} \jmath_0^\star\left(\frac{d}{dt}\theta_t^\star(\omega)\right) \\ &= \frac{d}{dt}\jmath_0^\star(\theta_t^\star(\omega)) \\ &= \frac{d}{dt}\jmath_t^\star(\omega). \end{aligned}$$

where $(\dagger)$ used Problem L.5. Thus by the (normal) Fundamental Theorem of Calculus we obtain

$$h(d\omega) + d(h(\omega)) = \int_0^1 \frac{d}{dt}\jmath_t^\star(\omega)\, dt = \jmath_1^\star(\omega) - \jmath_0^\star(\omega).$$

To see the last statement, we take a closed r-form on $M \times I$. Then

$$\jmath_1^\star[\omega] - \jmath_0^\star[\omega] = [h(d\omega) + d(h(\omega))] = 0.$$

This completes the proof. ∎

We can now prove the following key result.

THEOREM 23.17. *Let M and N be two smooth manifolds and suppose φ and ψ are two homotopic smooth maps from M to N. Then the induced maps $\varphi^\star$ and $\psi^\star$ on the de Rham cohomology groups are the same.*

Proof. To say that φ and ψ are homotopic means there is a *continuous* map $H\colon M \times [0,1] \to N$ such that $H(\cdot, 0) = \varphi$ and $H(\cdot, 1) = \psi$. In fact, by the Whitney Approximation Theorem, we may assume H is a *smooth* map (see Remark 6.17). We conclude with

$$\varphi^\star = (H \circ \jmath_0)^\star = \jmath_0^\star \circ H^\star = \jmath_1^\star \circ H^\star = (H \circ \jmath_1)^\star = \psi^\star.$$

This completes the proof. ∎

Recall in general two topological spaces X and Y are said to be **homotopy equivalent** if there exists continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that both $f \circ g$ and $g \circ f$ are homotopic to the respective identity maps.

Corollary 23.18 (Homotopy invariance of de Rham cohomology). *Let M and N be smooth manifolds that are homotopy equivalent. Then M and N have isomorphic de Rham cohomology groups.*

Proof. Let $f\colon M \to N$ and $g\colon N \to M$ be continuous maps such that $f \circ g$ and $g \circ f$ are homotopic to the identity maps. By the Whitney Approximation Theorem we may approximate f and g by smooth maps φ and ψ. Then $\varphi \circ \psi$ and $\psi \circ \varphi$ are homotopic to the identity maps. By Theorem 23.17, $(\varphi \circ \psi)^\star$ and $(\psi \circ \varphi)^\star$ coincide with the maps induced by the identity. Since $\mathrm{id}^\star$ is clearly the identity, we see that $\varphi^\star$ is an inverse to $\psi^\star$. The claim follows. ∎

Remark 23.19. A particular case of Corollary 23.18 tells us that the de Rham cohomology cannot see the smooth structure on a topological manifold M. This is surprising, since the data that defined it (the differential forms) very much depend on the choice of smooth structure.

A topological space is **contractible** if it is homotopy equivalent to a point.

Corollary 23.20. *Let M be contractible. Then $H^r_{\mathrm{dR}}(M) = 0$ for all $r \geq 1$.*

Proof. It is clear this is true for M equal to a point. Now apply Corollary 23.18. ∎

Remark 23.21. This shows that de Rham cohomology cannot distinguish Euclidean spaces: $H^r_{\mathrm{dR}}(\mathbb{R}^n)$ is independent of n (since all Euclidean spaces are contractible). Thus a lot of information is lost when passing to de Rham cohomology.

Perhaps the most useful corollary of this is the following statement, which is classically called the **Poincaré Lemma**.

Corollary 23.22 (The Poincaré Lemma). *Let M be a smooth manifold and let $\omega \in \Omega^r(M)$ be a closed differential form. For any point $x \in M$ there exists a neighbourhood U of x such that $\omega|_U$ is an exact form in $\Omega^r(U)$.*

Proof. Every point in a manifold admits a contractible neighbourhood. ∎

LECTURE 24

Principal bundles

In Lecture 13 we defined fibre bundles, and then for much of the course we focused on the special case where the fibre was a vector space, and the structure group was a matrix Lie group. In this lecture we will switch to another special case, where the fibre is a Lie group which acts freely on the total space. These are called *principal bundles*. Although principal bundles are not quite as ubiquitous in differential geometry as vector bundles, we will see that they are equally fundamental. Next semester in Differential Geometry II it will be convenient to switch back and forth from vector bundles to principal bundles as the mood so takes us, especially when it comes to defining connections.

The starting point for a principal bundle is a Lie group acting on a manifold on the *right*. The following definition is identical to Definition 10.17, except for the action is now a right action not a left action.

Definition 24.1. Let G be a Lie group and let P be a manifold. A smooth map $\mu\colon P \times G \to P$ satisfying

$$\mu(p, ab) = \mu(\mu(p, a), b), \qquad \mu(p, e) = p$$

for all $a, b \in G$ and $p \in P$ is called **a right action of G on P**. The action is **free** if $\mu(p, a) = p$ for some $p \in P$ and $a \in G$ implies $a = e$.

Remark 24.2. The difference between right actions and left actions is mainly for notational convenience, since we can always convert one into the other. Indeed, if $\mu\colon G \times P \to P$ is a left action then we can define a right action $\tilde{\mu}\colon P \times G \to P$ by $\tilde{\mu}(p, a) := \mu(a^{-1}, p)$, and conversely.

To keep the notation under control, for all principal bundles we will typically suppress the map μ and just write the action as $(p, a) \mapsto p \cdot a$. This should not be confusing, since there is no danger of overlap: points in G are always written with a, b etc and points in P are always written with p, q etc, and the multiplication in G itself is simply written as juxtaposition. This will become particularly important when there are multiple actions in play at the same time (see Theorem 25.3 in the next lecture).

Definition 24.3. Let $\pi\colon P \to M$ be a fibre bundle with fibre a Lie group G. Assume moreover that there exists a free fibre-preserving right action of G on P and a bundle atlas for P with the property that each bundle chart $\alpha\colon \pi^{-1}(U) \to G$ is **G-equivariant** in the sense that

$$\alpha(p \cdot a) = \alpha(p)a, \qquad \forall\, p \in \pi^{-1}(U),\ \forall\, a \in G.$$

Then we say that P is a **principal bundle** over M with group G (or a **G-principal bundle**), and we refer to the bundle atlas as a **principal bundle atlas**, and its charts as **principal bundle charts**.

Example 24.4. The simplest example of a principal is the trivial bundle $\mathrm{pr}_1 \colon M \times G \to M$, where the action is given by $(x, a) \cdot b := (x, ab)$.

Here are two basic properties of principal bundles.

Lemma 24.5. *Let $\pi \colon P \to M$ be a G-principal bundle. Then the structure group of P is G itself, where we let G act on itself via left translations.*

Proof. Suppose $\alpha \colon \pi^{-1}(U) \to G$ and $\beta \colon \pi^{-1}(V) \to G$ are two principal bundle charts such that $U \cap V \neq \emptyset$. Fix $a \in G$ and suppose $p \in P_x$ is the unique point such that $\beta|_{P_x}(p) = a$. Let $b := \alpha|_{P_x}(p)$. Then

$$\rho_{\alpha\beta}(x)(a) \overset{\text{def}}{=} \alpha|_{P_x} \circ \beta|_{P_x}^{-1}(a) = b.$$

Set $c := ba^{-1}$. We claim that $\rho_{\alpha\beta}(x) = l_c$. For this take an arbitrary element $a_1 \in G$. Then we can write $a_1 = aa_2$ for a unique a_2 (namely, $a_2 := aa_1^{-1}$). Then by G-equivariance, $\beta(p \cdot a_2) = \beta(p)a_2 = aa_2 = a_1$. Moreover $\alpha(p \cdot a_2) = \alpha(p)a_2 = ba_2 = ba^{-1}a_1 = ca_1 = l_c(a_1)$. Thus

$$\rho_{\alpha\beta}(x)(a_1) = \alpha|_{P_x} \circ \beta|_{P_x}^{-1}(a_1) = \alpha(p \cdot a_2) = l_c(a_1).$$

The group G acts on itself via left translation, and thus we can regard G as a subgroup of $\mathrm{Diff}(G)$ via $a \mapsto l_a$. We have thus shown that $\rho_{\alpha\beta}(x)$ belongs to this subgroup for each $x \in U \cap V$, which shows that we may take the structure group of P to be G (cf. Definition 13.9). This completes the proof. ■

Lemma 24.6. *Let $\pi \colon P \to M$ be a G-principal bundle. Then the fibres are exactly the orbits of the G-action (this means that G acts transitively on the fibres), and hence (as topological spaces) M is the quotient space P/G.*

Proof. Fix $x \in M$ and suppose $p, q \in P_x$. Let $\alpha \colon \pi^{-1}(U) \to G$ denote a principal bundle chart over a neighbourhood U of x. Let $a := \alpha(p)$ and $b := \alpha(q)$. Then since (π, α) is a diffeomorphism, we have

$$(\pi, \alpha)(p \cdot a^{-1}b) = (x, \alpha(p)a^{-1}b) = (x, b) = (\pi, \alpha)(q).$$

Thus $q = p \cdot a^{-1}b$. Conversely it is immediate that $p \cdot a \in P_x$ for all $a \in G$, and thus P_x is in bijection with G, as was to be proved. ■

The converse to Lemma 24.5 is also true. This requires the following auxiliary result. Recall the notion of a *local smooth section* from Corollary 12.5: if $\varphi \colon M \to N$ is a smooth map then a local smooth section is a map ψ defined on an open set V of N such that $\varphi \circ \psi$ is the identity on V.

Lemma 24.7. *Let $\varphi \colon M^n \to N^k$ be a surjective submersion. Then every point $x \in M$ is in the image of a local smooth section, and φ is a quotient map.*

Proof. Let $x \in M$. From the Implicit Function Theorem 5.3, we may choose a chart $\sigma \colon U \to O$ on M about x and a chart $\tau \colon V \to \Omega$ on N about $\varphi(x)$ such that $\tau \circ \varphi \circ \sigma^{-1}$ is of the form

$$(x^1, \dots, x^n) \mapsto (x^1, \dots, x^k). \tag{24.1}$$

By shrinking the domains if necessary we may assume $O = O_1 \times O_2 \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$ and $\Omega = O_1$. Let $z := \pi_2 \circ \sigma(x)$, where $\pi_2 \colon O_1 \times O_2 \to O_2$ is the second projection. Then let $i_z \colon O_1 \to O_1 \times O_2$ denote the map $i_z(w) = (w, z)$. A local smooth section for φ at x is given by $\psi := \sigma^{-1} \circ i_z \circ \tau$.

The fact that φ is an open map is clear from the representation (24.1), since a small open cube $\{|x^i| < \varepsilon \mid i = 1, \dots, n\}$ is mapped onto the small open cube $\{|x^i| < \varepsilon, \mid i = 1, \dots, k\}$. Finally, an surjective open map is necessarily a quotient map. ∎

Here is the promised converse to Lemma 24.5.

Proposition 24.8. *Let $\pi \colon P \to M$ denote a surjective submersion and let G denote a Lie group. Assume G acts freely on P in such a way that the orbit of a point $p \in P$ is exactly the fibre $\pi^{-1}(\pi(p))$. Then $\pi \colon P \to M$ is a principal G-bundle.*

Proof. Firstly, we may assume that G acts on P on the right, since if the action is a left action then we can convert it into a right action via Remark 24.2. Next, for each $x \in M$, there is a local smooth section $\psi \colon U \to P$ of π defined on a neighbourhood U of x. Consider the map

$$\Psi \colon U \times G \to \pi^{-1}(U), \qquad \Psi(y, a) := \psi(y) \cdot a. \tag{24.2}$$

By hypothesis the map Ψ is a smooth bijection. In fact, Ψ is a diffeomorphism by the Inverse Function Theorem 5.2, since its derivative is invertible (see Problem Sheet M). Thus we can write $\Psi^{-1} = (\pi, \alpha)$ for a uniquely determined smooth function $\alpha \colon \pi^{-1}(U) \to G$. This will form our desired principal bundle chart once we check G-equivariance. Let $p \in \pi^{-1}(U)$ and assume that $\pi(p) = y \in U$. Then for $a \in G$ we compute:

$$\begin{aligned} \Psi(y, \alpha(p)a) &= \psi(y) \cdot \alpha(p)a \\ &= \Psi\big(y, \alpha(p)\big) \cdot a \\ &= \Psi \circ (\pi, \alpha)(p) \cdot a \\ &\overset{(\dagger)}{=} p \cdot a \\ &= \Psi\big(y, \alpha(p \cdot a)\big), \end{aligned}$$

where $(\dagger)$ used that $\Psi^{-1} = (\pi, \alpha)$. Since Ψ is a diffeomorphism this shows that $\alpha(pa) = \alpha(p)a$. ∎

Thus the results of Lecture 12 tell us:

Corollary 24.9. *Let H be a closed Lie subgroup of a Lie group G. Then $\pi \colon G \to G/H$ is a principal H-bundle. More generally, if $M \cong G/H$ is a homogeneous space then $G \to M$ is a principal H-bundle.*

We will now state a more general version of Theorem 12.4 that applies in other situations.

Definition 24.10. A right action of a Lie group G on a manifold P is called **proper** if the map $P \times G \to P \times G$ given by $(p, a) \mapsto (p \cdot a, a)$ is a proper map (in the usual sense).

Theorem 24.11. *Let P be a manifold and let G be a Lie group acting freely and properly. Then P/G admits a unique smooth structure such that the projection $\pi\colon P \to P/G$ is a surjective submersion, where P/G is given the quotient topology.*

We won't prove Theorem 24.11, as the proof is technical and long and most of the interesting details were already contained in the proof of Theorem 12.4. However let us note that combining Theorem 24.11 and Proposition 24.8 gives us more examples of principal bundles:

Corollary 24.12. *Let P be a manifold and let G be a Lie group acting freely and properly. Then $\pi\colon P \to P/G$ is a principal G-bundle (after making the action a right action if necessary, cf Remark 24.2).*

We have introduced four different types of actions of Lie groups on manifolds:

- **free** actions (Definition 24.1),
- **effective** actions (Definition 13.8),
- **transitive** actions (Definition 12.9),
- **proper** actions (Definition 24.10).

Make sure you are aware of the difference in meaning of all of these words! A free action is necessarily effective, but other than that there are no relations between the four concepts.
Warning: In addition, be aware that the terminology is not entirely standard, and some texts define them differently.

We now move onto an extremely important example of a principal bundle, which will illustrate the deep link between principal bundles and vector bundles. For this let $\pi\colon E \to M$ be a vector bundle of rank k over M. Fix $x \in M$, and let $\mathrm{Fr}(E_x)$ denote the set of isomorphisms $T\colon \mathbb{R}^k \to E_x$. Any two isomorphisms $T, T_1\colon \mathbb{R}^k \to E_x$ differ by element of $\mathrm{GL}(k)$, i.e. $T_1 = T \circ A$. In fact, if we fix our favourite isomorphism T then the map $\mathrm{GL}(k) \to \mathrm{Fr}(E_x)$ given by $A \mapsto T \circ A$ is a bijection.

One can equivalently regard $\mathrm{Fr}(E_x)$ as the set of bases of the vector space E_x, since for any $T \in \mathrm{Fr}(E_x)$ the vectors (Te_i) form a basis of E_x, where e_i are the standard basis vectors in $\mathbb{R}^k$, and conversely given a basis (v_i) there is a uniquely determined linear isomorphism $T\colon \mathbb{R}^k \to E_x$ such that $Te_i = v_i$ for each i.

Definition 24.13. We now form the total space

$$\mathrm{Fr}(E) := \bigsqcup_{x \in M} \mathrm{Fr}(E_x),$$

and let $\hat{\pi}\colon \mathrm{Fr}(E) \to M$ denote the map that sends $\mathrm{Fr}(E_x)$ to x. We call $\mathrm{Fr}(E)$ the **frame bundle** of E.

Proposition 24.14. *Let $\pi\colon E \to M$ be a vector bundle of rank k over M. Then $\hat{\pi}\colon \mathrm{Fr}(E) \to M$ is a principal $\mathrm{GL}(k)$-bundle over M.*

Proof. Let $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ denote a vector bundle chart on E over U. Then $\alpha|_{E_x}^{-1}\colon \mathbb{R}^k \to E_x$ is a linear isomorphism, and thus $\alpha|_{E_x}^{-1} \in \mathrm{Fr}(E_x)$ for each $x \in U$. Define a map

$$\hat{\alpha}\colon \hat{\pi}^{-1}(U) \to \mathrm{GL}(k),$$

by declaring that

$$\hat{\alpha}(\alpha|_{E_x}^{-1} \circ A) = A.$$

We will show that $\hat{\alpha}$ is a principal bundle chart on $\mathrm{Fr}(E)$. For this, suppose $\beta\colon \pi^{-1}(V) \to \mathbb{R}^k$ is another vector bundle chart on E such that $U \cap V \neq \emptyset$. Let $\hat{\beta}$ denote the corresponding bundle chart on $\mathrm{Fr}(E)$. We compute the transition function $\rho_{\hat{\alpha}\hat{\beta}}$ as follows:

$$\begin{aligned}
\rho_{\hat{\alpha}\hat{\beta}}(x)(A) &= \hat{\alpha}|_{\mathrm{Fr}(E_x)} \circ \hat{\beta}|_{\mathrm{Fr}(E_x)}^{-1}(A) \\
&= \hat{\alpha}|_{\mathrm{Fr}(E_x)} \left(\beta|_{E_x}^{-1} \circ A\right) \\
&= \rho_{\alpha\beta}(x) \circ A
\end{aligned}$$

Thus the transition functions of $\mathrm{Fr}(E)$ is just left composition by the transition functions of E. In particular, the transition function $\rho_{\hat{\alpha}\hat{\beta}}$ is smooth, and thus $(\hat{\pi}, \hat{\alpha})$ and $(\hat{\pi}, \hat{\beta})$ defined the same smooth structure on $\hat{\pi}^{-1}(U \cap V)$. Thus we can use Remark 13.7 to endow $\mathrm{Fr}(E)$ with the structure of a smooth manifold such that that $\hat{\alpha}$ become bundle charts. We have thus proved that $\mathrm{Fr}(E)$ is a fibre bundle over M with fibre $\mathrm{GL}(k)$.

It remains to show that this is a principal bundle. For this first note that the map $\mathrm{Fr}(E) \times \mathrm{GL}(k) \to \mathrm{Fr}(E)$ given by $(T, A) \mapsto T \circ A$ is a smooth free action (it is free as we have already observed it is free on each fibre). It remains to check $\mathrm{GL}(k)$-invariance of the bundle charts. If $\hat{\alpha}$ is a chart and $A \in \mathrm{Fr}(E_x)$ then write $A = \alpha|_{E_x}^{-1} \circ A_1$. Then for $B \in \mathrm{GL}(k)$ one has

$$\hat{\alpha}(A \circ B) = \hat{\alpha}(\alpha|_{E_x}^{-1} \circ A_1 \circ B) = A_1 \circ B = \hat{\alpha}(A) \circ B.$$

■

Let us now address what it means for two principal bundles to be isomorphic. Note we only discussed isomorphisms of vector bundles (cf Definition 14.3), not for general fibre bundles, and thus the following definition is not a special case of another one. Nevertheless, it is fairly easy to "guess"—one just replaces "linear" with "equivariant" wherever appropriate.

Definition 24.15. Let $\pi_i\colon P_i \to M_i$ be two G-principal bundles. Suppose we are given two smooth maps $\Phi\colon P_1 \to P_2$ and $\varphi\colon M_1 \to M_2$. We say that Φ is a **principal bundle morphism along** φ if the restriction of Φ to each fibre $P_1|_x$ is a map $P_1|_x \to P_2|_{\varphi(x)}$. Thus the following commutes:

$$\begin{tikzcd}
P_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & P_2 \arrow[d, "\pi_2"] \\
M_1 \arrow[r, "\varphi"'] & M_2
\end{tikzcd}$$

Moreover Φ should be equivariant with respect to the two G-actions in the sense that

$$\Phi(p \cdot a) = \Phi(p) \cdot a, \qquad \forall p \in P_1, \ \forall a \in G. \tag{24.3}$$

Here the $\cdot$ on the left-hand side is the G-action in P_1, and the $\cdot$ on the right-hand side is the G-action in P_2. If Φ is a diffeomorphism then we say that Φ is a **principal bundle isomorphism along φ**.

The following lemma illustrates a key difference between vector bundle morphisms and principal bundle morphisms.

LEMMA 24.16. *Let $\pi_i \colon P_i \to M_i$ be two G-principal bundles. Suppose $\Phi \colon P_1 \to P_2$ is a principal bundle morphism along a diffeomorphism $\varphi \colon M_1 \to M_2$. Then Φ is also a diffeomorphism.*

Warning: this is *not* true for vector bundle morphisms! The proof of Lemma 24.16 is also on Problem Sheet M.

If $M_1 = M_2$ and $\varphi = \mathrm{id}$ then we just call Φ a **principal bundle isomorphism**. Note there is no point introducing a "principal bundle homomorphism" and then declaring a principal bundle isomorphism to be one which is in addition a diffeomorphism, because any principal bundle homomorphism is necessarily then an isomorphism as the identity map is a diffeomorphism.

DEFINITION 24.17. If $\pi \colon P \to M$ is a principal bundle and $\Phi \colon P \to P$ is a principal bundle isomorphism from P to itself then we call Φ a **gauge transformation**. We will come back to the study of gauge transformations extensively next semester.

One can also extend the notion of a principal bundle morphism for different Lie groups.

DEFINITION 24.18. Suppose G and H are two Lie groups. Let $\pi_1 \colon P_1 \to M_1$ be a principal G-bundle and let $\pi_2 \colon P_2 \to M_2$ be a principal H-bundle. Suppose $\psi \colon G \to H$ is a Lie group homomorphism. A **principal bundle morphism from P_1 to P_2 with respect to ψ** consists of a pair of smooth maps $\varphi \colon M_1 \to M_2$ and $\Phi \colon P_1 \to P_2$ such that the diagram from Definition 24.15 commutes, and with the equivariance condition (24.3) replaced by

$$\Phi(p \cdot a) = \Phi(p) \cdot \psi(a), \qquad \forall\, p \in P_1, \ a \in G.$$

If $M_1 = M_2$ and $\varphi = \mathrm{id}$ then we call Φ a **principal bundle homomorphism with respect to ψ**. In this case the analogue of Lemma 24.16 is *not* true, and so Φ does not need to be a principal bundle isomorphism with respect to ψ.

A special case of this gives rise to the notion of a principal subbundle.

DEFINITION 24.19. Let G be a Lie group and suppose $H \subset G$ is a Lie subgroup. Suppose $\pi_1 \colon Q \to M$ is a principal H-bundle and $\pi_2 \colon P \to M$ is a principal G-bundle such that $Q \subset P$. We say that Q is a **principal H-subbundle** of P if the inclusion $Q \hookrightarrow P$ is a principal bundle homomorphism with respect to the inclusion $H \hookrightarrow G$.

Here is a useful criterion for constructing principal subbundles.

PROPOSITION 24.20. *Let $\pi\colon P \to M$ be a principal G-bundle. Let $H \subset G$ be a Lie subgroup, and let $Q \subset P$ be a subset such that:*

(i) *The restriction $\pi|_Q \colon Q \to M$ is surjective.*

(ii) *If $q \in Q$ and $a \in H$ then $q \cdot a \in Q$.*

(iii) *For all $x \in M$, the action of H on $Q \cap P_x$ is transitive.*

(iv) *For all $x \in M$, there exists a neighbourhood U of x and a local section $s\colon U \to P$ such that $s(y) \in Q$ for all $y \in U$.*

Then $\pi|_Q \colon Q \to M$ is a principal H-bundle, and moreover Q is a principal H-subbundle of P.

Proof. The proof is a variation on the proof of Proposition 24.8. Fix $x \in M$ and let $U \subset M$ be a neighbourhood of x such that there exists a section $s\colon U \to P$ such that $s(y) \in Q$ for all $y \in U$. For every point $q \in Q \cap \pi^{-1}(U)$, there exists a unique $a \in H$ such that

$$q = s(\pi(q)) \cdot a.$$

Define $\alpha\colon Q \cap \pi^{-1}(U) \to H$ by

$$\alpha(\pi(q) \cdot a) = a.$$

Then $(\pi|_Q, \alpha)$ is a bijection from $Q \cap \pi^{-1}(U) \to U \times H$. We define a topology and smooth structure on Q by declaring this to be a diffeomorphism. The rest of the proof is now analogous to that of Proposition 24.8. ■

Let us also state explicitly the following result, whose proof is a special case of Theorem 14.1.

THEOREM 24.21. *Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open covering of a manifold M. Let G be a Lie group. Suppose for each $\mathtt{a}, \mathtt{b} \in \mathtt{A}$ such that $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$, we are given a smooth map $\rho_{\mathtt{ab}}\colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G$ such that the following cocycle condition is satisfied:*

$$\begin{cases} \rho_{\mathtt{ac}}(x) = \rho_{\mathtt{ab}}(x)\rho_{\mathtt{bc}}(x), & \forall x \in U_{\mathtt{a}} \cap U_{\mathtt{b}} \cap U_{\mathtt{c}}, \text{ if } U_{\mathtt{a}} \cap U_{\mathtt{b}} \cap U_{\mathtt{c}} \neq \emptyset, \\ \rho_{\mathtt{aa}}(x) = e, & \forall x \in U_{\mathtt{a}}, \ \forall \mathtt{a} \in \mathtt{A}. \end{cases} \tag{24.4}$$

Then there exists a G-principal bundle $\pi\colon P \to M$. Moreover there is a principal bundle atlas $\{\alpha_{\mathtt{a}}\colon \pi^{-1}(U_{\mathtt{a}}) \to G \mid \mathtt{a} \in \mathtt{A}\}$ such that the transition function $\rho_{\alpha_{\mathtt{a}}\alpha_{\mathtt{b}}}$ are given by $\rho_{\mathtt{ab}}$.

Proof. Apply Theorem 14.1 with $F = G$ and the action given by left translations. This gives a fibre bundle $\pi\colon P \to M$ with fibre G and the transition functions as stated. It thus remains to check that the corresponding bundle charts are G-equivariant, but this is immediate from the proof of Theorem 14.1. ■

Just as with vector bundles, principal bundles are determined up to isomorphism by their transition functions.

Proposition 24.22. *Let M be a smooth manifold and suppose $\pi_i \colon P_i \to M$ are principal G-bundles over M. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both*[1] *P_1 and P_2 admit principal bundle atlases over the $U_{\mathtt{a}}$. Let*

$$\rho^1_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G, \qquad \text{and} \qquad \rho^2_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G$$

denote the transition functions of P_1 and P_2 with respect to these bundle atlases. Then P_1 and P_2 are isomorphic principal bundles if and only if there exists a smooth family $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to G$ of functions such that

$$\nu_{\mathtt{a}}(x) \cdot \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \cdot \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}.$$

The proof of Proposition 24.22 is on Problem Sheet M.

Corollary 24.23. *The principal bundle constructed in Theorem 24.21 is unique up to isomorphism.*

Definition 24.24. Suppose now we are in the setting of Theorem 14.1, where in addition we are given an effective action μ of G on another manifold F, we can construct *two* bundles: a fibre bundle $E \to M$ with fibre F, and a principal bundle $P \to M$ with fibre G, such that the transition functions of P are given by (24.4) and the transition functions of E are given by $\tilde{\rho}_{\mathtt{ab}}(x)(v) := \mu(\rho_{\mathtt{ab}}(x), v)$ (compare (13.3)). Since we normally suppress the action μ when talking about transition functions, informally this is saying that both bundles have the *same* transition functions. In this case we call P **the principal bundle associated** to E and we call E **an associated bundle of** P.

Here is the key example of this process.

Example 24.25. Let E be a vector bundle. Then the principal bundle associated to E is the frame bundle $\mathrm{Fr}(E)$. Indeed, the proof of Proposition 24.14 showed that the transition functions of $\mathrm{Fr}(E)$ are the same as those of E.

Note the difference between "the" and "an"! A fibre bundle E with structure group G has exactly one principal bundle (up to isomorphism) associated to it by Corollary 24.23. However as we will see next lecture, if we start with the principal bundle then *any* effective action of G on another manifold F gives rise to an associated bundle.

[1] This can always be achieved by taking intersections.

LECTURE 25

Associated bundles

We begin this lecture by explaining how to build associated fibre bundles from a principal bundle.

DEFINITION 25.1. Let $\pi\colon P \to M$ be a principal G-bundle, and assume G acts effectively on another manifold F on the left via $\mu\colon G \times F \to F$. Define an equivalence relation $\sim$ on $P \times F$ by setting:

$$(p \cdot a, v) \sim (p, \mu(a, v)), \qquad p \in P,\ a \in G,\ v \in F, \tag{25.1}$$

Define $P \times_G F$ to be the quotient space $(P \times F)/\sim$. Writing $[p, v]$ for the equivalence class of (p, v), we define a map $\wp\colon P \times_G F \to M$ by setting $\wp[p, v] := \pi(p)$.

REMARK 25.2. The notation "$P \times_G F$" is somewhat ambiguous, since we really should specify the action we are using. When confusion is possible, we will occasionally write $P \times_{G,\mu} F$ or $P \times_\mu F$ instead.

The next result is the main one of today's lecture. It covers everything we could ever want to know about associated bundles, and the proof will take some time.

THEOREM 25.3. *Let $\pi\colon P \to M$ be a principal G-bundle, and assume G acts effectively on another manifold F on the left. Then:*

(i) *$P \times_G F$ is a smooth manifold.*

(ii) *$\wp\colon P \times_G F \to M$ is a fibre bundle with fibre F and structure group G.*

(iii) *P is the principal bundle associated to $P \times_G F$.*

(iv) *The quotient map $\Pi\colon P \times F \to P \times_G F$ given by $\Pi(p, v) := [p, v]$ is also a principal G-bundle.*

(v) *The first projection $\mathrm{pr}_1\colon P \times F \to P$ is a principal bundle morphism along $\wp$:*

$$\begin{tikzcd} P \times F \arrow[r, "\mathrm{pr}_1"] \arrow[d, "\Pi"'] & P \arrow[d, "\pi"] \\ P \times_G F \arrow[r, "\wp"'] & M \end{tikzcd}$$

(vi) *For each $p \in P$, the map $L_p\colon F \to P \times_G F$ given by $v \mapsto [p, v]$ is a diffeomorphism over the fibre $\pi(p)$.*

(vii) *If F is a vector space and G acts linearly on F (so that μ can be thought of as a representation $G \to \mathrm{GL}(F)$, cf. Definition 10.19) then $P \times_G F$ is a vector bundle over M and the map L_p from part (vi) is a linear isomorphism.*

Proof. We will prove the result in three steps.

1. In this first step, we will prove (i), (ii) and (iii). This is actually most of the work. Suppose $\alpha\colon \pi^{-1}(U) \to G$ is a principal bundle chart over an open set $U \subset M$. We define a map $\tilde{\alpha}\colon \wp^{-1}(U) \to F$ by

$$\tilde{\alpha}[p, v] := \mu(\alpha(p), v), \qquad (p, v) \in \pi^{-1}(U) \times F.$$

Our aim is to show that the collection of maps $\tilde{\alpha}$, as α runs over a principal bundle atlas for P, defines a bundle atlas for $P \times_G F$. Thus as a first step we must show that $(\wp, \tilde{\alpha})\colon \wp^{-1}(U) \to U \times F$ is bijective.

For each $x \in U$, let $q_x \in P_x$ denote the unique element such that $\alpha(q_x) = e$ (this is well defined as (π, α) is a diffeomorphism). Now define $\psi\colon U \times F \to \wp^{-1}(U)$ by $\psi(x, v) := [q_x, v]$. We claim that ψ is an inverse to $(\wp, \tilde{\alpha})$. Indeed,

$$(\wp, \tilde{\alpha}) \circ \psi(x, v) = (\wp, \tilde{\alpha})[q_x, v] = (x, \mu(\alpha(q_x), v)) = (x, v).$$

Going the other way round, if $x \in U$ and $p \in P_x$ then

$$(\pi, \alpha)(q_x \cdot \alpha(p)) = (x, \alpha(q_x)\alpha(p)) = (x, \alpha(p)) = (\pi, \alpha)(p)$$

and thus $q_x \cdot \alpha(p) = p$. We therefore have for $p \in P_x$ that

$$\begin{aligned} \psi \circ (\wp, \tilde{\alpha})[p, v] &= \psi(x, \mu(\alpha(p), v)) \\ &= [q_x, \mu(\alpha(p), v)] \\ &\overset{(\dagger)}{=} [q_x \cdot \alpha(p), v] \\ &= [p, v], \end{aligned}$$

where $(\dagger)$ used the defining relationship (25.1) for $\sim$. Thus $(\wp, \tilde{\alpha})$ is bijective. Now let us investigate the transition function. If α and β are two principal bundle charts on P over open sets U and V respectively such that $U \cap V \neq \emptyset$, then for $(x, v) \in (U \cap V) \times F$, the transition function $\rho_{\tilde{\alpha}\tilde{\beta}}$ satisfies

$$\begin{aligned} \rho_{\tilde{\alpha}\tilde{\beta}}(x)(v) &= \tilde{\alpha} \circ (\wp, \tilde{\beta})^{-1}(x, v) \\ &= \tilde{\alpha}[(\pi, \beta)^{-1}(x, e), v] \\ &= \mu(\alpha \circ (\pi, \beta)^{-1}(x, e), v) \\ &= \mu(\rho_{\alpha\beta}(x)(e), v), \end{aligned}$$

which under the convention from Definition 13.9 that we suppress the μ from the transition functions, we see that

$$\rho_{\tilde{\alpha}\tilde{\beta}} = \rho_{\alpha\beta},$$

that is, the transition functions of $P \times_G F$ are the same as the transition functions from P, and thus in particular are smooth. Thus as in Remark 13.7, we can endow $P \times_G F$ with a smooth structure by declaring all the maps $(\wp, \tilde{\alpha})$ to be diffeomorphisms—this gives a well-defined smooth structure as we just checked the transition functions are smooth. Then the collection $\{\tilde{\alpha}\}$ form a bundle atlas, and $P \times_G F$ is a fibre bundle. That the structure group is G is immediate from the fact that the transition functions coincide with those of P.

Since $P \times_G F$ and P have the same transition functions, it is clear that P is the principal bundle associated to $P \times_G F$ in the sense of Definition 24.24.

2. We now prove (iv) and (v). With respect to the charts on $P \times_G F$, if $\Pi(p, v) = [p, v]$ then the expression

$$(\wp, \tilde{\alpha}) \circ \Pi(p, v) = (\pi(p), \mu(\alpha(p), v)),$$

shows that locally Π is smooth. The right action of G on $P \times F$ is given by

$$(p, v) \cdot a = (p \cdot a, \mu(a^{-1}, v)).$$

This right action preserves the fibres of Π, since

$$\Pi((p, v) \cdot a) = \Pi(p \cdot a, \mu(a^{-1}, v)) = \big[p \cdot a, \mu(a^{-1}, v)\big] = [p, v] = \Pi(p, v)$$

by the defining relationship (25.1). It now follows from Proposition 24.8 that $P \times F \to P \times_G F$ is another principal G bundle. This proves (iv). The identity

$$\mathrm{pr}_1(p \cdot a, \mu(a^{-1}, v)) = p \cdot a = \mathrm{pr}_1(p, v) \cdot a$$

shows that pr_1 is a principal G-bundle morphism along $\wp$, which proves (v).

3. We now prove (vi) and (vii). Fix $x \in M$ and $p \in P_x$. The map $F \to \wp^{-1}(x)$ given by $v \mapsto [p, v]$ is smooth because Π is. Its inverse with respect to principal bundle charts α and $\tilde{\alpha}$ as given in Step 1 is given by $[p, w] \mapsto \mu(\alpha(p)^{-1}, \tilde{\alpha}[p, w])$. This proves (vi).

Finally, suppose F is a vector space and μ is a linear action. Given $p \in P$, let $[p, v]$ and $[p, w]$ be two points in $\wp^{-1}(\pi(p))$. Given $c \in \mathbb{R}$, we define

$$[p, v] + c[p, w] := [p, v + vc].$$

This is well defined, i.e. independent of the choice of p, since if $a \in G$ then $[p, v] = [p \cdot a, \mu(a^{-1}, v)]$ and $[p, w] = [p \cdot a, \mu(a^{-1}, w)]$ and then since μ is linear

$$[p \cdot a, \mu(a^{-1}, v) + c\mu(a^{-1}, w)] = [p \cdot a, \mu(a^{-1}, v + cw)] = [p, v + cw].$$

The charts $\tilde{\alpha}$ on $P \times_G F$ are now vector bundle charts. This completes the proof. ∎

As an application of Theorem 25.3, we give (yet) another way to view the different vector bundles we can build from one another. (Compare the Metatheorem from Lecture 14).

Corollary 25.4. *Let $\pi \colon E \to M$ be a vector bundle with standard fibre*[1] *V. Let $\hat{\pi} \colon \mathrm{Fr}(E) \to M$ denote the frame bundle of E. Then:*

[1] Of course, we could write $V = \mathbb{R}^k$, but as you will see from the statement of the corollary, writing the fibre as an abstract vector space makes it clearer what is going on.

(i) *The group* $\mathrm{GL}(V)$ *acts on the dual space* V^* *by* $\mu(A, p) = p \circ A^{-1}$*. This gives an associated vector bundle* $\mathrm{Fr}(E) \times_{\mathrm{GL}(V)} V^*$*, which is exactly the dual bundle* E^**.*

(ii) *The group* $\mathrm{GL}(V)$ *acts on the tensored space* $V \otimes V$ *by* $\mu(A, v \otimes w) = Av \otimes Aw$*. This gives an associated vector bundle* $\mathrm{Fr}(E) \times_{\mathrm{GL}(V)} (V \otimes V)$*, which is exactly the tensor bundle* $E \otimes E$*.*

(iii) *The group* $\mathrm{GL}(V)$ *acts on the exterior algebra* $\bigwedge(V)$ *by* $\mu(A, v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k$*. This gives an associated vector bundle* $\mathrm{Fr}(E) \times_{\mathrm{GL}(V)} \bigwedge(V)$*, which is exactly the exterior bundle* $\bigwedge(E)$*.*

The proof of Corollary 25.4 is immediate from the proof of Theorem 25.3 since in all cases we know what the transition functions are. *Fun Exercise*: Formulate an appropriate categorical statement and use this to recover Theorem 14.41 in full generality. Here is another application, this time to homogeneous spaces.

PROPOSITION 25.5. *Let* G *be a Lie group and let* μ *be a transitive and effective*[2] *left action on a connected smooth manifold* M*. Fix* $x \in M$ *and let* H *denote the isotropy group at* x*, so that* M *is the homogeneous space* G/H *(cf. Theorem 12.11). Assume also that* H *acts effectively on* T_xM *via the map*

$$\varphi \colon H \times T_xM \to T_xM, \qquad \varphi(c, v) := D\mu_c(x)[v]$$

(cf. Definition 12.10). Then with this action, the tangent bundle TM *is isomorphic as a vector bundle to the associated bundle* $G \times_H T_xM$*.*

Proof. Define a map $\Phi \colon G \times_H T_xM \to TM$ by $\Phi[a, v] := \big(\mu_a(x), D\mu_a(x)[v]\big)$. This map is visibly smooth and linear on each fibre, so we need only build an inverse. For this suppose $w \in T_yM$. Since the action of G is transitive, there exists $a \in G$ such that $\mu_a(x) = \mu(a, x) = y$. Define $\Psi \colon TM \to G \times_H T_xM$ by $\Psi(y, w) := [a, D\mu_a(x)^{-1}[w]]$. This is well defined, since if b is another element of G such that $\mu(b, x) = y$ then $c := a^{-1}b \in H$, and hence

$$\begin{aligned}
[a, D\mu_a(x)^{-1}[w]] &= [ac, \varphi(c^{-1}, D\mu_a(x)^{-1}[w])] \\
&= [ac, D\mu_{c^{-1}}(x) \circ D\mu_a(x)^{-1}[w]] \\
&= [ac, D\mu_c(x)^{-1} \circ D\mu_a(x)^{-1}[w]] \\
&= [ac, D\mu_{ac}(x)^{-1}[w]] \\
&= [b, D\mu_b(x)^{-1}[w]],
\end{aligned}$$

where the first equality used the definition of the equivalence relation (25.1) in the associated bundle $G \times_H T_xM$. ■

In fact, any homogeneous space can be realised as one arising from an effective action, as the following lemma shows. The proof is on Problem Sheet M.

[2] This really is an extra assumption, a homogeneous space only requires the action to be transitive. However see Lemma 25.6 below.

Lemma 25.6. *Suppose G is a Lie group acting transitively on a smooth manifold M, so that M is the homogeneous space G/H for an appropriate subgroup H of G. The subgroup of G acting trivially on M is the largest normal subgroup $N(H)$ of G contained in H. If $\bar{G}$ and $\bar{H}$ denote the quotient groups $G/N(H)$ and $H/N(H)$ respectively then $\bar{G}$ acts effectively and transitively on M, and M is the homogeneous space $\bar{G}/\bar{H}$.*

Here is another difference between vector bundles and principal bundles.

Proposition 25.7. *Let $\pi\colon P \to M$ be a principal G-bundle. Then P has a section if and only if P is trivial.*

Proof. If $P = M \times G$ is the trivial bundle, then for any $a \in G$ the map $s(x) := (x, a)$ is a section. Conversely, if $s\colon M \to P$ is a section then since p and $s(\pi(p))$ belong to the same fibre for each $p \in P$, there is a well-defined equivariant map $\alpha\colon P \to G$ such that

$$p = s(\pi(p)) \cdot \alpha(p), \qquad \forall\, p \in P.$$

We claim that α is a principal bundle chart, whence P is a trivial bundle. For this we need to prove that $(\pi, \alpha)\colon P \to M \times G$ is a diffeomorphism. But this follows from Lemma 24.16, since (π, α) is a principal bundle morphism along the identity map on M. ■

So far we have not made any use of the Lie group structure other than it being a group. But as we extensively studied earlier in the course, a Lie group G comes with a Lie algebra $\mathfrak{g}$. An element $v \in \mathfrak{g}$ determines a left-invariant vector field X_v on G (cf. Theorem 9.19) via $X_v(a) := Dl_a(e)[v]$. A similar thing works whenever a Lie group acts on a manifold, as we now explain.

Remark 25.8. For the rest of this lecture, all the results we prove are valid for an arbitrary right action of a Lie group on a manifold (i.e. we do *not* require a principal bundle action). Thus we will formulate the statements to come in this more general setting. Nevertheless, our only application of this material (which won't come until Lecture 39 next semester) will immediately restrict to a principal bundle.

Definition 25.9. Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose G acts on a manifold P on the right, which we write as $(p, a) \mapsto p \cdot a$. Given $v \in \mathfrak{g}$, we associate a vector field ξ_v on P via

$$\xi_v(p) := \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tv) \in T_pP.$$

Let us unpack this a bit. Fix $p \in P$. Then the curve $\gamma_p(t) := p \cdot \exp(tv)$ is a curve in P with initial point $\gamma_p(0) = p \cdot e = p$. Thus $\gamma_p'(0)$ belongs to T_pP, and this is the value of the vector field ξ_v:

$$\xi_v(p) = \gamma_p'(0).$$

If $f \in C^\infty(P)$ then (thought of a derivation), one has

$$\xi_v(f)(p) = \frac{d}{dt}\Big|_{t=0} f \circ \gamma_p(t) = \frac{d}{dt}\Big|_{t=0} f(p \cdot \exp(tv)).$$

Of course, calling something a "vector field" does not make it one. Certainly ξ_v is a section of TP, but it isn't immediate why it is smooth.

Lemma 25.10. *The fundamental vector field ξ_v is smooth (and hence a vector field on P).*

Proof. It suffices to show by Proposition 7.2 that $\xi_v(f)$ is a smooth function for each $f \in C^\infty(P)$. But this is clear from the formula above. To make it more transparent, let us write μ for the action. Then

$$\xi_v(f)(p) = \frac{d}{dt}\Big|_{t=0}(f \circ \mu)(p, \exp(tv))$$

is the composition of smooth functions in both p and t. ■

Proposition 25.11. *The flow of ξ_v is given by $\theta_t(p) := p \cdot \exp(tv)$. Thus ξ_v is always complete.*

Proof. With γ_p as above, we need only show that γ_p is the integral curve of ξ_v through p. This follows from:

$$\gamma_p'(t) = \frac{d}{ds}\Big|_{s=0}\gamma_p(t+s) = \frac{d}{ds}\Big|_{s=0} p \cdot \exp(tv)\exp(sv) = \xi_v(p \cdot \exp(tv)) = \xi_v(\gamma_p(t)).$$

■

An alternative way to define the fundamental vector field ξ_v is to consider the map

$$\eta_p \colon G \to P, \qquad \eta_p(a) := p \cdot a.$$

Then with γ_p as above,

$$D\eta_p(e)[v] = \frac{d}{dt}\Big|_{t=0}\eta_p(\exp(tv)) = \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tv) = \xi_v(p). \tag{25.2}$$

Example 25.12. Let G act on itself via right multiplication. Then by Proposition 10.9 the fundamental vector field associated to $v \in \mathfrak{g}$ is exactly the left-invariant vector field X_v.

On Problem Sheet M you will show:

Proposition 25.13. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose G acts on a manifold P on the right. Then the map $v \mapsto \xi_v$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(P)$.*

Our last result makes contact with the adjoint representation from Lecture 10. For this, let us write

$$r_a \colon P \to P, \qquad r_a(p) = p \cdot a$$

so that $r_a \in \mathrm{Diff}(P)$.

Proposition 25.14. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose G acts on a manifold P on the right. Then for $v \in \mathfrak{g}$ the following "infinitesimal" equivariance property holds:*

$$Dr_a(p)[\xi_v(p)] = \xi_{\mathrm{Ad}_{a^{-1}}(v)}(r_a(p)).$$

Proof. For any $b \in G$, one has

$$r_a \circ \eta_p(b) = p \cdot ba = \eta_{p \cdot a}(a^{-1}ba).$$

Differentiating this identity at $b = e$ and using the fact that Ad is the differential of the conjugation action $b \mapsto aba^{-1}$ at $b = e$, the claim follows from the chain rule and (25.2). ■

We will use this material again in Lecture 39 when discussing the *connection form* of a connection on a principal bundle (see Theorem 39.3 in particular).

Bundle-valued forms

In this lecture we will push our treatment of differential forms a little further and allow them to take values in an arbitrary vector space, or later, a vector bundle.

Unfortunately we will have to wait until Differential Geometry II next semester in order to understand the motivation behind this construction—for now let us just say that this material will be crucial to make sense of connections and curvature on vector and principal bundles, as well as to understand the relation between the two.

Let us start at the level of linear algebra. If V is a vector space, we have studied extensively the exterior wedge $\bigwedge^r(V^*)$, and its identification with the space $\mathrm{Alt}_r(V)$ of alternating multilinear maps

$$A \colon \overbrace{V \times \cdots \times V}^{r} \to \mathbb{R}.$$

Now suppose W is another vector space. In Definition 15.22 we actually originally introduced the space $\mathrm{Alt}_r(V, W)$ of alternating multilinear maps

$$A \colon \overbrace{V \times \cdots \times V}^{r} \to W.$$

Moreover Lemma 15.24 and Corollary 15.4 show that

$$\mathrm{Alt}_r(V, W) \cong \mathrm{L}\left(\bigwedge^r(V), W\right) \cong \left(\bigwedge^r(V)\right)^* \otimes W \cong \bigwedge^r(V^*) \otimes W.$$

This gives:

Lemma 26.1. *Let V and W be two vector spaces. For $r \geq 0$ there is a canonical isomorphism between $\mathrm{Alt}_r(V, W)$ and $\bigwedge^r(V^*) \otimes W$.*

We now generalise this idea. If E is a vector bundle over M and W is a vector space, we denote by $E \otimes W$ the bundle over M whose fibre is $(E \otimes W)_x := E_x \otimes W$ (equivalently, this is the bundle obtained by tensoring E with the trivial bundle $M \times W \to M$).

Definition 26.2. Let M be a smooth manifold and let W be a vector space. A **differential r-form on M with values in W** (also called a **vector-valued form**) is a section of the bundle $\bigwedge^r(T^*M) \otimes W \to M$. We denote the space of sections by

$$\Omega^r(M, W) := \Gamma\left(\bigwedge^r(T^*M) \otimes W\right).$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

This is not as scary as it looks (and reduces to the normal definition if $W = \mathbb{R}$). For instance, a W-valued one-form ω associates to every $x \in M$ a linear map $\omega_x \colon T_xM \to W$. Thus if we feed ω_x a tangent vector v we get an element of W, rather than an element of $\mathbb{R}$. If X is a vector field on M and $f \colon M \to W$ is a smooth function then

$$X(f) \colon M \to W, \qquad X(f)(x) := \mathcal{J}_{f(x)}^{-1}\left(Df(x)[X(x)]\right) \tag{26.1}$$

is another smooth function, where $\mathcal{J}_{f(x)} \colon W \to T_{f(x)}W$ is the map from Problem B.3. Thus the analogue of Proposition 7.2 holds for vector-valued functions as well. In fact, almost all of our earlier work on differential forms goes through without any changes (just insert W in appropriate places), and I will leave you to fill in the details, stating only a few pertinent results.

THEOREM 26.3 (The Vector-valued Differential Form Criterion). *Let M be a smooth manifold and let W be a vector space. Then there is a canonical identification between $\Omega^r(M, W)$ and alternating $C^\infty(M)$-multilinear functions*

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{r} \to C^\infty(M, W).$$

One can also define for an open subset $U \subset M$ the sections over U:

$$\Omega^r(U, W) := \Gamma\left(U, \bigwedge^r(T^*M) \otimes W\right),$$

and this gives us a sheaf in the same way as before.

REMARK 26.4. **Warning:** Unfortunately we tend to use the same letters U, V, W etc to denote both open sets in manifolds and arbitrary vector spaces. This means expressions like $\Omega^r(U, W)$ can be somewhat confusing... Oh well.

Assume now that W has dimension k, and let $(e_1, \dots, e_k)$ be a basis. If $\omega \in \Omega^r(M, W)$ and $x \in M$, then for any tangent vectors $v_1, \dots, v_r \in T_xM$, we can write $\omega_x(v_1, \dots, v_r)$ as a linear combination of the e_i. If we denote the coefficient of e_i by $\omega_x^i(v_1, \dots, v_r)$, we can thus write

$$\omega_x(v_1, \dots, v_r) = \omega_x^i(v_1, \dots, v_r)\, e_i.$$

Since ω_x is an alternating multilinear map, so is each ω_x^i. It follows that ω^i is a normal differential r-form on M, and we can write

$$\omega = \omega^i \otimes e_i.$$

This is of course, consistent with thinking of ω as a section of the tensored bundle $\bigwedge^r(T^*M) \otimes W$. This allows us to extend the exterior differential to a sheaf morphism $d \colon \Omega^r(M, W) \to \Omega^{r+1}(M, W)$ by declaring that

$$d\left(\omega^i \otimes e_i\right) := d\omega^i \otimes e_i.$$

Exercise: Why is this independent of the choice of basis of W?

The wedge product requires a little more thought to define, since this requires us to multiply vectors together. This isn't possible in an arbitrary vector space (only in algebras, cf. Definition 15.17). Thus in general we need to specify a bilinear map. This works as follows: suppose V, W_1, W_2 and Z are four vector spaces, and assume we are given a bilinear map $\beta\colon W_1 \times W_2 \to Z$ (equivalently, a linear map $W_1 \otimes W_2 \to Z$, cf. Lemma 15.2). Then motivated by Lemma 19.4, we make the following definition:

Definition 26.5. Let $\omega \in \mathrm{Alt}_r(V, W_1)$ and $\vartheta \in \mathrm{Alt}_s(V, W_2)$. We define $\omega \wedge_\beta \vartheta \in \mathrm{Alt}_{r+s}(V, Z)$ by

$$(\omega\wedge_\beta\vartheta)(v_1, \dots, v_{r+s}) = \frac{1}{r!s!} \sum_{\varrho \in \mathfrak{S}_{r+s}} \mathrm{sgn}(\varrho)\beta\left(\omega\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big), \vartheta\big(v_{\varrho(r+1)}, \dots, v_{\varrho(r+s)}\big)\right).$$

Equivalently we can think of $\wedge_\beta$ as defining a map

$$\Big(\bigwedge^r (V^*) \otimes W_1\Big) \times \Big(\bigwedge^s (V^*) \otimes W_2\Big) \to \Big(\bigwedge^{r+s} (V^*) \otimes Z\Big)$$

If for instance $W = W_1 = W_2$ is an algebra (and thus there is natural algebra multiplication $W \otimes W \to W$) we can regard the wedge product as a map

$$\Big(\bigwedge^r (V^*) \otimes W\Big) \times \Big(\bigwedge^s (V^*) \otimes W\Big) \to \Big(\bigwedge^{r+s} (V^*) \otimes W\Big),$$

and in this case we typically omit reference of the map β. Moreover if we have no convenient map β, we can always take $Z = W_1 \otimes W_2$ and have β be induced from the identity map $W_1 \times W_2 \to W_1 \times W_2$. Thus we always have a wedge product

$$\Big(\bigwedge^r (V^*) \otimes W_1\Big) \times \Big(\bigwedge^s (V^*) \otimes W_2\Big) \to \Big(\bigwedge^{r+s} (V^*) \otimes W_1 \otimes W_2\Big),$$

We can apply this to manifolds: if $\beta\colon W_1 \times W_2 \to Z$ is a bilinear map then we obtain a map

$$\Omega^r(M, W_1) \times \Omega^s(M, W_2) \xrightarrow{\wedge_\beta} \Omega^{r+s}(M, Z)$$

by applying the above construction pointwise:

$$(\omega \wedge_\beta \vartheta)_x = \omega_x \wedge_\beta \vartheta_x, \qquad x \in M.$$

Let $(e_1, \dots, e_k)$ be a basis of W_1 and $(e'_1, \dots, e'_l)$ be a basis of W_2. If we write $\omega = \omega^i \otimes e_i$ and $\vartheta = \vartheta^j \otimes e'_j$ then from the definition it follows that

$$\omega \wedge_\beta \vartheta = \omega^i \wedge \vartheta^j \beta(e_i, e'_j).$$

If $(f_1, \dots, f_m)$ is a basis of Z then we can write $\beta(e_i, e'_j) = a^h_{ij} f_h$ for real numbers a^h_{ij}, and thus

$$\omega \wedge_\beta \vartheta = a^h_{ij}\, \omega^i \wedge \vartheta^j\, f_h,$$

which also proves that $\omega \wedge_\beta \vartheta$ is smooth (if you were worried). The following result, whose proof is on Problem Sheet M, shows that the exterior differential on vector-valued forms is still skew-commutative.

Proposition 26.6. *Let M be a smooth manifold, and let W_1, W_2 and Z be vector spaces. Let $\omega \in \Omega^r(M, W_1)$ and let $\vartheta \in \Omega^s(M, W_2)$, and let $\beta \colon W_1 \times W_2 \to Z$ be a bilinear map. Then*

$$d(\omega \wedge_\beta \vartheta) = d\omega \wedge_\beta \vartheta + (-1)^r \omega \wedge_\beta d\vartheta.$$

Let us give an example of how this is useful.

Example 26.7. Let $\mathfrak{g}$ be a Lie algebra. Then $\mathfrak{g}$ is in particular a vector space, and the Lie bracket $(v, w) \mapsto [v, w]$ is a bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. Suppose M is a manifold. Given $\omega \in \Omega^r(M, \mathfrak{g})$ and $\vartheta \in \Omega^s(M, \mathfrak{g})$, we typically use the notation

$$[\omega, \vartheta] := \omega \wedge_{\beta = [\cdot,\cdot]} \vartheta.$$

We claim that this wedge product satisfies the following version of skew-commutativity:

$$[\omega, \vartheta] = (-1)^{rs+1} [\vartheta, \omega]. \tag{26.2}$$

To see this, let $(e_1, \dots, e_k)$ be a basis for $\mathfrak{g}$. Write $\omega = \omega^i \otimes e_i$ and $\vartheta = \vartheta^j \otimes e_j$. Then

$$[\omega, \vartheta] = \omega^i \wedge \vartheta^j [e_i, e_j] = (-1)^{rs+1} \vartheta^j \wedge \omega_i [e_j, e_i] = (-1)^{rs+1} [\vartheta, \omega],$$

where the $(-1)^{rs}$ came from swapping $\omega^i \wedge \vartheta^j$ to $\vartheta^j \wedge \omega^i$ and the other -1 came from $[e_i, e_j] = -[e_j, e_i]$. In particular, this shows that if $r = s = 1$ then $[\omega, \vartheta]$ is *symmetric* in ω and ϑ. (This should surprise you, since the normal Lie bracket is anti-symmetric). In particular, it is **not** (!) necessarily true that $[\omega, \omega] = 0$ for $\omega \in \Omega^1(M, \mathfrak{g})$. This will be important in Lecture 39 (see Theorem 39.10 in particular).

Let us also note the analogue of Theorem 20.7 holds.

Theorem 26.8. *Let M be a smooth manifold, $\omega \in \Omega^r(M, V)$ and $X_0, \dots X_r \in \mathfrak{X}(M)$. Then:*

$$\begin{aligned} d\omega(X_0, \dots, X_r) = &\sum_{i=0}^{r} (-1)^i X_i \big(\omega(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \sum_{0 \le i < j \le r} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_r). \end{aligned}$$

Proof. The proof is identical to the proof of Theorem 20.7—the only difference is that both sides are functions $M \to V$ rather than functions $M \to \mathbb{R}$. ∎

Similarly the proof of Lemma 19.19 goes through without any changes to give:

Lemma 26.9. *Let $\varphi \colon M \to N$ be a smooth map and let $\omega \in \Omega(N, V)$. Then*

$$\varphi^\star(d\omega) = d(\varphi^\star(\omega)).$$

Let us now take this one step further and look at differential forms with values in a vector bundle, rather than just a vector space.

Definition 26.10. Let M be a smooth manifold and $\pi\colon E \to M$ a vector bundle over M. A **differential r form with values in E** (or a **bundle-valued form**) is a section of $\bigwedge^r(T^*M) \otimes E$. As usual, we denote by $\Omega^r(M,E)$ the space of such sections.

Thus an element $\omega \in \Omega^r(M,E)$ defines for each $x \in M$ an alternating multilinear map

$$\omega_x \colon \overbrace{T_xM \times \cdots \times T_xM}^{r} \to E_x.$$

Again, this may seem confusing, but in reality is no more complicated than the case of a vector-valued form; the only difference is that the target vector space E_x now also depends on x. If $(e_1,\ldots,e_k)$ is a local frame for E over an open set U then any element $\omega \in \Omega^r(U,E)$ can be written as a sum

$$\omega = \omega^i \otimes e_i$$

where ω^i is a normal differential r-form on U.

Remark 26.11. **Warning:** Do not confuse $\Omega^r(U,E)$ and $\Gamma(U,\bigwedge^r(E))$!

Theorem 26.12 (The Bundle-valued Differential Form Criterion). *There is a natural $C^\infty(M)$-module isomorphism between $\Omega^r(M,E)$ and alternating $C^\infty(M)$-multilinear functions*

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{r} \to \Gamma(E).$$

This follows in the same way as Theorem 19.1 and Theorem 26.3, but since this is arguably the hardest of these sort of results, let us recap the details.

Proof. We use the Hom-Gamma Theorem 16.30:

$$\begin{aligned}\Omega^r(M,E) &= \Gamma\left(\textstyle\bigwedge^r(T^*M)\otimes E\right)\\ &= \Gamma\left(\mathrm{Hom}\left(\textstyle\bigwedge^r(TM),E\right)\right)\\ &= \mathrm{Hom}\left(\Gamma\left(\textstyle\bigwedge^r(TM)\right),\Gamma(E)\right).\end{aligned}$$

Now the argument used in the proof of Theorem 18.3 (which was proved as Problem I.3) shows that this latter space can be identified with alternating $C^\infty(M)$-multilinear functions

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{r} \to \Gamma(E).$$

Alternatively, one could use the alternating version of Theorem 18.5 to show that

$$\mathrm{Hom}\left(\Gamma\left(\textstyle\bigwedge^r(TM)\right),\Gamma(E)\right) \cong \mathrm{Alt}_r(\mathfrak{X}(M),\Gamma(E))$$

(this is more efficient, but harder, since $\mathfrak{X}(M)$ and $\Gamma(E)$ are infinite-dimensional vector spaces). ■

Thus we can think of an element of $\Omega^r(M,E)$ as a alternating map that eats vector fields and produces a section of E:

$$\omega(X_1,\dots,X_r)\in\Gamma(E).$$

Here is an example.

Example 26.13. Let $\varphi\colon M\to N$ be a smooth map. Then $D\varphi$ can be thought of as an element of $\Omega^1(M,\varphi^\star(TN))$.

We now return to associated bundles of principal bundles. We will only formulate this in the special case that will be of relevance to our considerations next semester. Let $\pi\colon P\to M$ denote a principal G-bundle, and let $\mu\colon G\to \mathrm{GL}(V)$ denote a effective linear representation on a vector space V, and let $E:=P\times_G V$ denote the associated vector bundle[1]. We are interested in the relation between the spaces

$$\Omega^r(P,V)\qquad\text{and}\qquad\Omega^r(M,E).$$

We will construct a subspace $\Omega^r_G(P,V)\subset\Omega^r(P,V)$ which will consist of *horizontal equivariant* forms. Then we will prove that $\Omega^r_G(P,V)\cong\Omega^r(M,E)$.

We begin with the following very general definition.

Definition 26.14. Recall from Problem I.5 that the **vertical bundle** $VP\subset TP$ is the vector subbundle of the tangent bundle of P whose fibre over $p\in P$ is $\ker D\pi(p)\colon T_pP\to T_{\pi(p)}M$. A tangent vector $\zeta\in T_pP$ is said to be **vertical** if $\zeta\in V_pP$, i.e. if $D\pi(p)[\zeta]=0$. A differential form $\omega\in\Omega^r(P,V)$ is said to be **horizontal** if ω vanishes whenever any of its variables is a vertical vector. If $r=0$, we declare all forms to be horizontal.

Remark 26.15. The vertical bundle is defined for an arbitrary fibre bundle, and thus we can define horizontal differential forms on arbitrary bundles (either real-valued, vector-valued, or bundle-valued) in the same way.

We write $r_a\colon P\to P$ for the map $p\mapsto p\cdot a$ and we write $\mu_a\colon V\to V$ for the map $v\mapsto\mu(a,v)$.

Definition 26.16. Let $\omega\in\Omega^r(P,V)$ denote an V-valued form. We say that ω is **equivariant** if

$$r_a^\star(\omega)=\mu_{a^{-1}}(\omega),\qquad\forall\, a\in G.$$

Explicitly, this means that for any $p\in P$ and $\zeta_1,\dots,\zeta_r\in T_pP$,

$$r_a^\star(\omega)_p(\zeta_1,\dots,\zeta_r)=\mu_{a^{-1}}(\omega_p(\zeta_1,\dots,\zeta_r)).$$

We set:

$$\Omega^r_G(P,V):=\{\omega\in\Omega^r(P,V)\mid\omega\text{ is horizontal and equivariant}\}.$$

Let us now explain how an element $\omega\in\Omega^r_G(P,V)$ gives rise to an element $\omega^\flat\in\Omega^r(M,E)$. Fix $x\in M$ and $v_1,\dots,v_r\in T_xM$. Choose any point $p\in P_x$, and

[1] Actually the "effective" hypothesis is not needed here, for all the relevant bits of Theorem 25.3 go through without this assumption.

choose any vectors $\zeta_i \in T_pP$ such that $D\pi(p)[\zeta_i] = v_i$ (such vectors exist as π is a submersion). Then $\omega_p(\zeta_1, \dots, \zeta_r)$ belongs to V. By part (vii) of Theorem 25.3, the map $v \mapsto [p, v]$ defines an linear isomorphism from V to E_x, which we write as

$$L_p \colon V \to E_{\pi(p)}, \qquad L_p(v) := [p, v].$$

We apply L_p to $\omega_p(\zeta_1, \dots, \zeta_r)$ to get an element of E_x, and then define $\omega^\flat_x(v_1, \dots, v_r)$ to be this element:

$$\omega^\flat_x(v_1, \dots, v_r) := L_p\big(\omega_p(\zeta_1, \dots, \zeta_r)\big).$$

Of course, it requires proof that $\omega^\flat$ is well defined. The next result is the main one of this lecture.

THEOREM 26.17. *If $\omega \in \Omega^r_G(P, V)$ then $\omega^\flat$ is well defined. Moreover the map*

$$\Omega^r_G(P, V) \mapsto \Omega^r(M, E), \qquad \omega \mapsto \omega^\flat$$

is a linear isomorphism.

Proof. We prove the result in four steps.

1. To show that $\omega^\flat$ is well defined we must check the value $L_p(\omega_p(\zeta_1, \dots, \zeta_r))$ does not depend on the choice of $p \in P_x$ and the choice of $\zeta_i \in T_pP$ such that $D\pi(p)[\zeta_i] = v_i$. In this step we deal with the ζ_i. If ξ_i was another tangent vector such that $D\pi(p)[\xi_i] = v_i$ then $\xi_i - \zeta_i$ is a vertical vector. Since ω is horizontal and r-linear we have

$$\begin{aligned} \omega_p(\zeta_1, \dots, \zeta_r) &= \omega_p(\zeta_1 - \xi_1 + \xi_1, \dots, \zeta_r - \xi_r + \xi_r) \\ &= \omega_p(\xi_1 + \text{vertical}, \dots, \xi_r + \text{vertical}) \\ &= \omega_p(\xi_1, \dots, \xi_r). \end{aligned}$$

2. Now let us deal with the choice of p. Suppose instead we choose $p \cdot a = r_a(p)$. Since $\pi \circ r_a = \pi$, we have

$$D\pi(r_a(p)) \circ Dr_a(p)[\zeta_i] = D\pi(p)[\zeta_i] = v_i, \tag{26.3}$$

so that $Dr_a(p)[\zeta_i] \in T_{r_a(p)}P$ is a tangent vector that maps onto v_i. Thus it suffices to show that

$$L_p\Big(\omega_p\big(\zeta_1, \dots, \zeta_r\big)\Big) = L_{p\cdot a}\Big(\omega_{p\cdot a}\big(Dr_a(p)[\zeta_1], \dots, Dr_a(p)[\zeta_r]\big)\Big). \tag{26.4}$$

But

$$\omega_{p\cdot a}(Dr_a(p)[\zeta_1], \dots, Dr_a(p)[\zeta_r]) = r^\star_a(\omega)_p(\zeta_1, \dots, \zeta_r) = \mu_{a^{-1}}(\omega_p(\zeta_1, \dots, \zeta_r),$$

by equivariance. To complete the proof of (26.4) we need only observe that:

$$L_{p\cdot a} = L_p \circ \mu_a \tag{26.5}$$

This is immediate from the definition: $L_{p\cdot a}(v) = [p \cdot a, v] = [p, \mu_a(v)] = L_p(\mu_a(v))$.

3. We now know that $\omega^\flat$ is well defined. To complete the proof, we build an inverse. Suppose $\vartheta \in \Omega^r(M,E)$. Define $\vartheta^\sharp \in \Omega^r(P,V)$ by

$$\vartheta^\sharp_p(\zeta_1,\dots,\zeta_r) := L_p^{-1}\Big(\vartheta_{\pi(p)}\big(D\pi(p)[\zeta_1],\dots,D\pi(p)[\zeta_r]\big)\Big).$$

It is obvious that $\vartheta^\sharp$ is horizontal, so let us check that $\vartheta^\sharp$ is equivariant. To see this we argue as follows:

$$\begin{aligned}(r_a^\star(\vartheta^\sharp))_p(\zeta_1,\dots,\zeta_r) &= \vartheta^\sharp_{p\cdot a}\big(Dr_a(p)[\zeta_1],\dots,Dr_a(p)[\zeta_r]\big)\\ &= L_{p\cdot a}^{-1}\Big(\vartheta_x\big(D\pi(r_a(p))\circ Dr_a(p)[\zeta_1],\dots,D\pi(r_a(p))\circ Dr_a(p)[\zeta_r]\big)\Big)\\ &\overset{(\dagger)}{=} \mu_{a^{-1}}\circ L_p^{-1}\Big(\vartheta_x\big(D\pi(p)[\zeta_1],\dots,D\pi(p)[\zeta_r]\big)\Big)\\ &= \mu_{a^{-1}}(\vartheta^\sharp_p(\zeta_1,\dots,\zeta_r)).\end{aligned}$$

where $(\dagger)$ used (26.3) and (26.5).

4. It is clear that $(\vartheta^\sharp)^\flat = \vartheta$ and $(\omega^\flat)^\sharp = \omega$. This completes the proof. ■

We conclude this lecture by studying the special case $r = 0$. Every zero-form is horizontal, and the equivariance condition for a function $f\colon P \to V$ becomes

$$f(p\cdot a) = \mu(a^{-1}, f(p)), \qquad \forall\, p \in P,\ a \in G. \tag{26.6}$$

Meanwhile $\Omega^0(M,E) = \Gamma(E)$. This proves:

Corollary 26.18. *Let P be a principal G bundle over M and let $E = P\times_G V$ denote an associated bundle. Then there is a one-to-one correspondence between $\Gamma(E)$ and functions $f\colon P\to V$ satisfying (26.6). Explicitly, given f satisfying (26.6) we define $s\colon M\to E$ via $s(x) := L_p(f(p))$, where p is any point in P_x. Conversely, given a section s, we define $f\colon P\to V$ by $f(p) = L_p^{-1}(s(x))$.*

Remark 26.19. Since any vector bundle can be seen as an associated bundle of a principal bundle (the frame bundle), this shows that any section of a vector bundle can be identified with an equivariant function on its frame bundle.

As explained at the start of the lecture, we will use Theorem 26.17 in Lecture 38 when discussing connections on principal bundles.

A proof of the de Rham Theorem

This entire lecture is completely **non-examinable**. Unlike the rest of the course, this lecture assumes you are familiar with singular (co)homology and some basic homological algebra.

The goal of this lecture is to prove that de Rham cohomology agrees with singular cohomology—this is usually referred to as the **de Rham Theorem**. There are many different ways to prove this result. Perhaps the neatest is via *sheaf cohomology*, but this is a little bit too far afield.

REMARK 27.1. All our homology and cohomology groups should be understood to have coefficients in $\mathbb{R}$ in this lecture. We will not comment on this further.

Let X be a topological space. You are hopefully familiar with the singular chain complex of X. This is normally defined by taking looking at singular *simplices* (i.e. continuous maps $\Delta^k \to X$, where Δ^k is the kth standard simplex). However one can equally well carry out the construction using singular *cubes* instead. The resulting algebraic invariant is the same (more on this later). Let us recall the definitions in the continuous category.

DEFINITION 27.2. Let X be a topological space. A **singular k-cube** in X is a *continuous* map $c\colon C^k \to X$. We let $Q_k(X)$ denote the (infinite-dimensional) vector space with basis all the singular k-cubes in X.

REMARK 27.3. Note in the continuous category there is no need to require c to extend to a map on an open neighbourhood.

DEFINITION 27.4. A singular k-cube $c\colon C^k \to X$ is said to be **degenerate** if there exists $1 \le i \le k$ such that c does not depend on x^i. Otherwise c is said to be **non-degenerate**. We let $D_k(X)$ denote the subspace of $Q_k(X)$ generated by the degenerate cubes, and we let

$$\overline{Q}_k(X) := Q_k(X) \Big/ D_k(X)$$

denote the quotient space.

Thus for instance

$$c\colon C^3 \to \mathbb{R}, \qquad c(x^1, x^2, x^3) := x^1 + x^3$$

is a degenerate singular 3-cube in $\mathbb{R}$.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Definition 27.5. Fix $1 \le i \le k$ and let $c\colon C^k \to X$ denote a singular k-cube. The ith **front face** of c, written $F_i c$, is the singular $(k-1)$-cube defined by

$$F_i c(x^1, \dots, x^{k-1}) := c(x^1, \dots, x^{i-1}, 0, x^i, \dots, x^{k-1}).$$

Similarly the ith **back face** is the singular $(k-1)$-cube defined by

$$B_i c(x^1, \dots, x^{k-1}) := c(x^1, \dots, x^{i-1}, 1, x^i, \dots, x^{k-1}).$$

Definition 27.6. Let $c\colon C^k \to X$ be a singular k-cube for $k > 0$. We define the **boundary** of c, written ∂c, to be the element of $Q_{k-1}(X)$ given by

$$\partial c := \sum_{i=1}^{k} (-1)^i (F_i c - B_i c).$$

We define the boundary of a 0-cube to be the real number 1. We then extend ∂ to arbitrary k-chains by linearity. Thus we may think of ∂ as a linear map $Q_k(X) \to Q_{k-1}(X)$ for all $k \ge 1$ (this works for $k = 0$ too if we define $Q_{-1}(X) := \mathbb{R}$).

Note that if a cube c is non-degenerate then so is ∂c. Thus we can also regard ∂ as a linear map

$$\partial\colon \overline{Q}_k(X) \to \overline{Q}_{k-1}(X), \qquad k \ge 0.$$

Proposition 27.7. *The boundary operator squares to zero: $\partial^2 = 0$. Thus $(\overline{Q}_\bullet(X), \partial)$ is a chain complex of vector spaces.*

Definition 27.8. The **cubical singular homology groups** $H_k^{\mathrm{cube}}(X; \mathbb{R})$ are defined to be the homology of this chain complex.

Remark 27.9. Why bother with quotienting out by the degenerate cubes? After all, $(Q_\bullet(X), \partial)$ is also a chain complex, so we could just take its homology instead. To see this why this quotienting out the degenerate cubes is superior, consider the case where X is a one point space $\{*\}$. It is easy to see that $H_k^{\mathrm{cube}}(\{\star\}) = 0$ for $k > 0$ and $H_0^{\mathrm{cube}}(\{*\}) = \mathbb{R}$, as one would hope (this is a necessary requirement in order for $H_\bullet^{\mathrm{cube}}$ to be "a homology theory" in the sense of Eilenberg-Steenrod). However if one does not quotient out by degenerate cubes, this ceases to be the case. *Exercise:* Why?

All the properties of the singular chain complex (built with singular simplices) continue to hold without change (homotopy invariance, long exact sequence, excision. Mayer-Vietoris, etc).

Remark 27.10. In fact, sometimes the proof gets *easier* for singular cubes. For instance, one of the key steps in establishing excision for singular simplices is the concept of **barycentric subdivision**, which allows one to chop up a singular simplex into smaller ones whose diameter can be made arbitrarily small. This is quite involved, and rather unpleasant. On the other hand, it is not very hard to work out how to chop up a singular cube into smaller ones!

Remark 27.11. For us the main reason we preferred cubes on simplices in Lecture 22 is that it is much easier to define an integral over a cube (it is just a nested sequence of integrals $\int_0^1 \cdots \int_0^1$), whereas this is messier for simplices.

Remark 27.12. At a much more advanced level, there are compelling reasons both to prefer using simplices and to prefer using cubes. This concerns **cubical sets** and **simplicial sets** in homotopy theory. However this all goes way beyond the course, so we won't discuss it.

It is not completely obvious why the resulting cubical singular homology groups agree with the normal singular homology groups. This can be proved directly using the[1] *Acyclic Models Theorem*, or it can deduced from the uniqueness result for Eilenberg-Steenrod homology theories.

Let us now return to manifolds. If M is a smooth manifold then the vector spaces $Q_k(M)$ defined today do *not* coincide with the vector spaces $Q_k(M)$ defined in Lecture 22. This is because in Lecture 22 we insisted on *smooth* maps. Let us temporarily write $Q_k^\infty(M)$ for the smooth singular k-cubes, and $H_k^{\mathrm{cube},\infty}(M)$ for the homology of the chain complex $(\overline{Q}_k^\infty(M), \partial)$. It is not obvious that the two groups coincide, but luckily they do:

Theorem 27.13. *Let M be a smooth manifold. Then*

$$H_k^{\mathrm{cube}}(M;\mathbb{R}) \cong H_k^{\mathrm{cube},\infty}(M;\mathbb{R}), \qquad \forall\, k \geq 0.$$

Proof (Sketch). The proof proceeds in six steps.

1. Suppose M is a single point. This is trivial.

2. Suppose M is an open contractible subspace of $\mathbb{R}^n$. This follows from Step 1 and the Whitney Approximation Theorem 6.14, which allows us to assume the contraction of M is smooth.

3. Suppose $M = U \cup V$, where U and V are open in M and the theorem is assumed to be true for U, V and $U \cap V$. We apply naturality of the Mayer-Vietoris sequence to see that the following diagram commutes, where we omit the coefficient group $\mathbb{R}$ so that the diagram fits on the page:

$$\begin{tikzcd}
H_k^{\mathrm{cube}}(U\cap V) \arrow[r] \arrow[d] & H_k^{\mathrm{cube}}(U)\oplus H_k^{\mathrm{cube}}(V) \arrow[r] \arrow[d] & H_k^{\mathrm{cube}}(M) \arrow[r] \arrow[d] & H_{k-1}^{\mathrm{cube}}(U\cap V) \arrow[d] \\
H_k^{\mathrm{cube},\infty}(U\cap V) \arrow[r] & H_k^{\mathrm{cube},\infty}(U)\oplus H_k^{\mathrm{cube},\infty}(V) \arrow[r] & H_k^{\mathrm{cube},\infty}(M) \arrow[r] & H_{k-1}^{\mathrm{cube},\infty}(U\cap V)
\end{tikzcd}$$

The Five Lemma then completes the proof.

4. Now assume $M = \bigcup_i U_i$, where $U_i \subset U_{i+1}$ is an open set, and the theorem is true for each U_i. Then the theorem follows for M via an abstract argument using filtered colimits[2]:

$$H_k^{\mathrm{cube}}(M;\mathbb{R}) \cong \varinjlim H_k^{\mathrm{cube}}(U_i;\mathbb{R}) = \varinjlim H_k^{\mathrm{cube},\infty}(U_i;\mathbb{R}) \cong H_k^{\mathrm{cube},\infty}(M;\mathbb{R})$$

5. Now assume M is an arbitrary open subset of $\mathbb{R}^n$. Then we can write M as a countable union of convex open subsets. For any finite union, the theorem holds by applying Step 3 and induction, and then Step 4 gives the result for M itself.

[1] I covered this my algebraic topology course here.

[2] I covered this my algebraic topology course here.

6. The general case: since we can cover M by charts, it follows from Step 5 and Zorn's Lemma that there exists a maximal open subset $U \subset M$ for which the theorem is true. If $U \neq M$, then we an find a chart domain V such that V is not contained in U. Then by Step 3 and Step 5, the theorem is true for $U \cup V$. This contradicts maximality of U. ∎

With this out of the way, we shall drop the ∞ from the notation and just write $\overline{Q}_k(M)$ for the groups defined in Lecture 22. Let us now recall how one constructs the *cohomology groups* from the homology groups.

Definition 27.14. Let X be a topological space. Set

$$Q^k(X) := \operatorname{Hom}(\overline{Q}_k(X); \mathbb{R})$$

and define $d \colon Q^k(X) \to Q^{k+1}(X)$ by

$$d(\alpha)(q) := \alpha(\partial q), \qquad \alpha \in Q^k(X),\ q \in \overline{Q}_{k+1}(X).$$

Then $d^2 = 0$, and hence $(Q^\bullet(X), d)$ is cochain complex. Its homology is denoted by $H^k_{\text{cube}}(X;\mathbb{R})$ and referred to as the **cubical singular cohomology** of X.

The next lemma is just a restatement of Corollary 22.27 (which itself was essentially a restatement of the Local Stokes' Theorem 22.25).

Lemma 27.15. *Let M be a smooth manifold. Then integration induces a cochain map*

$$\Phi \colon \Omega^\bullet(M) \to Q^\bullet(M), \qquad \Phi[\omega][q] := \int_q \omega.$$

We can now state the main result of the lecture.

Theorem 27.16 (The de Rham Theorem). *The integration cochain map Φ induces a natural isomorphism $H^k_{\mathrm{dR}}(M) \to H^k_{\text{cube}}(M;\mathbb{R})$.*

Proof (Sketch). The proof proceeds in the same fashion as Theorem 27.13.

1. Suppose M is an open convex subset of $\mathbb{R}^n$. Then the theorem follows from Corollary 23.20 and standard properties of singular cohomology.

2. Suppose $M = U \cup V$, where U and V are open in M and the theorem is assumed to be true for U, V and $U \cap V$. The theorem will gain follow for M via a standard argument using the Mayer-Vietoris sequences and the Five Lemma, apart from the fact that we have not constructed the Mayer-Vietoris sequence in de Rham cohomology. Let us rectify this. We denote by

$$\imath_U \colon U \cap V \hookrightarrow U, \qquad \imath_V \colon U \cap V \hookrightarrow V$$

and

$$\jmath_U \colon U \hookrightarrow M, \qquad \jmath_V \colon V \hookrightarrow M$$

the inclusions. One then defines

$$\alpha \colon \Omega^\bullet(M) \to \Omega^\bullet(U) \oplus \Omega^\bullet(V), \qquad \alpha(\omega) := (\jmath_U^\star(\omega), \jmath_V^\star(\omega))$$

and

$$\beta\colon \Omega^\bullet(U) \oplus \Omega^\bullet(V) \to \Omega^\bullet(U \cap V), \qquad \beta(\omega_1, \omega_2) := \imath_U^\star(\omega_1) - \imath_V^\star(\omega_2).$$

Then we claim that

$$0 \to \Omega^\bullet(M) \xrightarrow{\alpha} \Omega^\bullet(U) \oplus \Omega^\bullet(V) \xrightarrow{\beta} \Omega^\bullet(U \cap V) \to 0$$

is exact. The only claim that isn't clear is why β should be surjective. To see this, let $\{\lambda_U, \lambda_V\}$ be a partition of unity subordinate to the open cover $\{U, V\}$. Then if ω is an r-form on $U \cap V$, we can think of $\lambda_U\,\omega$ and $\lambda_V\,\omega$ as r-forms on U and V respectively, and $\beta(\lambda_U\,\omega, -\lambda_V\,\omega) = \omega$.

3. Now assume $M = \bigcup_i U_i$, where $U_i \subset U_{i+1}$ is an open set such that $\overline{U}_i$ is compact for each i, and assume the theorem is true for each U_i. Then the theorem is also true for M. The proof of this is considerably harder than the proof of Step 4 of Theorem 27.13, since now we are working with cohomology, and thus instead of filtered colimits we have filtered *limits*. It is a sad fact of life that limits are less well behaved than colimits, and are not exact functors from diagrams of vector spaces to diagrams of vector spaces. Consequently we needs to worry about the first right *derived* functors $R^1 \varprojlim$. But this is not too bad: since $\overline{U}_i$ is compact[3] one has $R^1 \varprojlim \Omega^k(U_i) = 0$ for all k and i, and thus we have a natural short exact sequence

$$0 \to R^1 \varprojlim H_{\mathrm{dR}}^{k-1}(U_i)) \to H_{\mathrm{dR}}^k(M) \to \varprojlim H_{\mathrm{dR}}^k(U_i) \to 0.$$

A similar sequence holds for $H^k_{\mathrm{cube}}(M)$, and naturality of the two sequences allow us to conclude this step.

4. Now assume M is an arbitrary open subset of $\mathbb{R}^n$. Then we can write M as a countable union of convex open subsets. For any finite union, the theorem holds by applying Step 2 and induction, and then Step 3 gives the result for M itself.

5. The general case: this follows from the previous step, since M has a countable basis of open sets diffeomorphic to open sets of Euclidean space. ■

We conclude this lecture with a brief discussion of **Poincaré duality**. This states that for a compact connected orientable manifold M of dimension n, one has

$$H^k_{\mathrm{cube}}(M; \mathbb{R}) \cong H_{n-k}^{\mathrm{cube}}(M; \mathbb{R}),$$

and hence there is a non-degenerate pairing

$$H^k_{\mathrm{cube}}(M; \mathbb{R}) \times H^{n-k}_{\mathrm{cube}}(M; \mathbb{R}) \to \mathbb{R}.$$

For de Rham cohomology, this pairing is particularly easy to understand: it is induced from the pairing

$$\Omega^k(M) \times \Omega^{n-k}(M) \to \mathbb{R}, \qquad (\omega, \vartheta) \mapsto \int_M \omega \wedge \vartheta$$

(this is well-defined since M is assumed to be compact and oriented). I will leave it to you to investigate how to prove this.

Enjoy your winter vacation, and see you next semester!

[3] I am skipping some details here, but the underlying arguments are not that hard.

LECTURE 28

Connections on vector bundles

Welcome to Differential Geometry II!

Hopefully you all had an enjoyable winter vacation and are now refreshed and ready to study again[1]. Our first major topic this semester is **connections**. To motivate the notion of a connection, let consider the following rather simple idea.

Let M be a smooth manifold. Suppose $f \in C^\infty(M)$ is a smooth function and $X \in \mathfrak{X}(M)$ is a vector field. We can feed f to X to get another smooth function $X(f) = df(X)$. Now consider the trivial one-dimensional vector bundle $M \times \mathbb{R} \to M$ over M. There is an obvious bijective correspondence between smooth functions f on M and sections $s \in \Gamma(M \times \mathbb{R})$. Explicitly, any section s can be uniquely written as $s(x) = s_f(x) = (x, f(x))$ for a smooth function f.

Thus the operation $f \mapsto X(f)$ can also be thought of as an operator on the space of sections of the trivial bundle. We write this operator as ∇_X:

$$\nabla_X \colon \Gamma(M \times \mathbb{R}) \to \Gamma(M \times \mathbb{R}),$$
$$\nabla_X(s_f)(x) := s_{X(f)}(x) = \big(x, X(f)(x)\big).$$

The operator ∇_X is local operator in the sense of Definition 16.17 but—provided X is not identically zero—it is not a point operator. (Exercise: Prove this!)

Next, note that the value of $\nabla_X(s)$ at a point x depends on X only via the tangent vector $X(x)$. Indeed, if $\gamma \colon (-\varepsilon, \varepsilon) \to M$ is any smooth curve with $\gamma(0) = x$ and $\gamma'(0) = X(x)$ then (up to identifying s_f with s) we have

$$\nabla_X(s_f)(x) = df|_x(X(x)) = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)) = \lim_{t \to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}. \tag{28.1}$$

This shows that we can think of $\nabla_X(s)$ as "differentiating s in the direction of X".

Let us now see what goes wrong with extending this idea to an arbitrary vector bundle. Let $\pi \colon E \to M$ be a vector bundle. As before, let X be a vector field on M and let $s \in \Gamma(E)$. Fix a point $x \in M$ and let γ denote a smooth curve with $\gamma(0) = x$ and $\gamma'(0) = X(x)$. We can again attempt to "define" a new section via (28.1)

$$\text{“}\nabla_X(s)\text{”}(x) = \lim_{t \to 0} \frac{s(\gamma(t)) - s(\gamma(0))}{t}. \tag{28.2}$$

A moment's thought reveals that (28.2) is nonsense: the vector $s(\gamma(t))$ belongs to the vector space $E_{\gamma(t)}$, and for different values of t these are *different* vector spaces. Thus is simply *does not make sense* to add or subtract them from one another.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] After all, it's been nearly 72 hours since the last Differential Geometry I exam finished. . . ugh. More holiday please.

Compare this to our original problem right at the beginning of Lecture 1 when we motivated manifolds: on an arbitrary topological space one cannot simply "add" points together. On a vector bundle whilst each fibre has a linear structure, in general each fibre is a different vector space, and thus we cannot add points.

The reason this worked on the trivial bundle $M \times \mathbb{R} \to M$ was that in this case each fibre $\{x\} \times \mathbb{R}$ was *canonically* isomorphic to $\mathbb{R}$ via the second projection. Equivalently, the identification $s = s_f$ of sections of $M \times \mathbb{R}$ with smooth functions on M was canonical—no choices were needed. This is also reflected in the fact that on the trivial bundle ∇_X can be identified with the Lie derivative $\mathcal{L}_X$.

More generally, the process we described at the start of the lecture works on any trivial bundle, and this leads us to the first definition of the course.

DEFINITION 28.1. Let M be a smooth manifold and let $E = M \times \mathbb{R}^k$ denote the trivial bundle over M. The **trivial connection** on E associates to every vector field X on M the operator

$$\nabla_X : \Gamma(E) \to \Gamma(E)$$

given by

$$\nabla_X(s_f)(x) := \lim_{t\to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t},$$

where we identify each section $s = s_f$ with a smooth function $f : M \to \mathbb{R}^k$.

Thus (by definition) any trivial vector bundle admits a trivial connection. In fact, the converse is true: if E admits a trivial connection then E is necessarily a trivial bundle, although this will take us some time to prove (see Proposition 32.14), and will require us to give an alternative definition of a connection that does not explicitly reference the trivialisation.

We first define the weaker notion of a *preconnection*, which will work in an arbitrary fibre bundle. As with many of the concepts we've seen in Differential Geometry, the relation between the formal definition of a connection and Definition 28.1 will at first sight not be so obvious. We will rectify this in Lecture 32.

DEFINITION 28.2. Let $\pi : E \to M$ be a fibre bundle over a smooth manifold M with fibre F. A **preconnection**[2] on E is a distribution $\mathcal{H}$ on E (i.e. a vector subbundle of the tangent bundle TE) with the additional property that for every $p \in E$ the map $D\pi(p)|_{\mathcal{H}_p} : \mathcal{H}_p \to T_{\pi(p)}M$ is a linear isomorphism.

Let us unpack this a bit. Requiring that $D\pi(p)|_{\mathcal{H}_p} : \mathcal{H}_p \to T_{\pi(p)}M$ is a linear isomorphism for every $p \in E$ is the same thing as saying that the restriction of $D\pi : TE \to TM$ to $\mathcal{H}$ is a vector bundle isomorphism along $\pi : E \to M$ in the sense of Definition 14.3. Equivalently, a preconnection is a vector subbundle of TE such that

$$\mathcal{H} \cong \pi^\star(TM). \tag{28.3}$$

Here $\pi^\star(TM)$ is the pullback of the bundle[3] $\pi_{TM} : TM \to M$ along $\pi : E \to M$. Explicitly:

$$\pi^\star(TM) = \{(p, v) \in E \times TM \mid \pi(p) = \pi_{TM}(v)\}$$

[2]This terminology is not standard. Some authors call what we call a preconnection an *Ehresmann connection*. However this won't matter, since we will shortly upgrade a preconnection to a genuine connection, and then will never have cause to speak about preconnections anymore.

[3]When there are multiple bundles in play, we will label the various footpoint maps where needed.

This is a vector bundle over E (even if E is not itself a vector bundle over M!)

Here is another way to express this condition. Recall we denote by $VE = \ker D\pi$ the vertical bundle of E (cf. Problem I.5). Then a preconnection is a distribution on E which is *complementary* to VE, i.e.

$$TE = \mathcal{H} \oplus VE. \tag{28.4}$$

as vector bundles over E.

So much for preconnections. If we instead start with a vector bundle, we can impose an additional condition on a preconnection, which gives rise to a *connection*.

DEFINITION 28.3. Let $\pi\colon E \to M$ be a vector bundle and let $a \in \mathbb{R}$. We denote by $\mu_a\colon E \to E$ scalar multiplication by a:

$$\mu_a\colon E \to E, \qquad \mu_a(p) = ap.$$

Since μ is fibre preserving, the map μ_a is a vector bundle homomorphism from E to itself (cf. Definition 14.5) for every $a \in \mathbb{R}$. For $a \neq 0$, μ_a is a diffeomorphism with inverse $\mu_{1/a}$.

DEFINITION 28.4. Let $\pi\colon E \to M$ be a vector bundle over a smooth manifold M. A **connection** $\mathcal{H}$ on E is a preconnection with the additional property that for every $p \in E$ and every $a \in \mathbb{R}$, one has

$$D\mu_a(p)[\mathcal{H}_p] = \mathcal{H}_{ap}. \tag{28.5}$$

The true significance of this condition won't become apparent until we discuss connections on principal bundles in Lecture 38 (see Proposition 38.6 in particular), although see Problem N.1 for one key consequence of (28.5).

For now, let us now prove that (pre)connections always exist.

THEOREM 28.5. *Every fibre bundle admits a preconnection. Every vector bundle admits a connection.*

Proof. We prove the result in two steps.

1. We first prove the result when $E = M \times F$ is the trivial bundle. Let $\imath_v\colon M \to M \times F$ denote the map $x \mapsto (x, v)$, and set

$$\mathcal{H}_{(x,v)} := D\imath_v(x)[T_xM].$$

This is a preconnection. If in addition $F = \mathbb{R}^k$ is a vector space then this is a connection, since $\mu_a \circ \imath_v = \imath_{av}$ and thus

$$D\mu_a(x,v)[\mathcal{H}_{(x,v)}] = D\mu_a(x,v) \circ D\imath_v(x)[T_xM] = D\imath_{av}(x)[T_xM] = \mathcal{H}_{(x,av)}.$$

2. For the general case, we use a partition of unity argument. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote an open cover of M such that E is trivial over each $U_{\mathtt{a}}$, and let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote a partition of unity subordinate to this cover. Let $\mathcal{H}^{\mathtt{a}}$ denote a (pre)connection on $\pi^{-1}(U_{\mathtt{a}})$, whose existence is guaranteed by Step 1. Given $x \in M$ and $p \in E_x$, define

$$H_p \colon T_x M \to T_p E, \qquad H_p(v) := \sum_{\{\mathtt{a} \in \mathtt{A} \mid x \in U_{\mathtt{a}}\}} \lambda_{\mathtt{a}}(x)\, v_{\mathtt{a}},$$

where $v_{\mathtt{a}}$ is the unique vector in $\mathcal{H}_p^{\mathtt{a}}$ such that $D\pi(p)[v_{\mathtt{a}}] = v$. Then H_p is a linear map such that $D\pi(p) \circ H_p = \mathrm{id}_{T_x M}$. We then define

$$\mathcal{H}_p := H_p[T_x M].$$

This is a (pre)connection. ■

We can use (pre)connections to lift vectors from TM to TE.

Definition 28.6. Let $\pi \colon E \to M$ be a fibre bundle, and let $\mathcal{H}$ be a preconnection on E. The splitting $TE = \mathcal{H} \oplus VE$ allows us to uniquely express any vector $\zeta \in TE$ as

$$\zeta = \zeta^{\mathtt{H}} + \zeta^{\mathtt{V}}$$

where if $\zeta \in T_p E$ then $\zeta^{\mathtt{H}} \in \mathcal{H}_p$ and $\zeta^{\mathtt{V}} \in V_p E$. We call $\zeta^{\mathtt{H}}$ the **horizontal component** of ζ and $\zeta^{\mathtt{V}}$ the **vertical component** of ζ. A vector is **horizontal** if $\zeta^{\mathtt{V}} = 0$ and **vertical** if $\zeta^{\mathtt{H}} = 0$.

This is consistent with Definition 26.14. The property of being horizontal depends on the specific choice of preconnection, but the property of being vertical does not.

Definition 28.7. Let $\pi \colon E \to M$ be a fibre bundle, and let $\mathcal{H}$ be a preconnection on E. Let $x \in M$, $p \in E_x$ and $v \in T_x M$. The **horizontal lift of v at p** is the unique vector $\overline{v} \in \mathcal{H}_p$ such that $D\pi(p)[\overline{v}] = v$.

Since $p \mapsto \mathcal{H}_p$ is smooth (this is true of any distribution) we can also lift vector fields.

Definition 28.8. Let $\pi \colon E \to M$ be a fibre bundle, and let $\mathcal{H}$ be a preconnection on E. If $X \in \mathfrak{X}(M)$ is a vector field then the **horizontal lift** of X is the unique vector field $\overline{X} \in \mathfrak{X}(E)$ such that $\overline{X}(p)$ is the horizontal lift of $X(\pi(p))$ at p for each $p \in E$.

The following result is almost immediate.

Lemma 28.9. *Let $\pi \colon E \to M$ be a fibre bundle and let $\mathcal{H}$ be a preconnection on E. Given $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, we have:*

(i) $\overline{X + Y} = \overline{X} + \overline{Y}$,

(ii) $\overline{fX} = (f \circ \pi)\overline{X}$,

(iii) $\overline{[X, Y]} = [\overline{X}, \overline{Y}]^{\mathtt{H}}$.

Proof. The first two statements are obvious. For the third we observe that

$$D\pi[\overline{X}, \overline{Y}] = [X, Y] = D\pi\overline{[X, Y]},$$

and thus $\overline{[X, Y]} = [\overline{X}, \overline{Y}]^{\mathtt{H}}$ by definition of a preconnection. ■

We next show that (pre)connections behave nicely under pullbacks. Suppose $\pi\colon E \to N$ is a fibre bundle and $\varphi\colon M \to N$ is a smooth map. The pullback bundle $\varphi^\star E \to M$ fits into the following commutative diagram.

$$\begin{tikzcd} \varphi^\star E \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\mathrm{pr}_1"'] & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd} \tag{28.6}$$

Recall from part (ii) of Problem G.7 that

$$T_{(x,p)}(\varphi^\star E) = \Big\{(v,\zeta) \in T_xM \times T_pE \mid D\varphi(x)[v] = D\pi(p)[\zeta]\Big\}. \tag{28.7}$$

Suppose $\mathcal{H}$ is a (pre)connection on E. We define

$$\varphi^\star\mathcal{H} := (D\,\mathrm{pr}_2)^{-1}[\mathcal{H}]$$

that is,

$$(\varphi^\star\mathcal{H})_{(x,p)} := \big\{(v,\zeta) \in T_{(x,p)}(\varphi^\star E) \mid D\,\mathrm{pr}_2(x,p)[v,\zeta] \in \mathcal{H}_p\big\}\,.$$

Proposition 28.10. *$\varphi^\star\mathcal{H}$ is a preconnection on $\varphi^\star E$. If E is a vector bundle and $\mathcal{H}$ is a connection on E then $\varphi^\star\mathcal{H}$ is a connection on the vector bundle $\varphi^\star E$.*

Proof. It follows from (28.7) that $V(\varphi^\star E) = \{0\} \times VE$ and that $\varphi^\star\mathcal{H}$ is given by

$$(\varphi^\star\mathcal{H})_{(x,p)} := \{(v,\zeta) \in T_xM \times \mathcal{H}_p \mid D\varphi(x)[v] = D\pi(p)[\zeta]\}\,.$$

Since any $\zeta \in TE$ decomposes uniquely as $\zeta^{\mathtt{H}} + \zeta^{\mathtt{V}} \in \mathcal{H} \oplus VE$, any $(v,\zeta) \in T(\varphi^\star E)$ decomposes uniquely as

$$(v,\zeta) = (v,\zeta^{\mathtt{H}}) + (0,\zeta^{\mathtt{V}}) \in \varphi^\star(\mathcal{H}) \oplus V(\varphi^\star E).$$

This shows that $\varphi^\star\mathcal{H}$ is complementary to $V(\varphi^\star E)$, and thus $\varphi^\star\mathcal{H}$ is a preconnection on $\varphi^\star E$. If E is a vector bundle and $\mathcal{H}$ is a connection then so is $\varphi^\star\mathcal{H}$, since if $(v,\zeta) \in \varphi^\star\mathcal{H}$ we have

$$D\mu_a(x,p)[v,\zeta] = (v, D\mu_a(p)[\zeta]).$$

This completes the proof. ∎

Parallel transport

In this lecture we define **parallel transport** axiomatically. Next week we will prove that the existence of a parallel transport system is equivalent to the existence of a connection, and thus, going forward we will view the two interchangeably. We begin with some more terminology. In the last lecture we defined what it meant for a tangent vector (or a vector field) to be horizontal. Now we explain what it means for a *section* to be horizontal.

DEFINITION 29.1. Let $\pi\colon E \to M$ be a fibre bundle with preconnection $\mathcal{H}$. A section $s \in \Gamma(E)$ is said to be **horizontal** if

$$Ds(x)[T_xM] = \mathcal{H}_{s(x)}, \qquad \forall\, x \in M.$$

Similarly a local section $s \in \Gamma(U, E)$ is horizontal if the above holds for all $x \in U$.

Next, we introduce the idea of a section *along* a map.

DEFINITION 29.2. Suppose $\pi\colon E \to N$ is a fibre bundle over a smooth manifold and $\varphi\colon M \to N$ is a smooth map. A **section of** E **along** φ is a smooth map $s\colon M \to E$ such that $s(x) \in E_{\varphi(x)}$. We denote by $\Gamma_\varphi(E)$ the space of such sections. If $U \subset M$ is an open set then we can also speak of the space $\Gamma_\varphi(U, E)$ of smooth maps $s\colon M \to E$ such that $s(x) \in E_{\varphi(x)}$ for all $x \in U$—we refer to these as **local sections along** φ.

LEMMA 29.3. *Suppose $\pi\colon E \to N$ is a fibre bundle over a smooth manifold and $\varphi\colon M \to N$ is a smooth map. There is a bijective correspondence between sections of the pullback bundle $\varphi^\star E \to M$ and sections of E along φ. Thus:*

$$\Gamma_\varphi(E) \cong \Gamma(\varphi^\star E).$$

The same is true for local sections.

Proof. Let $\mathrm{pr}_2\colon \varphi^\star E \to E$ denote the second projection (cf. (28.6) from the last lecture). If $\tilde{s} \in \Gamma(\varphi^\star E)$ then

$$s = \mathrm{pr}_2 \circ \tilde{s}$$

is a section of E along φ. Conversely a section s of E along φ uniquely determines a section $\tilde{s} \in \Gamma(\varphi^\star E)$ by the same equation. ■

As a result of Lemma 29.3, we will often simply identify elements of $\Gamma_\varphi(E)$ and $\Gamma(\varphi^\star E)$, and write them both with the same letter.

Definition 29.4. Let $\pi\colon E \to N$ be a fibre bundle and let $\mathcal{H}$ be a preconnection on E. Suppose $\varphi\colon M \to N$ is a smooth map and $s \in \Gamma_\varphi(E)$ is a section of E along φ. We say that s is **horizontal along** φ if the corresponding section of $\varphi^\star E$ is horizontal with respect to the pullback connection $\varphi^\star\mathcal{H}$. Explicitly, this just means that

$$Ds(x)[T_xM] \subset \mathcal{H}_{s(x)}, \qquad \forall x \in M.$$

Here are two examples to illustrate the versatility of this definition.

Example 29.5. Take $M = N$ and let φ be the identity. Then a section of E along φ is the same thing as a section of E, and a section s is horizontal along φ if and only s is horizontal in the sense of Definition 29.1.

Example 29.6. Take M to be an interval (a, b) and $\varphi = \gamma\colon (a, b) \to N$ to be a smooth curve[1] in N. We will usually use the special letter c (instead of s) to denote a section along a curve. Thus a section $c \in \Gamma_\gamma(E)$ is simply a smooth curve in E such that $c(t) \in E_{\gamma(t)}$ for all $t \in (a, b)$. Moreover c is horizontal along γ if

$$c'(t) \in \mathcal{H}_{c(t)}, \qquad \forall\, t \in (a, b).$$

Proposition 29.7. *Let $\pi\colon E \to M$ be a fibre bundle, and let $\mathcal{H}$ be a preconnection on E. Let $\gamma\colon [a, b] \to M$ be a smooth curve and let $t_0 \in [a, b]$. Then for any $p \in E_{\gamma(t_0)}$, there exists a unique horizontal section c of E along γ such that $c(t_0) = p$.*

Proof. Abbreviate by T the vector field $\frac{\partial}{\partial t}$ on $[a, b]$, and let $\overline{T}$ denote the horizontal lift of T to $\gamma^\star E$ with respect to the pullback connection $\gamma^\star\mathcal{H}$. Let δ denote the integral curve of $\overline{T}$ in $\gamma^\star E$ such that $\delta(t_0) = (t_0, p)$. Consider the diagram (28.6) again, specialised to the case in hand:

$$\begin{array}{ccc} \gamma^\star E & \xrightarrow{\mathrm{pr}_2} & E \\ {\scriptstyle \mathrm{pr}_1}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ [a, b] & \xrightarrow[\gamma]{} & M \end{array}$$

We claim that $\mathrm{pr}_1 \circ \delta$ is an integral curve of T. To see this we compute

$$\begin{aligned} \frac{d}{dt}\big(\mathrm{pr}_1 \circ \delta\big)(t) &= D\,\mathrm{pr}_1(\delta(t))[\delta'(t)] \\ &= D\,\mathrm{pr}_1(\delta(t))[\overline{T}(\delta(t))] \\ &= T(\mathrm{pr}_1(\delta(t)). \end{aligned}$$

Since $\mathrm{pr} \circ \delta(t_0) = t_0$ and $\mathrm{pr}_1 \circ \delta$ is an integral curve of T, we must have $\mathrm{pr}_1 \circ \delta(t) = t$ for all $t \in [a, b]$. Thus δ is actually a section of $\gamma^\star E$. Moreover it follows from the definition of δ and $\overline{T}$ that δ is a horizontal section of $\gamma^\star E$. Thus by Lemma 29.3, $c := \mathrm{pr}_2 \circ \delta$ is a horizontal section of E along γ with $c(t_0) = p$. Finally, uniqueness is immediate from the uniqueness of integral curves. ∎

[1] It will often be convenient to work with smooth curves defined on a *closed* interval $[a, b]$. Here "smooth" can be interpreted as either requiring that there exists a smooth extension to some interval $(a - \varepsilon, b + \varepsilon)$, or just by considering $[a, b]$ as a smooth manifold with boundary. Note also that if $\gamma\colon [a, b] \to N$ is a smooth curve then $\gamma^\star E \to [a, b]$ is a vector bundle over a smooth manifold with boundary.

Here is main definition of today's lecture.

DEFINITION 29.8. Let $\pi \colon E \to M$ be a vector bundle over a smooth manifold. A **parallel transport system** $\mathbb{P}$ on E assigns to every point $p \in E$ and every curve $\gamma \colon [a,b] \to M$ with $\gamma(a) = \pi(p)$, a unique section $\mathbb{P}_\gamma(p) \in \Gamma_\gamma(E)$ with initial condition p, i.e. such that $\mathbb{P}_\gamma(p)(a) = p$. One calls $\mathbb{P}_\gamma(p)$ the **parallel lift** of γ starting at p. This association should satisfy the following four axioms:

(i) **(Linear isomorphism)**: For every smooth curve $\gamma \colon [a,b] \to M$ the map

$$\widehat{\mathbb{P}}_\gamma \colon E_{\gamma(a)} \to E_{\gamma(b)}, \qquad \widehat{\mathbb{P}}_\gamma(p) := \mathbb{P}_\gamma(p)(b)$$

is a linear isomorphism. Moreover

$$\widehat{\mathbb{P}}_\gamma^{-1} = \widehat{\mathbb{P}}_{\gamma^-}$$

where $\gamma^- \colon [a,b] \to M$ is the reverse curve $t \mapsto \gamma(a - t + b)$.

(ii) **(Independence of parametrisation)**: If $\gamma \colon [a,b] \to M$ is a smooth curve and $h \colon [a_1, b_1] \to [a,b]$ is a diffeomorphism such that $h(a_1) = a$ and $h(b_1) = b$ then for every point $p \in E_{\gamma(a)}$ and every $t \in [a_1, b_1]$, we have

$$\mathbb{P}_{\gamma \circ h}(p)(t) = \mathbb{P}_\gamma(p)(h(t)).$$

(iii) **(Smooth dependence on initial conditions)**: The section $\mathbb{P}_\gamma(p)$ depends smoothly[2] on both γ and p.

(iv) **(Initial uniqueness)**: Suppose $\gamma, \delta \colon [a,b] \to M$ are two curves such that $\gamma(a) = \delta(a)$ and $\gamma'(a) = \delta'(a)$. Then for each $p \in E_{\gamma(a)}$, the two curves $t \mapsto \mathbb{P}_\gamma(p)(t)$ and $t \mapsto \mathbb{P}_\delta(p)(t)$ have the same initial tangent vector:

$$\left.\frac{d}{dt}\right|_{t=a} \mathbb{P}_\gamma(p)(t) = \left.\frac{d}{dt}\right|_{t=a} \mathbb{P}_\delta(p)(t)$$

REMARK 29.9. In general if $\gamma \colon [a,b] \to M$ is a smooth curve on M and $c \in \Gamma_\gamma(E)$ is any section along γ then we say c is **parallel along** γ if $c = \mathbb{P}_\gamma(p)$ for some $p \in E_{\gamma(a)}$.

REMARK 29.10. It follows from Axiom (i) that $p \mapsto \mathbb{P}_\gamma(p)$ is also linear (where now addition and scalar multiplication take place in the vector space of sections $\Gamma_\gamma(E)$). Thus in particular $p \mapsto \mathbb{P}_\gamma(p)$ is smooth. Therefore the only content of Axiom (iii) is the smooth dependence on γ.

[2]This condition is a little tedious to state precisely, and it is not too important to get hung up on this detail, but here is one way to formulate it rigorously: *For every open set $U \subset M$ and every smooth map $\Psi \colon TM|_U \to M$ such that $\Psi(x, 0_x) = x$ for every $x \in U$, where 0_x is the zero vector in T_xM, the map*

$$TM|_U \oplus \pi^{-1}(U) \to E, \qquad (v,p) \mapsto \widehat{\mathbb{P}}_\gamma(p), \qquad \textit{where } \gamma(t) := \Psi(tv)$$

is smooth.

Example 29.11. Let $E = M \times \mathbb{R}^k$ be a trivial bundle. We define the **trivial parallel transport system** on E by declaring that constant sections are parallel. Explicitly, if $\gamma\colon [a,b] \to M$ is any smooth curve with $\gamma(a) = x$ then we define

$$\mathbb{P}_\gamma(x,v)(t) \coloneqq (\gamma(t), v), \qquad v \in \mathbb{R}^k.$$

We will see in Lecture 32 that this is consistent with Definition 28.1.

Remark 29.12. We will explore this further in Lecture 32, but for now note that a parallel transport system gives us a way to identify two different fibres E_x and E_y of a vector bundle over M: simply take a curve γ from x to y and consider the linear isomorphism $\widehat{\mathbb{P}}_\gamma\colon E_x \to E_y$. This will allow us to make sense of (28.2) from the last lecture, and thus let us differentiate sections along vector fields for non-trivial vector bundles.

Next lecture we will prove that a parallel transport system $\mathbb{P}$ determines and is uniquely determined by a connection $\mathcal{H}$. We conclude this lecture with a preliminary lemma needed in for the proof (and which will also be useful later on in the course). To ease the notation given a subset $Z \subset M$ we write

$$TM|_Z \coloneqq \pi_{TM}^{-1}(Z) = \bigsqcup_{x \in Z} T_xM.$$

If Z is open in M then $TM|_Z \cong TZ$.

Lemma 29.13. *Let M be a smooth manifold and let $x \in M$. Then there is an open set U containing x and a smooth map*

$$\Psi\colon TM|_U \to M$$

such that

$$\left.\frac{d}{dt}\right|_{t=0} \Psi(x, tv) = v, \qquad \forall\, v \in T_xM. \tag{29.1}$$

This proof is non-examinable.

(♣) *Proof.* Let $\sigma\colon U \to O$ be a chart about x. Without loss of generality we may assume $0 \in O$ and $\sigma(x) = 0$. Let $\tilde{\sigma}\colon TM|_U \to O \times \mathbb{R}^n$ denote the associated chart on the tangent bundle TM over U (cf. the proof of Theorem 4.16), given explicitly by

$$\tilde{\sigma}(y,v) = \big(\sigma(y), dx^i|_y(v)\, e_i\big)$$

Abbreviate $\alpha(y,v) \coloneqq dx^i|_y(v)\, e_i$ so that $\tilde{\sigma} = (\sigma, \alpha)$ and α is a vector bundle chart on TM (cf. Example 13.15). Let us also write

$$\lambda\colon TM|_U \to T_xM, \qquad \lambda(y,v) \coloneqq \alpha|_{T_xM}^{-1}(\alpha(y,v)), \tag{29.2}$$

so that $\lambda|_{T_xM} = \mathrm{id}_{T_xM}$.

Next, up to shrinking O, we may assume that $V_x \coloneqq \alpha|_{T_xM}^{-1}(O)$ is a starshaped open set in T_xM containing 0_x. Set

$$\psi\colon V_x \to M, \qquad \psi(v) \coloneqq \sigma^{-1}(\alpha|_{T_xM}(v)).$$

Note that $\psi(0_x) = x$. We claim moreover that under the canonical isomorphism $T_{0_x}V_x \cong T_xM$ (Problem B.3) we have

$$D\psi(0_x) = \mathrm{id}_{T_xM} \colon T_xM \to T_xM. \tag{29.3}$$

To see this we observe that if $w = w^i \frac{\partial}{\partial x^i}\big|_x$ then

$$\begin{aligned} D\psi(0_x)[\mathcal{J}_{0_x}(w)] &= \frac{d}{dt}\Big|_{t=0} \psi(0_x + tw) \\ &= \frac{d}{dt}\Big|_{t=0} \sigma^{-1}(t dx^i|_x(w)e_i) \\ &= D\sigma^{-1}(0)[w^i e_i] \\ &= w. \end{aligned}$$

This proves (29.3). By the Inverse Function Theorem 5.2, up to shrinking V_x we may assume that ψ is a diffeomorphism onto its image. We now wish to extend ψ to a map defined on all of $TM|_U$; this map will be called Ψ. We require our extension to satisfy

$$\Psi(\psi(v), w) = \psi\big(v + \lambda \circ D\psi(v)^{-1}[w]\big), \tag{29.4}$$

whenever $v \in V_x$ is sufficiently close to 0_x and $w \in T_{\psi(v)}M$ is sufficiently close to $0_{\psi(v)}$. Equation 29.4 really is an extension of ψ, since if we take $v = 0_x$ and choose $w \in V_x$ then $\psi(v) = x$ and the defining equation for Ψ becomes

$$\Psi(x, w) = \psi(0_x + D\psi(0_x)^{-1}[w]) = \psi(w),$$

where we used (29.3). That (29.4) is well-defined follows from the fact that ψ is locally a diffeomorphism, and thus for an appropriate cutoff function $\eta(y, w)$ we can define Ψ as

$$\Psi(y, w) := \eta(y, w)\, \psi\big(\psi^{-1}(y) + \lambda \circ D\psi(\psi^{-1}(y))^{-1}[w]\big)$$

Finally, (29.1) follows from (29.3). ■

LECTURE 30

The equivalence of connections and parallel transport systems

The goal of this lecture is to prove that a parallel transport system $\mathbb{P}$ determines and is uniquely determined by a connection $\mathcal{H}$. This result is quite involved, and we prove each direction separately.

THEOREM 30.1. *Let $\pi\colon E \to M$ be a vector bundle, and let $\mathbb{P}$ be a parallel transport system on E. Then $\mathbb{P}$ determines a connection $\mathcal{H}$ on E. This connection has the property that a section c along a curve γ is parallel in the sense of Remark 29.9 if and only if c is horizontal with respect to $\mathcal{H}$ in the sense of Definition 29.1.*

This proof is non-examinable.

(♣) *Proof.* Throughout we assume that M has dimension n, and that E is a vector bundle of rank k. Given $p \in E$, define $\mathcal{H}_p \subset T_pE$ to be the set of all tangent vectors ζ such that there exists a smooth curve[1] $\gamma\colon [0,1] \to M$ with

$$\zeta = \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma(p)(t).$$

Then set

$$\mathcal{H} := \bigsqcup_{p \in E} \mathcal{H}_p \subset TE.$$

We will prove that $\mathcal{H}$ is a connection in four steps. In the fifth and final step we prove the last sentence of the theorem.

1. Fix a point $x \in M$, and let $\Psi\colon TM|_U \to M$ denote the map from Lemma 29.13, where U is a suitable neighbourhood of x. Choose a point $p \in E_x$. In this step we show that $\mathcal{H}_p$ is the image of a linear map $T_xM \to T_pE$, and thus in particular a vector space. Given $v \in T_xM$, let γ_v denote the curve $t \mapsto \Psi(x, tv)$, and let $c_v := \mathbb{P}_{\gamma_v}(p) \in \Gamma_{\gamma_v}(E)$. We then define

$$C\colon T_xM \times [0,1] \to E, \qquad C(v,t) := c_v(t)$$

By Axiom (iii) of Definition 29.8, C is a smooth map. By Axiom (ii), $c_v(t) = c_{vt}(1)$, and thus

$$c_v'(0) = \frac{d}{dt}\Big|_{t=0} c_{tv}(1) = DC(0_x, 1)[\mathcal{J}_{0_x}(v), 0].$$

Thus the map $T_xM \to T_pE$ that sends v to $c_v'(0)$ is linear, as it is the composition of linear maps. If we call this map $L\colon v \mapsto c_v'(0)$ then $\mathcal{H}_p$ is equal (by definition) to $\operatorname{im} L$. Thus $\mathcal{H}_p$ is a vector space, as claimed.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] The fact that we choose a fixed interval $[0,1]$ is just for convenience; by Axiom (ii) of Definition 29.8 any interval will work.

2. In this step we show that $T_pE = \mathcal{H}_p \oplus V_pE$. We already know that $\mathcal{H}_p$ is a vector space of dimension at most n by the previous step. With L as before, we have

$$\begin{aligned} D\pi(p) \circ L(v) &= D\pi(p)[c_v'(0)] \\ &= \frac{d}{dt}\Big|_{t=0} \pi(c_v(t)) \\ &= \frac{d}{dt}\Big|_{t=0} \Psi(x, tv) \\ &= v, \end{aligned}$$

where the last equality used (29.1).

3. In this step, we prove that $\mathcal{H} \to E$ is a vector subbundle of TE. This requires us to build bundle charts. For this we recall from the solution to part (ii) of Problem I.5 that the map

$$\mathcal{J} \colon \pi^\star E \to VE, \qquad (p, q) \mapsto \mathcal{J}_p(q) = \frac{d}{dt}\Big|_{t=0} p + tq$$

is a vector bundle isomorphism. If we denote by $\mathrm{pr}_2 \colon VE \to E$ the map $\mathrm{pr}_2(\mathcal{J}_p(q)) := q$ then pr_2 is a vector bundle isomorphism from VE to E along π, and the following diagram commutes:

$$\begin{tikzcd} VE \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\pi_{VE}"'] & E \arrow[d, "\pi"] \\ E \arrow[r, "\pi"'] & M \end{tikzcd} \tag{30.1}$$

Fix a chart $\sigma \colon U \to O$ on M and a vector bundle chart $\beta \colon \pi^{-1}(U) \to \mathbb{R}^k$ on E. Define a smooth map

$$\tilde{\beta} \colon TE|_{\pi^{-1}(U)} \to \mathbb{R}^k, \qquad \tilde{\beta}(p, \zeta) := \beta \circ \mathrm{pr}_2 \left(\zeta - \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma(p)(t) \right),$$

where $\gamma \colon [0, 1] \to M$ is any smooth curve with $\gamma(0) = \pi(p)$ and $\gamma'(0) = D\pi(p)[\zeta]$. The value of $\tilde{\beta}(p, \zeta)$ does not depend on the choice of γ by Axiom (iv) of Definition 29.8. If $\zeta \in \mathcal{H}_p$ then $\tilde{\beta}(p, \zeta) = 0$ by definition of $\mathcal{H}_p$. As in the proof of Lemma 29.13, let α be the vector bundle chart on TM associated to σ, so that $\alpha(y, v) = dx^i|_y(v)\, e_i$. Then the map

$$\chi \colon TE|_U \to \mathbb{R}^n \times \mathbb{R}^k, \qquad \chi(p, \zeta) := \big(\alpha \circ D\pi(p)[\zeta], \tilde{\beta}(p, \zeta)\big)$$

is a vector bundle chart for TE over U (i.e. (π_{TE}, χ) is a diffeomorphism). Since $\tilde{\beta}[\mathcal{H}_p] = 0$, we have

$$\big(\pi_{TE}(p, \zeta), \chi(p, \zeta)\big) = \big(p, \alpha \circ D\pi(p)[\zeta], 0\big)$$

for all $\zeta \in \mathcal{H}_p$. Thus the restriction $\alpha \circ D\pi$ can serve as a vector bundle chart on $\mathcal{H}$ that will turn $\mathcal{H}$ into a vector subbundle of TE, provided the transition function is smooth. This however is easy to check: if σ_1 is another chart with corresponding vector bundle chart α_1 on TM then if x is a point in their overlapping domains and $p \in E_x$ we compute

$$\begin{aligned}\big(\pi_{\mathcal{H}}, \alpha_1 \circ D\pi\big) \circ \big(\pi_{\mathcal{H}}, \alpha \circ D\pi\big)^{-1}(p, e_i) &= \big(\pi_{\mathcal{H}}, \alpha_1 \circ D\pi\big) \circ D\pi(p)|_{\mathcal{H}}^{-1} \left[\frac{\partial}{\partial x^i}\Big|_x \right] \\ &= \left(p, \alpha_1 \left(x, \frac{\partial}{\partial x^i}\Big|_x \right) \right) \\ &= \left(p, dy^j|_x \left(\frac{\partial}{\partial x^i}\Big|_x \right) e_j \right),\end{aligned}$$

where y^j are the local coordinates of σ_1. This shows (cf. the proof of Theorem 4.16) that the transition function is given simply by $D(\sigma_1 \circ \sigma^{-1}) \circ (\sigma \circ \pi)$, which is smooth.

4. The global splitting $TE = \mathcal{H} \oplus VE$ follows from the pointwise splitting already proved in Step 2. Thus we have now proved that $\mathcal{H}$ is a preconnection on E. It remains therefore to prove that $\mathcal{H}$ is a connection.

Fix $x \in M$ and $p \in E_x$. We need to show that for any $a \in \mathbb{R}$,

$$D\mu_a(p)[\mathcal{H}_p] = \mathcal{H}_{ap}$$

where μ_u is scalar multiplication in the fibres, as in Definition 28.3. Let $\gamma\colon [0,1] \to M$ denote a smooth curve with $\gamma(0) = x$, and let $c := \mathbb{P}_\gamma(p)$. By linearity of parallel transport (this is Axiom (i) of Definition 29.8), $\mu_a \circ c$ is also parallel along γ. Since

$$D\mu_a(p)[c'(0)] = \frac{d}{dt}\Big|_{t=0} (\mu_a \circ c)(t)$$

we see that $D\mu_a(p)[\mathcal{H}_p] \subset \mathcal{H}_{ap}$. Then since

$$D\pi(ap) \circ D\mu_a(p)[c'(0] = D\pi(p)[c'(0)]$$

and $D\pi(ap)$ maps $\mathcal{H}_{ap}$ isomorphically onto T_xM, it follows that $D\mu_a(p)[\mathcal{H}_p] = \mathcal{H}_{ap}$. This completes the proof that $\mathcal{H}$ is a connection.

5. Finally we prove that a section c along a curve γ is parallel in the sense of Remark 29.9 if and only if c is horizontal with respect to $\mathcal{H}$ in the sense of Definition 29.1. One direction is clear by definition of $\mathcal{H}$, so it suffices to show that if γ is a smooth curve and $c \in \Gamma_\gamma(E)$ is horizontal along γ then c is also parallel. Let $p = c(0)$ and let $c_1(t) := \mathbb{P}_\gamma(p)$. Since both c and c_1 are horizontal and

$$D\pi(c_1(t))[c_1'(t)] = \frac{d}{dt} \pi \circ c_1(t) = \gamma'(t) = D\pi(c(t))[c'(t)],$$

we have by the defining condition of a preconnection that $c'(t) = c_1'(t)$. Thus c and c_1 are two curves with the same initial condition and the same derivative, whence they are equal. This at last completes the proof of Theorem 30.1. ∎

We now get started on the proof of the opposite direction: how to go from a connection to a parallel transport system.

Theorem 30.2. *Let $\pi\colon E \to M$ be a vector bundle, and let $\mathcal{H}$ be a connection on E. The system of all horizontal lifts to E of smooth curves in M defines a parallel transport system $\mathbb{P}$ in E. Moreover the connection on E determined by $\mathbb{P}$ from Theorem 30.1 is just $\mathcal{H}$ again.*

Proof. As the statement of the theorem indicated, given a smooth curve $\gamma\colon [a,b] \to M$ and $p \in E_{\gamma(a)}$, we define $\mathbb{P}_\gamma(p) \in \Gamma_\gamma(E)$ to be the horizontal lift of γ with respect to $\mathcal{H}$, whose existence and uniqueness is guaranteed by Proposition 29.7. We must check that the four axioms of a parallel transport system are satisfied. We will do this in three steps.

1. In this step we check that our proposed parallel transport system satisfies Axiom (i) from Definition 29.8. Let $\gamma\colon [a,b] \to M$ be a smooth curve. Set $x = \gamma(a)$ and $y = \gamma(b)$. If c is a horizontal lift of γ to E then $a\,c$ is also horizontal since

$$\frac{d}{dt}(ac)(t) = D\mu_a(c(t))[c'(t)] \in \mathcal{H}_{ac(t)}$$

by (28.5). Since $(a\,c)(0) = a\,c(0)$, this shows that the map $\widehat{\mathbb{P}}_\gamma\colon E_x \to E_y$ is homogeneous (i.e. $\widehat{\mathbb{P}}_\gamma(a\,p) = a\widehat{\mathbb{P}}_\gamma(p)$ for $p \in E_x$ and $a \in \mathbb{R}$). Moreover it follows from the proof of Proposition 29.7 and the smooth dependence on initial conditions of integral curves (Theorem 8.1 applied to $\overline{T}$) that $\widehat{\mathbb{P}}_\gamma$ is differentiable as a map from the vector space E_x to the vector space E_y.

Note $\widehat{\mathbb{P}}_\gamma(0_x) = 0_y$ by homogeneity. If $p \in E_x$ then

$$\begin{aligned} D\widehat{\mathbb{P}}_\gamma(0_x)[\mathcal{J}_{0_x}(p)] &= \frac{d}{dt}\Big|_{t=0} \widehat{\mathbb{P}}_\gamma(tp) \\ &= \lim_{t\to 0} \frac{\widehat{\mathbb{P}}_\gamma(tp)}{t} \\ &= \lim_{t\to 0} \frac{t\widehat{\mathbb{P}}_\gamma(p)}{t} \\ &= \widehat{\mathbb{P}}_\gamma(p) \end{aligned}$$

Thus $D\widehat{\mathbb{P}}_\gamma(0_x)[\mathcal{J}_{0_x}(p)] = \widehat{\mathbb{P}}_\gamma(p)$, and this, coupled with homogeneity, implies that $\widehat{\mathbb{P}}_\gamma$ is linear[2].

Next, if $\gamma^-(t) := \gamma(1-t)$ is the reverse curve from y to x then $c^-(t) := c(1-t)$ is a horizontal section along γ^- with initial condition $c^-(1)$. It follows that $\widehat{\mathbb{P}}_\gamma$ is invertible, with inverse $\widehat{\mathbb{P}}_{\gamma^-}$. This proves that Axiom (i) from Definition 29.8 holds.

2. Let us now verify Axiom (ii) from Definition 29.8. Let $\gamma\colon [a,b] \to M$ be a smooth curve and $h\colon [a_1,b_1] \to [a,b]$ is a diffeomorphism such that $h(a_1) = a$ and $h(b_1) = b$. Set $\delta := \gamma \circ h$. Fix $p \in E_{\gamma(a)}$. Let c be the horizontal section of E along γ with $c(0) = p$ and let d be the horizontal section along δ such that $d(a_1) = p$. We claim that $d = c \circ h$. Indeed, $c \circ h$ is certainly a lift of δ (as $\pi \circ c \circ h = \gamma \circ h = \delta$) and

$$\frac{d}{dt}c(h(t)) = h'(t)c'(h(t)) \in \mathcal{H}_{c(h(t))}$$

by the chain rule. Thus by the uniqueness part of Proposition 29.7, we have $d = c\circ h$ as desired.

3. We now address the final two axioms, Axiom (iii) and Axiom (iv) from Definition 29.8. We will not say much about Axiom (iii) (given that we relegated

[2] Exercise: Verify that any continuous map $f\colon V \to W$ between two vector spaces which is differentiable at $0 \in V$ and homogeneous in the sense that $f(av) = af(v)$ for all $v \in V$ and $a \neq 0$ is necessarily a linear map.

the precise statement of this Axiom to a footnote!), other than that it essentially boils once again down to the fact that integral curves depend smoothly on initial conditions. Axiom (iv) is immediate, since if γ is a smooth curve in M, $p \in E_{\gamma(0)}$ and c is the horizontal section of E along γ with initial condition p then $c'(0)$ is the unique element of $\mathcal{H}_p$ which is mapped to $\gamma'(0)$ by $D\pi(p)$.

Thus $\mathbb{P}$ is indeed a parallel transport system. To complete the proof we must show that the connection obtained from $\mathbb{P}$ by applying Theorem 30.1 is simply $\mathcal{H}$ again. This however is immediate from Axiom (iv) of Definition 29.8. ∎

Remark 30.3. From now on we will usually work with connections, rather than parallel transport systems (this is mainly out of personal preference). Thus if a connection is specified and we refer to a section being "parallel", it should always be implicitly assumed that the parallel transport system in question is the one associated via Theorem 30.1 to the given connection.

This convention has the somewhat amusing consequence that the words "parallel" and "horizontal" can now often be used interchangeably. In general I will (usually) favour the word "parallel" when talking about sections, and "horizontal" when talking about vectors.

LECTURE 31

The connection map and covariant derivatives

In this lecture we introduce the *connection map* of a connection, and use this to define *covariant derivative* operators. First recall that if E is a vector bundle we denote by $\mathrm{pr}_2 \colon VE \to E$ the map $\mathrm{pr}_2(\mathcal{J}_p(q)) = q$ (cf. (30.1)).

DEFINITION 31.1. Let $\pi \colon E \to M$ be a vector bundle and let $\mathcal{H}$ be a connection on E. Define a map

$$\kappa \colon TE \to E, \qquad \kappa(\zeta) := \mathrm{pr}_2(\zeta^{\mathtt{V}}) = \mathrm{pr}_2(\zeta - \zeta^{\mathtt{H}}).$$

This makes sense, since $\zeta^{\mathtt{V}} \in VE$. We call κ the **connection map** of the connection $\mathcal{H}$.

REMARK 31.2. We can use the connection map κ and the parallel transport system $\mathbb{P}$ associated to $\mathcal{H}$ to give a new way to express the horizontal-vertical splitting of a tangent vector. Indeed, if $p \in E$ and $\zeta \in T_pE$ and $\gamma \colon [0,1] \to M$ is any smooth curve with $\gamma(0) = \pi(p)$ and $D\pi(p)[\zeta] = \gamma'(0)$, then it follows from Theorem 30.1 and Definition 31.1 that

$$\zeta^{\mathtt{H}} = \frac{d}{dt}\bigg|_{t=0} \mathbb{P}_\gamma(p)(t) \qquad \text{and} \qquad \zeta^{\mathtt{V}} = \mathcal{J}_p(\kappa(\zeta)).$$

It is immediate that κ is a vector bundle morphism along π, i.e. that the following commutes:

$$\begin{tikzcd} TE \arrow[r, "\kappa"] \arrow[d, "\pi_{TE}"'] & E \arrow[d, "\pi"] \\ E \arrow[r, "\pi"'] & M \end{tikzcd} \tag{31.1}$$

In fact, if we combine κ with $D\pi$ we can build a vector bundle isomorphism along π:

LEMMA 31.3. *Let $\pi \colon E \to M$ be a vector bundle and let $\mathcal{H}$ be a connection on E with connection map κ. Then $(D\pi, \kappa)$ is a vector bundle isomorphism along π:*

$$\begin{tikzcd} TE \arrow[r, "(D\pi{,}\kappa)"] \arrow[d, "\pi_{TE}"'] & TM \oplus E \arrow[d, "(\pi_{TM}{,}\pi)"] \\ E \arrow[r, "\pi"'] & M \end{tikzcd}$$

Proof. Since TE and $TM \oplus E$ have the same fibre dimension, it suffices to check that $\ker(D\pi, \kappa) = 0$. This is immediate from the definition of a (pre)connection, i.e. (28.4). ∎

Far less obviously, κ is also a vector bundle morphism from TE to E along $\pi \colon TM \to M$. Before proving this, let us recall[1] from the solution to part (iii) of Problem I.5 how to see that $D\pi \colon TE \to TM$ is also a vector bundle.

LEMMA 31.4. *Let $\pi \colon E \to M$ be a vector bundle of rank k. Then $D\pi \colon TE \to TM$ is a vector bundle of rank $2k$.*

Proof. If[2] $(x, v) \in TM$ then the fibre over (x, v) in TE are those pairs (p, ζ) where $\pi(p) = x$ and $D\pi(p)[\zeta] = v$. Let us now endow each fibre with a vector space structure. For this let $A \colon E \oplus E \to E$ denote the vector bundle homomorphism

$$A \colon E \oplus E \to E, \qquad A(p, q) = p + q \tag{31.2}$$

given by fibrewise addition. Then if (p, ζ) and (q, ξ) belong to the same fibre over (x, v) we define[3]

$$(p, \zeta) \boxplus (q, \xi) := \big(p + q, DA(p, q)[\zeta, \xi]\big). \tag{31.3}$$

Similarly if $a \in \mathbb{R}$ then we define

$$a \boxtimes (p, \zeta) := \big(ap, D\mu_a(p)[\zeta]\big) \tag{31.4}$$

where μ_a is the fibrewise scalar multiplication as in Definition 28.3. I will leave it up to you to check that if $\alpha \colon \pi^{-1}(U) \to \mathbb{R}^k$ is a vector bundle chart on E then the bundle $D\alpha \colon T(\pi^{-1}(U)) \to T\mathbb{R}^k = \mathbb{R}^{2k}$ is a linear isomorphism on each fibre, and hence may serve as a vector bundle chart. The result now follows from Proposition 13.16. ∎

The following lemma will appear on Problem Sheet O.

LEMMA 31.5. *Let $\pi \colon E \to M$ be a vector bundle of rank k with connection $\mathcal{H}$. Fix $x \in M$ and let $\{p_1, \dots, p_k\}$ be a basis of E_x. Let $\gamma \colon [a, b] \to M$ be a smooth curve with $\gamma(a) = x$. There exists a local frame $\{e_1, \dots, e_k\}$ of E over an open set U containing x such that $e_i(x) = p_i$ and such that $e_i \circ \gamma$ is a parallel along γ for each $i = 1, \dots, k$.*

We call $\{e_1, \dots, e_k\}$ a **parallel local frame along** γ. If $c \in \Gamma_\gamma(E)$ is any section along γ then we can write

$$c(t) = f^i(t)\, e_i(\gamma(t))$$

for some smooth functions $f^i(t)$. We claim:

[1] Here "recall" means: I am pretty certain most of you didn't do this problem… And even those of you who did have surely forgotten it by now, since it was non-examinable.

[2] Let me remind you that I (somewhat inconsistently) write points in TM either as pair (x, v) or sometimes simply as a single element v. The rule is roughly: if it's useful to explicitly say which fibre T_xM a vector v belongs to, I write (x, v). If it's not important, I just write v.

[3] We use the special notation $\boxplus$ and $\boxtimes$ in an attempt to minimise confusion later on.

Lemma 31.6. *Let $\pi\colon E \to M$ be a vector bundle of rank k with connection $\mathcal{H}$. Let γ be a curve in M with $\gamma(0) = x$, and let $\{e_1, \dots, e_k\}$ be a parallel local frame along γ. Fix $c \in \Gamma_\gamma(E)$ and write $c(t) = f^i(t)\, e_i(\gamma(t))$ as above. Then c is parallel along γ if and only if each f^i is a constant function.*

Proof. Set $p = c(0)$. Then c is parallel if and only if $c = \mathbb{P}_\gamma(p)$. If $p_i := e_i(\gamma(0))$ then we can write $p = a^i p_i$ for constants a^i, and $\mathbb{P}_\gamma(p)(t) = a^i \mathbb{P}_\gamma(p_i)(t) = a^i e_i(\gamma(t))$ (cf. Remark 29.10). ∎

Theorem 31.7. *Let $\pi\colon E \to M$ be a vector bundle and let $\mathcal{H}$ be a connection on E with connection map $\kappa\colon TE \to E$. Then*

$$\kappa(\zeta \boxplus \xi) = \kappa(\zeta) + \kappa(\xi), \qquad \kappa(a \boxtimes \zeta) = a\kappa(\zeta), \qquad \zeta, \xi \in TE,\ a \in \mathbb{R},$$

and hence κ is a vector bundle morphism along π_{TM}:

$$\begin{array}{ccc} TE & \xrightarrow{\kappa} & E \\ {\scriptstyle D\pi}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ TM & \xrightarrow[\pi_{TM}]{} & M \end{array}$$

This somewhat innocuous looking result is actually the lynchpin needed to define covariant derivatives, as we will see in the proof of Theorem 31.10 below. This proof is non-examinable.

(♣) *Proof.* Fix $(x, v) \in TM$, and let $\gamma\colon [0,1] \to M$ be a smooth curve with $\gamma(0) = x$ and $\gamma'(0) = v$. Let $(p, \zeta) \in TE$ belong to the fibre of TE over (x, v), so that $p \in E_x$ and $D\pi(p)[\zeta] = v$. Write

$$\zeta = \zeta^{\mathtt{H}} + \zeta^{\mathtt{V}}.$$

By Remark 31.2 we can write this as

$$\zeta^{\mathtt{H}} = \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma(p)(t) \qquad \text{and} \qquad \zeta^{\mathtt{V}} = \mathcal{J}_p(\kappa(\zeta)).$$

Now let $\{e_1, \dots, e_k\}$ be a parallel local frame along γ, and let α denote the corresponding vector bundle chart (cf. Remark (16.3)). Then we compute

$$\begin{aligned} D(\pi, \alpha)(p)[\zeta] &= \big(v, D\alpha(p)[\mathcal{J}_p(\kappa(\zeta))]\big) \\ &= \left(v, \frac{d}{dt}\Big|_{t=0} \alpha\big(p + t\kappa(\zeta)\big)\right) \\ &\overset{(\dagger)}{=} \frac{d}{dt}\Big|_{t=0} (\pi, \alpha)\Big(\mathbb{P}_\gamma(p)(t) + t\,\mathbb{P}_\gamma(\kappa(\zeta))(t)\Big) \\ &= D(\pi, \alpha)(p)\left[\frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma\big(p + t\kappa(\zeta)\big)(t)\right] \end{aligned}$$

where (†) used that $\pi \circ \mathbb{P}_\gamma(p) = \gamma$ and that α is constant along $\mathbb{P}_\gamma(p + t\kappa(\zeta))$ by Lemma 31.6. Since (π, α) is a diffeomorphism by definition of a vector bundle chart, this gives us the formula

$$\zeta = \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma\big(p + t\kappa(\zeta)\big)(t). \tag{31.5}$$

Now suppose (q, ξ) is another point in the fibre over (x, v). Then we can also write

$$\xi = \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma\big(q + t\kappa(\xi)\big)(t).$$

So using the definition of addition in TE, we have:

$$\begin{aligned}
\zeta \boxplus \xi &:= DA(p,q)[\zeta, \xi] \\
&= \frac{d}{dt}\Big|_{t=0} \Big(\mathbb{P}_\gamma\big(p + t\kappa(\zeta)\big)(t) + \mathbb{P}_\gamma\big(q + t\kappa(\xi)\big)(t)\Big) \\
&= \frac{d}{dt}\Big|_{t=0} \Big(\mathbb{P}_\gamma(p+q)(t) + t\,\mathbb{P}_\gamma\big(\kappa(\zeta) + \kappa(\xi)\big)(t)\Big) \\
&= \frac{d}{dt}\Big|_{t=0} \mathbb{P}_\gamma(p+q)(t) + \mathcal{J}_{p+q}\big(\kappa(\zeta) + \kappa(\xi)\big) \\
&= (\zeta + \xi)^{\mathtt{H}} + \mathcal{J}_{p+q}\big(\kappa(\zeta) + \kappa(\xi)\big),
\end{aligned}$$

where the last line used Remark 31.2. But Remark 31.2 also tells us that the vertical component of $\zeta \boxplus \xi$ is $\mathcal{J}_{p+q}\big(\kappa(\zeta \boxplus \xi)\big)$. Comparing this to the expression above and using the fact that $\mathcal{J}_{p+q}$ is an isomorphism, we see that

$$\kappa(\zeta \boxplus \xi) = \kappa(\zeta) + \kappa(\xi).$$

The proof that $\kappa(a \boxtimes \zeta) = a\kappa(\zeta)$ goes along similar lines, and is left as an exercise. ■

We now use the connection map to give a third interpretation of connections, via *covariant derivatives*. This point of view is the "usual" one, and many introductory accounts of connections only define them this way.

Definition 31.8. Let $\pi\colon E \to N$ be a vector bundle and let $\varphi\colon M \to N$ be a smooth map. An operator

$$\nabla\colon \mathfrak{X}(M) \times \Gamma_\varphi(E) \to \Gamma_\varphi(E),$$

written

$$(X, s) \mapsto \nabla_X(s)$$

is called a **covariant derivative operator in E along φ** if the following four conditions are satisfied for any $X, Y \in \mathfrak{X}(M)$, $s_1, s_2 \in \Gamma_\varphi(E)$, and $f \in C^\infty(M)$:

(i) $\nabla_{X+Y}(s) = \nabla_X(s) + \nabla_Y(s)$,

(ii) $\nabla_{fX}(s) = f\nabla_X(s)$,

(iii) $\nabla_X(s_1 + s_2) = \nabla_X(s_1) + \nabla_X(s_2)$,

(iv) $\nabla_X(fs) = X(f)s + f\nabla_X(s)$.

We call $\nabla_X(s)$ the **covariant derivative of s with respect to X**. If $M = N$ and $\varphi = \mathrm{id}$ then we call ∇ a **covariant derivative operator on E**.

Remark 31.9. Let $\pi\colon E \to N$ be a vector bundle and let $\varphi\colon M \to N$ be a smooth map. Suppose ∇ is a covariant derivative operator in E along φ. Let $s \in \Gamma_\varphi(E)$. By property (ii) the operator $X \mapsto \nabla_X(s)$ is $C^\infty(M)$-linear, and hence defines an element

$$\nabla s \in \operatorname{Hom}(\mathfrak{X}(M), \Gamma_\varphi(E)).$$

We call ∇s the **covariant differential of** s. The Hom-Γ Theorem 16.30 tells us that we can also think of ∇s as an element of $\Gamma\big(\operatorname{Hom}(TM, \varphi^\star E)\big)$. It thus follows that the value of $\nabla_X(s)(x)$ only depends on $X(x)$, and hence $\nabla_v(s)$ is well defined for $v \in TM$. Put differently, ∇s is a point operator. Note however that $s \mapsto \nabla_X(s)$ is *not* $C^\infty(M)$-linear, and thus $\nabla_X(p)$ is *not* well defined for $p \in E$, i.e. ∇_X is *not* a point operator. Compare Problem I.1.

Here then is the main result that links connections and covariant derivative operators. Just as with connections and parallel transport operators, the proof is quite involved, and we split it into two stages.

Theorem 31.10. *Let $\pi\colon E \to N$ be a vector bundle and let $\mathcal{H}$ be a connection on E with connection map κ. If $\varphi\colon M \to N$ is any smooth map then*

$$\nabla_X(s)(x) := \kappa(Ds(x)[X(x)]) \tag{31.6}$$

defines a covariant derivative operator in E along φ. This covariant derivative operator has the property that a section $s \in \Gamma_\varphi(E)$ is parallel if and only if $\nabla_X(s) = 0$ for all $X \in \mathfrak{X}(M)$. Moreover the chain rule holds: if $\psi\colon L \to M$ is a smooth map then

$$\nabla_w(s \circ \psi)(y) = \nabla_{D\psi(y)[w]}(s)(\psi(y)), \qquad y \in L,\ w \in T_yL \tag{31.7}$$

Remark 31.11. If $\psi\colon L \to M$ is actually a diffeomorphism then (31.7) can be written as

$$\nabla_Y(s \circ \psi) = \nabla_{\psi_\star(Y)}(s) \circ \psi, \qquad Y \in \mathfrak{X}(L),\ s \in \Gamma_\varphi(E). \tag{31.8}$$

This only makes sense for ψ a diffeomorphism, as otherwise $\psi_\star(Y)$ is not defined!

Proof. The formula (31.6) certainly defines an element of $\Gamma_\varphi(E)$. We show that ∇ really is a covariant derivative operator along φ in four steps.

1. In this step we show that a section $s \in \Gamma_\varphi(E)$ is parallel with respect to $\mathcal{H}$ if and only if $\nabla_X(s) = 0$ for every vector field $X \in \mathfrak{X}(M)$. This is clear, since $\mathcal{H} = \ker\kappa$ and by Definition 29.1 a section s is parallel (or horizontal—cf. Remark 30.3!) if and only if $Ds[TM] \subset \mathcal{H}$.

2. Let us now verify (31.7). Note $s \circ \psi \in \Gamma_{\varphi\circ\psi}(E)$. Fix $y \in L$ and $w \in T_yL$. Then

$$\begin{aligned}\nabla_w(s \circ \psi)(y) &= \kappa\big(D(s\circ\psi)(y)[w]\big)\\ &= \kappa\Big(Ds(\psi(y))[D\psi(y)[w]]\Big)\\ &= \nabla_{D\psi(y)[w]}(s)(\psi(y)).\end{aligned}$$

3. Let $s_1, s_2 \in \Gamma_\varphi(E)$ and $f \in C^\infty(M)$. Fix $x \in M$, $v \in T_xM$. In this step we show that

$$\nabla_v(s_1 + s_2)(x) = \nabla_v(s_1)(x) + \nabla_v(s_2)(x). \tag{31.9}$$

Let γ be a curve in M with $\gamma(0) = x$ and $\gamma'(0) = v$. Then with A as in (31.2) we have from (31.3) that

$$\begin{aligned} Ds_1(x)[v] + Ds_2(x)[v] &= DA(s_1(x), s_2(x))\big[Ds_1(x)[v], Ds_2(x)[v]\big] \\ &= \frac{d}{dt}\Big|_{t=0} s_1(\gamma(t)) + s_2(\gamma(t)) \\ &= \frac{d}{dt}\Big|_{t=0} (s_1 + s_2)(\gamma(t)) \\ &= D(s_1 + s_2)(x)[v]. \end{aligned}$$

Since κ is a vector bundle morphism along π_{TM} by Theorem 31.7, we obtain

$$\begin{aligned} \nabla_v(s_1 + s_2)(x) &= \kappa\big(D(s_1 + s_2)(x)[v]\big) \\ &= \kappa\big(Ds_1(x)[v] + Ds_2(x)[v]\big) \\ &= \kappa(Ds_1[v]) + \kappa(Ds_2(x)[v]) \\ &= \nabla_v(s_1)(x) + \nabla_v(s_2)(x). \end{aligned}$$

This proves (31.9).

4. Let $s \in \Gamma_\varphi(E)$ and $f \in C^\infty(M)$. Fix $x \in M$ and $v \in T_xM$. In this step we prove that

$$\nabla_v(fs)(x) = v(f)s(x) + f(x)\nabla_v(s)(x). \tag{31.10}$$

For this let

$$\mu\colon \mathbb{R} \times E \to E, \qquad \mu(a, p) := \mu_a(p) = ap.$$

Then

$$D\mu(a, p)\left[r\frac{\partial}{\partial t}\Big|_a, v\right] = D\mu_a(p)[v] + \mathcal{J}_{ap}(rp). \tag{31.11}$$

The section $x \mapsto f(x)s(x)$ can be written as the composition $\mu \circ (f, s)$, and hence using (31.11) we compute

$$\begin{aligned} D(fs)(x)[v] &= D\big(\mu \circ (f, s)\big)(x)[v] \\ &= D\mu(f(x), s(x)) \circ \big(Df(x)[v], Ds(x)[v]\big) \\ &= D\mu_{f(x)}(s(x))[Ds(x)[v]] + \mathcal{J}_{f(x)s(x)}(Df(x)[v]s(x)) \\ &= D\mu_{f(x)}(s(x))[Ds(x)[v]] + \mathcal{J}_{f(x)s(x)}(v(f)s(x)). \end{aligned}$$

Now by definition

$$D\mu_{f(x)}(s(x))[Ds(x)] = f(x) \boxtimes Ds(x).$$

Thus applying κ to both sides and using Theorem 31.7 we obtain

$$\kappa\big(D(fs)(x)[v]\big) = f(x)\kappa(Ds(x)[v]) + v(f)s(x),$$

which gives (31.10). This completes the proof. ∎

Corollary 31.12. *Let*[4] $\pi\colon E \to N$ *be a vector bundle with connection* $\mathcal{H}$, *and let* $\varphi\colon M \to N$ *be smooth. If* $s_1, s_2 \in \Gamma_\varphi(M)$ *are horizontal then so is* $c\, s_1 + s_2$ *for any* $c \in \mathbb{R}$.

Proof. For any vector field X on M,

$$\nabla_X(c\, s_1 + s_2) = c\nabla_X(s_1) + \nabla_X(s_2) = 0.$$

∎

[4]Thank you to "Miao" for reminding me of this!

LECTURE 32

Holonomy

We begin this lecture by completing the various chain of equivalences and proving that a covariant derivative operator uniquely determines a connection. We then introduce the notion of *holonomy*, which will give us a way to measure how "non-trivial" a connection is.

THEOREM 32.1. *Let ∇ be a covariant derivative operator on a vector bundle $\pi\colon E \to M$. Then there exists a connection $\mathcal{H}$ on E such that if $s \in \Gamma(E)$ and $(x,v) \in TM$ then $Ds(x)[v] \in \mathcal{H}_{s(x)}$ if and only if $\nabla_v(s)(x) = 0$.*

Proof. Given $p \in E_x$, we define

$$\mathcal{H}_p := \left\{ Ds(x)[v] - \mathcal{J}_p\big(\nabla_v(s)(x)\big) \mid \text{all } s \in \Gamma(E) \text{ such that } s(x) = p \text{ and all } v \in T_xM \right\},$$

and then we set $\mathcal{H} = \bigsqcup_{p\in E} \mathcal{H}_p$. In contrast to the proof of Theorem 30.1, this time it is clear that $\mathcal{H}_p$ is a linear subspace of T_pE. Moreover $D\pi(p)|_{\mathcal{H}_p}\colon \mathcal{H}_p \to T_{\pi(p)}M$ is a linear isomorphism by construction. The proof that $\mathcal{H}$ really is a vector subbundle goes along exactly the same lines as the proof of Step 3 of Theorem 30.1: If σ is a chart on M with local coordinates x^i, then if we set $\alpha := dx^i\, e_i$ then $\alpha \circ D\pi$ is a vector bundle chart on $\mathcal{H}$ that can be extended to a vector bundle chart on TE. Thus $\mathcal{H}$ is a preconnection. Finally if $a \in \mathbb{R}$ then

$$D(as)(x)[v] - \mathcal{J}_{ap}\big(\nabla_v(as)(x)\big) = D\mu_a(p)\Big(Ds(x)[v] - \mathcal{J}_p\big(\nabla_v(s)(x)\big)\Big),$$

and hence $\mathcal{H}$ is a connection. ■

REMARK 32.2. **Important convention:** Building on Remark 30.3, since we now know that connections, parallel transport systems and covariant derivative operators are really three different ways of expressing the same concept, we will refer to all of them as a "connection" and use the symbol[1] ∇. Thus for instance if $\pi\colon E \to N$ is a vector bundle and ∇ is a connection on E then we'll write $\varphi^\star(\nabla)$ for the connection on $\varphi^\star E$ from Proposition 28.10.

Next, we finally make rigorous the discussion from the beginning of Lecture 28 when we initially motivated the definition of a connection.

PROPOSITION 32.3. *Let $\pi\colon E \to N$ be a vector bundle with connection ∇. Let $\varphi\colon M \to N$ be a smooth map. Let $\gamma\colon [0,1] \to M$ be a smooth curve and abbreviate by*

$$\widehat{\mathbb{P}}_t\colon E_{\varphi(\gamma(0))} \to E_{\varphi(\gamma(t))}$$

the parallel transport along the curve $r \mapsto \varphi(\gamma(r))$ for $0 \le r \le t$. Then if $s \in \Gamma_\varphi(E)$ one has

$$\nabla_{\gamma'(0)}(s)(\gamma(0)) = \mathcal{J}^{-1}_{\varphi(\gamma(0))}\left(\frac{d}{dt}\Big|_{t=0} \widehat{\mathbb{P}}_t^{-1}(s(\gamma(t)))\right) \tag{32.1}$$

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] Yes, I know it makes no sense to mix words and notation like this. But it is consistent with the majority of the literature. Also, ∇ is easier to type than $\mathcal{H}$...

Proof. Let $\{e_i\}$ be a parallel local frame along $\varphi\circ\gamma$. We can write $s\circ\gamma = f^i(e_i\circ\varphi\circ\gamma)$ for smooth functions f^i. Then

$$\widehat{\mathbb{P}}_t^{-1}(s(\gamma(t)) = \widehat{\mathbb{P}}_t^{-1}\big(f^i(t)e_i(\varphi(\gamma(t)))\big) = f^i(t)e_i(\varphi(\gamma(0))). \tag{32.2}$$

Let T denote the vector field $\frac{\partial}{\partial t}$ on $[0,1]$. Then we have

$$\begin{aligned}\nabla_{\gamma'(0)}(s)(\gamma(0)) &\overset{(\dagger)}{=} \nabla_{T(0)}(s\circ\gamma)(0)\\ &= \nabla_{T(0)}\big(f^i(e_i\circ\varphi\circ\gamma)\big)(0)\\ &\overset{(\ddagger)}{=} (f^i)'(0)e_i(\varphi(\gamma(0)))\\ &\overset{(\heartsuit)}{=} \mathcal{J}^{-1}_{\varphi(\gamma(0))}\left(\frac{d}{dt}\Big|_{t=0}\widehat{\mathbb{P}}_t^{-1}(s(\gamma(t)))\right)\end{aligned}$$

where $(\dagger)$ used the chain rule (31.8), $(\ddagger)$ used (31.10), and $(\heartsuit)$ used (32.2). ■

Remark 32.4. The equation (32.1) shows how parallel transport allows us to make sense of (28.2). Indeed, if $\mathbb{P}$ is the trivial parallel transport system from Example 29.11 then this defines exactly what we called "the trivial connection" in Definition 28.1.

Remark 32.5. The proof of Proposition 32.3 used that we already knew that the parallel transport system $\mathbb{P}$ determined a covariant derivative operator ∇—we merely had to identify it. However a minor modification of the argument would allow us to *define* ∇ via (32.1). This would allow us to go directly from a parallel transport system to a covariant derivative operator and bypass connections entirely. Many introductory treatments of Differential Geometry do this. We will see one concrete advantage of why having the connection definition on hand is useful next lecture (Theorem 33.4).

We now move onto studying the *holonomy* of a connection. In the following, we will have cause to work with **piecewise smooth** curves. By definition a piecewise smooth curve $\gamma\colon [a,b]\to M$ in a manifold M is a continuous map γ such that there exist finitely many points $a_0 = a < a_1 < \dots < a_r = b$ such that $\gamma|_{[a_i,a_{i+1}]}\colon [a_i,a_{i+1}]\to M$ is smooth for each $i = 0,\dots r-1$ (thinking of $[a_i,a_{i+1}]$ as a one-dimensional manifold with boundary). The simplest way to manufacture such a curve is simply to glue two smooth curves together:

Example 32.6. Suppose $\gamma\colon [a,b]\to M$ and $\delta\colon [b,c]\to M$ are two smooth curves with $\gamma(b)=\delta(b)$. Then the **concatenation** of γ and δ is the piecewise smooth curve $\gamma*\delta\colon [a,c]\to M$ defined by

$$(\gamma*\delta)(t) := \begin{cases}\gamma(t), & a\le t\le b,\\ \delta(t), & b\le t\le c.\end{cases}$$

Definition 32.7. Let $\pi\colon E\to M$ be a vector bundle with connection ∇. Suppose $\gamma\colon [a,b]\to M$ and $\delta\colon [b,c]\to M$ are two smooth curves with $\gamma(b)=\delta(b)$. We define the **parallel transport**[2] **along the piecewise smooth curve** $\gamma*\delta$ to be the linear isomorphism

$$\widehat{\mathbb{P}}_{\gamma*\delta}\colon E_{\gamma(a)}\to E_{\delta(c)},\qquad \widehat{\mathbb{P}}_{\gamma*\delta} := \widehat{\mathbb{P}}_\delta\circ\widehat{\mathbb{P}}_\gamma.$$

[2]We could also define the non-hat version, but we won't need this.

The same definition works for any piecewise smooth curve; as the composition of finitely many linear isomorphisms, it is again a linear isomorphism.

Remark 32.8. More generally, suppose $\gamma\colon [a,b] \to M$ and $\delta\colon [b_1, c] \to M$ are two smooth curves with $\gamma(b) = \delta(b_1)$ but $b \neq b_1$. Then we cannot directly concatenate γ and δ, and thus we cannot directly define $\widehat{\mathbb{P}}_{\gamma*\delta}$. But this is easily rectified by reparametrising. Indeed, we can choose a diffeomorphism $h\colon [a, b_1] \to [a,b]$ such that $h(a) = a$ and $h(b) = b_1$ and replace γ with $\gamma \circ h$. Then $(\gamma \circ h) * \delta$ is defined. Alternatively, we could reparametrise δ. This reparametrisation will have no effect on the parallel transport thanks to Axiom (ii) from Definition 29.8. From now on we will often suppress the reparametrisation, and speak of the concatentated curve $\gamma * \delta$ and the parallel transport $\widehat{\mathbb{P}}_{\gamma*\delta}$ whenever γ and δ are two curves such that γ ends where δ begins.

Remark 32.9. It follows from Axiom (ii) that parallel transport along piecewise smooth curves is associative:

$$\widehat{\mathbb{P}}_{\gamma*(\delta*\varepsilon)} = \widehat{\mathbb{P}}_{(\gamma*\delta)*\varepsilon}$$

for three curves γ, δ, ε such that γ ends where δ begins, and δ ends where ε begins.

Since the inverse of $\widehat{\mathbb{P}}_{\gamma}$ is $\widehat{\mathbb{P}}_{\gamma^-}$, where γ^- is the reverse path—this is part of Axiom (i), it follows that if we fix a basepoint we get a group.

Definition 32.10. Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Fix $x \in M$. The **holonomy group of** ∇ at x is the subgroup $\mathrm{Hol}^{\nabla}(x) \subset \mathrm{GL}(E_x)$ consisting of all parallel transport maps $\widehat{\mathbb{P}}_{\gamma}\colon E_x \to E_x$ where γ is a piecewise smooth loop at x. We always consider $\mathrm{Hol}^{\nabla}(x)$ as carrying the subspace topology inherited from $\mathrm{GL}(E_x)$.

If the base manifold M is connected[3] then the holonomy group $\mathrm{Hol}^{\nabla}(x)$ is—up to isomorphism—independent of x.

Lemma 32.11. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Assume that M is connected. Fix $x, y \in M$ and let γ denote a piecewise smooth curve from x to y. Then the map*

$$\mathrm{Hol}^{\nabla}(x) \to \mathrm{Hol}^{\nabla}(y), \qquad \widehat{\mathbb{P}}_{\delta} \mapsto \widehat{\mathbb{P}}_{\gamma^- *\delta*\gamma} \tag{32.3}$$

is an isomorphism.

Proof. The map (32.3) is a group homomorphism by associativity of parallel transport, since parallel transport around $\gamma^- * (\delta * \varepsilon) * \gamma$ is the same as parallel transport around $\big(\gamma^- * \delta * \gamma\big) * \big(\gamma^- * \varepsilon * \gamma\big)$. Moreover it is an isomorphism as the inverse homomorphism is given by $\widehat{\mathbb{P}}_{\delta} \mapsto \widehat{\mathbb{P}}_{\gamma*\delta*\gamma^-}$. ∎

It is often useful to think of the holonomy group $\mathrm{Hol}^{\nabla}(x)$ as a subgroup of $\mathrm{GL}(k)$ rather than $\mathrm{GL}(E_x)$. This can be done, provided we only work up to conjugation.

[3] Recall a manifold is connected if and only if it is path connected, cf. part (ii) of Remark 1.9.

Recall the frame bundle $\mathrm{Fr}(E)$ from Definition 24.13. Fix $x \in M$ and $A \in \mathrm{Fr}(E_x)$; thus $A\colon \mathbb{R}^k \to E_x$ is a linear isomorphism. Then

$$\mathrm{Hol}^\nabla(x; A) := \left\{ A^{-1} \circ \widehat{\mathbb{P}}_\gamma \circ A \mid \widehat{\mathbb{P}}_\gamma \in \mathrm{Hol}^\nabla(x) \right\}$$

is a subgroup of $\mathrm{GL}(k)$. If $B \in \mathrm{Fr}(E_x)$ is another frame then the subgroup $\mathrm{Hol}^\nabla(x; B)$ is not equal to $\mathrm{Hol}^\nabla(x; A)$, but it is conjugate to it:

$$\mathrm{Hol}^\nabla(x; B) = \left\{ TST^{-1} \mid S \in \mathrm{Hol}^\nabla(x; A) \right\}$$

where $T := B^{-1} \circ A \in \mathrm{GL}(k)$. Moreover if M is connected then Lemma 32.11 shows that if x and y are two points in M and $A \in \mathrm{Fr}(E_x)$ and $B \in \mathrm{Fr}(E_y)$ then the subgroups $\mathrm{Hol}^\nabla(x; A)$ and $\mathrm{Hol}^\nabla(y; B)$ are also conjugate in $\mathrm{GL}(k)$. This proves:

Corollary 32.12. *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold M with connection ∇. Then for all $x \in M$, the holonomy group $\mathrm{Hol}^\nabla(x)$ can be regarded as a subgroup of $\mathrm{GL}(k)$, defined up to conjugation, and in this sense it is independent of x.*

Our first use of holonomy will be to define what it means for a connection on a connected manifold to be trivial.

Definition 32.13. Let $\pi\colon E \to M$ be a vector bundle over a connected manifold, and let ∇ be a connection on E. We say that ∇ is a **trivial connection** if Hol^∇ is the trivial group.

This definition is consistent with Definition 28.1 and Example 29.11.

Proposition 32.14. *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold, and let ∇ be a connection on E. Then ∇ is trivial if and only if E is a trivial vector bundle and the parallel transport system $\mathbb{P}$ is the trivial parallel transport system from Example 29.11.*

Proof. Suppose Hol^∇ is the trivial group, and fix $x \in M$. Define $\alpha\colon E \to E_x$ by

$$\alpha(p) := \mathbb{P}_\gamma(p)(1)$$

where γ is a smooth path in M from $\pi(p)$ to x. Then α is well-defined because $\mathrm{Hol}^\nabla(x)$ is trivial, and α is smooth by Axiom (iii) of Definition 29.8. By Axiom (i) it follows that α is a parallel[4] vector bundle chart on E, and thus E is the trivial bundle and $\mathbb{P}$ is the trivial parallel transport system.

Conversely, if E is the trivial bundle and $\mathbb{P}$ is the trivial parallel transport system, then if γ is path in M then any parallel section c along γ is of the form $c = s \circ \gamma$, where s is a global parallel section of E. Thus if $\gamma\colon [0,1] \to M$ is a loop then for any parallel c along γ,

$$c(1) = s(\gamma(1)) = s(\gamma(0)) = c(0).$$

This shows $\mathrm{Hol}^\nabla(\gamma(0))$ is the trivial group. ∎

[4]That is, the frame $\{e_i\}$ corresponding to α is a parallel frame.

It is often convenient to restrict to *contractible* loops.

Definition 32.15. Let $\pi\colon E \to M$ be a vector bundle over a connected manifold, and let ∇ be a connection on E. Fix $x \in M$. The **restricted holonomy group** $\mathrm{Hol}_0^\nabla(x)$ is the subgroup of $\mathrm{Hol}^\nabla(x)$ consisting of all parallel transports around contractible (i.e. null-homotopic) piecewise smooth loops at x.

Let $\pi_1(M,x)$ denote the fundamental group of M at x. It follows from the Whitney Approximation Theorem[5] 6.14 that any class $[\gamma] \in \pi_1(M,x)$ can be represented by a smooth map γ (and thus also a piecewise smooth γ).

Proposition 32.16. *The restricted holonomy group $\mathrm{Hol}_0^\nabla(x)$ is a path-connected normal subgroup of $\mathrm{Hol}^\nabla(x)$, and there exists a surjective group homomorphism*

$$\pi_1(M,x) \to \mathrm{Hol}^\nabla(x) \big/ \mathrm{Hol}_0^\nabla(x). \tag{32.4}$$

Proof. Suppose $\gamma\colon [0,1] \to M$ is a contractible piecewise smooth loop based at x. Thus there exists a continuous map $H\colon [0,1]\times[0,1] \to M$ such that $H(0,t) = \gamma(t)$, $H(1,t)$ is the constant loop $e_x(t) := x$ and such that $H(s,\cdot)$ is a piecewise smooth[6] loop based at x for each $s \in [0,1]$. Then $s \mapsto \widehat{\mathbb{P}}_{H(s,\cdot)}$ is a path in $\mathrm{Hol}_0^\nabla(x)$ from $\widehat{\mathbb{P}}_\gamma$ to the $\widehat{\mathbb{P}}_{e_x}$. Thus $\mathrm{Hol}_0^\nabla(x)$ is path-connected.

Next, if δ and γ are any two loops at x such that γ is nullhomotopic, then the concatenation $\delta^- * \gamma * \delta$ is also nullhomotopic. Thus if $\widehat{\mathbb{P}}_\gamma \in \mathrm{Hol}_0^\nabla(x)$ and $\widehat{\mathbb{P}}_\delta \in \mathrm{Hol}^\nabla(x)$ then $\widehat{\mathbb{P}}_\delta \circ \widehat{\mathbb{P}}_\gamma \circ \widehat{\mathbb{P}}_{\delta^-} = \widehat{\mathbb{P}}_{\delta^- * \gamma * \delta}$ belongs to $\mathrm{Hol}_0^\nabla(x)$. This shows that $\mathrm{Hol}_0^\nabla(x)$ is normal in $\mathrm{Hol}^\nabla(x)$.

Finally, the desired homomorphism (32.4) sends $[\gamma]$ to the equivalence class of $\widehat{\mathbb{P}}_\gamma$ in the quotient for γ a smooth representative of $[\gamma]$. This is a well-defined surjective group homomorphism by the argument above. ■

We conclude this lecture with the following important result.

Theorem 32.17. *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold, and let ∇ be a connection on E. Then $\mathrm{Hol}^\nabla(x)$ is a Lie group, and $\mathrm{Hol}_0^\nabla(x)$ is the connected component containing the identity.*

This proof is rather sketchy and is non-examinable.

(♣) *Proof.* A (difficult) theorem, proved independently by Kuranishi and Yamabe[7], says that any path connected subgroup of a Lie group is itself a Lie group. Applying this to $\mathrm{Hol}_0^\nabla(x) \subset \mathrm{GL}(x)$ shows that $\mathrm{Hol}_0^\nabla(x)$ is a Lie group. Since M is connected and second countable, its fundamental group is countable. Thus $\mathrm{Hol}^\nabla(x) \big/ \mathrm{Hol}_0^\nabla(x)$ is countable by Proposition 32.16. This implies that $\mathrm{Hol}^\nabla(x)$ is also a Lie group with $\mathrm{Hol}_0^\nabla(x)$ the connected component containing the identity. ■

[5] Quoting the Whitney Approximation Theorem here is overkill, since it easy to prove this directly.

[6] The fact that $H(s,\cdot)$ can be taken to be piecewise smooth again uses the Whitney Approximation Theorem (see Remark 6.17).

[7] Yamabe's proof is very short, and quite easy to understand. If you are interested, see here.

Curvature

In this lecture we explore what it means to say that a connection $\mathcal{H}$ forms an *integrable* distribution in the sense of Definition 11.11. This will lead us naturally to the concept of the *curvature* of a connection, which roughly speaking measures how far the connection is from being integrable.

DEFINITION 33.1. Let $\pi\colon E \to M$ be a vector bundle with connection ∇. We say that ∇ is a **flat connection** if the corresponding distribution $\mathcal{H}$ of E is integrable. The pair (E, ∇) is referred to as a **flat vector bundle**.

Trivial connections are always flat. To see this, let us first give an alternative criterion for a connection to be trivial.

LEMMA 33.2. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Then ∇ is the trivial connection if and only if for every point $p \in E$ there exist a global parallel section $s \in \Gamma(E)$ such that $s(\pi(p)) = p$.*

Proof. This is just a rephrasing of the last part of Proposition 32.14. It is clear that the trivial connection on the trivial vector bundle has this property. Meanwhile if such a section exists through every point then the argument in the last paragraph of the proof of Proposition 32.14 shows that the holonomy groups are trivial, whence Proposition 32.14 itself then shows that ∇ is the trivial connection. ∎

Why is this relevant? If $s \in \Gamma(E)$ is a global parallel section then $s(M) \subset E$ is an embedded submanifold of E (Lemma 13.4) with

$$D\imath_{s(x)}[T_{s(x)}s(M)] = \mathcal{H}_{s(x)}, \qquad \forall\, x \in M.$$

Thus $s(M)$ is an integral manifold for the distribution $\mathcal{H}$ passing through $s(x)$ in the sense of Definition 11.6. Therefore applying the easy half (Proposition 11.14) of the Frobenius Theorem, we obtain:

COROLLARY 33.3. *The trivial connection is flat.*

The converse is true locally. This uses the hard direction of the Frobenius Theorem.

THEOREM 33.4. *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold, and suppose ∇ is a flat connection on E. Then ∇ is a locally trivial connection and $\mathrm{Hol}_0^{\nabla}(x)$ is the trivial group for all $x \in M$.*

Here by a "locally trivial connection" we meant that every point $x \in M$ has a neighbourhood U such that the restriction of ∇ to the trivial subbundle $\pi^{-1}(U) \to U$ of E is the trivial connection. The proof uses a little bit of algebraic topology, and therefore is non-examinable.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

(♣) *Proof.* We prove the result in two steps.

1. In this step we show that ∇ is locally trivial. By the Global Frobenius Theorem 11.18, $\mathcal{H}$ induces a foliation of E. Let L be a leaf of the foliation, i.e. a maximal connected integral manifold of the distribution $\mathcal{H}$ corresponding to ∇. We claim that $\pi|_L \colon L \to M$ is surjective. Indeed, given $p \in L$ and $x \in M$, let $\gamma \colon [0,1] \to M$ be a smooth curve such that $\gamma(0) = \pi(p)$ and $\gamma(1) = x$. The section $\mathbb{P}_\gamma(p)$ is horizontal and thus has image contained in L. Since $\pi(\mathbb{P}_\gamma(p)(1)) = x$, this shows that $\pi|_L$ is surjective.

Since π is a submersion, the Inverse Function Theorem 5.2 tells us that $\pi|_L$ is a local diffeomorphism from L to M. Let $U \subset M$ be a connected and simply connected open subset over which E is trivial. Then the intersection $L \cap \pi^{-1}(U)$ is a disjoint union of connected embedded submanifolds of L such that for each component L_h, $\pi|_{L_h} \colon L_h \to U$ is a diffeomorphism. Thus $s_h := \pi|_{L_h}^{-1} \colon U \to L_h$ is a section of the vector subbundle $\pi^{-1}(U) \to U$. Since L_h is an integral submanifold of $\mathcal{H}|_{\pi^{-1}(U)}$, s_h is a parallel section. Thus for every point of $L \cap \pi^{-1}(U)$ there is a parallel section of $\pi^{-1}(U)$. By Lemma 33.2, the restriction of ∇ to $\pi^{-1}(U)$ is the trivial connection.

2. In this step we show that the restricted holonomy groups are always trivial. Fix a point $x \in M$, and let $\gamma \colon [0,1] \to M$ be a contractible piecewise smooth loop at x. Then as in the proof of Proposition 32.16, there exists a continuous map $H \colon [0,1] \times [0,1] \to M$ such that $H(0,t) = \gamma(t)$, $H(1,t)$ is the constant loop $e_x(t) := x$ and such that $H(s,\cdot)$ is a piecewise smooth contractible loop based at x for each $s \in [0,1]$. Fix $p \in E_x$, and let L be the maximal integral manifold of ∇ passing through p. Then as in the previous step, each section $\mathbb{P}_{H(s,\cdot)}(p)$ has image contained in L. Consider the map

$$\tilde{H} \colon [0,1] \times [0,1] \to L, \qquad \tilde{H}(s,t) := \mathbb{P}_{H(s,\cdot)}(p)(t).$$

This map is a lift of H to L in the sense that

$$\pi(\tilde{H}(s,t)) = H(s,t).$$

Since $H(s,1)$ is independent of s, so[1] is $\tilde{H}(s,1)$. Thus

$$\mathbb{P}_\gamma(p)(1) = \tilde{H}(0,1) = \tilde{H}(1,1) = \mathbb{P}_{e_x}(p)(1) = p$$

Thus parallel transport around γ is trivial. Since γ was arbitrary, it follows that $\mathrm{Hol}_0^\nabla(x)$ is the trivial group. This completes the proof. ∎

Corollary 33.5. *Let $\pi \colon E \to M$ be a vector bundle over a connected and simply connected manifold M, and let ∇ be a connection on E. Then ∇ is flat if and only if E is the trivial bundle and ∇ is the trivial connection.*

[1]The eagle-eyed readers might worry a bit here: why does $H(s,1)$ being independent of s imply that $\tilde{H}(s,1)$ is also independent of s? This argument is hiding a small amount of algebraic topology, which I am cleverly concealing in a footnote so as not to scare those of you who are unfamiliar with algebraic topology off. Indeed, the argument in Step 1 actually shows that $\pi|_L \colon L \to M$ is a **covering space**. Covering spaces enjoy the **unique homotopy lifting property**. One way to phrase this is as follows: if $\pi \colon Y \to X$ is a covering space and $\gamma, \delta \colon [0,1] \to X$ are two paths in X which are homotopic with fixed endpoints, then if $y \in Y$ is any point in Y such that $\pi(y) = \gamma(0)$ then there are unique lifts $\tilde{\gamma}, \tilde{\delta}$ of γ and δ that $\tilde{\gamma}(0) = \tilde{\delta}(0) = y$, and *moreover* these lifts also satisfy $\tilde{\gamma}(1) = \tilde{\delta}(1)$.

Proof. If M is simply connected then $\mathrm{Hol}^{\nabla}(x) = \mathrm{Hol}_0^{\nabla}(x)$ for all $x \in M$. Thus the claim is immediate from Proposition 32.14 and Theorem 33.4. ■

The Frobenius Theorem tells us that a connection is flat if and only if the vector space $\Gamma(\mathcal{H})$ of horizontal vector fields is a Lie subalgebra of the space $\mathfrak{X}(E) = \Gamma(TE)$ of all vector fields on E. The *curvature* of a connection gives a quantitative way to measure how far a given connection is from being flat.

Definition 33.6. Let $\pi \colon E \to M$ be a vector bundle with connection ∇ and connection map $\kappa \colon TE \to E$. The **curvature tensor** R^{∇} of ∇ is defined as follows. Fix vector fields $X, Y \in \mathfrak{X}(M)$ and $p \in E$. Let $\overline{X}$ and $\overline{Y}$ denote the horizontal lifts of X and Y to E (cf. Definition 28.8) and set

$$R^{\nabla}(X, Y)(p) := -\kappa\big([\overline{X}, \overline{Y}](p)\big). \tag{33.1}$$

That is, we take the vertical component of the tangent vector $[\overline{X}, \overline{Y}](p) \in T_pE$, which therefore belongs to V_pE, and then project it to E_p via the map $\mathrm{pr}_2 \colon VE \to E$ (cf. (30.1)).

The minus sign on the right-hand side of (33.1) may look a little unnatural, and indeed some authors define it with the other sign. Our preference for this sign convention will become clear next lecture when we give an alternative way of expressing the curvature (see Theorem 35.1).

Remark 33.7. The meaning of the word "curvature" will become apparent when we study Riemannian Geometry in the second half of the course. We will see that the curvature of a (Riemannian) manifold does indeed correspond to what you would naively guess it does. For example, the sphere S^n with its standard Euclidean metric is "positively" curved.

Since $\kappa(\zeta) = \mathrm{pr}_2(\zeta^{\mathsf{v}})$ and $\mathrm{pr}_2 \colon VE \to E$ is an vector bundle isomorphism along $\pi \colon E \to M$, we have:

Corollary 33.8. *Let $\pi \colon E \to M$ be a vector bundle with connection ∇. Then ∇ is flat if and only if the curvature R^{∇} is identically zero.*

If s is a section of E then the correspondence

$$x \mapsto R^{\nabla}(X, Y)(s(x))$$

defines another section of E, since it satisfies the section property and is smooth (being the composition of smooth maps). We write this section as $R^{\nabla}(X, Y)(s)$. Thus we can think of R^{∇} as defining a map

$$R^{\nabla} \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E).$$

The main result of this lecture proves that this map is a point operator in all three variables.

Theorem 33.9. *Let $\pi \colon E \to M$ be a vector bundle with connection ∇. Then R^{∇} is $C^{\infty}(M)$-linear in all three variables, and antisymmetric in the first two variables. Thus R^{∇} can be thought of as a section of the bundle* $\mathrm{Hom}\big(\bigwedge^2(TM), \mathrm{Hom}(E, E)\big) = \bigwedge^2(T^*M) \otimes E \otimes E^*$.

The proof will use the following lemma, which is a souped-up version of Problem E.2.

LEMMA 33.10. *Let M be a smooth manifold and let X, Y be vector fields on M with local flows θ_t^X and θ_t^Y respectively. Fix $x \in M$ and consider the curve*

$$\gamma\colon [0,\varepsilon) \to M, \qquad \gamma(t) := \theta^Y_{-\sqrt{t}} \circ \theta^X_{-\sqrt{t}} \circ \theta^Y_{\sqrt{t}} \circ \theta^X_{\sqrt{t}}(x),$$

which is well-defined for small enough ε. If $f \in C^\infty(U)$ is a smooth function on a neighbourhood U of x then

$$[X, Y](f)(x) = \lim_{t \to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}.$$

(♣) *Proof.* Let $\delta(t) := \gamma(t^2)$. Then we claim that

(i) $(f \circ \delta)'(0) = 0$,

(ii) $(f \circ \delta)''(0) = 2[X, Y](f)(x)$.

This implies

$$\begin{aligned} [X, Y](f)(x) &= \frac{1}{2}(f \circ \delta)''(0) \\ &= \lim_{t \to 0} \frac{f(\delta(t)) - f(\delta(0))}{t^2} \\ &= \lim_{t \to 0} \frac{f(\delta(\sqrt{t})) - f(\delta(0))}{t} \\ &= \lim_{t \to 0} \frac{(f(\gamma(t)) - f(\gamma(0))}{t}. \end{aligned}$$

To prove (i) and (ii), consider the rectangles

$$\begin{aligned} \Phi_1(s, t) &:= \theta^Y_s \circ \theta^X_t(x) \\ \Phi_2(s, t) &:= \theta^X_{-s} \circ \theta^Y_t \circ \theta^X_t(x) \\ \Phi_3(s, t) &:= \theta^Y_{-s} \circ \theta^X_{-t} \circ \theta^Y_t \circ \theta^X_t(x). \end{aligned}$$

Then $\delta(t) = \Phi_3(t, t)$ and $\Phi_3(0, t) = \Phi_2(t, t)$ and $\Phi_2(0, t) = \Phi_1(t, t)$. Abbreviate

$$\partial_s(f \circ \Phi_3)(0, 0) := D(f \circ \Phi_3)(0, 0)\left[\frac{\partial}{\partial s}\Big|_{s=0}, 0\right]$$

and similarly for the other partial derivatives. Then by the chain rule

$$\begin{aligned} (f \circ \delta)'(0) &= \partial_s(f \circ \Phi_3)(0, 0) + \partial_t(f \circ \Phi_3)(0, 0) \\ &= \partial_s(f \circ \Phi_3)(0, 0) + \partial_s(f \circ \Phi_2)(0, 0) + \partial_t(f \circ \Phi_2)(0, 0) \\ &= \partial_s(f \circ \Phi_3)(0, 0) + \partial_s(f \circ \Phi_2)(0, 0) + \partial_s(f \circ \Phi_1)(0, 0) + \partial_t(f \circ \Phi_1)(0, 0) \\ &= -Y(f)(x) - X(f)(x) + Y(f)(x) + X(f)(x) \\ &= 0. \end{aligned}$$

This proves (i). To prove (ii) we start from

$$(f\circ\delta)''(0) = \partial_{ss}(f\circ\Phi_3)(0,0) + 2\partial_{ts}(f\circ\Phi_3)(0,0) + \partial_{tt}(f\circ\Phi_3)(0,0). \tag{33.2}$$

Since $\partial_s(f\circ\Phi_3) = -Y(f)\circ\Phi_3$, the first term on the right-hand side of (33.2) is equal to

$$\partial_{ss}(f\circ\Phi_3)(0,0) = \partial_s(-Y(f)\circ\Phi_3)(0,0) = Y(Y(f))(x).$$

Similarly since

$$\partial_s(f\circ\Phi_1) = Y(f)\circ\Phi_1, \qquad \partial_s(f\circ\Phi_s) = -X(f)\circ\Phi_2,$$

and

$$\partial_t(f\circ\Phi_1)(0,t) = X(f)\circ\Phi_1(0,t),$$

we obtain

$$2\partial_{ts}(f\circ\Phi_3)(0,0) = -2Y(Y(f)(x))$$

and

$$\partial_{tt}(f\circ\Phi_3)(0,0) = Y(Y(f))(x) + 2[X,Y](f)(x).$$

Substituting these into (33.2) gives

$$(f\circ\delta)''(0) = 2Y(Y(f))(x) - 2Y(Y(f))(x) + 2[X,Y](f)(x),$$

which proves (ii). ∎

Remark 33.11. The curve γ from the statement of Lemma 33.10 is typically not differentiable (not even right differentiable) at 0. Thus strictly speaking, the tangent vector $\gamma'(0)$ is not defined. However if we formally define a tangent vector $\gamma'(0)$ by declaring that

$$\gamma'(0)(f) \stackrel{\text{def}}{=} \lim_{t\to 0}\frac{f(\gamma(t)) - f(\gamma(0))}{t}$$

then $\gamma'(0)$ is a well-defined element of T_xM. In this sense the conclusion of Lemma 33.10 can be restated as

$$[X,Y](x) = \gamma'(0).$$

We will use this convention without comment in the future.

Proof of Theorem 33.9. We prove the result in two steps.

1. By part (iii) of Lemma 28.9, if X,Y,Z are three vector fields on M and $s\in\Gamma(E)$ we have

$$\begin{aligned}[\overline{X+Y},\overline{Z}](s(x))^{\mathsf{v}} &= [\overline{X}+\overline{Y},\overline{Z}](s(x))^{\mathsf{v}} \\ &= [\overline{X},\overline{Z}](s(x))^{\mathsf{v}} + [\overline{Y},\overline{Z}](s(x))^{\mathsf{v}}.\end{aligned}$$

Since $\mathrm{pr}_2\colon VE\to E$ is a vector bundle homomorphism along π, this shows that for any section $s\in\Gamma(E)$, we have

$$R^\nabla(X+Y,Z)(s) = R^\nabla(X,Z)(s) + R^\nabla(Y,Z)(s).$$

Next, since the Lie bracket is anti-symmetric we certainly have

$$R^\nabla(X,Y)(s) = -R^\nabla(Y,X)(s).$$

Now suppose $f \in C^\infty(M)$. Then by part (ii) of Lemma 28.9 and Problem D.4, we have

$$\begin{aligned}[\overline{fX}, \overline{Y}](s(x))^{\mathsf{v}} &= [(f \circ \pi)\overline{X}, \overline{Y}](s(x))^{\mathsf{v}} \\ &= (f \circ \pi)(s(x))\,[\overline{X}, \overline{Y}](s(x))^{\mathsf{v}} - \overline{Y}(f \circ \pi)(s(x))\,\overline{X}(s(x))^{\mathsf{v}} \\ &= (f \circ \pi)(s(x))\,[\overline{X}, \overline{Y}](s(x))^{\mathsf{v}}\end{aligned}$$

since $\overline{X}(p)^{\mathsf{v}} = 0$ by definition of a horizontal lift. Thus

$$R^\nabla(fX, Y)(s) = fR^\nabla(X, Y)(s).$$

We have thus show that for a given section s, the map $R^\nabla(\cdot, \cdot)(s)$ is alternating and bilinear over $C^\infty(M)$. Thus it defines a section of the bundle $\operatorname{Hom}(\bigwedge^2(TM), \operatorname{Hom}(E, E))$ by the bundle-valued differential form criterion (Theorem 26.12).

2. It remains to show that R^∇ is $C^\infty(M)$-linear in the third argument, ie. that

$$R^\nabla(X, Y)(fs) = fR^\nabla(X, Y)(s).$$

This is a bit trickier. Since we already know that $R^\nabla(X, Y)$ maps sections of E to sections of E, by Proposition 16.26 it is sufficient to show that for fixed $x \in M$ and $v, w \in T_xM$, the map $R^\nabla(v, w)\colon E_x \to E_x$ is *linear* (over $\mathbb{R}$).

For this choose vector fields X, Y such that $X(x) = v$ and $Y(x) = w$. Since we already know R^∇ is a point operator in the first two variables, we may without loss of generality assume that $[X, Y] = 0$ on a neighbourhood of x. Let θ^X_t and θ^Y_t denote the local flows of X and Y.

Since $[X, Y] = 0$ near x, by either Lemma 33.10 above or Problem E.2, for sufficiently small $t > 0$ the curve γ_t obtained by concatenating the four curves:

(i) $s \mapsto \theta^X_s(x)$ for $0 \le s \le t$,

(ii) $s \mapsto \theta^Y_s \circ \theta^X_t(x)$ for $0 \le s \le t$,

(iii) $s \mapsto \theta^X_{-s} \circ \theta^Y_t \circ \theta^X_t(x)$ for $0 \le s \le t$,

(iv) $s \mapsto \theta^Y_{-s} \circ \theta^X_{-t} \circ \theta^Y_t \circ \theta^X_t(x)$ for $0 \le s \le t$,

is a piecewise smooth *loop* based at x. See Figure 33.1.

Now let $\theta^{\overline{X}}_t$ and $\theta^{\overline{Y}}_t$ denote the local flows of the horizontal lifts $\overline{X}$ and $\overline{Y}$. Then by Problem E.1, we have

$$\pi \circ \theta^{\overline{X}}_t = \theta^X_t \circ \pi, \qquad \pi \circ \theta^{\overline{Y}}_t = \theta^Y_t \circ \pi,$$

and by definition of the horizontal lift, for all $p \in E$ sufficiently close to E_x, we have

$$\theta^{\overline{X}}_t(p) = \mathbb{P}_{\delta^X}(p)(t), \qquad \text{where } \delta^X(t) := \theta^X_t(x).$$

and similarly

$$\theta^{\overline{Y}}_t(p) = \mathbb{P}_{\delta^Y}(p)(t), \qquad \text{where } \delta^Y(t) := \theta^Y_t(x).$$

Thus for $p \in E_x$ and $t > 0$ sufficiently small, the curve

$$\varepsilon_p(t) := \theta^{\overline{Y}}_{-\sqrt{t}} \circ \theta^{\overline{X}}_{-\sqrt{t}} \circ \theta^{\overline{Y}}_{\sqrt{t}} \circ \theta^{\overline{X}}_{\sqrt{t}}(p)$$

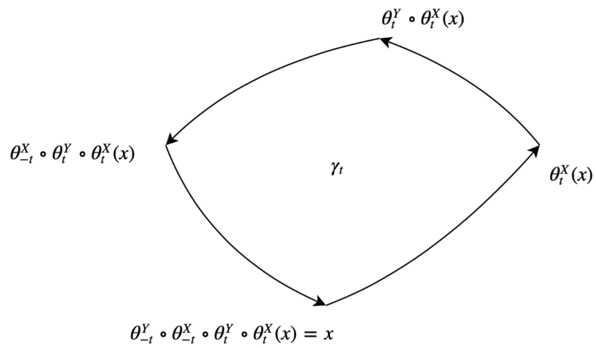

Figure 33.1: The piecewise smooth loop γ_t

is the parallel transport of p around the loop $\gamma_{\sqrt{t}}$:

$$\varepsilon_p(t) = \widehat{\mathbb{P}}_{\gamma_{\sqrt{t}}}(p). \tag{33.3}$$

Now we are in business: by Lemma 33.10 we have

$$[\overline{X}, \overline{Y}](p) = \varepsilon_p'(0).$$

Moreover since $\varepsilon_p(t)$ takes image in the fibre E_x, its tangent vector is vertical and hence $\mathrm{pr}_2(\varepsilon_p'(0)) = \mathcal{J}_p^{-1}(\varepsilon_p'(0))$. Thus

$$R^\nabla(v, w)(p) = -\mathcal{J}_p^{-1}(\varepsilon_p'(0)). \tag{33.4}$$

Now define a curve in $\mathrm{GL}(E_x)$ by

$$E(t)[p] := \varepsilon_p(t)$$

for small $t > 0$. Thus $E'(0) \in T_{\mathrm{id}}\,\mathrm{GL}(E_x) = \mathfrak{gl}(E_x) = \mathrm{Hom}(E_x, E_x)$. Then

$$R^\nabla(v, w) = -E'(0) \in \mathfrak{gl}(E_x)$$

is a linear operator, which implies what we wanted to prove. ■

Remark 33.12. The proof of Step 2 of Theorem 33.9 may have seemed somewhat roundabout, and an alternative proof is presented below. Nevertheless the geometric interpretation of the curvature in terms of parallel transport that our proof gave (specifically (33.3) and (33.4)) will turn out to be crucial next lecture when we formulate the *Ambrose-Singer Holonomy Theorem*, as well in our derivation of an alternative formula for R^∇ in Theorem 35.1.

Remark 33.13. Here[2] is a quicker way to prove $R^\nabla(X, Y)(\cdot)$ is a linear operator which is based on Theorem 31.7. Let $\mu_a \colon E \to E$ denote scalar multiplication, as in Definition 28.3. It follows from the defining condition (28.5) for a connection that any horizontal lift is μ_a-invariant:

$$\overline{X}(ap) = \overline{X}(\mu_a(p)) = D\mu_a(p)[\overline{X}(p)].$$

[2] Thank you to forum member "Miao" for suggesting I include this proof as well.

For $a \neq 0$, μ_a is a diffeomorphism, and thus we can write this as $(\mu_a)_\star(\overline{X}) = \overline{X}$. Thus by Proposition 7.20 for $a \neq 0$ we have

$$(\mu_a)_\star[\overline{X}, \overline{Y}] = [\overline{X}, \overline{Y}].$$

Now we use the fact that κ is a vector bundle morphism along $\pi_{TM}\colon TM \to M$ (Theorem 31.7), and thus in particular respects scalar multiplication, i.e.

$$\mu_a \circ \kappa = \kappa \circ D\mu_a$$

to obtain

$$\kappa\big([\overline{X}, \overline{Y}]\big) = \kappa\big((\mu_a)_\star[\overline{X}, \overline{Y}]\big) = \mu_a \circ \kappa\big([\overline{X}, \overline{Y}]\big).$$

This shows that $R^\nabla(v, w)\colon E_x \to E_x$ is a homogeneous map, i.e. $R^\nabla(v, w)(ap) = aR^\nabla(v, w)(p)$ for $a \neq 0$. Thus similarly to the proof of Step 1 of Theorem 30.2 (see the footnote in particular), since $R^\nabla(v, w)$ is differentiable at 0_x, it is then necessarily a linear map.

The holonomy algebra and the Ambrose-Singer Holonomy Theorem

We begin this lecture by defining the *holonomy algebra* of a connection, and stating the vector bundle version of the *Ambrose-Singer Holonomy Theorem*. We won't prove this theorem until Lecture 38, when we will first prove the principal bundle version, and then deduce the vector bundle version as a corollary.

DEFINITION 34.1. Let $\pi\colon E \to M$ be a vector bundle and let ∇ be a connection on E. We define the **holonomy algebra at** $x \in M$, written $\mathfrak{hol}^\nabla(x)$, to be the Lie algebra of $\mathrm{Hol}^\nabla(x)$. Since $\mathrm{Hol}^\nabla(x)$ is a Lie subgroup of $\mathrm{GL}(E_x)$, it follows that $\mathfrak{hol}^\nabla(x)$ is a Lie subalgebra of $\mathfrak{gl}(E_x) = \mathrm{L}(E_x, E_x)$, with Lie bracket given by matrix commutation (cf. Proposition 9.23):

$$[T_1, T_2] \coloneqq T_1 \circ T_2 - T_2 \circ T_1, \qquad T_1, T_2 \in \mathfrak{hol}^\nabla(x).$$

We then define

$$\mathfrak{hol}^\nabla \coloneqq \bigsqcup_{x \in M} \mathfrak{hol}^\nabla(x).$$

We call $\mathfrak{hol}^\nabla$ the **holonomy algebra** of ∇.

The holonomy algebra is itself a vector bundle over M. In fact, it is a (Lie) algebra bundle in the sense of Remark 15.27. Before proving this, we need another definition.

DEFINITION 34.2. Let $\pi\colon E \to M$ be a vector bundle and let ∇ be a connection on E. Suppose $E_0 \subset E$ is a vector subbundle of E. We say that the connection ∇ is **reducible** to E_0 if E_0 is invariant under parallel transport in the sense that if $\gamma\colon [0,1] \to M$ is a smooth curve and $p \in E_0|_{\gamma(0)}$ then $\mathbb{P}_\gamma(p)(1) \in E_0|_{\gamma(1)}$.

On Problem Sheet Q you will show that if ∇ is reducible to E_0 then ∇ induces a connection on E_0. In fact, the hypothesis that E_0 is a vector subbundle of E is superfluous, as the next lemma shows.

LEMMA 34.3. *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold and let ∇ be a connection on E. Assume $E_0 \subset E$ is a subset invariant under parallel transport with the property that there exists $x \in M$ such that $E_0 \cap E_x$ is a linear subspace of E_x. Then E_0 is a vector subbundle of E, and ∇ is reducible to E_0.*

Proof. Since $\widehat{\mathbb{P}}_\gamma\colon E_{\gamma(a)} \to E_{\gamma(b)}$ is a linear isomorphism for any smooth curve $\gamma\colon [a,b] \to M$, it follows that if $E_0 \cap E_x$ is a linear subspace of E_x for *some* point $x \in M$ then $E_0 \cap E_y$ is a linear subspace of E_y for *every* point $y \in M$. Vector subbundle charts on E_0 can be obtained by taking the restriction of the vector bundle charts on E built in Problem O.1. ∎

Next, by part (ii) of Problem P.1, we know that if ∇ is a connection on E then ∇ induces a connection[1] ∇^{Hom} on $\mathrm{Hom}(E,E)$.

THEOREM 34.4. *Let $\pi\colon E \to M$ be a vector bundle and let ∇ be a connection on E. Then the holonomy algebra $\mathfrak{hol}^{\nabla}$ is a vector subbundle of $\mathrm{Hom}(E,E)$ which in addition is a (Lie) algebra bundle in the sense of Remark 15.27. Moreover the induced connection ∇^{Hom} on $\mathrm{Hom}(E,E)$ is reducible to $\mathfrak{hol}^{\nabla}$.*

REMARK 34.5. The statement of Theorem 34.4 makes it look more complicated than it is, so here is the tl;dr version: *If $\gamma\colon [a,b] \to M$ is a smooth path and $T \in \mathfrak{hol}^{\nabla}(\gamma(a))$ then $\widehat{\mathbb{P}}_{\gamma} \circ T \circ \widehat{\mathbb{P}}_{\gamma}^{-1} \in \mathfrak{hol}^{\nabla}(\gamma(b))$.*

This proof is non-examinable.

(♣) *Proof.* We prove the result in three steps.

1. In this step we identify what parallel transport with respect to ∇^{Hom} in the bundle $\mathrm{Hom}(E,E)$ is. Suppose $x \in M$ and $\gamma\colon [0,1] \to M$ is a smooth curve with $\gamma(0) = x$. Abbreviate by $\widehat{\mathbb{P}}_t\colon E_x \to E_{\gamma(t)}$ parallel transport along the curve $r \mapsto \gamma(r)$ for $0 \le r \le t$ with respect to ∇ and similarly by

$$\widehat{\mathbb{P}}^{\mathrm{Hom}}_t\colon \mathrm{L}(E_{\gamma(0)}, E_{\gamma(0)}) \to \mathrm{L}(E_{\gamma(t)}, E_{\gamma(t)}).$$

the parallel transport with respect to ∇^{Hom}. Suppose $C \in \Gamma_{\gamma}(\mathrm{Hom}(E,E))$ is a section along γ, i.e.

$$C(t)\colon E_{\gamma(t)} \to E_{\gamma(t)}$$

is a linear map for each $t \in [0,1]$. It follows from Problem O.3, Problem P.1 and Proposition 32.3 that a section C is parallel along γ with respect to ∇^{Hom} if and only if for every section $c \in \Gamma_{\gamma}(E)$ which is parallel along γ with respect to ∇, the section $C[c] \in \Gamma_{\gamma}(E)$ defined by $t \mapsto C(t)[c(t)]$ is also parallel with respect to γ. This means that for $T \in \mathrm{L}(E_{\gamma(0)}, E_{\gamma(0)})$ we have

$$\widehat{\mathbb{P}}^{\mathrm{Hom}}_t(T) = \widehat{\mathbb{P}}_t \circ T \circ \widehat{\mathbb{P}}_t^{-1}. \tag{34.1}$$

2. In this step we use (34.1) to prove that $\mathfrak{hol}^{\nabla}$ is vector subbundle of $\mathrm{Hom}(E,E)$ and that ∇^{Hom} is reducible to $\mathfrak{hol}^{\nabla}$. Comparing (34.1) and Lemma 32.11, we see that the isomorphism $\mathrm{Hol}^{\nabla}(x) \cong \mathrm{Hol}^{\nabla}(\gamma(t))$ is exactly given by $\widehat{\mathbb{P}}^{\mathrm{Hom}}_t$:

$$\widehat{\mathbb{P}}^{\mathrm{Hom}}_t\colon \mathrm{Hol}^{\nabla}(x) \xrightarrow{\sim} \mathrm{Hol}^{\nabla}(\gamma(t)).$$

If we differentiate $\widehat{\mathbb{P}}^{\mathrm{Hom}}_t$ at $\mathrm{id} \in \mathrm{Hol}^{\nabla}(x)$, we get a linear map:

$$D\widehat{\mathbb{P}}^{\mathrm{Hom}}_t(\mathrm{id})\colon \mathfrak{hol}^{\nabla}(x) \to \mathfrak{hol}^{\nabla}(\gamma(t)).$$

But now as $\widehat{\mathbb{P}}^{\mathrm{Hom}}_t$ is a linear map, by Problem B.4 we have for all $T \in \mathfrak{hol}^{\nabla}(x)$ that

$$D\widehat{\mathbb{P}}^{\mathrm{Hom}}_t(\mathrm{id})\,[\mathcal{J}_{\mathrm{id}}(T)] = \mathcal{J}_{\mathrm{id}}(\widehat{\mathbb{P}}^{\mathrm{Hom}}_t(T)).$$

[1] In later lectures I will usually denote this connection as ∇ again (instead of ∇^{Hom}). However in this lecture for clarity I will keep the Hom superscript.

Suppressing the $\mathscr{J}$ maps, this says that $\widehat{\mathbb{P}}_t^{\mathrm{Hom}}$ defines a linear isomorphism

$$\widehat{\mathbb{P}}_t^{\mathrm{Hom}} \colon \mathfrak{hol}^\nabla(x) \to \mathfrak{hol}^\nabla(\gamma(t)),$$

which is exactly the statement that $\mathfrak{hol}^\nabla$ is invariant under parallel transport. We now apply Lemma 34.3 to deduce that $\mathfrak{hol}^\nabla$ is a vector subbundle of $\mathrm{Hom}(E,E)$ and the connection on $\mathrm{Hom}(E,E)$ induced by our original connection on E is reducible to $\mathfrak{hol}^\nabla$.

3. It remains to show that $\mathfrak{hol}^\nabla$ is actually a Lie algebra subbundle in the sense of Remark 15.27. For this the key observation is that if γ is a smooth curve in M and $C_1, C_2 \in \Gamma_\gamma(\mathrm{Hom}(E,E))$ are two parallel sections (with respect to ∇^{Hom}) then it follows from (34.1) that both $C_1 \circ C_2$ and $C_2 \circ C_1$ are also parallel, and thus so is the Lie bracket[2] $[C_1, C_2]$. Thus the vector bundle charts on $\mathrm{Hom}(E,E)$ constructed using Problem O.1 preserve the Lie bracket, and so can serve as Lie algebra charts for $\mathrm{Hom}(E,E)$. These charts restrict to Lie algebra charts on $\mathfrak{hol}^\nabla$ since the latter is invariant under parallel transport by Step 2. ∎

We now investigate how curvature affects the holonomy algebra. We first have:

LEMMA 34.6. *Let $\pi \colon E \to M$ be a vector bundle and suppose ∇ is a connection on E. Then for all $x \in M$ and $v, w \in T_xM$, the linear operator $R^\nabla(v,w) \in \mathfrak{gl}(E_x) = \mathrm{L}(E_x, E_x)$ actually belongs to $\mathfrak{hol}^\nabla(x)$.*

Proof. This is immediate from the proof of Step 2 of Theorem 33.9—specifically (33.3) and (33.4). ∎

DEFINITION 34.7. Let $\pi \colon E \to M$ be a vector bundle and let ∇ be a connection on M. We say that ∇ is **flat near** x if there exists an open set $U \subset M$ containing x such that $R^\nabla(v,w)$ is the zero operator[3] for all $y \in U$ and $v, w \in T_yM$.

One could naively hope that $\mathfrak{hol}^\nabla(x)$ is generated by all elements of the form $R^\nabla(v,w)$ for $v, w \in T_xM$. However it is perfectly possible for R^∇ to be flat near $x \in M$ without ∇ being globally flat, and thus in this situation $\mathfrak{hol}^\nabla(x)$ is non-zero but $R^\nabla(v,w)$ is the zero map for all $v, w \in T_xM$. Assume M is connected. Then it is easy to see how to construct a non-zero element of $\mathfrak{hol}^\nabla(x)$ using the curvature: choose some point $y \in M$ such that R^∇ is not flat near y, and choose $v, w \in T_yM$ such that $R^\nabla(v,w) \neq 0$. Thus $R^\nabla(v,w)$ determines a non-zero element in $\mathfrak{hol}^\nabla(y)$. Choose a smooth path $\gamma \colon [0,1] \to M$ such that $\gamma(0) = y$ and $\gamma(1) = x$. Then by Theorem 34.4,

$$\widehat{\mathbb{P}}_\gamma^{\mathrm{Hom}}(R^\nabla(v,w)) = \widehat{\mathbb{P}}_\gamma \circ R^\nabla(v,w) \circ \mathbb{P}_\gamma^{-1}$$

is a non-zero element of $\mathfrak{hol}^\nabla(x)$. The next theorem, which is one of the cornerstones of the subject, tells us that the entire Lie algebra $\mathfrak{hol}^\nabla(x)$ is generated by elements of this form.

---

[2]In slightly fancier language, this shows that the Lie bracket $[\cdot,\cdot]$ is itself a parallel section of the bundle $\mathrm{Hom}\left(\bigwedge^2(\mathrm{Hom}(E,E)), \mathrm{Hom}(E,E)\right)$, where the latter bundle is endowed with the connection induced from ∇ on E.

[3]By a standard abuse of notation we write $R^\nabla(v,w) \neq 0$ to mean $R^\nabla(v,w)$ is not the zero operator $E_x \to E_x$.

THEOREM 34.8 (The Ambrose-Singer Holonomy Theorem). *Let $\pi\colon E \to M$ be a vector bundle over a connected manifold M and let ∇ be a connection on E. Then for any $x \in M$, the holonomy algebra $\mathfrak{hol}^{\nabla}(x)$ at x is the vector subspace of $\mathfrak{gl}(E_x)$ spanned by all the elements of the form*

$$\mathbb{P}_\gamma \circ R^{\nabla}(v,w) \circ \mathbb{P}_\gamma^{-1}, \qquad y \in M,\ v,w \in T_yM$$

where γ is a piecewise smooth path in M from y to x.

The proof of Theorem 34.8 is deferred to Lecture 41, where we[4] will deduce it from a more general version for principal bundles (see Theorem 41.7). Instead, now we work towards deriving a more convenient formula for R^{∇}.

DEFINITION 34.9. Let $\pi\colon E \to N$ denote a vector bundle with connection ∇, and let $\varphi\colon M \to N$ denote a smooth map. Define for $X, Y \in \mathfrak{X}(M)$ and $s \in \Gamma_\varphi(E)$

$$R_\varphi^{\nabla}(X,Y)(s) = \nabla_X(\nabla_Y(s)) - \nabla_Y(\nabla_X(s)) - \nabla_{[X,Y]}(s). \tag{34.2}$$

We will prove next lecture that for $\varphi = \mathrm{id}$ we have

$$R_{\mathrm{id}}^{\nabla} = R^{\nabla}.$$

After doing so we will drop the φ subscript and just write R^{∇} for the operator defined in (34.2)—this is consistent with the fact we denote all covariant derivatives by ∇ and not, eg, ∇_φ. One can interpret (34.2) as

$$R_\varphi^{\nabla}(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}.$$

The first term measures the failure of ∇_X and ∇_Y to commute, and the second term is subtracted to make the following result true.

PROPOSITION 34.10. *The operator R_φ^{∇} is $C^\infty(M)$-linear in all three variables, and antisymmetric in the first two variables.*

Proof. We prove only that $R_\varphi^{\nabla}(fX,Y)(s) = fR_\varphi^{\nabla}(X,Y)(s)$; the remaining computations are similar and left as an exercise. By Problem D.4 we have $[fX,Y] = f[X,Y] - Y(f)X$ and hence

$$\begin{aligned}
R_\varphi^{\nabla}(fX,Y) &= \nabla_{fX}(\nabla_Y(s)) - \nabla_Y(\nabla_{fX}(s)) - \nabla_{[fX,Y]}(s)\\
&= f\nabla_X(\nabla_Y(s)) - \nabla_Y(f\nabla_X(s)) - \nabla_{f[X,Y]}(s) + \nabla_{Y(f)X}(s)\\
&= f\left(\nabla_X(\nabla_Y(s)) - \nabla_Y(\nabla_X(s)) - \nabla_{[X,Y]}(s)\right) - Y(f)\nabla_X(s) + Y(f)\nabla_X(s)\\
&= fR_\varphi^{\nabla}(X,Y)(s).
\end{aligned}$$

This completes the proof. ∎

Since R_φ^{∇} is $C^\infty(M)$-linear is all variables, it is a point operator in all three variables by Proposition 16.25, and hence $R_\varphi^{\nabla}(v,w)(p)$ is well defined for any $v,w \in T_xM$ and $p \in E_{\varphi(x)}$.

[4]Actually "you": the proof of Theorem 34.8 is on Problem Sheet S.

Proposition 34.11. *Let $\pi\colon E \to N$ denote a vector bundle with connection ∇, and let $\varphi\colon M \to N$ denote a smooth map. Then for all $x \in M$, $v, w \in T_xM$ and $p \in E_{\varphi(x)}$ we have*

$$R^{\nabla}_{\varphi}(v,w)(p) = R^{\nabla}_{\mathrm{id}}\big(D\varphi(x)[v], D\varphi(x)[w]\big)(p). \tag{34.3}$$

In particular, if $\varphi\colon M \to N$ is a diffeomorphism then for all $X, Y \in \mathfrak{X}(M)$ and $s \in \Gamma_{\varphi}(E)$ we have

$$R^{\nabla}_{\mathrm{id}}(\varphi_{\star}(X), \varphi_{\star}(Y))(s) = R^{\nabla}_{\varphi}(X,Y)(s).$$

Proof. Assume $X, Y \in \mathfrak{X}(M)$ are φ-related to vector fields $Z, W \in \mathfrak{X}(N)$, and assume $s \in \Gamma_{\varphi}(E)$ has the property that $s(x) = \tilde{s}(\varphi(x))$ for some section $\tilde{s}$ of E and all $x \in M$. Then by repeatedly applying the chain rule for covariant derivatives (31.7) we have

$$\begin{aligned}
\nabla_X(\nabla_Y(s)) &= \nabla_X(\nabla_Y(\tilde{s} \circ \varphi)) \\
&= \nabla_X(\nabla_{W \circ \varphi}(\tilde{s}))) \\
&= \nabla_X(\nabla_W(\tilde{s}) \circ \varphi) \\
&= \nabla_{Z \circ \varphi}(\nabla_W(\tilde{s})) \\
&= \nabla_Z(\nabla_W(\tilde{s})) \circ \varphi.
\end{aligned}$$

Similarly $\nabla_Y(\nabla_X(s)) = \nabla_W(\nabla_Z(\tilde{s})) \circ \varphi$. Moreover by Problem D.5 we have $\nabla_{[X.Y]}(s) = \nabla_{[Z,W]}(\tilde{s}) \circ \varphi$, and hence

$$R^{\nabla}_{\varphi}(X,Y)(s) = \Big(R^{\nabla}_{\mathrm{id}}(Z,W)(\tilde{s})\Big) \circ \varphi. \tag{34.4}$$

To prove the general case, we first claim that the module $\Gamma_{\varphi}(TN)$ of vector fields along φ is locally generated by elements of the form $Z \circ \varphi$ for Z a vector field defined on some open subset of N. Indeed, if $X \in \mathfrak{X}(M)$ and $x_0 \in M$, let $\tau\colon V \to \Omega$ be a chart on N about $\varphi(x_0)$ with local coordinates y^i. Set $U := \varphi^{-1}(V \cap \varphi(M))$ and define smooth functions $f^i\colon U \to \mathbb{R}$ by

$$f^i := X(y^i \circ \varphi).$$

Then for $x \in U$ we have

$$D\varphi(x)[X(x)] = f^i(x) \frac{\partial}{\partial y^i}\Big|_{\varphi(x)},$$

which proves the claim. Similarly sections of the form $\tilde{s} \circ \varphi$ locally generate $\Gamma_{\varphi}(E)$ over $C^{\infty}(M)$. Since both sides of (34.3) are point operators in all three variables, (34.3) follows from (34.4). ■

The exterior covariant differential

We begin today's lecture by completing our discussion of curvature and proving that for $\varphi = \mathrm{id}$ the operator R^{∇}_{id} from Definition 34.9 agrees with the curvature R^{∇}.

THEOREM 35.1. *Let $\pi \colon E \to M$ denote a vector bundle with connection ∇. Then $R^{\nabla}_{\mathrm{id}} = R^{\nabla}$.*

Proof. Consider $\mathbb{R}^2$ with coordinates[1] (s,t). Let S, T denote the vector fields $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$ on $\mathbb{R}^2$ respectively. Now fix $x \in M$, $v, w \in T_x M$ and $p \in E_x$. Choose a smooth map $\gamma \colon (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \to M$ such that $\gamma(0,0) = x$ and

$$D\gamma(0,0)[S(0,0)] = v, \qquad D\gamma(0,0)[T(0,0)] = w.$$

Now define a section[2] $\eta \in \Gamma_\gamma(E)$ such that $\eta(0,0) = p$ and such that:

(i) η is parallel along the curve $t \mapsto \gamma(0,t)$,

(ii) η is parallel along the curve $s \mapsto \gamma(s,t)$ for all $t \in (-\varepsilon, \varepsilon)$.

Such a section exists and is unique by Proposition 29.7. To see this first apply Proposition 29.7 along to curve to $t \mapsto \gamma(0,t)$ so that (i) is satisfied. Then define η along each curve $s \mapsto \gamma(s,t)$ again via Proposition 29.7. The fact that the resulting section η is smooth in both s and t is due to the fact that integral curves depend smoothly on initial conditions (see the proof of Proposition 29.7 and Theorem 8.1). Using (ii) and the fact that $[S,T] = 0$ we obtain

$$R^{\nabla}_{\gamma}(S(0,0), T(0,0))(p) = \nabla_{S(0,0)}(\nabla_T(\eta))(0,0).$$

Let $\widehat{\mathbb{P}}_s \colon E_x \to E_{\gamma(s,0)}$ denote parallel transport along $r \mapsto \gamma(r,0)$ for $0 \le r \le s$ and let $\widehat{\mathbb{P}}_{s,t} \colon E_{\gamma(s,0)} \to E_{\gamma(s,t)}$ denote parallel transport along $r \mapsto \gamma(s,r)$ for $0 \le r \le t$. Then by Proposition 32.3 we have

$$\nabla_T(\eta)(s,0) = \frac{d}{dt}\Big|_{t=0} \widehat{\mathbb{P}}_{s,t}^{-1}(\eta(s,t))$$

and thus

$$R^{\nabla}_{\gamma}(S(0,0), T(0,0))(p) = \frac{d^2}{dsdt}\Big|_{(s,t)=(0,0)} \widehat{\mathbb{P}}_s^{-1}\widehat{\mathbb{P}}_{s,t}^{-1}(\eta(s,t))$$

Thus by the definition of the derivative as a limit, the right-hand side is equal to

$$\lim_{s,t \to 0} \frac{\widehat{\mathbb{P}}_s^{-1}\widehat{\mathbb{P}}_{s,t}^{-1}(\eta(s,t)) - \widehat{\mathbb{P}}_s^{-1}\widehat{\mathbb{P}}_{s,0}^{-1}(\eta(s,0)) - \widehat{\mathbb{P}}_0^{-1}\widehat{\mathbb{P}}_{0,t}^{-1}(\eta(0,t)) + \widehat{\mathbb{P}}_0^{-1}\widehat{\mathbb{P}}_{0,0}^{-1}(\eta(0,0))}{st}$$

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] We break our usual convention that the coordinates on $\mathbb{R}^2$ are (x^1, x^2) here so as to simplify the notation in this proof.

[2] Sadly, using s as a coordinate means I can't use s as a section. . .

Since $\eta(s,0) = \widehat{\mathbb{P}}_s(p)$ by assumption (i), $\eta(0,t) = \widehat{\mathbb{P}}_{0,t}(p)$ by assumption (ii) and $\widehat{\mathbb{P}}_{s,0} = \mathrm{id}$ by definition we can simplify this to

$$\lim_{s,t\to 0} \frac{\widehat{\mathbb{P}}_s^{-1}\widehat{\mathbb{P}}_{s,t}^{-1}(\eta(s,t)) - p}{st}$$

Now take $s = t$ to obtain

$$R^\nabla_\gamma(S(0,0), T(0,0))(p) = \lim_{t\to 0} \frac{\widehat{\mathbb{P}}_t^{-1}\widehat{\mathbb{P}}_{t,t}^{-1}(\eta(t,t)) - p}{t^2}.$$

Finally set $r = \sqrt{t}$ and observe that the right-hand side is exactly the parallel transport of p along the inverse[3] of the loop γ_r from the proof of Step 2 of Theorem 33.9. Thus by (33.3) and (33.4) we obtain

$$R^\nabla_\gamma(S(0,0), T(0,0))(p)) = R^\nabla(v,w)(p).$$

Finally by Proposition 34.11 we have

$$R^\nabla_\gamma(S(0,0), T(0,0))(p)) = R^\nabla_{\mathrm{id}}(v,w)(p).$$

This completes the proof. ■

We now introduce the sheaf-theoretic definition of a connection. This requires us to use bundle-valued differential forms (Definition 26.10). We denote by $\Omega^r_{M,E}$ the sheaf $U \mapsto \Omega^r(U,E)$ of E-valued differential r-forms over $U \subset M$ open, and by $\Omega_{M,E}$ the sheaf $U \mapsto \Omega(U,E)$. Thus $\Omega^0(M,E) = \Gamma(E)$ and $\Omega^1(M,E)$ can be identified with $C^\infty(M)$-linear maps $\xi\colon \mathfrak{X}(M) \to \Gamma(E)$ (cf. Theorem 26.12). In particular, Remark 31.9 tells us that we can think of ∇s as belonging to $\Omega^1(M,E)$.

A **decomposable** element of $\Omega^r(U,E)$ is an element of the form $\xi = \omega \otimes s$ where $\omega \in \Omega^r(U)$ and $s \in \Gamma(U,E)$. Note that any $\mathbb{R}$-linear sheaf morphism $\Omega^r_{M,E} \to \Omega^r_{M,E}$ is entirely determined by what it does to decomposable elements on arbitrarily small open sets U.

Lemma 35.2. *A connection ∇ on E is equivalent to an $\mathbb{R}$-linear sheaf morphism $\nabla\colon \Omega^0_{M,E} \to \Omega^1_{M,E}$ which satisfies the Leibniz rule*

$$\nabla(fs) = df \otimes s + f\,\nabla s.$$

Proof. The axioms of a covariant derivative (Definition 31.8) tell us that we get an $\mathbb{R}$-linear map

$$\nabla\colon \Omega^0(M,E) \to \Omega^1(M,E)$$

satisfying the Leibniz rule. We therefore need only check that ∇ is a sheaf morphism. But this is immediate from Proposition 16.22: every local operator induces a sheaf morphism. ■

[3]The inverse is consistent with the minus sign in our original Definition 33.6 of R^∇.

So far all we have done is added notational complexity. Recall that the exterior differential $d\colon \Omega_M \to \Omega_M$ is a graded derivation of degree 1 (Definition 19.12), which extends the operation $f \mapsto df$ to higher differential forms. We now play the same game with connections.

First, some preliminaries. In Definition 26.5 we discussed how to define the wedge product for vector-valued forms. Now we will need a version of the wedge product for bundle-valued forms. Rather than work in maximal generality, we will give the relevant definitions only for the case we are interested in. Compare this to Definition 17.28.

DEFINITION 35.3. The sheaf $\Omega_{M,E}$ is a sheaf of Ω_M-**bimodules** in the sense that there are wedge products

$$\wedge\colon \Omega_M \times \Omega_{M,E} \to \Omega_{M,E}$$

and

$$\wedge\colon \Omega_{M,E} \times \Omega_M \to \Omega_{M,E}$$

which restrict to sheaf morphisms

$$\Omega^r_M \times \Omega^k_{M,E} \to \Omega^{r+k}_{M,E}, \qquad \Omega^k_{M,E} \times \Omega^r_M \to \Omega^{r+k}_{M,E}$$

and are compatible in the sense that

$$(\omega \wedge \xi) \wedge \vartheta = \omega \wedge (\xi \wedge \vartheta), \qquad \xi \in \Omega_{M,E},\ \omega, \vartheta \in \Omega_M.$$

Explicitly, these wedge product are defined on decomposable elements $\xi = \omega \otimes s$ as follows:

$$(\omega \otimes s) \wedge \vartheta := (\omega \wedge \vartheta) \otimes s.$$

where wedge product on the right-hand side is normal wedge product, and similarly

$$\vartheta \wedge (\omega \otimes s) := (\vartheta \wedge \omega) \otimes s,$$

and then extended by linearity. Just as the wedge product reduces to multiplication for 0-forms, we define

$$\omega \wedge s = s \wedge \omega := \omega \otimes s, \qquad \omega \in \Omega_M,\ s \in \Omega^0_{M,E}. \tag{35.1}$$

It follows from the definition that the wedge product is again graded commutative in the sense that

$$\omega \wedge \xi = (-1)^{rk} \xi \wedge \omega, \qquad \omega \in \Omega^r_M,\ \xi \in \Omega^k_{M,E}. \tag{35.2}$$

We can now formulate our main result.

THEOREM 35.4. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. There exists a unique $\mathbb{R}$-linear sheaf morphism*

$$d^\nabla\colon \Omega_{M,E} \to \Omega_{M,E}$$

of degree 1, i.e. that d^∇ restricts to define sheaf morphisms

$$d^\nabla\colon \Omega^r_{M,E} \to \Omega^{r+1}_{M,E}$$

such that:

(i) d^∇ *is a graded derivation with respect to the wedge products from Definition 35.3, i.e. for* $\omega \in \Omega^r_M$ and $\xi \in \Omega^k_{M,E}$ *we have*

$$\begin{aligned} d^\nabla(\omega \wedge \xi) &= d\omega \wedge \xi + (-1)^r \omega \wedge d^\nabla \xi. \\ d^\nabla(\xi \wedge \omega) &= d^\nabla \xi \wedge \omega + (-1)^k \xi \wedge d\omega. \end{aligned} \tag{35.3}$$

(ii) d^∇ *is equal to* ∇ *on* $\Omega^0_{M,E}$*:* $d^\nabla s = \nabla s$ *for* $s \in \Omega^0_{M,E}$.

We call d^∇ *the* **exterior covariant differential** *associated to the connection* ∇ *as refer to* $d^\nabla \xi$ *as the exterior covariant differential of* ξ.

The proof of Theorem 35.4 is very similar to the proof of Theorem 19.17. Indeed, if one takes E to be the trivial bundle $M \times \mathbb{R} \to \mathbb{R}$ and ∇ to be the trivial connection then $d^\nabla = d$ and the proof is of Theorem 35.4 reduces exactly to that of Theorem 19.17. The general case is only notationally different, and we leave it to the interested reader as an exercise.

We also have the following analogue of Theorem 20.7, which uses the Bundle-Valued Differential Form Criterion (Theorem 26.12) to make sense of its statement.

THEOREM 35.5. *Let* $\pi\colon E \to M$ *be a vector bundle with connection* ∇*. Let* $\xi \in \Omega^r(M,E)$ *and let* $X_0, \dots X_r \in \mathfrak{X}(M)$*. Then:*

$$\begin{aligned} d^\nabla \xi(X_0, \dots, X_r) = &\sum_{i=0}^r (-1)^i \nabla_{X_i}\big(\xi(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \sum_{0 \le i < j \le r} (-1)^{i+j} \xi([X_i, X_j], X_0, \dots \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_r). \end{aligned}$$

The proof is by induction on r, and proceeds in exactly the same was as Theorem 20.7. Similarly we have the following version of Lemma 19.19:

LEMMA 35.6. *Let* $\pi\colon E \to N$ *be a vector bundle with connection* ∇*. Let* $\varphi\colon M \to N$ *be a smooth map and let* $\xi \in \Omega_{N,E}$*. Then*

$$\varphi^\star(d^\nabla \xi) = d^\nabla(\varphi^\star(\omega)).$$

that is, $\varphi^\star$ *commutes with the exterior covariant differentials.*

REMARK 35.7. If $\pi\colon E \to M$ is an algebra bundle and ∇ is a connection such that the algebra multiplication $\beta\colon E \times E \to E$ is parallel in the sense that if γ is a curve in M and c_1, c_2 are two parallel sections then $\beta(c_1, c_2)$ is also parallel along γ then d^∇ satisfies a product rule

$$d^\nabla\big(\beta(\xi_1, \xi_2)\big) = \beta\big(d^\nabla \xi_1, \xi_2\big) + (-1)^r \beta\big(\xi_1, d^\nabla \xi_2\big), \qquad \xi_1 \in \Omega^r_{M,E},\ \xi_2 \in \Omega_{M,E}.$$

The proof of this assertion is similar to Proposition 26.6 (which was also Problem M.6).

Unlike the exterior differential however, the exterior covariant differential does *not* necessarily square to zero. For this we need a bit more formalism. Since the notation is rather cumbersome, we will effect the following convention.

DEFINITION 35.8. Let $\pi\colon E \to M$ be a vector bundle. We abbreviate

$$\mathcal{A}^r(U, E) := \Omega^r(U, \mathrm{Hom}(E, E))$$

the space of **endomorphism-valued r-forms** on E and by $\mathcal{A}^r_{M,E}$ the corresponding sheaf: thus $\mathcal{A}^r_{M,E} = \Omega^r_{M,\mathrm{Hom}(E,E)}$. We define $\mathcal{A}_{M,E}$ similarly.

EXAMPLE 35.9. Let $\pi\colon E \to M$ be a vector bundle and let ∇ be a connection on E. Then Theorem 33.9 tells us that the curvature of R^∇ is an element of $\mathcal{A}^2(M, E)$.

A decomposable element of $\mathcal{A}^r(M, E)$ is of the form $\omega \otimes T$ where $\omega \in \Omega^r(M)$ and $T \in \Gamma(\mathrm{Hom}(E, E))$. This allows us to extend Definition 35.3 and define (even) more wedge products.

DEFINITION 35.10. We turn the sheaf $\mathcal{A}_{M,E}$ into a sheaf of algebras over $\mathcal{C}^\infty_M$ via the multiplication

$$\mathcal{A}_{M,E} \times \mathcal{A}_{M,E} \to \mathcal{A}_{M,E}.$$

given by

$$(\omega \otimes T) \wedge (\vartheta \otimes S) := (\omega \wedge \vartheta) \otimes (T \circ S),$$

Moreover $\Omega_{M,E}$ is also a sheaf of left $\mathcal{A}_{M,E}$-modules via the wedge product

$$\wedge\colon \mathcal{A}_{M,E} \times \Omega_{M,E} \to \Omega_{M,E}$$

defined on decomposable elements by

$$(\omega \otimes T) \wedge (\vartheta \otimes s) := (\omega \wedge \vartheta) \otimes T(s).$$

Note this wedge products restrict to a sheaf morphism

$$\mathcal{A}^r_M \times \Omega^k_{M,E} \to \Omega^{r+k}_{M,E}.$$

This wedge product makes $\Omega_{M,E}$ into a a sheaf of $\mathcal{A}_{M,E}$-Ω_M bimodules, in the sense that

$$(A \wedge \xi) \wedge \omega = A \wedge (\xi \wedge \omega), \qquad A \in \mathcal{A}_{M,E},\ \xi \in \Omega_{M,E},\ \omega \in \Omega_M. \tag{35.4}$$

We conclude this lecture with the following result, which is guaranteed to make your head hurt.

THEOREM 35.11. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. For all $\xi \in \Omega^r(M, E)$ one has*

$$d^\nabla \circ d^\nabla(\xi) = R^\nabla \wedge \xi.$$

Thus $d^\nabla \circ d^\nabla = 0$ if and only if ∇ is flat.

Proof. We first prove the result in the special case $r = 0$, so that $\xi = s$ is just a section of E. Let $X, Y \in \mathfrak{X}(M)$. Then using Theorem 35.5 and Theorem 35.1 we compute:

$$\begin{aligned} d^\nabla \circ d^\nabla(s)(X, Y) &= \nabla_X(\nabla s(Y)) - \nabla_Y(\nabla s(X)) - \nabla s([X, Y])) \\ &= \nabla_X(\nabla_Y(s)) - \nabla_Y(\nabla_X(s)) - \nabla_{[X,Y]}(s) \\ &= R^\nabla(X, Y)(s), \end{aligned}$$

and hence

$$d^\nabla \circ d^\nabla s = R^\nabla \wedge s. \tag{35.5}$$

For the general case it suffices to take $\xi = \omega \otimes s$ to be a decomposable element. Then we compute

$$\begin{aligned}
d^\nabla \circ d^\nabla \xi &\overset{(35.1)}{=} d^\nabla \circ d^\nabla (\omega \wedge s) \\
&\overset{(35.3)}{=} d^\nabla \big(d\omega \wedge s + (-1)^r \omega \wedge d^\nabla s \big) \\
&= d(d\omega) \wedge s + (-1)^{r+1} d\omega \wedge d^\nabla (s) + (-1)^r d\omega \wedge d^\nabla s + (-1)^{2r} \omega \wedge \big(d^\nabla \circ d^\nabla s \big) \\
&\overset{(35.5)}{=} \omega \wedge (R^\nabla \wedge s) \\
&\overset{(35.2)}{=} (R^\nabla \wedge s) \wedge \omega \\
&\overset{(35.4)}{=} R^\nabla \wedge (s \wedge \omega) \\
&\overset{(35.2)}{=} R^\nabla \wedge (\omega \wedge s) \\
&\overset{(35.1)}{=} R^\nabla \wedge \xi.
\end{aligned}$$

This completes the proof. ∎

The Bianchi Identity and Riemannian vector bundles

We begin this lecture by stating and proving the *Bianchi identity*. This identity is the starting point for using connections to study de Rham cohomology of a manifold via *characteristic classes*. Let $\pi\colon E \to M$ denote a vector bundle and fix a connection ∇ on E. We denote by

$$d^{\nabla^{\mathrm{Hom}}}\colon \mathcal{A}_{M,E} \to \mathcal{A}_{M,E}$$

the exterior covariant differential associated to the connection ∇^{Hom} on $\mathrm{Hom}(E,E)$. By Example 35.9 the curvature R^{∇} of ∇ is an element of $\mathcal{A}^2(M,E)$, and hence $d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) \in \mathcal{A}^3(M,E)$. In fact, this element is always zero.

Theorem 36.1 (The Bianchi Identity). *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Then*

$$d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) = 0.$$

Proof. Let $\xi \in \Omega(M,E)$. We compute $(d^{\nabla})^3(\xi) := d^{\nabla} \circ d^{\nabla} \circ d^{\nabla}\xi$ in two ways. Firstly, by Theorem 35.11 we have

$$(d^{\nabla})^3(\xi) = (d^{\nabla})^2\big(d^{\nabla}\xi\big) = R^{\nabla} \wedge d^{\nabla}\xi. \tag{36.1}$$

However if we use Problem Q.3 in addition to Theorem 35.11 we alternatively have

$$\begin{aligned}
(d^{\nabla})^3(\xi) &= d^{\nabla}\big((d^{\nabla})^2(\xi)\big) \\
&= d^{\nabla}(R^{\nabla} \wedge \xi) \\
&= d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) \wedge \xi + (-1)^2 R^{\nabla} \wedge d^{\nabla}\xi \\
&= d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) \wedge \xi + R^{\nabla} \wedge d^{\nabla}\xi.
\end{aligned}$$

Comparing this with (36.1) tells us that

$$R^{\nabla} \wedge d^{\nabla}\xi = d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) \wedge \xi + R^{\nabla} \wedge d^{\nabla}\xi,$$

and hence

$$d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) \wedge \xi = 0, \qquad \forall\, \xi \in \Omega(M,E).$$

This implies that $d^{\nabla^{\mathrm{Hom}}}(R^{\nabla}) = 0$, and thus completes the proof. ∎

We now motivate the construction of characteristic classes by considering a simple—and ultimately, useless (see Proposition 36.6 below)—example.

Suppose $\pi\colon E \to M$ is a vector bundle and η is a section of the dual bundle E^*. If $\xi \in \Omega^r(M,E)$ is an E-valued differential r-form on M, then we can feed ξ to η to obtain a normal differential r-form $\eta(\xi) \in \Omega^r(M)$. Explicitly, if $\xi = \omega \otimes s$ is decomposable then $\eta(\xi) := \eta(s)\,\omega$.

Lemma 36.2. *Suppose $\pi\colon E \to M$ is a vector bundle with connection ∇. Suppose $\eta \in \Gamma(E^*)$ is a section of the dual bundle which is parallel with respect to the induced connection. Then for any $\xi \in \Omega^r(M, E)$, we have*

$$d(\eta(\xi)) = \eta\big(d^\nabla \xi\big)$$

as elements of $\Omega^{r+1}(M)$.

Proof. If $X_0, \dots, X_r$ are vector fields on M then by Theorem 20.7

$$\begin{aligned} d(\eta(\xi))(X_0, \dots, X_r) = &\sum_{i=0}^{r} (-1)^i X_i\big(\eta(\xi)(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \sum_{0 \le i < j \le r} (-1)^{i+j} \eta(\xi)([X_i, X_j], X_0, \dots \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_r). \end{aligned}$$

Using the definition of the induced connection on E^* (part (ii) of Problem O.3), we have

$$\begin{aligned} X_i\big(\eta(\xi)(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) = &(\nabla_{X_i}(\eta))\big(\xi(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \eta\big(\nabla_{X_i}(\xi(X_0, \dots, \widehat{X}_i, \dots, X_r))\big), \end{aligned}$$

and thus by Theorem 35.5 we have

$$\begin{aligned} d(\eta(\xi))(X_0, \dots, X_r) = &\sum_{i=0}^{r} (-1)^i \nabla_{X_i}(\eta)\big(\xi(X_0, \dots, \widehat{X}_i, \dots, X_r)\big) \\ &+ \eta\big(d^\nabla \xi(X_0, \dots, X_r)\big). \end{aligned}$$

If η is parallel then $\nabla \eta = 0$, and thus the result follows. ∎

Now consider the trace operator

$$\mathrm{tr}\colon \mathrm{Mat}(k) \to \mathbb{R}$$

that sends a matrix to its trace. We will show that tr induces a parallel section of the dual bundle to the homomorphism bundle. Recall the frame bundle $\mathrm{Fr}(E)$ associated to E from Definition 24.13. An element of the fibre $\mathrm{Fr}(E_x)$ is a linear isomorphism $F\colon \mathbb{R}^k \to E_x$.

Proposition 36.3. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. There is a well-defined section η of the bundle $(\mathrm{Hom}(E, E))^*$ given by the trace:*

$$\eta_x(T) := \mathrm{tr}(F^{-1} \circ T \circ F), \qquad T \in \mathrm{L}(E_x, E_x), \tag{36.2}$$

where $F \in \mathrm{Fr}(E_x)$ is any element. Morevoer this section η is parallel with respect to the dual connection on $\mathrm{Hom}(E, E)^*$ induced by the connection ∇^{Hom} on $\mathrm{Hom}(E, E)$.

Proof. To prove that η is well defined we observe that if $\tilde{F}\colon \mathbb{R}^k \to E_x$ was another element of $\mathrm{Fr}(E_x)$ then $L := F^{-1}\tilde{F} \in \mathrm{GL}(k)$, and

$$\mathrm{tr}(\tilde{F}^{-1} \circ T \circ \tilde{F}) = \mathrm{tr}(L^{-1}F^{-1} \circ T \circ FL) = \mathrm{tr}(F^{-1} \circ T \circ F).$$

To prove that η is parallel with respect to the dual connection on $(\mathrm{Hom}(E,E))^*$ induced by ∇^{Hom}, by part (i) of Problem O.3 we need to show that η is constant along parallel sections of $\mathrm{Hom}(E,E)$ with respect to ∇^{Hom}. Fix $x \in M$ and let $\gamma\colon [0,1] \to M$ be a curve with $\gamma(0) = x$. Let $\{e_i\}$ be a local frame of E over an open set U containing x which is parallel along γ (Lemma 31.5), and let $\alpha\colon \pi^{-1}(U) \to \mathbb{R}^k$ denote the associated vector bundle chart. Suppose $C \in \Gamma_\gamma(\mathrm{Hom}(E,E))$ is parallel with respect to ∇^{Hom}. Then by the discussion in Step 1 of the proof of Theorem 34.4, the curve

$$\tilde{C}(t) := \alpha|_{E_{\gamma(t)}} \circ C(t) \circ \alpha|_{E_{\gamma(t)}}^{-1}$$

is a constant curve in $\mathfrak{gl}(k)$. Thus

$$\eta_{\gamma(t)}(C(t)) = \mathrm{tr}(\tilde{C}(t))$$

is constant as required. ■

From now on by a slight abuse of notation we will denote the section η defined in (36.2) also by tr. What have we gained from this construction?

Corollary 36.4. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Then the differential 2-form $\mathrm{tr}(R^\nabla)$ is closed, and hence defines a de Rham cohomology class $[\mathrm{tr}(R^\nabla)] \in H^2_{\mathrm{dR}}(M)$.*

Proof. We apply Lemma 36.2 applied with “E” equal to $\mathrm{Hom}(E,E)$ and “∇” equal to ∇^{Hom}. Then using also the Bianchi Identity (Theorem 36.1) we have

$$d(\mathrm{tr}(R^\nabla)) = \mathrm{tr}\left(d^{\nabla^{\mathrm{Hom}}}(R^\nabla)\right) = 0.$$

Thus $\mathrm{tr}(R^\nabla)$ is closed, as required. ■

What is more surprising is that the cohomology class $[\mathrm{tr}(R^\nabla)]$ is actually independent of the choice of connection ∇.

Proposition 36.5. *Let $\pi\colon E \to M$ denote a vector bundle and let ∇_0 and ∇_1 denote two connections on E. Then as elements of $H^2_{\mathrm{dR}}(M)$, we have*

$$[\mathrm{tr}(R^{\nabla_0})] = [\mathrm{tr}(R^{\nabla_1})].$$

Proof. Let $\mathrm{pr}_1\colon M \times [0,1] \to M$ denote the first projection, and consider the pullback bundle $\mathrm{pr}_1^\star E$ over $M \times [0,1]$. Let $\overline{\nabla}_i$ denote the pullback connection $\mathrm{pr}_1^* \nabla_i$. If $\mathrm{pr}_2\colon M \times [0,1] \to [0,1]$ is the second projection, then

$$\nabla := (1 - \mathrm{pr}_2)\overline{\nabla}_0 + \mathrm{pr}_2 \overline{\nabla}_1$$

is a connection on $\mathrm{pr}_1^\star E$. If $\jmath_t\colon M \to M \times [0,1]$ is the map $\jmath_t(x) := (x,t)$ then

$$\jmath_t^\star \nabla = (1-t)\nabla_0 + t\nabla_1,$$

and thus in particular

$$\jmath_0^\star \nabla = \nabla_0, \qquad \jmath_1^\star \nabla = \nabla_1.$$

If R^{∇} denotes the curvature of ∇ and R^{∇_i} denotes the curvature of ∇_i then using Theorem 35.1 we obtain

$$\operatorname{tr}(R^{\nabla_0}) = \jmath_0^\star\big(\operatorname{tr}(R^{\nabla})\big), \qquad \operatorname{tr}(R^{\nabla_1}) = \jmath_1^\star\big(\operatorname{tr}(R^{\nabla})\big).$$

By Proposition 23.16 we obtain

$$[\operatorname{tr}(R^{\nabla_0})] = \jmath_0^\star[\operatorname{tr}(R^{\nabla})] = \jmath_1^\star[\operatorname{tr}(R^{\nabla})] = [\operatorname{tr}(R^{\nabla_1})].$$

This completes the proof. ■

We have thus shown that the trace of the curvature of a connection gives rise to a de Rham cohomology class in the base manifold that depends only on the vector bundle. Amusingly however, this cohomology class is not particularly interesting.

PROPOSITION 36.6. *Let $\pi\colon E \to M$ denote a vector bundle. Then* $[\operatorname{tr}(R^{\nabla})] = 0$.

Oh well, that was a waste.

Not!

The key idea behind characteristic classes is that we can play the same game with any "invariant polynomial", rather than just the trace. We will explore this further next lecture.

The proof of Proposition 36.6 is not particularly hard, but it requires us to introduce another concept, that of a *Riemannian metric*. After connections, this is the second most important idea of the entire course.

DEFINITION 36.7. Let $\pi\colon E \to M$ be a vector bundle. A **Riemannian metric** on E (often shorted to just "a metric on E") is a section $m \in \Gamma(E^* \otimes E^*)$ with the property that for all $x \in M$, the element $m_x \in E_x^* \otimes E_x^* \cong (E \otimes E)^*|_x$ is an inner product on the vector space E_x. We call the pair (E, m) a **Riemannian vector bundle**.

REMARK 36.8. **Warning:** It is common to use the symbol "g" for a metric. Since I like to use g to denote a smooth function, I won't do this, and instead use the (more logical) symbol "m" instead.

In the special case $E = TM$, we say that m is a **Riemannian metric on** M and refer to the pair (M, m) as a **Riemannian manifold**. The field of *Riemannian geometry* is the study of Riemannian metrics on manifold.

REMARK 36.9. **Warning:** Do not confuse a Riemannian metric with a normal metric in the sense of point-set topology. They are *not* the same thing! We will eventually prove that if (M, m) is a Riemannian manifold then the Riemannian metric m induces an actual metric d_m on M, which moreover induces the given topology on M.

We will often use the notation

$$\langle p, q \rangle_x := m_x(p, q)$$

to emphasise that m_x is an inner product. Often we will omit the subscript x and just write $\langle p, q \rangle$, and sometimes we will refer to the entire metric by $\langle \cdot, \cdot \rangle$. Similarly we abbreviate by

$$\|p\|_x := \sqrt{\langle p, p \rangle_x} = \sqrt{m_x(p, p)}$$

the associated norm on E_x, again sometimes omitting the subscript x.

DEFINITION 36.10. Let $\pi_i \colon E_i \to M_i$ be vector bundles equipped with Riemmanian metrics m_i for $i = 1, 2$. Suppose $\varphi \colon M_1 \to M_2$ is a smooth map and $\Phi \colon E_1 \to E_2$ is a vector bundle morphism along φ:

$$\begin{tikzcd} E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\ M_1 \arrow[r, "\varphi"'] & M_2 \end{tikzcd}$$

We say that Φ is an **isometric vector bundle morphism** if

$$m_1|_x(p, q) = m_2|_{\varphi(x)}(\Phi(p), \Phi(q)), \qquad \forall\, x \in M,\ p, q \in E_1|_x.$$

As with connections, every vector bundle admits a Riemannian metric.

PROPOSITION 36.11. *Every vector bundle $\pi \colon E \to M$ admits a Riemannian metric.*

Proof. This is a standard partition of unity argument. Suppose E has rank k. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that there exist a vector bundle chart $\alpha_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to \mathbb{R}^k$ for each $\mathtt{a} \in \mathtt{A}$. Let $\langle \cdot, \cdot \rangle$ denote the standard Euclidean inner product on $\mathbb{R}^k$, and define for $x \in U_{\mathtt{a}}$

$$m_{\mathtt{a}}|_x(p, q) := \langle \alpha_{\mathtt{a}}|_{E_x}(p), \alpha_{\mathtt{a}}|_{E_x}(q) \rangle$$

Then $m_{\mathtt{a}}$ is a Riemannian metric on the trivial bundle $\pi^{-1}(U_{\mathtt{a}}) \to U_{\mathtt{a}}$. To globalise this, let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote a partition of unity subordinate to $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ and extend the local section $\lambda_{\mathtt{a}} m_{\mathtt{a}}$ of $E^* \otimes E^*$ to be defined on all of M by setting it to be zero outside of $U_{\mathtt{a}}$. Then define

$$m := \sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}} m_{\mathtt{a}} \in \Gamma(E^* \otimes E^*).$$

This is a Riemannian metric on E as the sum is finite at every point. ■

We also have:

LEMMA 36.12. *Let $\pi \colon E \to M$ be a vector bundle and suppose m is a Riemannian metric on E. Then around any point $x \in M$ there exists a local frame $\{e_i\}$ for E which is orthonormal with respect to m.*

Proof. Apply the Gram-Schmidt process to an arbitrary local frame. ■

A corollary of this is that we can always reduce the structure group (cf. Remark 13.14) of a vector bundle to the orthogonal group.

COROLLARY 36.13. *If $\pi\colon E \to M$ is a vector bundle of rank k then the structure group G of E may be reduced to $\mathrm{O}(k) \subset \mathrm{GL}(k)$.*

Proof. Lemma 36.12 furnishes the necessary vector bundle charts. ■

Now we relate connections to metrics.

DEFINITION 36.14. Let $\pi\colon E \to M$ be a vector bundle with Riemannian metric m. A connection ∇ on E is said to be a **Riemannian connection with respect to** m if m is a parallel section with respect to the induced connection on $E^* \otimes E^*$.

We often simply write "a Riemannian connection" if the metric m is understood.

PROPOSITION 36.15. *Let $\pi\colon E \to M$ be a vector bundle with Riemannian metric $m = \langle \cdot, \cdot \rangle$. A connection ∇ on E is a Riemannian connection with respect to m if and only if the* **Ricci Identity** *holds:*

$$X \langle s_1, s_2 \rangle = \langle \nabla_X(s_1), s_2 \rangle + \langle s_1, \nabla_X(s_2) \rangle, \qquad \forall\, X \in \mathfrak{X}(M),\ s_1, s_2 \in \Gamma(E). \quad (36.3)$$

Proof. By (both parts of) Problem P.1, if ∇ (also) denotes the induced connection on $E^* \otimes E^*$ then for $X \in \mathfrak{X}(M)$ and $s_1, s_2 \in \Gamma(E)$, we have

$$\nabla_X(m)(s_1, s_2) = X \langle s_1, s_2 \rangle - \langle \nabla_X(s_1), s_2 \rangle - \langle s_1, \nabla_X(s_2) \rangle .$$

Thus $\nabla_X(m) = 0$ if and only if the Ricci Identity (36.3) holds. ■

The Ricci Identity also holds for the pullback of a Riemannian connection.

COROLLARY 36.16. *Let $\pi\colon E \to N$ be a vector bundle with Riemannian metric $m = \langle \cdot, \cdot \rangle$ and let ∇ denote a connection on E which is Riemannian with respect to m. Suppose $\varphi\colon M \to N$ is a smooth map. Then the pullback connection satisfies the Ricci identity too: holds:*

$$X \langle s_1, s_2 \rangle = \langle \nabla_X(s_1), s_2 \rangle + \langle s_1, \nabla_X(s_2) \rangle, \qquad \forall\, X \in \mathfrak{X}(M),\ s_1, s_2 \in \Gamma_\varphi(E),$$

where both sides are smooth functions on M.

Proof. Apply the chain rule (31.7) for covariant derivative operators. ■

On Problem Sheet R you will prove that if ∇ is a Riemannian connection then $\mathrm{Hol}^\nabla(x) \subset \mathrm{O}(E_x, m_x) \subset \mathrm{GL}(E_x)$, for every $x \in M$, where $\mathrm{O}(E_x, m_x)$ denotes the orthogonal transformations with respect to the inner product m_x.

PROPOSITION 36.17. *Let $\pi\colon E \to M$ be a vector bundle with Riemannian metric $m = \langle \cdot, \cdot \rangle$. Then there exists a Riemannian connection with respect to m.*

Proof. The argument is again via a partition of unity. Suppose E has rank k. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that there exist a orthonormal frame $\{e_i^{\mathtt{a}} \mid i = 1, \dots, k\}$ for E over $U_{\mathtt{a}}$. Define a covariant derivative operator on the trivial bundle $\pi^{-1}(U_{\mathtt{a}}) \to U_{\mathtt{a}}$ by

$$\nabla_X^{\mathtt{a}}(s) := \sum_{i=1}^k X \left\langle e_i^{\mathtt{a}}, s \right\rangle e_i^{\mathtt{a}}.$$

Now let $\{\lambda_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ denote a partition of unity subordinate to $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ and extend the local section $\lambda_{\mathtt{a}} \nabla^{\mathtt{a}}$ to be defined on all of M by setting it to be zero outside of $U_{\mathtt{a}}$. Then define

$$\nabla := \sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}} \nabla^{\mathtt{a}}.$$

This is a covariant derivative operator[1] on M. Moreover we claim that ∇ is Riemannian: indeed if $X \in \mathfrak{X}(M)$ and $s_1, s_2 \in \Gamma(E)$ then

$$\begin{aligned} X \left\langle s_1, s_2 \right\rangle &= \sum_{\mathtt{a} \in \mathtt{A}} \sum_{i=1}^k \lambda_{\mathtt{a}} X \left(\left\langle s_1, e_i^{\mathtt{a}} \right\rangle \left\langle e_i^{\mathtt{a}}, s_2 \right\rangle \right) \\ &= \sum_{\mathtt{a} \in \mathtt{A}} \lambda_{\mathtt{a}} \left(\left\langle \nabla_X^{\mathtt{a}}(s_1), s_2 \right\rangle + \left\langle s_1, \nabla_X^{\mathtt{a}}(s_2) \right\rangle \right) \\ &= \left\langle \nabla_X(s_1), s_2 \right\rangle + \left\langle s_1, \nabla_X(s_2) \right\rangle, \end{aligned}$$

where as usual the interchange of summation signs is justified as the sum is locally finite. ∎

We now prove that the curvature tensor of a Riemannian connection is skew-symmetric.

PROPOSITION 36.18. *Let $\pi \colon E \to M$ be a vector bundle with Riemannian metric m, and let ∇ be a Riemannian connection with respect to m. Then for all $X, Y \in \mathfrak{X}(M)$ and $s_1, s_2 \in \Gamma(E)$, one has*

$$\left\langle R^\nabla(X,Y)(s_1), s_2 \right\rangle + \left\langle s_1, R^\nabla(X,Y)(s_2) \right\rangle = 0.$$

Proof. It is sufficient to prove the result in the case $[X, Y] = 0$ since R^∇ is a point operator. Let $s \in \Gamma(E)$. Then by Theorem 35.1 and the Ricci Identity (36.3), we have

$$\begin{aligned} \left\langle R^\nabla(X,Y)(s), s \right\rangle &= \left\langle \nabla_X(\nabla_Y(s)), s \right\rangle - \left\langle \nabla_Y(\nabla_X(s)), s \right\rangle \\ &= X \left\langle \nabla_Y(s), s \right\rangle - \left\langle \nabla_Y(s), \nabla_X(s) \right\rangle - Y \left\langle \nabla_X(s), s \right\rangle + \left\langle \nabla_X(s), \nabla_Y(s) \right\rangle \\ &= \frac{1}{2} \left(XY \left\langle s, s \right\rangle - YX \left\langle s, s \right\rangle \right) \\ &= \frac{1}{2} [X, Y] \left\langle s, s \right\rangle \\ &= 0. \end{aligned}$$

This completes the proof. ∎

[1] If the associated distribution to $\nabla^{\mathtt{a}}$ is $\mathcal{H}^{\mathtt{a}}$ then the distribution $\mathcal{H}$ constructed in Step 2 of Theorem 28.5 has its associated covariant derivative equal to ∇.

If m is a metric on E then we denote by $\mathfrak{o}(E, m) \to M$ the orthogonal algebra bundle of (E, m). The fibre of $\mathfrak{o}(E, m)$ over $x \in M$ is the Lie subalgebra $\mathfrak{o}(E_x, m_x) \subset \mathrm{L}(E_x, E_x)$ of linear maps that are orthogonal with respect to the inner product m_x.

Corollary 36.19. *Let $\pi \colon E \to M$ be a vector bundle of rank k with Riemannian metric m, and let ∇ be a Riemannian connection with respect to m. Then for all $X, Y \in \mathfrak{X}(M)$, the curvature $R^\nabla(X, Y)$ belongs to the orthogonal algebra bundle $\mathfrak{o}(m) \subset \mathrm{Hom}(E, E)$.*

Proof. Proposition 36.18 shows us that $R^\nabla(v, w) \in \mathfrak{o}(E_x, m_x)$ for all $x \in M$ and $v, w \in T_xM$. ∎

We conclude this lecture by using Corollary 36.19 to prove Proposition 36.6.

Proof of Proposition 36.6. It suffices to find a single connection for which $[\mathrm{tr}(R^\nabla)] = 0$. Let m denote any Riemannian metric on E and let ∇ denote any Riemannian connection. Then Proposition 36.15 shows that $R^\nabla(X, Y)$ is skew-symmetric and hence has trace zero. ∎

Characteristic classes and the Chern-Weil homomorphism

In this lecture we construct the characteristic classes of a vector bundle in generality. The formalism is a little daunting, so I urge you to keep in mind the example of the trace from the last lecture. We begin at the level of linear algebra. Let $\mathbb{R}[X_1, \dots, X_k]$ denote the $\mathbb{R}$-algebra of polynomials in k indeterminates X_i. A polynomial $p \in \mathbb{R}[X_1, \dots, X_k]$ is said to be **homogeneous of degree** r if we can write

$$p(X_1, \dots, X_k) = \sum c_{i_1 \cdots i_r} X_{i_1} \cdots X_{i_r}.$$

where the sum is over all k^r tuples $(i_1, \dots, i_r)$ such that $1 \leq i_j \leq k$ for each i_j. We may without loss of generality always assume that the coefficients $c_{i_1 \cdots i_r}$ are *symmetric*[1] in the indices $i_1, \dots, i_r$.

Definition 37.1. Let V be a vector space of dimension k. A **homogeneous polynomial of degree** r on V is a map

$$\phi \colon V \to \mathbb{R}$$

such that for every basis $\{e^i\}$ of the dual space V^* then there exists a unique homogeneous $p \in \mathbb{R}[X_1, \dots, X_k]$ such that

$$\phi(v) = p(e^1, \dots, e^k)(v) = \sum c_{i_1 \cdots i_r} e^{i_1}(v) \cdots e^{i_r}(v). \tag{37.1}$$

It is easy to see that this property is independent of the choice of basis in the sense that we could replace "for every basis" with "there exists a basis".

Definition 37.2. Let V be a vector space. We let $P_r(V)$ denote the set of all homogeneous polynomials of degree r, and $P(V) = \bigoplus_{r \geq 0} P_r(V)$. Then $P(V)$ is an algebra under the usual pointwise product of functions.

Definition 37.3. Let V be a vector space, and suppose $\phi \in P_r(V)$. The **polarisation** of ϕ is the tensor $\mathrm{polar}(\phi) \in T^{0,r}(V) \cong \mathrm{Mult}_{r,0}(V)$ (cf. Proposition 15.9) defined by

$$\mathrm{polar}(\phi) = \sum c_{i_1 \cdots i_r} e^{i_1} \otimes \cdots \otimes e^{i_r}.$$

where $\{e^i\}$ is some basis of V^* and the coefficients $c_{i_1 \cdots i_r}$ are determined by (37.1).

As above, it is easy to see that definition of $\mathrm{polar}(\phi)$ does not depend on the choice of basis $\{e^i\}$ of V^*. Since we assumed that the original coefficients $c_{i_1 \cdots i_r}$ were symmetric in the indices i_j, the tensor $\mathrm{polar}(\phi)$ is actually a *symmetric* tensor in the following sense.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] **Warning:** Do not confuse this with requiring the polynomial p itself to be a *symmetric polynomial*!

DEFINITION 37.4. Let V be a vector space. A tensor $S \in T^{0,r}(V)$ is said to be **symmetric** if we can write

$$S = \sum c_{i_1 \cdots i_r} e^{i_1} \otimes \cdots \otimes e^{i_r}$$

for some basis $\{e^i\}$ such that the coefficients $c_{i_1\cdots i_r}$ are symmetric in the indices $i_1, \dots, i_r$. We denote by $\mathfrak{S}_r(V) \subset T^{0,r}(V)$ the set of symmetric tensors and by $\mathfrak{S}(V) = \bigoplus_{r\geq 0} \mathfrak{S}_r(V)$. Then $\mathfrak{S}(V)$ is an algebra under the product

$$(S \cdot T)(v_1, \dots, v_{r+s}) \coloneqq \frac{1}{(r+s)!} \sum_{\varrho \in \mathfrak{S}_{r+s}} S\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big)\, T\big(v_{\varrho(r+1)}, \dots, v_{\varrho(r+s)}\big)$$

for $S \in \mathfrak{S}_r(V)$ and $T \in \mathfrak{S}_s(V)$ (compare Lemma 19.4).

Thus we can think of polarisation as defining a degree-preserving algebra homomorphism

$$\mathrm{polar}\colon P(V) \to \mathfrak{S}(V).$$

Actually polar is an isomorphism, since an explicit inverse is given by

$$\mathrm{polar}^{-1}(S)(v) \coloneqq S(v, \dots, v),$$

as is easy to check.

DEFINITION 37.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$. A homogeneous polynomial $\phi\colon \mathfrak{g} \to \mathbb{R}$ is said to be **invariant** if

$$\phi(\mathrm{Ad}_a(v)) = \phi(v), \qquad \forall\, a \in G,\ v \in \mathfrak{g},$$

where $\mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$ was defined in Definition 10.21. We denote by $P_G(\mathfrak{g}) \subset P(\mathfrak{g})$ the subalgebra of all invariant polynomials.

EXAMPLE 37.6. Take $G = \mathrm{GL}(k)$. Then the adjoint action on $\mathfrak{gl}(k)$ is simply given by conjugation:

$$\mathrm{Ad}_T\colon \mathfrak{gl}(k) \to \mathfrak{gl}(k), \qquad \mathrm{Ad}_T(A) = TAT^{-1}.$$

and thus a polynomial $\phi\colon \mathfrak{gl}(k) \to \mathbb{R}$ is invariant if

$$\phi(TAT^{-1}) = \phi(A), \qquad \forall\, T \in \mathrm{GL}(k),\ A \in \mathfrak{gl}(k).$$

The following lemma is elementary linear algebra.

LEMMA 37.7. *The coefficients ϕ_i of the characteristic polynomial*

$$\det(tI + A) = \sum_{i=0}^{k} \phi_i(A) t^{k-i}, \qquad A \in \mathfrak{gl}(k),\ t \in \mathbb{R} \tag{37.2}$$

are invariant polynomials of degree i on $\mathfrak{gl}(k)$. In particular, the trace and determinant are invariant polynomials:

$$1 = \phi_0(A), \qquad \mathrm{tr}\, A = \phi_1(A), \qquad \det A = \phi_k(A)$$

In fact, polynomials of this form generate the entire algebra $P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$.

THEOREM 37.8. *The space $P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ is generated as an $\mathbb{R}$-algebra by the coefficients ϕ_i of the characteristic polynomial* (37.2).

Theorem 37.8 is a corollary of the so-called "Fundamental Theorem of Symmetric Polynomials[2]" which states that any symmetric polynomial can be written as a polynomial in the "elementary symmetric polynomials". This result is not hard to prove, but we will not carry it out here, since it has nothing to do with Differential Geometry and would take us too far afield.

DEFINITION 37.9. Let G be a Lie group with Lie algebra $\mathfrak{g}$. We say a symmetric tensor $S \in \mathfrak{S}_r(\mathfrak{g})$ is **invariant** if

$$S(\mathrm{Ad}_a(v_1), \dots, \mathrm{Ad}_a(v_r)) = S(v_1, \dots, v_r)$$

for all $a \in G$ and $v_i \in \mathfrak{g}$. We let $\mathfrak{S}_G(\mathfrak{g}) \subset \mathfrak{S}(\mathfrak{g})$ denote the invariant symmetric tensors.

If $\phi \in P_G(\mathfrak{g})$ then $\mathrm{polar}(\phi) \in \mathfrak{S}_G(\mathfrak{g})$, and we obtain

LEMMA 37.10. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then polarisation restricts to define an algebra isomorphism* $\mathrm{polar}\colon P_G(\mathfrak{g}) \to \mathfrak{S}_G(\mathfrak{g})$.

Let us now proceed to vector bundles. Suppose $\pi\colon E \to M$ is a vector bundle. Let us abbreviate

$$\mathrm{Hom}^r(E, E) := T^{r,0}(\mathrm{Hom}(E, E))$$

for $r \geq 0$. Thus the fibre of $\mathrm{Hom}^r(E, E)$ over $x \in M$ is

$$\underbrace{\mathfrak{gl}(E_x) \otimes \cdots \otimes \mathfrak{gl}(E_x)}_{r},$$

and $\mathrm{Hom}^1(E, E) = \mathrm{Hom}(E, E)$. The following lemma is the analogue of Proposition 36.3 to this new more complicated setting.

LEMMA 37.11. *Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Let $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ denote an invariant polynomial of degree r. Then ϕ induces a parallel section Φ of the dual bundle* $(\mathrm{Hom}^r(E, E))^*$.

Proof. Let $S = \mathrm{polar}(\phi)$ denote the polarisation of ϕ. For each $x \in M$, choose an element $F \in \mathrm{Fr}(E_x)$, i.e. a linear isomorphism $F\colon \mathbb{R}^k \to E_x$. We define

$$\Phi_x(T_1 \otimes \cdots \otimes T_r) := S\big(F^{-1} \circ T_1 \circ F, \dots, F^{-1} \circ T_r \circ F\big)$$

for $T_i \in \mathfrak{gl}(E_x) = \mathrm{L}(E_x, E_x)$. This definition is independent of the choice of F, since if $\tilde{F}\colon \mathbb{R}^k \to E_x$ was another element of $\mathrm{Fr}(E_x)$ then $L := F^{-1}\tilde{F} \in \mathrm{GL}(k)$, and

$$\begin{aligned} S\big(\tilde{F}^{-1} \circ T_1 \circ \tilde{F}, \dots, \tilde{F}^{-1} \circ T_r \circ \tilde{F}\big) &= S\big(L^{-1}F^{-1} \circ T_1 \circ FL, \dots, L^{-1}F^{-1} \circ T_r \circ FL\big) \\ &= S\big(F^{-1} \circ T_1 \circ F, \dots, F^{-1} \circ T_r \circ F\big) \end{aligned}$$

since S is invariant. The proof that Φ is parallel is identical to the proof of Proposition 36.3. ∎

[2] See for instance this Wikipedia entry.

Let us now consider differential forms with values in $\mathrm{Hom}^r(E,E)$. The tensor product gives us another way to multiply such forms together:

Definition 37.12. Let $A \in \Omega^h(M, \mathrm{Hom}^r(E,E))$ and $B \in \Omega^l(M, \mathrm{Hom}^s(E,E))$. We define an element $A \otimes B \in \Omega^{h+l}(M, \mathrm{Hom}^{r+s}(E,E))$ by wedging together the $\Omega(M)$ factors and tensoring the $\mathrm{Hom}(E,E)$ factors. Explicitly, on decomposable elements

$$A = \omega \otimes \big(T_1 \otimes \cdots \otimes T_r\big), \qquad B = \vartheta \otimes (S_1 \otimes \cdots \otimes S_s)$$

for $\omega \in \Omega^h(M)$, $\vartheta \in \Omega^l(M)$ and $T_i, S_j \in \Gamma(\mathrm{Hom}(E,E))$, we define

$$A \otimes B \coloneqq (\omega \wedge \vartheta) \otimes \big(T_1 \otimes \cdots \otimes T_r \otimes S_1 \otimes \cdots \otimes S_s\big).$$

(♣) Remark 37.13. If $E \to M$ is an algebra bundle in the sense of Remark 15.27, where the multiplication is given by $\beta \colon E \otimes E \to E$, then the module $\Omega(M,E)$ of E-valued differential forms is itself an algebra with the multiplication

$$\tilde{\beta} \colon \Omega(M,E) \times \Omega(M,E) \to \Omega(M,E)$$

via

$$\tilde{\beta}(\omega, \vartheta)_x(v_1, \dots, v_{r+s}) \coloneqq \frac{1}{r!s!} \sum_{\varrho \in \mathfrak{S}_{r+s}} \mathrm{sgn}(\varrho)\beta\left(\omega_x\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big) \otimes \vartheta_x\big(v_{\varrho(r+1)}, \dots v_{\varrho(r+s)}\big)\right).$$

(This is just the bundle-valued version of Definition 26.5.) Now, if we were to allow ourselves infinite-dimensional bundles, then Definition 37.12 would fit into this framework (and thus seem less ad-hoc). Namely, if we set

$$\mathrm{Hom}^\infty(E,E) \coloneqq \bigoplus_{r \geq 0} \mathrm{Hom}^r(E,E),$$

then $\mathrm{Hom}^\infty(E,E) \to M$ can be seen as an infinite-dimensional algebra bundle over M, cf. Remark 15.19, where β is simply $\otimes$. In this case Definition 37.12 is simply the natural algebra structure on $\mathrm{Hom}^\infty(E,E)$-valued forms coming from the algebra structure on $\mathrm{Hom}^\infty(E,E)$.

Just as in the discussion before Lemma 36.2, if we are given a section Φ of the dual bundle $(\mathrm{Hom}^r(E,E))^*$ then given an element $A \in \Omega^k(M, \mathrm{Hom}^r(E,E))$ we can feed it to Φ to obtain a normal differential k-form $\Phi(A)$. We then have the following generalisation of Lemma 36.2.

Lemma 37.14. *Suppose $\pi \colon E \to M$ is a vector bundle of rank k with connection ∇, and suppose $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ is an invariant polynomial of degree r. Let Φ denote the induced parallel section of the dual bundle $(\mathrm{Hom}^r(E,E))^*$. Suppose $A_i \in \mathcal{A}^{h_i}(M,E)$ for $i = 1, \dots, r$. Then denoting by ∇^{Hom} the induced connection on $\mathrm{Hom}(E,E)$, we have*

$$d\big(\Phi(A_1 \otimes \cdots \otimes A_r)\big) = \Phi\left(\sum_{j=1}^r (-1)^{h_1 + \cdots + h_{j-1}} A_1 \otimes \cdots \otimes d^{\nabla^{\mathrm{Hom}}} A_j \otimes \cdots \otimes A_r\right).$$

Proof. Use Remark 35.7 together with the argument from the proof of Lemma 36.2. ∎

We are now ready to state and prove our main result.

THEOREM 37.15 (The Chern-Weil Theorem). *Let $\pi\colon E \to M$ denote a vector bundle of rank k and let $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ have degree r. Then:*

(i) *If ∇ is a connection on E and Φ is the induced parallel section of $(\mathrm{Hom}^r(E,E))^*$ from Lemma 37.11 then the $2r$-form*

$$\phi(\nabla) := \Phi\left(\underbrace{R^\nabla \otimes \cdots \otimes R^\nabla}_{r}\right)$$

is closed.

(ii) *The cohomology class $[\phi(\nabla)] \in H^{2r}_{\mathrm{dR}}(M)$ is independent of ∇.*

(iii) *The map*[3]

$$\mathrm{CW}_E\colon P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)) \to H^*_{\mathrm{dR}}(M) := \bigoplus_{r\geq 0} H^r_{\mathrm{dR}}(M),$$

given by

$$\mathrm{CW}_E(\phi) := [\phi(\nabla)]$$

is an algebra homomorphism.

Proof. The proof of (i) is the same as Corollary 36.4, and uses the Bianchi Identity (Theorem 36.1) and Lemma 37.14:

$$\begin{aligned} d\big(\Phi(R^\nabla \otimes \cdots \otimes R^\nabla)\big) &= \Phi\left(\sum_i (-1)^{2+\cdots+2} R^\nabla \otimes \cdots \otimes \underbrace{d^{\nabla^{\mathrm{Hom}}}(R^\nabla)}_{=0} \otimes \cdots \otimes R^\nabla\right) \\ &= \Phi\left(R^\nabla \otimes \cdots \otimes 0 \otimes \cdots \otimes R^\nabla\right) \\ &= 0. \end{aligned}$$

The proof of (ii) is identical to that of Proposition 36.5: using the notation from that proof one has

$$\phi(\nabla_0) = \jmath_0^\star\big(\phi(\nabla)\big), \qquad \phi(\nabla_1) = \jmath_1^\star\big(\phi(\nabla)\big),$$

and hence by Proposition 23.16 again

$$[\phi(\nabla_0)] = \jmath_0^\star[\phi(\nabla)] = \jmath_1^\star[\phi(\nabla)] = [\phi(\nabla_1)].$$

The proof of (iii) is on Problem Sheet R. ■

We now prove that CW behaves nicely with respect to pullbacks.

[3] Here "CW" stands for "Chern-Weil" (don't confuse this with a CW complex in algebraic topology!)

Proposition 37.16. *Let $\pi\colon E \to N$ denote a vector bundle of rank k and let $\varphi\colon M \to N$ denote a smooth map. Then the following diagram commutes:*

$$\begin{tikzcd} & H^*_{\mathrm{dR}}(N) \arrow[dd, "\varphi^\star"] \\ P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)) \arrow[ur, "\mathrm{CW}_E"] \arrow[dr, "\mathrm{CW}_{\varphi^\star E}"'] & \\ & H^*_{\mathrm{dR}}(M) \end{tikzcd}$$

Proof. This follows from the equality $R^{\varphi^\star\nabla} = \varphi^\star(R^\nabla)$, which in turns follows from Proposition 34.11. ∎

Definition 37.17. Let $\pi\colon E \to M$ be a vector bundle of rank k. We call an element $\mathrm{CW}_E(\phi) \in H^*_{\mathrm{dR}}(M)$ a **characteristic class** of E. The map $\mathrm{CW}_E\colon P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)) \to H^*_{\mathrm{dR}}(M)$ is called the **Chern-Weil homomorphism**.

It follows from Proposition 37.16 that isomorphic vector bundles have the same characteristic classes. Turning this on its head, if E_1 and E_2 are any two vector bundles, then in order to show that E_1 and E_2 are *not* isomorphic, it suffices to find a single characteristic class which takes different values on E_1 and E_2.

The following generalisation of Proposition 36.6 is on Problem Sheet R.

Proposition 37.18. *If $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ is an invariant homogeneous polynomial of odd degree $2r+1$ then $\mathrm{CW}_E(\phi) = 0$ for any vector bundle of rank k.*

Combining this with Theorem 37.8 tells us that the ring of characteristic classes on M has two sets of generators:

(i) the trace polynomials of even degree: $\left[\operatorname{tr}\left(\underbrace{R^\nabla \otimes \cdots \otimes R^\nabla}_{2r}\right)\right]$,

(ii) the coefficients ϕ_{2r} of even degree of the characteristic polynomial (37.2).

Let us focus on the second case.

Definition 37.19. Let $\pi\colon E \to M$ be a vector bundle of rank k. We define the **rth Pontryagin class** of E to be

$$p_r(E) := \left[\phi_{2r}\left(\frac{i}{2\pi}\nabla\right)\right] \in H^{4r}_{\mathrm{dR}}(M).$$

The factor of $\frac{i}{2\pi}$ is not too important, it is just there to make certain other formulae prettier (note in particular that as ϕ_{2r} is homogeneous of degree $2r$, the factor of i disappears when fed to ϕ). Note $p_r(E) = 0$ if $r > \lfloor \frac{k}{2} \rfloor$. It is also formally useful to define $p_0(E) := 1$, where $1 \in H^0_{\mathrm{dR}}(M) = \mathbb{R}$ is the cohomology class containing the constant function 1. In order to state the next result, note that the wedge product on differential forms is also well defined on the level of de Rham cohomology.

DEFINITION 37.20. Let M be a smooth manifold. If $[\omega] \in H^r_{\mathrm{dR}}(M)$ and $\vartheta \in H^s_{\mathrm{dR}}(M)$ are two de Rham cohomology classes represented by closed forms ω and ϑ respectively, then we define

$$[\omega] \wedge [\vartheta] := [\omega \wedge \vartheta] \in H^{r+s}_{\mathrm{dR}}(M). \tag{37.3}$$

This is well defined as (a) the $(r+s)$-form $\omega \wedge \vartheta$ is closed, since $d(\omega \wedge \vartheta) = d\omega \wedge \vartheta + (-1)^r \omega \wedge d\vartheta = 0$, and (b) $[\omega \wedge \vartheta]$ is independent of the choice of representatives ω and ϑ, since if ω_1 and ϑ_1 were two more representatives (meaning that $\omega - \omega_1$ and $\vartheta - \vartheta_1$ are both exact) then the same formula shows that $\omega \wedge \vartheta - \omega_1 \wedge \vartheta_1$ is exact.

The wedge product (37.3) turns the total cohomology $H^*_{\mathrm{dR}}(M)$ into a graded ring. On Problem Sheet R you will prove[4]:

PROPOSITION 37.21 (The Whitney Product Formula). *If E_1 and E_2 are vector bundles over M then*

$$p_r(E_1 \oplus E_2) = \sum_{i=0}^{r} p_i(E_1) \wedge p_{r-i}(E_2).$$

We conclude this lecture with a sample application.

PROPOSITION 37.22. *Suppose M^n is a compact manifold which can be embedded in $\mathbb{R}^{n+1}$. Then $p_r(TM) = 0$ for $r > 0$.*

Proof. If M embeds in $\mathbb{R}^{n+1}$ then the normal bundle $\mathrm{Norm}(M)$ from Definition 6.8 is a one-dimensional vector bundle and hence has $p_r(\mathrm{Norm}(M)) = 0$ for $r > 0$. We have a vector bundle isomorphism:

$$T\mathbb{R}^{n+1}|_M \cong TM \oplus \mathrm{Norm}(M).$$

Proposition 37.16 applied to the embedding $M \hookrightarrow \mathbb{R}^{n+1}$ tells us that $p_r(T\mathbb{R}^{n+1}|_M) = 0$ for $r > 0$. Thus the Whitney Product Formula (Proposition 37.21) implies that $p_r(TM) = 0$ for $r > 0$. This completes the proof. ∎

Of course, the usefulness of Proposition 37.22 depends on our ability to compute the Pontryagin classes! But this is a topic best suited for a course on Algebraic Topology. We just state here one result.

COROLLARY 37.23. *$\mathbb{C}P^2$ does not embed in $\mathbb{R}^5$.*

(♣) *Proof.* We can think of $\mathbb{C}P^2$ as a compact manifold of (real) dimension four. One can show that the class $p_1(T\mathbb{C}P^2) \in H^4_{\mathrm{dR}}(\mathbb{C}P^2)$ is of the form $3c^2$, where $c \in H^2_{\mathrm{dR}}(\mathbb{C}P^2)$ is a generator (and thus in particular is non-zero). ∎

[4] For those of you who are familiar with Algebraic Topology: the statement would be more complicated if one worked with (singular) cohomology with coefficients in $\mathbb{Z}$, since then one would need to worry about 2-torsion elements.

Connections on principal bundles

In this lecture we look at connections on principal bundles. The definition of a preconnection is the same as Definition 28.2 (indeed, we defined preconnections for arbitrary fibre bundles, which thus includes principal bundles as a special case). Just as with vector bundles, a connection on a principal bundle is a preconnection which satisfies an additional condition. Recall that if $\pi\colon P \to M$ is a principal G-bundle we denote by $r_a\colon P \to P$ the right G-action

$$r_a(p) := p \cdot a, \qquad p \in P,\ a \in G.$$

DEFINITION 38.1. Let $\pi\colon P \to M$ be a principal G-bundle. A **connection** on P is a preconnection $\mathcal{H}$ which satisfies

$$Dr_a(p)[\mathcal{H}_p] = \mathcal{H}_{p\cdot a}, \qquad \forall\, p \in P,\ a \in G. \tag{38.1}$$

As before, given a connection $\mathcal{H}$ we denote by

$$\zeta = \zeta^{\mathtt{H}} + \zeta^{\mathtt{V}}$$

the horizontal-vertical splitting of a tangent vector $\zeta \in TP$. The condition (38.1) implies that

$$Dr_a(p)[\zeta^{\mathtt{H}}] = \left(Dr_a(p)[\zeta]\right)^{\mathtt{H}}, \qquad \forall\, \zeta \in T_pP,\ a \in G. \tag{38.2}$$

We will shortly investigate the relationship between connections on principal bundles and connections on vector bundles. Before doing so, however, we look at parallel transport systems in principal bundles. The following definition is easier to remember if you follow the general mantra that you simply take the vector bundle version and replace "linear" with "equivariant" at every opportunity.

DEFINITION 38.2. Let $\pi\colon P \to M$ be a principal G-bundle. A **parallel transport system** $\mathbb{P}$ on P assigns to every point $p \in P$ and every curve[1] $\gamma\colon [r_1, r_2] \to M$ with $\gamma(r_1) = \pi(p)$, a unique section $\mathbb{P}_\gamma(p) \in \Gamma_\gamma(P)$ with initial condition p, i.e. such that $\mathbb{P}_\gamma(p)(r_1) = p$. One calls $\mathbb{P}_\gamma(p)$ the **parallel lift** of γ starting at p. This association should satisfy the following four axioms:

(i) **(Equivariance)**: For every smooth curve $\gamma\colon [r_1, r_2] \to M$ the map

$$\widehat{\mathbb{P}}_\gamma\colon P_{\gamma(r_1)} \to P_{\gamma(r_2)}, \qquad \widehat{\mathbb{P}}_\gamma(p) := \mathbb{P}_\gamma(p)(r_2)$$

is a diffeomorphism which is equivariant with respect to the G-action:

$$\widehat{\mathbb{P}}_\gamma(p \cdot a) = \widehat{\mathbb{P}}_\gamma(p) \cdot a, \qquad \forall\, a \in G.$$

$$\widehat{\mathbb{P}}_\gamma^{-1} = \widehat{\mathbb{P}}_{\gamma^-}$$

where $\gamma^-\colon [r_1, r_2] \to M$ is the reverse curve $t \mapsto \gamma(r_1 - t + r_2)$.

[1] Oh dear, there are not enough letters in the alphabet! I cannot use $[a, b]$ for an interval since a and b are used to denote elements of G. *Screams internally*

(ii) **(Independence of parametrisation)**: If $\gamma\colon [r_1, r_2] \to M$ is a smooth curve and $h\colon [r_3, r_4] \to [r_1, r_2]$ is a diffeomorphism such that $h(r_3) = r_1$ and $h(r_4) = r_2$ then for every point $p \in P_{\gamma(r_1)}$ and every $t \in [r_3, r_4]$, we have

$$\mathbb{P}_{\gamma \circ h}(p)(t) = \mathbb{P}_{\gamma}(p)(h(t)).$$

(iii) **(Smooth dependence on initial conditions)**: The section $\mathbb{P}_\gamma(p)$ depends smoothly on both γ and p.

(iv) **(Initial uniqueness)**: Suppose $\gamma, \delta\colon [r_1, r_2] \to M$ are two curves such that $\gamma(r_1) = \delta(r_1)$ and $\gamma'(r_1) = \delta'(r_1)$. Then for each $p \in P_{\gamma(r_1)}$, the two curves $t \mapsto \mathbb{P}_\gamma(p)(t)$ and $t \mapsto \mathbb{P}_\delta(p)(t)$ have the same initial tangent vector:

$$\frac{d}{dt}\Big|_{t=r_1} \mathbb{P}_\gamma(p)(t) = \frac{d}{dt}\Big|_{t=r_1} \mathbb{P}_\delta(p)(t)$$

As you can guess, connections on principal bundles are the equivalent to parallel transport systems.

THEOREM 38.3. *Let $\pi\colon P \to M$ be a principal G-bundle. Then a connection on P (in the sense of Definition 38.1) determines and is uniquely determined by a parallel transport system on P (in the sense of Definition 38.2).*

The proof of Theorem 38.3 proceeds analogously to Theorem 30.1 and Theorem 30.2, and you will be pleased to learn that I will spare you the gory details. Let us now explain how connections on principal bundles are related to connections on vector bundles.

For this we recall how associated bundles work: Let $\pi\colon P \to M$ be a principal bundle. Suppose[2] $\rho\colon G \to \mathrm{GL}(V)$ is a smooth[3] effective representation of G on a vector space V. Let us write $\rho(P)$ for the associated vector bundle $P \times_G V$ over M. The salient properties we need to recall from Theorem 25.3 and Theorem 26.17 are as follows:

- An element of $\rho(P)$ is written $[p, v]$ where $p \in P$ and $v \in V$. This is the equivalence class of the pair $(p, v) \in P \times V$ under the equivalence relation

$$(p \cdot a, v) \sim (p, \rho_a(v))$$

- The footpoint map $\wp\colon \rho(P) \to M$ is given by $\wp([p, v]) := \pi(p)$.
- For any $p \in P$, the map $L_p\colon V \to \rho(P)|_{\pi(p)}$ given by $v \mapsto [p, v]$ is a linear isomorphism. Thus for $x \in M$ the vector space structure on $\rho(P)|_x$ is given by

$$[p, v] + r[p, w] \overset{\text{def}}{=} L_p(v + rw) = [p, v + rw]. \tag{38.3}$$

- The map $\Pi\colon P \times V \to \rho(P)$ given by $\Pi(p, v) := [p, v]$ exhibits $P \times V$ as another principal G-bundle over $\rho(P)$.

[2]Since I have been using μ to denote scalar multiplication on a vector bundle, I will switch notation from Lecture 25 and use ρ to denote a representation of a Lie group on a vector space.

[3]The smoothness is actually automatic provided ρ is a continuous group homomorphism (cf. Problem F.6).

- There is a bijective correspondence between horizontal G-equivariant V-valued forms on P and $\rho(P)$-valued forms on M (this is Theorem 26.17). This correspondence is written:

$$\begin{array}{rcl} \omega \in \Omega^k_G(P, V) & \mapsto & \omega^\flat \in \Omega^k(M, \rho(P)), \\ \xi \in \Omega^k(M, \rho(P)) & \mapsto & \xi^\sharp \in \Omega^k_G(P, V). \end{array}$$

In particular, a section $s \in \Gamma(\rho(P))$ determines and is uniquely determined by a smooth equivariant function $f\colon P \to V$ via

$$s(x) = [p, f(p)], \qquad \text{for any } p \in P_x. \tag{38.4}$$

REMARK 38.4. Fix a smooth manifold M. If E is a vector bundle of rank k then its frame bundle $\mathrm{Fr}(E)$ is a principal $\mathrm{GL}(k)$-bundle. Moreover if ρ is the canonical representation of $\mathrm{GL}(k)$ on $\mathbb{R}^k$ then $E \cong \rho(\mathrm{Fr}(E))$, and hence *there is a 1-1 correspondence between vector bundles of rank k over M and principal* $\mathrm{GL}(k)$*-bundles*. However any Lie group can be the fibre of a principal bundle over M (not just $\mathrm{GL}(k)$) and hence *principal bundles are more general than vector bundles*.

THEOREM 38.5. *Let $\pi\colon P \to M$ be a principal bundle. Suppose $\rho\colon G \to \mathrm{GL}(V)$ is a smooth representation of G on a vector space V, and let $E := \rho(P)$ denote the associated bundle. A connection $\mathcal{H}$ on P (in the principal bundle sense) induces a connection $\mathcal{H}_E$ on E (in the vector bundle sense).*

Proof. Although not strictly necessary, we will give three proofs, one from the point of view of parallel transport, one from the point of view of distributions, and one from the point of view of covariant derivatives.

- *Proof using parallel transport:* Let $\gamma\colon [0,1] \to M$ be a smooth curve in M, and suppose $c \in \Gamma_\gamma(P)$ is a section along γ. Then for any fixed $v \in V$, $t \mapsto \Pi(c(t), v)$ is a section of E along γ (not every section of E along γ is of this form though). We define a parallel transport system $\mathbb{P}^E$ on E by declaring that a section $\tilde{c}$ of E along γ is parallel if and only if $\tilde{c} = \Pi(c, v)$ for c a parallel section of P along γ. In terms of the hat maps, this means:

$$\widehat{\mathbb{P}}^E_\gamma[p, v] := \left[\widehat{\mathbb{P}}_\gamma(p), v\right].$$

This is well defined because $\widehat{\mathbb{P}}_\gamma$ is equivariant. Indeed, if $(q, w) \in P \times V$ is another representative of $[p, v]$ then there exists $a \in G$ such that $q = p \cdot a$ and $w = \rho_{a^{-1}}(v)$. Then

$$\begin{aligned} \left[\widehat{\mathbb{P}}_\gamma(q), w\right] &= \left[\widehat{\mathbb{P}}_\gamma(p \cdot a), w\right] \\ &= \left[\widehat{\mathbb{P}}_\gamma(p) \cdot a, w\right] \\ &= \left[\widehat{\mathbb{P}}_\gamma(p), \rho_a(w)\right] \\ &= \left[\widehat{\mathbb{P}}_\gamma(p), v\right]. \end{aligned}$$

All the axioms for $\mathbb{P}^E$ follow from those of $\mathbb{P}$. For instance, to see that $\widehat{\mathbb{P}}^E_\gamma$ is a linear map we simply observe by (38.3) that for $r \in \mathbb{R}$ and $v, w \in V$:

$$\begin{aligned} \widehat{\mathbb{P}}^E_\gamma\left([p,v] + r[p,w]\right) &= \widehat{\mathbb{P}}^E_\gamma[p, v+rw] \\ &= \left[\widehat{\mathbb{P}}_\gamma(p), v + rw\right] \\ &= \left[\widehat{\mathbb{P}}_\gamma(p), v\right] + r\left[\widehat{\mathbb{P}}_\gamma(p), w\right] \\ &= \widehat{\mathbb{P}}^E_\gamma[p,v] + r\widehat{\mathbb{P}}^E_\gamma[p,w]. \end{aligned}$$

- *Proof using distributions:* Alternatively, in terms of distributions, we define

  $$\mathcal{H}_E|_{[p,v]} := D\Pi(p,v)[\mathcal{H}_p \times \{0\}].$$

  It is clear this defines a preconnection on E. Let $\mu_r \colon E \to E$ denote the scalar multiplication $\mu_r[p,v] := [p, rv]$. Then

  $$\mu_r \circ \Pi(p,v) = [p, rv] = \Pi(p, rv),$$

  and hence

  $$\begin{aligned} D\mu_r[p,v]\left[\mathcal{H}_E|_{[p,v]}\right] &= D\mu_r[p,v] \circ D\Pi(p,v)\left[\mathcal{H}_p \times \{0\}\right] \\ &= D(\mu_r \circ \Pi)(p,v)\left[\mathcal{H}_p \times \{0\}\right] \\ &= D\Pi(p,rv)\left[\mathcal{H}_p \times \{0\}\right] \\ &= \mathcal{H}_E|_{[p,rv]}. \end{aligned}$$

- *Proof using covariant derivatives:* This is arguably the most interesting proof. Firstly, if $\omega \in \Omega^k(P, V)$ is any V-valued k-form on P, we define the **horizontal component** of ω (with respect to $\mathcal{H}$) to be the form $\omega^{\mathtt{H}} \in \Omega^k(P,V)$ given by

  $$\omega^{\mathtt{H}}_p(\zeta_1, \dots, \zeta_k) := \omega_p(\zeta_1^{\mathtt{H}}, \dots, \zeta_k^{\mathtt{H}}),$$

  where as usual $\zeta^{\mathtt{H}}$ denotes the horizontal component of ζ. Then $\omega^{\mathtt{H}}$ is a horizontal vector-valued form. Now we claim:

  $$\omega \in \Omega^k_G(P,V) \qquad \Rightarrow \qquad (d\omega)^{\mathtt{H}} \in \Omega^{k+1}_G(P,V). \tag{38.5}$$

  We split the proof of (38.5) into two parts:

  (i) If $\omega \in \Omega^k(P,V)$ is G-equivariant then so is $d\omega$.

  (ii) If $\omega \in \Omega^k(P,V)$ is G-equivariant then so is $\omega^{\mathtt{H}}$.

  To prove (i), fix $a \in G$. Then

  $$\begin{aligned} r_a^\star(d\omega) &= dr_a^\star(\omega) \\ &= d\rho_{a^{-1}}(\omega) \\ &= \rho_{a^{-1}}(d\omega) \end{aligned}$$

where the first line used Lemma 26.9 and the last line used the fact that $\rho_{a^{-1}}\colon V \to V$ is a linear map. Next, to prove (ii), we take $a \in G$, $p \in P$ and $\zeta_1, \dots, \zeta_k \in T_pP$ and compute

$$\begin{aligned}
(r_a^\star(\omega^{\mathtt{H}}))_p(\zeta_1, \dots, \zeta_k) &= \omega^{\mathtt{H}}_{p\cdot a}\big(Dr_a(p)[\zeta_1], \dots, Dr_a(p)[\zeta_k]\big) \\
&\overset{(\dagger)}{=} \omega_{p\cdot a}\left(Dr_a(p)[\zeta_1^{\mathtt{H}}], \dots, Dr_a(p)[\zeta_k^{\mathtt{H}}]\right) \\
&= (r_a^\star(\omega))_p(\zeta_1^{\mathtt{H}}, \dots \zeta_k^{\mathtt{H}}) \\
&= \rho_{a^{-1}}\left(\omega_p(\zeta_1^{\mathtt{H}}, \dots \zeta_k^{\mathtt{H}})\right) \\
&= \rho_{a^{-1}}\left(\omega_p^{\mathtt{H}}(\zeta_1, \dots, \zeta_k)\right),
\end{aligned}$$

where $(\dagger)$ used (38.2). Thus (38.5) is proved. We now use (38.5) to define an exterior covariant differential $d^\nabla : \Omega^k(M, E) \to \Omega^{k+1}(M, E)$ by

$$d^\nabla \xi := \Big(\big(d(\xi^\sharp)\big)^{\mathtt{H}}\Big)^\flat \tag{38.6}$$

In particular, for $k = 0$, if $s \in \Gamma(E)$ then

$$\nabla s := \big((df)^{\mathtt{H}}\big)^\flat,$$

where $f\colon P \to V$ is the function from (38.4). All the axioms of a covariant derivative operator are easy to check.

This completes the proof (three times over). *Wholesome exercise:* Check that all three proofs give rise to the same connection on E. ■

If $\pi\colon E \to M$ is a vector bundle of rank k, then its frame bundle $\hat{\pi}\colon \mathrm{Fr}(E) \to M$ is a principal $\mathrm{GL}(k)$-bundle. In this special case, the converse to Theorem 38.5 is true.

PROPOSITION 38.6. *Let $\pi\colon E \to M$ be a vector bundle of rank k. There is a bijective correspondence between connections on E (in the vector bundle sense) and connections on* $\mathrm{Fr}(E)$ *(in the principal bundle sense).*

Proof. We need only show that a connection on E determines one of $\mathrm{Fr}(E)$, thus providing an inverse to the construction from Theorem 38.5. This time we will give two proofs[4]:

- *Proof using parallel transport:* Suppose $\gamma\colon [0,1] \to M$ is a smooth curve in M. Let $F \in \mathrm{Fr}(E_{\gamma(0)})$. Write $v_i := F(e_i)$, where e_i is the standard basis of $\mathbb{R}^n$. Let $\tilde{F} \in \mathrm{Fr}(E_{\gamma(1)})$ denote the frame defined by

  $$\tilde{F}(e_i) := \widehat{\mathbb{P}}_\gamma(v_i).$$

  Then we define

  $$\widehat{\mathbb{P}}_\gamma^{\mathrm{Fr}(E)}(F) := \tilde{F}.$$

[4] Only two, since—so far—we only have two different ways to view connections on principal bundles. A third will be introduced next lecture...

Then we claim that $\widehat{\mathbb{P}}^{\mathrm{Fr}(E)}$ is a parallel transport system on $\mathrm{Fr}(E)$. Indeed, to prove that $\mathbb{P}^{\mathrm{Fr}(E)}$ satisfies the equivariance axiom, we must show that

$$\widehat{\mathbb{P}}^{\mathrm{Fr}(E)}_{\gamma}(F \circ T) = \widehat{\mathbb{P}}^{\mathrm{Fr}(E)}_{\gamma}(F) \circ T$$

for $T \in \mathrm{GL}(k)$. This is immediate from the fact that $\widehat{\mathbb{P}}_{\gamma}$ is a linear isomorphism.

- *Proof using distributions:* As before we consider the map $\Pi \colon \mathrm{Fr}(E) \times \mathbb{R}^k \to E$ is the map

  $$\Pi(F, v) \coloneqq F(v) \in E_x, \qquad F \in \mathrm{Fr}(E_x),\ v \in \mathbb{R}^k.$$

  We then define for $F \in \mathrm{Fr}(E)$

  $$\mathcal{H}^{\mathrm{Fr}(E)}_F \coloneqq \left\{ Z \in T_F \,\mathrm{Fr}(E) \mid D\Pi(F, 0)[Z, 0] \in \mathcal{H}_{\Pi(F,0)} \right\}.$$

  I will leave the verification that this defines a connection on $\mathrm{Fr}(E)$ as another instructive exercise.

This completes the proof (twice over). ■

Remark 38.7. We have shown that there is a bijective correspondence between connections on $\mathrm{Fr}(E)$ and connections on E. For a general principal G-bundle P however, the passage given by Theorem 38.5 from connections on P to connections on $\rho(P)$ may not be injective or surjective.

LECTURE 39

The connection form and the curvature form

In Lecture 31 we defined the connection map $\kappa\colon TE \to E$ associated to a connection on a vector bundle E. In this lecture we investigate the principal bundle analogue, and then use this to define the curvature of a principal bundle connection.

Let $\pi\colon P \to M$ be a principal G-bundle, and let $\mathfrak{g}$ denote the Lie algebra of G Recall from Definition 25.9 that for each $v \in \mathfrak{g}$ there is a vector field ξ_v on P, called the **fundamental vector field**, which is defined by

$$\xi_v(p) := \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tv) \in T_pP,$$

where $\exp\colon \mathfrak{g} \to G$ is the exponential map of G. Moreover if we denote by

$$\eta_p\colon G \to P, \qquad \eta_p(a) := p \cdot a \tag{39.1}$$

for $p \in P$ then by(25.2) we have

$$D\eta_p(e)[v] = \xi_v(p). \tag{39.2}$$

We now prove:

PROPOSITION 39.1. *Let $\pi\colon P \to M$ be a principal G-bundle. Then for any $p \in P$, the differential $D\eta_p(e)$ of the map η_p from (39.1) at e is an isomorphism*

$$D\eta_p(e)\colon \mathfrak{g} \to V_pP.$$

Proof. We first show that any fundamental vector field ξ_v is vertical. Indeed, note that $\pi \circ \eta_p$ is a constant map. Thus

$$D\pi(p)[\xi_v(p)] = D\pi(p) \circ D\eta_p(e)[v] = D(\pi \circ \eta_p)(e)[v] = 0.$$

Now suppose $v \in \ker D\eta_p(e)$. Then

$$0 = D\eta_p(e)[v] = \xi_v(p) = \frac{d}{dt}\Big|_{t=0} p \cdot \exp(tv)$$

implies by uniqueness of integral curves (cf. Proposition 25.11) that p is a fixed point of $\exp(tv)$. But G acts freely on P, whence $v = 0$. To complete the proof we note that both $\mathfrak{g}$ and V_pP have dimension equal to the dimension of G. Thus $D\eta_p(e)$ is an isomorphism, as claimed. ∎

We now define the principal bundle version of the connection map, which this time is called a *connection form*.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

Definition 39.2. Let $\pi\colon P \to M$ be a principal G-bundle, and let $\mathcal{H}$ be a connection on P. The **connection form** ϖ of $\mathcal{H}$ is the $\mathfrak{g}$-valued 1-form $\varpi \in \Omega^1(P, \mathfrak{g})$ defined by

$$\varpi_p(\zeta) := D\eta_p(e)^{-1}[\zeta^{\mathtt{V}}].$$

This does indeed define an element of $\mathfrak{g}$: $\zeta^{\mathtt{V}} \in V_pP$ and hence by Proposition 39.1 there is a unique element $\varpi_p(\zeta) \in \mathfrak{g}$ such that $D\eta_p(e)[\varpi_p(\zeta)] = \zeta^{\mathtt{V}}$.

Of course, it must be proved that ϖ really is smooth. The next result establishes this, and shows that ϖ uniquely determines $\mathcal{H}$. Recall that G acts on $\mathfrak{g}$ via the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$. In the following, whenever we talk about G acting on $\mathfrak{g}$, we will always implicitly assume that the action is the adjoint one.

Theorem 39.3 (Properties of the connection form). *Let $\pi\colon P \to M$ be a principal bundle with connection $\mathcal{H}$. Then the connection form ϖ is smooth and equivariant, i.e.*

$$r_a^\star(\varpi) = \mathrm{Ad}_{a^{-1}}(\varpi), \qquad \forall\, a \in G$$

and satisfies

$$\varpi(\xi_v) \equiv v, \qquad \forall\, v \in \mathfrak{g}. \tag{39.3}$$

Moreover if $\vartheta \in \Omega^1(P, \mathfrak{g})$ is any equivariant form satisfying (39.3) then $\ker \vartheta_p$ defines a connection on P.

Remark 39.4. The connection form does *not* belong to $\Omega^1_G(P, \mathfrak{g})$! Indeed, (39.3) is the "opposite" of being a horizontal form. We will see how that the *curvature form*, which is a $\mathfrak{g}$-valued 2-form, does belong to $\Omega^2_G(P, \mathfrak{g})$.

Proof of Theorem 39.3. We prove the theorem in three steps.

1. In this step we show that ϖ is equivariant and that (39.3) holds. We begin with the latter statement. By Proposition 39.1 for any $p \in P$ one has $\xi_v(p) \in V_pP$ and thus $\xi_v(p)^{\mathtt{V}} = \xi_v(p)$; thus

$$\varpi_p(\xi_v(p)) = D\eta_p(e)^{-1}[\xi_v(p)] = v$$

by (39.2). To verify equivariance, fix $p \in P$, $a \in G$, and $\zeta \in T_pP$. We wish to show that

$$\varpi_{p\cdot a}(Dr_a(p)[\zeta]) = \mathrm{Ad}_{a^{-1}}(\varpi_p(\zeta)). \tag{39.4}$$

Since both sides of (39.4) are $\mathbb{R}$-linear and $\zeta = \zeta^{\mathtt{H}} + \zeta^{\mathtt{V}}$ is the sum of a horizontal and vertical vector, it suffices to prove (39.3) when ζ is horizontal and when ζ is vertical.

If ζ is horizontal then by (38.2) so is $Dr_a(p)[\zeta]$. Thus $\varpi_p(\zeta)$ and $\varpi_{p\cdot a}(Dr_a(p)[\zeta])$ are both zero, and so (39.4) follows. If instead ζ is vertical then by Proposition 39.1 we may assume $\zeta = \xi_v(p)$ for some $v \in \mathfrak{g}$. We now use Proposition 25.14, which tells us that

$$Dr_a(p)[\xi_v(p)] = \xi_{\mathrm{Ad}_{a^{-1}}(v)}(r_a(p)).$$

Thus using (39.3) twice, we have:

$$\begin{aligned}
\varpi_{p\cdot a}\left(Dr_a(p)[\xi_v(p)]\right) &= \varpi_{p\cdot a}\left(\xi_{\mathrm{Ad}_{a^{-1}}(v)}(r_a(p))\right) \\
&= \mathrm{Ad}_{a^{-1}}(v) \\
&= \mathrm{Ad}_{a^{-1}}(\varpi_p(\xi_v(p)))
\end{aligned}$$

which proves (39.4) for the vertical case.

2. In this step we prove that ϖ is smooth. Choose a basis $\{v_i\}$ of $\mathfrak{g}$. Then by Proposition 39.1 the vector fields $\{\xi_{v_i}\}$ span the vertical subbundle. Now fix a point $p \in P$. Since $\mathcal{H}$ is a distribution, there exist vector fields Z_j on a neighbourhood of p that span $\mathcal{H}$. Since $\mathcal{H}$ is complementary to VP, the collection $\{\xi_{v_i}, Z_j\}$ span the entire tangent bundle to P near p. Thus if Z is any vector field on P we can write

$$Z = f^i \, \xi_{v^i} + g^j \, Z_j$$

near p for smooth functions f^i, g^j. Then by (39.3) one has near p that

$$\varpi(Z) = f^i \, v_i.$$

The right-hand side is smooth, and since Z was arbitrary this proves that ϖ is smooth at p (this is a special case of Theorem 26.3). Since p was also arbitrary, it follows that ϖ is smooth.

3. Finally we prove that any equivariant form $\vartheta \in \Omega^1(P, \mathfrak{g})$ satisfying (39.3) determines a connection via $\mathcal{H} := \ker \vartheta$. Indeed, $\ker \vartheta$ is automatically a subbundle (as ϑ is smooth), and (39.3) tells us that

$$TP = \ker \vartheta \oplus \ker D\pi = \ker \vartheta \oplus VP.$$

Thus $\mathcal{H}$ is a preconnection. Moreover since ϑ is equivariant we have

$$Dr_a[\ker \vartheta] \subseteq \ker \vartheta.$$

Applying this with $Dr_{a^{-1}}$ to both sides and using equivariance again we have

$$\ker \vartheta = Dr_{a^{-1}} \circ Dr_a[\ker \vartheta] \subset Dr_{a^{-1}}[\ker \vartheta] \subseteq \ker \vartheta$$

which shows we have equality. Thus $\ker \vartheta$ is a connection. This completes the proof. ■

Remark 39.5. We now have three different ways to specify a connection on a principal bundle: as a distribution, as a parallel transport system, and via a connection form. Just as with Remark 32.2, it is useful to have a single fixed notation to refer to a connection, which can then be used to mean whichever viewpoint is convenient at the time. Thus from now on we will typically refer to a connection on a principal bundle with the symbol ϖ.

We say a vector field on P is **horizontal** if $Z(p) \in \mathcal{H}_p$ for all p. Thus in particular given any vector field X on M, its horizontal lift (Definition 28.8) is horizontal.

Remark 39.6. We remark that for any $p \in P$ and any $\zeta \in T_pP$ there exists a horizontal vector field Z on P such that $Z(p) = \zeta^{\mathtt{H}}$. Indeed, we can even take Z to be a horizontal lift: let X denote any vector field on M such that $X(\pi(p)) = D\pi(p)[\zeta]$ (such X exists by Problem D.1). Then $\overline{X}(p) = \zeta^{\mathtt{H}}$. Similarly Proposition 39.1 shows that for any $p \in P$ and any $\zeta \in T_pP$ we can find $v \in \mathfrak{g}$ such that $\xi_v(p) = \zeta^{\mathtt{V}}$.

We now define the *curvature* of a connection on a principal bundle. Firstly, we define flatness in the same way.

Definition 39.7. A connection $\mathcal{H}$ on a principal bundle P is **flat** if $\mathcal{H}$ is integrable.

The curvature then measures how far away a connection is from being flat.

Definition 39.8. Let $\pi\colon P \to M$ be a principal G-bundle with connection ϖ. The **curvature form** $\Omega \in \Omega^2(P, \mathfrak{g})$ of ϖ is defined by

$$\Omega_p(\zeta_1, \zeta_2) := -\varpi_p\big([Z_1, Z_2](p)\big), \qquad p \in P,\ \zeta_1, \zeta_2 \in T_pP$$

where Z_1, Z_2 are any two horizontal vector fields on P such that $Z_i(p) = \zeta_i^{\mathtt{H}}$.

Such lifts exist by Remark 39.6. Of course, it must be proved that this is well-defined (i.e. independent of the choice of Z_1 and Z_2) and smooth.

Lemma 39.9. *The curvature form Ω is well defined and belongs to $\Omega^2(P, \mathfrak{g})$. Moreover the connection is flat if and only if Ω is identically zero.*

Proof. Fix $p \in P$ and $\zeta_1, \zeta_2 \in T_pP$. Suppose Z_1 and Z_2 are any two horizontal vector fields on P such that $Z_i(p) = \zeta_i^{\mathtt{H}}$. Let f be a smooth function on P such that $f(p) = 0$, and let W denote any horizontal vector field on P. Then any $Z := Z_1 + fW$ is another horizontal vector field on P such that $Z(p) = \zeta_1^{\mathtt{H}}$ (and moreover any horizontal vector field which agrees with Z_1 at p is locally a finite sum of vector fields of this form). Then

$$[Z, Z_2](p) = [Z_1, Z_2](p) + f(p)[W, Z_2](p) - Z_2(f)W(p),$$

and thus taking vertical components, we see that

$$[Z, Z_2](p)^{\mathtt{V}} = [Z_1, Z_2](p)^{\mathtt{V}},$$

and thus also

$$\varpi_p([Z, Z_2](p)) = \varpi_p([Z_1, Z_2](p)).$$

A similar argument shows that $\varpi_p([Z_1, Z_2](p))$ is independent of the choice of Z_2 as well. This proves that Ω is well defined. It is then obvious that Ω is smooth. It is clearly antisymmetric, and thus Ω does indeed define element of $\Omega^2(P, \mathfrak{g})$. Since $\mathcal{H} = \ker \varpi$, it is clear from the definition that the distribution $\mathcal{H}$ is integrable if and only if Ω is identically zero. ∎

In fact, the connection form belongs to $\Omega^2_G(P, \mathfrak{g}) \subset \Omega^2(P, \mathfrak{g})$. The next theorem summarises the main properties of Ω. Its proof is deferred to the next lecture.

Theorem 39.10 (Properties of the curvature form). *Let $\pi\colon P \to M$ denote a principal bundle, and let ϖ denote a connection on P. Then:*

(i) *The curvature Ω belongs to $\Omega^2_G(P, \mathfrak{g})$.*

(ii) *The two forms ϖ and Ω satisfy* **Cartan's Structure Equation***:*

$$\Omega = d\varpi + \frac{1}{2}[\varpi, \varpi] \tag{39.5}$$

(iii) *The* **Bianchi Identity** *holds:*

$$d\Omega = [\Omega, \varpi]. \tag{39.6}$$

Remark 39.11. The Lie bracket in (39.5) and (39.6) is the one from Example 26.7. Thus in particular if Z, W are two vector fields on P then

$$[\varpi, \varpi](Z, W) \stackrel{\text{def}}{=} [\varpi(Z), \varpi(W)] - [\varpi(W), \varpi(Z)].$$

But now note that by (26.2) this Lie bracket is *symmetric* on 1-forms, and hence

$$[\varpi, \varpi](Z, W) = 2[\varpi(Z), \varpi(W)].$$

This is the reason[1] for the factor of a $\frac{1}{2}$ on the right-hand side of (39.5).

Remark 39.12. The Bianchi Identity (39.6) for connections on principal bundles implies the Bianchi Identity for connections on vector bundles (Theorem 36.1), as you will prove on Problem Sheet S.

We conclude this lecture with a preliminary lemma that will be useful in the proof of Theorem 39.10. We say a vector field Z is **right-invariant** if $(r_a)_\star(Z) = Z$ for every $a \in G$, that is:

$$Dr_a(p)[Z(p)] = Z(p \cdot a), \qquad \forall\, p \in P,\ a \in G.$$

Lemma 39.13. *Let $\pi\colon P \to M$ be a principal bundle with connection ϖ. Then:*

(i) *If X is a vector field on M then the horizontal lift $\overline{X}$ of X is right-invariant.*

(ii) *If Z is a horizontal vector field on P then $[\xi_v, Z]$ is also horizontal for any $v \in \mathfrak{g}$.*

(iii) *If Z is any right-invariant vector field on P then $[\xi_v, Z] = 0$ for any $v \in \mathfrak{g}$,*

Proof. To prove (i) we take $p \in P$ and $a \in G$. Since $\pi \circ r_a = \pi$, we have

$$D\pi(p \cdot a)\big[Dr_a(p)[\overline{X}(p)]\big] = D\pi(p)[\overline{X}(p)] = X(\pi(p)) = D\pi(p \cdot a)[\overline{X}(p \cdot a)]$$

By uniqueness of horizontal lifts this implies that $Dr_a(p)[\overline{X}(p)] = \overline{X}(p \cdot a)$, and thus $\overline{X}$ is right-invariant, as claimed.

To prove (ii), we recall from Proposition 25.11 that the flow of ξ_v is given by $\theta_t(p) := p \cdot \exp(tv) = r_{\exp(tv)}(p)$. Thus using Theorem 8.25 we have

$$\begin{aligned} [\xi_v, Z](p) &= \mathcal{L}_{\xi_v}(Z)(p) \\ &= \lim_{t \to 0} \frac{Dr_{\exp(-tv)}(p \cdot \exp(tv))\big[Z(p \cdot \exp(tv))\big] - Z(p)}{t}. \end{aligned} \tag{39.7}$$

Since right translation preserves $\mathcal{H}$, if $Z(p) \in \mathcal{H}_p$ then also $Dr_{\exp(-tv)}(\exp(tv))\big[Z(p \cdot \exp(tv))\big] \in \mathcal{H}_p$. Thus if we set

$$h(t) := \frac{Dr_{\exp(-tv)}(p \cdot \exp(tv))\big[Z(p \cdot \exp(tv))\big] - Z(p)}{t}$$

then for all small t, $h(t)$ belongs to the vector space $\mathcal{H}_p$. Thus also $[\xi_v, Z](p) = \lim_{t \to 0} h(t)$ belongs to $\mathcal{H}_p$.

Finally to prove (iii), if Z is right-invariant then the numerator in (39.7) is identically zero, and thus $[\xi_v, Z]$ is too. ∎

[1] We warn the reader that some textbooks are inconsistent with how $[\cdot, \cdot]$ is defined, and thus sometimes the factor of $\frac{1}{2}$ is incorrectly omitted.

Cartan's Structure Equation

In this lecture we begin by proving Theorem 39.10. For convenience, we restate it again here.

THEOREM 40.1 (Properties of the curvature form). *Let $\pi\colon P \to M$ denote a principal bundle, and let ϖ denote a connection on P. Then:*

(i) *The curvature Ω belongs to $\Omega^2_G(P, \mathfrak{g})$.*

(ii) *The two forms ϖ and Ω satisfy* **Cartan's Structure Equation***:*

$$\Omega = d\varpi + \frac{1}{2}[\varpi, \varpi] \tag{40.1}$$

(iii) *The* **Bianchi Identity** *holds:*

$$d\Omega = [\Omega, \varpi]. \tag{40.2}$$

Proof. We will prove the result in three steps.

1. In this step we prove Cartan's Structure Equation (40.1). This means that for any two vector fields Z, W on P we must show that

$$\Omega(Z, W) = d\varpi(Z, W) + [\varpi(Z), \varpi(W)] \tag{40.3}$$

as functions $P \to \mathfrak{g}$ (cf. Remark 39.11). Since both sides of (40.3) are point operators, it suffices to consider separately the three cases where one or both Z and W are horizontal or vertical respectively. For this let X, Y denote two vector fields on M and let $v, w \in \mathfrak{g}$.

(i) *The case $Z = \xi_v$ and $W = \xi_w$ (both sides vertical):*

In this case $\Omega(\xi_v, \xi_w) = 0$ by definition. To compute the left-hand side we first start with:

$$\begin{aligned} d\varpi(\xi_v, \xi_w) &= \xi_v(\varpi(\xi_w)) - \xi_w(\varpi(\xi_v)) - \varpi([\xi_v, \xi_w]) \\ &= d(\varpi(\xi_w))[\xi_v] - d(\varpi(\xi_v))[\xi_w] - \varpi(\xi_{[v,w]}) \\ &= 0 - 0 - [v, w], \end{aligned}$$

where the first line used Theorem 26.8, the second line used Problem M.5, and the third line used the fact that $\varpi(\xi_w)$ is the constant function $p \mapsto w$ by (39.3), and thus $d(\varpi(\xi_w))$ is identically zero. Since

$$[\varpi(\xi_v), \varpi(\xi_w)] = [v, w]$$

by (39.3) again, this shows that the right-hand side of (40.3) is also identically zero.

(ii) *The case $Z = \xi_v$ and $W = \overline{Y}$ (one side vertical, one side horizontal:*

In this case by part (i) and part (ii) of Lemma 39.13 we have $[\xi_v, \overline{Y}] = 0$ and thus $\Omega(\xi_v, \overline{Y}) = 0$. Looking at the right-hand side again, we have:

$$\begin{aligned} d\varpi(\xi_v, \overline{Y}) &= \xi_v(\underbrace{\varpi(\overline{Y})}_{=0}) - \overline{Y}(\varpi(\xi_v)) - \varpi\big(\underbrace{[\xi_v, \overline{Y}]}_{=0}\big) \\ &= 0 - d(\varpi(\xi_v))[\overline{Y}] - 0 \\ &= 0 \end{aligned}$$

Similarly $[\varpi(\xi_v), \varpi(\overline{Y})] = 0$ as $\varpi(\overline{Y}) = 0$. This proves (40.3) in this case too.

(iii) *The case $Z = \overline{X}$ and $W = \overline{Y}$ (both sides horizontal):*

In this case we have by

$$\begin{aligned} \Omega(\overline{X}, \overline{Y}) &= -\varpi\big([\overline{X}, \overline{Y}]\big) \\ &= d\varpi(\overline{X}, \overline{Y}) - \overline{X}(\varpi(\overline{Y})) + \overline{Y}(\varpi(\overline{X})) \\ &= d\varpi(\overline{X}, \overline{Y}), \end{aligned}$$

where the second line used the Theorem 26.8 again and the last used $\varpi(\overline{X}) = \varpi(\overline{Y}) = 0$. But by this same logic we can also write

$$\Omega(\overline{X}, \overline{Y}) = d\varpi(\overline{X}, \overline{Y}) + [\varpi(\overline{X}), \varpi(\overline{Y})],$$

since we are just adding zero. This proves (40.1) in this case, and hence in general.

2. In this step we prove the Bianchi Identity (40.2). For this we argue as follows:

$$\begin{aligned} d\Omega &\overset{(\dagger)}{=} d^2\varpi + \frac{1}{2}d[\varpi, \varpi] \\ &\overset{(\ddagger)}{=} \frac{1}{2}\left([d\varpi, \varpi] - [\varpi, d\varpi]\right) \\ &\overset{(\heartsuit)}{=} [d\varpi, \varpi] \\ &\overset{(\dagger)}{=} [\Omega, \varpi] - \frac{1}{2}[[\varpi, \varpi], \varpi] \\ &\overset{(\diamondsuit)}{=} [\Omega, \varpi] \end{aligned}$$

where ($\dagger$) used the Cartan Structure Equation (both times), ($\ddagger$) used Problem M.6, ($\heartsuit$) used (26.2), and finally ($\diamondsuit$) used Problem S.1. This proves the Bianchi Identity.

3. To complete the proof we show that Ω is horizontal and equivariant, and hence defines an element of $\Omega^2_G(P, \mathfrak{g})$. The fact that Ω is horizontal was already proved in Step 1 (namely, the computation $\Omega(\xi_v, \cdot)$ is identically zero). Thus we need only prove equivariance. For this we use Cartan's Structure Equation and the fact that ϖ is equivariant to prove:

$$\begin{aligned} r_a^\star(\Omega) &= r_a^\star\left(d\varpi + \frac{1}{2}[\varpi, \varpi]\right) \\ &= dr_a^\star(\varpi) + \frac{1}{2}\big[r_a^\star(\varpi), r_a^\star(\varpi)\big] \\ &\overset{(\dagger)}{=} \mathrm{Ad}_{a^{-1}}(d\varpi) + \big[\,\mathrm{Ad}_{a^{-1}}(\varpi), \mathrm{Ad}_{a^{-1}}(\varpi)\big] \\ &= \mathrm{Ad}_{a^{-1}}(\Omega), \end{aligned}$$

where this time (†) used the fact that $d\varpi$ is also equivariant (see claim (ii) from our third proof of Theorem 38.3 in Lecture 38). This finally completes the proof of the theorem. ■

Since the curvature belongs to $\Omega^2_G(P, \mathfrak{g})$, we can interpret it also as a bundle-valued form on M.

DEFINITION 40.2. Let $\pi\colon P \to M$ denote a principal G-bundle. We denote by $\mathrm{Ad}(P) = P \times_G \mathfrak{g}$ the vector bundle over M corresponding to $\rho = \mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ and call this the **adjoint bundle** of P.

COROLLARY 40.3. *Let $\pi\colon P \to M$ be a principal G-bundle and let ϖ denote a connection on P. Then the curvature Ω induces a bundle-valued 2-form $\Omega^\flat \in \Omega^2(M, \mathrm{Ad}(P))$. Explicitly, $\Omega^\flat$ is defined by*

$$\Omega^\flat_x(w_1, w_2) := [p, \Omega_p(\zeta_1, \zeta_2)] \tag{40.4}$$

for $w_1, w_2 \in T_xM$, $p \in P_x$ and $\zeta_1, \zeta_2 \in T_pP$ satisfy $D\pi(p)[\zeta_i] = w_i$.

Proof. Apply Theorem 26.17 to Ω. ■

Let $\pi\colon P \to M$ denote a principal G-bundle, and let $\rho\colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. We have seen that a connection on a principal bundle P induces a connection on the associated bundle $\rho(P)$. We now examine the relationship between the connection form ϖ and the covariant derivative operator ∇, and also between the two curvatures Ω and R^∇.

The starting point for this discussion is the observation that the differential

$$\lambda := D\rho(e)\colon \mathfrak{g} \to \mathfrak{gl}(V)$$

of ρ is a Lie algebra representation of $\mathfrak{g}$ (this is a special case of Proposition 9.21). Thus for example if $\rho = \mathrm{Ad}$ is the adjoint representation of G on $\mathfrak{g}$ then $\lambda = \mathrm{ad}$. The following lemma is on Problem Sheet S.

LEMMA 40.4. *Let $\pi\colon P \to M$ denote a principal G-bundle, and let $\rho\colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. Let $\lambda := D\rho(e)$, and suppose $f\colon P \to V$ is an equivariant smooth function. Then for any $v \in \mathfrak{g}$, one has*

$$\xi_v(f) + \lambda(v)[f] = 0.$$

Here $\xi_v(f)$ should be interpreted as another smooth function $P \to V$, cf. (26.1).

DEFINITION 40.5. Suppose $\omega \in \Omega^r(P, \mathfrak{g})$ and $\vartheta \in \Omega^s(P, V)$ are vector-valued forms on $\mathfrak{g}$ and V respectively. Using $\lambda = D\rho(e)$, we can form the product of ω and ϑ via the formula

$$\begin{aligned}\big(\omega \wedge_\rho \vartheta\big)_p&(\zeta_1, \dots, \zeta_{r+s}) \\ &:= \frac{1}{r!s!} \sum_{\varrho \in \mathfrak{S}_{r+s}} \operatorname{sgn}(\varrho) \lambda\Big(\omega\big(\zeta_{\varrho(1)}, \dots, \zeta_{\varrho(r)}\big)\Big) \Big[\vartheta\big(\zeta_{\varrho(r+1)}, \dots \zeta_{\varrho(r+s)}\big)\Big].\end{aligned} \tag{40.5}$$

This is similar (but not quite the same) as the construction of the wedge product $\wedge_\beta$ in Lecture 26.

Our next result gives an explicit formula for d^∇ in terms of ρ and ϖ.

THEOREM 40.6. *Let $\pi\colon P \to M$ denote a principal G-bundle, and let $\rho\colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. Abbreviate by $E = \rho(P)$ the associated vector bundle. Suppose ϖ is a connection on P, and let d^∇ denote the corresponding exterior covariant differential on E. Then for any $\vartheta \in \Omega(M, E)$ we have*

$$d^\nabla \vartheta = \left(d\vartheta^\sharp + \varpi \wedge_\rho \vartheta^\sharp\right)^\flat .$$

REMARK 40.7. As the proof will show, $d\vartheta^\sharp + \varpi \wedge_\rho \vartheta^\sharp$ is horizontal and equivariant, and hence $\left(d\vartheta^\sharp + \varpi \wedge_\rho \vartheta^\sharp\right)^\flat$ is well-defined.

As a special case for $r = 0$, we obtain:

COROLLARY 40.8. *If $s \in \Gamma(E)$ corresponds to an equivariant function $f\colon P \to V$ (Corollary 26.18) then*

$$\nabla s = \left(df + \lambda(\varpi)[f]\right)^\flat \tag{40.6}$$

Thus if $X \in \mathfrak{X}(M)$ then section $\nabla_X(s)$ of E corresponds to the equivariant function $\overline{X}(f)$, where $\overline{X}$ is the horizontal lift of X.

Proof. If $f \in \Omega^0(P, V)$ is a zero-form, (40.5) simplifies to

$$\varpi \wedge_\rho f = \lambda(\varpi)[f].$$

and thus (40.6) is immediate from Theorem 40.6. Moreover if $X \in \mathfrak{X}(M)$ then $\nabla_X(s)$ is the section of E corresponding to the equivariant function

$$df[Z] + \lambda(\varpi(Z))[f],$$

where Z is any vector field on P such that $D\pi[Z] = X$ (this is independent of the choice of Z by equivariance). In particular, choosing $Z = \overline{X}$, the second term disappears, and thus $\nabla_X(s)$ corresponds to $\overline{X}(f)$, as required. ∎

Proof of Theorem 40.6. We will prove the theorem in three steps.

1. In this step we set up notation and outline the strategy of the proof. Suppose $\vartheta \in \Omega^r(M, E)$. By the definition (cf. (38.6)) of d^∇, we have

$$d^\nabla \vartheta := \left(\left(d(\vartheta^\sharp)\right)^{\mathtt{H}}\right)^\flat$$

Since the $\flat \leftrightarrow \sharp$ correspondence is bijective, it suffices to show that if $\psi := \vartheta^\sharp \in \Omega^r_G(P, V)$ then for all $p \in P$ and all $\zeta_1, \ldots, \zeta_{r+1} \in T_pP$, we have[1]:

$$\begin{aligned} d\psi|_p\left(\zeta_0^{\mathtt{H}}, \ldots, \zeta_r^{\mathtt{H}}\right) &= d\psi|_p\left(\zeta_0, \ldots, \zeta_r\right) \\ &\quad + \frac{1}{r!} \sum_{\varrho \in \mathfrak{S}_{r+1}} \operatorname{sgn}(\varrho) \lambda\left(\varpi_p(\zeta_{\varrho(0)})\right) \left[\psi\left(\zeta_{\varrho(1)}, \ldots \zeta_{\varrho(r)}\right)\right]. \end{aligned} \tag{40.7}$$

Since both sides of (40.7) are linear in each ζ_i, as in the previous theorem we may assume that each ζ_i is either vertical or horizontal. Moreover by Remark 39.6 we

[1] In the sum below, we think of elements $\varrho \in \mathfrak{S}_{r+1}$ as permutations of $\{0, 1, \ldots, r\}$ (instead of the more usual $\{1, 2, \ldots, r+1\}$.

may assume $\zeta_i = Z_i(p)$ for some vector field Z_i on P, which is either of the form $Z_i = \overline{X}_i$ for some vector field X_i on M (if ζ_i is horizontal) or of the form $Z_i = \xi_{v_i}$ for some $v_i \in \mathfrak{g}$ (if ζ_i is vertical). Define functions $\Psi_0, \Psi_1, \Psi_2 \colon P \to V$ by

$$\Psi_0(p) := d\psi|_p\big(Z_0(p)^{\mathtt{H}}, \dots, Z_r(p)^{\mathtt{H}}\big),$$

and

$$\Psi_1(p) := d\psi|_p\big(Z_0(p), \dots, Z_r(p)\big),$$

and

$$\Psi_2(p) := \frac{1}{r!} \sum_{\varrho \in \mathfrak{S}_{r+1}} \operatorname{sgn}(\varrho)\lambda\big(\varpi_p(Z_{\varrho(0)}(p))\big)\Big[\psi\big(Z_{\varrho(1)}(p), \dots Z_{\varrho(r)}(p)\big)\Big].$$

It suffices to show that

$$\Psi_0 = \Psi_1 + \Psi_2, \qquad \text{as functions } P \to V. \tag{40.8}$$

2. In this step we deal with the two easy cases. If every single Z_i is horizontal then $\Psi_0 = \Psi_1$ by definition, and $\Psi_2 = 0$ since $\varpi(Z_i) = 0$ for every i. Next, suppose two or more of the Z_i are vertical. In this case without loss of generality we may assume $Z_1 = \xi_{v_0}$ and $Z_2 = \xi_{v_1}$ are vertical. In this case we again have $\Psi_0 = 0$, since $Z_0^{\mathtt{H}} = Z_1^{\mathtt{H}} = 0$. Also $\Psi_2 = 0$ as at least one of the arguments $Z_{\varrho(i)}$ for $i = 0, \dots, r$ is vertical and $\psi = \vartheta^\sharp$ is horizontal. Thus we need only show that $\Psi_1 = 0$. By Theorem 26.8 we have

$$\begin{aligned}\Psi_1 = &\sum_{i=0}^{r}(-1)^i Z_i\big(\psi(Z_0, \dots, \widehat{Z}_i, \dots, Z_r)\big) \\ &+ \sum_{0 \le i < j \le r} (-1)^{i+j}\psi([Z_i, Z_j], Z_0, \dots \widehat{Z}_i, \dots, \widehat{Z}_j, \dots, Z_r).\end{aligned} \tag{40.9}$$

Every term in the first summand is zero, since at least one of the arguments is zero. The only term in the second summand that could possibly be non-zero is $i = 0$ and $j = 1$. But in this case by Problem M.5, $[Z_0, Z_1] = \xi_{[v_0, v_1]}$ is also vertical, and hence this term is zero too.

3. In this step we deal with the hardest case, where exactly one of the Z_i is vertical. Thus without loss of generality assume that $Z_0 = \xi_v$ and that $Z_i = \overline{X}_i$ for vector fields X_i on M for $i = 1, \dots, r$. As before, $\Psi_0 = 0$ (since $Z_0^{\mathtt{H}} = 0$). Now in (40.9) some of the terms survive, and we get

$$\Psi_1 = \xi_v\big(\psi(Z_1, \dots, Z_r)\big) + \sum_{i=1}^{r}(-1)^i\omega([\xi_v, Z_i], Z_1, \dots \widehat{Z}_i, \dots, Z_r).$$

But actually by part (iii) of Lemma 39.13, we have $[\xi_v, Z_i] = [\xi_v, \overline{X}_i] = 0$, and thus $\Psi_1 = \xi_v\big(\psi(Z_1, \dots, Z_r)\big)$. Now if we look at Ψ_2, all the terms die apart from those permutations ϱ such that $\varrho(0) = 0$. Since $\varpi(\xi_v) = v$, it follows that

$$\begin{aligned}\Psi_2 &= \frac{1}{r!} \sum_{\varrho \in \mathfrak{S}_{r+1} \text{ with } \varrho(0)=0} \operatorname{sgn}(\varrho)\lambda(v)\Big[\psi\big(Z_{\varrho(1)}, \dots Z_{\varrho(r)}\big)\Big] \\ &= \lambda(v)\left[\frac{1}{r!} \sum_{\varrho \in \mathfrak{S}_{r+1} \text{ with } \varrho(0)=0} \operatorname{sgn}(\varrho)\psi\big(Z_{\varrho(1)}, \dots Z_{\varrho(r)}\big)\right] \\ &= \lambda(v)\big[\psi(Z_1, \dots, Z_r)\big]\end{aligned}$$

Thus to complete the proof we need to show that

$$\xi_v(\psi(Z_1, \dots, Z_r)) + \lambda(v)[\psi(Z_1, \dots, Z_r)] = 0. \tag{40.10}$$

But since $Z_i = \overline{X}_i$ is right-invariant for each i and $\psi \in \Omega^r_G(P, V)$ is equivariant, it follows that $f := \psi(Z_1, \dots, Z_r)$ is itself equivariant. Thus (40.10) follows from Lemma 40.4. ■

We conclude this lecture by comparing the curvature forms Ω and R^∇. For this one starts by observing that $\lambda\colon \mathfrak{g} \to \mathfrak{gl}(V)$ induces a vector bundle homomorphism

$$\Lambda\colon \operatorname{Ad}(P) \to \operatorname{Hom}(\rho(P), \rho(P)),$$

given explicitly by

$$\Lambda([p, v])\big([p, w]\big) := [p, \lambda(v)[w]], \qquad p \in P,\ v \in \mathfrak{g},\ w \in V.$$

This in turn induces a $C^\infty(M)$-linear map $\Lambda_\star\colon \Gamma(\operatorname{Ad}(P)) \to \Gamma(\operatorname{Hom}(\rho(P), \rho(P))$ by

$$\Lambda_\star(s)(x) := \Lambda(s(x)), \qquad x \in M$$

(this is the easy direction of the Hom-Γ Theorem 16.30). Finally, we can also think of $\Lambda_\star$ as defining a map

$$\Lambda_\star\colon \Omega^r(M, \operatorname{Ad}(P)) \to \Omega^r(M, \operatorname{Hom}(\rho(P), \rho(P))),$$

for any $r \geq 0$, via

$$\Lambda_\star(\omega)(X_1, \dots, X_r) := \Lambda_\star(\omega(X_1, \dots, X_r))$$

for $\omega \in \Omega^r(M, \operatorname{Ad}(P))$ and $X_1, \dots, X_r \in \mathfrak{X}(M)$.

THEOREM 40.9. *Let $\pi\colon P \to M$ denote a principal G-bundle, and let $\rho\colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. Let ϖ denote a connection on P, and let ∇ denote the induced connection on $\rho(P)$. Let Ω denote the curvature form of ϖ, and consider $\Omega^\flat \in \Omega^2_G(M, \operatorname{Ad}(P))$ as in Corollary 40.3. Then*

$$\Lambda_\star(\Omega^\flat) = R^\nabla.$$

Proof. In this proof we will suppress the bijection between sections of $\rho(P)$ and equivariant functions $f\colon P \to V$, and treat it as an identification. Thus we write $s = f$ to indicate that a section s corresponds to f. Thus Corollary 40.8 can be stated more succinctly as

$$\nabla_X(s) = \overline{X}(f).$$

This will help keep the notation transparent. With this convention in mind, by Theorem 35.1 we have

$$R^\nabla(X, Y)(s) = \Big([\overline{X}, \overline{Y}] - \overline{[X, Y]}\Big)(f).$$

The vector field $[\overline{X}, \overline{Y}] - \overline{[X, Y]}$ is vertical by part (iii) of Lemma 28.9. In fact, by Problem S.2 and (39.2), one has

$$\Big([\overline{X}, \overline{Y}] - \overline{[X, Y]}\Big)(p) = -\xi_{\Omega_p(\overline{X}(p), \overline{Y}(p))}(p)$$

Thus applying Lemma 40.4 we see that[2] we have

$$R^{\nabla}(X,Y)(s) = \lambda\left(\Omega\big(\overline{X},\overline{Y}\big)\right)[f],$$

which—after unravelling the notation—is exactly what we wanted to prove. ■

[2] This is where it is crucial we defined Ω with a negative sign.

Holonomy and principal bundles

In this lecture we define holonomy in principal bundles, and prove the principal bundle version of the Ambrose-Singer Holonomy Theorem. The vector bundle version (Theorem 34.8) is a simple corollary of the principal version, as you will prove on Problem Sheet S.

DEFINITION 41.1. Let $\pi\colon P \to M$ be a principal G-bundle with connection ϖ. The **holonomy group** $\mathrm{Hol}^{\varpi}(x)$ of ϖ at $x \in M$ is the group of equivariant diffeomorphisms of the fibre P_x of the form $\widehat{\mathbb{P}}_{\gamma}$, where γ is a piecewise smooth loop in M based at x. The **restricted holonomy group** $\mathrm{Hol}^{\varpi}_0(x) \subset \mathrm{Hol}^{\varpi}(x)$ is the subgroup consisting of parallel transport around null-homotopic loops γ.

The following result is the principal bundle analogue of Proposition 32.11. The key difference is that we can view the holonomy group $\mathrm{Hol}^{\varpi}(x)$ as being a subgroup of G itself.

PROPOSITION 41.2. *Let $\pi\colon P \to M$ be a principal G-bundle with connection ϖ.*

(i) *For each $p \in P$, there is a subgroup $H^{\varpi}(p) \subset G$ and a group isomorphism*

$$\phi_p\colon \mathrm{Hol}^{\varpi}(\pi(p)) \to H^{\varpi}(p).$$

(ii) *The subgroups $H^{\varpi}(p)$ and $H^{\varpi}(p \cdot a)$ are conjugate in G.*

(iii) *If $p, q \in P$ can be joined by a horizontal path then $H^{\varpi}(p) = H^{\varpi}(q)$.*

(iv) *There is a subgroup $H^{\varpi}_0(p) \subset H^{\varpi}(p)$ such that ϕ_p restricts to define an isomorphism $\mathrm{Hol}^{\varpi}_0(x) \to H^{\varpi}_0(p)$. This subgroup again has the property that $H^{\varpi}_0(p \cdot a)$ is conjugate to $H^{\varpi}_0(p)$, and $H^{\varpi}_0(p) = H^{\varpi}_0(q)$ whenever p and q can be joined by a horizontal path.*

Proof. Let $x \in M$ and $p \in P_x$. If γ is a piecewise smooth loop based at x, we define $\phi_p\big(\widehat{\mathbb{P}}_{\gamma}\big)$ to be the unique element $b \in G$ such that

$$p \cdot b = \widehat{\mathbb{P}}_{\gamma}(p).$$

If $\phi_p(\widehat{\mathbb{P}}_{\gamma}) = e$ then we claim that $\widehat{\mathbb{P}}_{\gamma} = \mathrm{id}|_{P_x}$. Indeed, by the equivariance axiom from Definition 38.2 we have that for any $a \in G$,

$$\widehat{\mathbb{P}}_{\gamma}(p \cdot a) = \widehat{\mathbb{P}}_{\gamma}(p) \cdot a = p \cdot a.$$

We set $H^{\varpi}(p)$ to be the image of ϕ_p. Next, using the equivariance axiom again, if $\phi_p(\widehat{\mathbb{P}}_{\gamma}) = b$ then

$$(p \cdot a) \cdot (a^{-1} b a) = \widehat{\mathbb{P}}_{\gamma}(p) \cdot a = \widehat{\mathbb{P}}_{\gamma}(p \cdot a),$$

so that

$$\phi_{p\cdot a}\big(\widehat{\mathbb{P}}_\gamma\big) = a^{-1}ba.$$

Thus

$$H^\varpi(p\cdot a) = a^{-1}H^\varpi(p)a,$$

which proves the two subgroups are conjugate. This proves (ii). To prove (iii), let $\gamma\colon [0,1] \to M$ be a path in M from $x := \gamma(0)$ to $y := \gamma(1)$. Let $p \in P_x$ and set $q := \widehat{\mathbb{P}}_\gamma(p)$. We claim that $H^\varpi(p) \subseteq H^\varpi(q)$. Indeed, suppose $b \in H^\varpi(p)$. Then there exists a piecewise smooth loop δ based at x such that

$$\widehat{\mathbb{P}}_\delta(p) = p \cdot b.$$

Then $\gamma^- * \delta * \gamma$ is a loop based at y, and

$$\begin{aligned}
\widehat{\mathbb{P}}_{\gamma^- * \delta * \gamma}(q) &= \widehat{\mathbb{P}}_\gamma \circ \widehat{\mathbb{P}}_\delta \circ \widehat{\mathbb{P}}_{\gamma^-}(q) \\
&= \widehat{\mathbb{P}}_\gamma \circ \widehat{\mathbb{P}}_\delta(p) \\
&= \widehat{\mathbb{P}}_\gamma(p \cdot b) \\
&= \widehat{\mathbb{P}}_\gamma(p) \cdot b \\
&= q \cdot b
\end{aligned}$$

Thus $b \in H^\varpi(q)$. Applying the same argument with γ^- in place of γ shows that $H^\varpi(q) \subseteq H^\varpi(p)$, and thus $H^\varpi(p) = H^\varpi(q)$. Finally, (iv) is proved in the same way, and we leave the details as an exercise. ■

The holonomy groups $H^\varpi(p) \subset G$ enjoy the same properties that the holonomy groups did for vector bundles. The next theorem summarises the key properties we will need. The proofs all proceed analogously to the corresponding statements about vector bundles.

THEOREM 41.3. *Let $\pi\colon P \to M$ be a principal G-bundle with connection ϖ. Then the holonomy group $H^\varpi(p) \subset G$ is a Lie subgroup of G. The connected component of $H^\varpi(p)$ containing the identity is exactly $H^\varpi_0(p)$. If M is simply connected then $H^\varpi(p) = H^\varpi_0(p)$. Finally, $H^\varpi(p)$ is the trivial subgroup $\{e\}$ for all $p \in P$ if and only if P is a trivial bundle and ϖ is the trivial connection.*

Meanwhile the proof of the next result is on Problem Sheet S.

PROPOSITION 41.4. *Let $\pi\colon P \to M$ be a principal G-bundle. Let $\rho\colon G \to \mathrm{GL}(V)$ denote an effective representation. Let ϖ denote a connection on P and let ∇ denote the associated connection on $\rho(P)$. Fix $x \in M$. Then we can regard $\mathrm{Hol}^\varpi(x)$ and $\mathrm{Hol}^\nabla(x)$ as subgroups of G and $\mathrm{GL}(V)$ respectively, which are defined up to conjugation. Then (also up to conjugation)*

$$\rho\big(\mathrm{Hol}^\varpi(x)\big) = \mathrm{Hol}^\nabla(x).$$

DEFINITION 41.5. Let $\pi\colon P \to M$ be a principal G-bundle, let $H \subset G$ be a Lie subgroup and suppose $\pi_1\colon Q \to M$ is a principal H-subbundle of P (cf. Definition 24.19). A connection ϖ on P with associated distribution $\mathcal{H}$ is said to be **reducible** to Q if the distribution $\mathcal{H} \cap TQ$ defines a connection on Q. Equivalently, if $\imath\colon Q \hookrightarrow P$ denotes the inclusion this means that $\varpi|_Q = \imath^\star(\varpi)$ is a connection one-form on Q with curvature $\Omega|_Q = \imath^\star(\Omega)$ (compare Problem Q.1).

On Problem Sheet S you will investigate the relationship between this definition and the notion of a G-connection on a vector bundle (cf. Problem Q.4). The next result is similar to Theorem 34.4.

THEOREM 41.6 (The Reduction Theorem). *Let $\pi\colon P \to M$ be a principal G-bundle over a connected manifold M and let ϖ denote a connection on P. Fix a point $p \in P$ and set $H := H^{\varpi}(p) \subset G$. Let Q denote the set of all points $q \in P$ which can be joined to p via a piecewise smooth horizontal path. Then Q is a principal H-subbundle of P, and the connection ϖ is reducible to Q.*

Proof. The proof is an application of Proposition 24.20. Part (iii) of Proposition 41.2 tells us that Q is preserved by the action of H, and that the action of H on $P_y \cap Q$ for any point $y \in M$ is transitive. Moreover since M is connected the restriction of π to Q is surjective (compare this to the proof of Step 1 of Theorem 33.4). Thus to show that Q is a principal H-subbundle, by Proposition 24.20 we need only construct local sections of P that take values in Q.

For this, suppose $\pi(p) := x$, and fix any other point $y \in M$. Let $\sigma\colon U \to O$ denote a chart about y, and set

$$\gamma_v(t) := \sigma^{-1}\big(\sigma(y) + tD\sigma(y)[v]\big),$$

which is a well-defined smooth curve for t depending on v sufficiently small. Fix $q \in P_y$. Define a section $s_q \in \Gamma(U, P)$ by

$$s_q(\gamma_v(t)) := \mathbb{P}_{\gamma_v}(q)(t).$$

If the point $q \in P_y$ actually belongs to Q then s_q takes values in Q. Indeed, if $q = \widehat{\mathbb{P}}_{\gamma}(p)$ for some path $\gamma\colon [0,1] \to M$ such that $\gamma(0) = x$ and $\gamma(1) := y$ then

$$s_q(\gamma_v(t)) = \widehat{\mathbb{P}}_{\gamma * \gamma_{v,t}}(p),$$

where $\gamma_{v,t}(r) := \gamma_v(rt)$ for $0 \le r \le 1$.

Finally, to see that the connection is reducible to Q, we observe that by definition any horizontal curve starting in Q must remain in Q, and hence $\mathcal{H}_q \subset T_qQ$. Since clearly $V_qQ = V_qP \cap T_qQ$, it follows that $T_qQ = \mathcal{H}_q \oplus V_qQ$. Thus $\mathcal{H}$ is a preconnection on Q, and the equivariance condition is clear from above. This completes the proof. ■

We now state and prove the principal bundle version of the Ambrose-Singer Holonomy Theorem. The proof is another application of the Frobenius Theorem 11.18.

THEOREM 41.7 (The Ambrose-Singer Holonomy Theorem Redux). *Let $\pi\colon P \to M$ be a principal G-bundle over a connected manifold M. Let ϖ denote a connection on P, and let $\Omega \in \Omega^2_G(P, \mathfrak{g})$ denote the curvature form. Let $p \in P$, let $H = H^{\varpi}(p) \subset G$, and let Q denote the principal H-subbundle of P from the Reduction Theorem 41.6. Then the Lie algebra $\mathfrak{h}$ of H is the subalgebra of $\mathfrak{g}$ spanned by all elements of the form $\Omega_q(\zeta_1, \zeta_2)$ for $q \in Q$ and $\zeta_1, \zeta_2 \in T_qQ$.*

Proof. Let $\mathfrak{k}$ denote the Lie subalgebra of $\mathfrak{g}$ spanned by elements of the form $\Omega_q(\zeta_1, \zeta_2)$ for $q \in Q$ and $\zeta_1, \zeta_2 \in T_q Q$. Then certainly $\mathfrak{k} \subseteq \mathfrak{h}$. Let $k = \dim \mathfrak{k}$ and $n = \dim M$. Then $\dim \mathfrak{h} = \dim Q - n$. We will show that also $k = \dim Q - n$, which implies $\mathfrak{k} = \mathfrak{h}$.

Define a distribution Δ on Q by

$$\Delta_q := \mathcal{H}_q \oplus D\eta_q(e)[\mathfrak{k}],$$

where $\mathcal{H}$ is the connection distribution and $\eta_q \colon H \to Q$ is the map $\eta_q(a) = q \cdot a$. To see that this is indeed a distribution on Q, we argue as in the proof of Step 2 of Theorem 39.3. Take a basis $\{v_i \mid i = 1, \dots, k\}$ of $\mathfrak{k}$, and let ξ_{v_i} denote the fundamental vector fields associated to this basis. Fix $q \in Q$, and let $\{X_j \mid j = 1, \dots, n\}$ denote vector fields on M such that $\{X_j(y)\}$ is a basis of $T_y M$ for all y near $\pi(q)$, and let $\overline{X}_j$ denote the horizontal lifts of X_j. Then $\{\xi_{v_i}, \overline{X}_j\}$ spans Δ near q, and thus Δ is indeed a distribution of dimension $n + k$. Next. we claim Δ is integrable. Using Lemma 11.13, we need only check:

(i) $[\xi_{v_i}, \xi_{v_j}]$ belongs to Δ.

(ii) $[\xi_{v_i}, \overline{X}_j]$ belongs to Δ.

(iii) $[\overline{X}_i, \overline{X}_j]$ belongs to Δ.

Of these, (i) follows because $[\xi_{v_i}, \xi_{v_j}] = \xi_{[v_i, v_j]}$ by Problem M.5 and because $\mathfrak{k}$ is (by definition) a subalgebra. Next, (ii) is immediate, since by part (iii) of Lemma 39.13 such a bracket is always zero. Finally, by Problem S.2, we have

$$[\overline{X}_i, \overline{X}_j](q) = \overline{[X_i, X_j]}(q) - D\eta_q(e)[\Omega_q(\overline{X}_i(q), \overline{X}_j)(q)],$$

which belongs to Δ_q by definition. Thus by the Frobenius Theorem, Δ induces a foliation of Q. Let L denote the leaf containing p. We claim that in fact $L = Q$ (and thus this is not a particularly thrilling foliation). Indeed, if $c(t)$ is a horizontal curve starting at p then $c'(t) \in \mathcal{H}_{c(t)} \subset \Delta_{c(t)}$ for each t, and thus $\operatorname{im}(c)$ is contained in an integral manifold of Δ. By maximality, $\operatorname{im}(c)$ is also contained in L. Since c was arbitrary, this shows that $Q \subseteq L$. Since $L \subseteq Q$ by definition, we have $L = Q$ as claimed. Since $\dim L = n + k$, this shows that $k = \dim Q - n = \dim \mathfrak{h}$. This completes the proof. ■

We conclude this lecture by stating the following existence result. The proof is not too hard[1], but it is a long and somewhat uninspiring computation, and hence we will skip it.

THEOREM 41.8. *Let $\pi \colon P \to M$ denote a principal G-bundle. Assume $\dim M \geq 2$ and that G is connected. Then there exists a connection ϖ on P with $H^{\varpi}(p) = G$ for all $p \in P$.*

As a corollary, we obtain the following converse to the Reduction Theorem 41.6.

[1]See for instance Theorem 8.2 on p90 of Foundations of Differential Geometry Vol I. by Kobayashi and Nomizu.

Corollary 41.9. *Let* $\pi \colon P \to M$ *be a principal* G*-bundle, where* $\dim M \geq 2$*. Then for any connected Lie subgroup* $H \subset G$*, there exists a connection* ϖ *on* P *with* $H^{\varpi}(p) = H$ *(for some* $p \in P$*) if and only if* P *admits a principal* H*-subundle.*

Remark 41.10. Corollary 41.9 means that the question as to when a given principal G-bundle admits a connection with holonomy equal to a prescribed subgroup H of G is not very geometrically interesting. Indeed, the existence (or non-existence) of a principal H-subbundle is a purely topological issue, and can be answered using algebraic topology. We will see in Lecture 44 that the situation dramatically changes if we impose the additional condition that our connection is *torsion-free*.

LECTURE 42

Geodesics and sprays

In this lecture we study *geodesics* and *sprays*. These are concepts normally associated with Riemannian geometry. However, as we will see, they make perfect sense for an arbitrary connection on a manifold. The word "geodesic" needs to be understood carefully however—in this more general setting there is no relation between geodesics and shortest paths, see Remark 42.6 below.

REMARK 42.1. **Convention:** For the remainder of the course we will almost exclusively work on the tangent bundle TM of a manifold M, rather than an arbitrary vector bundle. Thus we adopt the convention that a **connection on** M is, by definition, a connection on the vector bundle $\pi\colon TM \to M$.

A connection ∇ on M induces a connection on all the associated tensor bundles (eg. the cotangent bundle $T^*M \to M$), which we will continue to denote by ∇. In fact, by Problem P.1, ∇ defines a tensor derivation $\mathcal{T}(M) \to \mathcal{T}(M)$ in the sense of Definition 18.14, which will also be denoted by ∇.

Let $\sigma\colon U \to O$ denote a chart on M with local coordinates x^i. When no confusion is possible, we will abbreviate

$$\partial_i := \frac{\partial}{\partial x^i}$$

for the corresponding vector field on U.

DEFINITION 42.2. We define the **Christoffel symbols** of the chart σ and the connection ∇ as

$$\Gamma^k_{ij}(x) := dx^k|_x \left(\nabla_{\partial_i}(\partial_j)(x)\right)$$

Thus $\Gamma^k_{ij}\colon U \to \mathbb{R}$ is a smooth function on U.

Pay attention to the indices—the Einstein Summation Convention is very useful here.

LEMMA 42.3. *The connection ∇ is uniquely determined on U by the Christoffel symbols.*

Proof. If X and Y are any two vector fields on U then we can write $X = f^i\,\partial_i$ and $Y = g^j\,\partial_j$ for smooth functions f^i, g^j. Abbreviate

$$\partial_i g^j := \frac{\partial g^j}{\partial x^i} = dg^j(\partial_i).$$

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

Then by the axioms for a covariant derivative operator (Definition 31.8) we have

$$\begin{aligned}\nabla_X(Y) &= \nabla_{f^i\,\partial_i}(g^j\,\partial_j)\\ &= f^i\nabla_{\partial_i}(g^j\,\partial_j)\\ &= f^ig^j\nabla_{\partial_i}(\partial_j) + f^i\partial_ig^j\partial_j\\ &= \left(f^i\,g^j\,\Gamma^k_{ij} + f^i\partial_ig^k\right)\partial_k,\end{aligned}$$

where on the last line we replaced the dummy variable j by k. ■

Lemma 42.3 gives yet another viewpoint on connections: they are determined locally by n^3 (where $n = \dim M$) smooth functions Γ^k_{ij}. On Problem Sheet T you will investigate how the Christoffel symbols of two charts with overlapping domains are related.

Definition 42.4. Let ∇ be a connection on M. A curve γ in M is called a **geodesic** of ∇ if γ' is a parallel curve.

Example 42.5. Consider the connection ∇ on S^n introduced in Problem N.3. By part (ii) of Problem O.2, if x, y are two points in S^n such that $x \perp y$ then the great circle $\gamma\colon [0, 2\pi] \to S^n$ defined by $\gamma(t) = (\cos t)x + (\sin t)y$ is a geodesic.

Remark 42.6. The word "geodesic" was originally used to mean the shortest path between two points on the Earth's surface. As we will see in Lecture 52, if M is endowed with a Riemannian metric m, and the connection ∇ is the *Levi-Civita* connection (see Theorem 45.1) of (M, m), then M admits a metric d_m (in the sense of point-set topology) for which every geodesic is locally a length-minimising curve. In this lecture however, we are working with arbitrary connections on manifolds, and thus geodesics do *not* need to locally minimise lengths (and indeed, without reference to a specific metric on M the idea of "length-minimising" simply does not make sense!)

Geodesics always exist with prescribed initial conditions. The next result is a variation of Proposition 29.7.

Proposition 42.7. *Let ∇ be a connection on M, and let $(x, v) \in TM$. There exists a geodesic γ of ∇ in M with initial condition $\gamma(0) = x$ and $\gamma'(0) = v$. Moreover γ is unique up to the domain of definition.*

Proof. By the chain rule for covariant derivatives (31.7), a curve γ is a geodesic if and only if $\nabla_T(\gamma') = 0$, where $T = \frac{\partial}{\partial t} \in \mathfrak{X}(\mathbb{R})$ and ∇ (also) denotes the pullback connection along γ.

Now let σ denote a chart on M that intersects the image of γ with local coordinates x^i. Abbreviate $\gamma^i = x^i \circ \gamma$ and $c_i := \partial_i \circ \gamma$. Then[1]

$$\nabla_T(c_i) = \nabla_{\gamma'}(\partial_i). \tag{42.1}$$

[1]This notation can be confusing, and it is hard to keep track of which connection is which, particularly as they are all written ∇! The thing to remember is that whilst $X \mapsto \nabla_X(s)$ is a point operator (and thus $\nabla_v(s)$ is defined for an individual tangent vector v), the operator $s \mapsto \nabla_X(s)$ is not a point operator, and hence this expression is *only* defined when s is a section of the correct vector bundle. Thus if you are not sure which connection is being used in a given expression, consult the second variable s. In (42.1), c_i belongs to $\Gamma_\gamma(TM)$; thus the left-hand side $\nabla_T(c_i)$ must refer to the covariant derivative operator along γ. Meanwhile on the right-hand side, ∂_i is a vector field on M, and thus $\nabla_{\gamma'}(\partial_i)$ refers to the covariant derivative operator on M itself. Easy, right? Right?

Since the notation $(\gamma^i)'$ is cumbersome (due to the superscript i), we will often use the dot notation and write $\dot{\gamma}^i$ for the function $t \mapsto (\gamma^i)'(t)$, so that $\dot{\gamma}^i = T(\gamma^i)$. Thus $\gamma' = \dot{\gamma}^i\, c_i$, and so by the chain rule (31.7) applied repeatedly, we have

$$\begin{aligned} \nabla_T(\gamma') &= T(\dot{\gamma}^i)c_i + \dot{\gamma}^i \nabla_T(c_i) \\ &= \ddot{\gamma}^i c_i + \dot{\gamma}^i \nabla_{\gamma'}(\partial_i) \\ &= \left(\ddot{\gamma}^k + \dot{\gamma}^i\, \dot{\gamma}^j\, \Gamma^k_{ij}(\gamma)\right) c_k. \end{aligned}$$

This means that the equation $\nabla_T(\gamma') = 0$ is locally equivalent to the second-order system of ordinary differential equations:

$$\ddot{\gamma}^k + \dot{\gamma}^i\, \dot{\gamma}^j\, \Gamma^k_{ij}(\gamma) = 0, \qquad \forall\, 1 \le i, j, k \le n. \tag{42.2}$$

We refer to (42.2) as the **geodesic equation**. The conclusion now follows from standard existence and uniqueness theorems for ordinary differential equations. Note that in general we only get short-term existence (unless $\Gamma^k_{ij} = 0$). ∎

Lemma 42.8. *Let $\gamma\colon (a, b) \to M$ be a non-constant geodesic, and let $h\colon (a_1, b_1) \to (a, b)$ be a smooth map. Then $\gamma \circ h$ is a geodesic if and only if h is an affine map, i.e. $h'' = 0$.*

Proof. By the chain rule for covariant derivative operators (31.7) we have

$$\nabla_T((\gamma \circ h)') = h''\,(\gamma' \circ h) + h' \nabla_T(\gamma' \circ h),$$

and $\nabla_T(\gamma' \circ h) = \nabla_{Dh[T]}(\gamma') = h' \nabla_{T \circ h}(\gamma') = 0$. Since γ' is non-constant we see that $\gamma \circ h$ is a geodesic if and only if $h'' = 0$. ∎

In general it may not be possible to extend a geodesic to be defined on all of $\mathbb{R}$. The following definition is analogous to Definition 8.12.

Definition 42.9. A connection ∇ on a manifold is called **complete** if all geodesics have maximal domain of definition equal to $\mathbb{R}$.

Example 42.10. The connection ∇ on S^n from Problem N.3 is complete. Indeed, by Example 42.5 and Proposition 42.7, we see that any geodesic on S^n is a great circle of the form $\gamma(t) = (\cos t)x + (\sin t)y$ for two perpendicular points on S^n, and moreover any such geodesic may be extended to all of $\mathbb{R}$ by periodicity.

In fact, Definition 42.9 is more than analogous to Definition 8.12—it is merely a special case, as we shall now see.

Remark 42.11. **Warning:** Do not confuse the bundle $\pi_{TM}\colon TTM \to TM$ (the tangent bundle of the tangent bundle) with the bundle $D\pi\colon TTM \to TM$ from Lemma 31.4. Despite the fact that these two bundles have the same total space and the same base space, they are not the same, as they have different fibres. We will adopt the convention that—unless explicitly stated otherwise—when referring to TTM we are always implicitly using the bundle structure arising π_{TM}.

In the following, we let $\mu_a\colon TM \to TM$ and $\tilde{\mu}_a\colon TTM \to TTM$ denote scalar multiplication in the fibres in TM and TTM respectively, i.e.

$$\mu_a(x, v) := (x, av), \qquad \tilde{\mu}_a(x, v, \zeta) := (x, v, a\zeta). \tag{42.3}$$

This should not be confused with scalar multiplication in the other vector bundle $D\pi\colon TTM \to TM$, which is given by $a \boxtimes (x, v, \zeta) := (x, av, D\mu_a(v)[\zeta])$ (cf. (31.4)).

Definition 42.12. Let M be a manifold. A vector field $\mathbb{S}$ on the tangent bundle TM is called a **spray** if $D\pi \circ \mathbb{S} = \mathrm{id}_{TM}$ and

$$\mathbb{S} \circ \mu_a = \tilde{\mu}_a \circ D\mu_a \circ \mathbb{S}. \tag{42.4}$$

Remark 42.13. Every vector field on TM satisfies (by definition) the section property for the bundle $\pi_{TM}\colon TTM \to TM$. The first condition $D\pi \circ \mathbb{S} = \mathrm{id}_{TM}$ in the definition of a spray is exactly the section property for the vector bundle $D\pi\colon TTM \to TM$ (cf. Lemma 31.4).

We now prove that geodesics can be seen as integral curves of a spray.

Theorem 42.14. *Let ∇ be a connection on M. There is a unique spray $\mathbb{S}$ on M which is horizontal with respect to ∇ (i.e. $\mathbb{S}(x, v) \in \mathcal{H}_{x,v}$ for all $(x, v) \in TM$, where $\mathcal{H}$ is the distribution associated to ∇). A curve γ in M is a geodesic if and only if γ' is an integral curve of $\mathbb{S}$.*

Remark 42.15. The spray $\mathbb{S}$ constructed in Theorem 42.14 is called the **geodesic spray** of the connection ∇. The converse to Theorem 42.14 is also true: if we are given any spray $\mathbb{S}$ then there exists a connection ∇ for which $\mathbb{S}$ is the geodesic spray of ∇. This is the content of the **Ambrose-Palais-Singer Spray Theorem**, and we will prove this next lecture as Theorem 43.5.

Proof. We prove the result in two steps.

1. Let $\kappa\colon TTM \to TM$ denote connection map of ∇. The requirement that $\mathbb{S}$ is horizontal is equivalent to asking that $\kappa \circ \mathbb{S} = 0$. Recall from Lemma 31.3 that $(D\pi, \kappa)\colon TTM \to TM \oplus TM$ is a vector bundle isomorphism. We can therefore define $\mathbb{S}$ by

$$\mathbb{S}(x, v) := (D\pi(x, v), \kappa)^{-1}\big((x, v), (x, 0_x)\big).$$

Then $\mathbb{S}$ is smooth, since it is the composition

$$\mathbb{S} = (D\pi, \kappa)^{-1}(\mathrm{id}_{TM}, o \circ \pi),$$

where $o\colon M \to TM$ is the zero section. Moreover $D\pi \circ \mathbb{S} = \mathrm{id}_{TM}$ and since $D\pi|_{\mathcal{H}}$ is an isomorphism, we see that $\mathbb{S}(x, v)$ is the only horizontal vector which is mapped to (x, v) under $D\pi(x, v)$. This shows there is at most one horizontal vector field on TM which satisfies the first condition of a spray. Thus if we can prove that $\mathbb{S}$ satisfies (42.4) we will have both existence and uniqueness for $\mathbb{S}$. For this, using Lemma 31.3 again, it suffices to show that

$$D\pi \circ \mathbb{S} \circ \mu_a = D\pi \circ \tilde{\mu}_a \circ D\mu_a \circ \mathbb{S} \tag{42.5}$$

and

$$\kappa \circ \mathbb{S} \circ \mu_a = \kappa \circ \tilde{\mu}_a \circ D\mu_a \circ \mathbb{S}. \tag{42.6}$$

To see (42.5) we observe that since $D\pi$ is a linear map,

$$D\pi \circ \tilde{\mu}_a = \mu_a \circ D\pi.$$

Next, since $\pi \circ \mu_a = \pi$ we have $D\pi \circ D\mu_a = D\pi$ and thus

$$D\pi \circ \tilde{\mu}_a \circ D\mu_a = \mu_a \circ D\pi.$$

Thus if we start with $(x, v) \in TM$, the right-hand side of (42.5) is

$$\begin{aligned} D\pi \circ \tilde{\mu}_a \circ D\mu_a \circ \mathbb{S}(x, v) &= \mu_a \circ D\pi \circ \mathbb{S}(x, v) \\ &= \mu_a(x, v) \\ &= (x, av). \end{aligned}$$

Similarly if we feed (x, v) to the left-hand side we get $D\pi \circ \mathbb{S}(x, av) = (x, av)$, and thus (42.5) is proved. To prove (42.6), we again start from the fact that κ is a linear map, and hence

$$\kappa \circ \tilde{\mu}_a = \mu_a \circ \kappa.$$

Moreover κ is also a vector bundle morphism along $D\pi$ (Theorem 31.7), which means that

$$\kappa \circ D\mu_a = \mu_a \circ \kappa.$$

Thus the right-hand side of (42.6) is equal to

$$\begin{aligned} \kappa \circ \tilde{\mu}_a \circ D\mu_a \circ \mathbb{S} &= \mu_a \circ \kappa \circ D\mu_a \circ \mathbb{S} \\ &= \mu_a \circ \mu_a \circ \kappa \circ \mathbb{S} \\ &= 0 \end{aligned}$$

since $\kappa \circ \mathbb{S} = 0$. Similarly the left-hand side of (42.6) is also zero. This proves that $\mathbb{S}$ is a spray.

2. It remains to show that the geodesics of ∇ are exactly the projections to M of integral curves of $\mathbb{S}$. Let δ be an integral curve of $\mathbb{S}$, and let $\gamma := \pi \circ \delta$. Since δ' is a curve in $\mathcal{H}$, δ is parallel along γ. But

$$\begin{aligned} \gamma'(t) &= \frac{d}{ds}\Big|_{s=t} \pi(\delta(s)) \\ &= D\pi(\delta(t))[\delta'(t))] \\ &= D\pi(\delta(t))[\mathbb{S}(\delta(t))] \\ &= \delta(t), \end{aligned}$$

and thus γ' is parallel along γ, so that γ is a geodesic. Conversely if γ is a geodesic then if δ denotes the unique integral curve of $\mathbb{S}$ with $\delta(0) = \gamma'(0)$ then the argument above shows that $\pi \circ \delta$ is another geodesic with the same initial condition as γ, and hence the uniqueness part of Proposition 42.7 shows that $\delta = \gamma'$. This completes the proof. ■

We conclude this lecture by defining the *geodesic flow* of a connection.

Definition 42.16. Let ∇ denote a connection on M. The **geodesic flow** of ∇ is the maximal flow Θ of the geodesic spray $\mathbb{S}$ of ∇.

In general by Theorem 8.10, the geodesic flow is a smooth map $\Theta \colon \mathcal{D} \to TM$, where $\mathcal{D} \subset \mathbb{R} \times TM$ is an open set containing $\{0\} \times TM$. We have $\mathcal{D} = \mathbb{R} \times TM$ if and only if ∇ is complete. Explicitly if we write $\Theta_t := \Theta(t, \cdot)$ then one has

$$\Theta_t(x, v) = \big(\gamma_{x,v}(t), \gamma'_{x,v}(t)\big),$$

where $\gamma_{x,v}$ is the unique geodesic from Proposition 42.7 with initial condition $\gamma_{x,v}(0) = x$ and $\gamma'_{x,v}(0) = v$.

LECTURE 43

The Ambrose-Palais-Singer Spray Theorem

In this lecture we define the exponential map associated to a spray. We warn the reader this is *not* the same "exponential map" as the one discussed previously for Lie groups. They are however related in some cases, see Remark 10.16 for more information.

DEFINITION 43.1. Let $\mathbb{S}$ denote a spray on M with maximal flow $\Theta^{\mathbb{S}}\colon \mathcal{D} \to TM$. Let $\mathcal{S}_x \subset TM$ denote the set of tangent vectors v such that that $(1,(x,v)) \in \mathcal{D}$, and set $\mathcal{S} = \bigcup_{x \in M} \mathcal{S}_x$. Thus $\mathcal{S}$ consists of those points $(x,v) \in TM$ for which the maximal integral curve of $\mathbb{S}$ with initial condition (x,v) is defined for at least $t = 1$.

Since $\mathcal{D}$ is open by Theorem 8.10, so is $\mathcal{S}$. We claim that $\mathcal{S}_x$ is non-empty for every $x \in M$. Indeed, it follows from (42.4) that

$$\mathbb{S}(x, 0_x) = 0_x$$

(see also the first part of the proof of Theorem 43.3 below), and thus 0_x is a fixed point of the flow $\Theta^{\mathbb{S}}_t$—in particular $\Theta^{\mathbb{S}}_t(x, 0_x)$ is defined for all $t \in \mathbb{R}$.

DEFINITION 43.2. We define the **exponential map** of $\mathbb{S}$ by

$$\exp^{\mathbb{S}}\colon \mathcal{S} \to M. \qquad \exp^{\mathbb{S}}(x,v) = \pi(\Theta^{\mathbb{S}}_1(x,v)).$$

Since $\Theta^{\mathbb{S}}$ is smooth (Theorem 8.10), the map $\exp^{\mathbb{S}}$ is smooth[1]. We write

$$\exp^{\mathbb{S}}_x := \exp^{\mathbb{S}}|_{\mathcal{S}_x}\colon \mathcal{S}_x \to M.$$

THEOREM 43.3 (Properties of the exponential map). *Let $\mathbb{S}$ be a spray on a smooth manifold M^n with exponential map $\exp^{\mathbb{S}}\colon \mathcal{S} \to M$. Then:*

(i) *For each $x \in M$, $\mathcal{S}_x$ is a star-shaped neighbourhood of 0_x. Moreover if $v \in \mathcal{S}_x$ then*
$$\exp^{\mathbb{S}}(x, tv) = \pi \circ \Theta^{\mathbb{S}}_t(x,v), \qquad \forall\, t \in [0,1].$$

(ii) *For each $x \in M$, $\exp^{\mathbb{S}}_x$ satisfies*
$$D\exp^{\mathbb{S}}_x(0_x) \circ \mathcal{J}_{0_x} = \mathrm{id}_{T_xM},$$
and so $\exp^{\mathbb{S}}_x$ has maximal rank n at 0_x. Thus $\exp^{\mathbb{S}}_x$ maps a neighbourhood of 0_x in T_xM diffeomorphically onto a neighbourhood of $x \in M$.

[1] Compare this to Theorem 10.10: there we had to work a bit here to prove smoothness, since we did not start with a vector field $\mathbb{S}$.

(iii) *For*[2] *every* $x \in M$, *the map* $(\pi, \exp^{\mathbb{S}})\colon \mathcal{S} \to M \times M$ *has rank* $2n$ *at* 0_x, *and thus maps a neighbourhood of* 0_x *in* T_xM *diffeomorphically onto a neighbourhood of* (x,x) *in* $M \times M$. *Moreover if* $o\colon M \to TM$ *denotes the zero section then there exists a neighbourhood* $\mathcal{U}$ *of* $o(M)$ *such that* $(\pi, \exp^{\mathbb{S}})$ *maps* $\mathcal{U}$ *diffeomorphically onto a neighbourhood of the diagonal* $\Delta = \{(x,x) \mid x \in M\}$ *in* $M \times M$.

Proof. During the proof we will drop the superscript $\mathbb{S}$ everywhere and just write Θ_t and exp. We prove the theorem in four steps.

1. In this step we prove part (i). Fix $(x,v) \in TM$ and let $\delta\colon (t^-, t^+) \to TM$ denote the maximal integral curve of $\mathbb{S}$ with initial condition (x,v). Let $\mu_a\colon TM \to TM$ and $\tilde{\mu}_a\colon TTM \to TTM$ denote scalar multiplication in the fibres in TM and TTM respectively (cf. (42.3)). For $a > 0$ we consider the curve

$$\delta_a\colon \left(\frac{t^-}{a}, \frac{t^+}{a}\right) \to TM, \qquad \delta_a(t) := \mu_a \circ \delta(at).$$

Then

$$\begin{aligned}
\delta_a'(t) &= D\mu_a(\delta(at))\,[a\delta'(at)] \\
&= aD\mu_a(\delta(at))[\mathbb{S}(\delta(at))] \\
&= \tilde{\mu}_a \circ D\mu_a(\delta(at))[\mathbb{S}(\delta(at))] \\
&\overset{(\dagger)}{=} \mathbb{S}(\mu_a(\delta(at))) \\
&= \mathbb{S}(\delta_a(t)),
\end{aligned}$$

where $(\dagger)$ used (42.4). Thus δ_a is an integral curve of $\mathbb{S}$. Since $\delta_a(0) = av$, it follows from uniqueness of integral curves that

$$\Theta_t(x, av) = \mu_a \circ \Theta_{at}(x,v), \qquad \text{for } at \in (t^-, t^+).$$

In particular, if $v \in \mathcal{S}_x$ (so that $t^+ > 1$) then $av \in \mathcal{S}_x$ for all $0 \le a \le 1$ and moreover

$$\begin{aligned}
\exp(x, av) &= \pi \circ \Theta_1(x, av) \\
&= \pi \circ \mu_a \circ \Theta_a(x,v) \\
&= \pi \circ \Theta_a(x,v).
\end{aligned}$$

This proves part (i).

2. In this step we prove (ii). Using (i), we have:

$$\begin{aligned}
D\exp_x(0_x)\,[\mathcal{J}_{0_x}(v)] &= \frac{d}{dt}\Big|_{t=0} \exp_x(0_x + tv) \\
&= \frac{d}{dt}\Big|_{t=0} \pi \circ \Theta_t(x,v) \\
&= D\pi(x,v)[\mathbb{S}(x,v)] \\
&= v,
\end{aligned}$$

Thus $D\exp_x(0_x) \circ \mathcal{J}_{0_x} = \mathrm{id}_{T_xM}$ as claimed. The Inverse Function Theorem 5.2 then completes the proof of (ii).

[2] The proof of this statement is non-examinable.

3. In this step[3] we investigate the map $(\pi, \exp)$ in local coordinates. Let $\sigma\colon U \to O$ denote a chart on M about x with local coordinates x^i. Let $\tilde{\sigma}\colon \pi^{-1}(U) \to O \times \mathbb{R}^n$ denote the corresponding chart on TM (cf. Theorem 4.16). Write (y^i) for the local coordinates of $\tilde{\sigma}$, so that

$$y^i = \begin{cases} x^i \circ \pi, & 1 \le i \le n, \\ dx^{i-n}, & n+1 \le i \le 2n. \end{cases}$$

Similarly we let $\tau := (\sigma \circ \mathrm{pr}_1, \sigma \circ \mathrm{pr}_2)\colon U \times U \to O \times O$, where $\mathrm{pr}_j\colon U \times U \to U$ is the projection onto the jth factor for $j = 1, 2$, so that τ is a chart on $M \times M$ about (x, x). If (z^i) denote the local coordinates of τ then

$$z^i = \begin{cases} x^i \circ \mathrm{pr}_1, & 1 \le i \le n, \\ x^{i-n} \circ \mathrm{pr}_2, & n+1 \le i \le 2n. \end{cases}$$

By Lemma 4.4, we have[4]

$$D(\pi, \exp)(0_x)\left[\frac{\partial}{\partial y^j}\Big|_{0_x} \right] = \sum_{i=1}^{2n} \frac{\partial}{\partial y^j}\Big|_{0_x} (z^i \circ (\pi, \exp)) \frac{\partial}{\partial z^i}\Big|_{(x,x)}$$

For $i \le n$ we have $z^i \circ (\pi, \exp) = x^i \circ \pi = y^i$ and for $i \ge n+1$ we have $z^i \circ (\pi, \exp) = x^{i-n} \circ \exp$. Thus if $1 \le j \le n$ we have

$$D(\pi, \exp)(0_x)\left[\frac{\partial}{\partial y^j}\Big|_{0_x} \right] = \frac{\partial}{\partial z^j}\Big|_{(x,x)} + \sum_{i=1}^{n} \frac{\partial}{\partial y^j}\Big|_{0_x} (x^i \circ \exp) \frac{\partial}{\partial z^{i+n}}\Big|_{(x,x)} \tag{43.1}$$

meanwhile if $n+1 \le j \le 2n$ we have

$$D(\pi, \exp)(0_x)\left[\frac{\partial}{\partial y^j}\Big|_{0_x} \right] = \sum_{i=n+1}^{2n} \frac{\partial}{\partial y^j}\Big|_{0_x} (x^{i-n} \circ \exp) \frac{\partial}{\partial z^i}\Big|_{(x,x)}. \tag{43.2}$$

We now claim that for $n+1 \le j \le 2n$ one has

$$D\exp(0_x)\left[\frac{\partial}{\partial y^j}\Big|_{0_x} \right] = \frac{\partial}{\partial x^{j-n}}\Big|_x. \tag{43.3}$$

Indeed, if $\imath_x\colon T_xM \hookrightarrow TM$ denotes the inclusion then[5]

$$\frac{\partial}{\partial y^j}\Big|_{0_x} = D\imath_x(0_x)\left[\mathcal{J}_{0_x}\left(\frac{\partial}{\partial x^{j-n}}\Big|_x \right) \right] \tag{43.4}$$

[3]Remember, this part of the proof is non-examinable!

[4]We write the summation signs in this proof to make it clear exactly what index range we are summing over.

[5]One way to see this is to consider the curve ε_j in T_xM given by

$$\varepsilon_j(t) := \imath_x\left(t \frac{\partial}{\partial x^{j-n}}\Big|_x \right).$$

The left-hand side of (43.4) is $\varepsilon_j'(0)$ computed using (4.5), and the right-hand side of (43.4) is $\varepsilon_j'(0)$ computed using (4.4).

Since $\exp_x = \exp \circ \imath_x$, (43.3) follows from (43.4) and part (ii). Now inserting (43.3) into (43.2) and simplifying tells us that for $n+1 \le j \le 2n$ we have

$$D(\pi, \exp)(0_x)\left[\frac{\partial}{\partial y^j}\Big|_{0_x}\right] = \frac{\partial}{\partial z^j}\Big|_{(x,x)} \tag{43.5}$$

4. We now complete the proof of (iii). By (43.1) and (43.5), the matrix of $D(\pi, \exp)(0_x)$ with respect to the bases $\{\frac{\partial}{\partial z^i}\big|_{0_x}\}$ and $\{\frac{\partial}{\partial y^j}\big|_{(x,x)}\}$ is of the form

$$D(\pi, \exp)(0_x) = \begin{pmatrix} \mathrm{id} & 0 \\ * & \mathrm{id} \end{pmatrix},$$

which has rank $2n$. Thus $(\pi, \exp)$ is a diffeomorphism on a neighbourhood of $(x, x$ by the Inverse Function Theorem 5.2. The final claim that $(\pi, \exp)$ is a diffeomorphism on a neighbourhood the zero-section is a formal point-set topological consequence of what we have already proved[6]. ■

Remark 43.4. We will use part (ii) of Theorem 43.3 many times throughout the rest of the course. The stronger statement given by part (iii) will only be needed once: during our proof that the *injectivity radius* of a compact manifold is positive (see Proposition 52.16).

We now prove a converse to Theorem 42.14.

Theorem 43.5 (The Ambrose-Palais-Singer Spray Theorem). *Let M be a smooth manifold and let $\mathbb{S}$ be a spray on TM. There exists a connection ∇ on M such that $\mathbb{S}$ is the geodesic spray of ∇.*

Remark 43.6. **Warning:** This theorem is *not* asserting that there exists a unique connection ∇ for which $\mathbb{S}$ is the geodesic spray of ∇. In general, there can be many connections with the same geodesics (and hence the same geodesic spray). As we will see next lecture, if we impose in addition that the connection ∇ is *torsion-free* then the correspondence becomes bijective, i.e. for each spray $\mathbb{S}$ on M there exists precisely one torsion-free connection ∇ for which $\mathbb{S}$ is the geodesic spray of ∇ (see Corollary 44.10).

This proof is non-examinable.

(♣) *Proof.* We prove the result in four steps.

1. In this step we define for each $(x, v) \in TM$ a subspace $\mathcal{H}_{(x,v)} \subset T_{(x,v)}TM$, which will form our desired connection distribution $\mathcal{H} \subset TM$. Write $\exp = \exp^{\mathbb{S}}$ for the exponential map of $\mathbb{S}$ with domain $\mathcal{S} \subset TM$. Fix $x \in M$. For any $w \in T_xM$, the curve

$$\gamma_w(t) := \exp_x(tw).$$

[6] To be precise: *Suppose X and Y are locally compact, Hausdorff, and paracompact topological spaces and $f\colon X \to Y$ is a local homeomorphism. If $A \subset X$ is any closed set such that $f|_A$ is a homeomorphism then there exists an open set U containing A such that $f|_U$ is also a homeomorphism.*

is well-defined on some interval about 0 and satisfies $\gamma_w(0) = x$. Now let $v \in T_xM$ denote another tangent vector at x (possibly equal to w). We define a section $c_{v,w} \in \Gamma_{\gamma_w}(TM)$ by

$$c_{v,w}(t) := D\exp_x(tw)\big[\mathcal{J}_{tw}(v)\big]. \tag{43.6}$$

This makes sense: $\exp_x$ is a map $\mathcal{S}_x \to M$, and thus for $tw \in \mathcal{S}_x$, its differential $D\exp_x(tw)$ is a map from $T_{(x,tw)}\mathcal{S}_x = T_{(x,tw)}T_xM = V_{(x,tw)}TM$ to $T_{\gamma_w(t)}M$:

$$D\exp_x(tw) \colon V_{(x,tw)}TM \to T_{\gamma_w(t)}M.$$

Moreover by part (ii) of Theorem 43.3, we have

$$c_{v,w}(0) = v,$$

and thus in particular

$$\gamma_w'(0) = c_{w,w}(0) = w. \tag{43.7}$$

We define our connection $\mathcal{H}$ by declaring that these sections are all parallel:

$$\mathcal{H}_{(x,v)} := \big\{ c_{v,w}'(0) \mid w \in T_xM \big\} \subset T_{(x,v)}TM, \tag{43.8}$$

(pay attention to the order of v and w!), and then set as usual

$$\mathcal{H} := \bigsqcup_{(x,v)\in TM} \mathcal{H}_{(x,v)}.$$

2. We now prove that $\mathcal{H}$ is a preconnection on TM. This proof is similar to the proof of Step 1 of Theorem 30.1, but simpler. Fix $(x,v) \in TM$ and consider the smooth map

$$C_v \colon \mathcal{S}_x \times [0,1] \to M, \qquad C_v(w,t) := c_{v,w}(t).$$

Since $C_v(w,t) = C_v(tw,1)$ we have

$$c_{v,w}'(0) = \frac{d}{dt}\Big|_{t=0} C_v(tw,1) = DC_v(0_x,1)\big[\mathcal{J}_{0_x}(w),0\big],$$

which shows that $\mathcal{H}_{(x,v)}$ is the image of a linear map $T_xM \to T_{(x,v)}TM$, and hence is a vector space of dimension at most $n := \dim M$. But since

$$D\pi(x,v)[c_{v,w}'(0)] = \frac{d}{dt}\Big|_{t=0} \pi(c_{v,w}(t)) = \gamma_w'(0) = w$$

by (43.7), we see that $D\pi(x,v)$ maps $\mathcal{H}_{(x,v)}$ surjectively onto T_xM. Thus $\mathcal{H}_{(x,v)}$ is a vector space of dimension n which is mapped isomorphically onto T_xM by $D\pi(x,v)$. Next, since exp and $\mathcal{J} \colon TM \oplus TM \to VTM$ are smooth, and exp has maximal rank near the zero section by part (ii) of Theorem 43.3, and $\mathcal{J}$ has maximal rank everywhere, it follows that $\mathcal{H}$ is a submanifold of TTM. The construction of vector bundle charts for $\mathcal{H}$ is similar (but again, easier) to Step 3 of Theorem 30.1, and we omit the details. Thus $\mathcal{H}$ is a distribution on TM, and the computation above shows it is complementary to VTM. This prove that $\mathcal{H}$ is a preconnection.

3. In this step we show that $\mathcal{H}$ is a genuine connection. Let $\mu_a \colon (x,v) \mapsto (x, av)$ denote the usual scalar multiplication on TM. We compute:

$$\begin{aligned} D\mu_a(x,v)\big[c'_{v,w}(0)\big] &= \frac{d}{dt}\Big|_{t=0} \mu_a(c_{v,w}(t)) \\ &= \frac{d}{dt}\Big|_{t=0} D\exp_x(tw)\big[\mathcal{J}_{tw}(av)\big] \\ &= c'_{av,w}(0). \end{aligned}$$

This shows that $D\mu_a(x,v)[\mathcal{H}_{(x,v)}] \subseteq \mathcal{H}_{(x,av)}$. Since $\pi \circ \mu_a = \pi$, we have $D\pi(x,av) \circ D\mu_a(x,v) = D\pi(x,v)$, and thus it follows that $D\pi(x,av)$ maps both $D\mu_a(x,v)[\mathcal{H}_{(x,v)}]$ and $\mathcal{H}_{(x,av)}$ isomorphically onto T_xM, and thus we must have equality:

$$D\mu_a(x,v)[\mathcal{H}_{(x,v)}] = \mathcal{H}_{(x,av)}.$$

4. It remains to show that $\mathbb{S}$ is the geodesic spray of $\mathcal{H}$. Since $\mathbb{S}$ is a spray and there is at most one horizontal spray with respect to a given connection by Theorem 42.14), it suffices to show that $\mathbb{S}(x,v) \in \mathcal{H}_{(x,v)}$ for each $(x,v) \in TM$. For this, let δ denote the integral curve of $\mathbb{S}$ with initial condition (x,v), and let $\gamma \coloneqq \pi \circ \delta$. Then $\delta = \gamma'$ by the argument from the last bit of the proof of Theorem 42.14, and thus

$$\begin{aligned} \delta(s) &= \gamma'(s) \\ &= \frac{d}{dt}\Big|_{t=0} \gamma(s+t) \\ &= \frac{d}{dt}\Big|_{t=0} \exp_x(sv + tv) \\ &= D\exp_x(sv)\big[\mathcal{J}_{sv}(v)\big] \\ &= c_{v,v}(s). \end{aligned}$$

Therefore

$$\mathbb{S}(x,v) = \delta'(0) = c'_{v,v}(0) \in \mathcal{H}_{(x,v)}.$$

This completes the proof. ∎

Torsion-free connections

As we remarked last lecture (Remark 43.6), the correspondence between connections on M and sprays on M is not bijective, since different connections can have the same geodesics (and hence also the same geodesic spray). The aim of this lecture is to introduce a special type of connection, called a *torsion-free* connection, which *is* uniquely determined by its geodesics.

Recall from Problem Q.2 that if ∇^1 and ∇^2 are two connections on M then their difference

$$A(X,Y) := \nabla^1_X(Y) - \nabla^2_X(Y)$$

is an element of $\mathcal{T}^{1,2}(M)$, i.e. a tensor of type $(1,2)$.

LEMMA 44.1. *Two connections ∇^1 and ∇^2 have the same geodesic spray if and only if their difference A is skew-symmetric.*

Proof. From the proof of Theorem 42.14, if $\mathbb{S}_i$ is the geodesic spray of ∇^i then

$$\mathbb{S}_i(x,v) = D\pi(x,v)|^{-1}_{\mathcal{H}_i|_{x,v}}(v),$$

where $\mathcal{H}_i \subset TTM$ is the connection distribution of ∇^i. By part (ii) of Problem Q.2, we have

$$\mathcal{H}_2|_{x,v} = \left\{ \zeta + \mathcal{J}_v\big(\tilde{A}(D\pi(x,v)[\zeta])\big) \mid \zeta \in \mathcal{H}_1|_{x,v} \right\},$$

where $\tilde{A}(X)(Y) := A(X,Y)$, and hence

$$\begin{aligned} \mathbb{S}_2(x,v) &= D\pi(x,v)|^{-1}_{\mathcal{H}_2|_{x,v}}(v) \\ &= D\pi(x,v)|^{-1}_{\mathcal{H}_1|_{x,v}}(v) + \mathcal{J}_v(A(v,v)) \\ &= \mathbb{S}_1(x,v) + \mathcal{J}_v(A(v,v)). \end{aligned}$$

Thus $\mathbb{S}_1 = \mathbb{S}_2$ if and only if $A(v,v) = 0$ for all v, i.e. that A is skew-symmetric. ∎

This motivates the following definition.

DEFINITION 44.2. Let ∇ be a connection on M. The **torsion tensor** T^∇ of ∇ is the tensor of type $(1,2)$ defined by

$$T^\nabla(X,Y) := \nabla_X(Y) - \nabla_Y(X) - [X,Y], \qquad X,Y \in \mathfrak{X}(M).$$

As with the curvature tensor, merely calling T^∇ a tensor does not make it one. In contrast to Theorem 33.9 however, the verification that T^∇ is a tensor is much easier.

LEMMA 44.3. *The torsion tensor T^∇ is an alternating tensor.*

Proof. By the Tensor Criterion (Theorem 18.3) we need only check that T^∇ is $C^\infty(M)$-linear in both variables. Take $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$. Then

$$\begin{aligned} T^\nabla(fX, Y) &= \nabla_{fX}(Y) - \nabla_Y(fX) - [fX, Y] \\ &= f\nabla_X(Y) - Y(f)X - f\nabla_Y(X) - f[X, Y] + Y(f)X \\ &= fT^\nabla(X, Y) + 0, \end{aligned}$$

where we used Problem D.4 and the now familiar properties of a covariant derivative operator (parts (ii) and (iv) of Definition 31.8). It is clear that T^∇ is alternating, and thus T^∇ is also $C^\infty(M)$-linear in the second variable. ■

DEFINITION 44.4. A connection ∇ is said to be **torsion-free** if $T^\nabla = 0$.

REMARK 44.5. Many textbooks call a torsion-free connection a "symmetric" connection. The motivation for this is the following: if ∇ is a torsion free connection then the Christoffel symbols Γ_{ij}^k associated to any chart σ on M (cf. Definition 42.2) are symmetric in i and j. Indeed, given any connection ∇, if $\sigma\colon U \to O$ is a chart on M with local coordinates (x^i) then the local expression for T^∇ on U with respect to σ is (cf. Definition 16.13):

$$T^\nabla = T_{ij}^k \, \partial_k \otimes dx^i \otimes dx^j$$

where the $T_{ij}^k\colon U \to \mathbb{R}$ are the smooth functions given by

$$T_{ij}^k = dx^k\big(T(\partial_i, \partial_j)\big).$$

But since $[\partial_i, \partial_j] = 0$ by Problem D.3, it follows that

$$T_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k.$$

In particular, $T^\nabla = 0$ if and only if for every local coordinate system (x^i) one has $\Gamma_{ij}^k = \Gamma_{ji}^k$.

REMARK 44.6. It is easy to turn any connection into a torsion-free one. Indeed, if ∇ is a connection then $\nabla_1 := \nabla - \frac{1}{2}T^\nabla$ is another connection by Problem Q.2, and it follows immediately from the definition that $T^{\nabla_1} = 0$.

The next theorem gives us yet another way to view connections: namely, specifying a connection on M is the same thing as specifying the geodesics and the torsion tensor.

THEOREM 44.7. *Let ∇^1 and ∇^2 denote two connections on M. Then $\nabla^1 = \nabla^2$ if and only if ∇^1 and ∇^2 have the same geodesics and the same torsion tensors.*

Proof. Let $A := \nabla^1 - \nabla^2$, and decompose A into its symmetric and alternating parts: $A = A^s + A^a$, i.e.

$$A^s(X, Y) := \frac{1}{2}\left(A(X, Y) + A(Y, X)\right), \qquad A^a(X, Y) := \frac{1}{2}\left(A(X, Y) - A(Y, X)\right)$$

In Lemma 44.1 we already showed that $A^s = 0$ if and only if ∇^1 and ∇^2 have the same geodesics. Thus if suffices to show that $A^a = 0$ if and only if $T^{\nabla^1} = T^{\nabla^2}$. But this is immediate from:

$$\begin{aligned} 2A^a(X,Y) &= A(X,Y) - A(Y,X) \\ &= \nabla^1_X(Y) - \nabla^2_X(Y) - \nabla^1_Y(X) + \nabla^2_Y(X) \\ &= T^{\nabla^1}(X,Y) - T^{\nabla^2}(X,Y). \end{aligned}$$

This completes the proof. ∎

If $\varphi \colon M \to N$ is a smooth map and ∇ is a connection on N, then the pullback connection (also denoted by ∇) on M is a map

$$\nabla \colon \mathfrak{X}(M) \times \Gamma_\varphi(TN) \to \Gamma_\varphi(TN).$$

If X is a vector field in M then $x \mapsto D\varphi(x)[X(x)]$ is a well-defined element of $\Gamma_\varphi(TN)$—this is true even if φ is not a diffeomorphism and so $\varphi_\star(X)$ is not defined!—which we write simply as $D\varphi[X]$. Thus the expression

$$T^\nabla_\varphi \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \Gamma_\varphi(TN)$$

given by

$$T^\nabla_\varphi(X,Y) := \nabla_X(D\varphi[Y]) - \nabla_Y(D\varphi[X]) - D\varphi[X,Y]$$

is well defined. Next, since T^∇ is a tensor, it is in particular a point operator, and hence $T^\nabla(v,w)$ is defined for individual tangent vectors v, w. Thus the expression $T^\nabla(D\varphi[X], D\varphi[Y])$ is also a well defined section along φ. The next result is the analogue of Proposition 34.11 for the torsion tensor..

Proposition 44.8. *Let ∇ denote a connection on a smooth manifold N, and let $\varphi \colon M \to N$ denote a smooth map. Then for any $X, Y \in \mathfrak{X}(M)$, one has*

$$T^\nabla(D\varphi[X], D\varphi[Y]) = T^\nabla_\varphi(X,Y)$$

as elements of $\Gamma_\varphi(TN)$.

The proof of Proposition 44.8 is almost identical to that of Proposition 34.11, and we leave the details to you. The Ambrose-Palais-Singer Spray Theorem 43.5 constructed a connection from a spray. In fact, this connection is torsion-free, as we now prove.

Proposition 44.9. *Let $\mathbb{S}$ denote a spray on M, and let ∇ denote the connection constructed in the proof of the Ambrose-Palais-Singer Spray Theorem 43.5. Then ∇ is torsion-free.*

This proof is non-examinable.

(♣) *Proof.* Fix $x \in M$ and $v, w \in T_xM$. We will prove that $T^\nabla(v,w) = 0$ in two steps.

1. In this step we derive an expression for $T^\nabla(v,w)$. There is a well-defined vector field $\mathcal{J}(v) \in \mathfrak{X}(T_xM)$ defined by

$$\mathcal{J}(v)(w) := \mathcal{J}_w(v) = \frac{d}{dt}\Big|_{t=0} w + tv.$$

Write $\exp = \exp^{\mathbb{S}}$ for the exponential map of $\mathbb{S}$ with domain $\mathcal{S} \subset TM$, with $\exp_x \colon \mathcal{S}_x \to M$ the restriction to the fibre over x. Then for any $v \in T_xM$, we may regard $\mathcal{J}(v)$ as a vector field on $\mathcal{S}_x$, and hence (using the notation above), $D\exp_x[\mathcal{J}(v)]$ is a vector field along $\exp_x$, which we abbreviate by $\mathbb{X}_v$. Moreover by part (ii) of Theorem 43.3, this vector field satisfies

$$\mathbb{X}_v(0_x) = D\exp_x(0_x)[\mathcal{J}_{0_x}(v)] = v.$$

By Proposition 44.8, we have

$$T^\nabla(\mathbb{X}_v, \mathbb{X}_w) = T^\nabla_{\exp_x}(\mathcal{J}(v), \mathcal{J}(w)),$$

and evaluating both sides at 0_x tells us that

$$T^\nabla(v,w) = \nabla_v(\mathbb{X}_w) - \nabla_w(\mathbb{X}_v) - [\mathbb{X}_v, \mathbb{X}_w](0_x). \tag{44.1}$$

2. In this step we compute the right-hand side of (44.1). Since $\mathcal{J}(v)$ and $\mathcal{J}(w)$ are constant vector fields, the Lie bracket $[\mathcal{J}(v), \mathcal{J}(w)]$ is zero by Problem D.3 (note this is a Lie bracket of vector fields on the vector space T_xM). Thus by Problem D.5 we also have $[\mathbb{X}_v, \mathbb{X}_w] = 0$. Now let $\mathcal{H}$ denote the connection distribution from (43.8) and let $\kappa \colon TTM \to TM$ denote the connection map of ∇. Then by definition (cf. Theorem 31.10) one has

$$\begin{aligned}\nabla_v(\mathbb{X}_w) &= \kappa\big(D\mathbb{X}_w(0_x)[\mathcal{J}(v)(0_x)]\big)\\ &= \kappa\left(\frac{d}{dt}\Big|_{t=0} D\exp_x(tv)[\mathcal{J}_{tv}(w)]\right).\end{aligned}$$

This last term is exactly $\kappa(c'_{w,v}(0))$ using the notation from the proof of Theorem 43.5 (cf. (43.6)). But now $c'_{w,v}(0) \in \mathcal{H}_{x,w}$ by definition, and thus as $\mathcal{H} = \ker\kappa$ we conclude that

$$\nabla_v(\mathbb{X}_w) = 0.$$

Similarly $\nabla_w(\mathbb{X}_v) = 0$ and thus by (44.1) we have $T^\nabla(v,w) = 0$. This completes the proof. ∎

This gives us the following strengthening of the Ambrose-Palais-Singer Spray Theorem.

Corollary 44.10. *Let $\mathbb{S}$ be a spray on M and let T be an alternating tensor of type $(1,2)$. There exists a unique connection on M with geodesic spray $\mathbb{S}$ and torsion tensor T.*

Proof. Let ∇ denote the connection on M given by the Ambrose-Palais-Singer Spray Theorem. Then ∇ has geodesic spray $\mathbb{S}$ and ∇ is torsion-free by Proposition 44.9. The desired connection is then given by $\nabla_1 := \nabla + \frac{1}{2}T$ (apply Remark 44.6 backwards). This connection is unique by Theorem 44.7. ∎

One of the most useful consequences of Corollary 44.10 is that if we start with a torsion-free connection we now have an explicit formula for the horizontal distribution in terms of the exponential map of the geodesic spray of ∇ (i.e. (43.8)). Here is an application of this, which will aid our forthcoming computations in Riemannnian geometry.

Proposition 44.11. *Let ∇ be a torsion-free connection on M^n. Fix $x \in M$ and let $\{v_1, \dots, v_n\}$ be a basis of T_xM. There exists a chart $\sigma \colon U \to O$ on M, where $x \in U$ and $0 \in O$ such that:*

(i) $\sigma(x) = 0$,

(ii) $\partial_i|_x = v_i$,

(iii) $\nabla_w(\partial_i) = 0$ *for all* $w \in T_xM$,

(iv) $\nabla_w(dx^i) = 0$ *for all* $w \in T_xM$ *(for ∇ is the induced connection on T^*M).*

Note that by (iii) we have that in these coordinates the Christoffel symbols vanish at x: $\Gamma^k_{ij}(x) = 0$ for all i, j, k.

Proof. Let $T \colon T_xM \to \mathbb{R}^n$ denote the linear isomorphism determined by $Tv_i = e_i$. Let exp denote the exponential map of the geodesic spray of ∇. Let $V \subset T_xM$ to be a neighbourhood of 0_x on which $\exp_x$ is a diffeomorphism (such V exists by part (ii) of Theorem 43.3). Let $U := \exp_x(V)$ and $O := T(V)$. Then if we define

$$\sigma := T \circ \exp_x|_V^{-1}$$

then (i) is clear. By construction we have

$$\partial_i\big|_{\exp_x(v)} = D\exp_x(v)[\mathcal{J}_v(v_i)],$$

and so taking $v = 0$ and applying part (ii) of Theorem 43.3 gives (ii). To prove (iii), we consider the curve $c(t) = tw$ in T_xM. Then thinking of ∂_i as a smooth map $U \to TU$, we have

$$D\partial_i(x)[w] = \frac{d}{dt}\bigg|_{t=0} D\exp_x(tw)[\mathcal{J}_{tw}(v_i)]$$

which belongs to the connection distribution $\mathcal{H}$ of ∇ at (x, v) by by (43.8). Thus if κ denotes the connection map of ∇ then

$$\nabla_w(\partial_i) = \kappa(D\partial_i(x)[w]) = 0$$

as $\ker\kappa = \mathcal{H}$. This proves property (iii). Finally property (iv) is immediate from (iii) and the definition (Problem O.3) of the induced connection on T^*M, since $dx^i(\partial_j) = \delta^i_j$. ■

A torsion-free connection enjoys some additional symmetry properties of its curvature tensor.

Proposition 44.12 (Additional symmetries of torsion-free curvature tensors). *Let ∇ be a torsion-free connection on M with curvature tensor R^∇. Then for all $X, Y, Z \in \mathfrak{X}(M)$, one has:*

(i) $R^\nabla(X,Y)(Z) + R^\nabla(Y,Z)(X) + R^\nabla(Z,X)(Y) = 0.$

(ii) $\big(\nabla_X R^\nabla\big)(Y,Z) + \big(\nabla_Y R^\nabla\big)(Z,X) + \big(\nabla_Z R^\nabla\big)(X,Y) = 0.$

Proof. Since R^∇ is a point operator in all three variables, it is sufficient to prove the result in the special case where $[X,Y] = [Y,Z] = [Z,X] = 0$. Then $\nabla_X(Y) = \nabla_Y(X)$, $\nabla_Y(Z) = \nabla_Z(Y)$, and $\nabla_Z(X) = \nabla_X(Z)$, and hence

$$
\begin{aligned}
&R^\nabla(X,Y)(Z) + R^\nabla(Y,Z)(X) + R^\nabla(Z,X)(Y) \\
&\quad = \nabla_X(\nabla_Y(Z)) - \nabla_Y(\nabla_X(Z)) + \nabla_Y(\nabla_Z(X)) \\
&\qquad - \nabla_Z(\nabla_Y(X)) + \nabla_Z(\nabla_X(Y)) - \nabla_X(\nabla_Z(Y)) \\
&\quad = \nabla_X\big(\nabla_Y(Z) - \nabla_Z(Y)\big) + \nabla_Y\big(\nabla_Z(X) - \nabla_X(Z)\big) + \nabla_Z\big(\nabla_X(Y) - \nabla_Y(X)\big) \\
&\quad = 0 + 0 + 0.
\end{aligned}
$$

This proves (i). The proof of (ii) is on Problem Sheet T. ■

REMARK 44.13. In the literature the two identities (i) and (ii) are often somewhat confusingly referred to as the "First Bianchi Identity" and the "Second Bianchi Identity" respectively. We will avoid this nomenclature since we already have two "Bianchi Identities" (Theorem 36.1 and (39.6) from Theorem 39.10)!

Proposition 38.6 tells us that there is a bijective correspondence between connections on M and connections on the principal bundle $\mathrm{Fr}(TM)$. Thus we can also unambiguously define a principal bundle connection ϖ on $\mathrm{Fr}(TM)$ to be **torsion-free** if the corresponding vector bundle connection ∇ on M is torsion-free. For the rest of this lecture we will suppress this bijection from Proposition 38.6 and simply regard connections on M as being the same as connections on $\mathrm{Fr}(TM)$.

Let us now briefly survey how the torsion-free condition affects the possible holonomy groups that can arise. As explained at the end of Lecture 41, the general question as to which Lie groups can arise as the holonomy group of a connection on a given principal bundle is not very interesting (see Remark 41.10). If however we work with torsion-free connections, this dramatically changes.

Consider the following question:

- *Let M be a connected smooth manifold of dimension n. What Lie subgroups $G \subset \mathrm{GL}(n,\mathbb{R})$ can occur as holonomy groups for torsion-free connections on M?*

This is an extremely difficult problem in general, and is an open problem for many manifolds M. We can simplify things by turning the question on its head and starting with the Lie group.

- *Let $G \subset \mathrm{GL}(n,\mathbb{R})$ be a Lie subgroup. Does there exist* **any** *manifold M^n and a torsion-free connection ∇ on M such that G is the holonomy group of ∇?*

This is still very hard, but a complete classification is (mostly) understood. We conclude this lecture by outlining why. The key starting point is the two additional symmetries from Proposition 44.12.

Definition 44.14. Let V be a vector space and suppose G is a Lie subgroup of $\mathrm{GL}(V)$ with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(V)$. We define two subspaces as follows:

$$K(\mathfrak{g}) := \left\{ R \in \bigwedge^2(V^*) \otimes \mathfrak{g} \mid R(u,v)w + R(v,w)(u) + R(w,u)(v) = 0, \ \forall\, u,v,w \in V \right\},$$

and

$$\tilde{K}(\mathfrak{g}) := \{\rho \in V^* \otimes K(\mathfrak{g}) \mid \rho(u)(v,w) + \rho(v)(w,u) + \rho(w)(u,v) = 0, \ \forall\, u,v,w \in V\}$$

Finally define

$$\mathfrak{k}(\mathfrak{g}) := \{R(u,v) \mid R \in K(\mathfrak{g}), \ u,v \in V\}.$$

Definition 44.15. We say that $G \subset \mathrm{GL}(V)$ is a **Berger subgroup** if its Lie algebra $\mathfrak{g}$ satisfies:

(i) $\tilde{K}(\mathfrak{g}) \neq \{0\}$.

(ii) $\mathfrak{k}(\mathfrak{g}) = \mathfrak{g}$.

The next result gives a necessary condition for a Lie subgroup to occur as the holonomy group of a torsion-free connection.

Theorem 44.16. *Let M be connected manifold of dimension n, and suppose $G \subset \mathrm{GL}(n, \mathbb{R})$. Assume that G is irreducible and M is not locally symmetric*[1]*. If G is the holonomy group of a torsion-free connection then G is necessarily a Berger group.*

(♣) *Proof.* The idea is very simple: if G is the holonomy group of a torsion-free connection ∇, then the curvature tensor defines an element of $\tilde{K}(\mathfrak{g})$ by Proposition 44.12. Thus $\tilde{K}(\mathfrak{g})$ is not zero. On the other hand, the Ambrose-Singer Holonomy Theorem 34.8 tells us that $\mathfrak{k}(\mathfrak{g})$ is all of $\mathfrak{g}$. ■

Theorem 44.16 allows us to rule out many Lie groups (i.e. all the non-Berger groups). This however is merely the "easy" half of answering the second question posed above—to show that a Lie group really does appear as a holonomy group, one needs to explicitly construct a connection. Unlike Theorem 41.8, there is no easy way to construct a connection "by hand". In 1999, a complete classification of those groups that could appear was obtained by Merkulov and Schwachhöfer. The list is rather long, and we will not attempt to enumerate it here.

We remark however that the list gets *much* shorter if we require that ∇ is not only torsion-free, but in addition is Riemannian with respect to some Riemannian metric on M (in other words, that ∇ is a Levi-Civita connection with respect to some Riemannian metric on M). The holonomy groups that can arise for such ∇ are the so-called **Riemannian holonomy groups**. We will come back to this at the end of the next lecture.

[1]For the purposes of our discussion here, just ignore these two conditions. Defining them precisely would take us too far afield, and it is not necessary to understand the general "idea".

LECTURE 45

The Fundamental Theorem of Riemannian Geometry

In this lecture we begin our study of Riemannian geometry proper, starting with the construction of the famous *Levi-Civita connection* of a Riemannian manifold. Let M be a smooth manifold, and suppose $m = \langle \cdot, \cdot \rangle$ is a Riemannian metric on M (i.e. a Riemannian metric on the vector bundle TM). Recall a connection ∇ on M is said to be Riemannian with respect to m if m is parallel with respect to the induced connection on $T^*M \otimes T^*M$: $\nabla m = 0$. By Proposition 36.15 this is equivalent to asking that the *Ricci Identity* holds:

$$X \langle Y, Z \rangle = \langle \nabla_X(Y), Z \rangle + \langle Y, \nabla_X(Z) \rangle, \qquad \forall\, X, Y, Z \in \mathfrak{X}(M). \tag{45.1}$$

In Proposition 36.17 we proved that Riemannian connections always exist. Meanwhile the Ambrose-Palais-Singer Theorem 43.5 (together with Proposition 44.9) proved that torsion-free connections exist. But can we satisfy both conditions simultaneously? The following somewhat grandiosely named theorem asserts that the answer is yes in the best possible way: there is a *unique* connection on M with both these properties.

THEOREM 45.1 (The Fundamental Theorem of Riemannian Geometry). *Let $m = \langle \cdot, \cdot \rangle$ be a Riemannian metric on M. There exists a unique connection ∇ on M which is Riemannian with respect to m and torsion-free. We call ∇ the* **Levi-Civita connection** *of m.*

Proof. We first deal with uniqueness. Suppose that ∇ is a torsion-free connection on M which is Riemannian with respect to m. Let $X, Y, Z \in \mathfrak{X}(M)$. We combine the Ricci Identity (45.1) together with the torsion-free condition:

$$\langle \nabla_X(Y), Z \rangle - \langle \nabla_Y(X), Z \rangle = \langle [X, Y], Z \rangle,$$

to obtain

$$\begin{aligned} X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle &= \langle \nabla_X(Y), Z \rangle + \langle Y, \nabla_X(Z) \rangle + \langle \nabla_Y(Z), X \rangle \\ &\quad + \langle Z, \nabla_Y(X) \rangle - \langle \nabla_Z(X), Y \rangle - \langle X, \nabla_Z(Y) \rangle \\ &= 2 \langle \nabla_X(Y), Z \rangle - \langle [X, Y], Z \rangle + \langle [X, Z], Y \rangle + \langle [Y, Z], X \rangle, \end{aligned}$$

and hence

$$\begin{aligned} \langle \nabla_X(Y), Z \rangle = \frac{1}{2}\Big(& X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle \\ & - \langle [Y, Z], X \rangle + \langle [Z, X], Y \rangle + \langle [X, Y], Z \rangle \Big). \end{aligned} \tag{45.2}$$

With this in mind, let us define a function

$$\omega_{X,Y} \colon \mathfrak{X}(M) \to C^\infty(M)$$

by declaring that $\omega_{X,Y}(Z)$ is the right-hand side of (45.2). We claim that $\omega_{X,Y}$ is actually a one-form on M. By Corollary 16.29 (which is itself a special case of the Differential Form Criterion Theorem 19.1) we must show that $\omega_{X,Y}$ is $C^\infty(M)$-linear. For this we compute:

$$\begin{aligned}
\omega_{X,Y}(fZ) =& \frac{1}{2}\Big(X\langle Y, fZ\rangle + Y\langle fZ, X\rangle - fZ\langle X, Y\rangle \\
&- \langle [Y, fZ], X\rangle + \langle [fZ, X], Y\rangle + \langle [X,Y], fZ\rangle\Big) \\
&= f\omega_{X,Y}(Z) + \frac{1}{2}\Big(X(f)\langle Y, Z\rangle + Y(f)\langle Z, X\rangle \\
&- X(f)\langle Y, Z\rangle - Y(f)\langle X, Z\rangle\Big) \\
&= f\omega_{X,Y}(Z) + 0.
\end{aligned}$$

Since $\omega_{X,Y}$ is a one-form, by Problem R.2 there is a unique well-defined vector field $(\omega_{X,Y})^\sharp$ on M obtained via the musical isomorphism with respect to m. Then

$$\nabla_X(Y) = (\omega_{X,Y})^\sharp. \tag{45.3}$$

Since $(\omega_{X,Y})^\sharp$ is defined independently of ∇, this establishes uniqueness.

For existence, we simply turn this argument on its head and *define* ∇ by (45.3). For this to make sense we need to prove that does indeed define a torsion-free connection which is Riemannian with respect to m. This is a series of straightforward, but rather lengthy computations. We must verify:

(i) $\nabla_{fX}(Y) = f\nabla_X(Y)$,

(ii) $\nabla_X(fY) = X(f)Y + f\nabla_X(Y)$,

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

(iv) $\langle \nabla_X(Y), Z\rangle + \langle Y, \nabla_X(Z)\rangle = X\langle Y, Z\rangle$,

as the remaining conditions are all trivial.

For (i), observe that

$$\begin{aligned}
2\langle \nabla_{fX}(Y), Z\rangle &= fX\langle Y, Z\rangle + Y\langle Z, fX\rangle - Z\langle fX, Y\rangle \\
&\quad - \langle [Y,Z], fX\rangle + \langle [Z, fX], Y\rangle + \langle [fX, Y], Z\rangle \\
&= f\Big(X\langle Y, Z\rangle - Y\langle Z, X\rangle - Z\langle X, Y\rangle \\
&\quad - \langle [Y,Z], X\rangle + \langle [Z,X], Y\rangle + \langle [X,Y], Z\rangle\Big) \\
&\quad + Y(f)\langle Z, X\rangle - Z(f)\langle X, Y\rangle + Z(f)\langle X, Y\rangle - Y(f)\langle X, Z\rangle \\
&= 2f\langle \nabla_X(Y), Z\rangle + 0.
\end{aligned}$$

To prove (ii), we see that

$$\begin{aligned}
2\left\langle\nabla_X(fY), Z\right\rangle &= X\left\langle fY, Z\right\rangle + fY\left\langle Z, X\right\rangle - Z\left\langle X, fY\right\rangle \\
&\quad - \left\langle [fY, Z], X\right\rangle + \left\langle [Z, X], fY\right\rangle + \left\langle [X, fY], Z\right\rangle \\
&= f\Big(X\left\langle Y, Z\right\rangle - Y\left\langle Z, X\right\rangle - Z\left\langle X, Y\right\rangle \\
&\quad - \left\langle [Y, Z], X\right\rangle + \left\langle [Z, X], Y\right\rangle + \left\langle [X, Y], Z\right\rangle\Big) \\
&\quad + X(f)\left\langle Y, Z\right\rangle - Z(f)\left\langle X, Y\right\rangle + Z(f)\left\langle Y, X\right\rangle + X(f)\left\langle Y, Z\right\rangle \\
&= 2f\left\langle\nabla_X(Y), Z\right\rangle + 2X(f)\left\langle Y, Z\right\rangle .
\end{aligned}$$

To prove (iii), we compute

$$\begin{aligned}
2\left\langle\nabla_X(Y), Z\right\rangle - 2\left\langle\nabla_Y(X), Z\right\rangle &= X\left\langle Y, Z\right\rangle - Y\left\langle Z, X\right\rangle - Z\left\langle X, Y\right\rangle \\
&\quad - \left\langle [Y, Z], X\right\rangle + \left\langle [Z, X], Y\right\rangle + \left\langle [X, Y], Z\right\rangle \\
&\quad - Y\left\langle X, Z\right\rangle + X\left\langle Z, Y\right\rangle + Z\left\langle Y, X\right\rangle \\
&\quad + \left\langle [X, Z], Y\right\rangle - \left\langle [Z, Y], X\right\rangle - \left\langle [Y, X], Z\right\rangle \\
&= - \left\langle [Y, Z], X\right\rangle - \left\langle [X, Z], Y\right\rangle + \left\langle [X, Y], Z\right\rangle \\
&\quad + \left\langle [X, Z], Y\right\rangle + \left\langle [Y, Z], X\right\rangle + \left\langle [X, Y], Z\right\rangle \\
&= 2\left\langle [X, Y], Z\right\rangle ,
\end{aligned}$$

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

Finally, to prove (iv) we compute

$$\begin{aligned}
2\left\langle\nabla_X(Y), Z\right\rangle + 2\left\langle Y, \nabla_X(Z)\right\rangle &= X\left\langle Y, Z\right\rangle + Y\left\langle Z, X\right\rangle - Z\left\langle X, Y\right\rangle \\
&\quad - \left\langle [Y, Z], X\right\rangle + \left\langle [Z, X], Y\right\rangle + \left\langle [X, Y], Z\right\rangle \\
&\quad + X\left\langle Z, Y\right\rangle + Z\left\langle Y, X\right\rangle - Y\left\langle X, Z\right\rangle \\
&\quad - \left\langle [Z, Y], X\right\rangle + \left\langle [Y, X], Z\right\rangle + \left\langle [X, Z], Y\right\rangle \\
&= 2X\left\langle Y, Z\right\rangle .
\end{aligned}$$

This completes the proof of existence. ■

We can use (45.2) to express the Levi-Civita connection in local coordinates. Suppose $\sigma\colon U \to O$ is a chart on M with local coordinates (x^i). Then we can write

$$m = m_{ij}\, dx^i \otimes dx^j$$

on U, where

$$m_{ij}\colon U \to \mathbb{R}, \qquad m_{ij} := \left\langle\partial_i, \partial_j\right\rangle$$

Note that the matrix $(m_{ij}(x))_{1\le i,j\le n}$ is symmetric and positive definite for every $x \in U$.

Lemma 45.2. *Let (M^n, m) be a Riemannian manifold and let ∇ denote the Levi-Civita connection of m. Let $\sigma\colon U \to O$ denote a chart on M with local coordinates x^i. Then the Christoffel symbols of ∇ are given by*

$$\Gamma^k_{ij} = \frac{1}{2} m^{kl}\left(\partial_i m_{jl} + \partial_j m_{li} - \partial_l m_{ij}\right),$$

where $(m^{ij})_{1\le i,j\le n}$ is the inverse matrix to $(m_{ij})_{1\le i,j\le n}$.

Proof. Firstly we have (this is true for any connection)

$$2\left\langle \nabla_{\partial_i}(\partial_j), \partial_l \right\rangle = 2\left\langle \Gamma^k_{ij}\partial_k, \partial_l \right\rangle = 2\Gamma^k_{ij} m_{kl}.$$

Now by (45.2) we have

$$\begin{aligned} 2\left\langle \nabla_{\partial_i}(\partial_j), \partial_l \right\rangle &= \partial_i \left\langle \partial_j, \partial_l \right\rangle + \partial_j \left\langle \partial_l, \partial_i \right\rangle - \partial_l \left\langle \partial_i, \partial_j \right\rangle \\ &= \partial_i m_{jl} + \partial_j m_{li} - \partial_l m_{ij}, \end{aligned}$$

since the Lie bracket terms $[\partial_i, \partial_j]$ all vanish by Problem D.3. Thus

$$2\Gamma^k_{ij} m_{kl} = \partial_i m_{jl} + \partial_j m_{li} - \partial_l m_{ij},$$

Multiply both sides by $\frac{1}{2} m^{pl}$ and sum over[1] l to get

$$\Gamma^k_{ij} m_{kl} m^{pl} = \frac{1}{2} m^{pl} \left(\partial_i m_{jl} + \partial_j m_{li} - \partial_l m_{ij} \right). \tag{45.4}$$

But

$$m_{kl} m^{lp} = \delta^p_k.$$

(this is the definition of the inverse matrix) and hence the left-hand side of (45.4) is

$$\Gamma^k_{ij} m_{kl} m^{pl} = \Gamma^k_{ij} \delta^p_k.$$

Thus in particular taking $p = k$ on the right-hand side of (45.4) gives

$$\Gamma^k_{ij} = \frac{1}{2} m^{kl} \left(\partial_i m_{jl} + \partial_j m_{li} - \partial_l m_{ij} \right)$$

as desired. ∎

COROLLARY 45.3. *Let (M, m) be a Riemannian manifold and let ∇ denote the Levi-Civita connection of m. For any point $x \in M$ there exists a chart σ about x with local coordinates (x^i) such that $\{\partial_i|_x\}$ is an orthonormal basis at x and such that the Christoffel symbols vanish at x: $\Gamma^k_{ij}(x) = 0$ for all i, j, k.*

Such coordinates are called **normal coordinates** at x.

Proof. Choose an orthonormal basis $\{v_i\}$ of $T_x M$ and apply Corollary 44.11. ∎

REMARK 45.4. One can alternatively characterise normal coordinates in terms of the first derivatives of the metric. Indeed, in any local coordinates (x^i) one has

$$\begin{aligned} \partial_k m_{ij} &= \partial_k \left\langle \partial_i, \partial_j \right\rangle \\ &\overset{(\dagger)}{=} \left\langle \nabla_{\partial_k}(\partial_i), \partial_j \right\rangle + \left\langle \partial_i, \nabla_{\partial_k}(\partial_j) \right\rangle \\ &= \left\langle \Gamma^l_{ki}\partial_l, \partial_j \right\rangle + \left\langle \partial_i, \Gamma^l_{kj}\partial_l \right\rangle \\ &= \Gamma^l_{ki} m_{lj} + \Gamma^l_{kj} m_{il}, \end{aligned}$$

where $(\dagger)$ used the Ricci identity. Thus coordinates (x^i) are normal at x if and only if $\{\partial_i|_x\}$ is an orthonormal basis at x and

$$\partial_k m_{ij}(x) = 0, \qquad \forall\, i, j, k. \tag{45.5}$$

[1]The summation over l is forced by the Einstein Summation Convention.

Remark 45.5. If (x^i) are normal coordinates at x and $v = a^i\,\partial_i$ is a tangent vector at T_xM then the unique geodesic $\gamma_{x,v}$ with $\gamma_{x,v}(0) = x$ and $\gamma'_{x,v}(0) = v$ is given by

$$\gamma_{x,v}(t) = \sigma^{-1}(ta^1, \dots, ta^n),$$

for all t sufficiently small, where σ is the chart corresponding to (x^i). This follows from the proof of Proposition 44.11.

The next result shows how the Levi-Civita connection behaves nicely with respect to pullbacks.

Proposition 45.6. *Let (N, m) be a Riemannian manifold, and let ∇ denote the Levi-Civita connection. Suppose $\varphi\colon M \to N$ is a smooth map. Then for $X, Y, Z \in \mathfrak{X}(M)$ the pullback connection satisfies*

$$\begin{aligned}\langle \nabla_X(D\varphi[Y]), D\varphi[Z]\rangle = \frac{1}{2}\Big(&X\,\langle D\varphi[Y], D\varphi[Z]\rangle + Y\,\langle D\varphi[Z], D\varphi[X]\rangle \\ &- Z\,\langle D\varphi[X], D\varphi[Y]\rangle - \big\langle D\varphi\big[[Y, Z]\big], D\varphi[X]\big\rangle \\ &+ \big\langle D\varphi\big[[Z, X]\big], D\varphi[Y]\big\rangle + \big\langle D\varphi\big[[X, Y]\big], D\varphi[Z]\big\rangle\Big).\end{aligned}$$

Proof. The pullback connection satisfies the Ricci Identity by Corollary 36.16. Thus the claim follows from the uniqueness of the Levi-Civita connection on (N, m) and Proposition 44.8. ■

In Proposition 45.6 the domain M of φ is not endowed with a Riemannian metric (only the target N is). Next lecture we will see that a stronger result holds if M is also Riemannian and φ preserves the metrics. For now we conclude this lecture by briefly discussing Riemannian holonomy groups.

Definition 45.7. Let (M^n, m) be a connected Riemannian manifold. We define the **holonomy group of** m, written as $\mathrm{Hol}(m)$, to be the holonomy group $\mathrm{Hol}^\nabla(x)$, where ∇ is the Levi-Civita connection of m. As in Corollary 32.12, we think of $\mathrm{Hol}(m)$ as a subgroup of $\mathrm{GL}(n)$, which is defined only up to conjugation. Similarly we define the **restricted holonomy group of** m, written $\mathrm{Hol}_0(m)$.

It follows from Problem R.1 that $\mathrm{Hol}(m)$ is actually a subgroup of $\mathrm{O}(n)$ (and thus $\mathrm{Hol}_0(m)$ is a subgroup of $\mathrm{SO}(n)$). On Problem Sheet U you will extend this to the following statement:

Proposition 45.8. *Let M^n be a connected manifold and suppose ∇ is a torsion-free connection on M. Then ∇ is the Levi-Civita connection of a Riemannian metric m on M if and only if Hol^∇ is conjugate in $\mathrm{GL}(n)$ to a subgroup of $\mathrm{O}(n)$.*

The following statement is much more difficult, and its proof goes beyond the scope of this course. It uses the Lie-theoretic fact that every connected Lie subgroup of $\mathrm{SO}(n)$ that acts *irreducibly* on $\mathbb{R}^n$ is in fact closed in $\mathrm{SO}(n)$.

Theorem 45.9. *Let (M^n, m) be a connected Riemannian manifold. Then $\mathrm{Hol}_0(m)$ is a closed connected subgroup of $\mathrm{SO}(n)$.*

Theorem 45.9, together with Theorem 44.16 (and lots and lots and lots of work) gives the following amazing result.

THEOREM 45.10 (The Berger Classification Theorem[2]). *Let M^n be a simply connected manifold and suppose m is an irreducible non-symmetric[3] Riemannian metric on M. Then exactly one of the following options holds for the holonomy group* $\mathrm{Hol}(m)$*:*

(i) $\mathrm{Hol}(m) = \mathrm{SO}(n)$.

(ii) $n = 2k$ *for* $k \geq 2$ *and* $\mathrm{Hol}(m) = \mathrm{U}(k) \subset \mathrm{SO}(2k)$.

(iii) $n = 2k$ *for* $k \geq 2$ *and* $\mathrm{Hol}(m) = \mathrm{SU}(k) \subset \mathrm{SO}(2k)$.

(iv) $n = 4k$ *for* $k \geq 2$ *and*[4] $\mathrm{Hol}(m) = \mathrm{Sp}^{\mathrm{c}}(k) \subset \mathrm{SO}(4k)$.

(v) $n = 4k$ *for* $k \geq 2$ *and* $\mathrm{Hol}(m) = \mathrm{Sp}(2k) \cdot \mathrm{Sp}^{\mathrm{c}}(1) \subset \mathrm{SO}(4k)$.

(vi) $n = 7$ *and*[5] $\mathrm{Hol}(m) = G_2 \subset \mathrm{SO}(7)$.

(vii) $n = 8$ *and*[6] $\mathrm{Hol}(m) = \mathrm{Spin}(7) \subset \mathrm{SO}(8)$.

Moreover all of these groups can occur as the holonomy group of an irreducible non-symmetric Riemannian metric.

As the name suggests, the fact that these are the only options is due to Berger in 1955. The proof that all of these groups really do occur took thirty more years to complete, and is the work of various mathematicians. This culminated in the work of Joyce, who in 1996 constructed compact Riemannian manifolds with holonomy the two so-called *exceptional holonomy groups* G_2 and $\mathrm{Spin}(7)$.

[2]This is the only result in the course where even the *statement* is non-examinable!

[3]As with Theorem 44.16, I won't define precisely what this means, as doing so would take us too far afield.

[4]Here $\mathrm{Sp}^{\mathrm{c}}(k)$ is the **compact symplectic group**, which is defined to be $\mathrm{Sp}^{\mathrm{c}}(k) := \mathrm{Sp}(2k;\mathbb{C}) \cap \mathrm{U}(2k)$. One can think of $\mathrm{Sp}^{\mathrm{c}}(k)$ as the quaternionic unitary group. In particular, $\mathrm{Sp}^{\mathrm{c}}(1)$ can be identified with $\mathrm{SU}(2)$.

[5]This Lie group is a little harder to concisely define—so I refer you to Wikipedia.

[6]The group $\mathrm{Spin}(n)$ is the double cover of $\mathrm{SO}(n)$ (recall $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}_2$). For $n \geq 3$ the group $\mathrm{Spin}(n)$ is simply connected, and thus is also the universal cover of $\mathrm{SO}(n)$.

Isometric maps and natural Riemannian connections

In this lecture we study isometric maps, which are maps between Riemannian manifolds that preserve the metrics. Recall from Definition 36.10 that a vector bundle morphism between two Riemannian vector bundles is said to be an *isometric* vector bundle morphism if it preserves the Riemannian metrics. The following definition specialises this to Riemannian metrics on manifolds.

DEFINITION 46.1. Let (M, m_1) and (N, m_2) be Riemannian manifolds. A smooth map $\varphi\colon M \to N$ is said to be **isometric** if $D\varphi\colon TM \to TN$ is an isometric vector bundle morphism in the sense of Definition 36.10. Explicitly, if we write $\langle \cdot, \cdot \rangle$ for (both) metrics, then φ is isometric if and only if

$$\langle v, w \rangle = \langle D\varphi(x)[v], D\varphi(x)[w] \rangle, \qquad \forall\, x \in M,\ v, w \in T_xM.$$

Equivalently, this means that the metric m_1 is equal to the pullback tensor $\varphi^\star(m_2)$ from Definition 18.7. Note that any isometric map is necessarily an immersion.

If φ is in addition a diffeomorphism, we say that φ is an **isometry**[1].

DEFINITION 46.2. We denote by $\mathrm{Iso}(M, m) \subset \mathrm{Diff}(M)$ the subgroup of isometries.

(♣) REMARK 46.3. If M has finitely many components then $\mathrm{Iso}(M, m)$ is itself a Lie group, which moreover is compact if M is. This is the content of the famous[2] Myers-Steenrod Theorem from 1939, which sadly we will not have time to prove in the course. We will however prove a simple result in this direction later this lecture (Corollary 46.22).

(♣) REMARK 46.4. Riemannian manifolds form a category Riem. The objects of this category are pairs (M, m) where M is a manifold and m is a Riemannian metric on M. The morphisms in this category are the isometric maps. The isomorphisms in this category are the isometries. There is a forgetful functor $\mathsf{Riem} \to \mathsf{Man}$ (where Man is the category of manifolds, cf. Example 14.20) that simply "forgets" the Riemannian metric.

As already remarked, any isometric map between Riemannian manifolds is necessarily an immersion. In fact, there is a partial converse to this, as we now explain.

Suppose (N, m) is a Riemannian manifold and $\varphi\colon M \to N$ is a smooth map. Consider the pullback tensor $\varphi^\star(m) \in \mathcal{T}^{0,2}(M)$. In general this will not define a metric on M—it will always be symmetric, but it need not be positive definite (for

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] Pay attention to the difference between "isometric" and "isometry".

[2] There is another result in Riemannian Geometry often called the Myers-Steenrod Theorem—we state this in Theorem 52.13 (see also Remark 52.14).

example, if φ is constant it is identically zero). If however φ is an immersion then $\varphi^\star(m)$ is positive definite, and hence a Riemannian metric on M. This proves the following useful statement.

LEMMA 46.5. *Let (N, m) be a Riemannian manifold and suppose $\varphi\colon M \to N$ is an immersion. Then $\varphi^\star(m)$ is a Riemannian metric on M, and $\varphi\colon (M, \varphi^\star(m)) \to (N, m)$ is an isometric map. Moreover $\varphi^\star(m)$ is the unique Riemannian metric on M with this property.*

DEFINITION 46.6. Let (N, m) be a Riemannian manifold. An embedded submanifold M of N is said to be a **Riemannian submanifold** if M is endowed with the pullback Riemannian metric $\imath^\star(m)$ (where $\imath\colon M \hookrightarrow N$ denote the inclusion).

EXAMPLE 46.7. The standard Riemannian metric m_{Eucl} on $\mathbb{R}^n$ is given by

$$\langle \mathcal{J}_x(v), \mathcal{J}_x(w) \rangle\rangle_{\mathrm{Eucl}} := \langle v, w \rangle_{\mathrm{Eucl}}\,, \qquad x, v, w \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle_{\mathrm{Eucl}}$ on the right-hand side denotes the Euclidean dot product. Let $\imath\colon S^n \to \mathbb{R}^{n+1}$ denote the inclusion, and let $m_{\mathrm{round}} := \imath^\star(m_{\mathrm{Eucl}})$. Then m_{round} is a Riemannian metric on S^n which we call the **round metric** (since the sphere looks "round" in this metric). This is the unique metric on S^n that makes S^n into a Riemannian submanifold of $\mathbb{R}^{n+1}$. Our favourite connection on S^n (introduced originally in Problem N.3) is in fact the Levi-Civita connection by Problem U.1.

If $\varphi\colon M \to N$ is an immersion then necessarily $\dim M \leq \dim N$. If $\dim M = \dim N$ then there are essentially two cases of interest:

(i) If φ is an *injective* immersion and $\dim M = \dim N$ then it follows from Proposition 5.6 and the Inverse Function Theorem 5.2 that φ is automatically[3] an embedding onto its image. Such a map is often called an **open embedding**, since $\varphi(M)$ is then open in N.

(ii) The other main case of interest is when φ is a **smooth covering map**. This means that φ is *surjective*, and moreover[4] every point $y \in N$ has a neighbourhood U_y such that φ maps each component of $\varphi^{-1}(U_y)$ diffeomorphically onto U_y.

Covering maps are important in Algebraic Topology. We will not really have any cause to use them, other than to note they provide us with further examples of isometric maps.

DEFINITION 46.8. A **Riemannian covering** $\varphi\colon (M, m_1) \to (N, m_2)$ is an isometric map between Riemannian manifolds which is in addition a smooth covering map. Note this necessarily implies $\dim M = \dim N$.

[3]Compare this to Problem C.4, which gives another (entirely unrelated) condition for an injective immersion to automatically be an embedding.

[4]Unlike in case (i), this really is an extra condition—not every surjective submersion between manifolds of the same dimension is a smooth covering map. Exercise: Find an example of this!

Recall a covering map[5] $f\colon X \to Y$ between topological spaces is **normal** if $f_\star(\pi_1(X,x))$ is a normal subgroup of $\pi_1(Y, f(x))$. In particular, a universal cover is always normal. It is a standard result in covering space theory that a covering is normal if and only if the deck transformation group acts transitively on the fibres. If $\varphi\colon M \to N$ is a smooth normal covering map then φ is necessarily a submersion, and the deck transformations are diffeomorphisms of M. The proof of the next result is deferred to Problem Sheet U.

PROPOSITION 46.9. *Let $\varphi\colon M \to N$ be a smooth normal covering map and let m be a Riemannian metric on M which is invariant under all deck transformations. Then there is a unique Riemannian metric on N such that φ is a Riemannian covering.*

EXAMPLE 46.10. We can think of the torus T^n as the quotient $\mathbb{R}^n/\mathbb{Z}^n$, and in fact this is the universal cover. By Proposition 46.9 there is a unique Riemannian metric on T^n such that the quotient map $\mathbb{R}^n \to T^n$ is a Riemannian covering, where $\mathbb{R}^n$ is equipped with its standard Euclidean metric m_{Eucl}. We call this metric the **flat metric** on the torus and write it as m_{flat}.

REMARK 46.11. **Warning:** Take $n = 2$. Then one can embed T^2 into $\mathbb{R}^3$ (think of a (hollow) doughnut). If $\imath\colon T^2 \to \mathbb{R}^3$ denotes the inclusion then $\imath^\star(m_{\mathrm{Eucl}})$ is another Riemannian metric on T^2. As we will see next lecture, m_{flat} is *not* the same Riemannian metric as $\imath^\star(m_{\mathrm{Eucl}})$. In fact, it is not possible to embed (T^2, m_{flat}) into $\mathbb{R}^3$.

EXAMPLE 46.12. The projection map $S^n \to \mathbb{R}P^n$ is a smooth normal covering. Thus there is a unique Riemannian metric m on $\mathbb{R}P^n$ such that $(S^n, m_{\mathrm{round}}) \to (\mathbb{R}P^n, m)$ is a Riemannian covering.

REMARK 46.13. Immersions are dual to submersions, and thus it won't surprise you to learn that there is a dual notion of a **Riemannian submersion** which allows for the case $\dim M \geq \dim N$. We won't have cause to study these in general (and they are a little messier to define), although see Problem U.5 for an important special case.

Suppose now that $\varphi\colon (M, m_1) \to (N, m_2)$ is an isometric map between Riemannian manifolds. Recall from the discussion just before Proposition 44.8 that if X is a vector field on M then $D\varphi[X]$ is a vector field along φ, i.e. an element of $\Gamma_\varphi(TN)$. In a similar vein, if $\vartheta \in \Gamma_\varphi(T^*N)$ is a one-form along φ then there is a well-defined one-form $\varphi^\star(\vartheta)$ given by

$$\varphi^\star(\vartheta)|_x(v) := \vartheta|_x(D\varphi(x)[v]), \qquad x \in M, v \in T_xM.$$

Warning: Despite the fact that we are using the same notation, this is *not* quite the same as the usual pullback operation $\varphi^\star\colon \Omega^1(N) \to \Omega^1(M)$ from Definition 19.6.

We can also define a musical isomorphism between vector fields along φ and one-forms along φ:

$$W \in \Gamma_\varphi(TN) \mapsto W^\flat \in \Gamma_\varphi(T^*N), \qquad W^\flat|_x(v) := \langle W(x), v\rangle, \qquad v \in T_{\varphi(x)}N,$$

[5] Don't worry if you are unfamiliar with covering space theory—we will not actually use any of this, it is just for interest.

and

$$\vartheta \in \Gamma_\varphi(T^*N) \mapsto \vartheta^\sharp \in \Gamma_\varphi(TN), \qquad \langle \vartheta^\sharp(x), v \rangle = \vartheta|_x(v), \qquad v \in T_{\varphi(x)}N.$$

Consider now the following picture[6]:

$$\begin{tikzcd}
\mathfrak{X}(M) \arrow[d, "X \mapsto D\varphi[X]"'] & \Omega^1(M) \arrow[l, "\sharp"'] \\
\Gamma_\varphi(TN) \arrow[r, "\flat"'] & \Gamma_\varphi(T^*N) \arrow[u, "\vartheta \mapsto \varphi^\star(\vartheta)"']
\end{tikzcd} \tag{46.1}$$

where the upper sharp refers to the metric m_1 on M, and the lower flat refers to the metric m_2 on N.

It is important to realise that going all the way round the square[7] is *not* the identity operator if $\dim M < \dim N$. This is because when $\dim M < \dim N$ the vertical maps are not invertible. Thus the following definition makes sense:

Definition 46.14. Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds. If $W \in \Gamma_\varphi(TN)$ is a vector field along φ then we define the **tangential component** of W to be the vector field $W^\top \in \Gamma_\varphi(TN)$ obtained from W by going all the way round the square (46.1):

$$W^\top := D\varphi\big[(\varphi^\star(W^\flat))^\sharp\big].$$

If $\dim M = \dim N$ then, as remarked above, going all the way around the diagram (46.1) is in this case the identity (as the vertical maps are then also isomorphisms), and thus

$$W^\top = W, \qquad \text{if } \dim M = \dim N. \tag{46.2}$$

Since W and $W^\top$ are both sections of the same vector bundle, we can subtract them:

Definition 46.15. We define the **orthogonal component** of W to be the vector field $W^\perp := W - W^\top$.

The next lemma is immediate, since the composition and difference of point operators is a point operator.

Lemma 46.16. *Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds. Then the two operators*

$$(\cdot)^\top \colon \Gamma_\varphi(TN) \to \Gamma_\varphi(TN), \qquad (\cdot)^\perp \colon \Gamma_\varphi(TN) \to \Gamma_\varphi(TN)$$

are both point operators.

[6] I confused many of you in lecture here, so I removed some of the arrows in the diagram to make it clearer. Written like this, it doesn't make sense to say that diagram commutes (or does not commute)—indeed, the statement "the diagram commutes" has no meaning, since there is only one way to pass from any given vertex to itself.

[7] Just to make things even more confusing: if one starts in the *top* left-hand corner and goes all the way round the square then one does get the identity operator—the reason for this is explained in Lemma 46.17 below.

Lemma 46.16 implies that we can speak of $w^\top$ and $w^\perp$ for a single vector $w \in T_{\varphi(x)}N$. On Problem Sheet U you will prove the following result, which gives the geometric intuition behind the name “tangential component”.

LEMMA 46.17. *Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds. Then for $x \in M$ the restriction of $(\cdot)^\top$ to $T_{\varphi(x)}N$ is the orthogonal projection onto $D\varphi(x)[T_xM]$. Thus if $\dim M = \dim N$ then $w^\top = w$ for all $w \in TN|_{\varphi(M)}$.*

PROPOSITION 46.18. *Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds. Let ∇^1 denote the Levi-Civita connection of (M, m_1), and let ∇^2 denote the Levi-Civita connection of (N, m_2), and let $\kappa_1\colon TTM \to TM$ and $\kappa_2\colon TTN \to TN$ denote the associated connection maps. Then:*

(i) *If $\zeta \in T_{(x,v)}TM$ then*

$$D\varphi(x)[\kappa_1(\zeta)] = \big(\kappa_2\big(D(D\varphi(x))(v)[\zeta]\big)\big)^\top,$$

where $D(D\varphi(x))[v]$ denotes the differential of the map $D\varphi(x)\colon T_xM \to T_{\varphi(x)}N$ at $v \in T_xM$.

(ii) *If in addition $\dim M = \dim N$ then the same holds without the “$\top$”:*

$$D\varphi(x)[\kappa_1(\zeta)] = \kappa_2\big(D(D\varphi(x))(v)[\zeta]\big),$$

and hence the following commutes:

$$\begin{tikzcd} TTM \arrow[r, "D(D\varphi)"] \arrow[d, "\kappa_1"'] & TTN \arrow[d, "\kappa_2"] \\ TM \arrow[r, "D\varphi"'] & TN \end{tikzcd}$$

Proof. Since φ is isometric, it follows from (45.2) and Proposition 45.6 that for $X, Y, Z \in \mathfrak{X}(M)$ that

$$\big\langle \nabla^2_X(D\varphi[Y]), D\varphi[Z]\big\rangle = \big\langle \nabla^1_X(Y), Z\big\rangle.$$

Moreover as φ is isometric we have

$$\big\langle \nabla^1_X(Y), Z\big\rangle = \big\langle D\varphi\big[\nabla^1_X(Y)\big], D\varphi[Z]\big\rangle,$$

which implies that

$$\big(\nabla^2_X(D\varphi[Y])\big)^\top = D\varphi\big[\nabla^1_X(Y)\big]$$

(both sides are elements of $\Gamma_\varphi(TN)$). This proves (i). The second statement is immediate from this and (46.2). ■

COROLLARY 46.19. *Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds of the same dimension. Let ∇^1 denote the Levi-Civita connection of (M, m_1), and let ∇^2 denote the Levi-Civita connection of (N, m_2). Let R^{∇^1} and R^{∇^2} denote their curvature tensors. Then for all $x \in M$ and all $u, v, w \in T_xM$, one has*

$$D\varphi(x)\big[R^{\nabla^1}(u, v)(w)\big] = R^{\nabla^2}(D\varphi(x)[u], D\varphi(x)[v])(D\varphi(x)[w]).$$

Proof. This follows from part (ii) of Proposition 46.18 together with Proposition 34.11. ∎

Definition 46.20. Let (M, m) be a Riemannian manifold. The **exponential map of** m is by definition the exponential map of the geodesic spray of the Levi-Civita connection of m.

The next result shows isometric maps between Riemannian manifolds of the same dimension behave similarly to Lie group homomorphisms for the exponential map of a Riemannian metric (compare this with Proposition 10.12).

Proposition 46.21. *Let* $\varphi\colon (M, m_1) \to (N, m_2)$ *be an isometric map between Riemannian manifolds of the same dimension. Let* ∇^1 *denote the Levi-Civita connection of* (M, m_1), *and let* ∇^2 *denote the Levi-Civita connection of* (N, m_2). *Let* $\exp^1$ *and* $\exp^2$ *denote the associated exponential maps. Then*

$$\exp^2 \circ D\varphi = \varphi \circ \exp^1$$

Proof. It follows from part (ii) of Proposition 46.18 that if c is a parallel vector field along a curve γ in M then $D\varphi[c]$ is a parallel vector field along $\varphi \circ \gamma$ in N. Taking $c = \gamma'$ shows that φ maps geodesics in M to geodesics in N. The claim now follows from the uniqueness part of Proposition 42.7. ∎

The next corollary shows how restrictive the condition of being an isometric map is when the manifolds have the same dimension.

Corollary 46.22. *Let* $\varphi, \psi\colon (M, m_1) \to (N, m_2)$ *be two isometric maps between Riemannian manifolds of the same dimension. Assume* M *is connected and that there exists* $x \in M$ *such that* $\varphi(x) = \psi(x)$ *and* $D\varphi(x) = D\psi(x)$. *Then* $\varphi = \psi$.

Proof. Let

$$A := \{y \in M \mid \varphi(y) = \psi(y) \text{ and } D\varphi(y) = D\psi(y)\}\,.$$

Then A is non-empty as $x \in A$. Moreover A is closed as TM is Hausdorff and $D\varphi$ and $D\psi$ are continuous (actually, smooth). If $y \in A$ then by part (ii) of Theorem 43.3 there exists a neighbourhood V_y of $0_y \in T_yM$ such that $\exp^1_y$ maps V_y diffeomorphically onto its image. If $v \in V_y$ then by Proposition 46.21 we have

$$\begin{aligned}\varphi(\exp^1_y(v)) &= \exp^2_{\varphi(y)}(D\varphi(y)[v])\\ &= \exp^2_{\psi(y)}(D\psi(y)[v])\\ &= \psi(\exp^1_y(v)),\end{aligned}$$

and hence on V_y one has (as smooth maps)

$$\varphi \circ \exp^1_y = \psi \circ \exp^1_y,$$

which in particular implies that $\exp^1_y(V_y) \subset A$. Since $\exp^1_y(V_y)$ is open and y was arbitrary, it follows that A is also open, and hence $A = M$ as M is connected. ∎

The rest of this lecture gives the necessary definitions required in order to state[8] a theorem of Epstein, which, roughly speaking, proves the converse to Proposition 46.18.

Definition 46.23. A **natural Riemannian connection** is an assignment of a connection $\nabla^{M,m}$ to every Riemannian manifold (M, m) which is natural in the following sense: If $\varphi\colon (M, m_1) \to (N, m_2)$ is an injective isometric map between Riemannian manifolds of the same dimension (i.e. an isometric open embedding) then

$$\varphi^\star(\nabla^{N,m_2}) = \nabla^{M,m_1}.$$

We denote such a natural connection by $\boldsymbol{\nabla}$. A natural Riemannian connection is said to be **homogeneous** if it is invariant under scaling:

$$\nabla^{M,m} = \nabla^{M,cm}$$

for any $c > 0$.

This is easiest to explain with an example.

Example 46.24. The assignment $\boldsymbol{\nabla}^{\mathrm{LC}}$ that assigns to each Riemannian manifold (M, m) its Levi-Civita connection is a natural Riemannian connection by Proposition 46.18. It is clear from (45.2) that the Levi-Civita connection is homogeneous.

(♣) Remark 46.25. The definition of a natural Riemannian connection can be phrased more concisely using categorical language. Here are the details. Consider the category $\mathsf{OpenEmb}$ whose objects are smooth manifolds and whose morphisms are *open embeddings*, i.e. embeddings that are diffeomorphisms onto their images (this is a subcategory of the category Man from Example 14.20—note there are no morphisms from M to N in this category if $\dim M \neq \dim N$). Consider the contravariant functor R on $\mathsf{OpenEmb}$ that assigns to a manifold M the space[9] $\mathsf{R}(M)$ of all Riemannian metrics on M, and assigns to an open embedding $\varphi\colon M \to N$ the induced map

$$\varphi^\star\colon \mathsf{R}(N) \to \mathsf{R}(M), \qquad m \mapsto \varphi^\star(m).$$

In a similar vein there is a contravariant functor C on $\mathsf{OpenEmb}$ that assigns to M the space $\mathsf{C}(M)$ of all connections on M, and on morphisms operates by pullback. Then a natural Riemannian connection $\boldsymbol{\nabla}$ is exactly a natural transformation from R to C.

Definition 46.26. Suppose $\boldsymbol{\nabla}$ is a natural Riemannian connection. We say that ∇ is of **polynomial type** if for each $n \geq 0$ there exist polynomials P^k_{ij} for $1 \leq i, j, k \leq n$ such that: For any n-dimensional Riemannian manifold (M, m), and for any chart $\sigma\colon U \to O$ on M with local coordinates x^i, the Christoffel symbols Γ^k_{ij} are given as polynomials in the components m_{ij} of m relative to x^i, together with their inverse m^{ij}, and all derivatives of m_{ij} up to some finite order d, i.e.

$$\Gamma^k_{ij} = P^k_{ij}\left(m_{pq}; m^{rs}; \frac{\partial^{|\alpha|}}{\partial x^\alpha} m_{hl}\right),$$

[8] Yes, only "state"! The proof goes well beyond the scope of this course; merely stating it accurately is a task in itself.

[9] The space $\mathsf{R}(M)$ is actually an infinite-dimensional (locally) Fréchet manifold, but we won't need or use this fact.

where α is any multi-index of degree at most d.

Again, this is easiest to explain with an example.

EXAMPLE 46.27. The natural Riemannian connection $\boldsymbol{\nabla}^{\mathrm{LC}}$ is a polynomial connection by Lemma 45.2 (with $d = 1$).

Here now is our promised theorem. One should think of it as a far-reaching complement of Theorem 45.1 (which in fact deserves the name "The Fundamental Theorem of Riemannian Geometry much better!)

THEOREM 46.28 (Epstein, 1978). *Let $\boldsymbol{\nabla}$ be a homogeneous natural Riemannian connection. Assume $\boldsymbol{\nabla}$ is of polynomial type. Then $\boldsymbol{\nabla} = \boldsymbol{\nabla}^{\mathrm{LC}}$ is the Levi-Civita connection.*

Sectional curvature and Schur's Theorem

In this lecture we investigate various other curvatures that can be associated to a Riemannian manifold. In doing so we will finally make contact with the geometric intuition of the word "curvature": as we will see, the sphere S^n thought of as a Riemannian submanifold of $\mathbb{R}^{n+1}$ is positively curved, whereas the hyperbolic plane with its natural metric (see Definition 49.6) is negatively curved.

Firstly, we show how a Riemannian metric allows us to view the curvature as a tensor of type $(0,4)$ instead of type $(1,3)$.

DEFINITION 47.1. Let $(M, m = \langle \cdot, \cdot \rangle)$ denote a Riemannian manifold, and suppose ∇ is a connection on M (not necessarily torsion-free or Riemannian with respect to m). Then $R^\nabla \in \mathcal{T}^{1,3}(M)$. We use m to define a new tensor $\mathcal{R}_m^\nabla \in \mathcal{T}^{0,4}(M)$ by

$$\mathcal{R}_m^\nabla(W, Z, X, Y) := \left\langle R^\nabla(X, Y)(Z), W \right\rangle, \qquad \forall\, X, Y, Z, W \in \mathfrak{X}(M).$$

Warning: Pay attention[1] to the ordering of W, Z, X and Y on the left-hand side!

Suppose $\sigma \colon U \to O$ is a chart on M with local coordinates (x^i). Then we can write

$$R^\nabla = R_{ijk}^l\, \partial_l \otimes dx^i \otimes dx^j \otimes dx^k$$

and

$$\mathcal{R}_m^\nabla = R_{ijkl}\, dx^i \otimes dx^j \otimes dx^k \otimes dx^l$$

where R_{ijk}^l and R_{ijkl} are smooth functions on U given by

$$R_{ijk}^l = dx^l\big(R^\nabla(\partial_i, \partial_j)(\partial_k)\big)$$

and

$$R_{ijkl} := \mathcal{R}_m^\nabla(\partial_i, \partial_j, \partial_k, \partial_l).$$

If we write $m = m_{ij}\, dx^i \otimes dx^j$ then

$$R_{ijkl} := \langle R\left(\partial_k, \partial_l\right)\left(\partial_j\right), \partial_i \rangle = \left\langle R_{klj}^h \partial_h, \partial_i \right\rangle = m_{hi} R_{klj}^h. \tag{47.1}$$

The next lemma clarifies the symmetries of $\mathcal{R}_m^\nabla$.

LEMMA 47.2 (Symmetries of $\mathcal{R}_m^\nabla$). *Let (M, m) be a Riemannian manifold and let ∇ be a connection on M. Then for any $X, Y, Z, W \in \mathfrak{X}(M)$:*

(i) $\mathcal{R}_m^\nabla(W, Z, Y, X) = -\mathcal{R}_m^\nabla(W, Z, X, Y)$.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1]Also: Be aware that practically all permutations of the indices are *someone's* favourite sign conventions! Thus when consulting textbooks, make sure you are aware which convention the author favours.

(ii) *If ∇ is Riemannian with respect to m then*

$$\mathcal{R}_m^\nabla(Z, W, X, Y) = -\mathcal{R}_m^\nabla(W, Z, X, Y).$$

(iii) *If ∇ is torsion-free then*

$$\mathcal{R}_m^\nabla(W, Z, X, Y) + \mathcal{R}_m^\nabla(W, X, Y, Z) + \mathcal{R}_m^\nabla(W, Y, Z, X) = 0.$$

(iv) *If ∇ is the Levi-Civita connection of m then*

$$\mathcal{R}_m^\nabla(W, Z, X, Y) = \mathcal{R}_m^\nabla(X, Y, W, Z).$$

Proof. Property (i) is clear as R^∇ is alternating. Property (ii) is a restatement of Proposition 36.18. Property (iii) is a restatement of part (i) of Proposition 44.12.

Finally, property (iv) is an algebraic consequence of the other properties. Indeed,

$$\begin{aligned}\mathcal{R}_m^\nabla(W, Z, X, Y) &= -\mathcal{R}_m^\nabla(W, Z, Y, X)\\ &= \mathcal{R}_m^\nabla(W, Y, X, Z) + \mathcal{R}_m^\nabla(W, X, Z, Y)\end{aligned}$$

and also

$$\begin{aligned}\mathcal{R}_m^\nabla(W, Z, X, Y) &= -\mathcal{R}_m^\nabla(Z, W, X, Y)\\ &= \mathcal{R}_m^\nabla(Z, X, Y, W) + \mathcal{R}_m^\nabla(Z, Y, W, X)\end{aligned}$$

and so

$$2\mathcal{R}_m^\nabla(W, Z, X, Y) = \mathcal{R}_m^\nabla(W, Y, X, Z)+\mathcal{R}_m^\nabla(W, X, Z, Y)+\mathcal{R}_m^\nabla(Z, X, Y, W)+\mathcal{R}_m^\nabla(Z, Y, W, X).$$

Similarly

$$2\mathcal{R}_m^\nabla(X, Y, W, Z) = \mathcal{R}_m^\nabla(X, Z, W, Y)+\mathcal{R}_m^\nabla(X, W, Y, Z)+\mathcal{R}_m^\nabla(Y, W, Z, X)+\mathcal{R}_m^\nabla(Y, Z, X, W).$$

Then using

$$\begin{aligned}\mathcal{R}_m^\nabla(X, Z, W, Y) &= (-1)^2\mathcal{R}_m^\nabla(Z, X, Y, W),\\ \mathcal{R}_m^\nabla(X, W, Y, Z) &= (-1)^2\mathcal{R}_m^\nabla(W, X, Z, Y),\\ \mathcal{R}_m^\nabla(Y, W, Z, X) &= (-1)^2\mathcal{R}_m^\nabla(W, Y, X, Z),\\ \mathcal{R}_m^\nabla(Y, Z, X, W) &= (-1)^2\mathcal{R}_m^\nabla(Z, Y, W, X),\end{aligned}$$

we see that

$$2\mathcal{R}_m^\nabla(X, Y, W, Z) = 2\mathcal{R}_m^\nabla(W, Z, X, Y),$$

and this completes the proof. ∎

Let ∇ be a torsion-free connection on M. Fix $x \in M$ and let (x^i) be local coordinates about x such that the Christoffel symbols vanish at x (possible by Proposition 44.11). The following computation is only valid at the point x, but to keep the notation simple in the following computation we omit the x from both sides:

$$\begin{aligned}
(\nabla_{\partial_i} R^\nabla)(\partial_j, \partial_k)(\partial_l) &\overset{(\dagger)}{=} \nabla_{\partial_i}\big(R^\nabla(\partial_j, \partial_k)\big)(\partial_l) - R^\nabla\big(\nabla_{\partial_i}\partial_j, \partial_k\big)(\partial_l) - R^\nabla\big(\partial_j, \nabla_{\partial_i}\partial_k\big)\partial_l \\
&= \nabla_{\partial_i}\big(R^\nabla(\partial_j, \partial_k)\big)(\partial_l) + 0 \\
&\overset{(\ddagger)}{=} \nabla_{\partial_i}\big(R^\nabla(\partial_j, \partial_k)\partial_l\big) - R^\nabla(\partial_j, \partial_k)\big(\nabla_{\partial_i}\partial_l\big) \\
&= \nabla_{\partial_i}\big(R^h_{jkl}\partial_h\big) + 0 \\
&= \partial_i\big(R^h_{jkl}\big)\partial_h + R^h_{jkl}\,\nabla_{\partial_i}(\partial_h) \\
&= \partial_i\big(R^h_{jkl}\big)\partial_h + 0,
\end{aligned}$$

where $(\dagger)$ used the definition of the induced connection on the tensor bundle $T^{1,3}(TM) \to M$ and $(\ddagger)$ used the definition of the induced connection on $T^{1,1}(TM)$ (cf. Problem P.2). Thus part (ii) of Proposition 44.12 tells us that in these coordinates we have

$$\partial_i\big(R^h_{jkl}\big)(x) + \partial_j\big(R^h_{kil}\big)(x) + \partial_k\big(R^h_{ijl}\big)(x) = 0. \tag{47.2}$$

Now suppose m is a Riemannian metric on m and ∇ is the Levi-Civita connection of m. Assume the (x^i) are normal coordinates at x. Then by Remark 45.4 we have $\partial_i m_{jk}(x) = 0$ and hence

$$\begin{aligned}
\partial_i R_{hljk}(x) &= \partial_i\big(m_{ph}R^p_{jkl}\big)(x) \\
&= \partial_i m_{ph}(x) R^p_{jkl}(x) + m_{ph}(x)\partial_i\big(R^p_{jkl}\big)(x) \\
&= \partial_i\big(R^h_{jkl}\big)(x).
\end{aligned}$$

Thus by (47.2) we see that

$$\partial_i R_{hljk}(x) + \partial_j R_{hlki}(x) + \partial_k R_{hlij}(x) = 0. \tag{47.3}$$

This formula will be used in the proof of Schur's Theorem 47.14 below.

From now on we will work exclusively with the Levi-Civita connection. This isn't strictly necessary, but it simplifies the discussion (and the notation), and is by far the most important case. Our next notion of curvature corresponds to the geometric intuition behind the word.

Definition 47.3. Let (M, m) be a Riemannian manifold. Let ∇ denote the Levi-Civita connection of m, and fix $x \in M$. Given two linearly independent tangent vectors $v_1, v_2 \in T_xM$ we define the **sectional curvature** of the 2-plane $\Pi = \operatorname{span}\{v_1, v_2\} \subseteq T_xM$ to be

$$\operatorname{sect}_m(x; \Pi) := \frac{\mathcal{R}^\nabla_m(v_1, v_2, v_1, v_2)}{\langle v_1, v_1\rangle\,\langle v_2, v_2\rangle - \langle v_1, v_2\rangle^2}. \tag{47.4}$$

Note that this depends only on the 2-plane Π and not the choice of basis $\{v_1, v_2\}$, since both $\mathcal{R}^\nabla_m$ and m are linear and thus both the numerator and denominator of

(47.4) are homogeneous of degree two. In particular, if $\{e_1, e_2\}$ are orthonormal, and $\Pi := \operatorname{span}\{e_1, e_2\}$ then

$$\operatorname{sect}_m(x; \Pi) = \mathcal{R}^{\nabla}_m(e_1, e_2, e_1, e_2).$$

REMARK 47.4. If $\dim M = 2$ then there is only one two-plane in each tangent space (namely, the entire tangent space), and thus in this case the sectional curvature is simply a function $\operatorname{sect}_m \colon M \to \mathbb{R}$. For historical reasons in this case the sectional curvature is often called the **Gaussian curvature**.

DEFINITION 47.5. Let (M, m) be a Riemannian manifold and let $\kappa \in \mathbb{R}$. We say that (M, m) has **constant curvature** κ if

$$\operatorname{sect}_m(x; \Pi) = \kappa, \qquad \forall\, x \in M,\ \forall\, 2\text{-planes } \Pi \subset T_x M.$$

EXAMPLE 47.6. If we consider $\mathbb{R}^n$ with its standard Euclidean metric (Example 46.7) then $\mathbb{R}^n$ has constant curvature with $\kappa = 0$.

EXAMPLE 47.7. If we consider the sphere S^n as a Riemannian submanifold of $\mathbb{R}^{n+1}$ then it follows from Problem P.4 that S^n has constant curvature with $\kappa = 1$. More generally, if $S^n(r)$ denotes the sphere of radius $r > 0$ then the same argument shows that $S^n(r)$ (as a Riemannian submanifold of $\mathbb{R}^{n+1}$) has constant curvature with $\kappa = \frac{1}{r^2}$.

We will discuss the case of $\kappa < 0$ in Lecture 48.

REMARK 47.8. The argument from Problem R.6 easily adapts to show that if M is any manifold that admits a metric of constant curvature then $p_r(TM) = 0$ for all $r > 0$.

In fact, the sectional curvature determines the full Riemannian curvature tensor. In order to prove this, we need the following algebraic lemma.

LEMMA 47.9. *Let V be a vector space and $R_1, R_2 \colon V \times V \times V \times V \to \mathbb{R}$ two quadrilinear maps such that for all $w, x, y, z \in V$ and $i = 1, 2$:*

(i) $R_i(w, z, y, x) = -R_i(w, z, x, y)$,

(ii) $R_i(z, w, x, y) = -R_i(w, z, x, y)$,

(iii) $R_i(w, z, x, y) + R_i(w, x, y, z) + R_i(w, y, z, x) = 0$.

(iv) $R_i(w, z, x, y) = R_i(x, y, w, z)$.

Then if for all $x, y \in V$ we also have $R_1(x, y, x, y) = R_2(x, y, x, y)$, then in fact $R_1 \equiv R_2$.

Proof. It suffices to show that if a quadrilinear map R satisfies the four conditions of the lemma and in addition satisfies $R(x, y, x, y) = 0$ for all $x, y \in V$ then $R \equiv 0$. So suppose this is the case. Then

$$\begin{aligned}
0 &= R(x, y+z, x, y+z) \\
&= R(x, y, x, y) + R(x, z, x, y) + R(x, z, x, z) + R(x, y, x, z) \\
&= R(x, z, x, y) + R(x, y, x, z) + 0 \\
&= 2R(x, y, x, z),
\end{aligned}$$

and hence R is alternating with respect to the first and third variables. Similarly R is alternating with respect to the second and fourth variables. Then

$$\begin{aligned} 0 &= R(w,z,x,y) + R(w,x,y,z) + R(w,y,z,x) \\ &= R(w,z,x,y) - R(w,z,y,x) - R(w,y,x,z) \\ &= 3R(w,z,x,y). \end{aligned}$$

This completes the proof. ■

COROLLARY 47.10. *The sectional curvatures determine the full curvature tensor.*

The next corollary tells us that if the sectional curvatures at a given point are independent of the choice of two-plane then the full curvature tensor takes a particularly nice form. First, a definition:

DEFINITION 47.11. Let (M,m) denote a Riemannian manifold. Define a tensor $\mathcal{S}_m \in \mathcal{T}^{0,4}(M)$ by

$$\mathcal{S}_m(W,Z,X,Y) \coloneqq \langle W,X\rangle\langle Z,Y\rangle - \langle W,Y\rangle\langle Z,X\rangle\,.$$

COROLLARY 47.12. *Suppose that (M,m) is a Riemannian manifold and ∇ is the Levi-Civita connection on M. Suppose there exists $x \in M$ and $\kappa \in \mathbb{R}$ such that*

$$\mathrm{sect}_m(x;\Pi) = \kappa, \qquad \forall\, \text{2-planes } \Pi \subset T_xM.$$

Then for all $W,X,Y,Z \in \mathfrak{X}(M)$ one has

$$\mathcal{R}^{\nabla}_m(W,Z,X,Y)(x) = \kappa\,\mathcal{S}_m(W,Z,X,Y)(x). \tag{47.5}$$

Proof. Apply Lemma 47.9 to $\mathcal{R}^{\nabla}_m|_x$ and $\kappa\,\mathcal{S}_m|_x$. ■

If M is 2-dimensional then the hypotheses of Corollary 47.12 are automatically satisfied (cf. Remark 47.4), and hence we obtain:

COROLLARY 47.13. *Let (M^2,m) be a two-dimensional Riemannian manifold, and let ∇ denote the Levi-Civita connection of m. Then*

$$\mathcal{R}^{\nabla}_m = \mathrm{sect}_m\,\mathcal{S}_m,$$

where $\mathrm{sect}_m \in C^\infty(M)$ denotes the sectional (or Gaussian) curvature.

In higher dimensions the situation dramatically changes: if M is connected then the hypotheses of Corollary 47.12 force m to have to constant curvature.

THEOREM 47.14 (Schur's Theorem). *Let (M,m) be a connected Riemannian manifold of dimension $n \geq 3$. Suppose there exists a function $f \in C^\infty(M)$ such that $\mathrm{sect}_m(x;\Pi) = f(x)$ for all 2-planes $\Pi \subseteq T_xM$. Then f is a constant function, and hence (M,m) is a space of constant curvature.*

Proof. Let $x \in M$, and (x^i) be normal coordinates on a neighbourhood U of x. Applying Corollary 47.12 to the coordinate vector fields ∂_i we see that on U we have

$$R_{ijkl}(y) = f(y)\big(m_{ik}m_{jl} - m_{il}m_{jk}\big).$$

Now by (47.3) and the fact that $m_{ij} = \delta_{ij}$, we obtain

$$\frac{\partial f}{\partial x^h}(x)\left(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right) + \frac{\partial f}{\partial x^k}(x)\left(\delta_{il}\delta_{jh} - \delta_{ih}\delta_{jl}\right) + \frac{\partial f}{\partial x^l}(x)\left(\delta_{ih}\delta_{jk} - \delta_{ik}\delta_{jh}\right) = 0.$$

Since $n \geq 3$, given h we can find i, j such that i, j, h are all distinct. Setting $k = i, l = j$ it then follows from the above that $\frac{\partial f}{\partial x^h}(x) = 0$. Since h was arbitrary, it follows $df|_x = 0$. Since x was arbitrary, it follows that f is locally constant. Since M is connected, f is constant. This completes the proof. ∎

LECTURE 48

Three equivalent definitions of the Laplacian

This lecture is an analytic "interlude". We introduce the divergence of a vector field and the gradient, Hessian, and Laplacian of a function on an oriented Riemannian manifold. These generalise the corresponding notions you learnt in multivariate calculus. Next lecture we will use this to investigate how the sectional curvature of an oriented two-dimensional Riemannian manifold behaves when we multiply the metric by a smooth function[1], which will naturally lead us to the hyperbolic metric.

Recall a manifold M^n is oriented if and only if there exists a volume form, i.e. a differential form $\mu \in \Omega^n(M)$ which is nowhere vanishing (cf. Corollary 20.23). The volume form is not unique—if $f\colon M \to (0,\infty)$ is any positive smooth function then $f\mu$ is another volume form which moreover defines the same orientation. However a choice of Riemannian metric on M gives us a way to normalise the volume form.

Definition 48.1. Let (M^n, m) be an oriented Riemannian manifold. The **Riemannian volume form** of m is the unique volume form $\mu_m \in \Omega^n(M)$ with the property that if $\{v_i\}$ is any positively oriented orthonormal basis of T_xM then

$$\mu_m|_x(v_1, \dots, v_n) = 1.$$

The notation μ_m is a little misleading, since μ_m depends both on the metric m and on the choice of orientation on M.

Remark 48.2. If M is not orientable, then it is still possible to define a **Riemannian density** (instead of a Riemannian volume form), which enjoys most of the same properties. This allows for nearly all of what we cover in this lecture to go through for non-orientable manifolds too. However, for simplicity we will restrict ourselves to the orientable case.

Now that the volume form is uniquely determined, we can integrate functions[2] on M.

Definition 48.3. Let (M, m) be an oriented Riemannian manifold with Riemannian volume form μ_m. If $f \in C_c^\infty(M)$ is a function with compact support we define the **integral of f** as

$$\int_{M,m} f := \int_M f\mu_m.$$

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1]This is of course not the most direct method for defining the hyperbolic plane! Nevertheless we will use the gradient and the Hessian multiple times during the rest of the course, and no course on differential geometry would be complete without at least defining the Laplacian.

[2]Compare this discussion to Problem L.7, which carries out the same idea for compact Lie groups.

We can use $\int_{M,m}$ to obtain an inner product on the space of smooth functions. For simplicity we will define this only in the case where M is compact[3].

Definition 48.4. Let (M, m) be a compact oriented Riemannian manifold. We define an inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on the vector space $C^\infty(M)$ by setting

$$\langle\!\langle f, h \rangle\!\rangle := \int_{M,m} fg$$

Warning: This does *not* turn $C_c^\infty(M)$ into a Hilbert space! This is because $C_c^\infty(M)$ is not complete under the norm $\|f\| := \sqrt{\langle\!\langle f, f \rangle\!\rangle}$. This annoyance can be overcome with a little bit of functional analysis (see Remark 48.6 below).

Definition 48.5. Let (M, m) be an orientable Riemannian manifold with Riemannian volume form μ_m. A diffeomorphism $\varphi \colon M \to M$ is **volume preserving** if $\varphi^\star(\mu_m) = \mu_m$. We let $\mathrm{Diff}_{\mathrm{vol}}(M, \mu_m) \subset \mathrm{Diff}(M)$ denote the volume preserving diffeomorphisms.

Any isometry is automatically volume-preserving, but the converse is not necessarily true, as you will prove on Problem Sheet V.

(♣) Remark 48.6. This remark uses a bit of measure theory and functional analysis. It can be ignored if you are not familiar with this material. For simplicity let us assume M is compact. The operation

$$\int_{M,m} : C^\infty(M) \to \mathbb{R}, \qquad f \mapsto \int_{M,m} f$$

extends uniquely[4] to an linear operator

$$\int_{M,m} : C^0(M) \to \mathbb{R}, \qquad h \mapsto \int_{M,m} h,$$

which is *positive* in the sense that if $h \geq 0$ then $\int_{M,m} h \geq 0$. This means that $\int_{M,m}$ is a positive Radon measure on M. As such, by the Riesz-Markov-Kakutani Representation Theorem we get a Borel measure vol_m on M such that

$$\int_{M,m} h = \int_M h \, d\, \mathrm{vol}_m,$$

where the integration on the right-hand side should be interpreted in the usual measure-theoretic sense (Lebesgue integration). The advantage of this viewpoint is that vol_m can eat any measurable subset of M, and $\int_M h \, d\, \mathrm{vol}_m$ is defined for any measurable function h on M. This allows us to define the **volume** of a measurable set $A \subset M$ to be $\mathrm{vol}_m(A)$. Then

$$\mathrm{vol}_m(A) = \int_M \mathbf{1}_A \, d\, \mathrm{vol}_m,$$

[3]This is mainly due to how we defined integration last semester. With a bit of work it can be dropped. Alternatively, one can equally well work with compactly supported functions here too.

[4]This is because $C^\infty(M)$ is dense in $C^0(M)$

where $\mathbf{1}_A\colon M \to \{0,1\}$ is the characteristic function of A, i.e.

$$\mathbf{1}_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

We define $L^2(M,m)$ to be those measurable functions with $\langle\langle f, f\rangle\rangle < \infty$. Then $L^2(M,m)$ is a Hilbert space under $\langle\langle \cdot,\cdot\rangle\rangle$, and $C^\infty(M)$ is a dense subspace. A volume-preserving diffeomorphism preserves the measure vol_m, i.e.

$$\mathrm{vol}_m(A) = \mathrm{vol}_m(\varphi^{-1}(A)), \qquad \forall\, A \subset M \text{ measurable.}$$

This motivates the name "volume-preserving". Much of the material you learnt in your elementary measure theory courses can now be extended to Riemannian manifolds—for example, there is a Fubini Theorem for Riemannian submersions. We will not need any of this, however.

If (x^i) are local coordinates on $U \subset M$ then any volume form μ can be written as $\mu = f dx^1 \wedge \cdots \wedge dx^n$ for some function $f \in C^\infty(U)$. The next lemma identifies this function when $\mu = \mu_m$ is the Riemannian volume form.

Lemma 48.7. *Let (M^n, m) be an oriented Riemannian manifold, and let (x^i) be positively oriented local coordinates (cf. Definition 20.22) on an open set $U \subset M$. Then on U,*

$$\mu_m = \sqrt{\det A}\, dx^1 \wedge \cdots \wedge dx^n,$$

where $A = (m_{ij})_{1\le i,j\le n}$.

Proof. Suppose $(v_1,\dots,v_n)$ is a positively oriented orthonormal basis of T_xM. By (15.8) we have

$$\big(dx^1 \wedge \cdots \wedge dx^n\big)|_x(v_1,\dots,v_n) = \det B,$$

where B is the matrix $\big(dx^i|_x(v_j)\big)_{1\le i,j\le n}$. Thus

$$\mu_m = \frac{1}{\det B}\, dx^1 \wedge \cdots \wedge dx^n. \tag{48.1}$$

Note that if we write

$$v_i = b_i^j \partial_j$$

then B is the matrix $(b_i^j)_{1\le i,j\le n}$. Now observe

$$\delta_{ij} = \langle v_i, v_j\rangle = \big\langle b_i^k \partial_k, b_j^l \partial_l\big\rangle = b_i^k b_j^l m_{kl},$$

which tells us that (as matrices)

$$I = BAB^{\mathrm{T}}.$$

Taking determinants gives

$$\sqrt{\det A} = \frac{1}{\det B},$$

which combined with (48.1) completes the proof. ∎

As a sanity check, let us observe that Lemma 48.7 does indeed specify a well-defined global n-form. Suppose (x^i) are positively oriented local coordinates over $U \subseteq M$ and (y^i) are positively oriented local coordinates over $V \subset M$ with $U \cap V \neq \emptyset$. Let

$$m_{ij} := \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle, \qquad m'_{ij} := \left\langle \frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j} \right\rangle,$$

and let $A := (m_{ij})_{1 \leq i,j \leq n}$ and $A' := (m'_{ij})_{1 \leq i,j \leq n}$. Define $\mu_U \in \Omega^n(U)$ and $\mu_V \in \Omega^n(V)$ by

$$\mu_U := \sqrt{\det A} dx^1 \wedge \cdots \wedge dx^n, \qquad \mu_V := \sqrt{\det A'} dy^1 \wedge \cdots \wedge dy^n.$$

We must show that

$$\mu_U|_{U \cap V} = \mu_V|_{U \cap V}.$$

Write

$$F^i_j = \frac{\partial x^i}{\partial y^j}, \quad H^j_i = \frac{\partial y^j}{\partial x^i};$$

note that $F = H^{-1}$. Then

$$dx^i = F^i_j dy^j, \qquad \frac{\partial}{\partial x^i} = H^j_i \frac{\partial}{\partial y^j}.$$

and

$$dx^1 \wedge \cdots \wedge dx^n = \det F \, dy^1 \wedge \cdots \wedge dy^n.$$

Moreover since

$$m_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = H^k_i H^l_j \left\langle \frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l} \right\rangle = H^k_i m'_{kl} H^j_l$$

we have

$$A = HA'H^{\mathrm{T}},$$

and thus

$$\begin{aligned} \sqrt{\det A} dx^1 \wedge \cdots \wedge dx^n &= \sqrt{\det HA'H^{\mathrm{T}}} \det F \, dy^1 \wedge \cdots \wedge dy^n \\ &= \sqrt{\det A'} \, dy^1 \wedge \cdots \wedge dy^n, \end{aligned}$$

since $\det HH^{\mathrm{T}} = \frac{1}{(\det F)^2}$.

Remark 48.8. The same argument as in the proof of Lemma 48.7 shows that if (x^i) are *any* local coordinates on $U \subset M$ (i.e. not necessarily positively oriented) then on U

$$\mu_m = \sqrt{|\det A|} \, dx^1 \wedge \cdots \wedge dx^n,$$

where as before $A = (m_{ij})_{1 \leq i,j \leq n}$.

Next, we move onto the divergence of a vector field.

Definition 48.9. Let (M, m) be an oriented Riemannian manifold with Riemannian volume form μ_m. Let $X \in \mathfrak{X}(M)$. We define the **divergence** of X to be the smooth function $\operatorname{div}_m(X) \colon M \to \mathbb{R}$ defined by requiring that

$$\operatorname{div}_m(X) \, \mu_m = \mathcal{L}_X(\mu_m).$$

Let θ_t denote the flow of X. Then by Problem L.5 we have

$$\frac{d}{dt}\Big|_{t=t_0} \theta_t^\star(\mu_m) = \theta_{t_0}^\star(\mathcal{L}_X(\mu_m)),$$

and hence the flow θ_t is volume-preserving if and only if X is **divergence-free** (meaning that $\operatorname{div}_m(X) = 0$).

THEOREM 48.10 (The Divergence Theorem). *Let (M, m) be an oriented Riemannian manifold. Then*

$$\int_{M,m} \operatorname{div}_m(X) = 0, \qquad \forall\, X \in \mathfrak{X}(M) \text{ with compact support.}$$

Proof. We use Cartan's Magic Formula (Theorem 20.6):

$$\operatorname{div}_m(X)\mu_m = \mathcal{L}_X(\mu_m) = di_X(\mu_m) + i_X d\mu_m = di_X(\mu_m) + 0,$$

since $d\mu_m = 0$ as μ_m is a top-dimensional form. Thus

$$\int_{M,m} \operatorname{div}_m(X) = \int_M d(i_X(\mu_m)) = 0$$

by Stokes' Theorem. ■

REMARK 48.11. There is also a version of the Divergence Theorem for manifolds with boundary—see Problem V.2.

If f is a function on M then df is a one-form, and hence $df^\sharp$ is a vector field.

DEFINITION 48.12. Let (M, m) be a Riemannian manifold. Let $f \in C^\infty(M)$. The **gradient** of f is the vector field $\operatorname{grad}_m(f) := df^\sharp$. In local coordinates (x^i) if we write $df = \partial_j f\, dx^j$ then

$$\operatorname{grad}_m(f) = (m^{ij}\partial_j f)\partial_i. \tag{48.2}$$

Warning: It is very common to use the notation "∇f" to denote the gradient of f. This is rather misleading, since it clashes with our notation for the connection ∇.

The divergence of the gradient is the Laplacian:

DEFINITION 48.13 (First definition of the Laplacian). Let (M, m) be an orientable Riemannian manifold. Given $f \in C^\infty(M)$, we define the **Laplacian** of f to be the smooth function

$$\Delta_m(f) := \operatorname{div}_m(\operatorname{grad}_m(f)).$$

We will shortly prove that for M an open subset of $\mathbb{R}^n$ this coincides with the "usual" definition of the Laplacian.

LEMMA 48.14. *Suppose $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$. Then*

$$\operatorname{div}_m(fX) = df(X) + f \operatorname{div}_m(X).$$

Proof. Both sides are functions, and thus it is sufficient to verify the identity pointwise. Thus fix $x \in M$. If $X(x) = 0$ then both sides are zero. Otherwise, choose a neighbourhood U of x such that $TM|_U$ is trivial (for instance, the domain of a chart). After shrinking U, we may assume that X is never zero on U. Thus we may pick a local frame $(X = X_1, X_2, \ldots, X_n)$. Let $\mu_m|_U \coloneqq \mu_U$. Then for any 1-form $\omega \in \Omega^1(U)$ we have

$$\begin{aligned}(\omega \wedge i_X \mu_U)(X_1, X_2, \ldots, X_n) &= \omega(X) i_X \mu_U(X_2, \ldots, X_n) \\ &= (\omega(X)\mu_U)(X_1, X_2, \ldots, X_n).\end{aligned}$$

Since (X_i) is a local frame, we conclude

$$\omega \wedge i_X \mu_U = \omega(X)\mu_U.$$

The result follows by applying this to $\omega = df$, and using the fact that $\mu_U(x) \neq 0$. ■

In a similar vein to Definition 48.4 we can define an inner product on $\Omega^1(M)$. As before, we will only make this definition in the compact case.

DEFINITION 48.15. Let (M, m) be a compact oriented Riemannian manifold. We define an inner product(also denoted by) $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ on $\Omega^1(M)$ by setting

$$\langle\!\langle \omega, \vartheta \rangle\!\rangle \coloneqq \int_{M,m} \left\langle \omega^\sharp, \vartheta^\sharp \right\rangle$$

DEFINITION 48.16. Let (M, m) be a compact oriented Riemannian manifold. The **adjoint**[5] of the exterior derivative $d \colon C^\infty(M) \to \Omega^1(M)$ with respect to $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is the map $\delta_m \colon \Omega^1(M) \to C^\infty(M)$ defined according to the recipe

$$\langle\!\langle f, \delta_m(\omega) \rangle\!\rangle = \langle\!\langle df, \omega \rangle\!\rangle .$$

This gives us a second way of defining the Laplacian. This definition will ultimately prove more useful for us when we discuss Hodge theory. A disadvantage is (at least as we have defined things) this one only works when M is compact[6].

DEFINITION 48.17 (Second definition of the Laplacian). Let (M, m) be a compact oriented Riemannian manifold. Given $f \in C^\infty(M)$, define the **Laplacian**

$$\Delta_m(f) \coloneqq -\delta_m(df).$$

Of course, it must be proved that Definition 48.13 is equivalent to Definition 48.17. We will do this shortly.

LEMMA 48.18. *It holds that*

$$\delta_m(\omega) = -\operatorname{div}_m(\omega^\sharp).$$

[5]Strictly speaking, this is only a "formal" adjoint as $C^\infty(M)$ and $\Omega^1(M)$ are not Hilbert spaces under $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ (see Remark 48.6). This can be rectified as follows: The exterior derivative extends to a linear operator from between the Hilbert space of L^2-functions on M to the Hilbert space of differential forms of class L^2. The adjoint δ_m is then the usual adjoint in the functional analysis sense.

[6]Actually as mentioned earlier, with a bit more work this still makes sense in the non-compact case too.

Proof. We show
$$\langle\!\langle df, \omega \rangle\!\rangle = \left\langle\!\!\left\langle f, -\operatorname{div}_m(\omega^\sharp) \right\rangle\!\!\right\rangle.$$
For this one observes that
$$df(\omega^\sharp) + f \operatorname{div}_m(\omega^\sharp) = \operatorname{div}_m(f\omega^\sharp)$$
by Lemma 48.14 and hence
$$\langle\!\langle df, \omega \rangle\!\rangle - \left\langle\!\!\left\langle f, -\operatorname{div}_m(\omega^\sharp) \right\rangle\!\!\right\rangle = \int_{M,m} \operatorname{div}_m(f\omega^\sharp) = 0$$
by the Divergence Theorem 48.10. ∎

Corollary 48.19. *When M is compact Definition 48.13 is equivalent to Definition 48.17.*

We now define the Hessian of a smooth function, which will lead us to a third equivalent definition of the Laplacian. This definition works for an arbitrary connection on an arbitrary manifold (not necessarily Riemannian).

Definition 48.20. Let M be a smooth manifold and let ∇ denote a connection on M. Let $f \in C^\infty(M)$. We define the **Hessian** $\operatorname{Hess}^\nabla(f)$ of f to be the tensor of type $(0,2)$ defined by
$$\begin{aligned}\operatorname{Hess}^\nabla(f)(X,Y) &:= \nabla_X(df)(Y) \\ &= X(Y(f)) - df(\nabla_X(Y)).\end{aligned}$$

In general the Hessian $\operatorname{Hess}^\nabla(f)$ depends on the choice of connection ∇ on M. However at a point $x \in M$ such that $df|_x = 0$ (such x is called a **critical point** of x), the operator $\operatorname{Hess}^\nabla(f)|_x$ is clearly independent of ∇. The next lemma clarifies this.

Lemma 48.21. *Let M be a smooth manifold and let ∇ denote a connection on M. Let $f \in C^\infty(M)$. Fix a point $x \in M$ and assume (x^i) are local coordinates about x. Write*
$$\operatorname{Hess}^\nabla(f) = H_{ij}\, dx^i \otimes dx^j.$$
Assume that either:

(i) *x is a critical point of f,* **or**

(ii) *∇ is the Levi-Civita connection of some Riemannian metric m on M, and the (x^i) are normal coordinates about x.*

Then
$$H_{ij} = \partial_{ij} f,$$
i.e. the Hessian is the matrix of second partial derivatives of f.

Proof. We have
$$\operatorname{Hess}^\nabla(f)(\partial_i, \partial_j) = \partial_{ij} f - df(\nabla_{\partial_i}\partial_j).$$
In both cases the second term vanishes at x. ∎

Next, we define the trace of a tensor. This again works for an arbitrary smooth manifold (compare to Proposition 36.3).

Definition 48.22. Let M be a smooth manifold and let $T \in \mathcal{T}^{1,1}(M)$ denote a tensor of type $(1,1)$. The **trace** of T is the smooth function $\operatorname{tr}(T) \in C^\infty(M) = \mathcal{T}^{0,0}(M)$ obtained by taking the usual trace pointwise

$$\operatorname{tr}(T)(x) = \operatorname{trace}(T_x \colon T_xM \to T_xM).$$

More generally, if T is a tensor of type $(r+1, s+1)$, we can define the trace of T to be the tensor $\operatorname{tr}(T)$ of type (r,s) by requiring that

$$\operatorname{tr}(T)(\omega_1, \dots \omega_r, X_1, \dots, X_s) = \operatorname{tr}\big(T(\square, \omega_1, \dots \omega_r, \square, X_1, \dots, X_s)\big),$$

which makes sense as $T(\square, \omega_1, \dots \omega_r, \square, X_1, \dots, X_s)$ is a tensor of type $(1,1)$.

If we start with a Riemannian manifold, we can take other traces:

Definition 48.23. Let (M, m) be a Riemannian manifold. We define

$$\operatorname{tr}_m \colon \mathcal{T}^{0,2}(M) \to C^\infty(M)$$

by first using the musical isomorphism to convert a tensor of type $(0,2)$ into a tensor of type $(1,1)$, and then taking the trace as above. Explicitly, if $x \in M$ and (X_i) is a local orthonormal frame of TM about x then near x we have

$$\operatorname{tr}_m(T) = \sum_{i=1}^n T(X_i, X_i). \tag{48.3}$$

A similar construction gives traces $\mathcal{T}^{r,s+2}(M) \to \mathcal{T}^{r,s}(M)$ and $\mathcal{T}^{r+2,s}(M) \to \mathcal{T}^{r,s}(M)$.

We now present our third (and final) definition of the Laplacian. This definition does not use orientability of M, but we include it anyway to be consistent with the other definitions (cf. Remark 48.2).

Definition 48.24 (Third definition of the Laplacian). Let (M, m) be an oriented Riemannian manifold. Given $f \in C^\infty(M)$, define the **Laplacian**

$$\Delta_m f \coloneqq \operatorname{tr}_m\big(\operatorname{Hess}^\nabla(f)\big),$$

where ∇ is the Levi-Civita connection of m.

In order to see the equivalence of Definition 48.24 with the other two definitions, we use the following lemma.

Lemma 48.25. *Let Δ_m be defined as in Definition 48.13 (our original definition). Then in local coordinates (x^i) on $U \subseteq M$ we have*

$$\Delta_m(f) = m^{ij}\partial_{ij} f + \text{lower order terms.}$$

where as usual $(m^{ij})_{1 \le i,j \le n}$ is the inverse matrix to $A \coloneqq (m_{ij})_{1 \le i,j \le n}$.

Proof. For convenience, put $a := \sqrt{\det A}$ so that

$$\mu_U = a dx^1 \wedge \cdots \wedge dx^n$$

by Lemma 48.7. Then if $X = h^i \partial_i$ is any vector field we have

$$\begin{aligned} i_X(\mu_U)(\partial_1, \ldots, \widehat{\partial}_i, \ldots, \partial_n) &= \mu_U(X, \partial_1, \ldots, \widehat{\partial}_i, \ldots, \partial_n) \\ &= (-1)^{i-1}\mu_U(\partial_1, \ldots, \partial_{i-1}, X, \partial_{i+1}, \ldots \partial_n), \end{aligned}$$

where as usual the $\widehat{\ }$ means we omit that entry. The last term is equal to $(-1)^{i-1}ah^i$, since all the other terms die. Since $\{dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n\}$ forms a local frame for $\Omega^{n-1}(M)$ we conclude that

$$i_X(\mu_U) = \sum_{i=1}^{n} (-1)^{i-1} ah^i \, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n,$$

and thus

$$\begin{aligned} d(i_X(\mu_U)) &= \sum_{i,j=1}^{n} (-1)^{i-1}\partial_j(ah^i) dx^j \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \\ &= \partial_i(ah^i)\frac{1}{a}\,\mu_U. \end{aligned}$$

Thus

$$\operatorname{div}_m(X) = \frac{1}{a}\,\partial_i(ah^i).$$

Now take $X = df^\sharp$ and use (48.2) to obtain

$$\Delta_m(f) = \operatorname{div}_m(\operatorname{grad}_m(f)) = \frac{1}{a}\partial_i(am^{ij}\partial_j f) = m^{ij}\partial_{ij} f + \text{ lower order terms.}$$

This completes the proof. ■

Corollary 48.26. *Definition 48.24 is equivalent to Definition 48.13.*

Proof. Apply Lemma 48.25 in the special case where the (x^i) are normal coordinates at a point $x \in M$, so $m^{ij}(x) = \delta^{ij}$, and hence in this case the lemma simplifies to give

$$\operatorname{div}_m(\operatorname{grad}_m(f))(x) = \sum_{i=1}^{n} \partial_{ii} f(x),$$

which is also equal to $\operatorname{tr}(\operatorname{Hess}^\nabla(f))(x)$ in these coordinates by Lemma 48.21. (Note this last expression also connects the Laplacian with the "usual" definition of the Laplacian of a smooth function.) ■

Ricci curvature and Einstein metrics

In this lecture we first investigate how the sectional curvatures change when one changes the metric. This will lead us to the hyperbolic plane.

Definition 49.1. Let M be a smooth manifold. Two Riemannian metrics m_1 and m_2 on M are **conformally equivalent** if there exists a smooth positive function $f\colon M \to (0, \infty)$ such that $m_2 = f\, m_1$.

Let us now compute how the Levi-Civita connection and its curvature tensor change under conformal equivalence. In the following we let $m = \langle \cdot, \cdot \rangle$ denote a Riemannian metric on M and we let $\tilde{m} = f\, m$ denote a conformally equivalent metric. Let

$$h := \log \sqrt{f} \qquad \text{so that} \qquad \tilde{m} = e^{2h}\, m$$

Lemma 49.2. *Let ∇ be the Levi-Civita connection of m and let $\tilde{\nabla}$ denote the Levi-Civita connection of $\tilde{m}$. Then for $X, Y \in \mathfrak{X}(M)$ one has*

$$\tilde{\nabla}_X(Y) - \nabla_X(Y) = X(h)Y + Y(h)X - \langle X, Y\rangle \operatorname{grad}_m(h).$$

Note that if h is a constant function then $\tilde{\nabla} = \nabla$—this once again shows that the Levi-Civita connection is homogeneous in the sense of Definition 46.23. Next, we have:

Lemma 49.3. *Let R^∇ be curvature tensor of the Levi-Civita connection ∇ of m and let $R^{\tilde{\nabla}}$ denote the curvature tensor of the Levi-Civita connection $\tilde{\nabla}$ of $\tilde{m}$. Then for $X, Y, Z \in \mathfrak{X}(M)$ one has*

$$\begin{aligned} R^{\tilde{\nabla}}(X,Y)(Z) - R^{\nabla}(X,Y)(Z) = {} & \langle \nabla_X(\operatorname{grad}_m(h)), Z\rangle\, Y - \langle \nabla_Y(\operatorname{grad}_m(h)), Z\rangle\, X \\ & - \langle X, Z\rangle\, \nabla_Y(\operatorname{grad}_m(h)) - \langle Y, Z\rangle\, \nabla_X(\operatorname{grad}_m(h)) \\ & + Y(h)Z(h)X - \langle Y, Z\rangle\, |\operatorname{grad}_m(h)|^2 X \\ & - X(h)Z(h)Y + \langle X, Z\rangle\, |\operatorname{grad}_m(h)|^2 Y \\ & + X(h)\, \langle Y, Z\rangle \operatorname{grad}_m(h) - Y(h)\, \langle X, Z\rangle \operatorname{grad}_m(h). \end{aligned}$$

The proof of Lemma 49.3 is an easy, albeit lengthy computation; I leave it to the conscientious reader as a wholesome exercise.

Corollary 49.4. *Let $x \in M$ and let $\Pi = \operatorname{span}\{e_1, e_2\} \subset T_xM$, where the e_i are orthonormal with respect to m. Then*

$$\begin{aligned} f(x) \operatorname{sect}_{\tilde{m}}(x; \Pi) - \operatorname{sect}_m(x; \Pi) = {} & - \langle \nabla_{e_1}(\operatorname{grad}_m(h)), e_1\rangle - \langle \nabla_{e_2}(\operatorname{grad}_m(h), e_2\rangle \\ & - |\operatorname{grad}_m(h)(x)|^2 + e_1(h)^2 + e_2(h)^2. \end{aligned}$$

In dimension 2 the formula is simpler.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

Corollary 49.5. *Suppose* $\dim M = 2$. *Then as functions on* M,

$$f\ \mathrm{sect}_{\tilde{m}} - \mathrm{sect}_m = -\Delta_m(h).$$

Proof. The right-hand side of the formula from Corollary 49.4 reduces to $-\operatorname{tr}(\mathrm{Hess}^\nabla(h))$. ■

Recall our notation for a half-space from Definition 21.2.

Definition 49.6. Let $\mathbb{H}^n := \mathbb{R}^n_{x^n>0}$. We equip $\mathbb{H}^n$ with the metric $m_{\mathrm{hyp}} := f\, m_{\mathrm{Eucl}}$ where f is the smooth positive function $f(x) = \frac{1}{(x^n)^2}$. Thus $h = \log \frac{1}{x^n}$ and Corollary 49.4 becomes

$$f\ \mathrm{sect}_{m_{\mathrm{hyp}}}(x; \Pi) - 0 = -f.$$

Thus $(\mathbb{H}^n, m_{\mathrm{hyp}})$ is a space with constant curvature $\kappa = -1$. We call $(\mathbb{H}^n, m_{\mathrm{hyp}})$ the n-dimensional **hyperbolic plane**. More generally if we take $f = \frac{r^2}{(x^n)^2}$ then we get a space with constant curvature $\kappa = -\frac{1}{r^2}$. We denote this metric by $m_{\mathrm{hyp};r}$.

We conclude the our discussion on sectional curvature with the following theorem. In the following we say a Riemannian manifold (M, m) is **complete**[1] if the Levi-Civita connection ∇ of m is complete in the sense of Definition 42.9.

Theorem 49.7 (Killing-Hopf). *Let* (M, m) *be a connected, simply connected and complete Riemannian metric with constant curvature* κ. *Then* (M, m) *is isometric to exactly one of the following three manifolds:*

(i) $(\mathbb{R}^n, m_{\mathrm{Eucl}})$ *if* $\kappa = 0$,

(ii) $(S^n(r), m_{\mathrm{round}})$ *if* $\kappa > 0$, *where* $r := \frac{1}{\sqrt{\kappa}}$.

(iii) $(\mathbb{H}^n, m_{\mathrm{hyp};r})$ *if* $\kappa < 0$, *where* $r := \frac{1}{\sqrt{-\kappa}}$.

Sadly we won't have enough time to prove Theorem 49.7. We will however prove several related results in Lecture 52, starting with the famous *Cartan-Hadamard Theorem* (Theorem 53.14).

Instead for now we move onto our next variant of the curvature tensor. For this consider the trace operator

$$\operatorname{tr}\colon \mathcal{T}^{1,3}(M) \to \mathcal{T}^{0,2}(M)$$

from Definition 48.22.

Definition 49.8. Let (M, m) be a Riemannian manifold with Levi-Civita connection ∇. The **Ricci** tensor of m is the $(0, 2)$-tensor Ric_m obtained by taking the trace of R^∇:

$$\mathrm{Ric}_m(X, Y) := \operatorname{tr}(u \mapsto R^\nabla(u, X)(Y)), \qquad X, Y \in \mathfrak{X}(M).$$

[1] In Lecture 52 we will see that this is equivalent to asking that M is complete as a metric space.

Assume $\dim M = n$. Take an orthonormal basis $\{e_1, \dots, e_n\}$ of T_xM and let $v, w \in T_xM$. Then we have[2]

$$\begin{aligned} \mathrm{Ric}_m(v,w) &= \sum_{i=1}^n \left\langle R^\nabla(e_i,v)(w), e_i \right\rangle \\ &= \sum_{i=1}^n \mathcal{R}_m^\nabla(e_i, w, e_i, v) \\ &= \sum_{i=1}^n \mathcal{R}_m^\nabla(e_i, v, e_i, w), \end{aligned}$$

where the last line used the symmetries of $\mathcal{R}_m^\nabla$ from Lemma 47.2. This proves:

LEMMA 49.9. *The Ricci tensor is symmetric.*

Moreover, the computation above shows that we can view Ric_m as being obtained by taking the trace of $\mathcal{R}_m^\nabla$ over the first and third variables, i.e. under the map $\mathrm{tr}_m \colon \mathcal{T}^{0,4}(M) \to \mathcal{T}^{0,2}(M)$ from Definition 48.23:

$$\mathrm{Ric}_m = \mathrm{tr}_m(\mathcal{R}_m^\nabla)$$

In local coordinates (x^i) we can write

$$\mathrm{Ric}_m = r_{ij}\, dx^i \otimes dx^j,$$

where $r_{ij} = \mathrm{Ric}_m(\partial_i, \partial_j)$. If the (x^i) are normal coordinates at $x \in M$ then we can choose $\{\partial_i|_x\}$ as our orthonormal basis of T_xM. Then by (47.1) we have at x:

$$r_{jl}(x) = \sum_{i=1}^n R_{ijil}(x) = \sum_{i=1}^n m_{hi}(x) R^h_{ilj}(x) = R^i_{ilj}(x) = R^i_{ijl}(x). \tag{49.1}$$

REMARK 49.10. Unlike the sectional curvatures, if $\dim M \geq 4$, the full curvature tensor $\mathcal{R}_m^\nabla$ is in general not completely determined by the Ricci tensors. This should not surprise you, as one typically not recover a matrix from its trace. When $\dim M = 2$ or $\dim M = 3$ however it is possible to recover $\mathcal{R}_m^\nabla$ from Ric_m, as you show on Problem Sheet V.

The Ricci tensor is a symmetric tensor of type $(0,2)$. The metric is another symmetric tensor of type $(0,2)$, and it therefore makes sense to ask whether the two are related. In general the answer is "no": for instance, there is no reason why Ric_m should be positive definite.

DEFINITION 49.11. We say that a metric m is an **Einstein metric** on M if there exists a constant $\lambda \in \mathbb{R}$ such that

$$\mathrm{Ric}_m = \lambda\, m.$$

[2] As you will see, the summation signs invariably need writing when discussing the Ricci tensor, since all expressions have a sum over e_i but both instances of the index i appear on the bottom. This is a general feature of taking the $(0,-2)$-trace of a tensor.

We will discuss the motivation for this condition (together with an explanation of the name) at the end of the lecture. However let us note now that this notion is only interesting when $\dim M \geq 4$. Indeed, on Problem Sheet V you will prove that if $\dim M = 2$ or $\dim M = 3$ then a metric m is Einstein if and only if m has constant curvature.

Definition 49.12. For any non-zero $v \in T_xM$ the **Ricci curvature in the direction** v is defined by

$$\mathrm{ric}_m(v) := \frac{\mathrm{Ric}(v,v)}{|v|^2}.$$

If $\|v\| = 1$ then $\mathrm{ric}_m(v) = \mathrm{Ric}_m(v,v)$. Moreover if $\|v\| = 1$ we may extend $\{v\}$ to an orthonormal basis $\{e_1 = v, e_2, \dots, e_n\}$ of T_xM. Then

$$\mathrm{ric}_m(v) = \sum_{i=1}^n \mathcal{R}_m^\nabla(e_i, v, e_i, v) = \sum_{i=2}^n \mathcal{R}_m^\nabla(e_i, v, e_i, v),$$

since $\mathcal{R}_m^\nabla(e_1, e_1, e_1, e_1) = 0$, and thus $\frac{\mathrm{ric}_m(v)}{n-1}$ is an average of sectional curvatures $\mathrm{sect}_m(x; \Pi_i)$ where

$$\Pi_i = \mathrm{span}\,\{v, e_i\}\,, \qquad i \geq 2.$$

The next lemma is analogous to Corollary 47.12.

Lemma 49.13. *The Ricci curvatures at x are all equal to a constant (say λ) if and only if $\mathrm{Ric}_m = \lambda\, m$ at x (i.e. m is Einstein "at x").*

Proof. One way is clear. For the converse, we simply note that the Ricci curvatures all being equal to λ imply that for any non-zero $v \in T_xM$ we have $\mathrm{Ric}_m(v,v) = \lambda \langle v, v\rangle$. Since $\mathrm{Ric}_m(\cdot,\cdot)$ is a symmetric bilinear form the standard polarisation identity gives

$$\begin{aligned} 2\,\mathrm{Ric}_m(v,w) &= \mathrm{Ric}_m(v+w, v+w) - \mathrm{Ric}_m(v,v) - \mathrm{Ric}_m(w,w) \\ &= \lambda \langle v+w, v+w\rangle - \lambda \langle v, v\rangle - \lambda \langle w, w\rangle \\ &= 2\lambda \langle v, w\rangle\,, \end{aligned}$$

since the polarisation identity also applies to the symmetric bilinear form $\langle \cdot, \cdot \rangle$. ■

We obtained the Ricci curvature by tracing the full curvature, thus reducing a tensor of type $(0,4)$ to one of type $(0,2)$. We can repeat the process to obtain a tensor of type $(0,0)$ (i.e. a smooth function). This function gets its own name:

Definition 49.14. The **scalar curvature** $\mathrm{scal}_m \in C^\infty(M)$ is the trace of the Ricci curvature:

$$\mathrm{scal}_m := \mathrm{tr}_m(\mathrm{Ric}_m) = \mathrm{tr}_m \circ \mathrm{tr}_m(\mathcal{R}_m^\nabla).$$

Fix $x \in M$. If $\{e_1, \dots, e_n\}$ is an orthonormal basis of T_xM then we have

$$\mathrm{scal}_m(x) = \sum_{i=1}^n \mathrm{ric}_m(e_i).$$

Thus $\frac{\mathrm{scal}_m(x)}{n}$ is an average of Ricci curvatures at x. Moreover if (x^i) are normal coordinates at x then by (48.3) and (49.1) we have (writing the summation signs for clarity):

$$\mathrm{scal}_m(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ijij}(x).$$

The following theorem is in a similar vein to Schur's Theorem 47.14.

THEOREM 49.15 (The Ricci curvature version of Schur's Theorem). *Let (M, m) be a connected Riemannian manifold of dimension $n \geq 3$. Then if the Ricci curvatures of M are pointwise constant, that is, $\mathrm{ric}_m(v) = \lambda(x)$ for all $v \neq 0 \in T_xM$, where $\lambda \in C^\infty(M)$, then λ is a constant function, and hence m is an Einstein metric.*

Proof. Fix $x \in M$ and let $(x^1, \dots, x^n)$ be normal coordinates about x. By Lemma 49.13 and the assumption we have

$$r_{ij}(x) = \lambda(x) m_{ij}(x).$$

In what follows, everything is to be evaluated at x; for notational simplicity however we will omit this from the notation. We will also once again suspend our use of the summation convention, as it will prove confusing in this proof. Fix some $p \in \{1, \dots, n\}$. Then by (47.3),

$$\partial_p R_{hjhi} + \partial_i R_{hjph} + \partial_h R_{hjip} = 0. \tag{49.2}$$

Using (49.1) together with the fact that the first derivatives of m_{ij} vanish at x (i.e. (45.5)) we obtain

$$\delta_{ij} = \partial_p r_{ij} = \sum_h R_{hihj}$$

and hence for *any* $1 \leq i \leq n$,

$$\partial_p \lambda = \sum_h \partial_p R_{hihi}. \tag{49.3}$$

Thus setting $i = j$ in (49.2) and summing over h we have

$$\partial_p \lambda + \sum_h \partial_i R_{hiph} + \sum_h \partial_h R_{hiip} = 0,$$

and so summing both sides over i,

$$n\partial_p \lambda + \sum_h \sum_{i \neq p} \partial_i R_{hihp} + \sum_h \partial_p R_{hpph} + \sum_{h \neq p} \sum_i \partial_h R_{ihpi} + \sum_i \partial_p R_{piip} = 0. \tag{49.4}$$

Now

$$\sum_h \sum_{i \neq p} \partial_i R_{hihp} = -\sum_{h \neq p} \sum_i \partial_h R_{ihpi},$$

and

$$\sum_h \partial_p R_{hpph} = -\partial_p \lambda = \sum_i \partial_p R_{piip}$$

we see that (49.4) becomes

$$(n-2)\partial_p \lambda = 0.$$

Since p was arbitrary we conclude $d\lambda|_x = 0$; thus λ is locally constant. Since M is connected, λ is constant. ∎

(♣) Remark 49.16. In several reasonable senses, Einstein metrics are the "best" sort of Riemannian metric a manifold can carry. Here are three explanations as to why:

(i) A naive guess as to what a "best" metric might look like would be to ask that m has constant curvature. But Theorem 49.7 (together with the Cartan-Hadamard Theorem 53.14) shows that this is too restrictive, in the sense that many manifolds M cannot admit such a metric. Indeed, if the universal cover $\widetilde{M}$ of M is not diffeomorphic to $\mathbb{R}^n$ or S^n, then no such metric exists. On the other hand, asking for a metric to have constant scalar curvature is not restrictive enough: one can show that if M is any compact manifold of dimension $n \geq 3$ then M admits an infinite dimensional family of metrics with constant scalar curvature. However the Einstein condition is "just right", in the sense that when Einstein metrics exist, they always occur in finite-dimensional families. It is known that some compact manifolds admit no Einstein metrics, but it is hoped that "most" high-dimensional manifolds do admit them. This is an active field of current research,

(ii) Consider the space $\mathsf{R}_1(M)$ of all Riemannian metrics m on M with volume 1 (i.e. metrics m such that $\mathrm{vol}_m(M) = 1$). This space can be seen as an infinite-dimensional Fréchet manifold. Now consider the functional

$$\mathcal{S}\colon \mathsf{R}_1(M) \to \mathbb{R}, \qquad \mathcal{S}(m) := \int_{M,m} \mathrm{scal}_m .$$

This functional is differentiable, and with a little bit of work one can show that a metric m is a critical point of $\mathcal{S}$ ($d\mathcal{S}|_m = 0$) if and only if m is an Einstein metric. Thus Einstein metrics are obtained by doing calculus of variations on the space of metrics.

(iii) The name "Einstein metric" comes from physics (no surprises there!) In general relativity, one posits that physical spacetime is a four-dimensional manifold equipped with a **Lorentz metric** (this is like a Riemannian metric, apart from instead of being positive definite, it has signature $(3, 1)$—it is negative definite on the time direction). The **Einstein Field Equation** states that

$$\mathrm{Ric}_m - \frac{1}{2}\,\mathrm{scal}_m\, m = T, \tag{49.5}$$

where T is the so-called **stress-energy tensor**. If $T \equiv 0$ then we obtain the **Einstein field equation in a vacuum**. In fact in this case one necessarily has $\mathrm{scal}_m = 0$, and thus the Einstein field equation in a vacuum is equivalent to asking that $\mathrm{Ric}_m = 0$. However from a mathematical point of view, it is then a natural generalisation the vacuum version of (49.5) to look at what we have deemed Einstein metrics.

A wonderful book on this subject (and a gateway to advanced Riemannian geometry in general) is the monograph Einstein Manifolds by Besse. I highly recommend it.

Jacobi fields and the Gauss Lemma

In this lecture we study *Jacobi fields*. These are an essential tool in Riemannian geometry, and much of the rest of the course will be devoted to exploiting them. For instance, next lecture we will use Jacobi fields to deduce that geodesics are locally length-minimising, thus finally justifying the name "geodesic".

Throughout this lecture we assume that M is a smooth manifold of dimension n, that m is a Riemannian metric on M, and that ∇ is the Levi-Civita connection of m. We let T denote the vector field $\frac{\partial}{\partial t} \in \mathfrak{X}(\mathbb{R})$. Thus a curve γ is a geodesic in M if and only if $\nabla_T(\gamma') = 0$. Throughout this lecture we will implicitly assume that all geodesics $\gamma(t)$ are defined on their maximal interval of definition (and thus in particular for $t = 0$). The exception to this is Proposition 50.13 below, where we need to restrict to a compact interval of definition.

DEFINITION 50.1. Let γ be a geodesic in M. A vector field c along γ is called a **Jacobi field** along γ if

$$\nabla_T(\nabla_T(c)) + R^{\nabla}(c, \gamma')(\gamma') = 0. \tag{50.1}$$

We let $\mathrm{Jac}(\gamma)$ denote the space of all Jacobi fields along γ.

The motivation for the Jacobi equation will become clear by the end of the lecture (see Proposition 50.11).

It is clear that $\mathrm{Jac}(\gamma)$ is a vector space, since if $c_1, c_2 \in \mathrm{Jac}(\gamma)$ and $a \in \mathbb{R}$ then by linearity one immediately sees that $ac_1 + c_2$ also satisfies (50.1). A priori however it is not clear that this vector space is finite-dimensional (or, going the other way round, non-trivial!) We will shortly prove in Corollary 50.9 below that $\mathrm{Jac}(\gamma)$ is a vector space of dimension $2n$. Let us first observe that $\mathrm{Jac}(\gamma)$ is non-trivial.

EXAMPLE 50.2. The simplest example of a Jacobi field is γ' itself. Indeed, γ is a geodesic if and only if $\nabla_T(\gamma') = 0$. and thus certainly $\nabla_T(\nabla_T(\gamma')) = 0$ for γ a geodesic. Since R^{∇} is alternating, it follows that γ' satisfies (50.1). More generally, any linear multiple $c(t) := a\gamma'(t)$ for $a \in \mathbb{R}$ is a Jacobi field.

EXAMPLE 50.3. The second simplest example of a Jacobi field is $c(t) := t\gamma'(t)$. Indeed, in this case by the Leibniz rule (part (iv) of Definition 31.8)we have $\nabla_T(c) = T(t)\gamma' + t\nabla_T(\gamma') = \gamma'$, and so $\nabla_T(\nabla_T(c)) = \nabla_T(\gamma') = 0$. Since R^{∇} is a point operator we have $R^{\nabla}(c, \gamma')(\gamma') = tR^{\nabla}(\gamma', \gamma')(\gamma') = 0$, and thus c satisfies (50.1) as claimed. More generally, any linear multiple $c(t) := at\gamma'(t)$ for $a \in \mathbb{R}$ is a Jacobi field.

Example 50.2 and Example 50.3 produce a 2-dimensional subspace of $\mathrm{Jac}(\gamma)$ (unless $\gamma' = 0$). Since these Jacobi fields always occur, when studying Jacobi fields

---

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

it is useful to separate off these examples and look only at Jacobi fields that are *not* of this form. This will take a little bit of work to set up.

LEMMA 50.4. *Let γ be a geodesic and suppose $c_1, c_2 \in \mathrm{Jac}(\gamma)$. Then the function*

$$t \mapsto \langle \nabla_T(c_1)(t), c_2(t)\rangle - \langle c_1(t), \nabla_T(c_2)(t)\rangle$$

is constant.

Proof. Let $f(t) := \langle \nabla_T(c_1)(t), c_2(t)\rangle - \langle c_1(t), \nabla_T(c_2)(t)\rangle$. Then by the Ricci Identity for pullback bundles (Corollary 36.16) we have

$$\begin{aligned} f' &= T\langle \nabla_T(c_1), c_2\rangle - T\langle c_1, \nabla_T(c_2)\rangle \\ &= \langle \nabla_T(\nabla_T(c_1)), c_2\rangle + \langle \nabla_T(c_1), \nabla_T(c_2)\rangle - \langle c_1, \nabla_T(\nabla_T(c_2))\rangle - \langle \nabla_T(c_1), \nabla_T(c_2)\rangle \\ &= \langle \nabla_T(\nabla_T(c_1)), c_2\rangle - \langle c_1, \nabla_T(\nabla_T(c_2))\rangle \end{aligned} \tag{50.2}$$

But by the Jacobi equation (50.1) we have

$$\begin{aligned} \langle \nabla_T(\nabla_T(c_1)), c_2\rangle &= -\left\langle R^{\nabla}(c_1, \gamma')(\gamma'), c_2\right\rangle \\ &= -\mathcal{R}^{\nabla}_m(c_2, \gamma', c_1, \gamma'). \end{aligned}$$

Part (iv) of Lemma 47.2 tells us that

$$\mathcal{R}^{\nabla}_m(c_2, \gamma', c_1, \gamma') = \mathcal{R}^{\nabla}_m(c_1, \gamma', c_2, \gamma'),$$

and thus

$$\langle \nabla_T(\nabla_T(c_1)), c_2\rangle - \langle c_1, \nabla_T(\nabla_T(c_2))\rangle = 0.$$

Combining this with (50.2) completes the proof. ∎

COROLLARY 50.5. *Let c denote a Jacobi field along γ. Then the function*

$$t \mapsto \langle c(t), \gamma'(t)\rangle$$

is affine.

Proof. Since $\nabla_T(\gamma') = 0$ the previous Lemma tells us that the function

$$t \mapsto \langle \nabla_T(c)(t), \gamma'(t)\rangle$$

is constant. By the same argument we have

$$\frac{d}{dt}\langle c(t), \gamma'(t)\rangle = \langle \nabla_T(c)(t), \gamma'(t)\rangle,$$

and thus

$$\frac{d^2}{dt^2}\langle c(t), \gamma'(t)\rangle = 0,$$

which is what we wanted to show. ∎

Thus for any Jacobi field c along γ, we can write

$$\langle c(t), \gamma'(t)\rangle = at + b \tag{50.3}$$

for constants $a, b \in \mathbb{R}$. In fact, we may take

$$a = \frac{d}{dt}\bigg|_{t=0}\langle c(t), \gamma'(t)\rangle, \qquad b = \langle c(0), \gamma'(0)\rangle. \tag{50.4}$$

Definition 50.6. We define the space $\mathrm{Jac}^{\perp}(\gamma)$ of Jacobi fields that are **orthogonal** to γ' as

$$\mathrm{Jac}^{\perp}(\gamma) := \{c \in \mathrm{Jac}(\gamma) \mid \langle c(t), \gamma'(t)\rangle = 0, \ \forall\, t\}.$$

Thus a Jacobi field belongs to $\mathrm{Jac}^{\perp}(\gamma)$ if and only if the constants a and b from (50.3) and (50.4) are both zero. Thus the Jacobi fields γ' and $t\gamma'$ from Examples 50.2 and 50.3 do not belong to $\mathrm{Jac}^{\perp}(\gamma)$. In fact, we will prove shortly that

$$\mathrm{Jac}(\gamma) = \mathrm{Jac}^{\perp}(\gamma) \oplus \mathrm{span}\{\gamma', t\gamma'\}, \tag{50.5}$$

and that $\mathrm{Jac}^{\perp}(\gamma)$ is $(2n-2)$-dimensional.

Remark 50.7. If γ is a geodesic, we can consider the domain of γ as a one-dimensional Riemannian manifold, under the usual Euclidean metric[1]. Thus if c is a Jacobi field then its tangential component $c^{\top}$ (Definition 46.14) and orthogonal component $c^{\perp} := c - c^{\top}$ (Definition 46.15) both make sense. Using Lemma 46.17 we see that

$$c^{\top} = \langle c, \gamma'\rangle\, \gamma' = (a + bt)\gamma'$$

where a, b are as in (50.3) and (50.4). As we have already observed in Examples 50.2 and 50.3, $c^{\top}$ satisfies the Jacobi equation 50.1. Thus as $\mathrm{Jac}(\gamma)$ is a vector space, the orthogonal component $c^{\perp}$ is also a Jacobi field, which in fact lies in $\mathrm{Jac}^{\perp}(\gamma')$.

Proposition 50.8. *Let γ be a geodesic in M, and let t_0 belong to the domain of γ. Let $v, w \in T_{\gamma(t_0)}M$. There exists a unique Jacobi field $c \in \mathrm{Jac}(\gamma)$ such that*

$$c(t_0) = v, \qquad \nabla_T(c)(t_0) = w.$$

An immediate corollary of Proposition 50.8 is:

Corollary 50.9. *The space $\mathrm{Jac}(\gamma)$ is a $2n$-dimensional vector space. Moreover $\mathrm{Jac}^{\perp}(\gamma)$ is a subspace of codimension 2, i.e. (50.5) holds.*

Proof. For any t_0 in the domain of γ the map

$$\mathrm{Jac}(\gamma) \mapsto T_{\gamma(t_0)}M \times T_{\gamma(t_0)}M, \qquad c \mapsto \big(c(t_0), \nabla_T(c)(t_0)\big)$$

is a linear isomorphism by Proposition 50.8. ■

The proof of Proposition 50.8 is by construction:

Proof of Proposition 50.8. We may assume $\gamma'(t)$ is non-zero for every t, as otherwise γ is a constant curve and the claim is trivial (cf. Remark 51.3). Let $e_1 := \gamma'$ and extend $\{e_1\}$ to a parallel local frame $\{e_1, \dots, e_n\}$ along γ such that $\{e_2(t_0), \dots e_n(t_0)\}$ form an orthonormal basis of the orthogonal complement $\gamma'(t_0)^{\perp} \subset T_{\gamma(t_0)}M$ (such a frame exists by Lemma 31.5—note $e_1 = \gamma'$ is parallel by assumption, since γ' is a geodesic.) Since $\nabla m = 0$ we also have that $\{e_2(t), \dots e_n(t)\}$ form an orthonormal

[1] Actually the choice of metric doesn't really matter: any two metrics on $\mathbb{R}$ are isometric, and the only Riemannian invariant of a circle S^1 is its length.

basis of the orthogonal complement $\gamma'(t)^{\perp} \subset T_{\gamma(t)}M$ for all t in their domain. Thus if $c \in \Gamma_\gamma(TM)$ then we can write

$$c = f^i\, e_i, \qquad \text{where} \qquad f^i = \begin{cases} \frac{1}{|\gamma'|^2}\left\langle c, \gamma' \right\rangle, & i = 1, \\ \left\langle c, e_i \right\rangle, & 2 \le i \le n. \end{cases}$$

Since each e_i is parallel we have $\nabla_T(\nabla_T(c)) = \ddot{f}^i\, e_i$, where as in (42.2) we temporarily use the dot notation for the derivative of the real-valued function f^i. Given $1 \le i, j \le n$ set

$$h_i^j := \begin{cases} 0, & , i = 1, \\ h_i^j := \mathcal{R}_m^\nabla(e_j, \gamma', e_i, \gamma'), & 2 \le i \le n. \end{cases}$$

Then c satisfies the Jacobi equation (50.1) if and only if

$$\ddot{f}^j + f^i h_i^j = 0, \qquad 1 \le j \le n. \tag{50.6}$$

The equations (50.6) are a homogeneous system of n linear second-order ordinary differential equations, and thus for a given set of initial conditions (on f^j and $\dot{f}^j$) there is a unique solution. ■

Remark 50.10. Suppose m has constant curvature κ. Then the Jacobi fields are easy to write down explicitly. Consider the ordinary differential equation

$$\ddot{f} + \kappa f = 0. \tag{50.7}$$

Let $a_\kappa(t)$ and $b_\kappa(t)$ denote the unique solutions of (50.7) with

$$\big(a_\kappa(0), \dot{a}_\kappa(0)\big) = (1, 0), \qquad \big(b_\kappa(0), \dot{b}_\kappa(0)\big) = (0, 1).$$

Thus for instance if $\kappa = 1$ then $a_1 = \cos$ and $b_1 = \sin$. Suppose γ is a geodesic. Given $v, w \in T_{\gamma(0)}M$ such that $\langle v, \gamma'(0)\rangle = \langle w, \gamma'(0)\rangle = 0$, let e_v and e_w denote the unique parallel sections along γ with $e_v(0) = v$ and $e_w(0) = w$. Then it follows from the proof of Proposition 50.8 that the unique Jacobi field c with initial conditions $c(0) = v$ and $\nabla_T(c)(0) = w$ is given by

$$c = a_\kappa\, e_v + b_\kappa\, e_w.$$

Proposition 50.8 implies that the Jacobi equation (50.1) is interesting in its own right. But it does not explain the motivation behind the equation—why, for instance, do we study (50.1) and not $\nabla_T(\nabla_T(c)) - R(c, \gamma')(\gamma')$? The next result clarifies this

Proposition 50.11. *The Jacobi equation* (50.1) *is the linearisation of the geodesic equation.*

For this to make sense, we should define precisely what "linearisation" means.

Definition 50.12. Let $\gamma\colon [a, b] \to M$ be a smooth curve. A **variation** of γ is a smooth map $\Gamma\colon (-\varepsilon, \varepsilon) \times [a, b] \to M$ such that $\Gamma(0, t) = \gamma(t)$.

Thus a variation is simply a homotopy of curves. The reason for the name "variation" rather than "homotopy" is that one usually regards a homotopy as running from one curve to another, whereas for a variation we regard the original curve γ as lying in the "middle" of the homotopy. In general we will write a the coordinates of a variation Γ as (s,t); thus s is the homotopy parameter and t is the curve parameter. Observe that

$$\partial_t\Gamma(0,t) = \gamma'(t).$$

Suppose now that Γ is a variation of a geodesic γ which in addition has the property that $t \mapsto \Gamma(s,t)$ is a geodesic for all $s \in (-\varepsilon, \varepsilon)$. In this case we call Γ a **variation of geodesics**. Proposition 50.11 is equivalent (by definition) to the following statement:

PROPOSITION 50.13. *Let $\gamma\colon [0,r] \to M$ be a geodesic and let Γ be a variation of geodesics along γ. Then $\partial_s\Gamma(0,t)$ is a Jacobi field along γ. Moreover if c is any Jacobi field along γ then there exists*[2] *a variation Γ of geodesics along γ such that $c = \partial_s\Gamma(0,\cdot)$.*

(♣) REMARK 50.14. A more satisfying definition of the word "linearisation" will be given next lecture (see Remark 51.19). This uses a little bit of infinite-dimensional differential geometry.

Proof of Proposition 50.13 (and Proposition 50.11). We prove the result in two steps.

1. Suppose $\Gamma\colon (-\varepsilon,\varepsilon) \times [0,r] \to M$ is a variation along γ. Set $c := \partial_s\Gamma(0,t)$. In this step we show that if Γ is a variation of geodesics then $c := \partial_s\Gamma(0,t)$ is a Jacobi field along γ.

Consider the pullback connection (also denoted by) ∇:

$$\nabla\colon \mathfrak{X}((-\varepsilon,\varepsilon)\times[0,r]) \times \Gamma_\Gamma(TM) \to \Gamma_\Gamma(TM).$$

Let $S = \frac{\partial}{\partial s}$ and $T = \frac{\partial}{\partial t}$, so that

$$\partial_s\Gamma = D\Gamma[S], \qquad \partial_t\Gamma = D\Gamma[T].$$

By assumption

$$\nabla_T(\partial_t\Gamma) = 0$$

(because Γ is a variation of geodesics). Next, since $[S,T]=0$, by Proposition 44.8 we have

$$\nabla_S(D\Gamma[T]) - \nabla_T(D\Gamma[S]) = T^\nabla(D\Gamma[S], D\Gamma[T]) = 0. \tag{50.8}$$

Thus

$$\begin{aligned} R^\nabla(\partial_s\Gamma,\partial_t\Gamma)(\partial_t\Gamma) &= \nabla_S(\nabla_T(\partial_t\Gamma) - \nabla_T(\nabla_S(\partial_t\Gamma)) - \nabla_{[S,T]}(\partial_t\Gamma) \\ &= 0 - \nabla_T(\nabla_S(D\Gamma[T]) - 0 \\ &\overset{(\dagger)}{=} -\nabla_T(D\Gamma(\nabla_T(S)) \\ &= -\nabla_T(\nabla_T(\partial_s\Gamma)), \end{aligned}$$

where $(\dagger)$ used (50.8). If we evaluate both sides at $s=0$ this gives

$$R^\nabla(c,\gamma')(\gamma') = -\nabla_T(\nabla_T(c)),$$

and thus c is a Jacobi field.

[2] This variation is *not* unique!

2. Now for the converse direction. We denote by $\exp\colon \mathcal{S} \to M$ the exponential map of m (Definition 46.20), and given $y \in M$ we let $\mathcal{S}_y := \mathcal{S} \cap T_yM$ denote the domain of $\exp_y$. Thus $\mathcal{S}_y$ is star-shaped neighbourhood of 0_y (part (i) of Theorem 43.3).

Suppose $\gamma\colon [0,r] \to M$ is a geodesic and $c \in \mathrm{Jac}(\gamma)$. Set $x := \gamma(0)$ and consider the three tangent vectors $u, v, w \in T_xM$ defined by

$$u := \gamma'(0), \qquad v := c(0), \qquad w := \nabla_T(c)(0).$$

Let $\delta\colon (-\varepsilon, \varepsilon) \to M$ denote any[3] smooth curve such that $\delta(0) = x$ and $\delta'(0) = v$. Let e_u and e_w be parallel vector fields along δ such that $e_u(0) = u$ and $e_w(0) = w$. Note that $u \in \mathcal{S}_x$ (since $\gamma(t) = \exp_x(tu)$). Since $\mathcal{S}$ is open in TM it follows that for $\varepsilon > 0$ sufficiently small one has

$$t(e_u(s) + se_w(s)) \in \mathcal{S}_{\delta(s)}, \qquad \forall\, (s,t) \in (-\varepsilon, \varepsilon) \times [0,r]$$

We now define
$$\Gamma(s,t) := \exp_{\delta(s)}\big(t(e_u(s) + se_w(s))\big).$$
For any fixed s the curve $t \mapsto \Gamma(s,t)$ is a geodesic by the definition of the exponential map. For $s = 0$ we have

$$\Gamma(0,t) = \exp_{\delta(0)}(te_u(0)) = \exp_x(tu) = \gamma(t).$$

Thus by Step 1 the vector field $\tilde{c}(t) := \partial_s\Gamma(0,t)$ is a Jacobi field along γ. To complete the proof we show that $\tilde{c} = c$. By Proposition 50.8 we need only check that
$$\tilde{c}(0) = v, \qquad \nabla_T(\tilde{c})(0) = w.$$
The first equality is clear, since

$$\tilde{c}(0) = \partial_s\Gamma(0,0) = \delta'(0) = v.$$

To see the second equality we compute

$$\nabla_T(\tilde{c})(0) = \nabla_T(D\Gamma[S])(0,0) = \nabla_S(D\Gamma[T])(0,0) = e_w(0) = w,$$

where the second equality used the same argument as (†) in the computation from Step 1. This completes the proof. ∎

The proof of Step 2 of Proposition 50.13 yields the following statement, which is useful in its own right.

Corollary 50.15. *Let γ be a geodesic with $\gamma(0) = x$ and fix $v \in T_xM$. The unique Jacobi field c along γ with $c(0) = 0_x$ and $\nabla_T(c)(0) = v$ is given by*

$$c(t) = D\exp_x(t\gamma'(0))\big[t\mathcal{J}_{t\gamma'(0)}(v)\big]$$

Before stating today's final result, we need one more definition,

[3] Any smooth curve satisfying these conditions will do. This shows why the variation we build is not unique.

Definition 50.16. Let m be a Riemannian metric on M. Given $x \in M$ the tangent space T_xM is a vector space, and thus also a smooth manifold. We now turn T_xM into a Riemannian manifold by defining a Riemannian metric $m_{T_xM} = \langle \cdot, \cdot \rangle_{T_xM}$. The idea is exactly the same as in Example 46.7, only instead of the Euclidean dot product we use the inner product coming from the metric m:

$$\langle \mathcal{J}_u(v), \mathcal{J}_u(w) \rangle_{T_xM} := \langle v, w \rangle , \qquad u, v, w \in T_xM.$$

We conclude this lecture with the following application of Jacobi fields, which is one of many different theorems across mathematics that bears Gauss' name.

Theorem 50.17 (The Gauss Lemma). *Fix $x \in M$ and suppose $v \in T_xM$ belongs to the domain $\mathcal{S}_x$ of $\exp_x$. Let $w \in T_xM$ denote any other tangent vector. Then*

$$\langle D\exp_x(v)[\mathcal{J}_v(v)], D\exp_x(v)[\mathcal{J}_v(w)] \rangle = \langle \mathcal{J}_v(v), \mathcal{J}_v(w) \rangle_{T_xM} . \tag{50.9}$$

In particular

$$\left| D\exp_x(tv)[\mathcal{J}_v(v)] \right| = |v| \tag{50.10}$$

and

$$D\exp_x(tv)[\mathcal{J}_v(v)] \perp D\exp_x(tv)[\mathcal{J}_v(w)] \qquad \text{if} \qquad v \perp w. \tag{50.11}$$

Theorem 50.17 shows that the exponential map at a point is "radially" isometric. It does *not* claim that $\exp_x \colon (\mathcal{S}_x, m_{T_xM}) \to (M, m)$ is an isometry, since we are only allowed to feed $D\exp_x(v)$ the tangent vector $\mathcal{J}_v(v)$ in the first variable, rather than an arbitrary tangent vector $\mathcal{J}_v(u)$.

Proof. The right-hand side of (50.9)is equal to $\langle v, w \rangle$ by definition of the metric m_x. Meanwhile if we let γ denote the geodesic $\exp_x(tv)$ and let c denote the Jacobi field along γ with $c(0) = 0$ and $\nabla_T(c)(0) = w$ then the left-hand side is exactly $\langle \gamma'(1), c(1) \rangle$ by Corollary 50.15. By (50.3) and (50.4) we have

$$\langle c(t), \gamma'(t) \rangle = t \langle v, w \rangle ,$$

and thus taking $t = 1$ completes the proof. ∎

The length and energy functionals of a Riemannian manifold

In this lecture we introduce the length and energy functionals of a Riemannian manifold. We show how geodesics can be detected via calculus of variations. This leads to a new viewpoint on Jacobi fields.

Throughout this lecture, (M, m) is a Riemannian manifold of dimension n, ∇ denotes the Levi-Civita connection of m, and $\exp\colon \mathcal{S} \to M$ denote the exponential map of m. We denote by T the vector field $\frac{\partial}{\partial t}$. In contrast to last lecture however, today we will restrict our attention to curves defined on compact intervals $[a, b]$. We begin by defining the length of a curve.

Definition 51.1. Let $\gamma\colon [a, b] \to M$ denote a smooth curve. We define the **length** of γ as

$$\mathbb{L}_m(\gamma) \colon = \int_a^b |\gamma'(t)| \, dt.$$

It is also useful to consider the **arc-length** α of a curve $\gamma\colon [a, b] \to M$:

$$\alpha\colon [a, b] \to [0, \infty), \qquad \alpha(t) := \int_a^t |\gamma'(t)| \, dt,$$

so that $\alpha(a) = 0$ and $\alpha(b) = \mathbb{L}_m(\gamma)$. Let us say a curve γ is **regular** if $\gamma'(t) \neq 0$ for all t. If γ is regular then its arc-length is a strictly monotone increasing function since

$$\alpha'(t) = |\gamma'(t)| > 0.$$

In this case the inverse function $\beta(t)$ is well-defined. The new curve

$$\delta(t)\colon [0, \mathbb{L}_m(\gamma)] \to M, \qquad \delta(t) := \gamma(\beta(t))$$

then satisfies

$$|\delta'(t)| = |\beta'(t)\gamma'(\beta(t))| = |\beta'(t)\alpha'(\beta(t))| = 1.$$

Thus in particular

$$\mathbb{L}_m(\gamma) = \mathbb{L}_m(\delta). \tag{51.1}$$

We say that a curve δ is of **constant speed** if $|\delta'(t)|$ is constant, and of **unit speed** if this constant is equal to one. We have thus proved:

Lemma 51.2. *If γ is any regular curve then we can reparametrise γ so that it is of unit speed.*

Remark 51.3. Any geodesic γ has constant speed. Indeed

$$\frac{d}{dt}|\gamma'|^2 = T\langle \gamma', \gamma' \rangle = 2\langle \nabla_T(\gamma'), \gamma' \rangle = 0,$$

and thus also $t \mapsto |\gamma'(t)|$ is constant. Thus if γ is a geodesic then either γ is a constant curve or γ is regular.

Remark 51.4. Recall the notion of a piecewise smooth curve from Lecture 32 (defined just before Example 32.6). It is useful to also have a notion of length for such a curve. If $\gamma\colon [a,b] \to M$ is a piecewise smooth curve and $a = a_0 < a_1 < \cdots < a_k = b$ are such that $\gamma|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \dots, k$ then we can define the length[1] of γ as

$$\mathbb{L}_m(\gamma) := \sum_{i=1}^{k} \mathbb{L}_m\big(\gamma|_{[a_{i-1},a_i]}\big).$$

Time for some more notation.

Definition 51.5. Let us abbreviate

$$\mathcal{P}([a,b]) := \{\gamma\colon [a,b] \to M \mid \gamma \text{ is piecewise smooth}\},$$

and

$$\mathcal{P} := \bigcup_{\text{all compact intervals } [a,b]} \mathcal{P}([a,b])$$

Moreover given $x, y \in M$ we set

$$\mathcal{P}_{xy}([a,b]) := \{\gamma \in \mathcal{P}([a,b]) \mid \gamma(a) = x,\ \gamma(b) = y\},$$

and

$$\mathcal{P}_{xy} := \bigcup_{\text{all compact intervals } [a,b]} \mathcal{P}_{xy}([a,b])$$

We will use the letter $\mathcal{C}$ instead of $\mathcal{P}$ when we want to restrict to genuine smooth curves (rather than just piecewise smooth curves). Thus for instance $\mathcal{C}_{xy}$ is the space of all smooth curves γ in M that start at x and end at y. The spaces $\mathcal{C}_{xy}$ are all collectively referred to as the **path spaces** of M.

We can regard $\mathbb{L}_m$ as a functional

$$\mathbb{L}_m\colon \mathcal{P} \to [0,\infty)$$

Lemma 51.6. *The length functional $\mathbb{L}_m\colon \mathcal{P} \to [0,\infty)$ satisfies:*

(i) *$\mathbb{L}_m(\gamma) = 0$ if and only if γ is a constant curve.*

(ii) *If the concatenation $\gamma * \delta$ (Example 32.6) of two curves $\gamma, \delta \in \mathcal{P}$ is defined then*

$$\mathbb{L}_m(\gamma * \delta) = \mathbb{L}_m(\gamma) + \mathbb{L}_m(\delta).$$

[1] Exercise: Why is this independent of the choice of a_i?

(iii) *If* $h\colon [a_1, b_1] \to [a, b]$ *is a monotone piecewise smooth map then*

$$\mathbb{L}_m(\gamma) = \mathbb{L}_m(\gamma \circ h)$$

for all $\gamma \in \mathcal{P}([a, b])$. *In particular*

$$\mathbb{L}_m(\gamma) = \mathbb{L}_m(\gamma^-),$$

where γ^- *is the reverse curve.*

The proof of Lemma 51.6 is easy. For example, part (ii) is an application of the usual change of variables formula for integration. (Note (51.1) was a special case.) This also shows that—as far as $\mathbb{L}_m$ is concerned—without loss of generality we may restrict attention to $[a, b] = [0, 1]$.

There is another functional which is usually more useful than the length functional.

Definition 51.7. We define the **energy functional** $\mathbb{E}_m\colon \mathcal{P} \to [0, \infty)$ by

$$\mathbb{E}_m(\gamma) := \frac{1}{2}\int_a^b |\gamma'(t)|^2\, dt, \qquad \gamma \in \mathcal{P}([a, b]).$$

Unlike the length functional, the energy functional is *not* invariant under reparametrisation.

Lemma 51.8. *For any* $\gamma \in \mathcal{P}([a, b])$ *we have*

$$\mathbb{L}_m(\gamma) \leq \sqrt{2(b-a)\mathbb{E}_m(\gamma)}$$

with equality if and only if γ *has constant speed.*

Proof. Apply the Cauchy-Schwarz inequality. ∎

The fact that the length functional is invariant under reparametrisation and the energy functional is not may seem to indicate the length functional is "better". Actually the converse is true, as we explain in Remark 51.25 at the end of the lecture.

Definition 51.9. Fix $\gamma \in \mathcal{C}_{xy}([a, b])$. We define

$$T_\gamma \mathcal{C}_{xy}([a, b]) := \{c \in \Gamma_\gamma(TM) \mid c(a) = 0_x,\ c(b) = 0_y\} \tag{51.2}$$

Thus $T_\gamma \mathcal{C}_{xy}([a, b])$ is the space of vector fields c along γ which vanish at the endpoints.

Why the tangent space notation?

Theorem 51.10. *The space* $\mathcal{C}_{xy}([a, b])$ *is an infinite dimensional (Fréchet) manifold. The tangent space to* $\mathcal{C}_{xy}([a, b])$ *at* γ *is given by* (51.2).

Since we have not even defined infinite dimensional manifolds properly, we cannot (of course) prove Theorem 51.10. But let us at least observe that the description of the tangent space from Definition 51.9 is reasonable. Firstly, $T_\gamma \mathcal{C}_{xy}([a,b])$ is an infinite dimensional vector space, so it at least plausibly might be a tangent space to an infinite dimensional manifold. Secondly, if $c \in T_\gamma \mathcal{C}_{xy}([a,b])$ then just as in the proof of Proposition 50.13, for $|s|$ sufficiently small the curve

$$\Gamma\colon (-\varepsilon,\varepsilon) \times [a,b] \to M, \qquad \Gamma(s,t) := \exp_{\gamma(t)}(sc(t)) \tag{51.3}$$

is well-defined. Moreover

$$\Gamma(0,t) = \gamma(t), \qquad \partial_s \Gamma(0,t) = c(t),$$

and since c vanishes at the endpoints,

$$\Gamma(s,a) = \gamma(a), \qquad \Gamma(s,b) = \gamma(b).$$

Thus we can think of Γ as a map

$$\tilde{\Gamma}\colon (-\varepsilon,\varepsilon) \to \mathcal{C}_{xy}([a,b]), \qquad \tilde{\Gamma}(s)(t) := \Gamma(s,t),$$

i.e. a curve in the manifold $\mathcal{C}_{xy}([a,b])$, and the tangent vector of this curve is

$$\tilde{\Gamma}'(0) = \partial_s \Gamma(0,\cdot) = c.$$

Thus elements of $T_\gamma \mathcal{C}_{xy}([a,b])$ are tangent vectors to curves in $\mathcal{C}_{xy}([a,b])$, and so Theorem 51.10 makes sense. In a similar vein, we define the piecewise-smooth version:

Definition 51.11. Let $\gamma \in \mathcal{P}_{xy}([a,b])$. Define

$$T_\gamma \mathcal{P}_{xy}([a,b]) := \{c \text{ a piecewise smooth vector field along } \gamma \text{ with } c(a) = 0_x,\ c(b) = 0_y\}.$$

Warning: The "tangent space" notation $T_\gamma \mathcal{P}_{xy}([a,b])$ should be taken with a grain of salt. Indeed, in contrast to Theorem 51.10, the space $\mathcal{P}_{xy}([a,b])$ typically does *not* admit the structure of an infinite-dimensional manifold.

If $c \in T_\gamma \mathcal{P}_{xy}([a,b])$ then just as (51.3) we can consider the (now only piecewise smooth) variation $\Gamma(s,t) = \exp_{\gamma(t)}(sc(t))$.

Definition 51.12. The differential of the energy functional at $\gamma \in \mathcal{P}_{xy}([a,b])$ is the linear map

$$d\mathbb{E}_m|_\gamma \colon T_\gamma \mathcal{P}_{xy}([a,b]) \to \mathbb{R}$$

given by

$$d\mathbb{E}_m|_\gamma(c) := \frac{d}{ds}\bigg|_{s=0} \mathbb{E}_m(\Gamma(s,\cdot))$$

Again, the word "differential" is a little naughty, since as $\mathcal{P}_{xy}([a,b])$ is not a manifold it does not make sense to say that the function $\mathbb{E}_m$ is (or is not) differentiable, and thus its differential is not defined. In fact, even if we replace $\mathcal{P}_{xy}([a,b])$ by $\mathcal{C}_{xy}([a,b])$ then Definition 51.12 can only be understood formally, since the functional $\mathbb{E}_m$ is typically not actually differentiable (!) with respect to this Fréchet manifold structure. This can be rectified by working with curves of lower regularity, as we explain in Remark 51.24 at the end of the lecture.

Let us now compute the differential of $\mathbb{E}_m$. We first give the result in the smooth case.

Proposition 51.13. *Let $\gamma \in \mathcal{C}_{xy}([a,b])$ and $c \in T_\gamma \mathcal{C}_{xy}([a,b])$. Then*

$$d\mathbb{E}_m|_\gamma(c) = -\int_a^b \langle c(t), \nabla_T(\gamma')(t)\rangle \, dt.$$

Proof. Let $\Gamma(s,t) = \exp_{\gamma(t)}(sc(t))$ be as in (51.3). As in the proof of Proposition 50.13 we let S denote the vector field $\frac{\partial}{\partial s}$. Then by differentiating under the integral sign we compute

$$\begin{aligned}
\frac{d}{ds}\mathbb{E}_m(\Gamma(s,\cdot)) &= \frac{1}{2}\int_a^b S\,\langle \partial_t\Gamma, \partial_t\Gamma\rangle \, dt \\
&= \int_a^b \langle \nabla_S(\partial_t\Gamma), \partial_t\Gamma\rangle \, dt \\
&\overset{(\dagger)}{=} \int_a^b \langle \nabla_T(\partial_s\Gamma), \partial_t\Gamma\rangle \, dt \\
&= \int_a^b T\,\langle \partial_s\Gamma, \partial_t\Gamma\rangle \, dt - \int_a^b \langle \partial_s\Gamma, \nabla_T(\partial_t\Gamma)\rangle \, dt \\
&\overset{(\ddagger)}{=} 0 - \int_a^b \langle \partial_s\Gamma, \nabla_T(\partial_t\Gamma)\rangle \, dt,
\end{aligned}$$

where $(\dagger)$ used $\nabla_T(S) = \nabla_S(T)$ and $(\ddagger)$ used the fact that

$$\Gamma(s,a) = \exp_{\gamma(a)}(sc(a)) = \gamma(a), \qquad \Gamma(s,b) = \exp_{\gamma(b)}(sc(b)) = \gamma(b)$$

for all s so that

$$\partial_s\Gamma(s,a) = \partial_s\Gamma(s,b) = 0.$$

Evaluating at $s = 0$ gives

$$d\mathbb{E}_m|_\gamma(c) = -\int_a^b \langle c, \nabla_T(\gamma')\rangle \; dt$$

as required. ■

In the piecewise smooth case the formula is a little less pleasant. Indeed, suppose $\gamma\colon [a,b] \to M$ is a piecewise smooth curve. Let $a = a_0 < a_1 < \cdots < a_k = b$ be such that $\gamma|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \ldots, k$. The two tangent vectors

$$\big(\gamma|_{[a_{i-1},a_i]}\big)'(a_i) \qquad \text{and} \qquad \big(\gamma|_{[a_i,a_{i+1}]}\big)'(a_i) \tag{51.4}$$

will typically not be the same. Thus the line $(\ddagger)$ in the proof of Proposition 51.13 is no longer true: when performing the integration by parts one gets additional error terms and we obtain

$$\begin{aligned}
d\mathbb{E}_m|_\gamma(c) = &-\int_a^b \langle c(t), \nabla_T(\gamma')(t)\rangle \, dt \\
&+ \sum_{i=1}^{k-1} \Big\langle c(a_i), \big(\gamma|_{[a_{i-1},a_i]}\big)'(a_i) - \big(\gamma|_{[a_i,a_{i+1}]}\big)'(a_i) \Big\rangle .
\end{aligned} \tag{51.5}$$

Corollary 51.14. *A curve $\gamma \in \mathcal{P}_{xy}([a,b])$ is a critical point of $\mathbb{E}_m$ (i.e. $d\mathbb{E}_m|_\gamma = 0$) if and only if γ is a geodesic (and thus in particular is smooth, not just piecewise smooth).*

Proof. It is clear from Proposition 51.13 that if $\gamma \in \mathcal{C}_{xy}([a,b])$ then γ is a critical point of $\mathbb{E}_m$ if and only if γ is a geodesic. Thus we need only show that if $\gamma \in \mathcal{P}_{xy}([a,b])$ is a critical point of $\mathbb{E}_m$ then γ is smooth. Using the notation from (51.4), we need to prove that in this case the two tangent vectors are the same. For this choose a smooth function $f\colon [a,b] \to \mathbb{R}$ such that $f(a_i) = 0$ for each a_i and $f(t) > 0$ for $t \neq a_i$. Then set $c(t) := f(t)\nabla_T(\gamma)(t)$. By (51.5) we obtain that

$$0 = d\mathbb{E}_m|_\gamma(c) = -\int_a^b f(t)|\nabla_T(\gamma')(t)|^2\, dt + 0,$$

and thus we conclude that $\nabla_T(\gamma')|_{[a_{i-1},a_i]} = 0$ for each i. Now choose a vector field $\tilde{c} \in T_\gamma \mathcal{P}_\gamma([a,b])$ such that

$$\tilde{c}(a_i) = \big(\gamma|_{[a_{i-1},a_i]}\big)'(a_i) - \big(\gamma|_{[a_i,a_{i+1}]}\big)'(a_i).$$

Then by (51.5) again we get

$$0 = d\mathbb{E}_m|_\gamma(\tilde{c}) = 0 - \sum_{i=1}^{k-1} |\big(\gamma|_{[a_{i-1},a_i]}\big)'(a_i) - \big(\gamma|_{[a_i,a_{i+1}]}\big)'(a_i)|^2$$

which implies that

$$\big(\gamma|_{[a_{i-1},a_i]}\big)'(a_i) = \big(\gamma|_{[a_i,a_{i+1}]}\big)'(a_i), \qquad \forall\, 1 \leq i \leq k-1$$

and hence γ is of class C^1 and satisfies the geodesic equation $\nabla_T(\gamma') = 0$ on all of $[a,b]$. But clearly any C^1 solution of the geodesic is automatically of class C^∞. ■

What happens if we differentiate again? For this we need the notion of a two-parameter family of variations. For this let $\gamma \in \mathcal{P}_{xy}([a,b])$ and let $c_1, c_2 \in T_\gamma \mathcal{P}_{xy}([a,b])$. For $|r|, |s|$ sufficiently small the map

$$\Gamma(r,s,t) := \exp_{\gamma(t)}\big(rc_1(t) + sc_2(t)\big) \tag{51.6}$$

is well-defined. Note $\partial_r\Gamma(0,0,\cdot) = c_1$ and $\partial_s\Gamma(0,0,t) = c_2$.

Definition 51.15. Let $\gamma \in \mathcal{C}_{xy}([a,b])$ be a geodesic. The **Hessian** of $\mathbb{E}_m$ is the symmetric bilinear form

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)\colon T_\gamma\mathcal{P}_{xy}([a,b]) \times T_\gamma\mathcal{P}_{xy}([a,b]) \to \mathbb{R}$$

given by

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1,c_2) := \frac{d^2}{drds}\Big|_{r=s=0} \mathbb{E}_m(\Gamma(r,s,\cdot)),$$

where Γ is defined as in (51.6).

This is (formally) consistent with Definition 48.20—we are only defining the Hessian at a critical point of $\mathbb{E}_m$, and thus we do not need to choose a connection on the infinite dimensional manifold $\mathcal{C}_{xy}([a,b])$. Then analogously to Proposition 51.13 we have the following computation, whose proof is deferred to Problem Sheet W.

Proposition 51.16. *Let $\gamma \in \mathcal{C}_{xy}([a,b])$ be a geodesic and let $c_1, c_2 \in T_\gamma \mathcal{C}_{xy}([a,b])$. Then*

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1,c_2) = -\int_a^b \left\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1,\gamma')(\gamma'), c_2 \right\rangle dt.$$

If instead we only require $c_1, c_2 \in T_\gamma \mathcal{P}_{xy}([a,b])$ then we have

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1,c_2) = &-\int_a^b \left\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1,\gamma')(\gamma'), c_2 \right\rangle dt \\ &+ \sum_{i=1}^{k-1} \left\langle \nabla_T(c_1)|_{[a_{i-1},a_i]}(a_i) - \nabla_T(c_1)|_{[a_i,a_{i+1}]}(a_i), c_2(a_i) \right\rangle,\end{aligned} \tag{51.7}$$

where $a = a_0 < a_1 < \cdots < a_k = b$ is any subdivision of $[a,b]$ such that $c_1|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \dots, k$.

Definition 51.17. Let $\gamma \in \mathcal{C}_{xy}([a,b])$ be a geodesic. We say an element $c \in T_\gamma \mathcal{P}_{xy}([a,b])$ belongs to the **null-space** of $\operatorname{Hess}(\mathbb{E}_m)(\gamma)$ if

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c,\tilde{c}) = 0 \qquad \forall \tilde{c} \in T_\gamma \mathcal{P}_{xy}([a,b]).$$

Corollary 51.18. *Let $\gamma \in \mathcal{C}_{xy}([a,b])$ be a geodesic. An element $c \in T_\gamma \mathcal{P}_{xy}([a,b])$ belongs to the null-space of $\operatorname{Hess}(\mathbb{E}_m)(\gamma)$ if and only if c is a Jacobi field along γ which vanishes at the endpoint (and thus in particular c is smooth).*

Proof. As with the proof of Corollary 51.14, if c is smooth then it is clear from Proposition 51.16 that c belongs to the null-space of the Hessian if and only if c is a Jacobi field along γ which vanishes at the endpoints. Thus we need only show that any element piecewise smooth element c of the null-space is in fact, smooth. For this assume $a = a_0 < a_1 < \cdots < a_k = b$ is a subdivision of $[a,b]$ such that $c_1|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \dots, k$. Let $f \colon [a,b] \to \mathbb{R}$ be a smooth function such that $f(a_i) = 0$ for each i and $f(t) > 0$ for $t \neq i$. Let

$$c_1(t) := f(t)\big(\nabla_T(\nabla_T(c)) + R^\nabla(c,\gamma')(\gamma')\big).$$

Then from (51.7) we obtain

$$0 = \operatorname{Hess}(\mathbb{E}_m)(\gamma)(c,c_1) = -\int_a^b f(t)|\nabla_T(\nabla_T(c)) + R^\nabla(c,\gamma')(\gamma')|^2 dt,$$

which shows that c satisfies the Jacobi equation on each interval $[a_{i-1},a_i]$. In particular for each $i = 1, \dots, k-1$ we have

$$\nabla_T\big(\nabla_T\big(c|_{[a_{i-1},a_i]}\big)\big)(a_i) = -R^\nabla(c(a_i),\gamma'(a_i))(\gamma'(a_i)) = \nabla_T\big(\nabla_T\big(c|_{[a_i,a_{i+1}]}\big)\big)(a_i).$$

Now choose $c_2 \in T_\gamma \mathcal{P}_{xy}([a,b])$ such that $c_2(a) = c_2(b) = 0$ and such that for each $i = 1, \dots, k-1$ we have

$$c_2(a_i) = \nabla_T\big(c|_{[a_{i-1},a_i]}\big)(a_i) - \nabla_T\big(c|_{[a_i,a_{i+1}]}\big)(a_i).$$

Then from (51.7) we obtain

$$0 = \mathrm{Hess}(\mathbb{E}_m)(\gamma)(c, c_2) = \sum_{i=1}^{k-1} |\nabla_T\big(c|_{[a_{i-1},a_i]}\big)(a_i) - \nabla_T\big(c|_{[a_i,a_{i+1}]}\big)(a_i)|^2$$

and hence for each $i = 1, \dots, k-1$ we have

$$\nabla_T\big(c|_{[a_{i-1},a_i]}\big)(a_i) = \nabla_T\big(c|_{[a_i,a_{i+1}]}\big)(a_i).$$

We conclude that c satisfies the Jacobi equation on $[a,b]$ and is of class C^2. It thus follows from Proposition 50.8 that c is of class C^∞. This completes the proof. ■

Remark 51.19. As promised in Remark 50.14, one can think of Corollary 51.14 and Corollary 51.18 as giving a less ad-hoc definition of the word “linearisation”.

Definition 51.20. Let $\gamma \in \mathcal{C}_{xy}([a,b])$ be a geodesic. Set

$$\mathrm{null}(\gamma) := \{c \in \mathrm{Jac}(\gamma) \mid c(a) = 0_x,\ c(b) = 0_y\}.$$

Thus by Corollary 51.18 the space $\mathrm{null}(\gamma)$ can be identified with the null-space of $\mathrm{Hess}(\mathbb{E}_m)(\gamma)$.

If γ is a regular geodesic then the Jacobi fields from Examples 50.2 and 50.3 never have this property, and thus

$$\dim \mathrm{null}(\gamma) \le 2n - 2.$$

We say that a geodesic γ is **non-degenerate** if $\dim \mathrm{null}(\gamma) = 0$, and **degenerate** if $\dim \mathrm{null}(\gamma) > 0$. It is of interest to know when a geodesic first becomes degenerate. This motivates the following definition.

Definition 51.21. Let $\gamma \colon [0,b] \to M$ be a regular geodesic. A point $t_0 \in (0,b]$ is said to be a **conjugate point** of γ if there exists a non-trivial Jacobi field $c \in \mathrm{Jac}(\gamma)$ such that $c(0) = c(t_0) = 0$.

Example 51.22. Assume m has constant curvature κ. Then from Remark 50.10 we see that a regular geodesic γ has a conjugate point at $\frac{\pi}{\sqrt{\kappa}}$ if $\kappa > 0$, meanwhile if $\kappa \le 0$ then no regular geodesic has conjugate points.

This is particularly illustrative for the sphere $(S^n, m_{\mathrm{round}})$, which has $\kappa = 1$. From Problem (ii) the geodesics are the great circles, and thus we see that the first conjugate point occurs at the antipodal point.

One can also connect the notion of a conjugate point with the exponential map. The next result is also on Problem Sheet W you will prove:

Proposition 51.23. *Let $\gamma \colon [0,b] \to M$ be a regular geodesic. A point $t_0 \in (0,b]$ is a conjugate point of γ if and only if $\exp_{\gamma(t_0)}$ does not have maximal rank at $t_0\gamma'(0)$. In fact,*

$$\dim \ker D\exp_{\gamma(0)}(t_0\gamma'(0)) = \dim \mathrm{null}(\gamma|_{[0,t_0]}).$$

We conclude this lecture with a couple of non-examinable remarks concerning regularity.

(♣) Remark 51.24. Whilst $\mathcal{C}_{xy}([a,b])$ admits the structure of an infinite dimensional manifold, it is only a Fréchet manifold. Without going into details, these are a rather badly behaved class of infinite dimensional manifolds (for example, the Implicit Function Theorem is not true for Fréchet manifold). A much better behaved class of infinite dimensional manifolds are Banach manifolds (or better still, Hilbert manifolds). We can turn $\mathcal{C}_{xy}([a,b])$ into a Hilbert manifold by relaxing the regularity assumption. For this let

$$\mathcal{W}^{1,2}_{xy}([a,b])$$

denote the set of absolutely continuous curves $\gamma\colon [a,b] \to M$ such that $|\gamma'|$ is square integrable (in other words, the set of maps $\gamma\colon [a,b] \to M$ that are of **Sobolev class** $W^{1,2}$. Then $\mathcal{W}^{1,2}_{xy}([a,b])$ is a Hilbert manifold with tangent space $T_\gamma \mathcal{W}^{1,2}_{xy}([a,b])$ equal to those sections c along γ that vanish at the endpoints and are also of Sobolev class $W^{1,2}$.

Whilst working with curves of lower regularity might seem like we are making life harder for ourselves, in fact things get much simpler. For instance, if we regard the energy functional as being defined on this space

$$\mathbb{E}_m\colon \mathcal{W}^{1,2}_{xy}([a,b]) \to [0,\infty)$$

then one can show that $\mathbb{E}_m$ is a differentiable function of class C^2, and in this case its differential $d\mathbb{E}_m|_\gamma$ at $\gamma \in \mathcal{W}^{1,2}_{xy}([a,b])$ is indeed given by the expression from Definition 51.12, and the Hessian of $\mathbb{E}_m$ (defined as in Definition 48.20) at a critical point γ is indeed given by Definition 51.15. Moreover, since the geodesic equation $\nabla_T(\gamma') = 0$ is an *elliptic* equation, one can use a powerful technique known as *elliptic regularity* to prove that if γ is a curve of Sobolev class $W^{1,2}$ such that $\nabla_T(\gamma') = 0$ then γ is automatically smooth. Thus the critical points of $\mathbb{E}_m$ when regarded as a functional on $\mathcal{W}^{1,2}_{xy}([a,b])$ are the same as the critical points of $\mathbb{E}_m$ on $\mathcal{C}_{xy}([a,b])$, i.e. the geodesics.

Finally let us justify why the energy functional is "better" than the length functional.

(♣) Remark 51.25. The length functional $\mathbb{L}_m$ is less well-behaved for two reasons:

(i) Since $\mathbb{L}_m$ is invariant under reparametrisation, critical points of $\mathbb{L}_m$ come in infinite-dimensional families. Indeed, a similar computation to Proposition 51.13 shows that $\gamma \in \mathcal{P}_{xy}([a,b])$ is a critical point of $\mathbb{L}_m$ if and only if γ is the reparametrisation of a geodesic. This is "bad": in general when one wants to do calculus of variations it is better to have as "few" critical points as possible.

(ii) Secondly, even if we consider $\mathbb{L}_m$ as a functional on the Hilbert manifold $\mathcal{W}^{1,2}_{xy}([a,b])$, it is not differentiable. (This is because $t \mapsto \sqrt{t}$ is not differentiable at $t = 0$.) One could fix this by restricting $\mathbb{L}_m$ to the submanifold of regular curves inside $\mathcal{W}^{1,2}_{xy}([a,b])$, since on such a curve $\mathbb{L}_m$ is differentiable. However, this creates additional technical complications elsewhere.

The metric structure of a Riemannian manifold

In this lecture we finally make good on our claim (Remark 42.6) that geodesics on a Riemannian manifold really can be thought of—at least locally—as length-minimising curves (and thus the name "geodesic" coincides with standard meaning outside of mathematics).

Throughout this lecture, (M, m) is a Riemannian manifold of dimension n, ∇ is the Levi-Civita connection of m, and $\exp\colon \mathcal{S} \to M$ is the exponential map of m. Here is our key definition for today:

Definition 52.1. Given $x, y \in M$ we define the m-**distance** between x and y by

$$d_m(x, y) := \inf \{\mathbb{L}_m(\gamma) \mid \gamma \in \mathcal{P}_{xy}([0, 1])\},$$

where by convention the infimum of the empty set is defined to be ∞.

If M is connected then $\mathcal{P}_{xy}([0, 1]) \neq \emptyset$ for all x, y, and thus $d_m(x, y) < \infty$. Note that since $\mathbb{L}_m$ is invariant under reparametrisations we could equally as well define

$$d_m(x, y) := \inf \{\mathbb{L}_m(\gamma) \mid \gamma \in \mathcal{P}_{xy}\}.$$

In fact, we can even restrict to smooth curves, as the next lemma shows.

Lemma 52.2. *Let $\gamma \in \mathcal{P}([a, b])$. Then there exists an increasing diffeomorphism $h\colon [a, b] \to [a, b]$ such that $\gamma \circ h \in \mathcal{C}([a, b])$.*

Proof. Let $a = a < a_1 < \cdots < a_k = b$ denote a subdivision such that $\gamma|_{[a_{i-1}, a_i]}$ is smooth for each $i = 1, \dots, k$. Let $f_i\colon [a, b] \to \mathbb{R}$ denote a smooth increasing function[1] such that

$$f_i(t) = \begin{cases} 0, & a \le t \le a_{i-1}, \\ 1, & a_i \le t \le b. \end{cases}$$

Then set

$$h(t) := a + \sum_{i=1}^{k} (a_i - a_{i-1}) f_i(t).$$

This is a smooth strictly increasing function such that $h(a_i) = a_i$ and such that $h'(a_i) = h''(a_i) = h'''(a_i) = \cdots = 0$ for each $i = 0, \dots, k$. Thus $\gamma \circ h$ is smooth at each point a_i, and hence smooth everywhere. ∎

Corollary 52.3. *One has*

$$d_m(x, y) = \inf \{\mathbb{L}_m(\gamma) \mid \gamma \in \mathcal{C}_{xy}([0, 1])\}$$

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] The existence of such a smooth function was proved in Step 1 of Lemma 3.11.

Here is today's first main result.

THEOREM 52.4. *Let (M, m) be a connected Riemannian manifold. Then the function*

$$d_m \colon M \times M \to [0, \infty)$$

is a metric in the sense of point-set topology. Moreover the topology that d_m induces on M is the same as the original topology on M.

The key to proving Theorem 52.4 is to first consider the case where M is a vector space and the metric comes from an inner product. This is completely trivial, but for completeness we spell out all the details.

So suppose V is a vector space and $\langle \cdot, \cdot \rangle$ is an inner product on V. We can define a Riemannian metric $m_V = \langle \cdot, \cdot \rangle_V$ on V by declaring that

$$\langle \mathcal{J}_x(v), \mathcal{J}_x(w) \rangle_V := \langle v, w \rangle, \qquad \forall\, x, v, w \in V.$$

We have already met two instances of this construction:

- The Euclidean metric m_{Eucl} on $\mathbb{R}^n$ is the special case where $V = \mathbb{R}^n$ and $\langle \cdot, \cdot \rangle$ is the standard Euclidean dot product.
- If (M, m) is a Riemannian manifold then the Riemannian structure $m_{T_x M}$ on the tangent space from Definition 50.16 corresponds to the case $V = T_x M$ and $\langle \cdot, \cdot \rangle = m|_x$.

Suppose $\eta \colon [a, b] \to V$ is a smooth curve. Let us temporarily write $\dot\eta$ for the derivative of η in the sense of multivariate calculus. Thus $\dot\eta \colon [a, b] \to V$ is another smooth curve in V and[2] one has

$$\eta' = \mathcal{J}_\eta(\dot\eta).$$

Set

$$r(t) := |\eta(t)|. \tag{52.1}$$

If $r(t) \neq 0$ then we can uniquely write

$$\eta(t) = r(t)e(t)$$

where $e \colon [a, b] \to V$ is a smooth curve such that

$$|e(t)| = 1.$$

The Leibniz rule gives us

$$\begin{aligned}
\eta'(t) &= \mathcal{J}_{\eta(t)}(\dot\eta(t)) \\
&= \mathcal{J}_{\eta(t)}(\dot r(t)e(t) + r(t)\dot e(t)) \\
&= \dot r(t)\mathcal{J}_{\eta(t)}(e(t)) + r(t)\mathcal{J}_{\eta(t)}(\dot e(t)) \\
&\overset{\text{def}}{=} \eta'_{\text{rad}}(t) + \eta'_{\text{pol}}(t).
\end{aligned}$$

[2] I really should have thought of this notation earlier... It would have made things far less confusing at the beginning of Differential Geometry I. Oh well. Note to self: Implement this change before I lecture the course again.

As the notation suggests, we call $\eta'_{\mathrm{rad}}(t)$ the **radial component** of $\eta'(t)$ and $\eta'_{\mathrm{pol}}(t)$ the **polar component** of $\eta'(t)$. Note that

$$|\eta'_{\mathrm{rad}}(t)|_V^2 = |\dot{r}(t)|^2 \left\langle \mathcal{J}_{\eta(t)}(e(t)), \mathcal{J}_{\eta(t)}(e(t)) \right\rangle_V = |\dot{r}|^2 \langle e(t), e(t) \rangle = |\dot{r}(t)|^2. \tag{52.2}$$

Moreover

$$\begin{aligned} \left\langle \eta'_{\mathrm{rad}}(t), \eta'_{\mathrm{pol}}(t) \right\rangle &= r(t)\dot{r}(t) \left\langle \mathcal{J}_{\eta(t)}(e(t)), \mathcal{J}_{\eta(t)}(\dot{e}(t)) \right\rangle \\ &= r(t)\dot{r}(t) \langle e(t), \dot{e}(t) \rangle . \end{aligned}$$

But

$$0 = \frac{d}{dt} \langle e(t), e(t) \rangle = 2 \langle e(t), \dot{e}(t) \rangle ,$$

and hence we see that the radial component is orthogonal to the polar component:

$$\eta'_{\mathrm{rad}}(t) \perp \eta'_{\mathrm{pol}}(t),$$

and hence

$$|\eta'(t)|_V^2 = |\eta'_{\mathrm{rad}}(t)|_V^2 + |\eta'_{\mathrm{pol}}(t)|_V^2.$$

Thus from (52.2) we see that

$$|\eta'(t)|_V \geq |\dot{r}(t)|, \qquad \text{with equality if and only if} \qquad \dot{e}(t) = 0. \tag{52.3}$$

The next lemma proves that such a ray is a length-minimising curve in the Riemannian manifold (V, m).

LEMMA 52.5. *One has*

$$d_m(0, v) = |v|.$$

Proof. We may assume that $v \neq 0$. Set $\eta(t) := tv$. Then for $t > 0$ the radial component is well defined, and in fact in this case one has

$$\eta'(t) = \mathcal{J}_{tv}(v) = \eta'_{\mathrm{rad}}(t),$$

and so

$$|\eta'(t)|_V = |v|.$$

Thus

$$\mathbb{L}_m(\eta) = \int_0^1 |\eta'(t)|_V \, dt = \int_0^1 |v| \, dt = |v|.$$

Thus $d_m(0, v) \leq |v|$. Now suppose ζ is any curve in $\mathcal{C}_{0v}([0,1])$. Suppose to begin with that $\zeta(t) \neq 0$ for $t > 0$. Then (52.3) shows that $\mathbb{L}_m(\zeta) \geq |v|$. For the general case if

$$a := \sup \{ t > 0 \mid \zeta(t) = 0 \} ,$$

then the argument above shows that

$$\mathbb{L}_m(\zeta) \geq \mathbb{L}_m\big(\zeta|_{[a,1]}\big) \geq |v|,$$

where the first equality used part (ii) of Lemma 51.6. This shows that $d_m(0, v) \geq |v|$, and hence $d_m(0, v) = |v|$. This completes the proof. ∎

Corollary 52.6. *Theorem 52.4 is true for* (V, m_V).

Now as in Definition 50.16, let us take $V = T_xM$, equipped with the inner product m_x. We have just seen that length minimising in T_xM passing through 0_x are straight lines. We now show that this property is preserved under the exponential map.

Proposition 52.7. *Suppose* $v \in \mathcal{S}_x$. *Let* $\eta_v \colon [0,1] \to T_xM$ *denote the curve* $\eta_v(t) := tv$. *Suppose* $\eta \colon [0,1] \to \mathcal{S}_x$ *is any other piecewise smooth curve such that* $\eta(0) = 0_x$ *and* $\eta(1) = v$. *Then*

$$\mathbb{L}_m(\exp_x \circ \eta) \geq \mathbb{L}_m(\exp_x \circ \eta_v).$$

Moreover this equality is strict if there exists some $t \in [0,1]$ *such that*

$$D\exp_x(\eta(t))\left[\eta'_{\mathrm{pol}}(t)\right] \neq 0. \tag{52.4}$$

Proof. By Lemma 52.2 we may assume that η is smooth. Moreover we may assume that $v \neq 0$ and that $\eta(t) \neq 0$ for $t > 0$. By the Gauss Lemma (Theorem 50.17) we have

$$\begin{aligned}
|D\exp_x(\eta(t))[\eta'(t)]|^2 &\overset{(50.11)}{=} |D\exp_x(\eta(t))[\eta'_{\mathrm{rad}}(t)]|^2 + |D\exp_x(\eta(t))[\eta'_{\mathrm{pol}}(t)]|^2 \\
&\overset{(\dagger)}{\geq} |D\exp_x(\eta(t))[\eta'_{\mathrm{rad}}(t)]|^2 \\
&\overset{(50.10)}{=} |\eta'_{\mathrm{rad}}(t)|^2_{T_xM} \\
&= |\dot{r}(t)|^2,
\end{aligned}$$

where $r(t) = |\eta(t)|$ is defined as in (52.1) and the last line used (52.2). Next, note that arguing as in the computation of (52.1) we have

$$\frac{d}{dt}|\eta(t)| = \frac{\langle \eta(t), \mathcal{J}_{\eta(t)}(\dot{\eta}(t))\rangle}{|\eta(t)|} = |\dot{r}(t)|.$$

Thus

$$\begin{aligned}
\mathbb{L}_m(\exp_x \circ \eta) &= \int_0^1 |D\exp_x(\eta(t))[\eta'(t)]|^2 \, dt \\
&\geq \int_0^1 |\dot{r}(t)| \, dt \\
&= \int_0^1 \frac{d}{dt}|\eta(t)| \, dt \\
&= |\eta(1)| \\
&= |v| \\
&= \mathbb{L}_m(\exp_x \circ \eta_v)
\end{aligned}$$

Finally, the last assertion is clear, as the only inequality was (†) in the equations above, and (52.4) is exactly the condition for this inequality to be strict. ∎

Given $x \in M$ and $\varepsilon > 0$ we denote by $O(x, \varepsilon)$ the **open** ball about 0_x in T_xM, measured with respect to the metric $m|_x$:

$$O(x, \varepsilon) \coloneqq \{v \in T_xM \mid |v| < \varepsilon\}. \tag{52.5}$$

We let $\overline{O}(x, \varepsilon)$ denote the closure of $O(x, \varepsilon)$:

$$\overline{O}(x, \varepsilon) \coloneqq \{v \in T_xM \mid |v| \leq \varepsilon\}.$$

DEFINITION 52.8. Fix $x \in M$. We denote by $s_m(x)$ the maximal radii of such a ball that lies in the domain of the exponential map:

$$s_m(x) \coloneqq \sup \{\varepsilon > 0 \mid O(x, \varepsilon) \subset \mathcal{S}_x\}.$$

The **injectivity radius of** $x \in M$ is the defined to be

$$\mathrm{inj}_m(x) \coloneqq \sup \left\{\varepsilon > 0 \mid \exp_x|_{O(x,\varepsilon)} \text{ is a diffeomorphism onto its image}\right\}.$$

Thus part (ii) of Theorem 43.3 tells us that

$$0 < \mathrm{inj}_m(x) \leq s_m(x).$$

REMARK 52.9. If m is a complete metric (which by the Hopf-Rinow Theorem 53.7 is equivalent to asking that (M, d_m) is a complete metric space) then one can show that

$$\mathrm{inj}_m(x) = \sup \left\{\varepsilon > 0 \mid \exp_x|_{O(x,\varepsilon)} \text{ is injective}\right\}.$$

This is the reason for the name "injectivity radius". Nevertheless, we will not need or use this fact.

Proposition 52.7 has the following consequence.

PROPOSITION 52.10. *Fix $x \in M$ and choose $0 < \varepsilon < \mathrm{inj}_m(x)$. Fix $v \in O(x, \varepsilon)$, and let $\gamma_v \coloneqq \exp_x(tv)$. Set $y \coloneqq \gamma_v(1)$. If $\gamma \in \mathcal{P}_{xy}([0, 1])$ then*

$$\mathbb{L}_m(\gamma) \geq \mathbb{L}_m(\gamma_v),$$

and the inequality is strict unless γ is a reparametrisation of γ_v.

Proof. Suppose γ has image in $\exp_x(O(x, \varepsilon))$. Then there exists a unique piecewise smooth curve $\eta \colon [0, 1] \to T_xM$ such that $\exp_x \circ \eta = \gamma$. By Proposition 52.7 in this case we have $\mathbb{L}_m(\gamma) \geq \mathbb{L}_m(\gamma_v)$. If equality holds then by Proposition 52.7 again we have

$$D \exp_x(\eta(t))\big[\eta'_{\mathrm{pol}}(t)\big] = 0, \qquad \forall\, t \in [0, 1],$$

and thus the same argument as in the proof of Proposition 52.7 tells us that for any $a \leq t_0 \leq t_1 \leq b$ one has

$$\mathbb{L}_m\big(\gamma|_{[t_0,t_1]}\big) = |\eta(t_1) - \eta(t_0)|.$$

Suppose for contradiction that γ is not a reparametrisation of γ_v. Then there exists $t_0 \in [0,1]$ such that $\eta(t_0) \notin \{tv \mid 0 \le t \le 1\}$. Then by part (ii) of Lemma 51.6 and the proof of Proposition 52.7 we have

$$\begin{aligned} \mathbb{L}_m(\gamma) &= \mathbb{L}_m(\gamma|_{[0,t_0]}) + \mathbb{L}_m(\gamma|_{[t_0,1]}) \\ &\ge |\eta(t_0)| + |v - \eta(t_0)| \\ &> |v| \\ &= \mathbb{L}_m(\gamma_v), \end{aligned}$$

which contradicts the assumption $\mathbb{L}_m(\gamma) = \mathbb{L}_m(\gamma_v)$. This deals with the case where γ has image inside $\exp_x(O(x,\varepsilon))$. If this is not the case, let

$$a := \sup\{t \mid \gamma([0,t]) \subset \exp_x(O(x,\varepsilon))\}.$$

Since $v \in O(x,\varepsilon)$, there exists some $0 < b < a$ such that $w := \big(\exp_x|_{O(x,\varepsilon)}\big)^{-1}(\gamma(b))$ satisfies $|w| > |v|$. Then using Proposition 52.7 again we have

$$\mathbb{L}_m(\gamma) \ge \mathbb{L}_m(\gamma|_{[0,b]}) \ge |w| > |v| = \mathbb{L}_m(\gamma_v).$$

This completes the proof. ■

Remark 52.11. It follows from Proposition 52.7 that every geodesic is locally length-minimising. We will explore this in more detail next lecture.

With these preparations out of the way we can prove Theorem 52.4 in the general case. Given $x \in M$ we denote by $U(x,\varepsilon)$ the set

$$U(x,\varepsilon) := \{y \in M \mid d_m(x,y) < \varepsilon\} \tag{52.6}$$

and by $\overline{U}(x,\varepsilon)$ the set

$$\overline{U}(x,\varepsilon) := \{y \in M \mid d_m(x,y) \le \varepsilon\}$$

Once the proof of Theorem 52.4 is complete, $U(x,\varepsilon)$ will be the open ball of radius ε about x in the d_m-metric, and $\overline{U}(x,\varepsilon)$ will be its closure in M. For now however, these are just sets that will prove useful in the proof of Theorem 52.4.

Proof of Theorem 52.4. We prove the result in three steps.

1. In this step we show that d_m is a metric. Since $\mathbb{L}_m(\gamma) = \mathbb{L}_m(\gamma^-)$ for any curve γ by part (iii) of Lemma 51.6 it is clear that d_m is symmetric. Moreover the triangle inequality is immediate from part (ii) of the same lemma. It remains to show that

$$d_m(x,y) = 0 \qquad \Rightarrow \qquad x = y.$$

Choose $0 < \varepsilon < \mathrm{inj}_m(x)$. Then we must have $y \in \exp_x(O(x,\varepsilon))$, since any point z not contained in $\exp_x(O(x,\varepsilon))$ has $d_m(x,z) \ge \varepsilon$ by Proposition 52.10. But also from the proof of Proposition 52.10 we have

$$d_m(x,z) = \left|\big(\exp_x|_{O(x,\varepsilon)}\big)^{-1}(z)\right|, \qquad \forall z \in \exp_x(O(x,\varepsilon)). \tag{52.7}$$

Thus if $d_m(x,y) = 0$ then $y = x$.

2. Thus (M, d_m) is a metric space. It remains to show that the topology d_m induces on M is the same as the original topology. In this step we show that d_m is continuous as a function $M \times M \to \mathbb{R}$. Indeed, if $x_h \to x$ then $d_m(x, x_h) \to 0$ by (52.7) and the fact that $z \mapsto \left|\left(\exp_x|_{O(x,\varepsilon)}\right)^{-1}(z)\right|$ is continuous. Then if $(x_h, y_h) \to (x, y)$ then by the triangle inequality we have

$$\begin{aligned} d_m(x,y) - d_m(x,x_h) - d_m(y,y_h) &\leq d_m(x_h,y_h) \\ &\leq d_m(x_h,x) + d_m(x,y) + d_m(y,y_h) \end{aligned}$$

and thus letting $h \to \infty$ we see that $d_m(x_h, y_h) \to d_m(x, y)$.

3. To show that d_m induces the same topology on M as we began with we must show that the balls[3] $U(x, \varepsilon)$ form a basis for the topology of M. Since d_m is continuous each such ball $U(x, \varepsilon)$ is open in M (with respect to the original topology on M). Thus we need to show that for any $x \in M$ and any neighbourhood U of x (with respect to the original topology on M) there exists $\varepsilon > 0$ such that $U(x, \varepsilon) \subset U$. In fact we will prove that for $0 < \varepsilon < \mathrm{inj}_m(x)$ one has

$$U(x,\varepsilon) = \exp_x(O(x,\varepsilon)), \tag{52.8}$$

which clearly implies the claim. To prove (52.8), first note that (52.7) implies that $\exp_x(O(x,\varepsilon)) \subset U(x,\varepsilon)$. Since $\exp_x(O(x,\varepsilon))$ is d_m-open in $U(x,\varepsilon)$, which is connected, it suffices to show that $\exp_x(O(x,\varepsilon))$ is d_m-closed in $U(x,\varepsilon)$. So suppose $y_h \in \exp_x(O(x,\varepsilon))$ with $y_h \to y$. Set

$$v_h := \left(\exp_x|_{O(x,\varepsilon)}\right)^{-1}(y_h).$$

Then (v_h) is a bounded sequence in T_xM, and hence by passing to a convergent subsequence if necessary, we may assume $v_h \to v$. Then

$$\begin{aligned} |v| &= \lim_{h\to\infty} |v_h| \\ &= \lim_{h\to\infty} d_m(x,y_h) \\ &= d_m(x,y) \\ &< \varepsilon \end{aligned}$$

by Step 2. Since $\exp_x$ is continuous, we also have $y = \exp_x(v)$, and thus $y \in U(x,\varepsilon)$. This completes the proof. ∎

Since it will be useful later let us state (52.8) again as a corollary.

Corollary 52.12. *Let $x \in M$ and $0 < \varepsilon < \mathrm{inj}_m(x)$. Then*

$$\exp_x(\overline{O}(x,\varepsilon)) = \overline{U}(x,\varepsilon)$$

and $\exp_x|_{O(x,\varepsilon)}$ is a diffeomorphism

$$\exp_x|_{O(x,\varepsilon)} \colon O(x,\varepsilon) \to U(x,\varepsilon).$$

[3] Since we have now proved d_m is a metric, it is okay to call them "balls"!

We will examine further properties of the the metric space (M, d_m) next lecture. But let us clear up a potential ambiguity in the terminology. If (X, d_X) and (Y, d_Y) are two metric spaces then an **isometry** $f: X \to Y$ is a continuous bijection such that $d_X(x, y) = d_Y(f(x), f(y))$ for all $x, y \in X$. Any isometry is necessarily a homeomorphism, and in fact with a bit of work one can show that a continuous map $f: X \to Y$ is an isometry if and only if f is surjective and distance-preserving.

Thus if (M, m_1) and (N, m_2) are two Riemannian manifolds, we now have two possible definitions of the word "isometry":

(i) A diffeomorphism $\varphi: M \to N$ such that $\varphi^\star(m_2) = m_1$ (Definition 46.1),

(ii) A continuous bijection $\varphi: M \to N$ which is an metric space isometry (with respect to the metric d_{m_1} on M and the metric d_{m_2} on N).

It is easy to see that an isometry in the sense of (i) is also an isometry in the sense of (ii). In fact the converse holds too, as the next theorem states.

THEOREM 52.13 (Myers-Steenrod). *An isometry in the sense of (ii) is also an isometry in the sense of (i) and hence the two definitions coincide.*

This theorem is surprising, since a priori there is no reason why an isometry in the sense of (ii) should even be differentiable! Sadly we will not have time to prove Theorem 52.13.

REMARK 52.14. Theorem 52.13 is one of two results in Riemannian Geometry often called the "Myers-Steenrod Theorem". The other was stated in Remark 46.3. They are both difficult, but Theorem 52.13 is the easier of the pair.

We can also define the injectivity radius of a set.

DEFINITION 52.15. If $A \subset M$ is an arbitrary subset we define the **injectivity radius of** A as

$$\mathrm{inj}_m(A) := \inf \{\mathrm{inj}_m(x) \mid x \in A\}.$$

In general it is possible for the injectivity radius of a subset to be zero. Nevertheless, the following holds.

PROPOSITION 52.16. *If $A \subset M$ is compact then $\mathrm{inj}_m(A) > 0$. In particular, if M is compact then $\mathrm{inj}_m(M) > 0$.*

Proof. By part (iii) of Theorem 43.3, there exists a neighbourhood V of the zero section $o(M)$ with the property that $(\pi, \exp)|_V$ is an embedding[4]. Since A is compact, there exist finitely many points $x_1, \dots, x_k$ and $\varepsilon_1, \dots, \varepsilon_k > 0$ such that

$$A \subset \bigcup_{i=1}^k U(x_i, \varepsilon_i),$$

and

$$U(x_i, 3\varepsilon_i) \times U(x_i, 3\varepsilon_i) \subset (\pi, \exp)(V), \qquad i = 1, \dots, k.$$

If $y \in A$ then there exists i such that $y \in U(x_i, \varepsilon_i)$. Moreover $\exp_y$ is an embedding on $O(x_i, 3\varepsilon_i)$. Thus $\mathrm{inj}_m(y) \geq 2\varepsilon_i$, and hence $\mathrm{inj}_m(A) \geq \min_{1 \leq i \leq k} 2\varepsilon_i > 0$. This completes the proof. ∎

[4] As mentioned in Remark 43.4, this is the only time in the course where we use the full strength of Theorem 43.3 (i.e. part (iii)) rather than the weaker statement given in part (ii).

The Hopf-Rinow Theorem and its friends

We conclude the course by proving three important foundational results in Riemannian Geometry, each of which is named after a pair of mathematicians:

- The Hopf-Rinow Theorem 53.7,
- The Cartan-Hadamard Theorem 53.14,
- The Bonnet-Myers Theorem 53.15.

Throughout this lecture, (M, m) is a connected Riemannian manifold of dimension n, ∇ denotes the Levi-Civita connection of m, and $\exp\colon \mathcal{S} \to M$ denotes the exponential map of m. We let $O(x, \varepsilon)$ and $U(x, \varepsilon)$ be defined as they were in (52.5) and (52.6) respectively. Recall that $s_m(x)$ is defined to be the maximal ε such that $O(x, \varepsilon)$ is contained in the domain of $\exp_x$, and $\mathrm{inj}_m(x)$ is the maximal ε such that $\exp_x|_{O(x,\varepsilon)}$ is a diffeomorphism.

DEFINITION 53.1. A piecewise smooth curve $\gamma\colon [a, b] \to M$ is said to be **minimal** if

$$d_m(\gamma(a), \gamma(b)) = \mathbb{L}_m(\gamma).$$

Minimal curves are geodesics.

LEMMA 53.2. *Let $\gamma\colon [a, b] \to M$ be minimal. Then γ is (up to reparametrisation) a geodesic.*

Proof. By Proposition 52.7, any minimal curve is locally (and hence also globally) a reparametrisation of a geodesic. ∎

REMARK 53.3. Alternatively one can argue using the analogue of Corollary 51.14 for $\mathbb{L}_m$ instead of $\mathbb{E}_m$. This tells us that[1] "critical points" of $\mathbb{L}_m$ are (up to reparametrisation) geodesics. A length-minimising curve is a special type of critical point (a local minimum), and thus Lemma 53.2 follows.

Moreover the proof of Proposition 52.7 shows that any geodesic is **locally minimal**, in the sense that if $\gamma\colon [a, b] \to M$ is a geodesic and $t \in (a, b)$ then there exists $\varepsilon > 0$ such that $\gamma|_{[t-\varepsilon,t+\varepsilon]}$ is minimal. In general however γ need not be length-minimising on its entire domain.

EXAMPLE 53.4. Consider $(S^n, m_{\mathrm{round}})$. The geodesics are the great circles by part (ii) of Problem O.2. Thus all geodesics are defined on all of $\mathbb{R}$. A geodesic is length-minimising until one reaches the antipodal point—after this, it is no longer length-minimising, since following the great circle in the opposite direction gives a shorter curve. The similarity of this example with Example 51.22 is not a coincidence (see Proposition 53.10 below).

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1]The quotation marks are there to remind you that $\mathbb{L}_m$ is not actually a differentiable function on the space of paths, cf. Remark 51.25.

The next technical result is the main step in the proof of the Hopf-Rinow Theorem 53.7

Proposition 53.5. *If $0 < \varepsilon < s_m(x)$ then for any $y \in U(x, \varepsilon)$ there exists a minimal geodesic γ joining x and y. Thus $\exp_x$ defines a surjective map from $O(x, \varepsilon)$ to $U(x, \varepsilon)$ for all $0 < \varepsilon < s_m(x)$.*

Note that in Proposition 53.5 the only hypothesis is that $\varepsilon < s_m(x)$. We are *not* assuming ε is less than the injectivity radius $\mathrm{inj}_m(x)$ (in which case the conclusion would be immediate from Corollary 52.12). Moreover the proposition is only asserting the existence of *some* minimal geodesic γ from x to y. This geodesic will typically not be entirely contained in $U(x, \varepsilon)$. The following proof is non-examinable—not because it is particularly hard, but rather because it is somewhat finnicky[2].

(♣) *Proof.* Given $0 < \delta < \varepsilon$, let

$$C(x, \delta) \coloneqq \left\{ y \in \overline{U}(x, \delta) \mid \text{ there exists a minimal geodesic joining } x \text{ and } y \right\}.$$

We will prove in three steps that one has

$$C(x, \delta) = \overline{U}(x, \delta), \tag{53.1}$$

from which the result clearly follows.

1. In this first step we show that $C(x, \delta)$ is a compact subset of M. Since certainly

$$C(x, \delta) \subset \exp_x(\overline{O}(x, \varepsilon)),$$

and $\exp_x(\overline{O}(x, \varepsilon))$ is compact, it is enough to show that $C(x, \delta)$ is closed. So suppose (y_h) is a sequence of points in $C(x, \delta)$ such that $y_h \to y$. Let $v_h \in \overline{O}(x, \delta)$ denote a vector such that $|v_h| = d_m(x, y_h)$ and such that $\exp_x(v_h) = y_h$. Since (v_h) is a bounded sequence, we may assume it converges to some $v \in \overline{O}(x, \varepsilon)$. Then as $\exp_x$ is continuous we have $y = \exp_x(v)$ and since d_m is continuous, $d_m(x, y) = |v|$. Thus $y \in C(x, \delta)$.

2. Let

$$I \coloneqq \left\{ \delta \in (0, \varepsilon) \mid C(x, \delta) = \overline{U}(x, \delta) \right\}.$$

Then I is non-empty, since $(0, \mathrm{inj}_m(x)) \subset I$ from the proof of Theorem 52.4. In this step we will show that I is also closed in $(0, \varepsilon)$. For this suppose $\delta_0 \in (0, \varepsilon)$ has the property that $\delta \in I$ for all $\delta < \delta_0$. Thus $C(x, \delta) = \overline{U}(x, \delta)$ for all $\delta < \delta_0$ and hence $U(x, \delta_0) \subseteq C(x, \delta_0)$. Since $C(x, \delta_0)$ is compact by Step 1, we have $\overline{U}(x, \delta_0) \subseteq C(x, \delta_0)$. The reverse inclusion always holds, and hence this establishes $\overline{U}(x, \delta_0) = C(x, \delta_0)$. Thus $\delta_0 \in I$, and so I is closed as desired.

3. In this final step we prove that I is also open in $(0, \varepsilon)$, from which it follows that $I = (0, \varepsilon)$ and (53.1) is proved. Fix $\delta_1 \in I$. Now let λ denote any number such that

$$0 < \lambda < \min \left\{ \varepsilon - \delta_1, \mathrm{inj}_m(C(x, \delta_1)) \right\}. \tag{53.2}$$

Here we are using Proposition 52.16 (together with Step 1) to guarantee that the right-hand side of (53.2) is positive. We shall show that $\delta_1 + \lambda \in I$. For this it

[2] And also because I skipped it in class.

suffices to show that $U(x, \delta_1 + \lambda) \subset C(x, \delta_1 + \lambda)$ as $C(x, \delta_1 + \lambda)$ is closed. Since $\delta_1 \in I$ we have $\overline{U}(x, \delta_1) = C(x, \delta_1) \subset C(x, \delta_1 + \lambda)$, and thus it suffices to show that

$$U(x, \delta_1 + \lambda) \setminus \overline{U}(x, \delta_1) \subset C(x, \delta_1 + \lambda). \tag{53.3}$$

Suppose y belongs to the left-hand side of (53.3). By definition of d_m as an infimum, we can find a sequence $(\gamma_h) \subset \mathcal{P}_{xy}([0, 1])$ such that

$$\mathbb{L}_m(\gamma_h) < d_m(x, y) + \frac{1}{h}.$$

By the intermediate value theorem there exists $t_h \in [0, 1]$ such that $d(x, \gamma_h(t_h)) = \delta_1$. Set $z_h := \gamma_h(t_h)$. Since $\delta_1 \in I$, the set $\overline{U}(x, \delta_1)$ is compact, and hence up to passing to a subsequence we may assume that $z_h \to z$. By continuity of d_m, we have $d_m(x, z) = \delta_1$. We will now prove that this point z has the special property that

$$d_m(x, y) = d_m(x, z) + d_m(z, y). \tag{53.4}$$

Since γ_h is a curve from x to y that passes through z_h, we have

$$d_m(x, z_h) + d_m(z_h, y) \le \mathbb{L}_m(\gamma_h) < d_m(x, y) + \frac{1}{h}. \tag{53.5}$$

Next, for each $h \in \mathbb{N}$ there exists $l(h) > h$ such that $d_m(z_{l(h)}, z) < \frac{1}{h}$. Then

$$\begin{aligned} d_m(x, z) + d_m(z, y) &\le d_m(x, z_{l(h)}) + d_m(z_{l(h)}, z) + d_m(z, z_{l(h)}) + d_m(z_{l(h)}, y) \\ &\le d_m(x, z_{l(h)}) + \frac{2}{h} + d_m(z_{l(h)}, y) \\ &\le d_m(x, y) + \frac{3}{h}, \end{aligned}$$

where the last line used (53.5) with $h = l(h)$. Since h was arbitrary we conclude

$$d_m(x, z) + d_m(z, y) \le d_m(x, y).$$

But $d_m(x, y) \le d_m(x, z) + d_m(z, y)$ by the triangle inequality for d_m, and this proves (53.4).

Why does this help? Well, since $z \in C(x, \delta_1)$ there exists a minimal geodesic $\tilde{\gamma}_0$ from x to z. Moreover since $d_m(z, y) < \operatorname{inj}_m(C(x, \delta_1))$ there exists a minimal geodesic $\tilde{\gamma}_1$ from z to y. The composition $\tilde{\gamma}_0 * \tilde{\gamma}_1$ is a piecewise smooth curve from x to y with

$$\begin{aligned} \mathbb{L}_m(\tilde{\gamma}_0 * \tilde{\gamma}_1) &= \mathbb{L}_m(\tilde{\gamma}_0) + \mathbb{L}_m(\tilde{\gamma}_1) \\ &= d_m(x, z) + d_m(z, y) \\ &= d_m(x, y) \end{aligned}$$

by (53.4). Thus by Lemma 53.2, $\tilde{\gamma}_0 * \tilde{\gamma}_1$ is (up to reparametrisation) a minimal geodesic from x to y. Thus $\delta_1 + \lambda \in I$, and hence I is open. This finally completes the proof. ■

Let us also note the following corollary from the proof of Proposition 53.5.

Corollary 53.6. *If $0 < \varepsilon < s_m(x)$ then $\overline{U}(x, \varepsilon)$ is compact in M.*

Recall a metric space is *complete* if every Cauchy sequence converges. This leads to following potential clash of terminology for what it means for a Riemannian manifold (M, m) to be complete: Does it mean that m is a complete metric, or that (M, d_m) is a complete metric space? Luckily, just as with the Myers-Steenrod Theorem 52.13, it turns out these two concepts coincide. This statement is usually referred to as the Hopf-Rinow Theorem, and we prove it next.

Theorem 53.7 (The Hopf-Rinow Theorem). *Let (M, m) be a connected Riemannian manifold. The following are equivalent.*

(i) *The Riemannian metric m is complete, i.e. the domain $\mathcal{S}$ of the exponential map is the entire tangent bundle TM, and hence $s_m(x) = \infty$ for all $x \in M$.*

(ii) *The Riemannian metric m is "complete at a single point", i.e. there exists a single point x such that $\mathcal{S}_x = T_xM$.*

(iii) *(M, d_m) is a complete metric space.*

(iv) *Any d_m-bounded set $A \subset M$ has compact closure.*

Proof. Firstly, (iv) $\Rightarrow$ (iii) is true of any metric space. Indeed, any Cauchy sequence is bounded, and so by (iv) it is contained in a compact set. It therefore admits a convergent subsequence, and hence it converges[3]. It is obvious that (i) $\Rightarrow$ (ii). Next, (ii) $\Rightarrow$ (iv) by Corollary 53.6, since if A is bounded then if $y \in A$ and

$$r := d_m(x, y) + \sup\{d_m(y, z) \mid z \in A\} + 1$$

then $A \subset U(x, r)$. It remains to show that (iii) $\Rightarrow$ (i). For this, let $\mathbb{S}$ denote the geodesic spray of m. Fix $(x, v) \in TM$ and let δ denote the maximal integral curve of $\mathbb{S}$ with $\delta(0) = (x, v)$. Thus

$$\delta(t) = (\gamma(t), \gamma'(t)),$$

where γ is the unique geodesic with initial conditions $\gamma(0) = x$ and $\gamma'(0) = v$, cf. Theorem 42.14. Let (t^-, t^+) denote the maximal domain of δ. We must show that $t^- = -\infty$ and $t^+ = \infty$. We do the latter case only, as the former is analogous. Assume for contradiction that $t^+ < \infty$, and choose $0 < t_h < t^+$ such that $t_h \to t^+$. Let $y_h := \pi(\delta(t_h))$. Then

$$d_m(y_h, y_l) = d_m(\exp(t_h v), \exp(t_l v)) \le |t_h - t_l||v|,$$

and thus (y_h) is a Cauchy sequence in M. By (iii), y_h converges to some point y. Moreover since

$$|\delta(t_h)| = |\gamma'(t_h)| = |\gamma'(0)| = v$$

(as geodesics have constant speed, cf. Remark 51.3), the sequence $(\delta(t_h))$ is contained in the set $\{w \in TM \mid |w| = |v|\}$. Let C denote any compact neighbourhood of y. Then for h sufficiently large,

$$\delta(t_h) \in \{w \in TM \mid \pi(w) \in C \text{ and } |w| = |v|\}.$$

[3] Recall if a subsequence of a Cauchy sequence converges then the entire sequence must converge.

The set on the right-hand side is compact. Thus $(\delta(t_h))$ admits a convergent subsequence. Now the argument from the proof of Lemma 8.21 shows that δ is actually defined on $(t^-, t^+ + \varepsilon)$ for some small $\varepsilon > 0$, which contradicts the maximality of t_0. This completes the proof. ■

From now on we can therefore use unambiguously use the terminology "complete Riemannian manifold" to mean any of the conditions in the Hopf-Rinow Theorem. Here are some easy corollaries.

COROLLARY 53.8. *Let M be a compact manifold. Then every Riemannian metric m on M is complete.*

Proof. Compact metric spaces are always complete. ■

COROLLARY 53.9. *Let (M, m) be a complete Riemannian manifold. Any two points in M can be joined by a (not necessarily unique) minimal geodesic.*

Proof. Immediate from Proposition 53.5. ■

We now relate minimality of geodesics with conjugate points.

PROPOSITION 53.10. *Let $\gamma\colon [0,b] \to M$ be a non-constant geodesic, and suppose there exists a conjugate point $0 < t_0 < b$. Then γ is not minimal.*

Warning: The converse to Proposition 53.10 is not true: there are examples of geodesics with no conjugate points that are not minimal. You are invited to find such an example on Problem Sheet W.

Proof. By assumption there exists a non-zero Jacobi field $c \in \mathrm{Jac}(\gamma)$ such that $c(0) = 0_{\gamma(0)}$ and $c(t_0) = 0_{\gamma(t_0)}$. We will produce from c a new piecewise smooth vector field $\tilde{c} \in T_\gamma \mathcal{P}_{\gamma(0)\gamma(b)}([0,b])$ which has the property that

$$\mathrm{Hess}(\mathbb{E}_m)(\gamma)(\tilde{c}, \tilde{c}) < 0. \tag{53.6}$$

Suppose for the time being we have constructed such a $\tilde{c}$. Let

$$\gamma_s(t) := \exp_{\gamma(t)}(s\tilde{c}(t)),$$

so that $\frac{d^2}{ds^2}\big|_{s=0} \mathbb{E}_m(\gamma_s) < 0$. Then by Lemma 51.8 we have for $s > 0$ sufficiently small that

$$\begin{aligned} \mathbb{L}_m(\gamma) &= \mathbb{L}_m(\gamma_0) \\ &= \sqrt{2b\mathbb{E}_m(\gamma_0)} \\ &> \sqrt{2b\mathbb{E}_m(\gamma_s)} \\ &\geq \mathbb{L}_m(\gamma_s), \end{aligned}$$

and hence γ is not minimal.

It thus remains to construct such a $\tilde{c}$. Since c is not identically zero, by the uniqueness part of Proposition 50.8 we must have

$$v := \nabla_T(c)(t_0) \neq 0.$$

Let e denote a parallel vector field along γ with $e(t_0) = -v$. Let $f\colon [0,b] \to \mathbb{R}$ denote a smooth function such that $f(0) = f(b) = 0$ and $f(t_0) = f'(t_0) = 1$. Then given $\varepsilon > 0$ set

$$\tilde{c}(t) := \begin{cases} c(t) + \varepsilon f(t)e(t), & 0 \le t \le t_0, \\ \varepsilon f(t)e(t), & t_0 \le t \le b. \end{cases}$$

We claim that for ε sufficiently small this choice of $\tilde{c}$ satisfies (53.6). Since the Hessian is bilinear and c lies in its null-space by Proposition 51.18, it follows from (51.7) that

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(\tilde{c},\tilde{c}) = \varepsilon^2 \operatorname{Hess}(\mathbb{E}_m)(\gamma)(fe, fe) + \left\langle \nabla_T(\tilde{c}|_{[0,t_0]})(t_0) - \nabla_T(\tilde{c}|_{[t_0,b]}), \tilde{c}(t_0) \right\rangle .$$

Let

$$C := \operatorname{Hess}(\mathbb{E}_m)(\gamma)(fe, fe).$$

The precise value of C is not too important. Let us compute the second term. Since e is parallel,

$$\begin{aligned} \nabla_T(\tilde{c}|_{[0,t_0]})(t) &= \nabla_T(c)(t) + \varepsilon f'(t)e(t) + \varepsilon f(t)\nabla_T(e)(t) \\ &= \nabla_T(c)(t) + \varepsilon f'(t)e(t). \end{aligned}$$

Evaluating at $t = t_0$ gives

$$\nabla_T(\tilde{c}|_{[0,t_0]})(t_0) = (1-\varepsilon)v,$$

Similarly

$$\nabla_T(\tilde{c}|_{[t_0,b]})(t_0) = -\varepsilon v,$$

and thus

$$\left\langle \nabla_T(\tilde{c}|_{[0,t_0]})(t_0) - \nabla_T(\tilde{c}|_{[t_0,b]}), \tilde{c}(t_0) \right\rangle = -\varepsilon |v|^2.$$

Hence

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(\tilde{c},\tilde{c}) = \varepsilon^2 C - \varepsilon |v|^2.$$

If $C \le 0$ then certainly $\operatorname{Hess}(\mathbb{E}_m)(\gamma)(\tilde{c},\tilde{c}) < 0$ (for any $\varepsilon > 0$). If instead $C > 0$ then for

$$0 < \varepsilon < \frac{C}{|v|^2}$$

we still have $\operatorname{Hess}(\mathbb{E}_m)(\gamma)(\tilde{c},\tilde{c}) < 0$. This completes the proof. ■

We conclude this course by briefly discussing two **comparison theorems** in Riemannian geometry. As the name suggests, these sort of results *compare* a Riemannian manifold with a simpler one (usually one of constant curvature). This is a vast and fruitful area of research, and we will only scratch the surface. Our particular plan of attack is the following:

- We saw in Example 51.22 that if m has constant curvature then it is easy to see when the first conjugate points appear.
- Proposition 53.10 tells us that a geodesic fails to be minimal after the first conjugate point.

- **Goal:** Suppose (M, m) is a Riemannian manifold satisfying

$$\mathrm{sect}_m(x; \Pi) \leq \kappa, \qquad \forall\, x \in M,\ \forall\, \text{2-planes } \Pi \subset T_x M.$$

  Let $(\tilde{M}, \tilde{m})$ denote a Riemannian manifold with constant curvature

$$\mathrm{sect}_{\tilde{m}} \equiv \kappa.$$

  We aim to relate the appearance of conjugate points in M with the appearance of conjugate points in $\tilde{M}$.

- **Profit:** Conclude something non-trivial about minimality of geodesics in M.

To carry out this program, we introduce the following construction. Let (M, m) and $(\tilde{M}, \tilde{m})$ be two Riemannian manifolds of the same dimension n with associated exponential maps $\exp$ and $\widetilde{\exp}$. Fix $x \in M$ and $\tilde{x} \in \tilde{M}$, and let $v \in T_x M$ and $\tilde{v} \in T_{\tilde{x}} \tilde{M}$ be two vectors of unit norm. Let $T \colon T_x M \to T_{\tilde{x}} \tilde{M}$ be any linear isomorphism such that $Tv = \tilde{v}$ and such that

$$\langle u, w \rangle = \langle Tu, Tw \rangle\,, \qquad \forall\, u, w \in T_x M.$$

Let $\gamma(t) := \exp_x(tv)$ and $\tilde{\gamma}(t) := \widetilde{\exp}_{\tilde{x}}(t\tilde{v})$. Let $b > 0$ be such that both γ and $\tilde{\gamma}$ are defined on $[0, b]$, and set $y := \gamma(b)$ and $\tilde{y} := \tilde{\gamma}(b)$. Finally, let

$$\widehat{\mathbb{P}}^m_t \colon T_x M \to T_{\gamma(t)} M, \qquad \widehat{\mathbb{P}}^{\tilde{m}}_t \colon T_{\tilde{x}} \tilde{M} \to T_{\tilde{\gamma}(t)} \tilde{M}$$

denote the parallel transport isomorphisms along γ and $\tilde{\gamma}$ (with respect to the Levi-Civita connections of m and $\tilde{m}$). For $0 \leq t \leq b$ consider the linear isomorphism

$$T_t \colon T_{\gamma(t)} M \to T_{\tilde{\gamma}(t)} \tilde{M},$$

defined by

$$T_t(v) := \mathbb{P}^{\tilde{m}}_t \circ T \circ (\mathbb{P}^m_t)^{-1}(v).$$

Define also the map

$$\tau \colon \Gamma_\gamma(TM) \to \Gamma_{\tilde{\gamma}}(T\tilde{M}) \tag{53.7}$$

using parallel transport:

$$\tau(c)(t) := T_t(c(t)).$$

The operator τ is an isomorphism. Indeed, as in the proof of Proposition 50.8, we can choose a parallel orthonormal frame $\{\gamma' = e_1, \dots, e_n\}$ along γ. Set $\tilde{e}_i(t) := T_t(e_i(t))$. Then $\{\tilde{\gamma}' = \tilde{e}_1, \dots, \tilde{e}_n\}$ is a parallel orthonormal frame along $\tilde{\gamma}$. If $c \in \Gamma_{\tilde{\gamma}}(TM)$ we can write

$$c = f^i e_i, \qquad \text{where} \qquad f^i := \langle c, e_i \rangle\,,$$

and similarly for elements of $\Gamma_{\tilde{\gamma}}(TM)$. The operator τ is then given by

$$\tau(f^i e_i) = f^i \tilde{e}_i,$$

which is plainly an isomorphism. Note that τ by construction also defines a map

$$\tau \colon T_\gamma \mathcal{P}_{xy}([0, b]) \to T_{\tilde{\gamma}} \mathcal{P}_{\tilde{x}\tilde{y}}([0, b]). \tag{53.8}$$

We then have the following result, whose proof is deferred to Problem Sheet W.

Proposition 53.11. *Suppose that for all $t \in [0,b]$ and for all 2-planes $\Pi \subset T_{\gamma(t)}M$ one has*

$$\mathrm{sect}_m(\gamma(t);\Pi) \le \mathrm{sect}_{\tilde{m}}\big(\tilde{\gamma}(t);T_t[\Pi]\big).$$

Then for all $c \in T_\gamma \mathcal{P}_{xy}([0,b])$ one has

$$\mathrm{Hess}(\mathbb{E}_m)(\gamma)(c,c) \ge \mathrm{Hess}(\mathbb{E}_{\tilde{m}})(\tilde{\gamma})(\tau(c),\tau(c)),$$

where τ is the map (53.8).

From now on let us write

$$\mathrm{sect}_m \le \kappa$$

to indicate that $\mathrm{sect}_m(x;\Pi) \le \kappa$ for all $x \in M$ and all 2-planes $\Pi \subset T_xM$.

Theorem 53.12. *Let (M,m) be a Riemannian manifold and let $\gamma\colon [0,b] \to M$ be a non-constant geodesic. Then:*

(i) *If $\mathrm{sect}_m \le 0$ then γ has no conjugate points.*

(ii) *If $\mathrm{sect}_m \le \kappa$ where $\kappa > 0$, and $\mathbb{L}_m(\gamma) < \frac{\pi}{\sqrt{\kappa}}$, then γ has no conjugate points.*

(iii) *If $\mathrm{sect}_m \ge \kappa > 0$ and $\mathbb{L}_m(\gamma) \ge \frac{\pi}{\sqrt{\kappa}}$, then γ has a conjugate point.*

Proof. Statement (i) follows from (ii) by taking $\kappa > 0$ arbitrarily small. To prove (ii) we apply Proposition 53.11 with $(\tilde{M},\tilde{m}) = (S^n(r), m_{\mathrm{round}})$ for $r = \frac{1}{\sqrt{\kappa}}$. The claim then follows from Example 51.22. Finally the (iii) is proved in the same fashion as (ii), after reversing the roles of M and $\tilde{M}$. ■

Before presenting our final two results, we need one more statement, whose proof is also deferred to Problem Sheet W.

Proposition 53.13. *Let $\varphi\colon (M,m_1) \to (N,m_2)$ be an isometric map between Riemannian manifolds of the same dimension. If (M,m_1) is complete then φ is automatically a Riemannian covering (Definition 46.8).*

The first application of Theorem 53.12 deals with the case where the sectional curvature is non-positive.

Theorem 53.14 (The Cartan-Hadamard Theorem). *Let (M^n,m) be a connected complete Riemannian manifold with $\mathrm{sect}_m \le 0$. Then the universal cover of M is diffeomorphic to $\mathbb{R}^n$.*

Proof. We claim that $\exp_x\colon T_xM \to M$ is a covering map for any $x \in M$. By part (i) of Theorem 53.12, the geodesics of M never have conjugate points, and hence by Proposition 51.23 the map $\exp_x$ has maximal rank everywhere. Thus we can define a Riemannian metric $\tilde{m}$ on T_xM by pulling back m:

$$\tilde{m} := \exp_x^\star(m).$$

Then (by definition) $\exp_x\colon (T_xM,\tilde{m}) \to (M,m)$ is an isometric map between Riemannian manifolds of the same dimension. Let us prove that $\tilde{m}$ is complete. If $v \in T_xM$ the ray $t \mapsto tv$ is mapped onto the geodesic $\gamma(t) = \exp_x(tv)$ in M, and thus is a geodesic in T_xM (cf. Proposition 46.21). Since γ is defined for all $t \in \mathbb{R}$ as M is complete, it follows that all geodesics in T_xM passing through 0_x are defined for all $t \in \mathbb{R}$. Thus $(T_xM,\tilde{m})$ is complete by part (ii) of the Hopf-Rinow Theorem 53.7. The claim now follows from Proposition 53.13. ■

The second application of Theorem 53.12 deals with the case where the curvature is uniformly positive. We define the **diameter** of a Riemannian manifold as

$$\operatorname{diam}(M, m) \coloneqq \sup \{ d_m(x, y) \mid x, y \in M \}.$$

THEOREM 53.15 (The Bonnet-Myers Theorem). *Let (M^n, m) be a complete Riemannian manifold with $\operatorname{sect}_m \geq \kappa > 0$. Then $\operatorname{diam}(M, m) \leq \frac{\pi}{\sqrt{\kappa}}$. Moreover M is necessarily compact and has finite fundamental group.*

Proof. Since (M, m) is complete, by Corollary 53.9 any two points can be joined by a minimal geodesic. But from part (iii) of Theorem 53.12, any geodesic γ with $\mathbb{L}_m(\gamma) \geq \frac{\pi}{\sqrt{\kappa}}$ has a conjugate point, and hence cannot be minimal by Proposition 53.10. Thus in fact any two points can be joined by a minimal geodesic with length less that $\frac{\pi}{\sqrt{\kappa}}$, and hence $\operatorname{diam}(M, m) \leq \frac{\pi}{\sqrt{\kappa}}$ as claimed. Then by part (iv) of the Hopf-Rinow Theorem 53.7, M is compact.

Finally, the same argument can be applied to the universal cover of M, equipped with the pullback metric. This allows us to conclude that the universal cover of M is compact, which implies that $\pi_1(M)$ is finite. This completes the proof. ∎

And also the course! Thank you all for attending!

Problem Sheet A

Problem A.1. Prove that the set $\mathrm{GL}(n)$ of invertible $n \times n$ matrices is a smooth manifold of dimension n^2.

Problem A.2. Let M and N be two smooth manifolds of dimension n and k respectively. Prove that $M \times N$ is a smooth manifold of dimension $n + k$. Deduce that the n-dimensional torus:

$$T^n := \underbrace{S^1 \times \cdots \times S^1}_{n \text{ times}} \subset \mathbb{R}^{2n}.$$

is a compact smooth manifold of dimension n.

Problem A.3. Let $\mathbb{R}P^n$ denote n-dimensional real projective space, i.e. the space of lines through the origin in $\mathbb{R}^{n+1}$. Prove that $\mathbb{R}P^n$ is a compact smooth manifold of dimension n.

(♣) Problem A.4. Let $G(k, n)$ denote the set of k-dimensional linear subspaces of $\mathbb{R}^n$. We call $G(k, n)$ a Grassmannian manifold. Prove that $G(k, n)$ is a compact smooth manifold and compute its dimension.

Problem A.5. Let X denote the union of the x-axis and the y-axis in $\mathbb{R}^2$. Prove that X is *not* a topological manifold.

Problem A.6. Let Y denote the "pinched 2-dimensional torus", as shown in Figure A.1. Prove that Y is *not* a topological manifold.

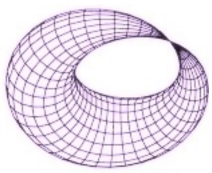

Figure A.1: The pinched torus Y.

Problem A.7. Show that the smooth atlas on $\mathbb{R}$ consisting of the single chart $\sigma\colon \mathbb{R} \to \mathbb{R}$ given by $\sigma(x) = x^3$ defines a smooth structure that is *different* to the "standard" smooth structure (the latter is the smooth structure containing the identity map as a chart). Prove however that both the smooth structures belong to the same diffeomorphism class.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Solutions to Problem Sheet A

Problem A.1. Prove that the set $\mathrm{GL}(n)$ of invertible $n \times n$ matrices is a smooth manifold of dimension n^2.

Solution. The space $\{n \times n \text{ matrices}\}$ is homeomorphic to $\mathbb{R}^{n^2}$, and hence is a smooth manifold of dimension n^2 by Example 1.10. The determinant function

$$\det\colon \{n \times n \text{ matrices}\} \to \mathbb{R}$$

is continuous, and $\mathrm{GL}(n) = \det^{-1}(\mathbb{R} \setminus \{0\})$ is thus an open subset. The claim follows from Lemma 1.20.

Problem A.2. Let M and N be two smooth manifolds of dimension n and k respectively. Prove that $M \times N$ is a smooth manifold of dimension $n + k$. Deduce that the n-dimensional torus:

$$T^n := \underbrace{S^1 \times \cdots \times S^1}_{n \text{ times}} \subset \mathbb{R}^{2n}.$$

is a compact smooth manifold of dimension n.

Solution. Indeed, if $\Sigma_1 = \{\sigma_{\mathtt{a}}\colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ and $\Sigma_2 = \{\tau_{\mathtt{b}}\colon V_{\mathtt{b}} \to \Omega_{\mathtt{b}} \mid \mathtt{b} \in \mathtt{B}\}$ are smooth atlases on M and N respectively then

$$\Sigma_1 \times \Sigma_2 := \{(\sigma_{\mathtt{a}}, \tau_{\mathtt{b}})\colon U_{\mathtt{a}} \times V_{\mathtt{b}} \to O_{\mathtt{a}} \times \Omega_{\mathtt{b}} \mid (\mathtt{a}, \mathtt{b}) \in \mathtt{A} \times \mathtt{B}\}$$

is a smooth atlas on $M \times N$, thus proving that $M \times N$ is a smooth manifold of the required dimension[1].

Therefore, the n-dimensional torus T^n is a smooth manifold as S^1 is a smooth manifold by Proposition 1.21.

Problem A.3. Let $\mathbb{R}P^n$ denote n-dimensional real projective space, that is, the space of lines through the origin in $\mathbb{R}^{n+1}$. Prove that $\mathbb{R}P^n$ is a compact smooth manifold of dimension n.

Solution. For convenience we use the definition of $\mathbb{R}P^n$ via equivalence classes, i.e. we define

$$\mathbb{R}P^n = \left\{[x]_\sim \mid x \in \mathbb{R}^{n+1} \setminus \{0\}\right\},$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

[1] Strictly speaking we should also check here that $M \times N$ is Hausdorff, has countably many connected components, and is paracompact, when endowed with the topology defined by declaring that all the maps $(\sigma_{\mathtt{a}}, \tau_{\mathtt{b}})$ are homeomorphisms. The first two conditions are obvious, but the fact that $M \times N$ is necessarily paracompact is less so , since the product of paracompact spaces is *not* always paracompact. One way to prove this would be to use the fact that metric spaces are paracompact (Theorem 1.4), that manifolds are metrisable (part (3) of Remark 1.9), and that the product of metric spaces is also a metric space.

where for x, y two elements in $\mathbb{R}^{n+1} \setminus \{0\}$ we define

$$x \sim y \colon \iff \exists \lambda \in \mathbb{R} \setminus \{0\} \text{ such that } x = \lambda y.$$

We will also adopt the notation $[x] = [x_1 \colon \ldots \colon x_{n+1}]$ for the equivalence class of $x = (x_1, \ldots, x_{n+1})$. We equip $\mathbb{R}P^n$ with the final topology of the projection

$$\pi \colon \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n,\ \pi(x) = [x]$$

and define an open cover $\{U_i\}_{i=1,\ldots,n+1}$ by setting

$$U_i = \{[x_1 \colon \ldots \colon x_{n+1}] \mid x_i \neq 0\},$$

for all $i = 1, \ldots, n+1$. Furthermore, for each i we define functions

$$\sigma_i \colon U_i \to \mathbb{R}^n,\ [x_1 \colon \ldots \colon x_{n+1}] \mapsto \frac{1}{x_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1})$$

and claim that

$$\Sigma := \{\sigma_i \colon U_i \to \mathbb{R}^n \mid i = 1, \ldots, n+1\}$$

defines a smooth atlas on $\mathbb{R}P^n$. It is straightforward to show that each σ_i is well defined and is a bijection. Continuity of σ_i can be shown using the universal property of the final topology on $\mathbb{R}P^n$. Thus, in order to conclude that the σ_i's are all homeomorphisms it suffices to find a continuous inverse τ_i, which can be defined as follows

$$\tau_i \colon \mathbb{R}^n \to U_i,\ \tau_i(x_1, \ldots, x_n) = [x_1 \colon \ldots \colon x_{i-1} \colon 1 \colon x_i \colon \ldots \colon x_n].$$

Note that all the τ_i's are indeed continuous, for they can all be written as a composition of continuous maps. Therefore the σ_i's are all homeomorphisms and we are only left to show the compatibility condition: Let σ_i and σ_j be two charts defined as above and assume without loss of generality that $i < j$. We need to show that the transition map

$$\sigma_i \circ \sigma_j^{-1} \colon \sigma_j(U_i \cap U_j) \to \sigma_i(U_i \cap U_j)$$

is a diffeomorphism, but plugging in the definitions shows that

$$\begin{aligned}
\sigma_i \circ \sigma_j^{-1}(x_1, \ldots, x_n) &= \sigma_i[x_1 \colon \ldots \colon x_{j-1} \colon 1 \colon x_j \colon \ldots \colon x_n] \\
&= \frac{1}{x_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, 1, x_j, \ldots, x_n),
\end{aligned}$$

which is clearly a diffeomorphism. This proves the claim that Σ is an atlas. The easiest way to see why $\mathbb{R}P^n$ is compact (thus paracompact) and Hausdorff is by viewing $\mathbb{R}P^n$ as the sphere S^n with the antipodal identification

$$\mathbb{R}P^n = S^n / (-x \sim x),$$

which readily finishes the proof.

(♣) Problem A.4. Let $G(k, n)$ denote the set of k-dimensional linear subspaces of $\mathbb{R}^n$. We call $G(k, n)$ a Grassmannian manifold. Prove that $G(k, n)$ is a compact smooth manifold and compute its dimension.

SOLUTION. We first determine a model for $G(k,n)$. We choose the obvious approach: Every element of $G(k,n)$ is given by the image $A(\mathbb{R}^k)$ of some injective linear map $\mathbb{R}^k \to \mathbb{R}^n$, represented by a $n \times k$ matrix A of rank k (i.e. maximal rank). We will denote the space of such matrices A by $\mathcal{M}(k,n)$. Of course, for any invertible $k \times k$ matrix B we will have $A(\mathbb{R}^k) = AB(\mathbb{R}^k)$, so in fact we have a map

$$\begin{aligned} \mathcal{M}(k,n)/\sim &\to G(k,n) \\ [A] &\mapsto A(\mathbb{R}^k), \end{aligned} \tag{A.1}$$

where $A \sim \widetilde{A}$ if and only if there exists an invertible $k \times k$ matrix B with $A = \widetilde{A}B$. We noted above that every element of $G(k,n)$ is in the image of (A.1) and by linear algebra (A.1) is also injective (think about why this is true). Hence, our model of $G(k,n)$ will be the space $\mathcal{M}(k,n)/\sim$.

Now, for integers $1 \le i_1 < \cdots < i_k \le n$ we denote by $U_{i_1,\ldots,i_k} \subset \mathcal{M}(k,n)/\sim$ the set of classes $[A]$ such that the i_1'th, i_2'th up to i_k'th row form a set of linearly independent vectors in $\mathbb{R}^k$.[2] Note that if we equip $\mathcal{M}(k,n)$ with the obvious topology it inherits from $\mathbb{R}^{k\cdot n}$, and $\mathcal{M}(k,n)/\sim$ with the quotient topology, then the $U_{i_1,\ldots,i_k}$s are open subsets (check this). Note also that $\{U_{i_1,\ldots,i_k}\}_{1\le i_1<\ldots<i_k\le n}$ is an open cover of $\mathcal{M}(k,n)/\sim$. By Gauss elimination every class $\alpha \in U_{1,\ldots,k} \subset \mathcal{M}(k,n)/\sim$ contains a unique element of the form

$$\begin{pmatrix} \mathrm{id}_{k\times k} \\ \hline Z_\alpha \end{pmatrix} \tag{A.2}$$

for some (unique) $(n-k) \times k$ matrix Z_α. Hence, we can define a bijective map

$$\begin{aligned} \sigma_{1,\ldots,k} \colon U_{1,\ldots,k} &\to \mathbb{R}^{k(n-k)} \\ \alpha &\mapsto Z_\alpha, \end{aligned}$$

which is clearly a homeomorphism. We have a bijection

$$\varphi_{i_1,\ldots,i_k} \colon U_{i_1,\ldots,i_k} \to U_{1,\ldots,k}$$

which is given by permuting the rows of every element of $U_{i_1,\ldots,i_k}$ with the unique permutation which takes $(i_1,\ldots,i_k)$ to $(1,\ldots,k)$ and preserves the relative order of $\{1,\ldots,n\}\backslash\{i_1,\ldots,i_k\}$. We can then define

$$\sigma_{i_1,\ldots,i_k} \colon U_{i_1,\ldots,i_k} \to \mathbb{R}^{k(n-k)}$$

as the composition $\sigma_{i_1,\ldots,i_k} := \sigma_{1,\ldots,k} \circ \varphi_{i_1,\ldots,i_k}$, which is again a homeomorphism. The transition maps

$$\sigma_{i_1,\ldots,i_k} \circ \sigma^{-1}_{j_1,\ldots,j_k} \colon \sigma_{j_1,\ldots,j_k}(U_{j_1,\ldots,j_k} \cap U_{i_1,\ldots,i_k}) \to \sigma_{i_1,\ldots,i_k}(U_{j_1,\ldots,j_k} \cap U_{i_1,\ldots,i_k})$$

are given by

$$\sigma_{1,\ldots,k} \circ \varphi_{i_1,\ldots,i_k} \circ \varphi^{-1}_{j_1,\ldots,j_k} \circ \sigma^{-1}_{1,\ldots,k}.$$

These maps are smooth simply because they are compositions of the following smooth maps: $\sigma^{-1}_{1,\ldots,k}$ takes a $(n-k) \times k$ matrix Z to the $n \times k$ matrix

$$\begin{pmatrix} \mathrm{id}_{k\times k} \\ \hline Z \end{pmatrix}.$$

[2] Note that $U_{i_1,\ldots,i_k}$ is well-defined because one element of a class in $\mathcal{M}(k,n)/\sim$ satisfies the mentioned condition if and only if all elements of that class do.

The map $\varphi_{i_1,\dots,i_k} \circ \varphi^{-1}_{j_1,\dots,j_k}$ simply permutes rows of this matrix to get a matrix

$$\begin{pmatrix} C \\ D \end{pmatrix}$$

for an invertible $k \times k$ matrix C and a $(n-k) \times k$ matrix D. The map $\sigma^{-1}_{1,\dots,k}$ sends this matrix to the matrix

$$DC^{-1} \in \mathbb{R}^{(n-k)k},$$

so it is clearly also smooth.

We now argue that $G(k,n)$ is in fact compact. We do this by induction on k. Note first that $G(1,n) = \mathbb{R}P^{n-1}$, which was seen to be compact in the previous exercise. Hence, it suffices to prove the induction step, i.e. that $G(k+1,n)$ is compact if $G(k,n)$ is. For this purpose, define the set[3]

$$\mathcal{S} := \bigcup_{V \in G(k,n)} (S^{n-1} \cap V^{\perp}) \times \{V\} \subset S^{n-1} \times G(k,n).$$

$\mathcal{S}$ is easily seen to be a compact subset of $S^{n-1} \times G(k,n)$. In particular, if $G(k,n)$ is compact then so is $\mathcal{S}$. The map

$$\begin{aligned} \mathcal{S} &\to G(k+1,n) \\ (v,V) &\mapsto \mathbb{R}v + V \end{aligned} \tag{A.3}$$

is both surjective (use the Gram-Schmidt orthogonalization procedure) and continuous. To see the latter, define S to be the preimage of $\mathcal{S}$ under the quotient map $\mathcal{M}(k,n) \to \mathcal{M}(k,n)/\sim$. Then the map (A.3) is the map on the quotient indiced by the (obviously continuous) composition of maps

$$\begin{aligned} S &\mapsto \mathcal{M}(k+1,n) \to G(k+1,n) \\ (v,A) &\mapsto \big(v \mid A\big) \mapsto \big[\big(v \mid A\big)\big] \end{aligned}$$

In particular $G(k+1,n)$ is compact if $\mathcal{S}$ is. Since (A.3) is both continuous and surjective, and $\mathcal{S}$ is compact it follows that $G(k+1,n)$ is compact.

Problem A.5. Let X denote the union of the x-axis and the y-axis in $\mathbb{R}^2$. Prove that X is *not* a topological manifold.

Solution. Since X has the subspace topology inherited from $\mathbb{R}^2$, it is clear that any *connected* neighbourhood $U \subset X$ of $(0,0)$ intersects both axes and that $U \setminus \{(0,0)\}$ has four connected components. Suppose that X is a topological manifold. Then X must be 1-dimensional, since every point outside of the origin is contained in a neighbourhood homeomorphic to an open subset of $\mathbb{R}$. We may thus choose U such that there exists a homeomorphism $\sigma \colon U \xrightarrow{\sim} O \subset \mathbb{R}^1$, where O is a connected open subset of $\mathbb{R}$, i.e., a bounded open interval. But then σ induces a homeomorphism from $U \setminus \{(0,0)\}$ to $O \setminus \{\sigma(0,0)\}$. Since the former has four connected components and latter has only two, this yields a contradiction.

[3] Here $V^{\perp} \subset \mathbb{R}^n$ denotes the set of vectors which are perpendicular to V with respect to the Euclidean inner product on $\mathbb{R}^{n+1}$ and $S^{n-1} \subset \mathbb{R}^n$ denotes the set of elements which have length 1.

PROBLEM A.6. Let Y denote the "pinched torus". Prove that Y is *not* a topological manifold.

SOLUTION. Let $y \in Y$ denote the "pinched point" on Y. Suppose Y is a topological manifold. Clearly $Y \setminus \{y\}$ is two-dimensional, so Y must be as well. There is therefore an open neighbourhood U of y which is homeomorphic to an open disc B in $\mathbb{R}^2$. But then $U \setminus \{y\}$ is homeomorphic to B minus a point. Since the former is disconnected and the latter is connected, this is a contradiction.

PROBLEM A.7. Show that the smooth atlas on $\mathbb{R}$ consisting of the single chart $\sigma\colon \mathbb{R} \to \mathbb{R}$ given by $\sigma(x) = x^3$ defines a smooth structure that is *different* to the "standard" smooth structure (the latter is the smooth structure containing the identity map as a chart). Prove however that both the smooth structures belong to the same diffeomorphism class.

SOLUTION. To prove that the maps

$$\sigma\colon \mathbb{R} \to \mathbb{R} \quad x \mapsto x^3 \quad \text{and} \quad \iota\colon \mathbb{R} \to \mathbb{R} \quad x \mapsto x$$

define charts on $\mathbb{R}$ belonging to different smooth structures, it is sufficient to check that either $\sigma \circ \iota^{-1}$ or $\iota \circ \sigma^{-1}$ are not smooth in the ordinary sense. Since

$$(\iota \circ \sigma^{-1})(x) = x^{1/3}$$

is not smooth (because its derivative $x^{-2/3}/3$ is singular at 0), the conclusion is reached.

To prove that the differentiable structures on $\mathbb{R}$ induced by σ and ι, which we denote by Σ_σ and Σ_ι respectively, belong to the same diffeomorphism class, we need to find a diffeomorphism $\varphi\colon (\mathbb{R}, \Sigma_\sigma) \to (\mathbb{R}, \Sigma_\iota)$. Since σ and ι provide global charts for $\mathbb{R}$, and it is not important which chart is chosen to check differentiability, it is sufficient to find a bijective map $\varphi\colon \mathbb{R} \to \mathbb{R}$ so that $\iota \circ \varphi \circ \sigma^{-1}$ and $\sigma \circ \varphi^{-1} \circ \iota^{-1}$ are smooth in the ordinary sense. By choosing $\varphi(x) = x^3$, we see that

$$(\iota \circ \varphi \circ \sigma^{-1})(x) = (x^{1/3})^3 = x \quad \text{and} \quad (\sigma \circ \varphi^{-1} \circ \iota)(x) = (x^{1/3})^3 = x$$

are smooth as required, hence the conclusion is reached.

Problem Sheet B

Problem B.1. Let $\{V_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be a family of vector spaces indexed by a set $\mathtt{A}$, and let W be a fixed set. Suppose that for each $\mathtt{a} \in \mathtt{A}$ we are given a bijection $T_{\mathtt{a}} \colon V_{\mathtt{a}} \to W$ such that for any $\mathtt{a}, \mathtt{b} \in \mathtt{A}$, the composition $T_{\mathtt{b}}^{-1} \circ T_{\mathtt{a}} \colon V_{\mathtt{a}} \to V_{\mathtt{b}}$ is a linear isomorphism. Prove that there is a unique vector space structure on W such that each $T_{\mathtt{a}}$ is a linear isomorphism.

Problem B.2. Let M be a smooth manifold of dimension n with maximal smooth atlas Σ. Given a point $x \in M$, let $\Sigma_x \subset \Sigma$ denote the set of charts $\sigma \in \Sigma$ such that x lies in the domain of σ. Define an equivalence relation on $\mathbb{R}^n \times \Sigma_x$ by saying

$$(v, \sigma) \sim (w, \tau) \qquad \Leftrightarrow \qquad D(\tau \circ \sigma^{-1})(\sigma(x))[v] = w.$$

(i) Prove that this is indeed a well-defined equivalence relation.

(ii) Let $\mathcal{T}_x$ denote the set of equivalence classes. Let $\sigma \in \Sigma_x$. Prove that the map $T_\sigma \colon \mathbb{R}^n \to \mathcal{T}_x$ given by

$$T_\sigma v \coloneqq [(v, \sigma)]$$

(where $[(v, \sigma)]$ denotes the equivalence class of (v, σ)) is a bijection. Deduce that $\mathcal{T}_x$ admits a unique vector space structure such that each T_σ is a linear isomorphism.

(iii) Let σ be a chart defined on a neighbourhood of x with local coordinates $x^i = u^i \circ \sigma$. Let[1] $\tilde{T}_\sigma \colon \mathbb{R}^n \to T_x M$ denote the linear isomorphism defined by

$$\tilde{T}_\sigma e_i = \frac{\partial}{\partial x^i}\Big|_x.$$

Prove[2] that there exists a linear isomorphism $\mathcal{S}_x \colon \mathcal{T}_x \to T_x M$ which in addition satisfies

$$\mathcal{S}_x \circ T_\sigma = \tilde{T}_\sigma,$$

for every chart σ about x.

Problem B.3. Let V be any vector space of dimension n, endowed with its standard smooth structure (cf. Example 1.19). Fix $x \in V$. Define a map

$$\mathcal{J}_x \colon V \to T_x V, \qquad \mathcal{J}_x(v) \coloneqq \gamma'(0), \qquad \text{where } \gamma(t) \coloneqq x + tv.$$

Prove that $\mathcal{J}_x$ is an isomorphism[3].

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] This map was denoted by T_x in (4.3), but here it is important to emphasise the dependence on σ so we use different notation.

[2] This provides yet another equivalent way of defining the tangent space.

[3] This proves that the tangent space to a vector space at any given point is *canonically* (i.e. independent of the choice of basis) identified with the vector space itself.

Problem B.4. Let V and W be vector spaces and assume that $T\colon V \to W$ is a linear map. Prove that the following commutes[4] for any $x \in V$:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\mathcal{J}_x"'] & W \arrow[d, "\mathcal{J}_{Tx}"] \\ T_xV \arrow[r, "DT(x)"'] & T_{Tx}W \end{tikzcd}$$

(♣) Problem B.5. Consider the subspace $S := \mathbb{R} \times \{-1\} \cup \mathbb{R} \times \{1\} \subseteq \mathbb{R}^2$ together with its subspace topology and define an equivalence on S by setting

$$(x_1, y_1) \sim (x_2, y_2) \iff x_1 = x_2 \text{ and } x_1, x_2 \neq 0.$$

Equip $M = S/\sim$ [5] with the quotient topology and define two functions

$$\sigma_1\colon M \setminus (0,-1) \to \mathbb{R} \text{ and } \sigma_2\colon M \setminus (0,1) \to \mathbb{R},$$

by setting $\sigma_1(x,y) = x$ and $\sigma_2(x,y) = x$. Show that M is paracompact and that $\Sigma = \{\sigma_1, \sigma_2\}$ defines a smooth atlas, but that M is *not* Hausdorff.

(♣) Problem B.6. Consider $\mathbb{R}^2$ as a set and equip it with the topology $\mathcal{T}$ generated by the basis $\mathcal{B} = \{U \times \{\mathtt{a}\} \mid U \subseteq \mathbb{R} \text{ open}, \mathtt{a} \in \mathbb{R}\}$. Define $\sigma_{\mathtt{a}}\colon \mathbb{R} \times \{\mathtt{a}\} \to \mathbb{R}$, $\sigma_{\mathtt{a}}(x, \mathtt{a}) = x$ and set $\Sigma = \{\sigma_{\mathtt{a}} \mid \mathtt{a} \in \mathbb{R}\}$. Prove that the topological space $(\mathbb{R}^2, \mathcal{T})$ is paracompact and that Σ defines a smooth atlas on it, but that it has an uncountable number of connected components.

[4] The commutative diagram thus gives a coordinate-free proof of (4.2).

[5] This topological space is also known as the "line with two origins". You should convince yourself that the name makes perfect sense!

Solutions to Problem Sheet B

PROBLEM B.1. Let $\{V_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be a family of vector spaces indexed by a set $\mathtt{A}$, and let W be a fixed set. Suppose that for each $\mathtt{a} \in \mathtt{A}$ we are given a bijection $T_{\mathtt{a}} \colon V_{\mathtt{a}} \to F$ such that for any $\mathtt{a}, \mathtt{b} \in \mathtt{A}$, the composition $T_{\mathtt{b}}^{-1} \circ T_{\mathtt{a}} \colon V_{\mathtt{a}} \to V_{\mathtt{b}}$ is a linear isomorphism. Prove that there is a unique vector space structure on W such that each $T_{\mathtt{a}}$ is a linear isomorphism.

SOLUTION. We must define a zero element, addition and scalar multiplication on W. Let $\mathtt{a} \in \mathtt{A}$. Since we want $T_{\mathtt{a}}$ to be a linear isomorphism, we must do the following:

- Define $0 \in W$ via $0 \coloneqq T_{\mathtt{a}}(0)$.[1]
- For each $v, w \in W$, define $v + w \coloneqq T_{\mathtt{a}}\big(T_{\mathtt{a}}^{-1}(v) + T_{\mathtt{a}}^{-1}(w)\big)$.
- For each $\lambda \in \mathbb{R}$ and $v \in W$, define $\lambda v \coloneqq T_{\mathtt{a}}\big(\lambda T_{\mathtt{a}}^{-1}(v)\big)$.

This satisfies the axioms of a vector space since $V_{\mathtt{a}}$ is itself a vector space. Moreover, the definitions are forced if we want $T_{\mathtt{a}}$ to be linear, whence the uniqueness of the resulting vector space structure on W.

It remains to show that the vector space structure on W is independent of the choice of $\mathtt{a} \in \mathtt{A}$. Let $\mathtt{b} \in \mathtt{A}$. The independence of "+" on F follows from the following computation:

$$\begin{aligned} T_{\mathtt{a}}\big(T_{\mathtt{a}}^{-1}(v) + T_{\mathtt{a}}^{-1}(w)\big) &= T_{\mathtt{a}}\big(T_{\mathtt{a}}^{-1} \circ T_{\mathtt{b}} \circ T_{\mathtt{b}}^{-1}(v) + T_{\mathtt{a}}^{-1} \circ T_{\mathtt{b}} \circ T_{\mathtt{b}}^{-1}(w)\big) \\ &= T_{\mathtt{a}} \circ T_{\mathtt{a}}^{-1} \circ T_{\mathtt{b}}\big(T_{\mathtt{b}}^{-1}(v) + T_{\mathtt{b}}^{-1}(w)\big) \\ &= T_{\mathtt{b}}\big(T_{\mathtt{b}}^{-1}(v) + T_{\mathtt{b}}^{-1}(w)\big). \end{aligned}$$

The proof that the definition of scalar multiplication on W is also independent from the choice of a is similar.

PROBLEM B.2. Let M be a smooth manifold of dimension n with maximal smooth atlas Σ. Given a point $x \in M$, let $\Sigma_x \subset \Sigma$ denote the set of charts $\sigma \in \Sigma$ such that x lies in the domain of σ. Define an equivalence relation on $\mathbb{R}^n \times \Sigma_x$ by saying

$$(v, \sigma) \sim (w, \tau) \qquad \Leftrightarrow \qquad D(\tau \circ \sigma^{-1})(\sigma(x))[v] = w.$$

(i) Prove that this is indeed a well-defined equivalence relation.

(ii) Let $\mathcal{T}_x$ denote the set of equivalence classes. Let $\sigma \in \Sigma_x$. Prove that the map $T_\sigma \colon \mathbb{R}^n \to \mathcal{T}_x$ given by

$$T_\sigma v \coloneqq [(v, \sigma)]$$

(where $[(v, \sigma)]$ denotes the equivalence class of (v, σ)) is a bijection. Deduce that $\mathcal{T}_x$ admits a unique vector space structure such that each T_σ is a linear isomorphism.

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

[1] Note that the zero element of W is already uniquely determined by the second bullet point since $v + w = v$ if and only if $T_{\mathtt{a}}^{-1}(v) + T_{\mathtt{a}}^{-1}(w) = T_{\mathtt{a}}^{-1}(v)$. We include its definition for clarity.

(iii) Let σ be a chart defined on a neighbourhood of x with local coordinates $x^i = u^i \circ \sigma$. Let[2] $\tilde{T}_\sigma \colon \mathbb{R}^n \to T_xM$ denote the linear isomorphism defined by

$$\tilde{T}_\sigma e_i = \frac{\partial}{\partial x^i}\Big|_x.$$

Prove[3] that there exists a linear isomorphism $\mathcal{S}_x \colon \mathcal{T}_x \to T_xM$ which in addition satisfies

$$\mathcal{S}_x \circ T_\sigma = \tilde{T}_\sigma,$$

for every chart σ about x.

Solution. We start by showing that $\sim$ is an equivalence relation. Reflexivity is quickly verified:

$$D(\sigma \circ \sigma^{-1})(\sigma(x))[v] = D(\mathrm{id})(\sigma(x))[v] = v.$$

Assuming that $(v, \sigma) \sim (w, \tau)$, we get symmetry by plugging in $w = D(\tau \circ \sigma^{-1})(\sigma(x))[v]$ and applying the chain rule

$$\begin{aligned} D(\sigma \circ \tau^{-1})(\tau(x))[w] &= D(\sigma \circ \tau^{-1})(\tau(x)) \circ D(\tau \circ \sigma^{-1})(\sigma(x))[v] \\ &= D(\underbrace{\sigma \circ \tau^{-1} \circ \tau \circ \sigma^{-1}}_{=\mathrm{id}})(\sigma(x))[v] = v. \end{aligned}$$

Transitivity follows from a similar chain rule argument. This concludes the proof of part (i).

For part (ii) it is convenient to invoke the result of (B.1) with $A = \Sigma_x$ and $F = \mathcal{T}_x$. Therefore we are only left to show that each $T_\sigma \colon \mathbb{R}^n \to \mathcal{T}_x$ defines a bijection and that $T_\tau^{-1} \circ T_\sigma$ is a linear isomorphism for every $\sigma, \tau \in \Sigma_x$. Injectivity readily follows from the observation

$$(v, \sigma) \sim (w, \sigma) \iff v = w.$$

For surjectivity let us pick an arbitrary $(w, \tau) \in \mathbb{R}^n \times \Sigma_x$. Simply setting $v := D(\sigma \circ \tau^{-1})(\tau(x))w$ grants $(v, \sigma) \sim (w, \tau)$ and hence

$$T_\sigma(v) = [(v, \sigma)] = [(w, \tau)].$$

A similar argument as in the surjectivity proof above shows that

$$T_\tau^{-1} \circ T_\sigma(v) = \underbrace{D(\tau \circ \sigma^{-1})(\sigma(x))}_{\text{linear isomorphism}}[v].$$

This proves part (ii).

For part (iii) we make an educated guess for $\mathcal{S}_x$, namely

$$\mathcal{S}_x \colon \mathcal{T}_x \to T_xM, \qquad [(v, \sigma)] \mapsto \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\Big|_x,$$

[2]This map was denoted by T_x in (4.3), but here it is important to emphasise the dependence on σ so we use different notation.

[3]This provides yet another equivalent way of defining the tangent space.

where $v = (v^1, \dots, v^n)$ and $x^i = u^i \circ \sigma$ denotes the local coordinates of σ. The fact that $\mathcal{S}_x$ is well defined follows immediately from Remark 3.3, but for the sake of completeness we will carry out the computation. Let $(w, \tau) \sim (v, \sigma)$. Let y^i denote the local coordinates of τ. We then have

$$\frac{\partial}{\partial y^j}\Big|_x (x^k) = D_j(x^k \circ \tau^{-1})(x) = \langle e_k, D(\sigma \circ \tau^{-1})(x)[e_j]\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean inner product. Using bilinearity of the inner product one obtains the identity

$$\sum_j w^j \frac{\partial}{\partial y^j}\Big|_x (x^k) = \langle e_k, D(\sigma \circ \tau^{-1})(x)[w]\rangle = v^k = \sum_i v^i \frac{\partial}{\partial x^i}\Big|_x (x^k),$$

which proves well-definedness of $\mathcal{S}_x$.

The relation $\mathcal{S}_x \circ T_\sigma = \tilde{T}_\sigma$ is easily verified. The map $\mathcal{S}_x$ is a linear isomorphism, for both T_σ and $\tilde{T}_\sigma$ are linear isomorphisms.

PROBLEM B.3. Let V be any vector space of dimension n, endowed with its standard smooth structure (cf. Example 1.19). Fix $x \in V$. Define a map

$$\mathcal{J}_x : V \to T_x V, \qquad \mathcal{J}_x(v) := \gamma'(0), \qquad \text{where } \gamma(t) := x + tv.$$

Prove that $\mathcal{J}_x$ is an isomorphism[4].

SOLUTION. The smooth structure on V is determined taking a chart which is a linear isomorphism $T : V \to \mathbb{R}^n$. Let $x^i = u^i \circ T$ denote the local coordinates of such a chart. The map T determines a basis of V, namely $w_i = T^{-1}e_i$, and if one writes an arbitrary vector in V in terms of this basis as $w = \sum_{i=1}^n a^i w_i$ then $a^i = x^i(w)$. Now with $\gamma(t) := x + tv$ one has

$$\begin{aligned}
\mathcal{J}_x v &= \gamma'(0) \\
&= \sum_{i=1}^n \gamma'(0)(x^i) \cdot \frac{\partial}{\partial x^i}\Big|_x \\
&= \sum_{i=1}^n (x^i \circ \gamma)'(0) \cdot \frac{\partial}{\partial x^i}\Big|_x \\
&= \sum_{i=1}^n x^i(v) \cdot \frac{\partial}{\partial x^i}\Big|_x .
\end{aligned}$$

This shows that the matrix of $\mathcal{J}_x$ with respect to the basis $\{w_i\}$ of V and $\left\{\frac{\partial}{\partial x^i}\big|_x\right\}$ of $T_x V$ is simply given by the map T itself. Since T is an linear isomorphism, it follows that $\mathcal{J}_x$ is too. (Remark: As is common in many arguments in linear algebra, we chose a basis of V to prove $\mathcal{J}_x$ was an isomorphism, although the claim is then independent of the basis.)

[4] This proves that the tangent space to a vector space at any given point is *canonically* (i.e. independent of the choice of basis) identified with the vector space itself.

Problem B.4. Let V and W be vector spaces and assume that $T\colon V \to W$ is a linear map. Prove that the following commutes[5] for any $x \in V$:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\mathcal{J}_x"'] & W \arrow[d, "\mathcal{J}_{Tx}"] \\ T_x V \arrow[r, "DT(x)"'] & T_{Tx} W \end{tikzcd}$$

Solution. Fix $x, v \in V$. With $\gamma(t) = x + tv$ we compute:

$$\begin{aligned} \mathcal{J}_{Tx} \circ Tv &= \widetilde{\gamma}'(0), \\ DT(x) \circ \mathcal{J}_x(v) &= DT(x) \circ D\gamma(0)\left[\frac{\partial}{\partial t}\Big|_t\right] = D(T \circ \gamma)(0)\left[\frac{\partial}{\partial t}\Big|_t\right] = \widehat{\gamma}'(0), \end{aligned}$$

with $\widetilde{\gamma}(t) = Tx + tTv$ and $\widehat{\gamma}(t) = T(x + tv)$. Note that in the last row of these equations we make use of the manifold version of the chain rule (Proposition 4.2) together with Proposition 4.10. Since T is linear we have $\widetilde{\gamma}(t) = Tx + tTv = T(x + tv) = \widehat{\gamma}(t)$, so it follows that $\mathcal{J}_{Tx} \circ T(v) = DT(x) \circ \mathcal{J}_x(v)$ for all $x, v \in V$, which exactly says that the diagram commutes.

(♣) Problem B.5. Consider the subspace $S \coloneqq \mathbb{R} \times \{-1\} \cup \mathbb{R} \times \{1\} \subseteq \mathbb{R}^2$ together with its subspace topology and define an equivalence on S by setting

$$(x_1, y_1) \sim (x_2, y_2) \iff x_1 = x_2 \text{ and } x_1, x_2 \neq 0.$$

Equip $M = S/\sim$ [6] with the quotient topology and define two functions

$$\sigma_1 \colon M \setminus (0, -1) \to \mathbb{R} \text{ and } \sigma_2 \colon M \setminus (0, 1) \to \mathbb{R},$$

by setting $\sigma_1(x, y) = x$ and $\sigma_2(x, y) = x$. Show that M is paracompact and that $\Sigma = \{\sigma_1, \sigma_2\}$ defines a smooth atlas, but that M is *not* Hausdorff.

Solution. Note first of all that for each element of $[(x, y)] \in M$ we have

$$[(x, y)] = \begin{cases} \{(x, y), (x, -y)\} & \text{if } x \neq 0, \\ \{(x, y)\} & \text{if } x = 0. \end{cases}$$

In what follows we denote for every $R > 0$ the "ball"

$$\beta_R((0, 1)) = \{[(x, y)] \in M \mid |x| < R, x \neq 0\} \cup (0, 1),$$

(here the inverted commas are used since the M does not have the structure of a metric space - if this were true it would be Hausdorff, and we are going to disprove this below), and similarly

$$\beta_R((0, -1)) = \{[(x, y)] \in M \mid |x| < R, x \neq 0\} \cup (0, -1).$$

[5] The commutative diagram thus gives a coordinate-free proof of (4.2).

[6] This topological space is also known as the "line with two origins". Convince yourself that the name makes perfect sense.

We claim that it is not possible to separate the points $(0,1)$ and $(0,-1)$. Indeed, if $U_{(0,1)}$ is any open neighbourhood of $(0,1)$ in M, then it must contain, for some $r>0$, the "ball" $\beta_r((0,1))$ and similarly if $U_{(0,-1)}$ is any open neighbourhood of $(0,-1)$ in M, then it must contain, for some $\rho>0$ the "ball" $\beta_\rho((0,-1))$. Setting $s=\min\{r,\rho\}$, we deduce in particular that $[(s/2,y)]\in U_{(0,1)}\cap U_{(0,-1)}$ and hence that $U_{(0,1)}$ and $U_{(0,-1)}$ have nonempty intersection. So M is not Hausdorff.

Let us prove that M is paracompact (this proof is very similar to the one for the paracompactness of a general metric space). If $\{X_{\mathtt{a}}\}_{\mathtt{a}\in\mathtt{A}}$ is an open cover for M, we let, for every $m\in\mathbb{N}_{\geq 1}$, $\gamma_m=\beta_m((0,1))\cup\beta_m((0,-1))$. Since $\overline{\gamma_m}$ is compact, we choose $\mathcal{C}_m$ to be a finite collection of $X_{\mathtt{a}}$'s that cover $\overline{\gamma_m}$, and finally $\mathcal{C}'_m$ to be

$$\mathcal{C}'_m=\begin{cases}\mathcal{C}_1, & \text{if } m=1,\\ \{X_{\mathtt{a}}\setminus\overline{\gamma_{m-1}}\mid X_{\mathtt{a}}\in\mathcal{C}_m\}, & \text{if } m>1.\end{cases}$$

We then set $\mathcal{C}'=\cup_{m\in\mathbb{N}_{\geq 1}}\mathcal{C}'_m$ and we claim that this is a locally finite refinement of the cover $\{X_{\mathtt{a}}\}_{\mathtt{a}\in\mathtt{A}}$. Indeed, each of its element is an open subset of some $X_{\mathtt{a}}$, it is by construction locally finite and it covers the whole space M since for every $p\in M$ there exist some $m\in\mathbb{N}_{\geq 1}$ so that $p\in\overline{\gamma_m}$ and hence p is an element of some set of the collection $\mathcal{C}'_m$.

Finally, each of the maps σ_i for $i=1,2$ is a homeomorphism and the transition map $\sigma_2\circ\sigma_1^{-1}\colon\mathbb{R}\setminus\{0\}\to\mathbb{R}\setminus\{0\}$ is the identity map, whence a diffeomorphism. Thus Σ defines a smooth atlas for M.

(♣) PROBLEM B.6. Consider $\mathbb{R}^2$ as a set and equip it with the topology $\mathcal{T}$ generated by the basis $\mathcal{B}=\{U\times\{\mathtt{a}\}\mid U\subseteq\mathbb{R}\text{ open}, \mathtt{a}\in\mathbb{R}\}$. Define $\sigma_{\mathtt{a}}\colon\mathbb{R}\times\{\mathtt{a}\}\to\mathbb{R}$, $\sigma_{\mathtt{a}}(x,\mathtt{a})=x$ and set $\Sigma=\{\sigma_{\mathtt{a}}\mid\mathtt{a}\in\mathbb{R}\}$. Prove that the topological space $(\mathbb{R}^2,\mathcal{T})$ is paracompact and that Σ defines a smooth atlas on it, but that it has an uncountable number of connected components.

SOLUTION. We claim that the connected components of $(\mathbb{R}^2,\mathcal{T})$ are the sets $\mathbb{R}_{\mathtt{a}}:=\mathbb{R}\times\{\mathtt{a}\}$ where $\mathtt{a}\in\mathbb{R}$, and in particular they are uncountably many. Indeed, any such set is open by definition, and since we may write

$$\mathbb{R}_{\mathtt{a}}=\bigcap_{\mathtt{b}\in\mathbb{R}\setminus\{\mathtt{a}\}}\mathbb{R}\times(\mathbb{R}\setminus\{\mathtt{b}\}),$$

$\mathbb{R}_{\mathtt{a}}$ is also closed because each of the $\mathbb{R}\times(\mathbb{R}\setminus\{\mathtt{b}\})$'s is closed. Since the union of the $\mathbb{R}_{\mathtt{a}}$'s is the total space, this implies the claim.

Paracompactness of $(\mathbb{R}^2,\mathcal{T})$ is equivalent to the paracompactness of each of its connected components. Since each of these is homeomorphic to $\mathbb{R}$, which we know to be paracompact, we deduce that the space is paracompact.

Finally, two different charts of above atlas always have disjoint domains, so the smoothness of the transition functions is trivially satisfied. As a consequence Σ defines a smooth atlas for $(\mathbb{R}^2,\mathcal{T})$.

Problem Sheet C

Problem C.1. Let M be a smooth manifold of dimension n. Prove that the cotangent bundle T^*M is naturally a smooth manifold of dimension $2n$.

Problem C.2. Let $\varphi\colon M \to N$ be a smooth map. Prove that $D\varphi\colon TM \to TN$ is also smooth. Prove that if $\varphi\colon M \to N$ is an embedding then so is $D\varphi\colon TM \to TN$.

Problem C.3. Let M be a smooth manifold of dimension n. Let (x^i) be local coordinates on M, and let[1] $v^i := dx^i$ so that (x^i, v^i) are local coordinates on TM. Suppose $f\colon TM \to \mathbb{R}$ is a smooth function. Define the **fibrewise derivative** of f at the point (x, v) to be the element

$$D^{\text{fibre}} f(x,v) \in T_x^* M, \qquad D^{\text{fibre}} f(x,v) := \sum_{i=1}^{n} \frac{\partial}{\partial v^i}\bigg|_{(x,v)} (f) dx^i|_x.$$

Prove that $D^{\text{fibre}} f$ is well defined (i.e. independent of the choice of local coordinates).

Problem C.4. Let $\varphi\colon M \to N$ be an injective immersion with M compact. Prove that φ is an embedding. Give an example to show that this need not be true if M is not compact.

Problem C.5. Let M and N be smooth manifolds. Prove that there is a canonical isomorphism:

$$T_{(x,y)}(M \times N) = T_x M \times T_y N, \qquad \forall\, (x,y) \in M \times N.$$

Problem C.6. Let O be an open subset in $\mathbb{R}^n$ and suppose $f\colon O \to \mathbb{R}$ is smooth. Define $g\colon O \to \mathbb{R}^{n+1}$ by

$$g(x) = (x, f(x)).$$

Prove that g is an smooth embedding, and hence that $g(O)$ is a smooth embedded n-dimensional submanifold of $\mathbb{R}^{n+1}$. We call $g(O)$ the **graph** of f.

Problem C.7. Let $\imath\colon S^n \hookrightarrow \mathbb{R}^{n+1}$ denote the inclusion. Prove that:

$$D\imath(x)[T_x S^n] = \mathcal{J}_x(x^\perp),$$

where $\mathcal{J}_x\colon \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ was defined in Problem B.3 and

$$x^\perp := \left\{ y \in \mathbb{R}^{n+1} \mid \langle x, y\rangle = 0 \right\},$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean dot product.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] We introduce the new notation v^i because we want to think of them as coordinates on TM, and thus consider expressions such as $\frac{\partial}{\partial v^i}$. Without the new notation we would end up with an expression like $\frac{\partial}{\partial (dx^i)}$, which would be horrendously confusing...

(♣) Problem C.8. Let M^n be an embedded submanifold of $\mathbb{R}^k$. We define the **normal space to** M **at** x to be the $(k-n)$-dimensional subspace $\operatorname{Norm}_x M \subset T_x\mathbb{R}^k$ consisting of all vectors that are orthogonal to T_xM with respect to the Euclidean dot product. We define the **normal bundle** of M as the set

$$\operatorname{Norm}(M) := \left\{(x, v) \in T\mathbb{R}^k = \mathbb{R}^k \times \mathbb{R}^k \mid x \in M, v \in \operatorname{Norm}_x M\right\}.$$

Prove that $\operatorname{Norm}(M)$ is an embedded k-dimensional submanifold of $T\mathbb{R}^k = \mathbb{R}^{2k}$.

Solutions to Problem Sheet C

Problem C.1. Let M be a smooth manifold of dimension n. Prove that the cotangent bundle T^*M is naturally a smooth manifold of dimension $2n$.

Solution. This proof is very similar to the tangent bundle proof. It is however interesting is to compare the behaviour of the transformation laws. We will use the Einstein Summation Convention. Let $\{e^i\}$ denote the basis of the dual space $(\mathbb{R}^n)^*$ which is dual to the standard basis $\{e_i\}$, i.e.

$$e^i(e_j) = \delta^i_j.$$

Of course, $(\mathbb{R}^n)^* \cong \mathbb{R}^n$ (an isomorphism is provided by $e^i \mapsto e_i$), but in order to make the formalism work we need to use the dual space[1]. If $\sigma \colon U \to O$ is a chart for M, we build a chart for T^*M as follows: if $\pi \colon T^*M \to M$ is the canonical projection, we define $\tilde{\sigma} \colon \pi^{-1}(U) \to \sigma(U) \times (\mathbb{R}^n)^*$ to be defined as

$$\tilde{\sigma}(x,p) = \left(\sigma(x), p\left(\frac{\partial}{\partial x^i}\bigg|_x\right) e^i\right),$$

This map is a bijection and its image is an open subset of $\mathbb{R}^n \times (\mathbb{R}^n)^* \cong \mathbb{R}^{2n}$. Suppose now that $\Sigma := \{\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ is an atlas on M. If $\sigma_{\mathtt{a}} \colon U_{\mathtt{a}} \to O_{\mathtt{a}}$ and $\sigma_{\mathtt{b}} \colon U_{\mathtt{b}} \to O_{\mathtt{b}}$ are charts for M whose domains have nonempty intersection, denoting by x^i and y^j the local coordinates induced by $\sigma_{\mathtt{a}}$ and $\sigma_{\mathtt{b}}$ respectively, if $(z, q = q_i\, e^i) \in \tilde{\sigma}_{\mathtt{a}}(U_{\mathtt{a}} \cap U_{\mathtt{b}}) \times (\mathbb{R}^n)^*$, then if $x := \sigma_{\mathtt{a}}^{-1}(z)$ we have

$$\tilde{\sigma}_{\mathtt{a}}^{-1}(z,q) = \left(x, q_i\, dx^i\big|_x\right),$$

and hence

$$\tilde{\sigma}_{\mathtt{b}} \circ \tilde{\sigma}_{\mathtt{a}}^{-1}(z,q) = \left(\sigma_{\mathtt{b}}(\sigma_{\mathtt{a}}^{-1}(z)), q_i\, dx^i\big|_x \left(\frac{\partial}{\partial y^j}\bigg|_x\right) e^j\right) \tag{C.1}$$

Since

$$dx^i\big|_x \left(\frac{\partial}{\partial y^j}\bigg|_x\right) = \frac{\partial x^i}{\partial y^j}\bigg|_x$$

by Remark 3.10 (where we are using the convention from Definition 7.4), we have

$$\tilde{\sigma}_{\mathtt{b}} \circ \tilde{\sigma}_{\mathtt{a}}^{-1}(z,q) = \left(\sigma_{\mathtt{b}}(\sigma_{\mathtt{a}}^{-1}(z)), q_i\, \frac{\partial x^i}{\partial y^j}\bigg|_x e^j\right)$$

Now $\frac{\partial x^i}{\partial y^j}\big|_x$ is the (i,j)th entry of the matrix $D(\sigma_{\mathtt{a}} \circ \sigma_{\mathtt{b}}^{-1})(\sigma_{\mathtt{b}}(x))$, again by Remark 3.10, and hence (C.1) shows that

$$\tilde{\sigma}_{\mathtt{b}} \circ \tilde{\sigma}_{\mathtt{a}}^{-1}(z,q) = \left(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1}(z), D(\sigma_{\mathtt{a}} \circ \sigma_{\mathtt{b}}^{-1})(\sigma_{\mathtt{b}} \circ \sigma_{\mathtt{a}}^{-1}(z))^T[q]\right),$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

[1] An easy way to check whether the formalism is correct is to see whether the Einstein Summation Convention works. If this indices are not in the correct position (upper versus lower) then something is wrong.

where we denote by A^T the *transpose* of the matrix A (pay attention to the ordering of $\mathtt{a}$ and $\mathtt{b}$ on the right-hand side!) Thus the charts $\{\tilde{\sigma}_{\mathtt{a}} \colon \pi^{-1}(U_{\mathtt{a}}) \to O_{\mathtt{a}} \times (\mathbb{R}^n)^* \mid \mathtt{a} \in \mathtt{A}\}$ then induce, thanks to Proposition 1.22, a smooth manifold structure on T^*M.

This computation and that done in the proof of Theorem 4.16 consequently summarise in the following sentence, somewhat dear to physicists: *while changing charts, (the local representatives of) vectors transform according to the Jacobian matrix of the transition map, while (the local representatives of) covectors transform according to the inverse-transposed Jacobian matrix of the transition map.*

Problem C.2. Let $\varphi \colon M \to N$ be a smooth map. Prove that $D\varphi \colon TM \to TN$ is also smooth. Prove that if $\varphi \colon M \to N$ is an embedding then so is $D\varphi \colon TM \to TN$.

Solution. To prove that the derivative of φ is a smooth map, let x be a point in M, let (U_M, σ_M) be a chart for M around x and let (U_N, σ_N) be a chart around $\varphi(x)$. If $(\pi_M^{-1}(U_M), \tilde{\sigma}_M)$ and $(\pi_N^{-1}(U_N), \tilde{\sigma}_M)$ are the associated charts for TM and TN respectively, the local expression for $D\varphi$ with respect to these charts is then

$$\tilde{\sigma}_N^{-1} \circ \varphi \circ \tilde{\sigma}_M^{-1}(x, v) = \big((\sigma_N^{-1} \circ \varphi \circ \sigma_M^{-1})(x), D(\sigma_N^{-1} \circ \varphi \circ \sigma_M^{-1}(x)[v]\big),$$

which, φ being smooth, is a smooth function of x and v. Hence $D\varphi$ is a smooth map.

Assume now that $\dim M = n$ and $\dim N = k$. If φ is an embedding, $\varphi(M)$ is an n-dimensional embedded submanifold of N, so in order to simplify the notations, we will show that if Σ is any embedded submanifold of N, then $T\Sigma$ is an embedded submanifold of TN. This will imply the result since $D\varphi$ defines a homeomorphism of TM onto $T(\varphi(M))$, where this latter is endowed with the induced topology of TN.

To say that a subset $\Sigma \subset N$ is an n-dimensional embedded submanifold is equivalent to say that, for every $x \in \Sigma$, there is a slice chart adapted to Σ, that is, a chart (U, σ) of N so that

$$\sigma(U \cap \Sigma) = \sigma(U) \cap (\mathbb{R}^n \times \{0, \ldots, 0\}), \tag{C.2}$$

where the number of zeroes is $k - n$. Any such slice chart induces a respective chart $(\pi_M^{-1}(U), \tilde{\sigma})$ for TN, and we claim that this is also a slice chart adapted to $T\Sigma$. Indeed, if x^i are the local coordinates relative to σ, we can write for any $v \in T_xN \subset \pi_M^{-1}(U)$,

$$\tilde{\sigma}(x, v) = \left(\sigma(x), \sum_{i=1}^{k} dx^i\big|_x(v) e_i\right),$$

but then condition (C.2) implies that the vector v_p is also tangent to Σ if and only if

$$0 = dx^{n+1}\big|_x(v) = dx^{n+2}\big|_x(v) = \ldots = dx^n\big|_x(v).$$

We obtain in other words that

$$\tilde{\sigma}(\pi_M^{-1}(U) \cap T\Sigma) = \big(\sigma(U) \cap (\mathbb{R}^n \times \{0, \ldots, 0\})\big) \times \mathbb{R}^n \times \{0, \ldots, 0\},$$

where each of the two chunks of zeroes contains $k-n$ elements. Up to a shift of the coordinates, this means precisely that we have obtained a slice chart adapted to $T\Sigma$. Since such chart can be produced around any $v \in T\Sigma$, we conclude that $T\Sigma$ is an embedded submanifold of TN by Proposition 5.10.

Problem C.3. Let M be a smooth manifold of dimension n. Let (x^i) be local coordinates on M, and let[2] $v^i := dx^i$ so that (x^i, v^i) are local coordinates on TM. Suppose $f\colon TM \to \mathbb{R}$ is a smooth function. Define the **fibrewise derivative** of f at the point (x,v) to be the element

$$D^{\text{fibre}} f(x,v) \in T_x^*M, \qquad D^{\text{fibre}} f(x,v) := \sum_{i=1}^n \frac{\partial}{\partial v^i}\bigg|_{(x,v)} (f) dx^i|_x.$$

Prove that $D^{\text{fibre}} f$ is well defined (i.e. independent of the choice of local coordinates).

Solution. Choose $(\tau,\beta) = (y^1,\dots,y^n,w^1,\dots,w^n)$ some local coordinates on TM just as in the statement above, i.e. y are local coordinates on M and $w^j = dy^j$, and denote by $\tilde\tau\colon TU \to \mathbb{R}^{2n}$ the corresponding chart on TM. We also denote $(\sigma,\alpha) = (x^1,\dots,x^n,v^1,\dots v^n)$ and let $\tilde\sigma$ be the corresponding chart in strict analogy to the above.

Observe that the expression $\frac{\partial}{\partial v^i}\big|_{(x,v)}$ is simply a vector in $T_{(x,v)}(TM)$ viewed as a derivation at (x,v). The tuple

$$\mathcal{B}_{(x,v)} := \left(\frac{\partial}{\partial y^1}\bigg|_{(x,v)}, \dots, \frac{\partial}{\partial y^n}\bigg|_{(x,v)}, \frac{\partial}{\partial w^1}\bigg|_{(x,v)}, \dots, \frac{\partial}{\partial w^n}\bigg|_{(x,v)} \right)$$

forms a basis of the tangent space $T_{(x,v)}(TM)$ and therefore we can express each $\frac{\partial}{\partial v^i}\big|_{(x,v)}$ in terms of $\mathcal{B}_{(x,v)}$ as follows:

$$\frac{\partial}{\partial v^i} = \sum_{j=1}^n \frac{\partial}{\partial v^i}(y^j)\frac{\partial}{\partial y^j} + \sum_{j=1}^n \frac{\partial}{\partial v^i}(w^j)\frac{\partial}{\partial w^j}.\text{[3]} \tag{C.3}$$

We compute the coefficients

$$\frac{\partial}{\partial v^i}(y^j) = D_{i+n}(y^j \circ \tilde\sigma^{-1}) = 0, \tag{C.4}$$

simply because $y^j \circ \tilde\sigma^{-1}\colon \mathbb{R}^{2n} \to \mathbb{R}$ is constant as a function in the $(n+i)$-th entry. This can be seen as follows: Let $z = (z_1,\dots,z_{2n}) \in \mathbb{R}^{2n}$ be an arbitrary vector, fix $x \in U \subseteq M$ and $v \in T_xM$ be the unique vector such that

$$\big(x^1(x),\dots,x^n(x),v^1(x,v)\dots,v^n(x,v)\big) = \tilde\sigma(x,v) = z.$$

[2] We introduce the new notation v^i because we want to think of them as coordinates on TM, and thus consider expressions such as $\frac{\partial}{\partial v^i}$. Without the new notation we would end up with an expression like $\frac{\partial}{\partial(dx^i)}$, which would be horrendously confusing...

[3] In order to keep the notation short we omit the basepoints from now on. Also, this expression follows from Remark (3.10)

We then observe that
$$y^j \circ \tilde{\sigma}^{-1}(z) = y^j(x),$$
which does not depend on z_{n+i} as we can leave x fixed and vary v to obtain any real value $z_{n+i} = v^i(x,v) = dx^i|_x(v) \in \mathbb{R}$. Therefore we conclude
$$D_{i+n}(y^j \circ \tilde{\sigma}^{-1})(\tilde{\sigma}(x,v)) = 0$$
as claimed above.

For the second coefficient we pick x, v and z as above. In view of Remark (4.13) we identify $\alpha(x,v) = D\sigma(x)[v]$ and $\beta(y,w) = D\tau(y)[w]$. We compute

$$\begin{aligned}(\beta \circ \tilde{\sigma}^{-1})z &= (\beta \circ (D\sigma(x))^{-1})z \\ &= D\tau(x) \circ D(\sigma^{-1})(\sigma(x))z \\ &= D(\tau \circ \sigma^{-1})(\sigma(x))z.\end{aligned}$$

This shows that the expression $(\beta \circ \tilde{\sigma}^{-1})$ is linear in $z \in \mathbb{R}^{2n}$ (for fixed x) and hence
$$D(\beta \circ \tilde{\sigma}^{-1})(\tilde{\sigma}(x,v)) = D(\tau \circ \sigma^{-1})(\sigma(x)).$$
From this we derive

$$\begin{aligned}\frac{\partial}{\partial v^i}\Big|_{(x,v)}(w^j) &= D_i(w^j \circ \tilde{\sigma}^{-1})(\tilde{\sigma}(x,v)) &\text{(C.5)}\\ &= u_j \circ D(\beta \circ \tilde{\sigma}^{-1})(\tilde{\sigma}(x,v))[e_i] \\ &= u_j \circ D(\tau \circ \sigma^{-1})(\sigma(x))[e_i] \\ &= D_i(y^j \circ \sigma^{-1}) \\ &= \frac{\partial}{\partial x^i}\Big|_x(y^j). &\text{(C.6)}\end{aligned}$$

Using (C.3), (C.4) and (C.5)=(C.6) we finally obtain

$$\begin{aligned}D^{\text{fibre}} f &= \sum_i \frac{\partial}{\partial v^i}(f) dx^i \\ &= \sum_{i,j} \frac{\partial}{\partial x^i}(y^j)\, \frac{\partial}{\partial w^j}(f) dx^i \\ &= \sum_j \frac{\partial}{\partial w^j}(f)\, \underbrace{\sum_i \frac{\partial}{\partial x^i}(y^j) dx^i}_{=dy^j} \\ &= \sum_j \frac{\partial}{\partial w^j}(f)\, dy^j.\end{aligned}$$

Problem C.4. Let $\varphi\colon M \to N$ be an injective immersion with M closed. Prove that φ is an embedding. Give an example to show that this need not be true if M is not closed.

SOLUTION. In order to prove that an injective immersion $\varphi\colon M \to N$ is an embedding if M closed we need to show that φ is a homeomorphism onto its image $\varphi(M)$, when the latter is equipped with he subspace topology it inherits form N. Clearly $\varphi\colon M \to \varphi(M)$ is bijective, so it suffices to show that it and its inverse are continuous. Given an open subset $O \subset \varphi(M)$ there exists an open subset $U \subset N$ such that

$$O = U \cap \varphi(M).$$

Here we simply use the definition of the subspace topology. Hence, $\varphi^{-1}(O) = \varphi^{-1}(U) \subset M$ is open because φ is assumed to be smooth (and in particular continuous) when viewed as a map $M \to N$. This shows that $\varphi\colon M \to \varphi(M)$ is continuous. To see that $\varphi^{-1}\colon \varphi(M) \to M$ is continuous it suffices to show that $(\varphi^{-1})^{-1}(C) = \varphi(C) \subset \varphi(M)$ is closed for every closed subset $C \subset M$. Since M is compact a closed subset $C \subset M$ is also compact (with the subspace topology). Since $\varphi\colon M \to \varphi(M)$ is continuous it maps compact subsets to compact subsets.[4] Hence $(\varphi^{-1})^{-1}(C) = \varphi(C) \subset \varphi(M)$ is closed for every closed subset $C \subset M$ and we conclude that $\varphi\colon M \to \varphi(M)$ is a homeomorphism which was exactly what we needed to show.[5]

The example we consider is that of $\mathbb{R}$ immersed (but not embedded!) into $\mathbb{R}^2$. The immersion $\varphi\colon \mathbb{R} \to \mathbb{R}^2$ takes the negative (infinite) end of $\mathbb{R}$ to the line $\{(0,y) \mid -1.5 < y < 1.5\}$, then "swipes around" and continuously into the curve $\{(t, \sin(\frac{1}{t}) \mid 0 < t < 1\}$ (see Figure C.1). It is not difficult to imagine that this can be done in such a way that φ is an immersion. Now note that $\varphi\colon \mathbb{R} \to \varphi(\mathbb{R})$ will not be a homeomorphism if the latter space is equipped with the subspace topology from $\mathbb{R}^2$. To see this, note that (in this topology) the point $(0,0) \in \varphi(\mathbb{R})$ has a basis of neighbourhoods all of which contain infinitely many connected components. This is not the case for any element in $\mathbb{R}$...

PROBLEM C.5. Let M and N be smooth manifolds. Prove that there is a canonical isomorphism:

$$T_{(x,y)}(M \times N) = T_xM \times T_yN, \qquad \forall\,(x,y) \in M \times N.$$

SOLUTION. Fix $(x,y) \in M \times N$ and let $\iota_M\colon M \to M \times N$, $\iota_M(x') = (x', y)$ denote the obvious inclusion and $\pi_M\colon M \times N \to M$ the obvious projection. Similarly we define $\iota_N\colon N \to M \times N, \iota_N(y') = (x, y')$, $\pi_N\colon M \times N \to N$ and define

$$T\colon T_xM \times T_yN \to T_{(x,y)}\left(M \times N\right), T(v,w) = D\iota_M(x)[v] + D\iota_N(y)[w].$$

To conclude that T is a canonical isomorphism we observe that the composition with $(D\pi_M(x,y), D\pi_N(x,y))$ is simply the identity:

$$\begin{aligned}(D\pi_M(x,y), D\pi_N(x,y)) \circ T(v,w) &= (D\pi_M(x,y), D\pi_N(x,y))\left(D\iota_M(x)[v] + D\iota_N(y)[w]\right)\\ &= \left(D(\underbrace{\pi_M \circ \iota_M}_{=\mathrm{id}})(x)[v], D(\underbrace{\pi_N \circ \iota_N}_{=\mathrm{id}})(y)[w]\right)\\ &= (v,w).\end{aligned}$$

[4]Check this if you don't remember the reason. Its not difficult...

[5]This proof never used the manifold structure. Indeed, it's a point-set topological fact that any continuous bijective map from a compact topological space into a Hausdorff space is automatically a homeomorphism.

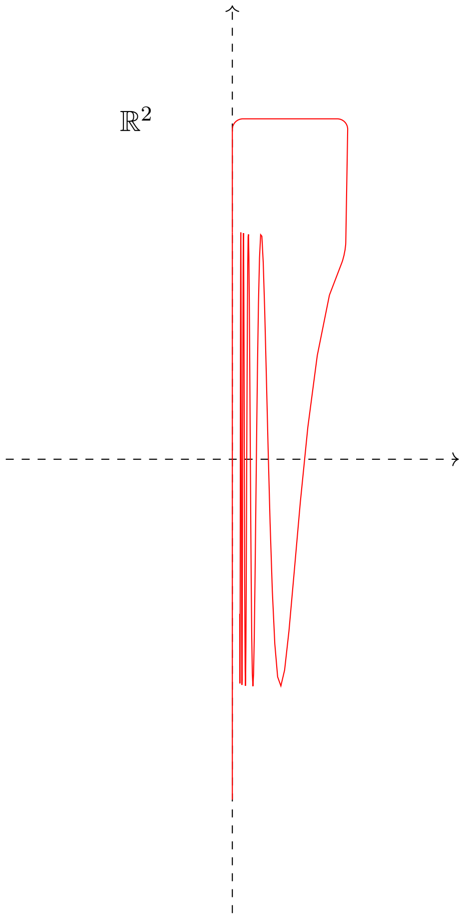

Figure C.1: A copy of $\mathbb{R}$ immersed in $\mathbb{R}^2$.

This forces T to be a linear isomorphism by dimension reasons[6] and also shows that T is canonical.

PROBLEM C.6. Let O be an open subset in $\mathbb{R}^n$ and suppose $f\colon O \to \mathbb{R}$ is smooth. Define $g\colon O \to \mathbb{R}^{n+1}$ by

$$g(x) = (x, f(x)).$$

Prove that g is an smooth embedding, and hence that $g(O)$ is a smooth n-dimensional submanifold of $\mathbb{R}^{n+1}$. We call g the **graph** of f.

SOLUTION. Define a map $\phi\colon O \times \mathbb{R} \to O \times \mathbb{R}$ by

$$\phi(x, y) = (x, y - f(x)).$$

Since f is smooth, it follows that ϕ is as well. Since the inverse

$$\phi^{-1}(u, v) = (u, v + f(u))$$

is also smooth, we deduce that ϕ is a diffeomorphism. We have $g = \iota \circ \phi^{-1}$, where $\iota\colon O \hookrightarrow O \times \mathbb{R}$ is the inclusion $x \mapsto (x, 0)$. Since ι and ϕ are both smooth embeddings, so is g. Since $\phi\big(g(O)\big) \cap (O \times \mathbb{R}) = \{(u, v) \in O \times \mathbb{R} \mid v = 0\}$, the map ϕ also defines an explicit slice chart for $g(O)$ in $O \times \mathbb{R}$.

[6] We do already know that $\dim(T_{(x,y)}(M \times N)) = \dim(M \times N) = \dim(T_x M \times T_y N)$.

Problem C.7. Let $\imath\colon S^n \hookrightarrow \mathbb{R}^{n+1}$ denote the inclusion. Prove that:

$$D\imath(x)[T_xS^n] = \mathcal{J}_x(x^\perp),$$

where $\mathcal{J}_x\colon \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ was defined in Problem B.3 and

$$x^\perp := \left\{ y \in \mathbb{R}^{n+1} \mid \langle x, y\rangle = 0 \right\},$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean dot product.

Solution. Previously we defined the isomorphism $\mathcal{J}_x : \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ by setting $\mathcal{J}_x(v) = \gamma'(0)$ with $\gamma(t) = x + tv$, where $\gamma'(0)$ is viewed as a linear derivation on the space of germs at x. What this is really saying is that the inverse $\mathcal{J}_x^{-1}\colon T_x\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ has the following simply description: Given any $v \in T_x\mathbb{R}^{n+1}$ there exists a smooth curve $\widetilde{\gamma}\colon (-\epsilon, \epsilon) \to \mathbb{R}^{n+1}$ such that $\widetilde{\gamma}(0) = x$ and $\widetilde{\gamma}'(0) = v$. Then

$$\mathcal{J}_x^{-1}(v) = \widetilde{\gamma}'(0),$$

where we view $\widetilde{\gamma}'(0)$ as an element of $\mathbb{R}^{n+1}$ in the sense of elementary calculus (differentiating each coordinate function).

Now fix $w \in T_xS^n$ and choose a smooth curve $\gamma\colon (-\epsilon, \epsilon) \to S^n$ with $\gamma(0) = x$ and $\gamma'(0) = w$. Since $S^n = \{y \in \mathbb{R}^{n+1} \mid |y| = 1\}$ we have we have $|\imath(\gamma(t))|^2 = 1$ for all $t \in (-\epsilon, \epsilon)$. From elementary calculus it now follows that

$$\begin{aligned} 0 &= \left.\frac{d}{dt}\right|_{t=0} |\imath(\gamma(t))|^2 = \left.\frac{d}{dt}\right|_{t=0} \langle \imath(\gamma(t)), \imath(\gamma(t))\rangle \\ &= \langle (\imath \circ \gamma)'(0), \imath(\gamma(0))\rangle + \langle \imath(\gamma(0)), (\imath \circ \gamma)'(0)\rangle \\ &= 2\langle (\imath \circ \gamma)'(0), x\rangle, \end{aligned}$$

where $(\imath \circ \gamma)'(0)$ is understood in the sense of elementary calculus. By the manifold version of the chain rule and the above remark we have $(\imath \circ \gamma)'(0) = \mathcal{J}_x^{-1} \circ D\imath(x)[w]$. So, since $w \in T_xS^n$ was arbitrary, the above computation implies

$$D\imath(x)[T_xS^n] \subset \mathcal{J}_x(x^\perp).$$

Moreover, since both $D\imath(x)[T_xS^n]$ and $\mathcal{J}_x(x^\perp)$ are linear subspaces of $T_x\mathbb{R}^{n+1}$ of dimension n it follows from elementary linear algebra that in fact

$$D\imath(x)[T_xS^n] = \mathcal{J}_x(x^\perp).$$

(♣) Problem C.8. Let M^n be an embedded submanifold of $\mathbb{R}^k$. We define the **normal space to M at x** to be the $(k-n)$-dimensional subspace $\mathrm{Norm}_x\, M \subset T_x\mathbb{R}^k$ consisting of all vectors that are orthogonal to M with respect to the Euclidean dot product. We define the **normal bundle** of M as the set

$$\mathrm{Norm}(M) := \left\{ (x, v) \in T\mathbb{R}^k = \mathbb{R}^k \times \mathbb{R}^k \mid x \in M, v \in \mathrm{Norm}_x\, M \right\}.$$

Prove that $\mathrm{Norm}(M)$ is an embedded k-dimensional submanifold of $T\mathbb{R}^k = \mathbb{R}^{2k}$.

Solution. Endow $\mathrm{Norm}(M) \subset \mathbb{R}^{2k}$ with the subspace topology. Let $x \in M$. Since M is an embedded submanifold of $\mathbb{R}^k$, there exists a slice chart $\sigma\colon U \subset \mathbb{R}^k \to O$ around x with corresponding coordinates $y^i := u^i \circ \sigma$ and

$$M \cap U = \{y \in U \mid y^{n+1}(y) = \cdots = y^k(y) = 0\}.$$

For each $y \in U$, the vectors $e_{j,y} := \partial/\partial y^j|_y$ form a basis of $T_y\mathbb{R}^k$, with $(e_{j,y})_{j=1,\ldots,n}$ forming a basis of the subspace $T_yM \subset T_y\mathbb{R}^k$. Let x^i denote the standard coordinates on $\mathbb{R}^k$. Then

$$e_{j,y} = \sum_{i=1}^{n} \frac{\partial}{\partial y^j}\bigg|_y (x^i) \cdot \frac{\partial}{\partial x^i}\bigg|_y.$$

Let $E(y) = \big(\partial/\partial y^j|_y(x^i)\big)_{ij}$ denote the corresponding change of basis matrix.

Define a function $\Sigma\colon U \times \mathbb{R}^k \to \sigma(U) \times \mathbb{R}^k$ by

$$\Sigma(y, v) = \big(y^1(y), \ldots, y^k(y), v \cdot e_{1,y}, \ldots, v \cdot e_{k,y}\big),$$

where $v \cdot e_{j,y}$ is the Euclidean dot product after identifying $T_y\mathbb{R}^k$ with $\mathbb{R}^k$ via the basis $(\partial/\partial x^i|_y)_{i=1,\ldots,k}$. This means that if $v = \sum_{i=1}^k v^i e^i$, then

$$v \cdot e_{j,y} = \sum_{i=1}^{k} v^i \frac{\partial}{\partial y^j}\bigg|_y (x^i).$$

Since as y varies each partial derivative $\partial/\partial y^j|_y(x^i)$ defines a smooth function on U, it follows that Σ is smooth. The total derivative of Σ at a point (y, v) is

$$D\Sigma(y, v) = \begin{pmatrix} \left(\frac{\partial y^i}{\partial x^j}(y)\right)_{ij} & 0 \\ * & E(y) \end{pmatrix}.$$

Since this is invertible, it follows from the Inverse Function Theorem that Σ is a local diffeomorphism. If $\Sigma(y, v) = \Sigma(y', v')$, then $y = y'$ because σ is injective. By assumption, we have $v \cdot e_{i,y} = v' \cdot e_{i,y}$ for each i, which implies that $v - v'$ is orthogonal to the span of $(e_{1,y}, \ldots, e_{k,y})$. Since the $e_{i,y}$ form a basis of $\mathbb{R}^k$, it follows that $v - v' = 0$. Thus Σ is injective and hence defines a smooth chart on $U \times \mathbb{R}^k$. By construction, we have $(y, v) \in \mathrm{Norm}(M)$ if and only if $\Sigma(y, v)$ is contained in

$$\{(z, w) \in \mathbb{R}^k \times \mathbb{R}^k \mid z^{n+1} = \cdots = z^k = 0, w^1 = \cdots = w^n = 0\}.$$

It follows that Σ is a slice chart for $\mathrm{Norm}(M)$. Proposition 5.10 implies that $\mathrm{Norm}(M)$ is an embedded k-dimensional submanifold of $\mathbb{R}^{2k}$, as desired.

Problem Sheet D

Problem D.1. Let M be a smooth manifold, let $x \in M$, and let $v \in T_xM$. Let U be any open set containing x. Prove that there exists a vector field $X \in \mathfrak{X}(U)$ such that $X(x) = v$.

Problem D.2. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Prove that the Lie bracket $[\cdot, \cdot]$ on $\mathfrak{X}(W)$ satisfies the Jacobi identity.

Problem D.3. Let M be a smooth manifold and let $\sigma \colon U \to O$ be a chart on M with local coordinates x^i, and let $X, Y \in \mathfrak{X}(U)$. Write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$. Prove that

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j},$$

where $\frac{\partial Y^j}{\partial x^i}$ and $\frac{\partial X^j}{\partial x^i}$ are the functions from Definition 7.4.

Problem D.4. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Let $X, Y \in \mathfrak{X}(W)$, and let $f, g \in C^\infty(W)$. Prove that

$$[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X.$$

Problem D.5. Let $\varphi \colon M \to N$ be a smooth map. Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$. We say that X and Y are **φ-related** if

$$D\varphi(x)[X(x)] = Y(\varphi(x)), \qquad \forall\, x \in M.$$

Of course if φ is a diffeomorphism then any $X \in \mathfrak{X}(M)$ is φ-related to $\varphi_\star(X)$.

(i) Prove that $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are φ-related if and only if for every open set $V \subset N$ and every smooth function $f \in C^\infty(V)$, one has

$$X(f \circ \varphi) = Y(f) \circ \varphi.$$

(ii) Let $X_i \in \mathfrak{X}(M)$ and $Y_i \in \mathfrak{X}(N)$ for $i = 1, 2$ be vector fields. Assume X_i is φ-related to Y_i for each $i = 1, 2$. Prove that $[X_1, X_2]$ is φ-related to $[Y_1, Y_2]$.

(♣) Problem D.6. Let $M \subset N$ be an (immersed or embedded) submanifold and let $x \in M$. We say that a vector field $Y \in \mathfrak{X}(N)$ is **tangent to M at x** if $Y(x) \in T_xM \subset T_xN$. We say Y is **tangent to M** if it is tangent to M at every point $x \in M$.

(i) Assume $M \subset N$ is an embedded submanifold. Prove that $Y \in \mathfrak{X}(N)$ is tangent to M if and only if $(Yf)|_M \equiv 0$ for every function $f \in C^\infty(N)$ such that $f|_M \equiv 0$.

(ii) Now assume $M \subset N$ is merely an immersed submanifold. Let $\imath \colon M \hookrightarrow N$ denote the inclusion. Assume that $Y \in \mathfrak{X}(N)$ is tangent to M. Prove there exists a unique $X \in \mathfrak{X}(M)$ such that X is $\imath$-related to Y.

(iii) Continue to assume that $M \subset N$ is an immersed submanifold. Suppose $Y_1, Y_2 \in \mathfrak{X}(N)$ are tangent to M. Prove that $[Y_1, Y_2]$ is tangent to M.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Solutions to Problem Sheet D

Problem D.1. Let M be a smooth manifold, let $x \in M$, and let $v \in T_xM$. Let U be any open set containing x. Prove that there exists a vector field $X \in \mathfrak{X}(U)$ such that $X(x) = v$.

Solution. First consider an arbitrary subset $A \subset M$ endowed with the subspace topology. A **smooth vector field along** A is a continuous map $X \colon A \to TM$ such that

1. the section property $\pi(X(x)) = x$ holds for all $x \in A$, and
2. for each $x \in A$, there is a neighbourhood V of x in M and a smooth vector field $\tilde{X}$ on V that agrees with X on $V \cap A$.

The statement in the exercise follows from the following more general statement:

Lemma. Let $A \subset M$ be a closed subset, and let X be a smooth vector field along A. Given any open subset $U \subset M$ containing A, there exists a smooth global vector field $\tilde{X}$ on M such that $\tilde{X}|_A = X$ and supp $\tilde{X} \subset U$.

Proof. By assumption, for each $x \in A$, we may choose a neighbourhood $V_x \subset U$ of x and a smooth vector field $\tilde{X}_x$ of V_x that agrees with X on $V_x \cap A$. Then the family of sets $\{V_x \mid x \in A\} \cup \{M \setminus A\}$ is an open cover of M. Let $\{\lambda_x \mid x \in A\} \cup \{\lambda_0\}$ be a partition of unity subordinate to this open cover, with supp $\lambda_x \subset V_x$ and supp $\lambda_0 \subset M \setminus A$.

For each $x \in A$, the product $\lambda_x \tilde{X}_x$ is a smooth vector field on V_x. We consider it as a smooth vector field on M by extending by 0 on $M \setminus \operatorname{supp} \lambda_x$. Consider the function $\tilde{X} \colon M \to TM$ defined by

$$\tilde{X}(x) = \sum_{y \in A} \lambda_y(x) \tilde{X}_y(x).$$

Each term in the above sum is smooth. Since $\{\operatorname{supp} \lambda_x \mid x \in A\}$ is locally finite, the sum has only finitely many terms in a neighbourhood of any point of M and thus defines a smooth function. If $x \in A$, then $\lambda_0(x) = 0$ and $\lambda_y(x)\tilde{X}_y(x) = X(x)$ for each y such that $\lambda_y(x) \neq 0$. Thus

$$\tilde{X}(x) = \sum_{y \in A} \lambda_y(x) X(x) = \Big(\lambda_0(x) + \sum_{y \in A} \lambda_y(x)\Big) X(x) = X(x);$$

hence $\tilde{X}$ is a global smooth vector field extending X. Finally, consider the following equalities:

$$\operatorname{supp} \tilde{X} = \overline{\bigcup_{x \in A} \operatorname{supp} \lambda_x} = \bigcup_{x \in A} \operatorname{supp} \lambda_x \subset U.$$

Thus supp $\tilde{X} \subset U$. This concludes the proof. ∎

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

The function $X\colon \{x\} \to TM,\ x \mapsto v$ is a smooth vector field along $\{x\}$ since it can be extended, for example, to a constant-coefficient vector field in a coordinate neighbourhood of x. We thus obtain the desired result by applying the above proposition with $A = \{x\}$ and $U = M$.

Note that in order to prove the statement in the exercise, one can simply extend to a constant coefficient vector field in a coordinate neighbourhood U of x with value v at x, multiply by a bump function on U, and then extend the resulting vector field on U by zero to a vector field on M. We include the general statement above as a nice illustration of the usefulness of partitions of unity, and one can see that its proof generalises the aforementioned simpler argument.

PROBLEM D.2. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Prove that the Lie bracket $[\cdot,\cdot]$ on $\mathfrak{X}(W)$ satisfies the Jacobi identity.

SOLUTION. Let $X, Y, Z \in \mathfrak{X}(W)$ and let $f \in C^\infty(W)$. We compute:

$$\begin{aligned}
[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] =& X[Y,Z] - [Y,Z]X + Y[Z,X] \\
&- [Z,X]Y + Z[X,Y] - [X,Y]Z \\
=& XYZ - XZY - YZX + ZYX + YZX \\
&- ZXY + XZY + ZXY - ZYX - XYZ + YXZ.
\end{aligned}$$

Note that the product of vector fields in the above equations is given by composition, i.e., $XY := X \circ Y$, where X and Y are viewed as functions from $C^\infty(W)$ to itself. In the last expression above, all of the terms cancel in pairs. Thus $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ and $\mathfrak{X}(W)$ satisfies the Jacobi identity.

PROBLEM D.3. Let M be a smooth manifold and let $\sigma\colon U \to O$ be a chart on M with local coordinates x^i, and let $X, Y \in \mathfrak{X}(U)$. Write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$. Prove that

$$[X,Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j},$$

where $\frac{\partial Y^j}{\partial x^i}$ and $\frac{\partial X^j}{\partial x^i}$ are the functions from Definition (7.4).

SOLUTION. We recall that the Lie bracket $[X,Y]$ of two vector fields $X, Y \in \mathfrak{X}(U)$ is a derivation defined by $[X,Y]\colon C^\infty(U) \to C^\infty(U)$, $[X,Y](f) = X(Y(f)) - Y(X(f))$. Let us write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^j \frac{\partial}{\partial x^j}$ and view them both as derivations. We pick some arbitrary smooth function $f \in C^\infty(U)$ and feed it to Y to obtain, for each $x \in U$:

$$Y(f)(x) = \left(Y^j \frac{\partial}{\partial x^j} \right)(f)(x) = Y^j(x)\, \frac{\partial f}{\partial x^j}(x).$$

Note that in the last step we used the notation introduced in Definition (7.4). For notations reasons it is convenient to set g_j[1] $= \frac{\partial f}{\partial x^j} \in C^\infty(U)$. Invoking the derivation property grants us

$$\begin{aligned}
X(Y(f))(x) &= X^i(x) \frac{\partial}{\partial x^i} \left(Y^j\, g_j \right)(x) && \text{(D.1)} \\
&= X^i(x) \left(\frac{\partial Y^j}{\partial x^i}(x)\, g_j(x) + \frac{\partial g_j}{\partial x^i}(x)\, Y^j(x) \right), && \text{(D.2)}
\end{aligned}$$

[1]Since we are using the Einstein Summation Convention we are forced to use a lower subscript.

and by symmetry, after swapping the i's and j's indices, we also get

$$Y(X(f))(x) = Y^i(x)\frac{\partial}{\partial x^i}\left(X^j\, g_j\right)(x) \tag{D.3}$$

$$= Y^i(x)\left(\frac{\partial X^j}{\partial x^i}(x)\, g_j(x) + \frac{\partial g_j}{\partial x^i}(x)\, X^j(x)\right). \tag{D.4}$$

Now we compute the difference (D.2)-(D.4):

$$X(Y(f)) - Y(X(f)) = \left(X^i\,\frac{\partial Y^j}{\partial x^i}\,\frac{\partial f}{\partial x^j} - Y^i\,\frac{\partial X^j}{\partial x^i}\,\frac{\partial f}{\partial x^j}\right) + \left(\underbrace{X^i\,\frac{\partial g_j}{\partial x^i}\,Y^j - Y^i\,\frac{\partial g_j}{\partial x^i}\,X^j}_{=0}\right).$$

This concludes the proof.

Problem D.4. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Let $X, Y \in \mathfrak{X}(W)$, and let $f, g \in C^\infty(W)$. Prove that

$$[fX, gY] = fg[X,Y] + fX(g)Y - gY(f)X.$$

Solution. In order to make the next computation more readable we will use dots "$\cdot$" to emphasise that we are taking the product of two smooth functions on W. This has the advantage that it becomes very clear where the derivation property (i.e. the Leibniz-rule) is used.

We pick some $h \in C^\infty(W)$ and finish the proof with the following computation:

$$\begin{aligned}
[f\,X, g\,Y]\,(h) &= f\cdot X\left((g\,Y)(h)\right) - g\cdot Y\left((f\,X)(h)\right)\\
&= f\cdot X\left(g\cdot Y(h)\right) - g\cdot Y\left(f\cdot X(h)\right)\\
&= f\cdot\{X(g)\cdot Y(h) + g\cdot X(Y(h))\} - g\cdot\{Y(f)\cdot X(h) + f\cdot Y(X(h))\}\\
&= (f\cdot g)\left([X,Y]\,(h)\right) + f\cdot X(g)\cdot Y(h) - g\cdot Y(f)\cdot X(h).
\end{aligned}$$

Problem D.5. Let $\varphi\colon M \to N$ be a smooth map. Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$. We say that X and Y are **φ-related** if

$$D\varphi(x)[X(x)] = Y(\varphi(x)), \qquad \forall\, x \in M.$$

Of course if φ is a diffeomorphism then any $X \in \mathfrak{X}(M)$ is φ-related to $\varphi_\star(X)$.

(i) Prove that $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ are φ-related if and only if for every open set $V \subset N$ and every smooth function $f \in C^\infty(V)$, one has

$$X(f\circ\varphi) = Y(f)\circ\varphi.$$

(ii) Let $X_i \in \mathfrak{X}(M)$ and $Y_i \in \mathfrak{X}(N)$ for $i = 1,2$ be vector fields. Assume X_i is φ-related to Y_i for each $i = 1,2$. Prove that $[X_1, X_2]$ is φ-related to $[Y_1, Y_2]$.

Solution.

(i) Recall that the derivative of φ satisfies, for every x in M and every neighbourhood $V \subset N$ of $\varphi(x)$,

$$X(f \circ \varphi)(x) = D\varphi(x)[X(x)](f) \quad \text{for every } f \in C^\infty(V).$$

Since two tangent vectors $W_1, W_2 \in T_{\varphi(x)}N$ coincide if and only if

$$W_1(f) = W_2(f) \quad \text{for every } f \in C^\infty(V),$$

where $V \subset N$ is any (equiv. some) neighbourhood of $\varphi(x)$, we deduce that the vector field Y is φ-related to X if and only if $Y(f) \circ \varphi(x) = X(f \circ \varphi)(x)$ for every $x \in M$ and $f \in C^\infty(V)$, where V is any (equiv. some) neighbourhood of $\varphi(x)$ in N.

(ii) By (i) it suffices to prove that, for every open set $V \subset N$ (so that $V \cap \varphi(M) \neq \emptyset$, of course) and every $f \in C^\infty(V)$,

$$[X_1, X_2](f \circ \varphi) = [Y_1, Y_2](f) \circ \varphi.$$

Since Y_i is φ-related to X_i we see that

$$X_1(X_2(f \circ \varphi)) = X_1(Y_2(f) \circ \varphi) = Y_1(Y_2(f)) \circ \varphi,$$

and similarly we see that $X_2(X_1(f \circ \varphi)) = Y_2(Y_1(f)) \circ \varphi$. Consequently we deduce that

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

as desired.

(♣) Problem D.6. Let $M \subset N$ be an (immersed or embedded) submanifold and let $x \in M$. We say that a vector field $Y \in \mathfrak{X}(N)$ is **tangent to** M **at** x if $Y(x) \in T_xM \subset T_xN$. We say Y is **tangent to** M if it is tangent to M at every point $x \in M$.

(i) Assume $M \subset N$ is an embedded submanifold. Prove that $Y \in \mathfrak{X}(N)$ is tangent to M if and only if $(Yf)|_M \equiv 0$ for every function $f \in C^\infty(N)$ such that $f|_M \equiv 0$.

Solution. Let's define $n = \dim M$ and $k := \dim N$ so that $k \geq n$. We apply Proposition 5.6, according to which there exists a chart $\varphi\colon U \to O$ on N for every $x \in M$ with $x \in U \overset{\text{open}}{\subset} N$ and $0 = \varphi(x) \in O \overset{\text{open}}{\subset} \mathbb{R}^k$ such that

$$\varphi(M \cap U) = O \cap (\mathbb{R}^n \times \{0\}^{k-n}).$$

We will sometime use the notation $y = (z_1, z_2) \in \mathbb{R}^n \times \mathbb{R}^{k-n} = \mathbb{R}^k$. Now let $Y^\varphi \in \mathfrak{X}(O)$ be the pushforward of Y:

$$Y^\varphi(y) = D\varphi(x)[Y(x)],$$

where $y = \varphi(x)$. We can expand Y^φ in the basis $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}$:

$$Y^\varphi(y) = Z^i(y) \left.\frac{\partial}{\partial x^i}\right|_y \quad \forall\, y \in O.$$

Here we use the Einstein summation convention and $Z^i \in C^\infty(O)$ for every $i = 1, \dots, k$. Clearly Y being tangent to M is equivalent to $Z^i(z_1, 0) = 0$ for all $(z_1, 0) \in O$ and every $i = n+1, \dots, k$.

Assume first that Y is tangent to M, i.e.

$$Z^i(z_1, 0) = 0 \quad \forall\, (z_1, 0) \in O,\ i = n+1, \dots, k. \tag{D.5}$$

If $f \in C^\infty(N)$ satisfies $f|_M \equiv 0$ then

$$f \circ \varphi^{-1}(z_1, 0) = 0 \quad \forall\, (z_1, 0) \in O. \tag{D.6}$$

In particular we can compute that for $x \in M$ and $\varphi(x) = y = (z_1, 0)$:

$$\begin{aligned} Y(f)|_x &= Y^\varphi(f \circ \varphi^{-1})|_{(z_1,0)} \\ &= \sum_{i=1}^{n} Z^i(z_1, 0) \left.\frac{\partial (f \circ \varphi^{-1})}{\partial x^i}\right|_{(z_1,0)} + \sum_{i=n+1}^{k} Z^i(z_1, 0) \left.\frac{\partial (f \circ \varphi^{-1})}{\partial x^i}\right|_{(z_1,0)} \\ &= 0 \end{aligned}$$

The first sum vanishes because of condition (D.6) and the second vanishes because of condition (D.5).

Assume now that $Y(f)|_M \equiv 0$ for every $f \in C^\infty(N)$ with $f|_M \equiv 0$. We need to show that Y is tangent to M. We do so by contradiction. Hence, assume for contradiction that Y were not tangent to M. Then there exists some chart φ as above about x and some $l \in \{m+1, \dots, n\}$ such that

$$Z^l(\varphi(x)) \neq 0.$$

Now define $\widetilde{f}(x_1, \dots, x_n) = x_l$. By Lemma 3.2 there exists a compactly supported function $\eta \in C^\infty(O)$ such that $\mathrm{supp}(\eta) \subset O$ and $\eta \equiv 1$ on $\overline{B} \subset O$, where B is a small open ball in O containing $0 = \varphi(x) \in O$. The function

$$f(y) := \eta(y) \widetilde{f}(y)$$

is a compactly supported function on O, so by extending $f \circ \varphi|_U$ by 0 outside of U we can view $f \circ \varphi \in C^\infty(N)$. Note that $f \circ \varphi|_M \equiv 0$ because $\widetilde{f}|_{\mathbb{R}^{n-m} \times \{0\}^m} \equiv 0$. Now we can compute

$$\begin{aligned} Y(f \circ \varphi)(x) = Y^\varphi(f)(\varphi(x)) &= Z^i(\varphi(x)) \left.\frac{\partial f}{\partial x^i}\right|_{\varphi(x)} \\ &= Z^i(\varphi(x)) \left.\frac{\partial x^l}{\partial x^i}\right|_{\varphi(x)} = Z^l(\varphi(x)) \neq 0. \end{aligned}$$

In this computation we use that $f = x_l$ on a neighbourhood of x. This contradicts our assumption that $Y(f)|_M \equiv 0$ for all $f \in C^\infty(N)$ with $f|_M \equiv 0$ and hence finishes the proof.

(ii) Now assume $M \subset N$ is merely an immersed submanifold. Let $\imath\colon M \hookrightarrow N$ denote the inclusion. Assume that $Y \in \mathfrak{X}(N)$ is tangent to M. Prove there exists a unique $X \in \mathfrak{X}(M)$ such that X is $\imath$-related to Y.

Solution. That $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$ be $\imath$-related means that

$$D\imath(x)[X(x)] = Y(\imath(x)) \quad \forall\, x \in M. \tag{D.7}$$

Since $\imath\colon M \to N$ is an immersion, $D\imath(x)\colon T_xM \to T_{\imath(x)}N$ is injective for every $x \in M$. The assumption that Y is tangent to M means that $Y(\imath(x)) \in \mathrm{Image}(D\imath(x))$ for every $x \in M$. Hence, (D.7) defines X pointwise uniquely. That X is smooth (i.e. $X \in \mathfrak{X}(M)$) follows from the Implicit Function Theorem 5.3.

(iii) Continue to assume that $M \subset N$ is an immersed submanifold. Suppose $Y_1, Y_2 \in \mathfrak{X}(N)$ are tangent to M. Prove that $[Y_1, Y_2]$ is tangent to M.

Solution. $[Y_1, Y_2]$ being tangent to M is clearly a local property: We need to show that for a given $x \in M \subset N$ we have $[Y_1, Y_2](x) \in T_xM$. Since any immersion is a local embedding, there exists a neighbourhood $O \subset M$ of x such that $O \subset N$ is an *embedded* submanifold. Since $T_xO = T_xM$ it suffices to show that $[Y_1, Y_2](x) \in T_xO$. Given $f \in C^\infty(N)$ with $f|_O \equiv 0$ we compute using (i) and the fact that Y_1 and Y_2 are tangent to O:

$$\begin{aligned}
[Y_1, Y_2](f)(x) &= Y_1(Y_2(f))(x) - Y_2(Y_1(f))(x) \\
&= Y_1(0)(x) - Y_2(0)(x) = 0
\end{aligned}$$

This shows that $[Y_1, Y_2](x) \in T_xO$ and thus finishes the exercise.

Problem Sheet E

Problem E.1. Let $\varphi\colon M \to N$ be a smooth map between two smooth manifolds. Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$, and assume X and Y are φ-related in the sense of Problem D.5. Let θ^X_t and θ^Y_t denote the respective flows, and define $M_t \subset M$ and $N_t \subset N$ as Remark 8.11. Prove that $\varphi(M_t) \subset N_t$ and that

$$\theta^Y_t \circ \varphi = \varphi \circ \theta^X_t, \qquad \text{on } M_t.$$

Deduce that if φ is a diffeomorphism then for any vector field X on M one has $\theta^{\varphi_\star(X)}_t = \varphi \circ \theta^X_t \circ \varphi^{-1}$.

(♣) Problem E.2. Let X and Y be vector fields on a smooth manifold M with flows θ^X_t and θ^Y_t respectively. Prove that $[X, Y] \equiv 0$ if and only if the two flows commute, i.e. $\theta^X_t \circ \theta^Y_s = \theta^Y_s \circ \theta^X_t$ for all s, t small.

Problem E.3. Let $J_0 \in \mathrm{Mat}(2n)$ denote the matrix

$$J_0 := \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix},$$

where I is the $n \times n$ identity matrix. The **symplectic linear group** $\mathrm{Sp}(2n)$ consists of the matrices A such that $A^T J_0 A = J_0$. Prove that $\mathrm{Sp}(2n)$ is a Lie group. Compute its dimension, and compute its Lie algebra $\mathfrak{sp}(2n)$.

(♣) Problem E.4. Let G be a Lie group and suppose H is a subgroup of G which is also an embedded submanifold. Prove that H is closed in G (as a subspace).

Problem E.5. Prove that if Lie group is abelian then its Lie algebra is abelian. (*Remark*: On Problem Sheet G you will prove that if a *connected* Lie group has abelian Lie algebra, then it is an abelian Lie group.)

Problem E.6. Prove that the Lie bracket on $\mathfrak{gl}(n)$ is given by matrix commutation, i.e.

$$[A, B] = AB - BA, \qquad \forall\, A, B \in \mathfrak{gl}(n) = \mathrm{Mat}(n).$$

(♣) Problem E.7. Let $\varphi\colon M^n \to N^k$ be smooth, and let $L^r \subset N$ be an embedded submanifold. We say that φ is **transverse** and **regular** at L if

$$D\varphi(x)[T_x M] + T_{\varphi(x)} L = T_{\varphi(x)} N, \qquad \forall\, x \in \varphi^{-1}(L).$$

Prove that if φ is transverse and regular at P then if $\varphi^{-1}(L) \neq \emptyset$ then $\varphi^{-1}(L)$ is a smooth embedded submanifold of M of dimension $n - k + r$. (*Remark*: The Implicit Function Theorem 5.13 is the special case where L is a point. As a hint, try to reduce this problem to Theorem 5.13.)

Solutions to Problem Sheet E

Problem E.1. Let $\varphi\colon M \to N$ be a smooth map between two smooth manifolds. Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(N)$, and assume X and Y are φ-related in the sense of Problem D.5. Let θ_t^X and θ_t^Y denote the respective flows, and define $M_t \subset M$ and $N_t \subset N$ as Remark 8.11. Prove that $\varphi(M_t) \subset N_t$ and that

$$\theta_t^Y \circ \varphi = \varphi \circ \theta_t^X, \qquad \text{on } M_t.$$

Deduce that if φ is a diffeomorphism then for any vector field X on M one has $\theta_t^{\varphi_\star(X)} = \varphi \circ \theta_t^X \circ \varphi^{-1}$.

Solution. Recall that, by definition, a point x belongs to M_t when the solution of the Cauchy problem

$$\begin{cases} \gamma_x'(\tau) = X(\gamma_x(\tau)), \\ \gamma_x(0) = x, \end{cases}$$

is defined for $\tau = t$. If we want to prove that $\varphi(x)$ belongs to N_t we need to show that the solution of the Cauchy problem

$$\begin{cases} \zeta_x'(\tau) = Y(\zeta_x(\tau)), \\ \zeta_x(0) = \varphi(x), \end{cases} \tag{E.1}$$

is defined for $\tau = t$. We claim that the solution to such problem is the curve $\zeta_x(\tau) = \varphi(\gamma_x(\tau))$. Indeed, there holds $\zeta_x(0) = \varphi(\gamma_x(0)) = \varphi(x)$, and since X and Y are φ-related, we can compute:

$$\zeta_x'(\tau) = D\varphi(\gamma_x(\tau))[X(\gamma_x(\tau))] = Y(\varphi(\gamma_x(\tau))) = Y(\zeta_x(\tau)).$$

Hence η is the (unique) solution of (E.1), and, since γ_x is defined at $\tau = t$, so is ζ_x. Note moreover, by definition of the flows θ_t^X and θ_t^Y, we simply have

$$\theta_t^X(x) = \gamma_x(t) \quad \text{and} \quad \theta_t^Y(\varphi(x)) = \zeta_x(t),$$

hence since $\zeta_x(t) = \varphi(\gamma_x(t))$, we also showed that there holds

$$(\theta_t^Y \circ \varphi)(x) = (\varphi \circ \theta_t^X)(x) \quad \text{for } x \in M_t.$$

Finally note that if φ is a diffeomphism, we can pre-compose on the left both hand-sides of the above relation with φ^{-1} and notice that Y is φ-related to X precisely if $Y = \varphi_\star(X)$.

(♣) Problem E.2. Let X and Y be vector fields on a smooth manifold M with flows θ_t^X and θ_t^Y respectively. Prove that $[X, Y] \equiv 0$ if and only if the two flows commute, i.e. $\theta_t^X \circ \theta_s^Y = \theta_s^Y \circ \theta_t^X$ for all s, t small.

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

SOLUTION. Recall first of all the fundamental identity

$$[X,Y](x) = \mathcal{L}_X Y(x) = \frac{d}{d\varepsilon}\Big(D(\theta^X_{-\varepsilon})(\theta^X_\varepsilon(x))[Y(\theta^X_\varepsilon(x))]\Big)\Big|_{\varepsilon=0} \quad \text{for } x \in M. \tag{E.2}$$

If the flows of X and Y commute, we deduce that Y *is X-invariant*, that is, for every $x \in M$ and every sufficiently small values s, t so that the expressions below are defined, we have

$$D(\theta^X_s)(x)[Y(x)] = \frac{d}{dt}(\theta^X_s \circ \theta^Y_t)(x)\Big|_{t=0} = \frac{d}{dt}(\theta^Y_t \circ \theta^X_s)(x)\Big|_{t=0} = Y(\theta^X_s(x)).$$

In particular, for any $p \in M$, if ε is sufficiently small we may choose $x = \theta^X_{-\varepsilon}(p)$ and thus deduce that

$$Y(p) = D(\theta^X_{-\varepsilon})(\theta^X_\varepsilon(p))[Y(\theta^X_\varepsilon(p)].$$

This means that that the vector field on the right-hand-side is independent of ε. Since the derivative of a constant function is 0, from (E.2) we deduce that $[X,Y] = 0$.

Conversely, let us assume that the bracket of X and Y vanishes. Let us prove first that Y is X-invariant, that is, for every $x \in M$ there hols

$$(\theta^X_\varepsilon)(x)[Y(x)] = Y(\theta^X_\varepsilon(x)),$$

for every sufficiently small value of ε so that this expression is defined. Replacing ε with $-\varepsilon$ and setting $p = \theta^X_\varepsilon(x)$, this is equivalent to prove that the vector field

$$D(\theta^X_{-\varepsilon})(\theta^X_\varepsilon(p))[Y(\theta^X_\varepsilon(p))]$$

is constant with respect to ε. In the computation that follows, we use the so-called semigroup property of the flow:

$$\theta^X_{\tau+\eta}(p) = \theta^X_\tau(\theta^X_\eta(p)),$$

which implies, when differentiated and applied to Y, that

$$D(\theta^X_{\tau+\eta})(p)[Y(p)] = D(\theta^X_\tau)(\theta^X_\eta(p))[D\theta^X_\eta(p)[Y(p)]],$$

for every τ, η so that the above expressions are defined. We may then compute:

$$\begin{aligned}
&\frac{d}{d\varepsilon}\Big(D(\theta^X_{-\varepsilon})(\theta^X_\varepsilon(p))[Y(\theta^X_\varepsilon(p))]\Big)\Big|_{\varepsilon=\varepsilon_0} \\
&= \frac{d}{d\varepsilon}\Big(D(\theta^X_{-\varepsilon-\varepsilon_0})(\theta^X_{\varepsilon+\varepsilon_0}(p))[Y(\theta^X_{\varepsilon+\varepsilon_0}(p))]\Big)\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\Big(D(\theta^X_{-\varepsilon_0})(\theta^X_{\varepsilon_0}(p))\big[D(\theta^X_{-\varepsilon})(\theta^X_{-\varepsilon+\varepsilon_0}(p))[Y(\theta^X_{\varepsilon+\varepsilon_0}(p))]\big]\Big)\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\Big(D(\theta^X_{\varepsilon_0}(\theta^X_\varepsilon(p))[\mathcal{L}_X Y(\theta^X_{\varepsilon_0}(p))]\Big)\Big|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\Big(D(\theta^X_{\varepsilon_0}(\theta^X_\varepsilon(p))\big[[X,Y](\theta^X_{\varepsilon_0}(p))\big]\Big)\Big|_{\varepsilon=0} = 0,
\end{aligned}$$

and hence Y is X-invariant. As a consequence, if for fixed t and p we consider the curve $\sigma(s) = (\theta_t^X \circ \theta_s^X)(p)$, differentiating σ it with respect to s yields

$$\begin{aligned}\sigma'(s) &= D(\theta_t^X)(\theta_s^Y(p))[Y(\theta_s^Y(p))] \\ &= Y(\theta_t^X(\theta_s^Y(p))) \\ &= Y(\sigma(s)),\end{aligned}$$

and since $\sigma(0) = \theta_s^X(p)$, we deduce that σ is the integral curve for Y at the point $\theta_t^X(p)$, that is $\sigma(s) = \theta_s^Y(\theta_t^X(p))$, and from the definition of σ, this means precisely that tha flows of X and Y commute.

PROBLEM E.3. Let $J_0 \in \mathrm{Mat}(2n)$ denote the matrix

$$J_0 := \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix},$$

where I is the $n \times n$ identity matrix. The **symplectic linear group** $\mathrm{Sp}(2n)$ consists of the matrices A such that $A^T J_0 A = J_0$. Prove that $\mathrm{Sp}(2n)$ is a Lie group. Compute its dimension, and compute its Lie algebra $\mathfrak{sp}(2n)$.

SOLUTION. One way to show that the symplectic linear group defines a Lie group is by expressing it in terms of a regular level of some smooth function[1]. First of all we observe that the matrix $A^T J_0 A$ is antisymmetric, i.e. its transpose is equal to minus itself. In other words, the expression $A^T J_0 A$ lives inside the vector space

$$\mathrm{Antsym}(2n) = \left\{ B \in \mathrm{Mat}(2n) \middle| \, B^T = -B \right\}.$$

It is straightforward to see that the (real) vector space dimension is given by

$$\dim_{\mathbb{R}} \mathrm{Antsym}(2n) = \frac{2n(2n+1)}{2} - 2n = \frac{2n(2n-1)}{2}.$$

We now define the smooth function

$$\varphi\colon \mathrm{Mat}(2n) \to \mathrm{Antsym}(2n),\ A \mapsto A^T J_0 A.$$

In view of $\varphi^{-1}(J_0) = \mathrm{Sp}(2n)$, we want to show that J_0 is a regular value of φ so that $\mathrm{Sp}(2n)$ inherits a smooth manifold structure by means of the Implicit Function Theorem, more precisely Theorem (5.2).

Let $A_0 \in \varphi^{-1}(J_0)$, $B \in \mathrm{Mat}(2n)$ and pick $A(t)$ a smooth path in $\mathrm{Mat}(2n)$ such that

$$A(0) = A_0,\ \frac{d}{dt}\Big|_{t=0} A(t) = B.$$

Using the chain rule we can compute the differential of φ at A_0 evaluated on B:

$$\begin{aligned}D\varphi(A_0)[B] &= \frac{d}{dt}\Big|_{t=0} (\varphi \circ A)(t) \\ &= \frac{d}{dt}\Big|_{t=0} \left(A^T(t) J_0 A(t)\right)(t) \\ &= B^T J_0 A_0 + A_0^T J_0 B.\end{aligned}$$

[1]This is very similar to the proof strategy used in Proposition (9.13)

For any antisymmetric matrix C we can now define $B = -\frac{1}{2}A_0 J_0 C$ and plug it into the differential above. Using the fact that $A_0 = A_0^T J_0 A_0$ one then obtains

$$D\varphi(A_0)[B] = C,$$

which shows that $D\varphi(A_0)$ is surjective[2] for any element A_0 in the preimage $\varphi^{-1}(J_0)$ and equivalently that J_0 is a regular value of φ. Thus the Implicit Function Theorem tells us that $\mathrm{Sp}(2n)$ is a smooth manifold of dimension

$$\dim(\mathrm{Sp}(2n)) = 4n^2 - \dim(\mathrm{Antsym}(2n)) = 4n^2 - \frac{4n^2 - 2n}{2} = 2n^2 + n.$$

Proposition (9.10) tells us that $\mathrm{Sp}(2n)$ is automatically a Lie group and therefore we are only left to compute the Lie algebra $\mathfrak{sp}(2n)$, but this follows immediately from Proposition (5.15) and the computation above:

$$\mathfrak{sp}(2n) = T_e \,\mathrm{Sp}(2n) \cong \ker D\varphi(e) = \left\{ B \in \mathrm{Mat}(2n) \middle| \, B^T J_0 + J_0 B = 0 \right\}.$$

(♣) Problem E.4. Let G be a Lie group and suppose H is a subgroup of G which is also an embedded submanifold. Prove that H is closed in G (as a subspace).

Solution. Choose a sequence $(h_k)_{k\in\mathbb{N}}$ such that

$$h_k \to g \in G$$

for $k \to \infty$. We need to show that $g \in H$. Since the identity element $e \in H \subset G$ we can choose a slice chart $\varphi\colon U \to O$ with $e \in U \subset G$ and $0 = \varphi(e) \in O \subset \mathbb{R}^n$. Recall that this means that $\varphi(H \cap U) = O \cap (\mathbb{R}^k \times \{0\}^{n-k})$ for k equal to the dimension of H (as a submanifold). Choose an open subset $e \in W \subset U$ such that $\overline{W} \subset U$. Note that this implies (by the properties of a slice chart) that

$$\overline{W} \cap H \text{ is a closed subset of } G. \tag{E.3}$$

Now, since the map $f(g_1, g_2) = g_1^{-1} g_2$ is continuous, there exists an open non-empty set $e \in V \subset G$ such that

$$V \times V \subset f^{-1}(W).$$

Now, since f is continuous we have

$$g^{-1}h_k = f(g, h_k) \overset{k\to\infty}{\longrightarrow} e.$$

In particular, there exists $K \in \mathbb{N}$ such that $g^{-1}h_k \in V$ for all $k \geq K$. For $j \geq K$ we therefore have

$$h_K^{-1} h_j = (g^{-1}h_K)^{-1}(g^{-1}h_j) = f(g^{-1}h_K, g^{-1}h_j) \in W,$$

which implies

$$h_K^{-1} g = \lim_{j\to\infty} h_K^{-1} h_j \in \overline{W}.$$

Since each $h_j^{-1} h_K \in H$ (for $j \geq K$) it follows from (E.3) that $h_K^{-1} g \in H$. In particular, $g = (h_K)(h_K^{-1} g) \in H$ which finishes the proof.

[2] We are implictly using the identification $T_{J_0}(\mathrm{Antsym}(2n)) \cong \mathrm{Antsym}(2n)$ as vector spaces.

Problem E.5. Prove that if Lie group is abelian then its Lie algebra is abelian. (*Remark*: On Problem Sheet G you will prove that if a *connected* Lie group has abelian Lie algebra, then it is an abelian Lie group.)

Solution. Let G be an abelian Lie group with Lie algebra $\mathfrak{g}$. We first show that the inversion map $i\colon G \to G$ is a Lie group isomorphism. First, observe that

$$i(ab) = (ab)^{-1} = b^{-1}a^{-1} = a^{-1}b^{-1} = i(a)i(b),$$

so i is a group homomorphism; hence a Lie group homomorphism since it is smooth by assumption. Since i is clearly bijective, it follows from Corollary 9.8 that i is a Lie group isomorphism, as desired. We now apply Proposition 9.21 to deduce that $Di(e)\colon \mathfrak{g} \to \mathfrak{g}$ is a Lie algebra isomorphism.

Recall that we can identify $T_{e,e}(G \times G)$ with $\mathfrak{g} \oplus \mathfrak{g}$ as in Problem C.5. We need the following lemma:

Lemma. Let G be a Lie group with multiplication map $m\colon G \times G \to G$ and inversion map $i\colon G \to G$. Then

a) $Dm(e,e)\colon \mathfrak{g} \oplus \mathfrak{g} \to \mathfrak{g}$ is the map $(X, Y) \mapsto X + Y$.

b) $Di(e)\colon \mathfrak{g} \to \mathfrak{g}$ is the map $X \mapsto -X$.

Proof. Using the fact that $Dm(e,e)$ is linear as well as the chain rule, we obtain

$$\begin{aligned} Dm(e,e)(X,Y) &= Dm(e,e)(X,0) + Dm(e,e)(Y,0) \\ &= Dm_1(e)(X) + Dm_2(e)(Y), \end{aligned}$$

where m_1 and m_2 are the smooth maps from G to itself defined respectively by $x \mapsto m(a,e)$ and $b \mapsto m(e,b)$. Since $m_1 = m_2 = \mathrm{Id}_G$, this proves a).

Now consider the maps $p\colon G \times G \to G \times G$ given by $p(a,b) = (a, b^{-1})$ and $\triangle\colon G \to G \times G$ given by $\triangle(a) = (a,a)$. Then the composite $n := m \circ p \circ \triangle$ is the constant map $a \mapsto e$. We thus have

$$\begin{aligned} 0 &= Dn(e)(X) \\ &= Dm(e,e)\Big(Dp(e,e)\big(D\triangle(e)(X)\big)\Big) \\ &= Dm(e,e)\big(Dp(e,e)(X,X)\big) \\ &= Dm(e,e)\big(X, Di_e(X)\big) \\ &= X + Di(e)(X), \end{aligned}$$

where we have used a) and the chain rule for manifolds. This proves b). ∎

Using part b) of the lemma and the fact that $Di(e)$ is an isomorphism of Lie algebras, we obtain

$$[X,Y] = [-X,-Y] = [Di(e)(X), Di(e)(Y)] = Di(e)[X,Y] = -[X,Y],$$

and hence $[X,Y] = 0$ for all $X, Y \in \mathfrak{g}$. Thus $\mathfrak{g}$ is abelian.

PROBLEM E.6. Prove that the Lie bracket on $\mathfrak{gl}(n)$ is given by matrix commutation, i.e.

$$[A, B] = AB - BA, \qquad \forall\, A, B \in \mathfrak{gl}(n) = \mathrm{Mat}(n).$$

SOLUTION. Fix $A, B \in \mathrm{Mat}(n)$. Using the canonical map $\mathcal{J}_I$ from Problem B.3, we want to prove that

$$[\mathcal{J}_I(A), \mathcal{J}_I(B)] = \mathcal{J}_I(AB - BA),$$

where the left-hand side $[\cdot, \cdot]$ is the Lie bracket in $\mathfrak{gl}(n)$. Let X_A and X_B denote the left-invariant vector fields on $\mathrm{GL}(n)$ such that $X_A(I) = \mathcal{J}_I(A)$ and $X_B(I) = \mathcal{J}_I(B)$. By definition of the Lie bracket on $\mathfrak{gl}(n)$ one has

$$[\mathcal{J}_I(A), \mathcal{J}_I(B)] = [X_A, X_B](I).$$

Fix an arbitrary matrix $C \in \mathrm{GL}(n)$, Since $l_C \colon \mathrm{GL}(n) \to \mathrm{GL}(n)$ is itself a linear map, one has $X_A(C) = \mathcal{J}_C(CA)$ by Problem B.4. Now let $u^i_j \colon \mathrm{Mat}(n) \to \mathbb{R}$ denote the function that assigns to a matrix C its (i, j)th entry. Then

$$X_B(u^i_j)(C) = X_B(C)(u^i_j) = \mathcal{J}_C(CB)(u^i_j),$$

which is equal to the (i, j)th entry of CB. Thus as functions we have

$$X_B(u^i_j) = u^i_j \circ r_B,$$

where r_B is right-translation by B. Let $\gamma(t) = I + tA$, so that $\mathcal{J}_I(A) = \gamma'(0)$. Then

$$X_A(I)(X_B(u^i_j)) = \gamma'(0)(u^i_j \circ r_B) = \frac{d}{dt}\Big|_{t=0} u^i_j(B + tAB),$$

which is the (i, j)th entry of AB. Thus

$$[X_A, X_B](I)(u^i_j) = X_A(I)(X_B(u^i_j)) - X_B(I)(X_A(u^i_j))$$

is the (i, j)th entry of $AB - BA$. This shows that

$$[\mathcal{J}_I(A), \mathcal{J}_I(B)](u^i_j) = \mathcal{J}_I(AB - BA)(u^i_j).$$

Since for any C, one has[3]

$$\mathcal{J}_I(C) = \mathcal{J}_I(C)(u^i_j)\, \frac{\partial}{\partial u^i_j},$$

the claim follows.

(♣) PROBLEM E.7. Let $\varphi \colon M^n \to N^k$ be smooth, and let $L^r \subset N$ be an embedded submanifold. We say that φ is **transverse** and **regular** at L if

$$D\varphi(x)[T_x M] + T_{\varphi(x)} L = T_{\varphi(x)} N, \qquad \forall\, x \in \varphi^{-1}(L).$$

Prove that if φ is transverse and regular at P then if $\varphi^{-1}(L) \neq \emptyset$ then $\varphi^{-1}(L)$ is a smooth embedded submanifold of M of dimension $n - k + r$. (*Remark*: The Implicit Function Theorem 5.13 is the special case where L is a point. As a hint, try to reduce this problem to Theorem 5.13.)

[3] In this formula, in order to make the Einstein Summation Convention work, in the term $\frac{\partial}{\partial u^i_j}$, the i is a lower index and the j is an upper index...

Solution. First of all we make the crucial observation that transversality, as well as regularity, are both *local* properties. In particular, it suffices to show the statement for

$$M = \mathbb{R}^n,\ N = \mathbb{R}^k,\ L = \left\{(x^1, x^2, \dots, x^k) \in N \middle|\, x^{r+1} = x^{r+2} = \cdots = x^k = 0\right\}^{4}.$$

We define the auxiliary function

$$\pi\colon N \to \mathbb{R}^{k-r},\ (x^1, \dots, x^k) \mapsto (x^{r+1}, \dots, x^k).$$

The map $\pi\colon N \to \mathbb{R}^{k-r}$ is obviously smooth.

The claim now is that $\varphi\colon M \to N$ is transverse to $L \subset N$ *if and only if* the zero vector $0 \in \mathbb{R}^{k-r}$ is a regular value of the composition $\pi \circ \varphi\colon M \to \mathbb{R}^{k-r}$. In order to prove the above claim we first observe that

$$\pi^{-1}(0) = L,$$

in particular

$$(\pi \circ \varphi)^{-1}(0) = \varphi^{-1}(\pi^{-1}(0)) = \varphi^{-1}(L).$$

Furthermore, since we are viewing N as $\mathbb{R}^k$, the differential of π at any any point $y \in N$ is equal to π again, i.e.

$$\forall y \in N\colon D\pi(y) = \pi\colon T_yN = \mathbb{R}^k \to \mathbb{R}^{k-r}.$$

By applying the chain rule[5] we then see that $D(\pi \circ \varphi)(x) = \pi \circ D\varphi(x)\colon \mathbb{R}^n \to \mathbb{R}^{k-r}$ is surjective for each $x \in (\pi \circ \varphi)^{-1}(0)$ if and only if the image of the differential $D\varphi(x)$ contains the subspace $\left\{(x^1, \dots, x^k) \in \mathbb{R}^k \middle|\, x^{r+1}, \dots, x^k \in \mathbb{R}\right\}$, i.e.

$$D\varphi(x)[\mathbb{R}^n] \supseteq \left\{(x^1, \dots, x^k) \in \mathbb{R}^k \middle|\, x^{r+1}, \dots, x^k \in \mathbb{R}\right\}.$$

On the other hand, this is precisely the case whenever one has

$$D\varphi(x)[\mathbb{R}^n] + L = \mathbb{R}^k.$$

This proves the claim.

With the claim at hands plus the transversality assumption, we can invoke the Implicit Function Theorem (cf. Theorem (5.2)) in order to deduce that $(\pi \circ \varphi)^{-1}(0) = \varphi^{-1}(L)$ is an embedded submanifold of M of dimension

$$\dim \varphi^{-1}(L) = n - (k - r) = n - k + r,$$

which concludes the proof.

[4]This is the local picture of M, N and L, where for L we are using the slice charts introduced in Lecture 5.

[5]For notational simplicity we suppress the maps $\mathcal{J}_x\colon \mathbb{R}^n \to T_x\mathbb{R}^n$ in this proof—this is relatively harmless.

Problem Sheet F

(♣) Problem F.1. Let $A \in \mathfrak{gl}(n) = \mathrm{Mat}(n)$. Prove that the matrix exponential

$$\exp(A) := \sum_{h=0}^{\infty} \frac{1}{h!} A^h$$

converges and defines an element of $\mathrm{GL}(n)$. Prove that $A \mapsto \exp(A)$ is the exponential map of $\mathrm{GL}(n)$.

Problem F.2. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Given $v, w \in \mathfrak{g}$, let X_v denote (as usual) the left-invariant vector field on G with $X_v(e) = v$ and let $\tilde{X}_w$ denote the right-invariant vector field on G with $\tilde{X}_w(e) = w$ (cf. Remark 10.7). Prove that $[X_v, \tilde{X}_w] = 0$.

Problem F.3. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Prove that for $v, w \in \mathfrak{g}$ one has $\mathrm{ad}_v(w) = [v, w]$.

Problem F.4. Let M be a manifold of dimension n. Assume there exist vector fields $X_1, \dots, X_n \in \mathfrak{X}(M)$ such that $\{X_i(x)\}$ is a basis of T_xM for every $x \in M$. Prove that the tangent bundle TM is diffeomorphic to $M \times \mathbb{R}^n$.

Problem F.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Prove that TG is diffeomorphic to $G \times \mathfrak{g}$.

(♣) Problem F.6. A **topological group** G is a topological space that is also a group in the algebraic sense, with the property that the group multiplication

$$m \colon G \times G \to G, \qquad m(a, b) = ab,$$

and group inversion

$$i \colon G \to G, \qquad i(a) = a^{-1},$$

are both **continuous** maps. The goal of this problem is to show that if G is a topological space that simultaneously carries the structure of a topological manifold and a topological group, then G admits at most one diffeomorphism class of smooth structures[1] that turns G into a Lie group.

(i) Let G be a Lie group. Suppose $\gamma \colon \mathbb{R} \to G$ is a continuous group homomorphism. Prove that γ is necessarily smooth, and hence is a one-parameter subgroup. *Hint:* It suffices to prove that γ is smooth on a neighbourhood of $0 \in \mathbb{R}$. Use the fact that the exponential map of G is a diffeomorphism from a neighbourhood of $0 \in \mathfrak{g}$ to a neighbourhood of $e \in G$ (Corollary 10.11.)

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] The converse to this problem was Hilbert's Fifth Problem, posed originally by David Hilbert in 1900. It was eventually proved in 1952 by Montgomery and Zippen, although von Neumann and Pontryagin proved important special cases earlier, and Yamabe proved a stronger result in 1953.

(ii) Let G and H be Lie groups, and suppose $\varphi\colon G \to H$ is a continuous group homomorphism. Prove that φ is automatically smooth, and hence is a Lie group homomorphism. *Hint:* Use the previous part.

(iii) Let G be a topological space which is simultaneously a topological group and a topological manifold. Prove that G admits at most one diffeomorphism class of smooth structures that turns G into a Lie group.

Problem F.7. Let $\varphi\colon M^n \to N^k$ be a surjective submersion. Prove that the connected components of the pre-images $\varphi^{-1}(x)$ as x ranges over N defines an $(n-k)$-dimensional foliation of M.

Solutions to Problem Sheet F

(♣) Problem F.1. Let $A \in \mathfrak{gl}(n) = \mathrm{Mat}(n)$. Prove that the matrix exponential

$$\exp(A) := \sum_{h=0}^{\infty} \frac{1}{h!} A^h$$

converges and defines an element of $\mathrm{GL}(n)$. Prove that $A \mapsto \exp(A)$ is the exponential map of $\mathrm{GL}(n)$.

Solution. We first check that the series converges. The norm on $\mathfrak{gl}(n)$, obtained via identification with $\mathbb{R}^{n^2}$, is given by $|A| = \sqrt{\sum_{i,j} a_{ij}^2}$. This is called the *Frobenius norm* on $\mathfrak{gl}(n)$. The norm is submultiplicative, i.e., $|AB| \leq |A||B|$, and hence by induction $|A^k| \leq |A|^k$. Using this, we apply the Weierstrass M-test to deduce that the matrix exponential $\exp(A)$ converges uniformly on any bounded subset of $\mathfrak{gl}(n)$, by comparison with the series $\sum_k (1/k!)c^k = e^c$ for suitable $c \in \mathbb{R}$.

Fix $A \in \mathfrak{gl}(n)$. The matrix A corresponds to a left-invariant vector field $X_A \in \mathfrak{X}^\ell(\mathrm{GL}(n))$. Recall that we can identify the Lie algebra $\mathfrak{gl}(n) = T_I\,\mathrm{GL}(n)$ of $\mathrm{GL}(n)$ with $\mathrm{Mat}(n)$ via the canonical isomorphism $\mathcal{J}_I\colon \mathrm{Mat}(n) \xrightarrow{\sim} \mathfrak{gl}(n)$ from Problem B.3. The exponential map, which we denote by $\widetilde{\exp}$ in order to distinguish it from the matrix exponential, is defined via $\widetilde{\exp}\big(\mathcal{J}_I(A)\big) := \gamma(1)$, where γ is the unique integral curve of X_A at the identity I. The integral curve γ is determined by the initial value problem:

$$\gamma'(t) = X_A(\gamma(t)), \;\; \gamma(0) = I_n. \tag{F.1}$$

One can show easily that $X_A(B) = BA$ for each $B \in \mathrm{GL}(n)$ (after identifying $T_B\,\mathrm{GL}(n)$ with $\mathrm{Mat}_{n\times n}(\mathbb{R})$). Thus (F.1) becomes the matrix equation

$$\gamma'(t) = \gamma(t)A. \tag{F.2}$$

We claim that $\gamma(t) := \exp(tA)$ satisfies (F.2). To prove the claim, we differentiate the series $\exp(tA)$ formally term-by-term to obtain

$$\gamma'(t) = \sum_{k=1}^{\infty} \frac{k}{k!} A^k = \Big(\sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} A^{k-1}\Big) A = \gamma(t)A.$$

The differentiated series also converges uniformly on bounded sets since it only differs from the series for $\gamma(t)$ by a factor of A. The term-by-term differentiation is thus justified, and we see that γ satisfies (F.2), as desired.

It only remains to show that $\gamma(t)$ is invertible for all t so that, in particular, $\exp(A) \in \mathrm{GL}(n)$. Define $\sigma(t) := \gamma(t)\gamma(-t)$. Then σ is a smooth curve in $\mathfrak{gl}(n)$ and satisfies

$$\sigma'(t) = \big(\gamma(t)A\big)\gamma(-t) - \gamma(t)\big(A\gamma(-t)\big) = 0.$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.

Last modified: June 28, 2019.

It follows that σ is the constant curve $\sigma(t) = I_n$. It follows that $\gamma(t)$ is invertible with inverse $\gamma(-t)$. Finally, we have $\widetilde{\exp}\big(\mathcal{J}_I(A)\big) = \gamma(1) = \exp(A) \in \mathrm{GL}(n)$, as desired.

PROBLEM F.2. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Given $v, w \in \mathfrak{g}$, let X_v denote (as usual) the left-invariant vector field on G with $X_v(e) = v$ and let $\tilde{X}_w$ denote the right-invariant vector field on G with $\tilde{X}_w(e) = w$ (cf. Remark 10.7). Prove that $[X_v, \tilde{X}_w] = 0$.

SOLUTION. By Problem E.2, we have $[X_v, \tilde{X}_w] = 0$ if and only if their flows commute, i.e., for all s and t, we have

$$\theta^v_t \circ \tilde{\theta}^w_s = \tilde{\theta}^w_s \circ \theta^v_t. \tag{F.3}$$

Let $\gamma^v(t)$ (resp. $\tilde{\gamma}^w(t)$) be the unique integral curve of X_v (resp. $\tilde{X}_w$) through the identity. By Proposition 10.6, we have $\theta^v_t = r_{\gamma^v(t)}$ and $\tilde{\theta}^w_s = l_{\tilde{\gamma}^w(t)}$. Since left and right multiplication commute, we have the following equality:

$$r_{\gamma^v(t)} \circ l_{\tilde{\gamma}^w(t)} = l_{\tilde{\gamma}^w(t)} \circ r_{\gamma^v(t)}.$$

This is precisely (F.3); hence $[X_v, \tilde{X}_w] = 0$, as desired.

PROBLEM F.3. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Prove that for $v, w \in \mathfrak{g}$ one has $\mathrm{ad}_v(w) = [v, w]$.

SOLUTION. First note that

$$\begin{aligned}
\mathrm{ad}_v(w) &= D(\mathrm{Ad})(e)[v][w] \\
&= \left(\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}(\exp(tv))\right)[w] \\
&= \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{\exp(tv)}(w) \\
&= \frac{d}{dt}\bigg|_{t=0} D(\mu_{\exp(tv)})(e)[w],
\end{aligned}$$

where $\mu_a(b) = aba^{-1}$. Now writing $\mu_a = r_{a^{-1}} \circ l_a$ and letting X_v and θ^v_t be the left-invariant vector field associated to v and the flow thereof, respectively, we have:

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=0} D(\mu_{\exp(tv)})(e)[w] &= \frac{d}{dt}\bigg|_{t=0} Dr_{\exp(-tv)} \circ Dl_{\exp(tv)}(e)[X_w(e)] \\
&= \frac{d}{dt}\bigg|_{t=0} Dr_{\exp(-tv)}[X_w(\exp(tv)] \\
&= \frac{d}{dt}\bigg|_{t=0} D\theta^v_{-t}(\theta^v_t(e))[X_w(\theta^v_t(e)] \\
&= \mathcal{L}_{X_v} X_w(e).
\end{aligned}$$

where the second equality used the definition of X_w, the third inequality used part *(iii)* and the last equality is the definition of the Lie derivative. Then finally by Theorem 8.25 we have

$$\mathcal{L}_{X_v} X_w(e) = [X_v, X_w](e) = [v, w].$$

Problem F.4. Let M be a manifold of dimension n. Assume there exist vector fields $X_1, \dots, X_n \in \mathfrak{X}(M)$ such that $\{X_i(x)\}$ is a basis of T_xM for every $x \in M$. Prove that the tangent bundle TM is diffeomorphic to $M \times \mathbb{R}^n$.

Solution. For each point in the tangent bundle $(x, v) \in TM$ we write

$$v = v^i(x)\, X_i(x) \in T_xM,$$

where $v^i(x) \in \mathbb{R}$ are the unique coordinates of v with respect to the basis $\{X_i(x)\}$ of T_xM. This allows us to define

$$\varphi\colon TM \to M \times \mathbb{R}^n,\ (x, v) \mapsto (x, v^i(x)\, e_i).$$

The smoothness of the map φ is a direct consequence of the smoothness of the vector fields X_i. Alternatively, one can check directly that this map is smooth, which boils down to the proof of Theorem 4.16.

In order to see that φ defines a diffeomorphism we consider the function

$$\psi\colon M \times \mathbb{R}^n \to TM,\ (y, w) = (y, w^j\, e_j) \mapsto (y, w^j\, X_j(x)).$$

The map ψ is also smooth (same reason as before) and it is an inverse to φ:

$$(\psi \circ \varphi)(x, v) = \psi(x, v^j(x)\, e_j) = (x, v^j(x)\, X_j(x)) = (x, v).$$

Problem F.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Prove that TG is diffeomorphic to $G \times \mathfrak{g}$.

Solution. In view of the previous exercise it suffices to find n-many vector fields $X_i\colon G \to TG$ that form a basis of T_aG at every point $a \in G$. We fix a basis

$$(v_1, \dots, v_n) \subset \mathfrak{g} = T_eG$$

and set

$$X_i(a) = Dl_a(e)[v_i],$$

for all $i = 1, \dots, n$. Recall that $l_a\colon G \to G$, $b \mapsto ab$ is a diffeomorphism, in particular $Dl_a(e)\colon T_eG \to T_aG$ is a vector space isomorphism and therefore[1] our vector fields do indeed satisfy the assumptions of (F.4). We can also write down the diffeomorphism φ from (F.4) explicitly

$$\varphi\colon TG \to G \times \mathfrak{g} = G \times T_eG,\ (a, u) = (a, u^i\, X_i(a)) \mapsto (a, u^i\, v_i).$$

(♣) Problem F.6. A **topological group** G is a topological space that is also a group in the algebraic sense, with the property that the group multiplication

$$m\colon G \times G \to G, \qquad m(a, b) = ab,$$

and group inversion

$$i\colon G \to G, \qquad i(a) = a^{-1},$$

[1] For more details on why X_i are vector fields, we refer to the proof of Theorem (9.19) and notice $X_i = X_{v_i}$.

are both **continuous** maps. The goal of this problem is to show that if G is a topological space that simultaneously carries the structure of a topological manifold and a topological group, then G admits at most one diffeomorphism class of smooth structures[2] that turns G into a Lie group.

(i) Let G be a Lie group. Suppose $\gamma \colon \mathbb{R} \to G$ is a continuous group homomorphism. Prove that γ is necessarily smooth, and hence is a one-parameter subgroup. *Hint:* It suffices to prove that γ is smooth on a neighbourhood of $0 \in \mathbb{R}$. Use the fact that the exponential map of G is a diffeomorphism from a neighbourhood of $0 \in \mathfrak{g}$ to a neighbourhood of $e \in G$ (Corollary 10.11.)

Solution. Let $\gamma \colon \mathbb{R} \to G$ be a continuous homomorphism. Then $\gamma(0) = e \in G$. If there exists $t_0 > 0$ such that γ is smooth on $(-t_0, t_0)$ then for every $T \in \mathbb{R}$, the map

$$t \mapsto \gamma(T+t) = \gamma(T)\gamma(t) = l_{\gamma(T)}(\gamma(t))$$

is also smooth on $(-t_0, t_0)$, being the composition of smooth maps. I.e., to show that γ is smooth, it suffices to find $t_0 > 0$ with the above property. Let $U \subset \mathfrak{g}$ be a neighbourhood of the origin such that $\exp|_U$ is a diffeomorphism onto its image $V := \exp(U) \subset G$. We may w.l.o.g. choose such a U which is convex. We will use the notation $TU = \{Tv \mid v \in U\}$ for $T \in \mathbb{R}$. Since γ is continuous we can find $t_0 > 0$ such that $\gamma(t) \in \exp(\frac{U}{2})$ for all $|t| \leq t_0$. In particular there exists a unique $v \in \frac{U}{2}$ such that $\exp(v) = \gamma(t_0)$. We claim that $\gamma(t) = \exp(\frac{tv}{t_0})$ for all $|t| \leq t_0$ which proves the claim. By continuity it in fact suffices to prove that $\gamma(\frac{mt_0}{n}) = \exp(\frac{mv}{n})$ for all integers $0 \leq |m| < n$. Since $\gamma(\frac{mt_0}{n}) = \gamma(\frac{t_0}{n})^m$ and $\exp(\frac{mv}{n}) = \exp(\frac{v}{n})^m$ it in fact suffices to show that

$$\gamma(\tfrac{t_0}{n}) = \exp(\tfrac{v}{n}) \quad \forall\, n \in \mathbb{N}.$$

For a fixed $n \in \mathbb{N}$ there exists a unique (because $\exp|_U$ is injective) $w \in \frac{U}{2}$ such that $\exp(w) = \gamma(\frac{t_0}{n})$ and our job is to show that $w = \frac{v}{n}$, or $nw = v$. Since $v \in \frac{U}{2}$ and $\exp|_{\frac{U}{2}}$ is injective it suffices to show that $nw \in \frac{U}{2}$. Suppose now $kw \in \frac{U}{2}$ for all $1 \leq k < j$ for some $j \leq n$. Then $jw = (1 + (j-1))w \in U$, so since

$$\exp(jw) = \exp(w)^j = \gamma(\tfrac{t_0}{n})^j = \gamma(\tfrac{jt_0}{n}) \in \exp(\tfrac{U}{2}),$$

and $\exp|_U$ is injective we have $jw \in \frac{U}{2}$. This is the induction step which shows $nw \in \frac{U}{2}$ and finishes the proof.

(ii) Let G and H be Lie groups, and suppose $\varphi \colon H \to G$ is a continuous group homomorphism. Prove that φ is automatically smooth, and hence is a Lie group homomorphism. *Hint:* Use the previous part.

[2] The converse to this problem was Hilbert's Fifth Problem, posed originally by David Hilbert in 1900. It was eventually proved in 1952 by Montgomery and Zippen, although von Neumann and Pontryagin proved important special cases earlier, and Yamabe proved a stronger result in 1953.

SOLUTION. Denote by $\mathfrak{h}$ the Lie algebra of H and choose a basis $v_1, \dots, v_m$ for $\mathfrak{h}$ (so $\dim(H) = m$ as a smooth manifold). The map

$$\mathbb{R}^m \ni (t_1, \dots, t_m) \mapsto \varphi(\exp(t_1 v_1) \cdots \exp(t_m v_m))$$

is clearly smooth, since each of the functions $\mathbb{R} \ni t_k \mapsto \varphi(\exp(t_k v_k))$ is smooth by (i) above (this is clearly a continuous homomorphism $\mathbb{R} \to G$). From the lectures we know that the map

$$f(t_1, \dots, t_m) = \exp(t_1 v_1) \cdots \exp(t_m v_m)$$

is non-singular at the origin. In particular, by the implicit function theorem there exist neighbourhoods $U \subset \mathfrak{h}$ of the origin and $V \subset H$ of $e \in H$ such that $f|_U : U \to V$ is a diffeomorphism. It follows that $\varphi = (\varphi \circ f) \circ f^{-1}$ is smooth on V. This shows that φ is smooth on a neighbourhood of the origin. Given $h_0 \in H$ we have

$$H \ni h \mapsto \varphi(h_0 h) = \varphi(h_0)\varphi(h) = l_{\varphi(h_0)} \circ \varphi(h),$$

so in fact φ is a smooth on a neighbourhood of h_0. This finishes the proof.

(iii) Let G be a topological space which is simultaneously a topological group and a topological manifold. Prove that G admits at most one diffeomorphism class of smooth structures that turns G into a Lie group.

SOLUTION. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two smooth structures on G. We must show that $\mathcal{G}_1 = \mathcal{G}_2$. The maps

$$(G, \mathcal{G}_1) \to (G, \mathcal{G}_2) \quad \& \quad (G, \mathcal{G}_2) \to (G, \mathcal{G}_1)$$

given by the identity are clearly continuous homomorphisms. In particular they are both smooth by (ii). Since they are each others inverses, it follows that $(G, \mathcal{G}_1)$ is diffeomorphic to $(G, \mathcal{G}_2)$. I.e. $\mathcal{G}_1 = \mathcal{G}_2$.

PROBLEM F.7. Let $\varphi \colon M^n \to N^k$ be a surjective submersion. Prove that the collection of preimages $\varphi^{-1}(x)$ as x ranges over N defines an $(n-k)$-dimensional foliation of M.

SOLUTION. Since φ is a submersion, as a consequence of the Implicit Function Theorem, for every $x \in N$, each connected component of $\varphi^{-1}(x)$ is an embedded $(m-k)$-dimensional submanifold of M and, if $w \in \varphi^{-1}(x)$, there holds

$$T_w(\varphi^{-1}(x)) = \ker(D\varphi(x)),$$

where we are identifying $T_w(\varphi^{-1}(x))$ as a subspace of $T_w M$. As a consequence, if we set

$$\Delta_w = T_w(\varphi^{-1}(x)),$$

we have a mapping that to each $x \in M$ assigns a $(n-k)$-dimensional subspace of $T_w M$. To verify that this mapping is smooth, we proceed as follows: given $w_0 \in M$, by the Implicit Function Theorem we may choose a local chart (U,σ)about w_0 in M and a local chart (W,τ) around $\varphi(w_0)$ in N so that the local representative of φ with respect to these charts is given by the projection onto the first k-coordinates, namely (recall that $n \geq k$)

$$(\tau \circ \varphi \circ \sigma^{-1})(w^1, \ldots, w^n) = (w^1, \ldots, w^k), \quad \text{for } v \in \sigma(U).$$

Hence, Δ may be written for every $w \in \varphi(U)$ as

$$\Delta_w = \operatorname{span}\left\{ \left.\frac{\partial}{\partial w^{k+1}}\right|_w, \ldots, \left.\frac{\partial}{\partial w^n}\right|_w \right\} \quad \textit{for } w \in \varphi(U),$$

which implies that $w \mapsto \Delta_w$ is smooth about w_0, and hence that Δ is a $(n-k)$-dimensional distribution on M.

Finally, let Σ be any connected integral manifold of Δ and denote by $\iota\colon \Sigma \to M$ the inclusion map. For fixed $w_0 = \iota(v_0) \in M \cap \iota(\Sigma)$, choosing a local chart (U,σ) about w_0 as above, we have that, for every v in a sufficiently small, connected neighbourhood V of v_0 so that $\iota(V) \subset U$, there holds

$$\begin{aligned} X \in T_v\Sigma &\iff D\iota(v)[X] \in \Delta_w \\ &\iff D\iota(v)[X](w^i) = 0 \quad \text{for } i = k+1, \ldots, n, \end{aligned}$$

and consequently, that

$$v \mapsto w^i(\iota(v)) \quad \text{is constant for every } i = k+1, \ldots, n.$$

This implies $\iota(V)$ is contained in the connected component of $\varphi^{-1}(\varphi(w_0))$ containing w_0. Since w_0 was arbitrarily chosen, we deduce that $\iota(\Sigma)$ is contained in the connected component of $\varphi^{-1}(\varphi(w_0))$ containing w_0.

This proves that the family $\{\varphi^{-1}(x)\}_{x\in N}$ defines a $(n-k)$-dimensional foliation on M.

Problem Sheet G

Problem G.1. Let $\varphi\colon H \to G$ be a Lie group homomorphism. Let $K := \ker \varphi = \{a \in H \mid \varphi(a) = e\}$ and let $\mathfrak{k} := \ker D\varphi(e)$. Prove that K is a closed Lie subgroup of H with Lie algebra $\mathfrak{k}$.

(♣) Problem G.2. Let $\mathfrak{h}$ be a Lie subalgebra of a Lie algebra $\mathfrak{g}$. We say that $\mathfrak{h}$ is an **ideal** in $\mathfrak{g}$ if:

$$v \in \mathfrak{h},\ w \in \mathfrak{g} \qquad \Rightarrow \qquad [v, w] \in \mathfrak{h}.$$

Now let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Let $H \subset G$ be a closed connected subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Prove that H is a normal subgroup if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

(♣) Problem G.3. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Prove that the centre of G is the kernel of the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$. Deduce that G is abelian if and only if $\mathfrak{g}$ is abelian.

Problem G.4. Show that the real projective space $\mathbb{R}P^{n-1}$ can be seen as the homogeneous space $\mathrm{SO}(n)/\,\mathrm{O}(n-1)$.

Problem G.5. Prove that the quotient map $\pi\colon \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n$ (see Problem A.3) is a fibre bundle with fibre $\mathbb{R} \setminus \{0\}$. By adding a single point to each fibre, construct a bundle $E_n \to \mathbb{R}P^n$ whose fibre is $\mathbb{R}$ and for which π is a subbundle. E_n is called the **universal line bundle over** $\mathbb{R}P^n$. Prove that E_n is naturally a subbundle of E_{n+1}.

Problem G.6. Prove that the Klein bottle is a fibre bundle over S^1 with fibre S^1.

Problem G.7. Let $\pi\colon E \to N$ be a fibre bundle with fibre F and structure group G, and let $\varphi\colon M \to N$ be a smooth map.

(i) Prove that $\varphi^\star E$ is a fibre bundle with fibre F and structure group a Lie subgroup of G.

(ii) Prove that

$$T_{(x,p)}(\varphi^\star E) = \{(v, \zeta) \in T_x M \times T_p E \mid D\varphi(x)[v] = D\pi(p)[\zeta]\}\,.$$

(iii) Now suppose that E is actually a vector bundle and $\psi\colon L \to M$ is another smooth map. Prove that $\psi^\star(\varphi^\star E)$ and $(\varphi \circ \psi)^\star E$ are isomorphic as vector bundles.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

Solutions to Problem Sheet G

Problem G.1. Let $\varphi\colon H \to G$ be a Lie group homomorphism. Let $K \coloneqq \ker\varphi = \{a \in H \mid \varphi(a) = e\}$ and let $\mathfrak{k} \coloneqq \ker D\varphi(e)$. Prove that K is a closed Lie subgroup of H with Lie algebra $\mathfrak{k}$.

Solution. Since $\{e\} \subset G$ is closed and φ is continuous, the kernel K is a closed subgroup of H and hence an embedded Lie subgroup of H by the Closed Subgroup Theorem. Let $\iota\colon K \to H$ denote the inclusion map. By Proposition 10.12, we have the following commutative diagram:

$$\begin{tikzcd}
K \arrow[r, "\iota"] & H \arrow[r, "\varphi"] & G \\
\mathrm{Lie}(K) \arrow[u, "\exp"] \arrow[r, "D\iota(e)"'] & \mathfrak{h} \arrow[u, "\exp"] \arrow[r, "D\varphi(e)"'] & \mathfrak{g} \arrow[u, "\exp"']
\end{tikzcd}$$

We identify $\mathrm{Lie}(K)$ with its image under $D\iota(e)$. Let $v \in \mathfrak{h}$. Then $v \in \mathrm{Lie}(K)$ if and only if $\exp(tv) \in K$ for all $t \in \mathbb{R}$, which occurs if and only if $\varphi\big(\exp(tv)\big) = e$ for all $t \in \mathbb{R}$. By the commutativity of the above diagram, this is equivalent to $\exp\big(tD\varphi(e)[v]\big) = e$ for all $t \in \mathbb{R}$. This happens precisely when $D\varphi(e)[v] = 0$, i.e., $v \in \mathfrak{k}$. It follows that $\mathrm{Lie}(K) = \mathfrak{k}$.

(♣) Problem G.2. Let $\mathfrak{h}$ be a Lie subalgebra of a Lie algebra $\mathfrak{g}$. We say that $\mathfrak{h}$ is an **ideal** in $\mathfrak{g}$ if:

$$v \in \mathfrak{h},\ w \in \mathfrak{g} \qquad \Rightarrow \qquad [v, w] \in \mathfrak{h}.$$

Now let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Let $H \subset G$ be a closed connected subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Prove that H is a normal subgroup if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Solution. We will need a few lemmas. In all of the following statements, the groups G and H are as in the statement of the problem.

Lemma 1. Let $U \subset G$ be an open neighbourhood of the identity element such that

$$U = U^{-1} \coloneqq \{g^{-1} \mid g \in U\}.$$

Then for all $g \in G$, there exists a positive integer k and elements $g_1, \dots, g_k \in U$ such that $g = g_1 \cdots g_k$.

Proof. For each $k \geq 1$, let U_k denote the set of all elements of G that can be expressed as the product of k elements in U. We claim that $H \coloneqq \bigcup_{i=1}^{\infty} U_i$ is an open subgroup of G. For the openness, we note that $U_1 = U$ is open and

$$U_k = \bigcup_{g \in U_1} l_g(U_{k-1}).$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

Since each l_g is a diffeomorphism, it follows by induction that each U_k is open; hence H is open. Now if $g \in H$, we may write $g = g_1 \cdots g_k$ for some $g_i \in U$. By assumption, each g_i^{-1} is in U and thus $g^{-1} = g_k^{-1} \cdots g_1^{-1}$ is in H, which is hence a subgroup of G. Since H is an open subgroup of G, it is also closed (this is a general property of topological groups). Since G is connected, it follows that $H = G$. ■

LEMMA 2. The subgroup H is normal in G if and only if

$$(\exp v)(\exp w)\big(\exp(-v)\big) \in H \quad \text{for all } v \in \mathfrak{g} \text{ and } w \in \mathfrak{h}. \tag{G.1}$$

Proof. Note that $\exp(-v) = \exp(v)^{-1}$, so that (G.1) holds when H is normal by the definition of normality. Conversely, suppose (G.1) holds. Choose open subsets $0 \in V \subset \mathfrak{g}$ and $e \in U \subset G$ such that the restriction $\exp\colon V \to U$ is a diffeomorphism (this is possible via Theorem 10.10 and the Inverse Function Theorem). Since the exponential map of H is the restriction of that of G, after shrinking V if necessary, we may assume that $\exp|_{V\cap\mathfrak{h}}$ is a diffeomorphism to a neighbourhood $U_0 \subset H$ of the identity in H. Shrinking V even further, we may assume the $v \in V$ if and only if $-v \in V$. Then (G.1) implies that $ghg^{-1} \in H$ whenever $g \in U$ and $h \in U_0$.

Let let h be an arbitrary element of H. By Lemma 1, we may write $h = h_1 \cdots h_m$ for $h_i \in U_0$. Then for any $g \in U$, we have

$$ghg^{-1} = (gh_1g^{-1}) \cdots (gh_mg^{-1}) \in H.$$

Now let $g \in G$ be arbitrary and write $g = g_1 \cdots g_k$ for $g_j \in U$. It follows by induction on k that $ghg^{-1} \in H$. This proves the lemma. ■

Consider the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$, $\quad g \mapsto \mathrm{Ad}_g$, and recall that for any $g \in G$, we have a commutative diagram

$$\begin{tikzcd} G \arrow[r, "\mu_g"] & G \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "\mathrm{Ad}_g"'] & \mathfrak{g} \arrow[u, "\exp"'] \end{tikzcd} \tag{G.2}$$

where μ_g is the action of g on G by conjugation.

Suppose that $\mathfrak{h}$ is an ideal. Let $v \in \mathfrak{g}$ and $w \in \mathfrak{h}$. Then substituting $\exp v = g$ into (G.2) yields

$$\exp\big(\mathrm{Ad}_{\exp v}(w)\big) = \mu_{\exp v}(\exp w) = (\exp v)(\exp w)\big(\exp(-v)\big). \tag{G.3}$$

On the other hand, we also have the commutative diagram

$$\begin{tikzcd} G \arrow[r, "\mathrm{Ad}"] & \mathrm{GL}(\mathfrak{g}) \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "\mathrm{ad}"'] & \mathfrak{gl}(\mathfrak{g}) \arrow[u, "\exp"'] \end{tikzcd} \tag{G.4}$$

which yields

$$\mathrm{Ad}_{\exp v} = \exp(\mathrm{ad}_v).$$

By Problem F.1, the exponential on the right-hand side is just matrix exponentiation,[1] and we obtain

$$\mathrm{Ad}_{\exp v}(w) = \exp(\mathrm{ad}_v)(w) = \sum_{k=0}^{\infty} \frac{1}{k!}(\mathrm{ad}_v)^k(w). \tag{G.5}$$

Recall from Proposition 10.23 that $\mathrm{ad}_v(w) = [v, w]$. Since $\mathfrak{h}$ is an ideal in $\mathfrak{g}$, we have $\mathrm{ad}_v(w) = [v, w] \in \mathfrak{h}$, and by induction $(\mathrm{ad}_v)^k(w) \in \mathfrak{h}$ for all k. Thus (G.5) implies that $\mathrm{Ad}_v(w) \in \mathfrak{h}$. It follows that $\exp\big(\mathrm{Ad}_{\exp v}(w)\big) \in H$ and hence H is normal by (G.3) and Lemma 2.

Conversely, suppose H is normal. Given $v \in \mathfrak{g}$ and $w \in \mathfrak{h}$, we again use (G.2) to deduce that

$$\exp\big(\mathrm{Ad}_{\exp tv}(sw)\big) = \mu_{\exp tv}(\exp sw) = (\exp tv)(\exp sw)\big(\exp(tv)\big)^{-1},$$

and the left-hand side is in H for all $s, t \in \mathbb{R}$ by assumption. Since $\mathrm{Ad}_{\exp tv}$ is $\mathbb{R}$-linear, we have

$$\exp\big(\mathrm{Ad}_{\exp tv}(sw)\big) = \exp\big(s\,\mathrm{Ad}_{\exp tv}(sw)\big),$$

which we have just shown to be in H for all s. It follows from Corollary 10.13 that $\mathrm{Ad}_{\exp tv}(w) \in \mathfrak{h}$ for all $t \in \mathbb{R}$. Finally, since $\gamma(t) := \exp tv$ is a curve in G with $\gamma(0) = e$ and $\gamma'(0) = v$, we have

$$\frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}_{\exp tv}(w) = D(\mathrm{Ad})(e)[v][w] = \mathrm{ad}_v(w) = [v, w].$$

By the above argument, the left-hand side is in $\mathfrak{h}$. Thus $[v, w] \in \mathfrak{h}$ and $\mathfrak{h}$ is an ideal.

(♣) PROBLEM G.3. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Prove that the centre of G is the kernel of the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$. Deduce that G is abelian if and only if $\mathfrak{g}$ is abelian.

SOLUTION. We denote by $Z(G)$ the centre of G and by $\mu_a(b) = aba^{-1}$ the inner automorphism of G induced by a. Recall finally that $\mathrm{Ad}(a) = D\mu_a(e)$, and that saying that $ab = ba$ in G is exactly the same as saying that $\mu_a(b) = b$.

If $a \in Z(G)$, then $\mu_a(b) = b$ for every $b \in G$ and in particular it follows that $D\mu_a(e)$ is the identity map of $\mathfrak{g}$, that is, $\mathrm{Ad}(a) = \mathrm{id}_{\mathfrak{g}}$.

Conversely suppose that $a \in G$ is so that $D\mu_a(e) = \mathrm{id}_{\mathfrak{g}}$. By property of the exponential map (Proposition 10.12) we then deduce

$$\mu_a \circ \exp = \exp \circ D\mu_a(e) = \exp \quad \text{in } \mathfrak{g},$$

but since exp is a diffeomorphism when restricted to a sufficiently small neighbourhood V of 0 in $\mathfrak{g}$, the above equality also shows that, in the neighbourhood of e in

[1] Or rather *homomorphism exponentiation* since we haven't taken a basis of $\mathfrak{g}$. The argument in the solution to Problem F.1 applies just as well to this setting. Note that $(ad_v)^k$ means the k-fold composition of ad_v with itself.

G given by $U = \exp(V)$, μ_a is the identity map. To conclude that μ_a is the identity not only when restricted to U but to the whole G, recall that, as proved in Lemma 1, since G is connected we have $G = \bigcup_{k\geq 1} U^k$, that is to say, every element $g \in G$ can be written as $g = b_1 \cdots b_n$ for some $b_j \in U$ and some $n \in \mathbb{N}$. But then, arguing by finite induction, we see that

$$\begin{aligned}
\mu_a(g) &= a(b_1 \cdots b_n)a^{-1} \\
&= b_1\, a(b_2 \cdots b_n)a^{-1} \\
&= b_1 b_2\, a(b_3 \cdots b_n)a^{-1} \\
&= \dots \\
&= (b_1 \cdots b_n)aa^{-1} \\
&= g,
\end{aligned}$$

which proves that μ_a is the identity on the whole G, and hence that $a \in Z(G)$.

We can draw the conclusions. If G is abelian, then Ad is the constant map: $\mathrm{Ad}(a) = \mathrm{id}_{\mathfrak{g}}$ for every $a \in G$; it follows that $\mathrm{ad} = D\,\mathrm{Ad}(e) = 0$, but then $[v, w] = \mathrm{ad}_v\, w = 0$ for every $v, w \in \mathfrak{g}$, and hence $\mathfrak{g}$ is abelian. Vice versa, if $\mathfrak{g}$ is abelian, then $[v, w] = 0$ for every $v, w \in \mathfrak{g}$, but then ad is the zero map, and hence Ad is constant; since $\mathrm{Ad}(e) = \mathrm{id}_{\mathfrak{g}}$ and G is connected, Ad is the constant map and thus by what proved above, G is abelian.

Problem G.4. Show that the real projective space $\mathbb{R}P^{n-1}$ can be seen as the homogeneous space $\mathrm{SO}(n)\big/\,\mathrm{O}(n-1)$.

Solution. The Lie group action of $SO(n)$ on $\mathbb{R}P^{n-1}$ is defined as follows: for $A \in SO(n)$ and $[x] = [x^1, \dots, x^n] \in \mathbb{R}P^{n-1}$, we set

$$SO(n) \times \mathbb{R}P^{n-1} \to \mathbb{R}P^{n-1} \quad (A, [x]) \mapsto A[x] := [Ax]$$

Since A is a linear map, this map is well defined and it is a Lie group action.

It is a transitive action: if $[x], [y]$ are elements of $\mathbb{R}P^{n-1}$, let x_0 be one on the the points in the equivalence class of $[x]$ so that $x_0 \in S^{n-1}$ (recall that there are precisely two of them, and in fact $\mathbb{R}P^{n-1} \simeq S^{n-1}/\{p \mapsto -p\}$) and likewise let y_0 be one on the the points in the equivalence class of $[x]$ so that $y_0 \in S^{n-1}$. The map $A \in SO(n)$ so that $Ax_0 = y_0$ will then map $[x_0]$ to $[y_0]$ via the above action.

We deduce that, if H is the isotropy group of a fixed element $[x]$ for the above action, then $\mathbb{R}P^{n-1} \simeq SO(n)/H$, so let us compute the isotropy group of the n-th canonical vector $[0, \dots, 1] = [e_n]$. We have that $A[e_n] = [e_n]$ precisely if $Ae_n = e_n$ or $Ae_n = -e_n$. Since $A \in SO(n)$, it must then have the form

$$A = \begin{pmatrix} \begin{pmatrix} & & \\ & B & \\ & & \end{pmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 \end{matrix} & 1 \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} \begin{pmatrix} & & \\ & B' & \\ & & \end{pmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 \end{matrix} & -1 \end{pmatrix}$$

where $B, B' \in O(n-1)$ with $\det(B) = 1$ and $\det(B') = -1$. But then we may identify H with $O(n-1)$ via the map

$$A \mapsto A|_{\mathbb{R}^{n-1}},$$

where $A|_{\mathbb{R}^{n-1}}$ is the restriction of A to the first $n-1$ components. Such identification is compatible with the group action we are considering, and consequently we conclude that $\mathbb{R}P^{n-1} \simeq SO(n)/O(n-1)$.

Problem G.5. Prove that the quotient map $\pi\colon \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{R}P^n$ (see Problem A.3) is a fibre bundle with fibre $\mathbb{R} \setminus \{0\}$. By adding a single point to each fibre, construct a bundle $E_n \to \mathbb{R}P^n$ whose fibre is $\mathbb{R}$ and for which π is a subbundle. E_n is called the **universal line bundle over** $\mathbb{R}P^n$. Prove that E_n is naturally a subbundle of E_{n+1}.

Solution. Recall that the equivalence relation $\sim$ on $\mathbb{R}^{n+1}$ is defined by $x \sim y$ if any only if there exists $\lambda \neq 0$ such that $x = \lambda y$. We write an element of $\mathbb{R}P^n$ as $[x_0 : \cdots : x_n]$ with the understanding that $[x_0 : \cdots : x_n] = [\lambda x_0 : \cdots : \lambda x_n]$ for all $\lambda \neq 0$. The canonical map $\pi\colon \mathbb{R}^{n+1} \to \mathbb{R}P^n$ is given by

$$\pi(x_0, \ldots, x_n) = [x_0 : \cdots : x_n].$$

Set $\widetilde{U}_i \coloneqq \{(x_0, \ldots, x_n) \mid x_i \neq 0\}$ for $i = 0, \ldots, n$ and define $U_i \coloneqq \pi(\widetilde{U}_i)$. We have seen in previous exercises that the U_i's are open subsets and that $\varphi_i\colon U_i \to \mathbb{R}^n$ given by

$$\varphi_i([x_0 : \cdots : x_n]) = \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots \frac{x_n}{x_i}\right)$$

are charts of $\mathbb{R}P^n$. We define the map $\alpha_i\colon \widetilde{U}_i \to \mathbb{R}\backslash\{0\}$ by

$$\alpha_i(x_o, \ldots, x_n) = x_i,$$

so that $(\pi, \alpha_i)\colon \widetilde{U}_i \to U_i \times \mathbb{R}\backslash\{0\}$ is given by

$$(\pi, \alpha_i)(x_0, \ldots, x_n) = ([x_0 : \cdots : x_n], x_i).$$

Note that this map is a diffeomorphism if and only if $(\varphi_i \circ \pi, \alpha_i)\colon \widetilde{U}_i \to \mathbb{R}^n \times \mathbb{R}\backslash\{0\}$ is a diffeomorphism. We have

$$(\varphi_i \circ \pi, \alpha_i)(x_0, \ldots, x_n) = \left(\left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots \frac{x_n}{x_i}\right), x_i\right)$$

and it is easy to see that this map is a diffeomorphism: It is clearly smooth and its inverse map is given by

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}\backslash\{0\} &\to \widetilde{U}_i \\ ((x_1, \ldots x_n), \lambda) &\mapsto (\lambda x_1, \ldots, \lambda x_{i-1}, \lambda, \lambda x_{i+1}, \lambda x_n). \end{aligned}$$

This finishes the proof that $\pi\colon \mathbb{R}^{n+1} \to \mathbb{R}P^n$ is a fibre bundle with fibre $\mathbb{R}\backslash\{0\}$.

We now "glue" the element $0 \in \mathbb{R}$ into each fibre of this fibre bundle to obtain a vector bundle. The precise procedure for doing this is the following: The total set of the vector bundle is going to be

$$E_n \coloneqq (U_0 \times \mathbb{R} \sqcup \cdots \sqcup U_n \times \mathbb{R}) \,/\, \sim, \tag{G.6}$$

where the equivalence relation $\sim$ is defined as follows: We have well-defined maps

$$\varphi_{ji} \coloneqq (\pi|_{\widetilde{U}_j}, \alpha_j) \circ (\pi|_{\widetilde{U}_i}, \alpha_i)^{-1} \colon U_i \cap U_j \times (\mathbb{R}\backslash\{0\}) \to U_i \cap U_j \times (\mathbb{R}\backslash\{0\})$$
$$\varphi_{ji}([x_0\colon \cdots \colon x_n], \lambda) = \left([x_0\colon \cdots \colon x_n], \lambda \frac{x_j}{x_i}\right)$$

which clearly extend in a unique way to a smooth map

$$\varphi_{ji} \colon U_i \cap U_j \times \mathbb{R} \to U_i \cap U_j \times \mathbb{R}$$

such that for fixed $[x] \in U_i \cap U_j$, the map $\mathbb{R} \ni \lambda \mapsto p_2 \circ \varphi_{ji}([x], \lambda)$ is linear (here $p_2 \colon U_i \cap U_j \times \mathbb{R} \to \mathbb{R}$ denotes the natural projection). By definition, the equivalence relation in (G.6) identifies an element $([x], \lambda_1) \in U_i \times \mathbb{R}$ with the element $([y], \lambda_2) \in U_j \times \mathbb{R}$ exactly if $\varphi_{ji}([x], \lambda_1) = ([y], \lambda_2)$. This defines a smooth structure on E_n in such a way that we have an induced projection map $\widetilde{\pi} \colon E_n \to \mathbb{R}P^n$ which gives a vector bundle structure to E_n. From the construction it is clear that we have a commuting diagram of smooth maps

$$\begin{tikzcd}
\mathbb{R}^{n+1}\backslash\{0\} \arrow[rr] \arrow[dr, "\pi"'] & & E_n \arrow[dl, "\widetilde{\pi}"] \\
& \mathbb{R}P^n &
\end{tikzcd}$$

I.e. $\pi \colon \mathbb{R}^{n+1}\backslash\{0\} \to \mathbb{R}P^n$ is a subbundle of $\widetilde{\pi} \colon E_n \to \mathbb{R}P^n$. The inclusion $\imath \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+2}$ given by $\imath(x_0, \ldots, x_n) = (x_0, \ldots, x_n, 0)$ descends to an inclusion

$$i \colon \mathbb{R}P^n \to \mathbb{R}P^{n+1}$$

such that the following diagram commutes

$$\begin{tikzcd}
\mathbb{R}^{n+1}\backslash\{0\} \arrow[r, "\imath"] \arrow[d, "\pi_{n+1}"] & \mathbb{R}^{n+2}\backslash\{0\} \arrow[d, "\pi_{n+1}"'] \\
\mathbb{R}P^n \arrow[r, "i"] & \mathbb{R}P^{n+1}
\end{tikzcd}$$

It is now immediate to check that $\imath$ induces a map $E_n \to E_{n+1}$, making E_n into a subbundle of E_{n+1}.

Problem G.6. Prove that the Klein bottle is a fibre bundle over S^1 with fibre S^1.

Solution. We define the Klein bottle

$$K = [0,1]^2/\sim$$

by making the following identifications on the boundary of $[0,1]^2$

- $(x, 0) \sim (x, 1)$
- $(0, y) \sim (1, 1-y)$

and denote equivalence classes in K with square brackets, i.e. $[x, y] \in K$. Checking that K is a smooth manifold is not hard and is left to reader. Now we define the projection

$$\pi\colon K \to S^1 = \mathbb{R}/\mathbb{Z},\ [x, y] \mapsto x$$

and claim that this is a fibre bundle with fibre S^1[2].

Indeed, $\pi\colon K \to S^1$ is a smooth surjection and we can construct two bundle charts α_1 and α_2 as follows: Let $U_1 := (0, 1)$ be the open arc viewed inside S^1 and $U_2 = \left[0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right]$ also viewed inside S^1. The map

$$\alpha_1\colon \pi^{-1}(U_1) \to S^1,\ [x, y] \mapsto y$$

is obviously well defined as $[x, 1] = [x, 0]$ and $0 = 1$ in $S^1 = \mathbb{R}/\mathbb{Z}$, and also smooth. It readily follows that

$$(\pi, \alpha_1)\colon \pi^{-1}(U_1) \to U_1 \times S^1$$

defines a diffeomorphism, thus proving that α_1 is a bundle chart.

For $\alpha_2\colon \pi^{-1}(U_2) \to S^1$ one has to take the identification $(0, y) \sim (1, 1-y)$ into account, which is the reason why the same choice as for α_1 does not work. However, we claim that the following choice works:

$$\alpha_2([x, y]) = \begin{cases} y, & \text{if } x < \frac{1}{2} \\ -y, & \text{if } x > \frac{1}{2}. \end{cases}$$

The only classes $[x, y] \in \pi^{-1}(U_2)$ that have several representatives and actually matter for the definition of α are those of the form

$$[0, y] = \{(0, y), (1, 1-y)\}.$$

Then we can check

$$\alpha_2([0, y]) = y \text{ and } \alpha_2([1, 1-y]) = y - 1,$$

but $y - 1$ equals y if viewed in S^1 and therefore α_2 is well defined (note that it is crucial to exclude $\frac{1}{2}$ to make this definition work!). Just as for α_1, smoothness of α_2 is clear. It's also not hard to see that

$$(\pi, \alpha_2)\colon \pi^{-1}(U_2) \to U_2 \times S^1$$

defines a diffeomorphisms by taking a closer look at what α_2 does on the two preimages $V_+ = \pi^{-1}\left[0, \frac{1}{2}\right)$ and $V_- = \pi^{-1}\left(\frac{1}{2}, 1\right]$, namely:

$$(\pi, \alpha_2)\Big|_{V_\pm} = \pm\,\mathrm{id}_{V_\pm}.$$

Problem G.7. Let $\pi\colon E \to N$ be a fibre bundle with fibre F and structure group G, and let $\varphi\colon M \to N$ be a smooth map.

(i) Prove that $\varphi^\star E$ is a fibre bundle with fibre F and structure group a Lie subgroup of G.

[2] Sanity check: $\pi^{-1}(x) = \{[x, y] \mid y \in [0, 1]\} \cong S^1$ as $[x, 1] = [x, 0]$.

(ii) Prove that

$$T_{(x,p)}(\varphi^\star E) = \{(v,\zeta) \in T_xM \times T_pE \mid D\varphi(x)[v] = D\pi(p)[\zeta]\}\,.$$

(iii) Now suppose that E is actually a vector bundle and $\psi\colon L \to M$ is another smooth map. Prove that $\psi^\star(\varphi^\star E)$ and $(\varphi \circ \psi)^\star E$ are isomorphic as vector bundles.

Solution. **Part (i)**:

We start part (i) by proving that $\varphi^\star E$ carries a smooth manifold structure. By defining the smooth map

$$\varphi\colon M \times E \to N \times N,\ \varphi(x,p) = (\varphi(x), \pi(p))$$

and the diagonal

$$\Delta = \{(y,y) \in N \times N\} \subset N \times N$$

one can write

$$\varphi^\star E = \varphi^{-1}(\Delta).$$

In view of problem E.7 it suffices to show that φ is transverse (and regular) at the diagonal Δ to conclude that $\varphi^\star E$ is a smooth manifold[3]. This follows from the fact that $\pi\colon E \to N$ is a submersion (cf. Lemma 13.4) and $T_{(y,y)}\Delta = \{(v,v) \in T_yN \times T_yN\}$.

The projection $\mathrm{pr}_1\colon \varphi^\star E \to M$, $(x,p) \mapsto x$ is obliviously a continuous surjection. For the bundle charts, we make the same choice as in Example 13.19 by picking $\alpha\colon \pi^{-1}(U) \to F$ a bundle chart on $\pi\colon E \to N$ and setting

$$\alpha^\star := \alpha \circ \mathrm{pr}_2 \colon U^\star \to F, \text{ with } U^\star := \mathrm{pr}_1^{-1}(\varphi^{-1}(U))^4.$$

The map $(\mathrm{pr}_1, \alpha^\star)\colon U^\star \to \varphi^{-1}(U) \times F$ is a diffeomorphism since

$$(\mathrm{pr}_1, \alpha^\star)(x,p) = (x, \alpha^\star(x,p)) = (x, \alpha(p)).$$

We finish part (i) by proving that G is a structure group of the fibre bundle $\mathrm{pr}_1\colon \varphi^\star E \to M$. Let $\mu\colon G \times F \to F$ denote the G-action on F coming from the bundle $\pi\colon E \to N$. Any two bundle charts $\alpha^\star\colon U^\star \to F$, $\beta^\star\colon V^\star \to F$, with $U^\star \cap V^\star \neq \emptyset$ and the notation from above, are (G,μ)-compatible via the smooth function

$$\tilde\rho_{\alpha^\star\,\beta^\star}\colon U^\star \cap V^\star \to G,\ x \mapsto \tilde\rho_{\alpha\,\beta}(\varphi(x)).$$

Indeed, for any $x \in U^\star \cap V^\star$ and $z \in F$ we have

$$\begin{aligned}
\rho_{\alpha^\star\,\beta^\star}(x)(z) &= \alpha^\star\Big|_{\varphi^\star E_x} \circ \beta^\star\Big|_{\varphi^\star E_x}^{-1}(z) \\
&= \left(\alpha \circ \mathrm{pr}_2\Big|_{\varphi^\star E_x}\right) \circ \left(\beta \circ \mathrm{pr}_2\Big|_{\varphi^\star E_x}\right)^{-1}(z) \\
&= \left(\alpha\Big|_{E_{\varphi(x)}} \circ \beta\Big|_{E_{\varphi(x)}}^{-1}\right)(z) \\
&= \rho_{\alpha\,\beta}(\varphi(x))(z),
\end{aligned}$$

[3]We even get the dimension of $\varphi^\star E$, namely $\dim \varphi^\star E = \dim M + \dim E - \dim N$. This shows $\dim \varphi^\star E \geq \dim M$, which is necessary if we want $\varphi^\star E \to M$ to be a fibre bundle (why ?).

[4]**Warning:** This notation is not standard and it's only purpose is to prevent cumbersome notation throughout the proof.

and therefore

$$\rho_{\alpha^\star\,\beta^\star}(x)(z) = \mu(\tilde{\rho}_{\alpha\,\beta}(\varphi(x)), z) = \mu(\tilde{\rho}_{\alpha^\star\,\beta^\star}(x), z).$$

This finishes the proof of (i).

Part (ii):

We proceed by showing that the LHS is included in the RHS and deduce the equality by a dimension argument. Let $(x,p) \in \varphi^\star E$, $(v,\zeta) \in T_{(x,p)}\varphi^\star E$ and pick a smooth curve $\gamma\colon (-\epsilon,\epsilon) \to \varphi^\star E$ satisfying

$$\begin{cases} \gamma(0) = (x,p), \\ \dot{\gamma}(0) = (v,\zeta), \end{cases}$$

and write $\gamma(t) = (x(t), p(t)) \in \varphi^\star E$. By definition of the pullback bundle we have the relation

$$\varphi(x(t)) = \pi(p(t))$$

and differentiating this at time $t = 0$ gives

$$D\varphi(x)[v] = D\pi(p)[\zeta],$$

which proves the desired inclusion.

For the dimension argument we observe that the RHS can be written as the kernel of the following linear operator

$$A\colon T_xM \times T_pE \to T_{\varphi(x)}N,\ A(v,\zeta) = D\varphi(x)[v] - D\pi(p)[\zeta].$$

This is well defined because $\varphi(x) = \pi(p)$. The strategy now consists in showing that

$$\dim\ker A = \dim T_{(x,p)}(\varphi^\star E), \tag{G.7}$$

where the latter is equal to

$$\dim \varphi^\star E = \dim M + \dim E - \dim N,$$

as we have already seen (cf. footnote in Part (i)). Indeed, A is surjective as $\pi\colon E \to N$ is a submersion and basic linear algebra then tells us that

$$\dim\ker A = \underbrace{\dim(T_xM \times T_pE)} - \dim\operatorname{im} A = \dim M + \dim E - \dim N.$$

This finishes Part (ii).

Part (iii):

The goal is to find a smooth map between the total spaces

$$\varphi\colon (\varphi \circ \psi)^\star E \to \psi^\star(\varphi^\star E)$$

that covers the identity $\mathrm{id}\colon L \to L$. Unravelling the definitions of both the total space $(\varphi \circ \psi)^\star E$ and $\psi^\star(\varphi^\star E)$ will naturally lead us to the right candidate for φ. Observe:

$$(a,p) \in (\varphi \circ \psi)^\star E \iff (a,p) \in E \times L \text{ and } (\varphi \circ \psi)(a) = \pi(p),$$

and

$$
\begin{aligned}
(b, x, p) \in \psi^\star(\varphi^\star E) &\iff (b, (x, p)) \in L \times \varphi^\star E \text{ and } \psi(b) = \mathrm{pr}_1(x, p) = x \\
&\iff (b, x, p) \in L \times M \times E \text{ and } \varphi(x) = \pi(p),\ \psi(b) = x.
\end{aligned}
$$

We claim that

$$
\varphi\colon (\varphi \circ \psi)^\star E \to \psi^\star(\varphi^\star E),\ \varphi(a, p) \coloneqq (a, (\varphi \circ \psi)(a), p),
$$

is well defined and satisfies the bundle isomorphism properties. Indeed, well-definedness follows immediately form the equivalences above. Smoothness and the fact that φ cover the identity on L are both obvious by construction of φ, so we are only left to show that φ maps fibres to fibres isomorphically. But this is also clear by noting that

$$
\begin{aligned}
(\varphi \circ \psi)^\star E_a &= \{a\} \times E_{\varphi(\psi(a))}, \\
\psi^\star(\varphi^\star) E_a &= \{a\} \times \varphi^\star E_{\psi(a)} = \{a\} \times \{(\varphi \circ \psi)(a)\} \times E_{\varphi(\psi(a))}.
\end{aligned}
$$

This finishes Part (iii) and the proof.

Problem Sheet H

Problem H.1. Let V, W and U be vector spaces. Prove there are natural isomorphisms $V \otimes W \cong W \otimes V$ and $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$.

Problem H.2. Let V and W be vector spaces. For any $A \in \mathrm{Alt}_r(V, W)$ prove there is a unique linear map $T \colon \bigwedge^r(V) \to W$ such that the following diagram commutes:

$$\begin{tikzcd} \overbrace{V \times \cdots \times V}^{r} \arrow[rr, "\wedge"] \arrow[dr, "A"'] & & \bigwedge^r(V) \arrow[dl, dashed, "T"] \\ & W & \end{tikzcd}$$

Moreover $\bigwedge^r(V)$ is uniquely characterised by this property.

Problem H.3. Let V be a vector space of dimension k with basis $\{e_1, \dots, e_k\}$. Prove that

$$\{e_{i_1} \wedge \cdots \wedge e_{i_r} \mid 1 \le i_1 < \cdots < i_r \le k\}$$

is a basis of $\bigwedge^r(V)$. Prove that $\bigwedge^r(V) = 0$ for $r > k$. Thus $\dim \bigwedge^r(V) = \binom{k}{r}$ and $\dim \bigwedge(V) = 2^k$.

Problem H.4. Let M be a smooth manifold and suppose $\pi_i \colon E_i \to M$ are two vector bundles over M of the same rank k. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both[1] E_1 and E_2 admit $\mathrm{GL}(k)$-bundle atlases over the $U_{\mathtt{a}}$. Let

$$\rho^1_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k), \qquad \text{and} \qquad \rho^2_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k)$$

denote the transition functions of E_1 and E_2 with respect to these bundle atlases. Prove that E_1 and E_2 are isomorphic if and only if there exists a smooth family $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to \mathrm{GL}(k)$ of functions such that

$$\nu_{\mathtt{a}}(x) \circ \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \circ \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}.$$

(♣) Problem H.5. Let $\varphi \colon M \to N$ be a smooth map and suppose $\pi \colon E \to N$ is a vector bundle, which we illustrate pictorially as:

$$\begin{tikzcd} & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd} \qquad (\delta)$$

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] This can always be achieved by taking intersections.

A **solution** of the diagram (δ) is a vector bundle $\pi_1 \colon E_1 \to M$ over M together with a vector bundle morphism $\Phi \colon E_1 \to E$ along φ. Thus a solution is any pair $\pi_1 \colon E_1 \to M$ and Φ such that the following commutes:

$$\begin{tikzcd}
E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E \arrow[d, "\pi"] \\
M \arrow[r, "\varphi"'] & N
\end{tikzcd}$$

As we have seen, one possible solution is the pullback bundle $\varphi^\star E$:

$$\begin{tikzcd}
\varphi^\star E \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\mathrm{pr}_1"'] & E \arrow[d, "\pi"] \\
M \arrow[r, "\varphi"'] & N
\end{tikzcd}$$

The aim of this exercise is to prove that $\varphi^\star E$ is the "most efficient" solution in the following sense: Suppose $\pi_1 \colon E_1 \to E$ and Φ is *any* solution to (δ). Prove there exists a *unique* vector bundle homomorphism $\Psi \colon E_1 \to \varphi^\star E$ such that the following diagram commutes:

$$\begin{tikzcd}
E_1 \arrow[dr, dashed, "\Psi"] \arrow[drr, bend left, "\Phi"] \arrow[ddr, bend right, "\pi_1"'] & & \\
& \varphi^\star E \arrow[r, "\mathrm{pr}_2"'] \arrow[d, "\mathrm{pr}_1"] & E \arrow[d, "\pi"] \\
& M \arrow[r, "\varphi"'] & N
\end{tikzcd}$$

Prove moreover that $\varphi^\star E$ is uniquely determined by this property. Explicitly this means that if $\tilde{\pi} \colon \tilde{E} \to M$ and $\tilde{\Phi}$ is another solution to the diagram (δ) with the property that for *any* solution $\pi_1 \colon E_1 \to M$ and Φ of (δ) there exists a *unique* vector bundle homomorphism $\tilde{\Psi} \colon E_1 \to \tilde{E}$ such that the following commutes:

$$\begin{tikzcd}
E_1 \arrow[dr, dashed, "\tilde{\Psi}"] \arrow[drr, bend left, "\Phi"] \arrow[ddr, bend right, "\pi_1"'] & & \\
& \tilde{E} \arrow[r, "\tilde{\Phi}"'] \arrow[d, "\tilde{\pi}"] & E \arrow[d, "\pi"] \\
& M \arrow[r, "\varphi"'] & N
\end{tikzcd}$$

then in fact $\tilde{E}$ is isomorphic as a vector bundle over M to $\varphi^\star E$. *Hint:* Argue as in the proof of Lemma 15.2.

Problem H.6. Let $\pi_i\colon E_i \to M_i$ be vector bundles for $i = 1, 2$. Suppose $\Theta\colon E_1 \to E_2$ is any smooth map that maps each fibre $\pi_1^{-1}(x)$ linearly onto some fibre $\pi_2^{-1}(y)$ for $x \in M_1$ and $y \in M_2$. Prove that $\Theta = \Psi \circ \Phi$ where Ψ is a vector bundle homomorphism and Φ is a vector bundle morphism along a map $M_1 \to M_2$.

Problem H.7. Let $\pi_i\colon E_i \to M$ be two vector bundles over the same manifold M of ranks k_i. Let $\Phi\colon E_1 \to E_2$ be a vector bundle homomorphism.

(i) Assume Φ is injective on each fibre. Consider the quotient vector space

$$\bar{E}_x \coloneqq E_2|_x / \Phi|_{E_1|_x}(E_1|_x).$$

Prove that $\bar{E} \coloneqq \bigsqcup_{x \in M} \bar{E}_x$ is a vector bundle of rank $k_2 - k_1$.

(ii) Assume that Φ is surjective on each fibre. Let

$$K_x \coloneqq \ker \Phi|_{E_x} \subset E_1|_x.$$

Prove that $K \coloneqq \bigsqcup_{x \in M} K_x$ is a vector bundle over M of rank $k_1 - k_2$.

Solutions to Problem Sheet H

Problem H.1. Let V, W and U are vector spaces. Prove there are natural isomorphisms $V \otimes W \cong W \otimes V$ and $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$.

Solution. To prove that $V \otimes W \cong W \otimes V$, consider the bilinear map $F : V \times W \to W \otimes V$ given by $F(v, w) = w \otimes v$. By the universal property of the tensor product (lemma 15.2) there is an associated linear map $T_F : V \otimes W \to W \otimes V$ defined by $T_F(v \otimes w) = w \otimes v$ and then extended by linearity. Such map is invertible, the inverse being defined by $T_F^{-1}(w \otimes v) = v \otimes w$ and then extended by linearity. Such map is then the required isomorphism.

To prove that $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$, consider the map $G : (U \otimes V) \times W \to U \otimes (V \otimes W)$ (which we regard as a function of two variables, the first on $U \otimes V$ and the second on W) defined by $G(u \otimes v, w) = u \otimes (v \otimes w)$, and extended by linearity with respect to its first argument. Such map is bilinear and so by the universal property of the tensor product (lemma 15.2) there is an associated linear map $T_G : (U \otimes V) \otimes W \to U \otimes (V \otimes W)$ defined by $T_G((u \otimes v) \otimes w) = u \otimes (v \otimes w)$ and extended by linearity. Since such map is invertible, (the inverse being defined by $T_G^{-1}(u \otimes (v \otimes w)) = (u \otimes v) \otimes w$ and then extended by linearity), and so it provides the required isomorphism.

Problem H.2. Let V and W be vector spaces. For any $A \in \mathrm{Alt}_r(V, W)$ prove there is a unique linear map $T : \bigwedge^r(V) \to W$ such that the following diagram commutes:

$$\begin{tikzcd}
\overbrace{V \times \cdots \times V}^{r} \arrow[rr, "\wedge"] \arrow[dr, "A"'] & & \bigwedge^r(V) \arrow[dl, dashed, "T"] \\
& W &
\end{tikzcd}$$

Moreover $\bigwedge^r(V)$ is uniquely characterised by this property.

Solution. Any alternating r-linear is by definition multilinear which allows us to define a map

$$\tilde{T} : T^{r,0}(V) = V \otimes \cdots \otimes V \to W,\ \tilde{T}(v_1 \otimes \cdots \otimes v_r) := A(v_1, \dots, v_r).$$

This map is well defined and linear by the very definition of the tensor product (cf. Definition 15.1) and multilinearity of A.

Now we claim that our $\tilde{T} : T^{r,0}(V) \to W$ factors through $\bigwedge^r(V) \cong T^{r,0}(V) \big/ I^r(V)$. Indeed, any element of the form $v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_i \otimes \cdots \otimes v_r$ gets mapped to 0 by $\tilde{T}$ simply because the v_i appears twice and A is alternating, i.e.

$$A(v_1, \dots, v_i, \dots, v_i, \dots, v_r) = 0.$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

Therefore
$$I^r(V) \subset \ker \tilde{T}$$
and we end up with a linear map $T\colon \bigwedge^r(V) \to W$ such that the following diagram commutes

$$\begin{tikzcd}
T^{r,0}(V) \arrow[rr, "\tilde{T}"] \arrow[dr, "p"'] & & W \\
& \bigwedge^r(V) \arrow[ur, "T"'] &
\end{tikzcd}$$

where p the obvious projection.

Moreover, the linear map T fits into the desired commutative diagram ($T \circ \wedge = T \circ (p \circ \otimes) = \tilde{T} \circ \otimes = A$) and we are only left to show that such a T is unique, but that's straightforward: The decomposable elements $v_1 \wedge \cdots \wedge v_r$ generate $\bigwedge^r(V)$ and as $T(v_1 \wedge \cdots \wedge v_r)$ is already determined to be equal to $A(v_1, \ldots, v_r)$ the uniqueness of T follows.

The proof of the second statement about the unique characterisation of $\bigwedge^r(V)$ is identical to the second part in the proof of Lemma 15.2.

Problem H.3. Let V be a vector space of dimension k with basis $\{e_1, \ldots, e_k\}$. Prove that
$$\{e_{i_1} \wedge \cdots \wedge e_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq k\}$$
is a basis of $\bigwedge^r(V)$. Thus $\dim \bigwedge^r(V) = \binom{k}{r}$ and $\dim \bigwedge(V) = 2^k$.

Solution. We start by noticing that an element $A \in \mathrm{Alt}_r(V)$ has components (with respect to the chosen basis on V) given by

$$A_{i_1 \cdots i_r} = A(e_{i_1}, \ldots, e_{i_r}) \quad \text{for every } 1 \leq i_1 \leq \cdots \leq i_r \leq k.$$

Since A is alternating, we deduce that if σ is a permutation of any set of indices $\{i_1, \ldots, i_r\}$, then $A_{\sigma(i_1) \cdots \sigma(i_r)} = \mathrm{sign}(\sigma) A_{i_1 \cdots i_r}$ ($\mathrm{sign}(\sigma)$ is the signature of σ). In particular A is determined completely by the values $A_{i_1 \cdots i_r}$ for $1 \leq i_1 < i_2 < \cdots i_r \leq k$ and consequently a basis for $\mathrm{Alt}_r(V)$ is given, for this choice of indices, by the maps

$$E^{i_1, \ldots, i_r}(e_{j_1}, \ldots, e_{j_r}) = \begin{cases} \mathrm{sign}(\eta) & \text{if } \{i_1, \ldots, i_r\} = \{j_1, \ldots, j_r\}, \\ 0 & \text{else,} \end{cases}$$

for every $1 \leq j_1 \leq \ldots \leq j_r$ where $\mathrm{sign}(\eta)$ is the signature of the permutation sending $(i_1, \ldots, i_r)$ to $(j_1, \ldots j_r)$. With the canonical vector-space identification between $\mathrm{Alt}_r(V)$ and $\bigwedge^r(V^*)$ given by Proposition 15.23, $E^{i_1, \ldots, i_r}$ corresponds to the element $e^{i_1} \wedge \cdots \wedge e^{i_r}$. Consequently, $\{e^{i_1} \wedge \cdots \wedge e^{i_r}\}_{1 \leq i_1 < \cdots < i_r \leq k}$ must then be a basis for $\bigwedge^r(V^*)$.

The same reasoning can be applied to $\mathrm{Alt}_r(V^*)$ and $\bigwedge^r(V)$ (or also simply recalling the canonical identification $V \cong V^{**}$), and so the conclusion is reached.

Finally, it is clear that $\bigwedge^r(V) = 0$ for $r > k$ since eg. $e_1 \wedge \cdots \wedge e_n \wedge e_i = 0$ for any i as the e_i term appears twice.

PROBLEM H.4. Let M be a smooth manifold and suppose $\pi_i \colon E_i \to M$ are two vector bundles over M of the same rank k. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both[1] E_1 and E_2 admit $\mathrm{GL}(k)$-bundle atlases over the $U_{\mathtt{a}}$. Let

$$\rho^1_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k), \qquad \text{and} \qquad \rho^2_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to \mathrm{GL}(k)$$

denote the transition functions of E_1 and E_2 with respect to these bundle atlases. Prove that E_1 and E_2 are isomorphic if and only if there exists a smooth family $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to \mathrm{GL}(k)$ of functions such that

$$\nu_{\mathtt{a}}(x) \circ \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \circ \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}. \tag{H.1}$$

SOLUTION. Suppose first that we are given a vector bundle isomorphism $\Phi \colon E_1 \xrightarrow{\sim} E_2$. For $i = 1, 2$ and each $\mathtt{a} \in \mathtt{A}$, we denote the corresponding bundle chart $\pi_i^{-1}(U_{\mathtt{a}}) \to \mathbb{R}^k$ by $\mathtt{a}^i$. Let $x \in U_{\mathtt{a}}$ and define

$$\nu_{\mathtt{a}}(x) := \mathtt{a}^2 \circ \Phi \circ (\mathtt{a}^1|_{E_{1,x}})^{-1}.$$

Since each $\mathtt{a}^i|_{E_{i,x}} \colon E_{i,x} \xrightarrow{\sim} \mathbb{R}^k$ is a linear isomorphism, it follows that $\nu_{\mathtt{a}}(x) \in \mathrm{GL}(k)$. Let $\mathtt{a}, \mathtt{b} \in \mathtt{A}$ be such that $U_{\mathtt{a}} \cap U_{\mathtt{b}} \neq \emptyset$. We compute

$$\begin{aligned}
\nu_{\mathtt{a}}(x) \circ \rho^1_{\mathtt{ab}}(x) &= \big(\mathtt{a}^2 \circ \Phi \circ (\mathtt{a}^1|_{E_{1,x}})^{-1}\big) \circ \big(\mathtt{a}^1 \circ (\mathtt{b}^1|_{E_{1,x}})^{-1}\big) \\
&= \mathtt{a}^2 \circ \Phi \circ (\mathtt{b}^1|_{E_{1,x}})^{-1} \\
&= \mathtt{a}^2 \circ (\mathtt{b}^2|_{E_{2,x}})^{-1} \circ \mathtt{b}_2 \circ \Phi \circ (\mathtt{b}^1|_{E_{1,x}})^{-1} \\
&= \rho^2_{\mathtt{ab}}(x) \circ \nu_{\mathtt{b}}(x).
\end{aligned}$$

It follows that the $\nu_{\mathtt{a}}$ satisfy (H.1), as desired.

Conversely, suppose we are given a smooth family of functions $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to \mathrm{GL}(k)$ satisfying (H.1). For each $\mathtt{a} \in \mathtt{A}$, define $\tilde{\Phi}_{\mathtt{a}} \colon U_{\mathtt{a}} \times \mathbb{R}^k \to U_{\mathtt{a}} \times \mathbb{R}^k$ by $(x, v) \mapsto (x, \nu_{\mathtt{a}}(x)v)$. We then define $\Phi \colon E_1 \to E_2$ as follows: for $e_1 \in E_1$, choose $\mathtt{a} \in \mathtt{A}$ such that $e_1 \in \pi_1^{-1}(U_{\mathtt{a}})$ and define

$$\Phi(e_1) := \big((\pi_2, \mathtt{a}^2)^{-1} \circ \tilde{\Phi}_{\mathtt{a}} \circ (\pi_1, \mathtt{a}^1)\big)(e_1). \tag{H.2}$$

We first check that Φ is well-defined, i.e., independent of the choice of $\mathtt{a} \in \mathtt{A}$. Let $\mathtt{b} \in \mathtt{A}$ be another element such that $e_1 \in \pi_1^{-1}(U_{\mathtt{b}})$. Let $x := \pi_1(e_1) \in U_{\mathtt{a}} \cap U_{\mathtt{b}}$. Then it follows from the definition of each map in the composite that the right hand side of (H.2) is equal to $(a^2|_{E_{2,x}})^{-1}\nu_{\mathtt{a}}(x)a^1|_{E_{1,x}}(e_1)$. Checking that Φ is well-defined therefore amounts to checking that the following equality holds:

$$(\mathtt{a}^2|_{E_{2,x}})^{-1}\nu_{\mathtt{a}}(x)\mathtt{a}^1|_{E_{1,x}}(e_1) = (\mathtt{b}^2|_{E_{2,x}})^{-1}\nu_{\mathtt{b}}(x)\mathtt{b}^1|_{E_{1,x}}(e_1) \tag{H.3}$$

Precomposing both sides by $\mathtt{a}^2|_{E_{2,x}}$ and recalling that $\rho^i_{\mathtt{ab}}(x) := \mathtt{a}^i|_{E_{i,x}} \circ (\mathtt{b}^i|_{E_{i,x}})^{-1}$, we find that (H.3) is equivalent to (H.1), which holds by assumption. Thus Φ is well-defined.

It remains to show that the smooth map $\Phi \colon E_1 \to E_2$ is a vector bundle isomorphism. This amounts to checking that $\Phi|_{E_{1,x}}$ is a linear isomorphism from $E_{1,x}$ to $E_{2,x}$ for all $x \in M$. Fix $x \in M$ and choose $\mathtt{a} \in \mathtt{A}$ such that $x \in U_{\mathtt{a}}$. Then by construction we have $\Phi|_{E_{1,x}} = (\mathtt{a}^2|_{E_{2,x}})^{-1}\nu_{\mathtt{a}}(x)\mathtt{a}^1|_{E_{1,x}}$, and the right hand side is clearly a linear isomorphism. This concludes the proof.

[1]This can always be achieved by taking intersections.

(♣) Problem H.5. Let $\varphi\colon M \to N$ be a smooth map and suppose $\pi\colon E \to N$ is a vector bundle, which we illustrate pictorially as:

$$\begin{tikzcd} & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd} \tag{δ}$$

A **solution** of the diagram (δ) is a vector bundle $\pi_1\colon E_1 \to M$ over M together with a vector bundle morphism $\Phi\colon E_1 \to E$ along φ. Thus a solution is any pair $\pi_1\colon E_1 \to M$ and Φ such that the following commutes:

$$\begin{tikzcd} E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

As we have seen, one possible solution is the pullback bundle $\varphi^\star E$:

$$\begin{tikzcd} \varphi^\star E \arrow[r, "\mathrm{pr}_2"] \arrow[d, "\mathrm{pr}_1"'] & E \arrow[d, "\pi"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

The aim of this exercise is to prove that $\varphi^\star E$ is the "most efficient" solution in the following sense: Suppose $\pi_1\colon E_1 \to E$ and Φ is *any* solution to (δ). Prove there exists a *unique* vector bundle homomorphism $\Psi\colon E_1 \to \varphi^\star E$ such that the following diagram commutes:

$$\begin{tikzcd} E_1 \arrow[dr, dashed, "\Psi"] \arrow[drr, "\Phi"] \arrow[ddr, "\pi_1"'] & & \\ & \varphi^\star E \arrow[r, "\mathrm{pr}_2"'] \arrow[d, "\mathrm{pr}_1"] & E \arrow[d, "\pi"] \\ & M \arrow[r, "\varphi"'] & N \end{tikzcd} \tag{Δ}$$

Prove moreover that $\varphi^\star E$ is uniquely determined by this property. Explicitly this means that if $\tilde{\pi}\colon \tilde{E} \to M$ and $\tilde{\Phi}$ is another solution to the diagram (δ) with the property that for *any* solution $\pi_1\colon E_1 \to M$ and Φ of (δ) there exists a *unique* vector bundle homomorphism $\tilde{\Psi}\colon E_1 \to \tilde{E}$ such that the following commutes:

$$\begin{tikzcd}
E_1 \arrow[dr, dashed, "\tilde{\Psi}"] \arrow[drr, "\Phi"] \arrow[ddr, "\pi_1"'] & & \\
& \tilde{E} \arrow[r, "\tilde{\Phi}"'] \arrow[d, "\tilde{\pi}"] & E \arrow[d, "\pi"] \\
& M \arrow[r, "\varphi"'] & N
\end{tikzcd}$$

then in fact $\tilde{E}$ is isomorphic as a vector bundle over M to $\varphi^\star E$. *Hint:* Argue as in the proof of Lemma 15.2.

Solution. In the language of category theory we are looking for a certain bundle morphism (Ψ, id_M) between the two objects $\pi_1\colon E_1 \to M$ and $\mathrm{pr}_1\colon \varphi^* E \to M$ in the category of vector bundles $\mathsf{VectBundles}$. Suppose first that there exists a bundle morphism (Ψ, id_M) fitting into the diagram from the statement (modulo the uniqueness part), i.e.

$$\Psi\colon E_1 \to \varphi^* E,$$

is smooth and satisfies

$$\begin{cases} \mathrm{pr}_2 \circ \Psi & = \Phi \\ \mathrm{pr}_1 \circ \Psi & = \pi_1. \end{cases}$$

The definition of the pullback bundle $\varphi^* E$ and the fact that Ψ takes values in $\varphi^* E$ imply

$$\forall p \in E_1\colon\; \Psi(p) = (\mathrm{pr}_1(\Psi(p)), \mathrm{pr}_2(\Psi(p)))\,.$$

With the commutative relations above this becomes

$$\Psi(p) = (\pi_1(p), \Phi(p))\,,$$

which proves uniqueness. Existence follows by reading the last few lines backwards and finishes the first part of the exercise.

For the second part of the exercise we pick the vector bundle $\tilde{\pi}\colon \tilde{E} \to M$ together with the bundle morphism $(\tilde{\Phi}, \varphi)$ just as in the statement. In particular, the vector bundle $\tilde{\pi}\colon \tilde{E} \to M$ is a solution of (δ) and so the first part of the exercise grants us the following commutative diagram

$$\begin{tikzcd}
\tilde{E} \arrow[dr, dashed, "\tilde{\Psi}"] \arrow[drr, "\Phi"] \arrow[ddr, "\tilde{\pi}"'] & & \\
& \varphi^\star E \arrow[r, "\mathrm{pr}_2"'] \arrow[d, "\mathrm{pr}_1"] & E \arrow[d, "\pi"] \\
& M \arrow[r, "\varphi"'] & N
\end{tikzcd} \tag{$\tilde{\Delta}$}$$

Due to the assumptions on $\tilde{\pi}\colon \tilde{E} \to M$ however, we can "stick" yet another commutative diagram on top of ($\tilde{\Delta}$), namely:

$$\begin{tikzcd}
\varphi^* E \arrow[dr, dashed, "\tilde\Theta"] \arrow[ddrrr, bend left, "\mathrm{pr}_2"] \arrow[dddrr, bend right, "\mathrm{pr}_1"'] & & & \\
& \tilde E \arrow[dr, dashed, "\tilde\Psi"] \arrow[drr, bend left, "\Phi"] \arrow[ddr, "\tilde\pi"'] & & \\
& & \varphi^\star E \arrow[r, "\mathrm{pr}_2"'] \arrow[d, "\mathrm{pr}_1"] & E \arrow[d, "\pi"] \\
& & M \arrow[r, "\varphi"'] & N
\end{tikzcd} \tag{$\tilde\Delta$}$$

This means that $\mathrm{pr}_1\colon \varphi^* E \to M$ and pr_2 is the *unique* solution of (δ) by the first part of the proof which shows

$$\tilde{\Psi} \circ \tilde{\Theta} = \mathrm{id}_{\varphi^* E},$$

simply because $\mathrm{pr}_1\colon \varphi^* E \to M$ and $\mathrm{id}_{\varphi^* E}$ also solves (δ).

Similarly, reversing the roles of $\varphi^* E$ and $\tilde{E}$[2] then proves

$$\tilde{\Theta} \circ \tilde{\Psi} = \tilde{\mathrm{E}}\,.$$

This concludes the prove.

Problem H.6. Let $\pi_i\colon E_i \to M_i$ be vector bundles for $i = 1, 2$. Suppose $\Theta\colon E_1 \to E_2$ is any smooth map that maps each fibre $\pi_1^{-1}(x)$ linearly onto some fibre $\pi_2^{-1}(y)$ for $x \in M_1$ and $y \in M_2$. Prove that $\Theta = \Phi \circ \Psi$ where Ψ is a vector bundle homomorphism and Φ is a vector bundle morphism along a map $M_1 \to M_2$.

Solution. Note first that any vector bundle $\pi_1\colon E_1 \to M$ has a (canonical) section $s \in \Gamma(E_1 \to M)$ given by $s(x) = (x, 0) \in \pi_1^{-1}(x)$. I.e. s maps every point in M to the 0-element in $\pi_1^{-1}(x)$. The map $\varphi \coloneqq \pi_2 \circ \Phi \circ s$ is a smooth map $M_1 \to M_2$ which is "covered" by Θ. Now note that there is a natural induced map $\Psi\colon E_1 \to \varphi^\star E_2$ which covers the identity $M_1 \to M_1$. This map is simply given by $\Psi(x, v) = (x, P \circ \Theta(x, v))$. Here P is the map which sends an element of $z \in E_2$ to the vector sitting over $\pi_2(z)$. With this definition we clearly have

$$\begin{tikzcd}
E_1 \arrow[rr, "\Psi"] \arrow[dr, "\pi_1"'] & & \varphi^\star E_2 \arrow[dl, "\pi"] \\
& M_1 &
\end{tikzcd}$$

where $\pi\colon \varphi^\star E_2 \to M$ is induced projection map. Now of course we have a map $\Phi\colon \varphi^\star E_2 \to E_2$ covering φ. This map is simply given by $\varphi^\star E_2 \ni (x, v) \mapsto (\varphi(x), v) \in$

[2] Note that reversing the roles also uses the first part of the proof. Why?

E_2 (here we use the notation that x is the basepoint coordinate and v is the fibre coordinate). Now everything fits into a commutative diagram as wanted:

$$\begin{tikzcd} E_1 \arrow[r, "\Psi"] \arrow[dr, "\pi_1"'] & \varphi^\star E_2 \arrow[r, "\Phi"] \arrow[d, "\pi"'] & E_2 \arrow[d, "\pi_2"] \\ & M_1 \arrow[r, "\varphi"] & M_2. \end{tikzcd}$$

PROBLEM H.7. Let $\pi_i\colon E_i \to M$ be two vector bundles over the same manifold M of ranks k_i. Let $\Phi\colon E_1 \to E_2$ be a vector bundle homomorphism.

(i) Assume Φ is injective on each fibre. Consider the quotient vector space

$$\bar{E}_x := E_2|_x / \Phi|_{E_1|_x}(E_1|_x).$$

Prove that $\bar{E} := \bigsqcup_{x\in M} \bar{E}_x$ is a vector bundle of rank $k_2 - k_1$.

(ii) Assume that Φ is surjective on each fibre. Let

$$K_x := \ker \Phi|_{E_x} \subset E_1|_x.$$

Prove that $K := \bigsqcup_{x\in M} K_x$ is a vector bundle over M of rank $k_1 - k_2$.

SOLUTION. (i) Choose a covering $\mathcal{U} = \{U\}$ of open subsets of M such that E_2 is trivial over each U. I.e. we have diffeomorphisms

$$\begin{tikzcd} \pi_2^{-1}(U) \arrow[rr, "\cong"] \arrow[dr, "\pi_2"'] & & U \times \mathbb{R}^{k_2} \arrow[dl, "\text{projection}"] \\ & U & \end{tikzcd}$$

which restrict to linear isomorphisms on fibres. Since the restriction of Φ to fibres is assumed to be an injective map, $\Phi(E_1|_x) \subset E_2|_x$ is a k_1-dimensional linear subspace. Over a subset U we may therefore view $\Phi(E_1|_x)$ as a k_1-dimensional linear subspace of $\mathbb{R}^{k_2}$. Hence, we may for each point $x \in U$ choose a basis $v_1(x), \dots, v_{k_1}(x)$ of $\Phi(E_1|_x)$. By perhaps restricting to an open subset $V \subset U$, this basis can in fact be chosen in such a way that they depend smoothly on $x \in V$. By the Gram-Schmidt orthogonalisation procedure this set of vectors $v_1(x), \dots, v_{k_1}(x)$ can be extended to a basis $v_1(x), \dots, v_{k_1}(x), v_{k_1+1}(x), \dots v_{k_2}(x)$ for $\mathbb{R}^{k_2}$ and this basis depends smoothly on $x \in V$.[3] The set of vectors $v_{k_1+1}(x), \dots v_{k_2}(x)$ spans $E_2|_x/\Phi(E_1|_x)$ for each $x \in V$. Since we can cover M by such subsets V we conclude that $\overline{E}$ is a vector bundle of rank $k_2 - k_1$ over M.

(ii) Fix a point $x_0 \in M$ and choose an open neighbourhood U of x such that both E_1 and E_2 are trivial over U. For all $x \in U$ we can then view $\Phi|_x\colon E_1|_x \to E_2|_x$ as a surjective linear map $\mathbb{R}^{k_1} \to \mathbb{R}^{k_2}$, and hence it is represented by a matrix A_x. In particular, $A_{x_0}\colon \mathbb{R}^{k_1} \to \mathbb{R}^{k_2}$ is surjective, so there exists $v_1, \dots v_{k_2} \in \mathbb{R}^{k_1}$

[3] These sets V are determined by an open covering on M over which E_1 is trivial.

such that $A_{x_0}(v_1), \dots A_{x_0}(v_{k_2}) \in \mathbb{R}^{k_2}$ is a basis. By continuity it follows that also

$$A_x(v_1), \dots A_x(v_{k_2}) \in \mathbb{R}^{k_2}$$

is a basis for all $x \in V$, where V is a sufficiently small neighbourhood of x_0. Set $E \coloneqq \operatorname{Span}(v_1, \dots v_{k_2}) \leq \mathbb{R}^{k_1}$. The concatenation (over V)

$$K \hookrightarrow V \times \mathbb{R}^{k_1} \to V \times (\mathbb{R}^{k_1}/E)$$

of the inclusion and the quotient map is a fibrewise linear isomorphism. Since $x_0 \in M$ was arbitrary we have shown that K is locally trivial and admits a smooth structure, so it is a vector bundle of rank $k_1 - k_2$.

Problem Sheet I

Problem I.1. Let $\pi\colon E \to M$ be a vector bundle. An operator $\chi\colon \Gamma(E) \to \Gamma(E)$ is said to satisfy the **Leibniz rule** if there exists a vector field X on M such that for any $f \in C^\infty(M)$ and $s \in \Gamma(E)$ one has

$$\chi(fs) = (Xf)s + f\chi(s).$$

Prove that an operator satisfying the Leibniz rule is a local operator but not a point operator.

Problem I.2. Let M be a smooth manifold and let $E_1, \dots, E_r$ and E be vector bundles over M. Let $\chi\colon \Gamma(E_1) \times \cdots \times \Gamma(E_r) \to \Gamma(E)$ be a $C^\infty(M)$-multilinear operator. Prove that for each $x \in M$ there is a unique $\mathbb{R}$-multilinear map

$$\Phi_x \colon E_1|_x \times \cdots \times E_r|_x \to E_x$$

such that for all $s_i \in \Gamma(E_i)$ one has

$$\Phi_x(s_1(x), \dots s_r(x)) = \chi(s_1, \dots, s_r)(x).$$

Problem I.3. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Let $r, s \geq 0$. Prove that there is a one-to-one correspondence between tensor fields on W of type (r, s) and $C^\infty(W)$-multilinear functions

$$A\colon \underbrace{\Omega^1(W) \times \cdots \times \Omega^1(W)}_{r} \times \overbrace{\mathfrak{X}(W) \times \cdots \times \mathfrak{X}(W)}^{s} \to C^\infty(W).$$

This generalises Corollary 16.29. *Hint:* Use the previous problem.

Problem I.4. Let M be a smooth manifold and let $\pi\colon E \to M$ be a vector bundle over M. Prove that both the presheaf $\mathcal{C}^\infty_M$ of smooth functions on M and the presheaf $\mathcal{E}_E$ of sections of E are in fact sheaves.

(♣) Problem I.5. This problem introduces the notion of a **vertical bundle**.

(i) Let $\pi\colon E \to M$ be a fibre bundle with fibre F. Assume F has dimension k as a manifold. Let

$$VE := \bigsqcup_{p \in E} \{\ker D\pi(p)\colon T_pE \to T_{\pi(p)}M\}$$

with projection map $\pi_V\colon VE \to E$. Prove that VE is a vector bundle over E of rank k.

(ii) Assume now that $\pi\colon E \to M$ is a vector bundle. Prove that the vertical bundle VE is isomorphic as a vector bundle to the pullback bundle $\pi^\star E \to E$.

(iii) Continue to assume that $\pi\colon E \to M$ is a vector bundle. Prove that the composite vector bundle[1] $\pi \circ \pi_V\colon VE \to M$ is isomorphic as a vector bundle over M to the direct sum bundle $E \oplus E$.

(iv) Continue to assume that $\pi\colon E \to M$ is a vector bundle. View M as an embedded submanifold of TM via the zero section. Prove that the composite bundle $\pi \circ \pi_V\colon VE \to M$ is a vector subbundle of $D\pi\colon TE \to TM$ in the sense of Example 14.8.

[1] See Example 13.20.

Solutions to Problem Sheet I

Problem I.1. Let $\pi\colon E \to M$ be a vector bundle. An operator $\chi\colon \Gamma(E) \to \Gamma(E)$ is said to satisfy the **Leibniz rule** if there exists a vector field X on M such that for any $f \in C^\infty(M)$ and $s \in \Gamma(E)$ one has

$$\chi(fs) = (Xf)s + f\chi(s).$$

Prove that an operator satisfying the Leibniz rule is a local operator but not a point operator.

Solution. To prove that χ is a local operator, let $U \subset M$ be an open set and let $s \in \Gamma(E)$ be a section of E vanishing on U. Fix any $x_0 \in U$ and let $f \in C^\infty(M)$ be so that

$$f(x) = \begin{cases} 0 & \text{at } x = x_0, \\ 1 & \text{in } M \setminus U. \end{cases}$$

Such function can be constructed since $\{x_0\}$ is a compact subset of U, and since s vanishes on U, we have

$$f\,s \equiv s \quad \text{in } M.$$

Consequently,

$$\chi(s)(x_0) = \chi(fs)(x_0) = \underbrace{X(f)(x_0)}_{=0}\,s(x_0) + \underbrace{f(x_0)}_{=0}\,\chi(s)s(x_0) = 0,$$

and since x_0 is arbitrarily chosen on U, we conclude that $\chi(s)$ vanishes on U.

To prove that, unless $X \equiv 0$, χ is never a point operator, let x_0 be any fixed point in M where $X(x_0) \neq 0$ and let e be a section of E so that $e(x_0) \neq 0$. Since X is not zero at x_0, we can always construct a function $f \in C^\infty(M)$ so that $f(x_0) = 0$ and $X(f)(x_0) \neq 0$, as follows: let (U, σ) be local chart at x_0 so that $\sigma(x_0) = 0$ and in whose local coordinates $x^1, \ldots, x^m$ X is written as $X(x) = \frac{\partial}{\partial x^1}\big|_x$ (Corollary 11.2). The required function f is then defined in this coordinate patch as $f(x^1, \ldots, x^m) = x^1$, and then extended arbitrarily to a smooth function on all of M by means of a partition-of-unity argument. Now consider the section $s \in \Gamma(E)$ given by $s = f\,e$: we see that

$$s(x_0) = f(x_0)\,e(x_0) = 0, \quad \text{but} \quad \chi(s)(x_0) = X(f)(x_0)\,e_0(x_0) = e(x_0) \neq 0,$$

so χ is not a point operator.

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

Problem I.2. Let M be a smooth manifold and let $E_1, \ldots, E_r$ and E be vector bundles over M. Let $\chi\colon \Gamma(E_1) \times \cdots \times \Gamma(E_r) \to \Gamma(E)$ be a $C^\infty(M)$-multilinear operator. Prove that for each $x \in M$ there is a unique $\mathbb{R}$-multilinear map

$$\Phi_x\colon E_1|_x \times \cdots \times E_r|_x \to E_x$$

such that for all $s_i \in \Gamma(E_i)$ one has

$$\Phi_x(s_1(x), \ldots s_r(x)) = \chi(s_1, \ldots, s_r)(x).$$

Solution. Fix $x \in M$ and let $v_k \in E_k|_x$. Extend each v_k to a global section $s_k \in \Gamma(E_k)$ (here we use Lemma 16.16). Define

$$\Phi_x(v_1, \ldots, v_r) \coloneqq \chi(s_1, \ldots s_r)(x). \tag{I.1}$$

We claim that this definition is independent of the extensions s_k of v_k. To see this, let $\hat{s}_k$ be another extension of v_k. Then

$$\chi(s_1, \ldots, s_k - \hat{s}_k, \ldots s_r)(x) = 0$$

because χ is a point operator (Proposition 16.25) and $s_k(x) - \hat{s}_k(x) = v_k - v_k = 0$. Now multilinearity of χ implies

$$\chi(s_1, \ldots, s_k, \ldots s_r)(x) = \chi(s_1, \ldots, \hat{s}_k, \ldots s_r)(x).$$

This shows that (I.1) defines Φ_x independently of the extensions s_k. Clearly, $\mathbb{R}$-multilinearity of Φ_x follows from $C^\infty(M)$-multilinearity of χ.

Problem I.3. Let M be a smooth manifold and let $W \subset M$ be a non-empty open set. Let $r, s \geq 0$. Prove that there is a one-to-one correspondence between tensor fields on W of type (r, s) and $C^\infty(W)$-multilinear functions

$$\tilde{T}\colon \underbrace{\Omega^1(W) \times \cdots \times \Omega^1(W)}_{r} \times \overbrace{\mathfrak{X}(W) \times \cdots \times \mathfrak{X}(W)}^{s} \to C^\infty(W).$$

Hint: Use the previous problem.

Solution. The idea is to mimic the proof of Corollary 16.29 with Problem I.2 taking the role of Proposition 16.28. We start with $\tilde{T}$ a $C^\infty(W)$-multilinear function as in the statement. Recalling $\Gamma(TW) = \mathfrak{X}(W)$, $\Gamma(T^*W) = \Omega^1(W)$ and $\Gamma(W \times \mathbb{R}) = C^\infty(W)$ we can invoke Problem I.2 to get, for every $x \in W$, a *unique* $\mathbb{R}$-multilinear map

$$\Phi_x\colon \underbrace{T_x^*W \times \cdots T_x^*W}_{r} \times \overbrace{T_xW \times \cdots \times T_xW}^{s} \to \mathbb{R},$$

such that for every tuple of 1-forms $(\omega_1, \ldots, \omega_r)$ and vector fields $(X_1, \ldots, X_s)$ on W one has

$$\Phi_x\left(\omega_1(x), \ldots, \omega_r(x), X_1(x), \ldots, X_s(x)\right) = \tilde{T}\left(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s\right)(x).$$

For each $x \in W$ we have $\Phi_x \in \mathrm{Mult}_{r,s}(T_x^*W) \cong \mathrm{Mult}_{s,r}(T_xW) \cong T^{r,s}(T_xW)$ (cf. Proposition 15.9). Just as in Corollary 16.29 we can use Remark 16.9 to deduce that the map

$$\Phi\colon W \to T^{r,s}(TW),\ x \mapsto \Phi_x$$

is smooth and satisfies the section property, thus proving that Φ is a tensor field on W of type (r, s), i.e.

$$\Phi \in \mathcal{T}^{r,s}(W).$$

PROBLEM I.4. Let M be a smooth manifold and let $\pi\colon E \to M$ be a vector bundle over M. Prove that both the presheaf $\mathcal{C}_M^\infty$ of smooth functions on M and the presheaf $\mathcal{E}_E$ of sections of E are in fact sheaves.

SOLUTION. We start with proving that the $\mathbb{R}$-valued presheaf $\mathcal{C}_M^\infty$ on M defines a sheaf. Let $U \subset M$ by any open set, $\{U_a \mid a \in \mathtt{A}\}$ an open cover of U and

$$f\colon U \to \mathbb{R}$$

any function such that the all the restrictions $f|_{U_a}\colon U_a \to \mathbb{R}$ are smooth, i.e. $f|_{U_a} \in \mathcal{C}_M^\infty(U_a)$. In order to conclude that $\mathcal{C}_M^\infty$ defines a sheaf we need to show that $f\colon U \to \mathbb{R}$ is smooth (cf. Remark 17.10), but this basically follows from smoothness being a local property. Indeed, for any $x \in U$ we need to show that for any chart $\sigma\colon U_x \to \mathbb{R}^n$ around x the map

$$f \circ \sigma^{-1}\colon \mathbb{R}^n \to \mathbb{R}$$

is of class C^∞ in the usual calculus sense. But for $b \in A$ with $x \in U_b$ we already know that

$$f|_{U_b \cap U_x} \circ \sigma^{-1}|_{\sigma(U_b \cap U_x)}\colon \sigma(U_b \cap U_x) \to \mathbb{R}$$

is of class C^∞ as

$$f|_{U_a} \in \mathcal{C}_M^\infty(U_a) \implies \mathrm{res}^{U_a}_{U_a \cap U_x}(f|_{U_a}) = f_{U_a \cap U_x} \in \mathcal{C}_M^\infty(U_a \cap U_x)^{1},$$

which implies smoothness of f at x. This then implies smoothness of f and hence proves that $\mathcal{C}_M^\infty$ is a sheaf.

The proof that $\mathcal{E}_E$ defines a sheaf on M is very similar. Adopt the notation from above and pick any function

$$s\colon U \to E$$

such that for all $a \in \mathtt{A}$

$$s|_{U_a} \in \mathcal{E}_E(U_a).$$

We want to show that s belongs to $\mathcal{E}_E(U)$ which is equivalent to

$$s\colon U \to E \text{ smooth and } \pi \circ s = \mathrm{id}_M.$$

The smooth bit follows in an analogous way to what we did the first part of the proof. The section property is trivial as for every $x \in U$ there is a $b \in A$ such that $x \in U_b$ and therefore

$$(\pi \circ s)(x) = (\pi \circ s|_{U_b})(x) = x.$$

[1] This uses the fact that $\mathcal{C}_M^\infty$ is a presheaf. Of course, one could argue directly as to why $f|_{U_b \cap U_x}$ is smooth.

(♣) Problem I.5. This problem introduces the notion of a **vertical bundle**.

(i) Let $\pi\colon E \to M$ be a fibre bundle with fibre F. Assume F has dimension k as a manifold. Let

$$VE := \bigsqcup_{p\in E}\{\ker D\pi(p)\colon T_pE \to T_{\pi(p)}M\}$$

with projection map $\pi_V\colon VE \to E$. Prove that VE is a vector bundle over E of rank k.

(ii) Assume now that $\pi\colon E \to M$ is a vector bundle. Prove that the vertical bundle VE is isomorphic as a vector bundle to the pullback bundle $\pi^\star E \to E$.

(iii) Continue to assume that $\pi\colon E \to M$ is a vector bundle. Prove that the composite vector bundle[2] $\pi \circ \pi_V\colon VE \to M$ is isomorphic as a vector bundle over M to the direct sum bundle $E \oplus E$.

(iv) Continue to assume that $\pi\colon E \to M$ is a vector bundle. View M as an embedded submanifold of TM via the zero section. Prove that the composite bundle $\pi \circ \pi_V\colon VE \to M$ is a vector subbundle of $D\pi\colon TE \to TM$ in the sense of Example 14.8.

Solution. For (i), we define bundle charts on VE as follows: let $\alpha\colon \pi^{-1}(U) \to F$ be a bundle chart for E, and let $\sigma\colon \tilde{U} \xrightarrow{\sim} O \subset \mathbb{R}^k$ be a coordinate chart on F. We identify the tangent bundle TO with $O \times \mathbb{R}^k$ and define

$$\big(\alpha_V := \mathrm{pr}_2 \circ D\sigma \circ D\alpha\big)\colon \pi_V^{-1}\big(\pi^{-1}(U) \cap \alpha^{-1}(\tilde{U})\big) \to \mathbb{R}^k. \tag{I.2}$$

One then follows the procedure outlined in Remark 13.7 to endow VE with a smooth structure. To check that the resulting smooth manifold is a vector bundle over E, we must verify that the transition functions obtained from the α_V are linear isomorphisms. Consider a bundle chart $\beta\colon \pi^{-1}(V) \to F$ and coordinate chart $\tau\colon \tilde{V} \xrightarrow{\sim} O'$ such that

$$\big(\pi^{-1}(U) \cap \alpha^{-1}(\tilde{U})\big) \cap \big(\pi^{-1}(V) \cap \beta^{-1}(\tilde{V})\big) \neq \emptyset, \tag{I.3}$$

and define β_V in analogy to α_V. Consider an element p in the left hand side of (I.3). Let $x := \pi(p)$. By Proposition 5.15, the inclusion $\iota_x\colon E_x \to E$ yields an isomorphism $D\iota_x(p)\colon T_pE_x \xrightarrow{\sim} \ker D\pi(p) \subset T_pE$. We have the following commutative diagram

$$\begin{tikzcd}
T_pE_x \arrow[rr, "D\iota_x(p)", "\sim"'] \arrow[dr, "D(\alpha|_{E_x})(p)"'] & & \ker D\pi(p)^{D\alpha(p)} \arrow[dl, "D\alpha(p)"] \\
& T_{\alpha(p)}F. &
\end{tikzcd}$$

Since $\alpha|_{E_x}$ is a diffeomorphism, it follows that $D(\alpha|_{E_x})(p)|_{\ker D\pi(p)}$ is an isomorphism. We deduce from the above diagram that $D\alpha(p)|_{\ker D\pi(p)}$ is also an isomorphism. Since $\alpha_V|_{(VE)_p}\colon \big((VE_p) = \ker D\pi(p)\big) \to \mathbb{R}^k$ is simply the composite

[2]See Example 13.20.

$D\sigma\big(\alpha(p)\big) \circ D\alpha(p)|_{\ker D\pi(p)}$, we conclude that $\alpha_V|_{(VE)_p}$ is a linear isomorphism. The map $\beta_V|_{(VE)_p}$ has a similar description, and thus

$$\rho_{\alpha_V \beta_V}(p) = \alpha_V|_{(VE)_p} \circ (\beta_V|_{(VE)_p})^{-1}$$

is a linear isomorphism. We conclude that VE is a rank k vector bundle over E, as desired.

For (ii), we first note that by definition

$$\pi^* E = \{(p, q) \in E \times E \mid \pi(p) = \pi(q)\}. \tag{I.4}$$

In other words, elements of the pullback bundle π^*E consist of pairs $(p, q) \in E \times E$ such that $p, q \in E_x$ for some $x \in M$. Since each E_x is a vector space, for any $(p, q) \in \pi^*E$ the map $t \mapsto \pi(p + tq)$ from $\mathbb{R}$ to M is constant in t. It follows that $\frac{d}{dt}|_{t=0}(p + tq) \in \ker D\pi(p)$. We thus have a map

$$\begin{aligned} \mathcal{J} \colon \pi^* E &\to VE \\ (p, q) &\mapsto \frac{d}{dt}\Big|_{t=0} (p + tq). \end{aligned}$$

To see that $\mathcal{J}$ is an isomorphism of vector bundles, it suffices to show that it restricts to a linear isomorphism $(\pi^*E)_p \xrightarrow{\sim} (VE)_p$ for every $p \in E$. This is easy to see since the map $(p, q) \mapsto q$ identifies $(\pi^*E)_p$ with $E_{\pi(p)}$ and, again by Proposition 5.15, we may identify $(VE)_p = \ker D\pi(p)$ with $T_p E_{\pi(p)} \cong E_{\pi(p)}$. All of these identifications are linear, and after making these identifications, the map $\mathcal{J}|_{(\pi^*E)_p}$ becomes the identity map on $E_{\pi(p)}$. It follows that $\mathcal{J}|_{(\pi^*E)_p}$ is a linear isomorphism; hence $\mathcal{J}$ is an isomorphism of vector bundles.

For (iii), we first observe that $\pi^*E = E \oplus E$ as vector bundles over M. Indeed, as stated in the proof of (ii), for each element $(p, q) \in \pi^*E$, we have $p, q \in E_x$, where $x := \pi(p) = \pi(q)$. Hence we can view (p, q) as an element of $E_x \oplus E_x = (E \oplus E)_x$. Conversely, any $(p, q) \in E_x \oplus E_x$ can be viewed as an element of π^*E. We may thus identify π^*E and E as sets. To see that they are equal as vector bundles, it suffices to observe that they have the same bundle charts. Indeed, given a bundle chart $\alpha \colon \pi^{-1}(U) \to \mathbb{R}^k$ on E, the corresponding chart on $\mathrm{pr}_1 \colon \pi^*E \to E$ is given by

$$\alpha \circ \mathrm{pr}_2 \colon \mathrm{pr}_1^{-1}\big(\pi^{-1}(U)\big) \to \mathbb{R}^k,$$

which in turn corresponds to the bundle chart

$$(\alpha \circ \mathrm{pr}_1, \alpha \circ \mathrm{pr}_2) \colon \mathrm{pr}_1^{-1}\big(\pi^{-1}(U)\big) \to \mathbb{R}^k \times \mathbb{R}^k \tag{I.5}$$

on the composite bundle $\pi \circ \mathrm{pr}_1 \colon \pi^*E \to M$. Since

$$\mathrm{pr}_1^{-1}\big(\pi^{-1}(U)\big) = \pi^*\pi^{-1}(U) = \pi^{-1}(U) \oplus \pi^{-1}(U),$$

we see that (I.5) is precisely the bundle chart on $E \oplus E$ given by

$$(\alpha, \alpha) \colon \pi^{-1}(U) \oplus \pi^{-1}(U) \to \mathbb{R}^k \times \mathbb{R}^k.$$

Hence $\pi^*E = E \oplus E$ as vector bundles over M, as claimed.

We may thus view the map $\mathcal{J}$ from part (ii) as morphism $\mathcal{J}\colon E\oplus E\to VE$ of vector bundles over M. Since $\mathcal{J}$ is smooth, it only remains to show that $\mathcal{J}$ restricts to a linear isomorphism $E_x\oplus E_x\to(VE)_x$ for each $x\in M$. We first describe the vector space structure on $(VE)_x$ as a subspace of $(TE)_x$. The latter has vector space structure given by

$$\zeta+\xi\coloneqq Da(p,q)(\zeta,\xi)\in T_{p+q}E,\qquad p,q\in E_x,\quad \zeta\in T_pE,\quad \xi\in T_qE, \tag{I.6}$$

where $a\colon E\oplus E\to E$ is the bundle map given on fibers by $(p,q)\mapsto p+q$. For $(p,v),(q,w)\in E_x\oplus E_x$ and $\lambda\in\mathbb{R}$, we need to show that

$$\mathcal{J}(p+\lambda q,v+\lambda w)=\mathcal{J}(p,v)+\lambda\mathcal{J}(q,w).$$

This amounts to showing the equality

$$\frac{d}{dt}\Big|_{t=0}(p+\lambda q+t(v+\lambda w))=\frac{d}{dt}\Big|_{t=0}(p+tv)+\lambda\frac{d}{dt}\Big|_{t=0}(q+tw), \tag{I.7}$$

where the addition on the right-hand side is as defined in (I.6). We denote the path $t\mapsto p+tv$ by $\gamma^{p,v}\colon\mathbb{R}\to E$. Then by definition

$$\frac{d}{dt}\Big|_{t=0}(p+tq)=D\gamma^{p,v}(0)\Big(\frac{d}{dt}\Big|_{t=0}\Big).$$

Using the chain rule, we compute

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0}(p+tv)+\lambda\frac{d}{dt}\Big|_{t=0}(q+tw)&=Da(p,\lambda q)\Big[D\gamma^{p,v}(0)\Big(\frac{d}{dt}\Big|_{t=0}\Big),D\gamma^{\lambda q,\lambda w}(0)\Big(\frac{d}{dt}\Big|_{t=0}\Big)\Big]\\&=D\big(a\circ(\gamma^{p,v},\gamma^{\lambda q,\lambda w})\big)(0)\Big(\frac{d}{dt}\Big|_{t=0}\Big)\\&=D(\gamma^{p+\lambda q,v+\lambda w})(0)\Big(\frac{d}{dt}\Big|_{t=0}\Big)\\&=\frac{d}{dt}\Big|_{t=0}(p+\lambda q+t(v+\lambda w)).\end{aligned}$$

This yields (I.7). Thus $\mathcal{J}|_{E_x\oplus E_x}$ is linear. Since it is also bijective onto $(VE)_x$, it must be a linear isomorphism. Hence $\mathcal{J}\colon E\oplus E\to VE$ is a vector bundle isomorphism over M, as desired.

For (iv), first recall that we can view M as an embedded submanifold of TM via the 0-section

$$s_0\colon M\to TM,\quad x\mapsto(x,0).$$

Let $\alpha\colon\pi^{-1}(U)\to\mathbb{R}^k$ be a bundle chart on U. Note that in the case of vector bundles, the associated chart as defined in (I.2) has the simpler description

$$\alpha_V\coloneqq \mathrm{pr}_2\circ D\alpha\colon \pi_V^{-1}\big(\pi^{-1}(U)\big)\to\mathbb{R}^k.$$

The associated bundle chart of $\pi\circ\pi_V$ is defined to be $(\alpha\circ\pi_V,\mathrm{pr}_2\circ D\alpha)$. Since the π_V is just the restriction of the natural map $TE\to E$, it follows that $\alpha\circ\pi_V=\mathrm{pr}_1\circ D\alpha$. Thus the bundle chart $(\alpha\circ\pi_V,\mathrm{pr}_2\circ D\alpha)$ on the composite is simply

$$D\alpha\colon\pi_V^{-1}\big(\pi^{-1}(U)\big)\to T\mathbb{R}^k=\mathbb{R}^k\times\mathbb{R}^k.$$

Since $D\alpha$ is also a bundle chart for $D\pi\colon TE\to TM$, this shows that the natural inclusion $VE\to TM$ is a vector bundle morphism along $s_0\colon M\to TM$, i.e., that $\pi\circ\pi_V\colon VE\to M$ is a vector subbundle of $D\pi\colon TE\to TM$.

Problem Sheet J

(♣) PROBLEM J.1. Let M be a smooth manifold of dimension n.

(i) Let $A \in \mathcal{T}^{r,s}(M)$ denote a tensor of type (r,s). Let $\sigma\colon U \to O$ and $\tau\colon V \to \Omega$ denote two charts on M with $U \cap V \neq \emptyset$. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$A = f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \qquad \text{on } U$$

and

$$A = g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial y^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial y^{i_r}} \otimes dy^{j_1} \otimes \cdots \otimes dy^{j_s}, \qquad \text{on } V,$$

for smooth functions $f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(U)$ and $g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(V)$. Investigate the relationship between

$$f^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V} \qquad \text{and} \qquad g^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V}.$$

(ii) Let $\omega \in \Omega^r(M)$ denote a differential r-form. Let $\sigma\colon U \to O$ and $\tau\colon V \to \Omega$ denote two charts on M with $U \cap V \neq \emptyset$. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\omega = f_{i_1 \cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}, \qquad \text{on } U$$

and

$$\omega = g_{i_1 \cdots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}, \qquad \text{on } V,$$

for smooth functions $f_{i_1 \cdots i_r} \in C^\infty(U)$ and $g_{i_1 \cdots i_r} \in C^\infty(V)$. Investigate the relationship between

$$f_{i_1 \cdots i_r}|_{U \cap V} \qquad \text{and} \qquad g_{i_1 \cdots i_r}|_{U \cap V}.$$

(iii) Suppose $\varphi\colon M \to N$ is a diffeomorphism. Let $A \in \mathcal{T}^{r,s}(N)$. Let $\sigma\colon U \to O$ denote a chart on M and $\tau\colon V := \varphi(U) \to \Omega$ denote a chart on N. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\varphi^\star(A) = f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \qquad \text{on } U$$

and

$$A = g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial y^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial y^{i_r}} \otimes dy^{j_1} \otimes \cdots \otimes dy^{j_s}, \qquad \text{on } V,$$

for smooth functions $f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(U)$ and $g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(V)$. Investigate the relationship between

$$f^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V} \qquad \text{and} \qquad g^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V}.$$

(iv) Suppose $\varphi\colon M \to N$ is a smooth map. Let $\omega \in \Omega^r(N)$. Let $\sigma\colon U \to O$ denote a chart on M and $\tau\colon V := \varphi(U) \to \Omega$ denote a chart on N. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\varphi^\star(\omega) = f_{i_1\cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}, \qquad \text{on } U$$

and

$$\omega = g_{i_1\cdots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}, \qquad \text{on } V,$$

for smooth functions $f_{i_1\cdots i_r} \in C^\infty(U)$ and $g_{i_1\cdots i_r} \in C^\infty(V)$. Investigate the relationship between

$$f_{i_1\cdots i_r}|_{U\cap V} \qquad \text{and} \qquad g_{i_1\cdots i_r}|_{U\cap V}.$$

(v) Conclude that local coordinates are horrible.

PROBLEM J.2. Let $\varphi\colon M \to N$ denote a smooth map. Let $A \in \mathcal{T}^{0,s}(N)$. Using the Tensor Criterion (Theorem 18.3), regard A as a $C^\infty(N)$-multilinear function

$$\underbrace{\mathfrak{X}(N) \times \cdots \times \mathfrak{X}(N)}_{s} \to C^\infty(N).$$

and similarly regard $\varphi^\star(A)$ as a $C^\infty(M)$-multilinear function

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{s} \to C^\infty(M).$$

Suppose $X_i \in \mathfrak{X}(M)$ is φ-related to $Y_i \in \mathfrak{X}(N)$ for $i = 1, \dots, s$. Prove that

$$\varphi^\star(A)(X_1, \dots, X_s) = A(Y_1, \dots, Y_s) \circ \varphi$$

as functions $M \to N$.

PROBLEM J.3. Let V be a vector space and suppose $\omega \in \bigwedge^r(V^*)$ and $\vartheta \in \bigwedge^s(V^*)$. Let $v_i \in V$ for $i = 1, \dots, r+s$. Identify ω with an element of $\mathrm{Alt}_r(V)$, ϑ with an element of $\mathrm{Alt}_s(V)$ and $\omega \wedge \vartheta$ with an element of $\mathrm{Alt}_{r+s}(V)$ (using Proposition 15.23). Prove that:

$$(\omega \wedge \vartheta)(v_1, \dots, v_{r+s}) = \frac{1}{r!s!} \sum_{\varrho \in \mathfrak{S}_{r+s}} \operatorname{sgn}(\varrho)\omega\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big)\, \vartheta\big(v_{\varrho(r+s)}, \dots v_{\varrho(r+s)}\big)$$

or equivalently

$$(\omega \wedge \vartheta)(v_1, \dots, v_{r+s}) = \sum_{\varrho \in \mathrm{Shuffle}(r,s)} \operatorname{sgn}(\varrho)\omega\big(v_{\varrho(1)}, \dots, v_{\varrho(r)}\big)\, \vartheta\big(v_{\varrho(r+s)}, \dots v_{\varrho(r+s)}\big),$$

where $\mathrm{Shuffle}(r, s)$ was defined in Definition 19.3.

PROBLEM J.4. Let (r, s), (r', s') and (r'', s'') be three pairs of non-negative integers. Suppose we are given a $\mathcal{C}^\infty_M$-bilinear sheaf homomorphism

$$\mathcal{A}\colon \mathcal{T}^{r,s}_M \times \mathcal{T}^{r',s'}_M \to \mathcal{T}^{r'',s''}_M.$$

Assume in addition that $\mathcal{A}$ has the property that if $\varphi\colon U \to V$ is a local diffeomorphism between open sets of M then

$$\varphi^\star(\mathcal{A}_V(A,B)) = \mathcal{A}_U(\varphi^\star(A), \varphi^\star(B)).$$

Prove that for every vector field X on M, one has

$$\tilde{\mathcal{L}}_X(\mathcal{A}(A,B)) = \mathcal{A}(\tilde{\mathcal{L}}_X(A), B) + \mathcal{A}(A, \tilde{\mathcal{L}}_X(B)),$$

where $\tilde{\mathcal{L}}_X$ is defined as in (18.8) from Definition 18.19. (*Remark*: This Problem is used in Lecture 18 to show that $\tilde{\mathcal{L}}_X = \mathcal{L}_X$.)

Problem J.5. Let M be a smooth manifold.

(i) Suppose $A \in \mathcal{T}^{1,1}(M) \cong \Gamma(\mathrm{End}(TM))$. Prove there exists a unique tensor derivation $\mathcal{D}^A$ on M with the property that $\mathcal{D}^A(Y)(x) = A_x(Y(x))$ for any vector field Y and satisfies $\mathcal{D}^A(f) = 0$ for any function f.

(ii) Let $\mathcal{D}$ be an arbitrary tensor derivation. Prove that there exists a vector field X on M and $A \in \mathcal{T}^{1,1}(M)$ such that $\mathcal{D} = \mathcal{L}_X + \mathcal{D}^A$.

Solutions to Problem Sheet J

(♣) Problem J.1. Let M be a smooth manifold of dimension n.

(i) Let $A \in \mathcal{T}^{r,s}(M)$ denote a tensor of type (r,s). Let $\sigma \colon U \to O$ and $\tau \colon V \to \Omega$ denote two charts on M with $U \cap V \neq \emptyset$. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$A = f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \qquad \text{on } U$$

and

$$A = g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial y^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial y^{i_r}} \otimes dy^{j_1} \otimes \cdots \otimes dy^{j_s}, \qquad \text{on } V,$$

for smooth functions $f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(U)$ and $g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(V)$. Investigate the relationship between

$$f^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V} \qquad \text{and} \qquad g^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V}.$$

(ii) Let $\omega \in \Omega^r(M)$ denote a differential r-form. Let $\sigma \colon U \to O$ and $\tau \colon V \to \Omega$ denote two charts on M with $U \cap V \neq \emptyset$. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\omega = f_{i_1 \cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}, \qquad \text{on } U$$

and

$$\omega = g_{i_1 \cdots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}, \qquad \text{on } V,$$

for smooth functions $f_{i_1 \cdots i_r} \in C^\infty(U)$ and $g_{i_1 \cdots i_r} \in C^\infty(V)$. Investigate the relationship between

$$f_{i_1 \cdots i_r}|_{U \cap V} \qquad \text{and} \qquad g_{i_1 \cdots i_r}|_{U \cap V}.$$

(iii) Suppose $\varphi \colon M \to N$ is a diffeomorphism. Let $A \in \mathcal{T}^{r,s}(N)$. Let $\sigma \colon U \to O$ denote a chart on M and $\tau \colon V := \varphi(U) \to \Omega$ denote a chart on N. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\varphi^\star(A) = f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \qquad \text{on } U$$

and

$$A = g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial y^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial y^{i_r}} \otimes dy^{j_1} \otimes \cdots \otimes dy^{j_s}, \qquad \text{on } V,$$

for smooth functions $f^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(U)$ and $g^{i_1 \cdots i_r}_{j_1 \cdots j_s} \in C^\infty(V)$. Investigate the relationship between

$$f^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V} \qquad \text{and} \qquad g^{i_1 \cdots i_r}_{j_1 \cdots j_s}|_{U \cap V}.$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

(iv) Suppose $\varphi\colon M \to N$ is a smooth map. Let $\omega \in \Omega^r(N)$. Let $\sigma\colon U \to O$ denote a chart on M and $\tau\colon V := \varphi(U) \to \Omega$ denote a chart on N. Let x^i denote the local coordinates of σ and y^i denote the local coordinates of τ. Then one can write

$$\varphi^\star(\omega) = f_{i_1\cdots i_r} dx^{i_1} \wedge \cdots \wedge dx^{i_r}, \qquad \text{on } U$$

and

$$\omega = g_{i_1\cdots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}, \qquad \text{on } V,$$

for smooth functions $f_{i_1\cdots i_r} \in C^\infty(U)$ and $g_{i_1\cdots i_r} \in C^\infty(V)$. Investigate the relationship between

$$f_{i_1\cdots i_r}|_{U\cap V} \qquad \text{and} \qquad g_{i_1\cdots i_r}|_{U\cap V}.$$

(v) Conclude that local coordinates are horrible.

SOLUTION. We start with (i). First of all observe that each $\frac{\partial}{\partial x^i} \in T(U\cap V)$ can be written in terms of the local frame field $\frac{\partial}{\partial y^1}, \ldots \frac{\partial}{\partial y^n}$, i.e.

$$\frac{\partial}{\partial x^i} = dy^k\left(\frac{\partial}{\partial x^i}\right)\frac{\partial}{\partial y^k} = \frac{\partial y^k}{\partial x^i}\frac{\partial}{\partial y^k},$$

and similarly

$$dx^i = dx^i\left(\frac{\partial}{\partial y^k}\right)dy^k = \frac{\partial x^i}{\partial y^k}\,dy^k.$$

Thus on $U\cap V$ we get

$$\begin{aligned} A &= f^{i_1\cdots i_r}_{j_1\cdots j_s}\frac{\partial}{\partial x^{i_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{i_r}}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_s} \\ &= f^{i_1\cdots i_r}_{j_1\cdots j_s}\left(\frac{\partial y^{k_1}}{\partial x^{i_1}}\frac{\partial}{\partial y^{k_1}}\right)\otimes\cdots\otimes\left(\frac{\partial y^{k_r}}{\partial x^{i_r}}\frac{\partial}{\partial y^{k_r}}\right)\otimes\left(\frac{\partial x^{j_1}}{\partial y^{l_1}}dy^{l_1}\right)\otimes\cdots\otimes\left(\frac{\partial x^{j_s}}{\partial y^{l_s}}dy^{l_s}\right) \\ &= f^{i_1\cdots i_r}_{j_1\cdots j_s}\left(\frac{\partial y^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial y^{k_r}}{\partial x_{i_r}}\right)\cdot\left(\frac{\partial x^{j_1}}{\partial y^{l_1}}\cdots\frac{\partial x^{j_s}}{\partial y^{l_s}}\right)\frac{\partial}{\partial y^{k_1}}\otimes\cdots\otimes\frac{\partial}{\partial y^{k_r}}\otimes dy^{l_1}\otimes\cdots\otimes dy^{l_s}. \end{aligned}$$

This then proves

$$g^{k_1\cdots k_r}_{l_1\cdots l_s} = f^{i_1\cdots i_r}_{j_1\cdots j_s}\left(\frac{\partial y^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial y^{k_r}}{\partial x^{i_r}}\right)\cdot\left(\frac{\partial x^{j_1}}{\partial y^{l_1}}\cdots\frac{\partial x^{j_s}}{\partial y^{l_s}}\right), \qquad \text{on } U\cap V.$$

The relation for (ii) follows in a similar fashion, namely

$$\begin{aligned} \omega &= f_{i_1\cdots i_r} dx^{i_1}\wedge\cdots\wedge dx^{i_r} \\ &= f_{i_1\cdots i_r}\frac{\partial x^{i_1}}{\partial y^{j_1}}\cdots\frac{\partial x^{i_r}}{\partial y^{j_r}}\,dy^{j_1}\wedge\cdots\wedge dy^{j_r}, \end{aligned}$$

and hence

$$g_{j_1\cdots j_r} = f_{i_1\cdots i_r}\frac{\partial x^{i_1}}{\partial y^{j_1}}\cdots\frac{\partial x^{i_r}}{\partial y^{j_r}}, \qquad \text{on } U\cap V.$$

For (iii) it will be convenient to view A and its pullback $\varphi^* A$ as multilinear maps at every footpoint (see Proposition 15.9). We first observe that $f^{i_1\cdots i_r}_{j_1\cdots j_s}$ is given by

$$\begin{aligned} f^{i_1\cdots i_r}_{j_1\cdots j_s} &= \varphi^*(A)\left(dx^{i_1},\cdots,dx^{i_r},\frac{\partial}{\partial x^{j_1}},\cdots,\frac{\partial}{\partial x^{j_s}}\right) \\ &= A_\varphi\left(dx^{i_1}\circ D\varphi^{-1},\cdots,dx^{i_r}\circ D\varphi^{-1},D\varphi\left[\frac{\partial}{\partial x^{j_1}}\right],\cdots,D\varphi\left[\frac{\partial}{\partial x^{j_s}}\right]\right), \end{aligned}$$

where the second equality is just definition. Now let us denote by $(D\varphi^{-1})^i_j(y)$ the (i,j)-th entry of the matrix representative $D\varphi^{-1}(y)\colon T_yN \to T_{\varphi^{-1}(y)}M$ with respect to the two bases $\frac{\partial}{\partial y^k}\big|_y$ and $\frac{\partial}{\partial x^l}\big|_{\varphi^{-1}(y)}$. Adopting the analogue notation for $D\varphi^i_j(x)$ and omitting footpoints again we get

$$f^{i_1\cdots i_r}_{j_1\cdots j_s} = A_\varphi\left((D\varphi^{-1})^{i_1}_{k_1}\,dy^{k_1},\cdots,(D\varphi^{-1})^{i_r}_{k_r}\,dy^{k_r},D\varphi^{l_1}_{j_1}\frac{\partial}{\partial y^{l_1}},\cdots,D\varphi^{l_s}_{j_s}\frac{\partial}{\partial y^{l_s}}\right).$$

Now by multilinearity of A we conclude

$$f^{i_1\cdots i_r}_{j_1\cdots j_s} = (D\varphi^{-1})^{i_1}_{k_1}\cdots(D\varphi^{-1})^{i_r}_{k_r}\cdot D\varphi^{l_1}_{j_1}\cdots D\varphi^{l_s}_{j_s}\,g^{k_1,\cdots k_r}_{l_1\cdots l_s}.$$

Part (iv) follows from a very similar argument to the one in (iii) [1]

$$\begin{aligned} f_{i_1\cdots i_r} &= \varphi^*\omega\left(\frac{\partial}{\partial x^{i_1}},\cdots\frac{\partial}{\partial x^{i_r}}\right) \\ &= \omega\left(D\varphi\left[\frac{\partial}{\partial x_{i_1}}\right],\cdots D\varphi\left[\frac{\partial}{\partial x_{i_r}}\right]\right) \\ &= D\varphi^{j_1}_{i_1}\cdots D\varphi^{j_r}_{i_r}\,g_{j_1,\cdots g_r}. \end{aligned}$$

Problem J.2. Let $\varphi\colon M\to N$ denote a smooth map. Let $A\in\mathcal{T}^{0,s}(N)$. Using the Tensor Criterion (Theorem 18.3), regard A as a $C^\infty(N)$-multilinear function

$$\underbrace{\mathfrak{X}(N)\times\cdots\times\mathfrak{X}(N)}_{s}\to C^\infty(N).$$

and similarly regard $\varphi^\star(A)$ as a $C^\infty(M)$-multilinear function

$$\underbrace{\mathfrak{X}(M)\times\cdots\times\mathfrak{X}(M)}_{s}\to C^\infty(M).$$

Suppose $X_i\in\mathfrak{X}(M)$ is φ-related to $Y_i\in\mathfrak{X}(N)$ for $i=1,\dots,s$. Prove that

$$\varphi^\star(A)(X_1,\dots,X_s) = A(Y_1,\dots,Y_s)\circ\varphi$$

as functions $M\to N$.

[1] Actually both (ii) and (iv) are just special cases of (i) and (iii), respectively.

SOLUTION. We compute: For a point $x \in M$ we set $y = \varphi(x)$, so that

$$\begin{aligned}\varphi^\star(A)(X_1,\dots,X_s)(x) &= A_{\varphi(x)}(D\varphi(x)X_1(x),\dots,D\varphi(x)X_s(x))\\ &= A_y(D\varphi(\varphi^{-1}(y))X_1(\varphi^{-1}(y)),\dots,D\varphi(\varphi^{-1}(y))X_s(\varphi^{-1}(y)))\\ &= A_y(Y_1(y),\dots,Y_s(y))\\ &= A(Y_1,\dots,Y_s)\circ\varphi(x)\end{aligned}$$

This solves the problem.

PROBLEM J.3. Let V be a vector space and suppose $\omega \in \bigwedge^r(V^*)$ and $\vartheta \in \bigwedge^s(V^*)$. Let $v_i \in V$ for $i = 1,\dots,r+s$. Identify ω with an element of $\mathrm{Alt}_r(V)$, ϑ with an element of $\mathrm{Alt}_s(V)$ and $\omega\wedge\vartheta$ with an element of $\mathrm{Alt}_{r+s}(V)$ (using Proposition 15.23). Prove that:

$$(\omega\wedge\vartheta)(v_1,\dots,v_{r+s}) = \frac{1}{r!s!}\sum_{\varrho\in\mathfrak{S}_{r+s}} \mathrm{sgn}(\varrho)\omega\big(v_{\varrho(1)},\dots,v_{\varrho(r)}\big)\,\vartheta\big(v_{\varrho(r+s)},\dots v_{\varrho(r+s)}\big) \tag{J.1}$$

or equivalently

$$(\omega\wedge\vartheta)(v_1,\dots,v_{r+s}) = \sum_{\varrho\in\mathrm{Shuffle}(r,s)} \mathrm{sgn}(\varrho)\omega\big(v_{\varrho(1)},\dots,v_{\varrho(r)}\big)\,\vartheta\big(v_{\varrho(r+s)},\dots v_{\varrho(r+s)}\big), \tag{J.2}$$

where $\mathrm{Shuffle}(r,s)$ was defined in Definition 19.3.

SOLUTION. We first observe that given any permutation $\pi \in \mathfrak{S}_{r+s}$, there exists a unique permutation σ of $\mathfrak{S}_r$ and a unique permutation τ on the letters $r+1,\dots,r+s$ such that $\varrho := \pi\circ\sigma\circ\tau$ is an (r,s)-shuffle. Let us write $\mathfrak{S}'_s$ for the permutations on the letters $r+1,\dots,r+s$. We now verify (J.2). Both sides are linear in ω and ϑ, and hence we may assume that both ω and ϑ are decomposable, say:

$$\omega = p^1\wedge\cdots\wedge p^r, \qquad \vartheta = q^1\wedge\cdots\wedge q^s.$$

Then if $v_1,\dots,v_{r+s}\in V$, we have (using the Leibniz formula for the determinant) that

$$\begin{aligned}&\big(p^1\wedge\cdots\wedge p^r\wedge q^1\wedge\cdots\wedge q^s\big)\big(v_1,\dots,v_{r+s}\big)\\ &\qquad= \sum_{\pi\in\mathfrak{S}_{r+s}} \mathrm{sgn}(\pi)p^1\big(v_{\pi(1)}\big)\cdots p^r\big(v_{\pi(r)}\big)\,q^1\big(v_{\pi(r+1)}\big)\cdots q^s\big(v_{\pi(r+s)}\big)\\ &\qquad= \sum_{\varrho\in\mathrm{Shuffle}(r,s)}\sum_{\sigma\in\mathfrak{S}_r}\sum_{\tau\in\mathfrak{S}'_s} \mathrm{sgn}(\varrho)\,\mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)\\ &\qquad\quad p^1\big(v_{\varrho(\sigma(1))}\big)\cdots p^r\big(v_{\varrho(\sigma(r))}\big)\,q^1\big(v_{\varrho(\tau(r+1))}\big)\cdots q^s\big(v_{\varrho(\tau(r+s))}\big)\end{aligned} \tag{J.3}$$

Set $w_i := v_{\varrho(i)}$. Then $w_{\sigma(i)} = v_{\varrho(\sigma(i))}$ and hence

$$\begin{aligned}\sum_{\sigma\in\mathfrak{S}_r} p^1\big(v_{\varrho(\sigma(1))}\big)\cdots p^r\big(v_{\varrho(\sigma(r))}\big) &= \sum_{\sigma\in\mathfrak{S}_r}\mathrm{sgn}(\sigma)p^1\big(w_{\sigma(1)}\big)\cdots p^r\big(w_{\sigma(r)}\big)\\ &= \big(p^1\wedge\cdots\wedge p^r\big)\big(w_1,\dots,w_r\big)\\ &= \big(p^1\wedge\cdots\wedge p^r\big)\big(v_{\varrho(1)},\dots,v_{\varrho(r)}\big)\end{aligned}$$

It follows that the sum in (J.3) is

$$\sum_{\varrho\in\text{Shuffle}(r,s)} \operatorname{sgn}(\rho)\,\big(p^1\wedge\cdots\wedge p^r\big)\big(v_{\varrho(1)},\dots,v_{\varrho(r)}\big)\,\big(q^1\wedge\cdots\wedge q^s\big)\big(v_{\varrho(r+1)},\dots,v_{\varrho(r+s)}\big)$$
$$= \sum_{\varrho\in\text{Shuffle}(r,s)} \operatorname{sgn}(\rho)\,\omega\big(v_{\varrho(1)},\dots,v_{\varrho(r)}\big)\,\vartheta\big(v_{\varrho(r+1)},\dots,v_{\varrho(r+s)}\big)$$

This proves (J.2). Finally, (J.1) is a formal consequence of (J.2).

PROBLEM J.4. Let (r,s), (r',s') and (r'',s'') be three pairs of non-negative integers. Suppose we are given a $\mathcal{C}^\infty_M$-bilinear sheaf homomorphism

$$\mathcal{A}\colon \mathcal{T}^{r,s}_M\times\mathcal{T}^{r',s'}_M\to\mathcal{T}^{r'',s''}_M.$$

Assume in addition that $\mathcal{A}$ has the property that if $\varphi\colon U\to V$ is a local diffeomorphism between open sets of M then

$$\varphi^\star(\mathcal{A}_V(A,B)) = \mathcal{A}_U(\varphi^\star(A),\varphi^\star(B)).$$

Prove that for every vector field X on M, one has

$$\tilde{\mathcal{L}}_X(\mathcal{A}(A,B)) = \mathcal{A}(\tilde{\mathcal{L}}_X(A),B)+\mathcal{A}(A,\tilde{\mathcal{L}}_X(B)),$$

where $\tilde{\mathcal{L}}_X$ is defined as in (18.8) from Lecture 18.

SOLUTION. Denote by $\varphi_t\colon M\to M$ the flow generated by the vector field X. For notational reasons we will omit the basepoints in the following computation:

$$\begin{aligned}
\tilde{\mathcal{L}}_X(\mathcal{A}(A,B)) &= \lim_{t\to0}\frac{\varphi_t^*\mathcal{A}(A,B)-\mathcal{A}(A,B)}{t}\\
&= \lim_{t\to0}\frac{\varphi_t^*\mathcal{A}_V(A,B)-\mathcal{A}_U(A,B)}{t}\\
&\overset{(1)}{=} \lim_{t\to0}\frac{\mathcal{A}_U(\varphi_t^*A,\varphi_t^*B)-\mathcal{A}_U(A,B)}{t}\\
&\overset{(2)}{=} \lim_{t\to0}\frac{\mathcal{A}(\varphi_t^*A-A,\varphi_t^*B)+\mathcal{A}(A,\varphi_t^*B)-\mathcal{A}(A,B)}{t}\\
&\overset{(3)}{=} \lim_{t\to0}\mathcal{A}\left(\frac{\varphi_t^*A-A}{t},\varphi_t^*B\right)+\mathcal{A}\left(A,\frac{\varphi_t^*B-B}{t}\right)\\
&\overset{(4)}{=} \mathcal{A}(\tilde{\mathcal{L}}_X(A),B)+\mathcal{A}(A,\tilde{\mathcal{L}}_X(B)).
\end{aligned}$$

Equality (1) is justified by our assumptions, (2) and (3) are just bilinearity and (4) uses the $\mathcal{C}^\infty_M$-bilinear assumption on $\mathcal{A}$, or more precisely the fact that $\mathcal{A}$ is smooth in both entries so that we can take the limit inside.

PROBLEM J.5. Let M be a smooth manifold.

(i) Suppose $A\in\mathcal{T}^{1,1}(M)\cong\Gamma(\operatorname{End}(TM))$. Prove there exists a unique tensor derivation $\mathcal{D}^A$ on M with the property that $\mathcal{D}^A(Y)(x)=A_x(Y(x))$ for any vector field Y and satisfies $\mathcal{D}^A(f)=0$ for any function f.

(ii) Let $\mathcal{D}$ be an arbitrary tensor derivation. Prove that there exists a vector field X on M and $A \in \mathcal{T}^{1,1}(M)$ such that $\mathcal{D} = \mathcal{L}_X + \mathcal{D}^A$.

SOLUTION.

(i) Let us suppose first that $\mathcal{D}^A$ exists and let us see how it must behave on functions, that is to say how $\mathcal{D}^A(f)$ for a fixed $f \in C^\infty(M)$. On the one hand, since A is a tensor, its C^∞-multilinearity implies that for every $X \in \mathfrak{X}(M)$ it must satisfy

$$\mathcal{D}^A(fX) = A(fX) = fA(X).$$

On the other hand, since $\mathcal{D}^A$ is a derivation it must also hold that

$$\mathcal{D}^A(fX) = \mathcal{D}^A(f)X + f\mathcal{D}^A(X) = \mathcal{D}^A(f)X + fA(X).$$

For both condition to hold it is then necessary that

$$\mathcal{D}^A(f)X = 0 \quad \text{for every } X \in \mathfrak{X}(M),$$

and consequently that $\mathcal{D}^A(f) = 0$. In other words, $\mathcal{D}^A$ has to vanish identically on functions. To prove that $\mathcal{D}^A$ exists we now set, by definition, for every open set $U \subseteq M$,

$$\begin{aligned} \mathcal{D}^A_U(f) = 0 \quad &\text{for every } f \in C^\infty(U), \\ \mathcal{D}^A_U(X) = A(X) \quad &\text{for every } X \in \mathfrak{X}(U), \end{aligned}$$

and note that the class of maps $\mathcal{D}^A_U$ defines a sheaf morphism since tensors are point operators and the derivation rule is satisfied. The existence of $\mathcal{D}^A$ is then given by Proposition 18.17.

(ii) If $\mathcal{D}$ is any derivation, it is in particular a derivation on functions and consequently, by Proposition 7.7, there exists some vector field $Z \in \mathfrak{X}(M)$ so that

$$\mathcal{D}(f) = \mathcal{L}_Z(f) = X(f) \quad \text{for every } f \in C^\infty(M).$$

In particular, it is sufficient to check that the difference $\mathcal{D}' = \mathcal{D} - \mathcal{L}_Z$ defines a $(1,1)$-tensor and then we can conclude tanks to point (i) above. By Theorem 18.3, we only need to check that it is C^∞-bilinear, and indeed for every $f \in C^\infty(M)$ and every $X \in \mathfrak{X}(M)$ we have:

$$\begin{aligned} \mathcal{D}'(fX) &= \mathcal{D}(fX) - \mathcal{L}_Z(fX) \\ &= \mathcal{D}(f)X + f\mathcal{D}(X) - \mathcal{L}_Z(f)X - f\mathcal{L}_Z(X) \\ &= f(\mathcal{D}(X) - \mathcal{L}_Z(X)) \\ &= f\mathcal{D}'(X), \end{aligned}$$

and for every $\omega \in \Omega^1(M)$ we have

$$\mathcal{D}'(X)(f\omega) = (f\omega)(\mathcal{D}'(X)) = f(\omega\mathcal{D}'(X)),$$

and this yields the required multilinearity of $\mathcal{D}'$.

Problem Sheet K

Problem K.1.

(i) Prove that S^n is orientable.

(ii) Prove that any Lie group is orientable.

(iii) Prove that $\mathbb{R}P^n$ is orientable if and only if n is odd. *Hint:* Consider the antipodal map $x \mapsto -x$ on S^n.

Problem K.2. Let

$$\mathbb{R}^n_- := \mathbb{R}^n_{u^1 \le 0}, \qquad \mathbb{H}^n := \mathbb{R}^n_{u^n \ge 0}.$$

We can identify both $\partial\mathbb{R}^n_-$ and $\partial\mathbb{H}^n$ with $\mathbb{R}^{n-1}$. Endow both $\mathbb{R}^n_-$ and $\mathbb{H}^n$ with their standard orientation they inherit from $\mathbb{R}^n$. Show that the induced orientation on $\partial\mathbb{R}^n_-$ is equal to standard orientation on $\mathbb{R}^{n-1}$ for all n, but that the induced orientation on $\partial\mathbb{H}^n$ agrees with the standard orientation of $\mathbb{R}^{n-1}$ only when n is even. *Remark:* This is the main reason we take our "standard" half-space to be $\mathbb{R}^n_-$, not $\mathbb{H}^n$, cf. Remark 21.4.

(♣) Problem K.3.

(i) Let V be a vector space of dimension r. A **symplectic form** on V is an element $\omega \in \mathrm{Alt}_2(V) \cong \bigwedge^2(V^*)$ which is **non-degenerate** in the sense that $i_v(\omega) \equiv 0$ if and only if $v = 0$. Prove that if a symplectic form exists then $r = 2n$ is necessarily an even number.

(ii) A **symplectic manifold** is a smooth manifold M equipped with a *closed* differential 2-form ω such that ω_x is a symplectic form on T_xM for every $x \in M$. Prove that any symplectic manifold is orientable.

(iii) Let M be a smooth manifold. Define a 1-form $\lambda \in \Omega^1(T^*M)$ on the cotangent bundle via the formula:

$$\lambda_{x,p}(\zeta) = p\big(D\pi(x,p)[\zeta]\big), \qquad x \in M, p \in T^*_xM, \zeta \in T_{(x,p)}(T^*M),$$

where $\pi\colon T^*M \to M$ is the projection[1]. Prove that $\omega := d\lambda$ is a symplectic form on T^*M. Thus every cotangent bundle is a symplectic manifold.

(♣) Problem K.4. Let M and N be smooth manifolds. Prove that if M has boundary and N does not, them $M \times N$ is a smooth manifold with boundary. Prove that if both M and N have non-empty boundary then $M \times N$ is not a smooth manifold with boundary,

(♣) Problem K.5. After making appropriate modifications, reprove all results in the course for manifolds with boundary.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1]This makes sense, as $D\pi(x,p)$ is a linear map $T_{(x,p)}(T^*M) \to T_xM$, and thus p can eat $D\pi(x,p)[\zeta]$ to produce a real number.

Solutions to Problem Sheet K

Problem K.1.

(i) Prove that S^n is orientable.

(ii) Prove that any Lie group is orientable.

(iii) Prove that $\mathbb{R}P^n$ is orientable if and only if n is odd. *Hint:* Consider the antipodal map $x \mapsto -x$ on S^n.

Solution.

(i) We will use a concrete model for S^n:

$$S^n := \{x = (x^0, \dots, x^n) \in \mathbb{R}^{n+1} \mid |x| = 1\}.$$

On $\mathbb{R}^{n+1}$ we have the canonical differential $(n+1)$-form $\omega_{\mathbb{R}^{n+1}} := dx^0 \wedge \dots \wedge dx^n$, which is clearly a volume form. On $\mathbb{R}^{n+1} \backslash \{0\}$ we have a canonical nowhere vanishing radial vector field $X \in \mathfrak{X}(\mathbb{R}^{n+1})$ which assigns to every point $x \neq 0$ the vector $X(x) = x \in \mathbb{R}^{n+1} = T_x\mathbb{R}^{n+1}$. The interior product $\imath_X \omega_{\mathbb{R}^{n+1}}$ is a differential n-form $\mathbb{R}^{n+1} \backslash \{0\}$, so in particular the restriction $\omega_{S^n} := (\imath_X \omega_{\mathbb{R}^{n+1}})|_{S^n}$ is a differential n-form on S^n.[1]

We claim that ω_{S^n} is a volume form on S^n. To see this, recall that the differential of the inclusion $S^n \hookrightarrow \mathbb{R}^{n+1}$ identifies the tangent space of S^n at $x \in S^n$ with the orthogonal complement $x^\perp \subset \mathbb{R}^{n+1} = T_x\mathbb{R}^{n+1}$. In particular, if $v_1, \dots, v_n$ denotes a basis for T_xS^n, then $x, v_1, \dots, v_n$ is a basis for $\mathbb{R}^{n+1} = T_x\mathbb{R}^{n+1}$. Hence,

$$\omega_{S^n}(v_1, \dots v_n) = \omega_{\mathbb{R}^{n+1}}(x, v_1, \dots, v_n) \neq 0.$$

Since this is true for every $x \in S^n$ we see that ω_{S^n} is indeed a volume form on S^n and thus S^n is orientable. Note that the above proof can be adapted to show the following more general statement: The boundary of an orientable manifold is orientable. In the above case, S^n is the boundary of the ball $\{x \mid |x| < 1\}$, which is clearly orientable, simply because it is an open subset of the orientable $\mathbb{R}^{n+1}$.

(ii) Let G denote an n-dimensional Lie group. The Lie algebra $\mathfrak{g} = T_eG$ is an n-dimensional real vector space and hence admits a volume form of the type $\omega_e := e_1^* \wedge \dots \wedge e_n^*$, where $e_1, \dots, e_n$ is some basis for $\mathfrak{g}$ with dual basis $e_1^*, \dots, e_n^*$. At some point $g \in G$ we can define

$$\omega_g := (Dl_{g^{-1}}(g))^* \omega_e$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

[1] Recall that taking the restriction means pulling back by the inclusion map, so $(\imath_X \omega_{\mathbb{R}^{n+1}})|_{S^n} = i^*(\imath_X \omega_{\mathbb{R}^{n+1}})$, where $i \colon S^n \hookrightarrow \mathbb{R}^{n+1}$ denotes the inclusion.

which is the (linear) pull back of ω_e by the linear map $Dl_{g^{-1}}(g)\colon T_gG \to T_eG$. The map $g \mapsto \omega_g$ is clearly a smooth differential n-form which vanishes nowhere, because ω_e doesn't vanish and $Dl_{g^{-1}}(g)$ is a linear isomorphism. This shows that G is orientable.

(iii) Denote by $q\colon S^n \to \mathbb{R}P^n = S^n/(x \sim -x)$ the quotient map and by $a\colon S^n \to S^n$ the antipodal map

$$a(x) = -x.$$

$\mathbb{R}P^n$ is orientable if and only if there exists a nowhere vanishing differential n-form $\omega_{\mathbb{R}P^n}$ on $\mathbb{R}P^n$. Since the differential of q is an isomorphism (q is a covering map) at every point it follows that $\omega_{\mathbb{R}P^n}$ vanishes nowhere if and only if $q^*\omega_{\mathbb{R}P^n}$ vanishes nowhere. Since $q \circ a = q$ we also have

$$a^*q^*\omega_{\mathbb{R}P^n} = (q \circ a)^*\omega_{\mathbb{R}P^n} = q^*\omega_{\mathbb{R}P^n} \tag{K.1}$$

Putting these things together we conclude that, if $\mathbb{R}P^n$ is orientable, then $q^*\omega_{\mathbb{R}P^n}$ is a volume form on S^n, so it is a generator for the 1-dimensional real vector space $H^n_{\mathrm{dR}}(S^n)$, and (K.1) says that the induced linear map $a^*\colon H^n_{\mathrm{dR}}(S^n) \to H^n_{\mathrm{dR}}(S^n)$ is the identity map (i.e. multiplication by 1). Hence, to see that $\mathbb{R}P^n$ is *not* orientable when n is even it suffices to check that a^* is *not* the identity in this case. In (i) we saw that

$$x \mapsto \omega_{S^n}|_x := \imath_x(dx^0 \wedge \cdots \wedge dx^n)|_{S^n}$$

defines a volume form on the sphere. Denote by $A\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ the antipodal map $A(x) = -x$, so that if $i\colon S^n \to \mathbb{R}^{n+1}$ denotes the inclusion then $i \circ a = A \circ i$ as a map $S^n \to \mathbb{R}^{n+1}$. Then

$$\begin{aligned}
(a^*\omega_{S^n})|_x &= a^*(\omega_{S^n}|_{-x}) \\
&= a^*i^*(\imath_{-x}[dx^0 \wedge \cdots \wedge dx^n]) \\
&= (i \circ a)^*(\imath_{-x}[dx^0 \wedge \cdots \wedge dx^n]) \\
&= (A \circ i)^*(\imath_{-x}[dx^0 \wedge \cdots \wedge dx^n]) \\
&= i^*A^*(\imath_{-x}[dx^0 \wedge \cdots \wedge dx^n]) \\
&= i^*(\imath_x[-dx^0 \wedge \cdots \wedge -dx^n]) \\
&= (-1)^{n+1}i^*(\imath_x[dx^0 \wedge \cdots \wedge dx_n]) \\
&= (-1)^{n+1}\omega_{S^n}|_x
\end{aligned}$$

Hence, we see that if n is even (so $(n+1)$ is odd) then $a^*\colon H^n_{\mathrm{dR}}(S^n) \to H^n_{\mathrm{dR}}(S^n)$ is multiplication by -1 instead of 1, which concludes the proof that $\mathbb{R}P^n$ is *not* orientable when n is even. To see that $\mathbb{R}P^n$ is orientable when n is odd, note that the above computation shows that in this case ω_{S^n} is an a^*-invariant volume form (in the sense that $a^*\omega_{S^n} = \omega_{S^n}$). This implies that ω_{S^n} descends to a volume form on $\mathbb{R}P^n$ when n is odd, finishing the proof that $\mathbb{R}P^n$ is orientable in this case.

Problem K.2. Let

$$\mathbb{R}^n_- := \mathbb{R}^n_{u^1 \leq 0}, \qquad \mathbb{H}^n := \mathbb{R}^n_{u^n \geq 0}.$$

We can identify both $\partial\mathbb{R}^n_-$ and $\partial\mathbb{H}^n$ with $\mathbb{R}^{n-1}$. Endow both $\mathbb{R}^n_-$ and $\mathbb{H}^n$ with their standard orientation they inherit from $\mathbb{R}^n$. Show that the induced orientation on $\partial\mathbb{R}^n_-$ is equal to standard orientation on $\mathbb{R}^{n-1}$ for all n, but that the induced orientation on $\partial\mathbb{H}^n$ agrees with the standard orientation of $\mathbb{R}^{n-1}$ only when n is even. *Remark:* This is the main reason we take our "standard" half-space to be $\mathbb{R}^n_-$, not $\mathbb{H}^n$, cf. Remark 21.4.

SOLUTION. Both on $\mathbb{R}^n_-$ and $\mathbb{H}^n$ the orientation is the one induced standard volume form $\mu = dx^1 \wedge \cdots \wedge dx^n$.

To determine the orientation induced on $\partial\mathbb{R}^n_-$, it is sufficient (see the exercise at the end of Definition 21.21) to compute $i_X(\mu)$ for an arbitrarily chosen outward-pointing section $X(x) = X^i(x)\partial_{x^i}|_x$ of $T\mathbb{R}^n_-|_{\partial\mathbb{R}^n_-}$. By Definition 21.18, X is outward-pointing if and only if

$$dx^1(X)|_x = X^1(x) > 0 \quad \text{for every } x \in \partial\mathbb{R}^n_-,$$

consequently the obvious choice is the constant section $X = \partial_1|_x$, for which using the formula given in Definition 20.1 we see that

$$i_X(\mu) = dx^2 \wedge \cdots \wedge dx^n,$$

and this volume form on $\partial\mathbb{R}^n_-$ is exactly the standard one on $\mathbb{R}^{n-1} \simeq \{0\} \times \mathbb{R}^{n-1}$.

Similarly, on $\mathbb{H}^n$ a section $Y(x) = Y^i(x)\partial_{x^i}|_x$ of $T\mathbb{H}^n|_{\partial\mathbb{H}^n}$ is outward-pointing if and only if

$$dx^n(Y)|_x = Y^n(x) < 0 \quad \text{for every } x \in \partial\mathbb{H}^n,$$

so choosing $Y = -\partial_{x^n}|_x$ we compute

$$i_Y(\mu) = (-1)^{n+1} dx^1 \wedge \cdots \wedge dx^{n-1},$$

and so this volume form induces the same orientation on $\partial\mathbb{H}^n$ as the standard one in $\mathbb{R}^{n-1} \simeq \mathbb{R}^{n-1} \times \{0\}$ if and only if n is odd.

(♣) PROBLEM K.3.

(i) Let V be a vector space of dimension r. A **symplectic form** on V is an element $\omega \in \mathrm{Alt}_2(V) \cong \bigwedge^2(V^*)$ which is **non-degenerate** in the sense that $i_v(\omega) \equiv 0$ if and only if $v = 0$. Prove that if a symplectic form exists then $r = 2n$ is necessarily an even number.

(ii) A **symplectic manifold** is a smooth manifold M equipped with a *closed* differential 2-form ω such that ω_x is a symplectic form on T_xM for every $x \in M$. Prove that any symplectic manifold is orientable.

(iii) Let M be a smooth manifold. Define a 1-form $\lambda \in \Omega^1(T^*M)$ on the cotangent bundle via the formula:

$$\lambda_{x,p}(\zeta) = p\big(D\pi(x,p)[\zeta]\big), \qquad x \in M, p \in T^*_xM, \zeta \in T_{(x,p)}(T^*M),$$

where $\pi\colon T^*M \to M$ is the projection[2]. Prove that $\omega := d\lambda$ is a symplectic form on T^*M. Thus every cotangent bundle is a symplectic manifold.

[2]This makes sense, as $D\pi(x,p)$ is a linear map $T_{(x,p)}(T^*M) \to T_xM$, and thus p can eat $D\pi(x,p)[\zeta])$ to produce a real number.

Solution. We will prove part (i) by proving the following stronger statement[3]. There exists a basis $e_1, \ldots, e_r, f_1, \ldots, f_r$ of V satisfying

$$\forall i, j = 1, \ldots, r\colon\ \omega(e_i, e_j) = \omega(f_i, f_j) = 0, \text{ and } \omega(e_i, f_j) = \delta_{ij}.$$

We will refer to such a basis as a *symplectic basis* of (V, ω). If $\dim V < 1$ then there is nothing to prove, so we assume $\dim V \geq 1$. Pick a non-zero vector $e_1 \in V$. Non-degeneracy of ω and $e_1 \neq 0$ imply the existence of a non-zero $f_1 \in V$ such that

$$\omega(e_1, f_1) \neq 0.$$

By bilinearity of ω we can assume $\omega(e_1, f_1) = 1$ after rescaling f_1. Since ω is alternating and bilinear we know that f_1 cannot lie in the span of e_1 which in turn implies that the dimension of V is *strictly* greater than 1. If $\dim V = 2$ then we are done, otherwise we proceed as follows. For any subspace $W \subset V$ there exists the so called *symplectic complement*

$$W^\omega = \{v \in V \mid \omega(v, w) = 0 \text{ for all } w \in W\},$$

which is itself a subspace of V giving rise to the splitting

$$V = W \oplus W^\omega.$$

Since $\dim V > 2$ we get for $W_1 := \operatorname{span}(e_1, f_1)$ a non-trivial symplectic complement

$$W_1^\omega \neq 0.$$

Again, pick a non-zero vector e_2, this time lying in W_1^ω, and another non-zero vector $f_2 \in V$ such that $\omega(e_2, f_2) = 1$. By definition of the symplectic complement, $e_2 \in W_1$ and $V = W_1 \oplus W_1^\omega$ it follows that $f_2 \in W_1^\omega \setminus \operatorname{span}(e_2)$ and hence

$$\dim \operatorname{span}(e_2, f_2) = 2,$$

which again implies

$$\dim V \geq 4.$$

Proceeding inductively (with $W_i = \operatorname{span}(e_1, \ldots, e_i, f_1, \ldots, f_i)$ and so on) and observing that the procedure terminates at some point ($\dim V < +\infty$!) finishes the proof.

We continue with part (ii). In Lecture 20 we have seen that a manifold M is orientable if and only if there exists a volume form (see 20.23), i.e. a everywhere non-vanishing top form

$$\mu \in \Omega^{2r}(M),$$

where $\dim M = 2r$ is even by part (i) since ω_x is a symplectic form on T_xM and $\dim M = \dim T_xM$. We claim that the r-wedge

$$\omega^r := \underbrace{\omega \wedge \cdots \wedge \omega}_{r-\text{times}} \in \Omega^{2r}(M)$$

[3]This makes the proof of (ii) and (iii) much more pleasant.

defines a volume form. Indeed, by part (i) we know that for each $x \in M$ there exists a symplectic basis $(e_{1,x}, \cdots, e_{r,x}, f_{1,x}, \cdots f_{r,x})$ for (T_xM, ω_x). By evaluating ω_x on this basis it is easy to see that

$$\omega_x = \sum_{i=1}^{r} e^*_{i,x} \wedge f^*_{i,x},$$

where the $*$'s indicate the dual basis of T^*_xM. A straightforward combinatorial argument shows

$$\omega^r_x = r!\, e^*_{1,x} \wedge f^*_{1,x} \wedge \cdots \wedge e^*_{r,x} \wedge f^*_{r,x},$$

thus proving $\omega^r_x \neq 0$ since

$$\omega^r_x\,(e_{1,x}, f_{1,x}, \ldots, e_{r,x}, f_{r,x}) = r! \neq 0.$$

Finally we move to part (iii). The quickest way to see why $d\lambda$ defines a symplectic form on T^*M is by computing it in local coordinates and invoking part (i). Let $(x^1, \cdots x^n)$ be the local coordinates on M associated to a chart $\sigma\colon U \to \mathbb{R}^n$ and define, for every $i = 1, \ldots, n$,

$$y^i\colon T^*U \to \mathbb{R},\ y^i(x,p) := p\left(\frac{\partial}{\partial x^i}\Big|_x\right).$$

We can view $x^i\colon U \to \mathbb{R}$ also as smooth functions on T^*U by setting $x^i(x,p) = x^i(x)$ which then leads us to local coordinates $(x^1, \ldots, x^n, y^1, \ldots, y^n)$ on T^*M (see Problem C.1 for more details). Let us show that on T^*U one has

$$\lambda = \sum_{i=1}^{n} y^i\, dx^i.$$

To prove this it suffices to compute λ on the local frame field associated to our local coordinates, i.e.

$$\begin{aligned}
\lambda_{(x,p)}\left(\frac{\partial}{\partial x^i}\Big|_{(x,p)}\right) &= p\left(D\pi(x,p)\left[\frac{\partial}{\partial x^i}\Big|_{(x,p)}\right]\right)\\
&\overset{(1)}{=} p\left(\frac{\partial}{\partial x^i}\Big|_x\right)\\
&= y^i(x,p),
\end{aligned}$$

where in (1) we used the fact that $\sigma \circ \pi \circ \sigma^{-1} = \mathrm{id}$ (this is clear as $\pi(x,p) = x$ and $\sigma(x,p) = \sigma(x)$). On the other hand we have

$$\lambda_{(x,p)}\left(\frac{\partial}{\partial y^i}\Big|_{(x,p)}\right) = 0,$$

simply because the differential $D\pi$ maps all the $\frac{\partial}{\partial y^i}$'s to 0. All in all this proves the claim.

Now differentiating $\lambda = \sum_{i=1}^n y^i\, dx^i$ and using the Leibniz rule grants

$$d\lambda = \sum_{i=1}^{n} dy^i \wedge dx^i.$$

This 2-form is certainly closed as it is exact (remember: $d^2 = 0$) and non-degeneracy follows immediately by either a direct check or comparing it with the proof of part (ii).

(♣) PROBLEM K.4. Let M and N be smooth manifolds. Prove that if M has boundary and N does not, them $M \times N$ is a smooth manifold with boundary. Prove that if both M and N have non-empty boundary then $M \times N$ is not a smooth manifold with boundary,

SOLUTION. Suppose M has boundary and N does not. Let $m := \dim M$ and $n := \dim N$. We claim that $M \times N$ is a smooth manifold with boundary given by $\partial M \times N$. We may always choose an atlas $\Sigma = \{\sigma_a \colon U_a \to Q_a \mid a \in A\}$ on M such that Q_a is an open subset of $\mathbb{R}^m_+$. Let $\mathcal{T} = \{\tau_b \colon V_b \to O_b \mid b \in B\}$ be an atlas on N. Then

$$\{\sigma_a \times \tau_b \colon U_a \times V_b \to Q_a \times O_b \mid (a,b) \in A \times B\}$$

is an atlas on $M \times N$ making it into a smooth manifold with boundary (since $Q_a \times O_b \subset \mathbb{R}^m_+ \times \mathbb{R}^n = \mathbb{R}^{m+n}_+$).

Showing that $\partial(M \times N) = \partial M \times N$ is equivalent to showing that

$$\operatorname{int}(M \times N) = \operatorname{int}(M) \times N. \tag{K.2}$$

Let $(p,q) \in M \times N$ and consider charts $\sigma \colon U \to \mathbb{R}^m_+$ and $\tau \colon V \to \mathbb{R}^n$ around p and q respectively. Then $\sigma \times \tau \colon U \times V \to \mathbb{R}^{m+n}_+$ is a chart around (p,q) and $(\sigma \times \tau)(p,q) \in \operatorname{int}(\mathbb{R}^{m+n}_+)$ if and only if $\sigma(p) \in \operatorname{int}(\mathbb{R}^m_+)$. It follows that $(p,q) \in \operatorname{int}(M \times N)$ if and only if $p \in \operatorname{int}(M)$. Thus (K.2) holds.

Now let M and N be smooth manifolds with non-empty boundary. Note that the boundary of a 0-dimensional manifold is empty, so our assumptions imply that $m = \dim M$ and $n = \dim N$ are both ≥ 1. Suppose first that $m = n = 1$. It suffices to show that $\mathbb{R}^2_+$ and $\mathbb{R}_+ \times \mathbb{R}_+$ are not diffeomorphic. Suppose there exists a diffeomorphism $\phi \colon \mathbb{R}_+ \times \mathbb{R}_+ \xrightarrow{\sim} \mathbb{R}^2_+$ and let $x_0 := (0,0)$. Since a diffeomorphism is also a homeomorphism, it follows that ϕ must send $\partial(\mathbb{R}_+ \times \mathbb{R}_+)$ to $\partial\mathbb{R}^2_+$, where we mean boundaries in the sense of topological manifolds with boundary. By assumption, the map ϕ extends on an open neighbourhood $U \subset \mathbb{R}^2$ of x_0 to a diffeomorphism $\tilde{\phi} \colon U \xrightarrow{\sim} V$, where V is open in $\mathbb{R}^2$. Consider the curves

$$\gamma_1 \colon (-\epsilon_1, \epsilon_1) \to U, \quad t \mapsto (0,t)$$

and

$$\gamma_2 \colon (-\epsilon_2, \epsilon_2) \to U, \quad t \mapsto (t,0).$$

Then the $\gamma_i'(0) \in T_{x_0}U$ are linearly independent for $i = 1,2$. Since $\tilde{\phi}$ is a diffeomorphism, the vectors $D\tilde{\phi}(x_0)\big(\gamma_i'(0)\big)$ for $i = 1,2$ must also be linearly independent. But for negative t, the curves $\tilde{\phi} \circ \gamma_i$ lie on the y-axis of $\mathbb{R}^2$, from which it follows that the tangent vectors $(\tilde{\phi} \circ \gamma_i)'(0)$ for $i = 1,2$ are linearly dependent, a contradiction.
Note. *One can generalise this argument to higher dimensions. In that case, one must instead take $n+m$ curves passing through a point $(p,q) \in \partial\mathbb{R}^m_+ \times \partial\mathbb{R}^n_+$ such that the tangent vectors are linearly independent, but whose images under the differential of a diffeomorphism are contained in an $(n+m-1)$-dimensional subspace. We present a different argument of the same flavor below.*

Now suppose that $m+n > 2$. Let $(p,q) \in \partial M \times \partial N$ and consider charts $\sigma\colon U \to Q \subset \mathbb{R}^m_+$ and $\tau\colon V \to O \subset \mathbb{R}^n$ around p and q respectively. As a consequence of the fact that the set of interior points and the set of boundary points on a manifold are disjoint, it follows that the image of (p,q) under $\sigma \times \tau$ is contained in $\partial\mathbb{R}^m_+ \times \partial\mathbb{R}^n_+$. In order to obtain a contradiction, suppose that $M \times N$ is a smooth manifold with boundary. Then there exists a smooth chart $\rho\colon U' \to Q' \subset \mathbb{R}^{m+n}_+$ around (p,q). Without loss of generality, we may assume that $U' = U \times V$. We thus obtain a diffeomorphism $\rho \circ (\sigma \times \tau)^{-1}\colon Q \times O \xrightarrow{\sim} Q'$. Let $x = (x^i)_i := (\sigma \times \tau)^{-1}(p,q) \in \mathbb{R}^m \times \mathbb{R}^n$. By our assumptions, we have $x^1 = x^{m+1} = 0$. Let $S \subset \mathbb{R}^{m+n}$ be the $(n+m-2)$-dimensional subspace consisting of points whose first and $(m+1)$-st coordinates are zero. Then $\tilde{U} := (Q \times O) \cap S$ is open in S and $x \in \tilde{U}$. Consider the composite

$$\alpha\colon \tilde{U} \hookleftarrow Q \times O \to Q' \hookrightarrow \mathbb{R}^{m+n},$$

where the middle map is $\rho \circ (\sigma \times \tau)^{-1}$. By construction, the map α is smooth and the composite $(\sigma \times \tau) \circ \rho^{-1} \circ \alpha$ is the identity on $\tilde{U}$, so $D\alpha(x)$ is injective. Let $T := D\alpha(x)(T_x S) \subset \mathbb{R}^{m+n}$. Because T is $(n+m-2)$ dimensional, it must contain a vector v such that one of the first three components, v^1, v^2 or v^3 is non-zero. Renumbering coordinates and replacing v by $-v$ if necessary, we may assume the $v^1 < 0$. Let $\gamma\colon (-\epsilon, \epsilon) \to S$ be a smooth curve such that $\gamma(0) = x$ and $D\alpha(\gamma'(0)) = v$. Then $\alpha \circ \gamma(t)$ has negative x^1 coordinate for all small $t > 0$, which contradicts the fact that α takes values in $\mathbb{R}^{m+n}_+$.

(♣) PROBLEM K.5. After making appropriate modifications, reprove all results in the course for manifolds with boundary.

SOLUTION. Nope.

Problem Sheet L

Problem L.1. A singular k-cube $c\colon C^k \to M$ is said to be **degenerate** if there exists $1 \le i \le k$ such that c does not depend on x^i. Prove that if $c\colon C^k \to M$ is a degenerate singular k-cube then $\int_c \omega = 0$ for any $\omega \in \Omega^k(M)$.

Problem L.2. Let $c\colon C^k \to M$ be a smooth singular k-cube in M and let $\varphi\colon C^k \to C^k$ be an orientation preserving diffeomorphism[1]. Let $\tilde{c} := c \circ \varphi$. Prove that for any $\omega \in \Omega^k(M)$, one has

$$\int_c \omega = \int_{\tilde{c}} \omega.$$

Problem L.3. Prove that there does not exist a compact symplectic manifold (M, ω) (without boundary) with the property that ω is exact. (See Problem K.3 if you forgot the definition of a symplectic manifold.)

Problem L.4. Find a closed $(n-1)$-form on $\mathbb{R}^n \setminus \{0\}$ that is not exact.

Problem L.5. Let M be a smooth manifold, let $X \in \mathfrak{X}(M)$, and let A be a tensor field. Let θ_t denote the flow of X. Prove that

$$\frac{d}{dt}\Big|_{t=t_0} \theta_t^\star(A) = \theta_{t_0}^\star(\mathcal{L}_X(A)).$$

Problem L.6. Let $\varphi\colon M^n \to N^n$ be a diffeomorphism of connected oriented manifolds and let $\omega \in \Omega_c^n(N)$. Prove that

$$\int_M \varphi^\star(\omega) = \pm \int_N \omega,$$

where the $+$ signs occurs if and only if φ is orientation preserving (cf. Definition 20.21).

Problem L.7. Let G be a compact connected Lie group.

(i) G is orientable by part (ii) of Problem K.1. Prove there exists a unique **normalised left-invariant** volume form μ on G, i.e. a volume form μ such that $\int_G \mu = 1$ and $l_a^\star(\mu) = \mu$ for all $a \in G$.

(ii) This allows us to define the integral of a *function* on G via:

$$\int_G f := \int_G f\,\mu, \qquad f \in C^\infty(G).$$

Prove that

$$\int_G f = \int_G (f \circ l_a) = \int_G (f \circ r_a), \qquad \forall\, f \in C^\infty(G),\ a \in G.$$

Hint: Use Problem L.6.

[1] As usual, think of this as meaning that φ is the restriction to C^k of an orientation preserving diffeomorphism of some neighbourhood.

(♣) Problem L.8. In this problem you may assume that for any compact connected orientable smooth manifold M^n, one has $H^n_{\mathrm{dR}}(M) \cong \mathbb{R}$, and that an explicit isomorphism is given by

$$\int : H^n_{\mathrm{dR}}(M) \to \mathbb{R}, \qquad [\omega] \mapsto \int_M \omega.$$

(This will be justified in Lecture 27. Let $\varphi\colon M \to N$ be a smooth map between compact connected orientable smooth manifolds of dimension n. Then $\varphi^\star\colon H^n_{\mathrm{dR}}(N) \to H^n_{\mathrm{dR}}(M)$ is a linear map between one-dimension vector spaces, and hence is multiplication by a number. We call this number the **degree** of φ. Explicitly,

$$\int_M \varphi^\star(\omega) = \deg(\varphi) \int_N \omega, \qquad \omega \in \Omega^n(N).$$

The purpose of this question is to investigate how to compute $\deg(\varphi)$.

(i) Let $y \in N$ denote a regular value of φ. Given $x \in \varphi^{-1}(y)$, let

$$\operatorname{sgn}_x(f) := \begin{cases} +1, & \text{if } D\varphi(x) \text{ is orientation preserving,} \\ -1, & \text{if } D\varphi(x) \text{ is not orientation preserving.} \end{cases}$$

Prove that

$$\deg(\varphi) = \sum_{x \in \varphi^{-1}(y)} \operatorname{sgn}_x(f).$$

Thus $\deg(\varphi)$ is an integer. *Hint:* Use Problem L.6 again.

(ii) Prove the **Hairy Ball Theorem**: if n is even then any vector field on S^n has at least one zero. *Hint:* Recall from part (iii) of Problem K.1 that the antipodal map $x \mapsto -x$ is orientation reversing if n is even.

Solutions to Problem Sheet L

PROBLEM L.1. A singular k-cube $c\colon C^k \to M$ is said to be **degenerate** if there exists $1 \le i \le k$ such that c does not depend on x^i. Prove that if $c\colon C^k \to M$ is a degenerate singular k-cube then $\int_c \omega = 0$ for any $\omega \in \Omega^k(M)$.

SOLUTION. Let $c\colon C^k \to M$ be a degenerate singular k-cube and let $\omega \in \Omega^k(M)$. Recall that we can write c^*w as $hdx^1 \wedge \cdots \wedge dx^k$, where $h \in C^\infty(C^k)$ is the function

$$h = c^*(\omega)\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\Big).$$

Fix $p \in \mathrm{Int(C^k)}$. By definition,

$$c^*(\omega)_p\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\Big) = \omega_{c(p)}\Big(Dc(p)\Big[\frac{\partial}{\partial x^1}\Big], \dots, Dc(p)\Big[\frac{\partial}{\partial x^k}\Big]\Big) \tag{L.1}$$

Let $1 \le i \le k$ be such that c does not depend on x^i. Consider the map

$$j^i_p\colon C^1 \to C^k, \quad t \mapsto (p^1, \dots, p^{i-1}, t, p^{i+1}, \dots, p^k).$$

Since $c \circ j^i_p$ is constant, it follows that

$$Dc(p)\Big[\frac{\partial}{\partial x^i}\Big] = D(c \circ j^i_p)(p^i)\Big[\frac{\partial}{\partial t}\Big] = 0.$$

As $\omega_{c(p)}$ is alternating, we conclude that (L.1) is zero. Since this holds for all $p \in \mathrm{Int(C^k)}$, this implies that $\int_c \omega = 0$.

PROBLEM L.2. Let $c\colon C^k \to M$ be a smooth singular k-cube in M and let $\varphi\colon C^k \to C^k$ be an orientation preserving diffeomorphism[1]. Let $\tilde c := c \circ \varphi$. Prove that for any $\omega \in \Omega^k(M)$, one has

$$\int_c \omega = \int_{\tilde c} \omega.$$

SOLUTION. We write $c^*\omega = hdx^1 \wedge \cdots \wedge dx^k$ for $h \in C^\infty(C^k)$. We have

$$\begin{aligned}
\int_c \omega := \int_{C^k} c^*\omega \quad &= \quad \int_{C^k} h \\
&\overset{(*)}{=} \quad \int_{C^k} (h \circ \phi)|\det D\phi| \\
&\overset{(**)}{=} \quad \int_{C^k} (h \circ \phi)(\det D\phi) \\
&\overset{(***)}{=} \quad \int_{C^k} \phi^* c^* \omega \\
&= \quad \int_{\tilde c} \omega,
\end{aligned}$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick.
Last modified: June 28, 2019.

[1] As usual, think of this as meaning that φ is the restriction to C^k of an orientation preserving diffeomorphism of some neighbourhood.

where $D\phi$ is the Jacobian matrix of ϕ. Here $(*)$ is just the classical change of variable formula from multivariable calculus, and $(**)$ follows from the assumption that ϕ is orientation preserving. Finally, equality $(***)$ comes from pullback formula

$$\phi^*(hdx^1 \wedge \cdots \wedge dx^n) = (h \circ \phi)(\det D\phi)dx^1 \wedge \cdots \wedge dx^n.$$

Problem L.3. Prove that there does not exist a compact symplectic manifold (M, ω) (without boundary) with the property that ω is exact. (See Problem K.3 if you forgot the definition of a symplectic manifold.)

Solution. Suppose for contradiction that there exists a closed exact symplectic manifold

$$(M^{2n}, \omega = d\lambda).$$

By the second part of problem K.3, or rather its proof, we do know that ω^n defines a volume form on M. Without loss of generality we therefore assume

$$\omega^n > 0,$$

which readily implies

$$\int_M \omega^n > 0.$$

On the other hand exactness of ω implies exactness of ω^n as one can see by the computation

$$\omega^n = (d\lambda)^n = \overbrace{d\lambda \wedge \cdots \wedge d\lambda}^{n} = d(\lambda \cdot \overbrace{d\lambda \wedge \cdots \wedge d\lambda}^{n-1}).$$

The last inequality is a consequence of $d^2 = 0$ and the Leibniz-rule. Thus, for $\alpha := \lambda \cdot (d\lambda \wedge \cdots \wedge d\lambda)$ we have

$$\omega^n = d\alpha.$$

We finally obtain our desired contradiction via Stokes' Theore) and the fact that M has empty boundary, i.e. $\partial M = \emptyset$, as it is a *closed* manifold by assumption. We have

$$0 < \int_M \omega^n = \int_M d\alpha = \int_{\partial M} \alpha = 0,$$

which is a contradiction and concludes the proof.

Problem L.4. Find a closed $(n-1)$-form on $\mathbb{R}^n \setminus \{0\}$ that is not exact.

Solution. Our strategy consists in pulling back a volume form on S^{n-1} to $\mathbb{R}^n \setminus \{0\}$: Pick a volume form ω on S^{n-1} inducing the standard orientation and define the smooth function

$$r \colon \mathbb{R}^n \setminus \{0\} \to S^{n-1},\ r(x) := \frac{x}{\|x\|}.$$

Observe that r is a left inverse of the inclusion $i \colon S^{n-1} \hookrightarrow \mathbb{R}^n \setminus \{0\}$. Now the claim is that the pullback

$$r^*\omega \in \Omega^{n-1}(\mathbb{R}^n \setminus \{0\})$$

defines a closed non-exact $(n-1)$-form. The closed bit is straightforward by dimension reasons of the sphere, more precisely

$$dr^*\omega = r^* \underbrace{d\omega}_{=0} = 0.$$

To see why $r^*\omega$ is non-exact we argue by contradiction. Assume for contradiction that there exists $\lambda \in \Omega^{n-2}\left(\mathbb{R}^n \setminus \{0\}\right)$ such that

$$d\lambda = r^*\omega.$$

From this we deduce that ω is also exact as

$$\omega = (r \circ i)^*\omega = i^*\left(r^*\omega\right) = i^* d\lambda = di^*\lambda.$$

But this already leads to a contradiction using the fact that ω is a volume form and Stokes Theorem

$$0 < \int_{S^{n-1}} \omega = \int_{S^{n-1}} di^*\lambda = \int_{\partial S^{n-1}} i^*\lambda = 0,$$

hence finishes the proof.

Problem L.5. Let M be a smooth manifold, let $X \in \mathfrak{X}(M)$, and let A be a tensor field. Let θ_t denote the flow of X. Prove that

$$\left.\frac{d}{dt}\right|_{t=t_0} \theta_t^\star(A) = \theta_{t_0}^\star(\mathcal{L}_X(A)).$$

Solution. Changing variable in the limit of the difference quotient, we can write

$$\left.\frac{d}{dt}\right|_{t=t_0} \theta_t^\star(A) = \left.\frac{d}{dt}\right|_{t=0} \theta_{t_0+t}^\star(A).$$

By the properties of the flow, we then have $\theta_{t_0+t} = \theta_{t_0} \circ \theta_t$, and consequently by the properties of the pull-back, this implies $\theta_{t_0+t}^\star(A) = \theta_{t_0}^\star(\theta_t^\star(A))$. Moreover, since the pull-back is a linear operation, it commutes with derivatives and thus we conclude that

$$\left.\frac{d}{dt}\right|_{t=0} \theta_{t_0+t}^\star(A) = \left.\frac{d}{dt}\right|_{t=0} \theta_{t_0}^\star(\theta_t^\star(A)) = \theta_{t_0}^\star\left(\left.\frac{d}{dt}\right|_{t=0} \theta_t^\star(A)\right) = \theta_{t_0}^\star(\mathcal{L}_X(A)),$$

as desired.

Problem L.6. Let $\varphi \colon M^n \to N^n$ be a diffeomorphism of connected oriented manifolds and let $\omega \in \Omega_c^n(N)$. Prove that

$$\int_M \varphi^\star(\omega) = \pm \int_N \omega,$$

where the $+$ signs occurs if and only if φ is orientation preserving (cf. Definition 20.21).

Solution. Assume first that ω and $\varphi^\star(\omega)$ are compactly supported in the domains of single charts whose coordinates we denote $y = (y^1, \dots, y^n)$ and $x = (x^1, \dots, x^n)$ respectively. We may then write $\omega(y) = f(y)dx^1 \wedge \cdots \wedge dx^n$ for some smooth function f, and consequently (by the properties of the exerior derivative and of the wedge product)

$$\varphi^\star(\omega)(x) = f(\varphi(x))d\varphi^1(x) \wedge \cdots \wedge \varphi^n(x) = f(\varphi(x)) \det(D\varphi(x))dx^1 \wedge \cdots \wedge dx^n,$$

and so, with the standard change of variable formula for the integrals and the rule on the derivative of the inverse function, setting $y = \varphi(x)$, we see that

$$\begin{aligned}
\int \varphi^\star(\omega) &= \int f(\varphi(x)) \det(D\varphi(x))dx^1 \cdots dx^n \\
&= \int f(y) \det D\varphi(\varphi^{-1}(y)) \, |\det D(\varphi^{-1})(y)| dy^1 \cdots dy^n \\
&= \int f(y) \frac{|\det D(\varphi^{-1})(y)|}{\det D(\varphi^{-1})(y)} dy^1 \cdots dy^n \\
&= \pm \int f(y) dy^1 \cdots dy^n \\
&= \pm \int \omega,
\end{aligned}$$

where the plus sign occurs if φ (and so φ^{-1}) is orientation-preserving and the minus sign if if it is orientation-reversing. This proves the result in this special single-chart case.

For the general case, we choose an positively oriented atlas $(U_i, \sigma_i)_{i \in I}$ for M so that, for each U_i, $\varphi(U_i)$ is contained in a single coordinate chart of of N (this is possible because φ is smooth and in particular continuous), and $(\psi_i)_{i \in I}$ a correspondent partition of unity. Since φ is a diffeomorphism, $(\varphi(U_i), \sigma_i \circ \varphi^{-1})_{i \in I}$ will then constitute a (positively or negatively, depending on φ) oriented atlas for N, and $(\psi_i \circ \varphi^{-1})_{i \in I}$ a corresponding partition of unity. We can then compute, thanks to the special case above:

$$\int_N \varphi^\star(\omega) = \sum_i \int_{\varphi(U_i)} (\psi_i \circ \varphi^{-1}) \, \varphi^\star(\omega) = \sum_i \pm \int_{U_i} \psi_i \, \omega = \pm \int_M \omega,$$

where again, the plus sign occurs if φ is orientation-preserving and the minus sign if if it is orientation-reversing.

Problem L.7. Let G be a compact connected Lie group.

(i) G is orientable by part (ii) of Problem K.1. Prove there exists a unique **normalised left-invariant** volume form μ on G, i.e. a volume form μ such that $\int_G \mu = 1$ and $l_a^\star(\mu) = \mu$ for all $a \in G$.

(ii) This allows us to define the integral of a *function* on G via:

$$\int_G f \coloneqq \int_G f \, \mu, \qquad f \in C^\infty(G).$$

Prove that

$$\int_G f = \int_G (f \circ l_a) = \int_G (f \circ r_a), \qquad \forall\, f \in C^\infty(G),\ a \in G.$$

Hint: Use Problem L.6.

Solution. Ad (i): Observe, that two volume forms on G which are left-invariant and coincide at the identity $e \in G$, coincide everywhere. This is because, if μ is such a form then $\mu|_{T_gG} \equiv Dl_{g^{-1}}(g)^\star \mu|_{T_eG}$ for all $g \in G$, so $\mu|_{T_gG}$ is uniquely determined by $\mu|_{T_eG}$. Hence, it suffices to show that two left-invariant volume forms on G coincide on T_eG up to scaling, but this follows from basic linear algebra: We know that

$$\bigwedge^{\dim G} T_e^* G$$

is 1-dimensional.

Ad (ii) By definition we have

$$\begin{aligned}\int_G (f \circ l_a) &= \int_G (f \circ l_a)\mu \\ &= \int_G (f \circ l_a) l_a^* \mu \\ &= \int_G l_a^*(f\mu) \\ &= \int_G f\mu \\ &= \int_G f,\end{aligned}$$

where at the second to last equality we make use of L.6 and the fact that $l_a \colon G \to G$ is orientation preserving (since $l_a^*\mu = \mu$). The equality $\int f = \int (f \circ r_a)$ follows from the same computation if only $r_a^*\mu = \mu$. But this is the case because

$$r_a^*\mu = \mu \Longleftrightarrow \mu = l_{a^{-1}}^* r_a^* \mu = l_{a^{-1}}^* \mu$$

and the later is clearly the case because of the assumption that μ is left-invariant.

(♣) Problem L.8. In this problem you may assume that for any compact connected orientable smooth manifold M^n, one has $H^n_{\mathrm{dR}}(M) \cong \mathbb{R}$, and that an explicit isomorphism is given by

$$\int : H^n_{\mathrm{dR}}(M) \to \mathbb{R}, \qquad [\omega] \mapsto \int_M \omega.$$

(This will be justified in Lecture 27. Let $\varphi \colon M \to N$ be a smooth map between compact connected orientable smooth manifolds of dimension n. Then $\varphi^\star \colon H^n_{\mathrm{dR}}(N) \to H^n_{\mathrm{dR}}(M)$ is a linear map between one-dimension vector spaces, and hence is multiplication by a number. We call this number the **degree** of φ. Explicitly,

$$\int_M \varphi^\star(\omega) = \deg(\varphi) \int_N \omega, \qquad \omega \in \Omega^n(N).$$

The purpose of this question is to investigate how to compute $\deg(\varphi)$.

(i) Let $y \in N$ denote a regular value of φ. Given $x \in \varphi^{-1}(y)$, let

$$\operatorname{sgn}_x(f) := \begin{cases} +1, & \text{if } D\varphi(x) \text{ is orientation preserving,} \\ -1, & \text{if } D\varphi(x) \text{ is not orientation preserving.} \end{cases}$$

Prove that

$$\deg(\varphi) = \sum_{x \in \varphi^{-1}(y)} \operatorname{sgn}_x(f).$$

Thus $\deg(\varphi)$ is an integer. *Hint:* Use Problem L.6 again.

(ii) Prove the **Hairy Ball Theorem**: if n is even then any vector field on S^n has at least one zero. *Hint:* Recall from part (iii) of Problem K.1 that the antipodal map $x \mapsto -x$ is orientation reversing if n is even.

Solution. Ad part (i): By Sard's Theorem (cf. Theorem 5.17) there exists a regular value $y \in N$ of $\varphi\colon M \to N$. The Implicit Function Theorem (cf. Theorem 5.13) then implies that $\varphi^{-1}(y) \subset M$ is a 0-dimensional manifold, i.e. a collection of isolated points in M. Since M is compact and $\varphi^{-1}(y)$ does not have any accumulation points, it follows that $\varphi^{-1}(y)$ is finite. We enumerate

$$\varphi^{-1}(y) = \{x_1, \dots, x_l\} \subset M.$$

By definition each x_i defines a regular point of φ which then readily implies that each differential

$$d\varphi(x_i)\colon T_{x_i}M \to T_y M$$

is a linear isomorphism (remember $\dim M = \dim N$!). Therefore the Inverse Function Theorem (cf. Theorem 5.2) grants the existence of pairwise disjoint charts

$$U_1, \dots, U_l \subset M,\ \phi_i\colon U_i \to \mathbb{R}^n$$

such that φ restricted to each one of these is a diffeomorphism onto its image, i.e.

$$\varphi\big|_{U_i} \xrightarrow{\sim} \varphi(U_i), \text{ for all } i = 1, \dots, l.$$

Now we define the open sets

$$V = \bigcap_{i=1}^{l} \varphi(U_i) \text{ and } U_i' :== U_i \cap \varphi^{-1}(V).$$

By shrinking U_i (and hence V) if necessary, we can assume that (V, ψ) defines a chart on N. Let

$$g \in C^\infty(N)$$

be a smooth function with support in V and define the top form

$$\omega := g\, dy^1 \wedge \cdots \wedge dy^n$$

on N, where the y^i's are the local coordinates coming from (V, ψ). Due to our choices the pullback form $\varphi^*\omega$ on M has support lying inside the union $\bigcup_i U_i'$ and

$$\varphi(U_i') = V,$$

hence

$$\int_M \varphi^*\omega = \sum_{i=1}^{l} \int_{U_i'} \varphi^*\omega \overset{(1)}{=} \sum_{i=1}^{l} \operatorname{sgn}_{x_i}(\varphi) \int_V \omega = \sum_{i=1}^{l} \operatorname{sgn}_{x_i}(\varphi) \int_N \omega,$$

where step (1) follows from Problem L.6. This proves

$$\deg(\varphi) = \sum_{x \in \varphi^{-1}(y)} \operatorname{sgn}_x(\varphi)$$

and finishes the first part.

For part (ii) we make the crucial observation that the degree is a homotopy invariant, i.e. two smoothly homotopic maps $\varphi, \psi\colon M \to N$ have the same degree

$$\deg(\varphi) = \deg(\psi).$$

This follows immediately from the definition of deg as the scalar of the induced map $\varphi^*\colon H^n_{\mathrm{dR}}(M) \cong \mathbb{R} \to H^n_{\mathrm{dR}}(N) \cong \mathbb{R}$ and Theorem 23.17 which asserts that two smoothly homotopic maps induce the same linear map on the de Rham cohomology.

We proceed by contradiction and assume that for n even there exists a vector field X on S^n with no zeros. The strategy is to build a homotopy between the antipodal map and the identity, which would lead to a contradiction as the former has degree -1 whereas $\deg(\mathrm{id}_{S^n}) = 1$ (see the hint above).

For every $x \in S^n$, there exists a *unique* semicircle $\gamma_x \approx S^1$ on S^n determined by the direction $X(x) \in T_x S^n \setminus \{0\}$. Each such semicircle can be viewed as a smooth loop

$$\gamma_x\colon [0,2] \to S^n$$

satisfying

$$\gamma_x(0) = \gamma_x(2) = x \text{ and } \gamma_x(-1) = -x.$$

With this we are able to define a smooth homotopy

$$H\colon [0,1] \times S^n \to S^n,\ H(t,x) := \gamma_x(t)$$

that connects the identity

$$H(0,x) = \gamma_x(0) = x$$

to the antipodal map

$$H(1,x) = \gamma_x(1) = -x.$$

The fact that H is smooth follows from γ_x being uniquely defined by $X(x)$, where the later is smooth by definition. This finishes the proof

Problem Sheet M

Problem M.1. Let $\pi\colon P \to N$ be a G-principal bundle, and let $\varphi\colon M \to N$ be a smooth map. Prove that the fibre bundle $\varphi^\star P \to M$ is also a G-principal bundle.

Problem M.2. Let $\pi_i\colon P_i \to M_i$ be two G-principal bundles. Suppose $\Phi\colon P_1 \to P_2$ is a principal bundle morphism along a diffeomorphism $\varphi\colon M_1 \to M_2$. Prove that Φ is also a diffeomorphism.

Problem M.3. Let M be a smooth manifold and suppose $\pi_i\colon P_i \to M$ are principal G-bundles over M. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both[1] P_1 and P_2 admit principal bundle atlases over the $U_{\mathtt{a}}$. Let

$$\rho^1_{\mathtt{ab}}\colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G, \qquad \text{and} \qquad \rho^2_{\mathtt{ab}}\colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G$$

denote the transition functions of P_1 and P_2 with respect to these bundle atlases. Prove that P_1 and P_2 are isomorphic principal bundles if and only if there exists a smooth family $\nu_{\mathtt{a}}\colon U_{\mathtt{a}} \to G$ of functions such that

$$\nu_{\mathtt{a}}(x) \cdot \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \cdot \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}.$$

Problem M.4. Suppose G is a Lie group acting transitively on a smooth manifold M, so that M is the homogeneous space G/H for an appropriate subgroup H of G. Prove that the subgroup of G acting trivially on M is the largest normal subgroup $N(H)$ of G contained in H. Let $\bar{G}$ and $\bar{H}$ denote the quotient groups $G/N(H)$ and $H/N(H)$ respectively. Prove that $\bar{G}$ acts effectively and transitively on M, and M is the homogeneous space $\bar{G}/\bar{H}$.

Problem M.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose G acts on a manifold P on the right. Prove that the map $v \mapsto \xi_v$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(P)$.

Problem M.6. Let M be a smooth manifold, and let W_1, W_2 and Z be vector spaces. Let $\omega \in \Omega^r(M, W_1)$ and let $\vartheta \in \Omega^s(M, W_2)$, and let $\beta\colon W_1 \times W_2 \to Z$ be a bilinear map. Prove that the exterior differential satisfies

$$d(\omega \wedge_\beta \vartheta) = d\omega \wedge_\beta \vartheta + (-1)^r \omega \wedge_\beta d\vartheta.$$

(♣) Problem M.7. Compute the derivative of the map Ψ from (24.2) (used in the proof of Proposition 24.8), and show that its derivative is invertible.

Will J. Merry, Diff. Geometry I, Autumn 2018, ETH Zürich. Last modified: June 28, 2019.

[1] This can always be achieved by taking intersections.

Solutions to Problem Sheet M

Problem M.1. Let $\pi\colon P \to N$ be a G-principal bundle, and let $\varphi\colon M \to N$ be a smooth map. Prove that the fibre bundle $\varphi^\star P \to M$ is also a G-principal bundle.

Solution. Recall that the pullback bundle $\mathrm{pr}\colon \varphi^\star P \to M$ is given by

$$\varphi^\star P \coloneqq \{(x,p) \in M \times P \mid \varphi(x) = \pi(p)\} \text{ and } \mathrm{pr}(x,p) \coloneqq x.$$

We define a right G-action on the total space $\varphi^\star P$ by setting

$$(x,p)\cdot a \coloneqq (x, p\cdot a),$$

where $(x,p) \in \varphi^\star P$, $a \in G$ and $p \cdot a$ denotes the G-action on P. Freeness of the G-action on $\varphi^\star P$ readily follows from the freeness of the G-action on P. Therefore we are only left to show the G-equivariance of the bundle charts on $\mathrm{pr}\colon \varphi^\star P \to M$.

Let

$$\alpha\colon \pi^{-1}(V) \to G$$

be a bundle chart on $\pi\colon P \to N$. The corresponding bundle chart on the pullback bundle is defined via

$$\alpha^\star\colon \varphi^{-1}(V) \to G,\ \alpha^\star(x,p) \coloneqq \alpha(p).$$

The following computation finishes the proof:

$$\begin{aligned}
\alpha^\star((x,p)\cdot a) &= \alpha^\star(x, p\cdot a) \\
&= \alpha(p\cdot a) \\
&= \alpha(p)a \\
&= \alpha^\star(x,p)a.
\end{aligned}$$

Problem M.2. Let $\pi_i\colon P_i \to M_i$ be two G-principal bundles. Suppose $\Phi\colon P_1 \to P_2$ is a principal bundle morphism along a diffeomorphism $\varphi\colon M_1 \to M_2$. Prove that Φ is also a diffeomorphism.

Solution. We begin by showing that Φ is surjective. Let $q \in P_2$ be any element lying over, say, $y \in M_2$. Now let $p \in P_1$ be some element in the fibre over $\varphi^{-1}(y)$. Then $\Phi(p) \in (P_2)_y$ and since the fibres are precisely the G-orbits (cf. Lemma 24.6) there exists some $a \in G$ such that

$$\Phi(p)\cdot a = q.$$

Using the G-equivariance of principle bundle morphisms we deduce

$$\Phi(p\cdot a) = q,$$

Solutions written by Mads Bisgaard, Francesco Palmurella, Alessio Pellegrini, and Alexandre Puttick. Last modified: June 28, 2019.

which proves that Φ is surjective.

For the injectivity we pick two distinct elements $p, q \in P_1$ and let us assume wlog that they lie in the same fibre - otherwise their images also lie in distinct fibres since φ is injective and then there is nothing to prove. Thus

$$\Phi(p), \Phi(q) \in (P_2)_{\varphi(x)}.$$

Since G acts freely and transitively on $(P_1)_x$ there exists a unique $a \in G \setminus \{e\}$ such that $p = q \cdot a$ and consequently

$$\Phi(p) = \Phi(q \cdot a) = \Phi(q) \cdot a.$$

But G also acts freely on P_2 and since $a \neq e$ we conclude

$$\Phi(q) \neq \Phi(p).$$

In order to conclude the proof we will show that at any point $p \in P$ the differential $D\Phi(p)$ is an isomorphism - we admit this for a second and see how one can finish the proof using this: By the Implicit Function Theorem it follows that Φ is a local diffeomorphism around any p and since Φ itself is already bijective one obtains that Φ is a (global) diffeomorphism. We go back to the proof of "$D\Phi(p)$ linear isomorphism": We fix any point $p \in P_1$ and pick a bundle chart

$$\alpha\colon \pi_1^{-1}(U) \to G,$$

where U is an open neighbourhood of $x := \pi_1(p)$. Denote $a = \alpha(p)$ and observe that

$$D\Phi(p) = D(\Phi \circ (\pi_1, \alpha)^{-1})(x, a) = \left(D\varphi(x), D\left(\Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1}\right)(a)\right)$$

can be seen as a linear map on $T_xM_1 \times T_aG \cong T_pP_1$. We already know that $D\varphi(x)\colon T_xM_1 \to T_{\varphi(x)}M_2$ is a linear isomorphism as $\varphi\colon M_1 \to M_2$ is a diffeomorphism by assumption, thus it suffices to show that

$$\Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1}\colon G \to (P_1)_x$$

is a diffeomorphism. Set

$$c := \Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1}(e) \in G$$

and define the unique G-equivariant smooth map

$$\Psi_c\colon (P_1)_x \to G \text{ with } \Psi_c(c) = e^{1}.$$

Since the bundle chart α is G-equivariant by definition, we have that the whole composition

$$\Psi_c \circ \Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1}\colon G \to G$$

is G-equivariant, satisfying

$$\Psi_c \circ \Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1}(e) = e.$$

Therefore

$$\Psi_c \circ \Phi\big|_{(P_1)_x} \circ \alpha\big|_{(P_1)_x}^{-1} = \mathrm{id}_G$$

by uniqueness of such a G-equivariant map, which then concludes the proof.

[1] Uniqueness of such a map follows from Lemma 24.6.

PROBLEM M.3. Let M be a smooth manifold and suppose $\pi_i \colon P_i \to M$ are principal G-bundles over M. Let $\{U_{\mathtt{a}} \mid \mathtt{a} \in \mathtt{A}\}$ be an open cover of M such that both[2] P_1 and P_2 admit principal bundle atlases over the $U_{\mathtt{a}}$. Let

$$\rho^1_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G, \qquad \text{and} \qquad \rho^2_{\mathtt{ab}} \colon U_{\mathtt{a}} \cap U_{\mathtt{b}} \to G$$

denote the transition functions of P_1 and P_2 with respect to these bundle atlases. Prove that P_1 and P_2 are isomorphic principal bundles if and only if there exists a smooth family $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to G$ of functions such that

$$\nu_{\mathtt{a}}(x) \cdot \rho^1_{\mathtt{ab}}(x) = \rho^2_{\mathtt{ab}}(x) \cdot \nu_{\mathtt{b}}(x), \qquad \forall\, x \in U_{\mathtt{a}} \cap U_{\mathtt{b}},\ \forall\, \mathtt{a}, \mathtt{b} \in \mathtt{A}. \tag{M.1}$$

SOLUTION. Suppose first that we are given a principal G-bundle isomorphism $\Phi \colon P_1 \xrightarrow{\sim} P_2$. For $i = 1, 2$ and each $\mathtt{a} \in \mathtt{A}$, we denote the corresponding bundle chart $\pi_i(U_{\mathtt{a}}) \to G$ by $\mathtt{a}^i$. Let $x \in U_{\mathtt{a}}$ and define

$$\nu_{\mathtt{a}}(x) \coloneqq \mathtt{a}^2 \circ \Phi \circ (\mathtt{a}^1|_{P_{1,x}})^{-1}.$$

A priori, we have $\nu_{\mathtt{a}}(x) \in \mathrm{Diff}(G)$, but we claim that $\nu_{\mathtt{a}}(x)$ is in fact left translation by an element of G. To see this, let $g \in G$ and let $p \in P_{2,x}$ be the unique element such that $\mathtt{a}^2(p) = g$. Define $h \coloneqq \mathtt{a}^1\big(\Phi^{-1}(p)\big)$. Fix $g_1 \in G$. Then we can write $g_1 = g g_2$ for $g_2 = g^{-1} g_1$. By G-equivariance, we have

$$\mathtt{a}^2(p \cdot g_2) = \mathtt{a}^2(p) g_2 = g g_2 = g_1.$$

Similarly,

$$\mathtt{a}^1\big((\Phi^{-1}(p) \cdot g_2\big) = \mathtt{a}^1\big(\Phi^{-1}(p)\big) g_2 = h g_2 = h g^{-1} g_1 = \ell_{hg^{-1}}(g_1).$$

It follows that $\nu_{\mathtt{a}}(x)(g_1) = \ell_{hg^{-1}}(g_1)$. Identifying $\nu_{\mathtt{a}}(x)$ with hg^{-1} and varying $x \in U_{\mathtt{a}}$, we thus obtain a map $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to G$. It remains to show that the $\nu_{\mathtt{a}}$ satisfy (M.1). Note that in (M.1), we could also write $\circ$ instead of $\cdot$ after identifying G with the subgroup of $\mathrm{Diff}(G)$ given by left-translations. Checking that (M.1) holds is done in precisely the same way as in the solution to Problem H.4.

Conversely, suppose we are given a smooth family of function $\nu_{\mathtt{a}} \colon U_{\mathtt{a}} \to G$ satisfying (M.1). For each $\mathtt{a} \in \mathtt{A}$, define $\tilde{\Phi}_{\mathtt{a}} \colon U_{\mathtt{a}} \times G \to U_{\mathtt{a}} \times G$ by $(x, g) \mapsto (x, \nu_{\mathtt{a}}(x) g)$. We define $\Phi \colon P_1 \to P_2$ as follows: for $p_1 \in P_1$, choose $\mathtt{a} \in \mathtt{A}$ such that $p_1 \in \pi_1^{-1}(U_{\mathtt{a}})$ and define

$$\Phi(p_1) \coloneqq \big((\pi_2, \mathtt{a}^2)^{-1} \circ \tilde{\Phi}_{\mathtt{a}} \circ (\pi_1, \mathtt{a}^1)\big)(p_1).$$

Checking that Φ is well-defined is done exactly in the same manner as in the solution to Problem H.4. This uses equation (M.1).

It remains to show that the smooth map $\Phi \colon P_1 \to P_2$ is an isomorphism of principal G-bundles. It is clear the Φ maps $P_{1,x}$ to $P_{2,x}$ for all $x \in M$, and by Problem M.2, it suffices to show that Φ as a principle G-bundle homomorphism, which follows if we show that Φ is G-equivariant. Let $p \in P_1$ and let $x \coloneqq \pi_1(p)$ and let $\mathtt{a} \in \mathtt{A}$ be such that $x \in U_{\mathtt{a}}$. We note that fact that the G-actions are fiber-preserving and the G-equivariance of the $\mathtt{a}^i$ for $i = 1, 2$ implies that

$$(\pi_i, \mathtt{a}^i) \colon \pi_i^{-1}(U_{\mathtt{a}}) \xrightarrow{\sim} U_{\mathtt{a}} \times G$$

[2]This can always be achieved by taking intersections.

and its inverse are also G-equivariant, where we define the G-action on the right-hand side via $(x, g_1) \cdot g := (x, g_1 g)$. Let $g \in G$. We compute:

$$\begin{aligned}
\Phi(pg) &= \big((\pi_2, \mathtt{a}^2)^{-1} \circ \tilde{\Phi}_{\mathtt{a}} \circ (\pi_1, \mathtt{a}^1)\big)(pg) \\
&= (\pi_2, \mathtt{a}^2)^{-1}\big(x, \nu_a(x)\mathtt{a}^1(pg)\big) \\
&= (\pi_2, \mathtt{a}^2)^{-1}\Big(\big(x, \nu_{\mathtt{a}}(x)\mathtt{a}^1(p)\big)g\Big) \\
&= (\pi_2, \mathtt{a}^2)^{-1}\Big(\big(x, \nu_{\mathtt{a}}(x)\mathtt{a}^1(p)\big)\Big)g \\
&= \Phi(p)g,
\end{aligned}$$

as desired.

Problem M.4. Suppose G is a Lie group acting transitively on a smooth manifold M, so that M is the homogeneous space G/H for an appropriate subgroup H of G. Prove that the subgroup of G acting trivially on M is the largest normal subgroup $N(H)$ of G contained in H. Let $\bar{G}$ and $\bar{H}$ denote the quotient groups $G/N(H)$ and $H/N(H)$ respectively. Prove that $\bar{G}$ acts effectively and transitively on M, and M is the homogeneous space $\bar{G}/\bar{H}$.

Solution. Denote by $N \subset G$ the subset of elements which act trivially on M. Suppose $g' \in N$ and fix any $g \in G$. Then for all $x \in M$ we have

$$\mu(g^{-1}g'g, x) = \mu(g^{-1}, g'gx) = \mu(g^{-1}, gx) = \mu(g^{-1}g, x) = x,$$

where $\mu\colon G \times M \to M$ denotes the right action of G on M. This shows that N is a normal subgroup. Clearly $N \subset H$, since $H = \{g \in G \mid \mu(g, x_0) = x_0\}$ for some fixed $x_0 \in M$. Suppose $N_0 \leq G$ is a normal subgroup satisfying $N_0 \subset H$. We need to show that this implies $N_0 \subset N$. Choose $g_0 \in N_0$ and $y \in M$. Then, since G acts transitively on M there is a $g \in G$ such that $gx_0 = y$. Since N_0 is normal in G we have $gg_0g^{-1} \in N_0 \subset H$, which implies $g^{-1}g_0gx_0 = x_0$, or $g_0y = y$. This shows that every element of N_0 fixes every element of M. Hence, $N_0 \subset N$ which proves the first claim. Since $N = N(H)$ is a normal subgroup of G it is well-known from algebra that $\bar{G} := G/N$ carries a canonical group structure. It is obvious that $\bar{G}$ acts on M in a canonical way. This action is said to be effective if $\bar{g}x = x$ for all $x \in M$ implies $\bar{g} = \mathrm{id}$. Suppose $\bar{g} \in \bar{G}$ satisfies $\bar{g}x = x$ for all $x \in M$. Then the representative $g \in G$ of the class $\bar{g} \in G$ fixes every element of M, so that $g \in N$. Hence, $\bar{g}$ is the multiplicative unit in $\bar{G}$, which proves that the action $\bar{G} \times M \to M$ is effective. $\bar{H}$ is the isotropy group of the (transitive) action, so $M = \bar{G}/\bar{H}$.

Problem M.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$, and suppose G acts on a manifold P on the right. Prove that the map $v \mapsto \xi_v$ is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(P)$.

Solution. Denote by

$$\mu\colon G \times P \to P,\ \mu(g, p) = p \cdot g$$

the smooth right action of G on P and set $\mu^p = \mu(\cdot, p)\colon G \to P$ for any $p \in P$. The map

$$\mathfrak{g} \to \mathfrak{X}(P),\ v \mapsto \xi_v$$

is defined (see (25.2)) by

$$\xi_v(p) := D(\mu^p)_e[v] \in T_{p\cdot e}P = T_pP.$$

By smoothness of μ and the fact that $p \cdot e = p$ it readily follows that ξ_v defines a vector field on P. In order to conclude that this is a Lie algebra homomorphism[3] we need to prove that

$$\xi_{[v,w]} = [\xi_v, \xi_w],$$

for all $v, w \in \mathfrak{g}$. Recall that the Lie bracket $[v, w]$ on $\mathfrak{g}$ is defined by

$$[v, w] := \mathrm{eval}_e[X_v, X_w],$$

where X_v and X_w are the unique left-invariant vector fields associated to v and w. The following computation proves the desired identity:

$$\begin{aligned}
[\xi_v, \xi_w](p) &= [D(\mu^p)(e)[v], D(\mu^p)(e)[w]](p) \\
&= [D(\mu^p)(e)[X_v(e)], D(\mu^p)(e)[X_w(e)]](p) \\
&\overset{(1)}{=} D(\mu^p)(e)\left([X_v, X_w](e)\right) \\
&= \xi_{[v,w]},
\end{aligned}$$

where in step (1) we used Problem (ii).

Problem M.6. Let M be a smooth manifold, and let W_1, W_2 and Z be vector spaces. Let $\omega \in \Omega^r(M, W_1)$ and let $\vartheta \in \Omega^s(M, W_2)$, and let $\beta\colon W_1 \times W_2 \to Z$ be a bilinear map. Prove that the exterior differential satisfies

$$d(\omega \wedge_\beta \vartheta) = d\omega \wedge_\beta \vartheta + (-1)^s \omega \wedge_\beta d\vartheta.$$

Solution. We quickly recall that for two bundle-valued forms ω and ϑ as above one defines their β-wedge product by

$$\omega \wedge_\beta \vartheta = \omega^i \wedge \vartheta^j\, \beta(e_i, e'_j),$$

where (e_i) (resp. (e'_j)) is a basis of W_1 (resp. W_2). Let us fix a basis (f_h) for Z and write $\beta(e_i, e'_j) = a^h_{ij}\, f_h$. With this we can compute

$$\begin{aligned}
d(\omega \wedge_\beta \vartheta) &= d(a^h_{ij}\, \omega^i \wedge \vartheta^j) \\
&= a^h_{ij} d(\omega^i \wedge \vartheta^j) \\
&= a^h_{ij}\left(d\omega^i \wedge \vartheta^j + (-1)^s \omega^i \wedge d\vartheta^j\right) \\
&= d\omega^i \wedge \vartheta^j\, \beta(e_i, e'_j) + (-1)^s \omega^i \wedge d\vartheta^j\, \beta(e_i, e'_j).
\end{aligned}$$

On the other hand, by definition of the exterior differential on bundle-valued forms:

$$d\omega \wedge_\beta \vartheta = (d\omega^i \otimes e_i) \wedge_\beta (\vartheta^j \otimes e'_j) = d\omega^i \wedge \vartheta^j\, \beta(e_i, e'_j)$$

and similarly

$$\omega \wedge_\beta d\vartheta = \omega^i \wedge d\vartheta^j\, \beta(e_i, e'_j),$$

thus proving the desired formula and finishing the proof.

(♣) Problem M.7. Compute the derivative of the map Ψ from (24.2) (used in the proof of Proposition 24.8), and show that its derivative is invertible.

Solution. A wholesome exercise like this is best left unsolved.

[3]Linearity is obvious.

Problem Sheet N

Problem N.1. Let $\pi\colon E \to M$ be a vector bundle, let $\mathcal{H}$ denote a connection on E, and let $o\colon M \to E$ denote the zero section. Prove that

$$\mathcal{H}_{0_x} = Do(x)[T_x M], \qquad \forall x \in M,$$

where 0_x is the zero element of the vector space E_x.

Problem N.2. Let $\pi\colon E \to M$ be a vector bundle. Prove that a preconnection $\mathcal{H}$ on E is a vector subbundle of TE such that $(\pi_{TE}, D\pi)|_{\mathcal{H}}$ is a fibre-preserving diffeomorphism from the composite bundle $\mathcal{H} \xrightarrow{\pi \circ \pi_{TE}} M$ to the bundle $E \oplus TM$.

Problem N.3. Recall from Problem C.7 that if we let $\imath\colon S^n \hookrightarrow \mathbb{R}^{n+1}$ denote the inclusion then

$$D\imath(x)[T_x S^n] = \mathcal{J}_x(x^\perp),$$

where $\mathcal{J}_x\colon \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ was defined in Problem B.3 and

$$x^\perp := \left\{ y \in \mathbb{R}^{n+1} \mid \langle x, y\rangle = 0 \right\},$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean dot product. Use this to prove that one can identify

$$T_{(x,v)}TS^n = \left\{ (u, w) \in \mathbb{R}^{2n+2} \mid \langle x, u\rangle = 0 = \langle x, w\rangle + \langle v, u\rangle \right\}.$$

Prove that

$$\mathcal{H}_{(x,v)} := \left\{ \big(u, -\langle v, u\rangle x\big) \mid u \in \mathbb{R}^{n+1},\ \langle x, u\rangle = 0 \right\} \subset T_{(x,v)}TS^n$$

defines a connection on TS^n.

(♣) Problem N.4. Take $n = 2$ and use the connection on TS^2 from Problem N.3. Let $x_N = (0,0,1)$ denote the North pole.

1. Let γ be a great circle. Compute $\widehat{\mathbb{P}}_\gamma\colon T_{\gamma(0)}S^2 \to T_{\gamma(0)}S^2$.
2. Given $s \in (-\pi.\pi)$, let

   $$\gamma_s(t) := \big(\cos t \sin s, \sin t \sin s, \cos s \big)$$

   Compute $\widehat{\mathbb{P}}_{\gamma_s}\colon T_{\gamma_s(0)}S^2 \to T_{\gamma_s(0)}S^2$.

(♣) Problem N.5. Let $\pi\colon E \to M$ be a fibre bundle with fibre F and structure group G. Let $\gamma\colon (a,b) \to M$ be a smooth curve. Prove that $\gamma^\star E \to (a,b)$ (which is another fibre bundle with fibre F and structure group contained in G, c.f. Problem G.7) is a trivial bundle.

Solutions to Problem Sheet N

Problem N.1. Let $\pi\colon E \to M$ be a vector bundle, let $\mathcal{H}$ denote a connection on E, and let $o\colon M \to TM$ denote the zero section. Prove that

$$\mathcal{H}_{0_x} = Do(x)[T_xM], \qquad \forall x \in M,$$

where 0_x is the zero element of the vector space E_x.

Solution. Fix $x \in M$ and let $p \in E_x$. Given $\zeta \in T_pE$, consider a curve $\gamma\colon (-\varepsilon, \varepsilon) \to E$ with $\gamma(0) = p$ and $\gamma'(0) = \zeta$. Then $\mu_0 \circ \gamma$ haa image in $o(M)$, so that $D\mu_0(p)[\zeta] \in Do(x)[T_xM]$. By (28.5) we have

$$\mathcal{H}_{0_x} = D\mu_0(p)[\mathcal{H}_p],$$

and thus $\mathcal{H}_{0_x} \subset Do(x)[T_xM]$. But by the definition of a preconnection

$$\dim \mathcal{H}_{0_x} = \dim Do(x)[T_xM],$$

and thus these two spaces must coincide.

Problem N.2. Let $\pi\colon E \to M$ be a vector bundle. Prove that a preconnection $\mathcal{H}$ on E is a vector subbundle of TE such that $(\pi_{TE}, D\pi)|_{\mathcal{H}}$ is a fibre-preserving diffeomorphism from the composite bundle $\mathcal{H} \xrightarrow{\pi_{TM}\circ\pi} M$ to the bundle $E \oplus TM$.

Solution. Recall that, by definition of preconnection, the map $D\pi(p)|_{\mathcal{H}_p}\colon \mathcal{H}_p \to T_{\pi(p)}M$ is a linear isomorphism. Therefore the map $(D\pi(p)|_{\mathcal{H}_p})^{-1}\colon T_{\pi(p)}M \to \mathcal{H}_p$ is a well-defined isomorphism.

More generally, the map

$$\begin{array}{rcl} \mathscr{F}\colon E \oplus TM & \to & \mathcal{H} \\ (p, v) & \mapsto & (D\pi(p)|_{\mathcal{H}_p})^{-1}(v) \end{array}$$

is well-defined and smooth. Indeed, by definition of sum of fibre bundles, for every $(p, v) \in E \oplus TM$ it holds that $v \in T_{\pi(p)}M$.

Let us now prove that the map $\mathscr{F}$ is the inverse of $(\pi_{TE}, D\pi)|_{\mathcal{H}}$, that is $\mathscr{F} \circ (\pi_{TE}, D\pi)|_{\mathcal{H}} = \mathrm{id}_{\mathcal{H}}$ and $(\pi_{TE}, D\pi)|_{\mathcal{H}} \circ \mathscr{F} = \mathrm{id}_{E\oplus M}$. This would prove that $(\pi_{TE}, D\pi)|_{\mathcal{H}}$ is a diffeomorphism.

This is straightforward to check. For $z \in \mathcal{H}$, let us denote $p = \pi_{TE}(z)$, then

$$\mathscr{F} \circ (\pi_{TE}, D\pi)|_{\mathcal{H}}(z) = \mathscr{F}(p, D\pi|_{\mathcal{H}_p}(z)) = (D\pi(p)|_{\mathcal{H}_p})^{-1}(D\pi|_{\mathcal{H}_p}(z)) = z.$$

On the other hand, given $(p, v) \in E \oplus TM$, it holds

$$(\pi_{TE}, D\pi)|_{\mathcal{H}} \circ \mathscr{F}(p, v) = (\pi_{TE}, D\pi)|_{\mathcal{H}}((D\pi(p)|_{\mathcal{H}_p})^{-1}(v)) = (p, v).$$

It is only left to check that $(\pi_{TE}, D\pi)|_{\mathcal{H}}$ is fibre-preserving as diffeomorphism from the composite bundle $\mathcal{H} \xrightarrow{\pi\circ\pi_{TE}} M$ to the bundle $E \oplus TM$. For this purpose, it is sufficient to check that $z \in \mathcal{H}$ and $(\pi_{TE}, D\pi)|_{\mathcal{H}}(z)$ have the same base point in M, which is however patently true.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati. Last modified: June 28, 2019.

Problem N.3. Recall from Problem C.7 that if we let $\imath\colon S^n \hookrightarrow \mathbb{R}^{n+1}$ denote the inclusion then

$$D\imath(x)[T_xS^n] = \mathcal{J}_x(x^\perp),$$

where $\mathcal{J}_x\colon \mathbb{R}^{n+1} \to T_x\mathbb{R}^{n+1}$ was defined in Problem B.3 and

$$x^\perp := \left\{ y \in \mathbb{R}^{n+1} \mid \langle x, y\rangle = 0 \right\},$$

where $\langle\cdot,\cdot\rangle$ is the standard Euclidean dot product. Use this to prove that one can identify

$$T_{(x,v)}TS^n = \left\{ \big((x,v),(u,w)\big) \in TS^n \times \mathbb{R}^{2n+2} \mid \langle x,u\rangle = 0 = \langle x,w\rangle + \langle v,u\rangle \right\}.$$

Prove that

$$\mathcal{H}_{(x,v)} := \left\{ \big((x,v),(u,-\langle v,u\rangle x\big) \mid u \in \mathbb{R}^{n+1},\ \langle x,u\rangle = 0 \right\}$$

defines a connection on TS^n.

Solution. Thanks to the canonical isomorphism $D\imath(x)$ above, we may write

$$T_xS^n = \{v \in \mathbb{R}^{n+1} : \langle x,v\rangle = 0\}$$

and

$$TS^n = \{(x,v) \in \mathbb{R}^{2n+2} : |x|^2 = 1, \langle x,v\rangle = 0\}.$$

To deduce the expression for $T_{(x,v)}TS^n$ given above, consider an arbitrary curve $\alpha : (-\varepsilon,\varepsilon) \to TS^n$, $\alpha(t) = (x(t),v(t))$ with $\alpha(0) = (x(0),x(0)) = (x,v)$ and set $(u,w) = \alpha'(0) = (x'(0),v'(0))$. Note now that differentiating at $t = 0$ the expressions

$$|x(t)|^2 = 1 \quad \text{and} \quad \langle x(t),v(t)\rangle = 0,$$

we deduce that

$$\langle u,x\rangle = 0 \quad \text{and} \quad \langle u,v\rangle + \langle x,w\rangle = 0.$$

Since α is arbitrary, (u,v) defines an arbitrary element of $T_{(x,v)}TS^n$ and we then deduce the inclusion

$$T_{(x,v)}TS^n \subseteq \{(u,w) \in \mathbb{R}^{2n+2} : \langle u,x\rangle = 0,\ \langle u,v\rangle + \langle x,w\rangle = 0\}.$$

On the other hand, the set on the right-hand side is also a vector subspace of $\mathbb{R}^{2n+2}$ whose dimension equals $2n$ because (for example) it is the preimage of (0,0) of the function

$$F : \mathbb{R}^{2n+2} \to \mathbb{R}^2, \quad (u,w) \mapsto (\langle u,x\rangle, \langle u,v\rangle + \langle x,w\rangle),$$

and (0,0) is a regular value for F. The two spaces have the same dimension and must then coincide.

Now let us prove that $\mathcal{H}$ defines a preconnection on TS^n. The fact that, for every (x,v), $\mathcal{H}_{(x,v)}$ is a vector subspace of $T_{(x,v)}TS^n$ follows because the relations

$$\langle x,u\rangle = 0 \quad \text{and} \quad w = -\langle v,u\rangle x$$

are linear in u and w and moreover each couple (u, w) which satisfies such relations also satisfies

$$\langle u, v\rangle + \langle x, w\rangle = 0,$$

and so (u, w) is an element of $T_{(x,v)}TS^n$ by what we proved above. Let us now compute the differential of $\pi : TS^n \to S^n$ at a point (x, v) evaluated at a vector (u, w), by picking the same curve α given above:

$$D\pi(x, v)[(u, w)] = \left.\frac{d\pi(\alpha(t))}{dt}\right|_{t=0} = \left.\frac{dx(t)}{dt}\right|_{t=0} = u.$$

It follows that, when restricted to $\mathcal{H}_{(x,v)}$, $D\pi(x, v)$ is bijective, and consequently $\mathcal{H}$ is a preconection on TS^n.

Finally, to prove that $\mathcal{H}$ is a connection let us compute the differential of the scalar multiplication map $\mu_a(x, v) = (x, av)$, for a fixed $a \in \mathbb{R}$, similarly as before:

$$D\mu_a(x, v)[(u, w)] = \left.\frac{d\mu_a(\alpha(t))}{dt}\right|_{t=0} = \left.\frac{d}{dt}(x(t), av(t))\right|_{t=0} = (u, aw).$$

Now, for $a = 0$, we directly see that $D\mu_0[\mathcal{H}_{(x,v)}] = \mathcal{H}_{(x,0)}$, and for $a \neq 0$,

$$w = -\langle v, u\rangle x \quad \text{if and only if} \quad aw = -\langle av, u\rangle x,$$

so we conclude that $D\mu_a(x, v)[\mathcal{H}_{(x,v)}] = \mathcal{H}_{(x,av)}$.

(♣) PROBLEM N.4. Take $n = 2$ and use the connection on TS^2 from Problem N.3.

1. Let γ be a great circle. Compute $\widehat{\mathbb{P}}_\gamma \colon T_{\gamma(0)}S^2 \to T_{\gamma(0)}S^2$.
2. Given $s \in (-\pi.\pi)$, let

$$\gamma_s(t) := \big(\cos t \sin t, \sin t \sin s, \cos s\big)$$

   Compute $\widehat{\mathbb{P}}_{\gamma_s} \colon T_{\gamma_s(0)}S^2 \to T_{\gamma_s(0)}S^2$.

SOLUTION. First of all, for any curve $\gamma \colon [a, b] \to S^2$ we have

$$\begin{aligned} \gamma^\star(TS^2) &= \{(t, v) \in [a, b] \times \mathbb{R}^3 \mid \langle\gamma(t), v\rangle = 0\}, \\ (\gamma^\star\mathcal{H})_{(t,v)} &= \{(u, v) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \langle\gamma(t), u\rangle = 0,\ w = -\langle v, u\rangle\gamma(t)\}, \end{aligned}$$

Moreover from the previous exercise we recall that, for every $(x, v) \in TS^2$, the isomorphism between T_xS^2 and $\mathcal{H}_{(x,v)}$ induced by the projection $\pi \colon TS^2 \to S^2$ is given by

$$D\pi(x, v)[(u, w)] = u \quad \text{with inverse} \quad D\pi(x, v)^{-1}[V] = (V, -\langle v, V\rangle x),$$

consequently, the induced isomorphism on the pulled-back bundles is given by

$$T_t[a, b] \to (\gamma^\star\mathcal{H})_{(t,v)}, \qquad \frac{\partial}{\partial t} \mapsto (\gamma'(t), -\langle v, \gamma'(t)\rangle\gamma(t))$$

For any fixed $v \in T_{\gamma(a)}S^2$, the parallel transport of v along γ, $\widehat{\mathbb{P}}_\gamma(v)$ will be given by the solution, evaluated at $t = b$, of the following Cauchy problem:

$$\begin{cases} v'(t) = -\langle v(t), \gamma'(t)\rangle \gamma(t), & \text{for } t \in (a,b), \\ v(0) = v. \end{cases}$$

1. Let us first compute the parallel transport map in the case of the great circle in the xz-plane starting from the north pole, namely

$$\gamma\colon [0, 2\pi] \to S^2, \qquad \gamma(t) = \begin{pmatrix} \sin t \\ 0 \\ \cos t \end{pmatrix},$$

which, being $\gamma(0) = x_N = \gamma(2\pi)$, we will be an endomorphism $\mathbb{P}_\gamma\colon T_{x_N}S^2 \to T_{x_N}S^2$. Now, a basis of $T_{x_N}S^2$ is given by the vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

and if we know how these two vector are mapped under $\mathbb{P}_\gamma$, the transformation on any other vector follows by linearity. For $\mathbf{e}_1$, we have to solve

$$\begin{cases} (v^1)'(t) = \big(-\cos t\, v^1(t) + \sin t\, v^3(t) \big) \sin t, \\ (v^2)'(t) = 0, \\ (v^3)'(t) = \big(-\cos t\, v^1(t) + \sin t\, v^3(t) \big) \cos t, \\ \quad v^1(0) = 1, \\ \quad v^2(0) = 0, \\ \quad v^3(0) = 0, \end{cases}$$

and a moment of thought reveals that the (necessarily unique) solution to this system is given by $(\cos t, 0, -\sin t)$. Consequently, we have $\widehat{\mathbb{P}}_\gamma(\mathbf{e}_1) = \mathbf{e}_1$. Similarly, for $\mathbf{e}_2$ we have to solve:

$$\begin{cases} (v^1)'(t) = \big(-\cos t\, v^1(t) + \sin t\, v^3(t) \big) \sin t, \\ (v^2)'(t) = 0, \\ (v^3)'(t) = \big(-\cos t\, v^1(t) + \sin t\, v^3(t) \big) \cos t, \\ \quad v^1(0) = 0, \\ \quad v^2(0) = 1, \\ \quad v^3(0) = 0, \end{cases}$$

and the only solution to this problems is given by the constant map $v(t) = (0, 1, 0)$, thus yielding that $\widehat{\mathbb{P}}_\gamma(\mathbf{e}_2) = \mathbf{e}_2$. Necessarily than $\widehat{\mathbb{P}}_\gamma$ has to be the identity from $T_{x_N}S^2$ onto itself.

If γ is a parametrisation (in arc-length without loss of generality) of any other great circle of S^2, then we may suppose that its domain is $[0, 2\pi]$ and consequently, we may find an orthogonal transformation $A \in O(3)$ so that $A\gamma = (\sin t, 0, \cos t)$. Since

$$\begin{cases} v'(t) = -\langle v(t), \gamma'(t)\rangle \gamma(t), \\ v(0) = v., \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} Av'(t) = -\langle Av(t), A\gamma'(t)\rangle A\gamma(t), \\ Av(0) = Av, \end{cases}$$

it follows that $\widehat{\mathbb{P}}_\gamma(v) = A^{-1}\widehat{\mathbb{P}}_{A\gamma}(Av)$. Since above we computed that $\widehat{\mathbb{P}}_{A\gamma}$ is the identity, so is $\widehat{\mathbb{P}}_\gamma$ for any great circle γ.

2. For every s, a basis of the tangent plane $T_{\gamma_s(0)}S^2$ is given by

$$\mathbf{f}_1 = \begin{pmatrix} \cos s \\ 0 \\ -\sin s \end{pmatrix} \quad \text{and} \quad \mathbf{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

so it is enough to know how $\widehat{\mathbb{P}}_{\gamma_s}$ transforms these two vectors. As before, to find $\widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_1)$ we have to solve the problem

$$\begin{cases} (v^1)'(t) = (\sin s)^2 \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big) \cos t, \\ (v^2)'(t) = (\sin s)^2 \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big) \sin t, \\ (v^3)'(t) = \sin s\, \cos s\, \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big), \\ \quad v^1(0) = \cos s, \\ \quad v^2(0) = 0, \\ \quad v^3(0) = -\sin s, \end{cases}$$

while to find $\widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_2)$ we have to solve the problem

$$\begin{cases} (v^1)'(t) = (\sin s)^2 \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big) \cos t, \\ (v^2)'(t) = (\sin s)^2 \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big) \sin t, \\ (v^3)'(t) = \sin s\, \cos s\, \big(\sin t\, v^1(t) - \cos t\, v^2(t)\big), \\ \quad v^1(0) = 0, \\ \quad v^2(0) = 1, \\ \quad v^3(0) = 0, \end{cases}$$

while In this case we need to sweat considerably more than before, however it is still possible to find explicit solutions, namely

$$\begin{pmatrix} (\cos s)\cos(t\, \cos s) \\ -\sin(t\, \cos s) \\ -(\cos s)\sin(t\, \cos s) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} (\cos s)\sin(t\, \cos s) \\ \cos(t\, \cos s) \\ -(\sin s)\sin(t\, \cos s) \end{pmatrix}$$

respectively. We thus deduce that

$$\widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_1) = \begin{pmatrix} (\cos s)\cos(2\pi\, \cos s) \\ -\sin(2\pi\, \cos s) \\ -(\cos s)\sin(2\pi\, \cos s) \end{pmatrix} \quad \text{and} \quad \widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_2) = \begin{pmatrix} (\cos s)\sin(2\pi\, \cos s) \\ \cos(2\pi\, \cos s) \\ -(\sin s)\sin(2\pi\, \cos s) \end{pmatrix},$$

a fact that we can rewrite more meaningfully in matrix-form as

$$(\widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_1), \widehat{\mathbb{P}}_{\gamma_s}(\mathbf{f}_2)) = (\mathbf{f}_1, \mathbf{f}_2)\begin{pmatrix} \cos(2\pi\cos s) & -\sin(2\pi\cos s) \\ \sin(2\pi\cos s) & \cos(2\pi\cos s) \end{pmatrix}$$

and which allows us to conclude that $\widehat{\mathbb{P}}_{\gamma_s} : T_{\gamma_s(0)}S^2 \to T_{\gamma_s(0)}S^2$ is a counter-clockwise rotation of angle $2\pi \cos s$ on that plane (with the orientation taken in accordance with the basis $(\mathbf{f}_1, \mathbf{f}_2)$).

(♣) PROBLEM N.5. Let $\pi \colon E \to M$ be a fibre bundle with fibre F and structure group G. Let $\gamma \colon (a, b) \to M$ be a smooth curve. Prove that $\gamma^\star E \to (a, b)$ (which is another fibre bundle with fibre F and structure group contained in G, c.f. Problem G.7) is a trivial bundle.

SOLUTION. More generally, we show that any fibre bundle E' over (a, b), with structure group G, is trivial.[1]

We call $\mu \colon G \times F \to F$ the given action of G on the fibre F and we call $\alpha \colon U_\alpha \to F$ a generic bundle chart in the given G-bundle atlas.

Showing that E' is trivial amounts to find a map $\beta \colon E' \to F$ such that $\pi \times \beta \colon E' \to (a, b) \times F$ is an isomorphism of fibre bundles with structure group G, meaning that:

(i) $\pi \times \beta$ is a diffeomorphism;

(ii) there exist smooth maps $\psi_\alpha \colon U_\alpha \to G$ such that $\beta(x) = \mu(\psi_\alpha \circ \pi(x), \alpha(x))$.

Indeed, the trivial bundle $(a, b) \times F$ has projection map $(t, z) \mapsto t$. Hence, given a fibre bundle homomorphism $\Phi \colon E' \to (a, b) \times F$ and writing $\Phi = p \times \beta$, we have necessarily $p = \pi$ (since Φ is fibre-preserving).

Notice that condition (ii) does not depend on the G-bundle atlas, provided our atlas is replaced with a G-compatible one (meaning that all the bundle charts in the first one are G-compatible with all the bundle charts in the second one). Condition (ii) can be dropped if one is merely interested in showing the triviality of E' as a fibre bundle with fibre F, but it will come for free from the construction of β.

It is easy to find an increasing sequence $(t_k)_{k\in\mathbb{Z}} \subseteq (a, b)$ with $\lim_{k\to-\infty} t_k = a$ and $\lim_{k\to\infty} t_k = b$, together with bundle charts $\alpha_k \colon U_{\alpha_k} \to G$ with $(t_{k-1}, t_{k+1}) \subseteq U_{\alpha_k}$. For instance, take first an increasing sequence $(s_j)_{j\in\mathbb{Z}}$ with $\lim_{j\to-\infty} s_j = a$ and $\lim_{j\to\infty} s_j = b$, and take a further subdivision of the intervals $[s_j, s_{j+1}]$ into (finitely many) subintervals whose size is less than half the Lebesgue number of $[s_{j-1}, s_{j+1}]$ with respect to the open cover (U_α). The sequence (t_k) is then constructed starting from $t_0 := s_0$ and enumerating the endpoints of the new intervals at the left and right of s_0 in an increasing way. By construction, each interval $[t_{k-1}, t_{k+1}]$ is then covered by the domain U_{α_k} of some bundle chart.

Observe that $\rho_{\alpha_0\alpha_1}$ is given by left multiplication by some smooth function $\widehat{\rho} \colon (t_0, t_1) \to G$, namely

$$\rho_{\alpha_0\alpha_1}(t)(z) = \mu(\widehat{\rho}_{\alpha_0\alpha_1}(t), z)$$

[1] Actually, this fact is equivalent to the statement of the exercise: to see the converse implication, we can just choose $E := E'$, $M := (a, b)$, $\gamma := \mathrm{id}$ and observe that the pullback bundle becomes $\gamma^* E = \mathrm{id}^* E'$, which is canonically isomorphic to E'.

for all $t \in (t_0, t_1)$ and all $z \in F$. It would then make sense to define β piecewise, with $\beta \coloneqq \alpha_0$ on $\pi^{-1}((t_{-1}, t_1))$ and $\beta(x) \coloneqq \mu(\psi_{0,1} \circ \pi(x), \alpha_1(x))$ on $\pi^{-1}((t_0, t_2))$, where $\psi_{0,1}$ is some smooth extension of $\widehat{\rho}_{\alpha_0\alpha_1}$. The problem is that such extension may not always exist, since $\widehat{\rho}_{\alpha_0\alpha_1}(t)$ could oscillate too wildly as $t \uparrow t_1$!

To overcome this difficulty, we need to twist the maps α_k in such a way that this becomes possible. We claim that there exists a curve $\sigma_1 \colon (t_0, t_1) \to G$ such that

(i') $\sigma_1(t) = e$ for t close to t_0, say $t < t_0 + \epsilon_1$ (for some $\epsilon_1 > 0$);

(ii') $\sigma_1(t)\widehat{\rho}_{\alpha_0\alpha_1}$ is constant for t close to t_1, say equal to some $g_1 \in G$ for $t > t_1 - \epsilon_1$.

For instance, let $\epsilon_1 \coloneqq \frac{t_1 - t_0}{3}$ and $\sigma_1(t) \coloneqq g_1 \cdot \widehat{\rho}_{\alpha_0\alpha_1}(\tau_1(t))^{-1}$, where $g_1 \coloneqq \widehat{\rho}_{\alpha_0\alpha_1}(t_0 + \epsilon_1)$ and $\tau_1 \colon (t_0, t_1) \to \mathbb{R}$ is a smooth nondecreasing function such that $\tau_1(t) = t_0 + \epsilon_1$ for $t < t_0 + \epsilon_1$ and $\tau_1(t) = t$ for $t > t_1 - \epsilon_1$.[2]

Similarly, for all $k > 0$, we can find $\sigma_k \colon (t_{k-1}, t_k) \to G$, ϵ_k and $g_k \in G$ such that $\sigma_k(t) = e$ for $t < t_{k-1} + \epsilon_k$ and $\sigma_k(t)\widehat{\rho}_{\alpha_{k-1}\alpha_k} = g_k$ for $t > t_k - \epsilon_k$.

Analogously we construct $\sigma_{-k} \colon (t_{-k}, t_{-(k-1)}) \to G$ and g_{-k} (again for $k > 0$), asking that $\sigma_{-k}(t) = e$ for $t > t_{-(k-1)} - \epsilon_{-k}$ and $\sigma_{-k}(t)\widehat{\rho}_{\alpha_{-(k-1)}\alpha_{-k}} = g_{-k}$ for $t < t_{-k} + \epsilon_{-k}$. Now on $\pi^{-1}((-t_1, t_1))$ we let

$$\beta(x) \coloneqq \begin{cases} \mu(\sigma_1 \circ \pi(x), \alpha_0(x)) & \text{if } \pi(x) \geq 0, \\ \mu(\sigma_{-1} \circ \pi(x), \alpha_0(x)) & \text{if } \pi(x) \leq 0, \end{cases}$$

and, on $\pi^{-1}((t_k - \epsilon_k, t_{k+1}))$ (for $k > 0$), we define

$$\beta(x) \coloneqq \begin{cases} \mu(g_1 \cdots g_k, \alpha_k(x)) & \text{if } \pi(x) \leq t_k, \\ \mu(g_1 \cdots g_k \cdot \sigma_{k+1} \circ \pi(x), \alpha_k(x)) & \text{if } \pi(x) \geq t_k, \end{cases}$$

and symmetrically on $\pi^{-1}((t_{-(k+1)}, t_{-k} + \epsilon_k))$. The smoothness of β, as well as properties (i) and (ii), are clear. We only have to check that β is well defined. If $\pi(x) \in (t_k - \epsilon_k, t_k)$ with $k > 0$ then we have to check that

$$\mu(g_1 \cdots g_{k-1} \cdot \sigma_k \circ \pi(x), \alpha_{k-1}(x)) = \mu(g_1 \cdots g_k, \alpha_k(x))$$

(where $g_1 \cdots g_{k-1} = e$ if $k = 1$), but this holds as

$$\begin{aligned} &\mu(g_1 \cdots g_{k-1} \cdot \sigma_k \circ \pi(x), \alpha_{k-1}(x)) \\ &= \mu(g_1 \cdots g_{k-1} \cdot \sigma_k \circ \pi(x), \mu(\rho_{\alpha_{k-1}\alpha_k} \circ \pi(x), \alpha_k(x))) \\ &= \mu(g_1 \cdots g_{k-1} \cdot (\sigma_k \rho_{\alpha_{k-1}\alpha_k}) \circ \pi(x), \alpha_k(x)) \\ &= \mu(g_1 \cdots g_k, \alpha_k(x)) \end{aligned}$$

by construction and by definition of left action.

[2] Given a nonnegative bump function $\varphi \colon \mathbb{R} \to \mathbb{R}$ compactly supported in $(t_0 + \epsilon_1, t_1 - \epsilon_1)$ and with $\int_{-\infty}^{\infty} \varphi(t)\, dt = 1$, an example of such τ_1 is $\tau_1(t) \coloneqq t_0 + \epsilon_1 + (t - t_0 - \epsilon_1) \int_{-\infty}^{t} \varphi(s)\, ds$.

Problem Sheet O

Problem O.1. Let $\pi\colon E \to M$ be a vector bundle of rank k with parallel transport structure $\mathbb{P}$. Fix $x \in M$ and let $\{p_1, \dots, p_k\}$ be a basis of E_x. Suppose $\psi\colon V_x \to U$ is a diffeomorphism, where V_x is a starshaped open set in T_xM about 0_x and U is a neighbourhood of x in M. Define for $v \in V_x$ a smooth curve

$$\gamma_v\colon [0,1] \to M, \qquad \gamma_v(t) := \psi(tv).$$

Prove there exists a local frame $\{e_1, \dots, e_k\}$ of E over U such that $e_i(x) = p_i$ and such that $e_i(\gamma_v(t))$ is a parallel along γ_v for each $i = 1, \dots, k$ and all $v \in V_x$. **Remark:** Lemma 31.5 is a special case of this problem.

(♣) Problem O.2. Let $\mathcal{H}$ denote the connection on TS^n from Problem N.3.

(i) Find an explicit formula for the connection map $\kappa\colon T(TS^n) \to TS^n$ and for the covariant derivative operator $\nabla\colon \mathfrak{X}(S^n) \times \mathfrak{X}(S^n) \to \mathfrak{X}(S^n)$.

(ii) Let x, y be two points in S^n such that $x \perp y$. Let $\gamma\colon [0, 2\pi] \to S^n$ denote the great circle $\gamma(t) = (\cos t)x + (\sin t)y$. Prove that the covariant derivative operator along γ satisfies $\nabla_T(\gamma') = 0$, where $T \in \mathfrak{X}([0, 2\pi])$ is the vector field $\frac{\partial}{\partial t}$.

Problem O.3. Let $\mathcal{H}$ be a connection in a vector bundle $\pi\colon E \to M$ with associated parallel transport system $\mathbb{P}$ and covariant derivative $\nabla\colon \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E)$.

(i) Define the **dual parallel transport system** in the dual bundle E^* by declaring that a section $\rho \in \Gamma_\gamma(E^*)$ is parallel if and only if $\rho(c)$ is constant for every parallel section $c \in \Gamma_\gamma(E)$. Prove *directly* that this defines a parallel transport system. (You may skip the verification of Axiom (iii) of Definition 29.8!).

(ii) Define the **dual covariant derivative operator** $\nabla^*\colon \mathfrak{X}(M) \times \Gamma(E^*) \to \Gamma(E^*)$ defined by

$$(\nabla^*_X \rho)(s) = X(\rho(s)) - \rho(\nabla_X(s)).$$

Prove *directly* that this is a covariant derivative operator in E^*.

(iii) The **dual connection** on E^* is the connection $\mathcal{H}^*$ whose associated parallel transport system is the dual parallel transport system from part (i) and whose associated covariant derivative operator is the dual covariant derivative operator from part (ii). How does one define $\mathcal{H}^*$ explicitly?

(♣) Problem O.4. Let $\pi_i\colon E_i \to M_i$ be two vector bundles, and let $\varphi\colon M_1 \to M_2$ denote a smooth map. Suppose $\Phi\colon E_1 \to E_2$ is a smooth map such that the following diagram commutes:

$$\begin{tikzcd} E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\ M_1 \arrow[r, "\varphi"'] & M_2 \end{tikzcd}$$

Note we are *not* assuming that Φ is linear on the fibres, and hence Φ need not be a vector bundle morphism along φ. For each $x \in M_1$, Φ defines a smooth map

$$\Phi_x := \Phi|_{E_1|_x} \colon E_1|_x \to E_2|_{\varphi(x)}.$$

This is a map between two vector spaces, so for any $p \in E_1|_x$ we can take its derivative

$$D\Phi_x(p) \colon T_p E_1|_x \to T_{\Phi(p)} E_2|_{\varphi(x)}.$$

Composing with the $\mathcal{J}$ maps from Problem B.3, we get a linear map from

$$\widehat{D}\Phi_x(p) := \mathcal{J}^{-1}_{\Phi(p)} \circ D\Phi_x(p) \circ \mathcal{J}_p \colon E_1|_x \to E_2|_{\varphi(x)}.$$

Consider the vector bundle $\tilde{\pi} \colon \mathrm{Hom}(E_1, \varphi^\star E_2) \to M_1$ over M_1. The fibre of this bundle over $x \in M_1$ is $\mathrm{L}(E_1|_x, E_2|_{\varphi(x)})$. Since Φ is smooth, the map $p \mapsto \widehat{D}\Phi_x(p)$ defines a smooth map $D^{\mathrm{fibre}}\Phi \colon E_1 \to \mathrm{Hom}(E_1, \varphi^\star E_2)$ which we call the **fibrewise derivative** of Φ:

$$\begin{tikzcd} E_1 \arrow[rr, "D^{\mathrm{fibre}}\Phi"] \arrow[dr, "\pi_1"'] & & \mathrm{Hom}(E_1, \varphi^\star E_2) \arrow[dl, "\tilde{\pi}"] \\ & M_1 & \end{tikzcd}$$

Note again that $D^{\mathrm{fibre}}\Phi$ is *not* necessarily linear on the fibres (i.e. $\widehat{D}\Phi_x(p)$ does not have to depend linearly on p), and thus $D^{\mathrm{fibre}}\Phi$ is *not* necessarily a vector bundle morphism along φ.

(i) Show that the normal derivative $D\Phi \colon TE_1 \to TE_2$ of Φ restricts to define a map $D\Phi|_{VE_1} \colon VE_1 \to VE_2$. Prove that if $p, q \in E_1|_x$ then

$$D^{\mathrm{fibre}}\Phi(p)[q] = \mathrm{pr}_2^{E_2} \circ D\Phi|_{VE_1} \circ \mathcal{J}_p(q),$$

where $\mathrm{pr}_2^{E_2} \colon VE_2 \to E_2$ is the "projection onto the second factor" map (see (30.1) from Lecture 30 or Problem I.5).

(ii) Now take $M_1 = M_2 = M$ and φ to be the identity. Let $E_1 = TM$ denote the tangent bundle and let $E_2 = M \times \mathbb{R}$ denote the trivial bundle. Then a fibre preserving map $\Phi \colon E_1 \to E_2$ can be identified with a smooth function $f \colon TM \to \mathbb{R}$. Prove that the definition of $D^{\mathrm{fibre}}f$ given above is consistent with the fibrewise derivative $D^{\mathrm{fibre}}f \colon TM \to T^*M$ given in Problem C.3.

(iii) Now return to the general setup, and assume that both E_1 and E_2 are endowed with connections $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Let $\kappa_i \colon TE_i \to E_i$ denote the connection map of $\mathcal{H}_i$. We define the **parallel derivative** of Φ to be the map

$$D^{\mathrm{parallel}}\Phi \colon E_1 \to \mathrm{Hom}(TM_1, \varphi^\star E_2)$$

by

$$D^{\mathrm{parallel}}\Phi(p)[v] := \kappa_2 \circ D\Phi(p)[\overline{v}], \qquad p \in E_1, \ v \in TM_1,$$

where $\overline{v}$ is the horizontal lift of v at p with respect to $\mathcal{H}_1$ (see Definition 28.7). Prove that for $x \in M_1$, $p \in E_1|_x$ and $\zeta \in T_pE_1$ that the following formula holds:

$$\kappa_2\big(D\Phi(p)[\zeta]\big) = D^{\mathrm{fibre}}\Phi(p)[\kappa_1(\zeta)] + D^{\mathrm{parallel}}\Phi(p)\big[D\pi_1(p)[\zeta]\big].$$

(iv) Conclude that $D\Phi$ is entirely determined by $D\varphi$, $D^{\mathrm{fibre}}\Phi$ and $D^{\mathrm{parallel}}\Phi$. That is, under the vector bundle isomorphism $(D\pi_i, \kappa_i) \colon TE_i \to TM_i \oplus E_i$ along π_i given by Lemma 31.3, $D\Phi$ takes matrix form:

$$D\Phi = \begin{pmatrix} D\varphi & 0 \\ D^{\mathrm{parallel}}\Phi & D^{\mathrm{fibre}}\Phi \end{pmatrix}$$

This formula is often very useful in computations.

Solutions to Problem Sheet O

Problem O.1. Let $\pi\colon E \to M$ be a vector bundle of rank k with parallel transport structure $\mathbb{P}$. Fix $x \in M$ and let $\{p_1, \dots, p_k\}$ be a basis of E_x. Suppose $\psi\colon V_x \to U$ is a diffeomorphism, where V_x is a starshaped open set in T_xM about 0_x and U is a neighbourhood of x in M. Define for $v \in V_x$ a smooth curve

$$\gamma_v\colon [0,1] \to M, \qquad \gamma_v(t) \coloneqq \psi(tv).$$

Prove there exists a local frame $\{e_1, \dots, e_k\}$ of E over U such that $e_i(x) = p_i$ and such that $e_i(\gamma_v(t))$ is a parallel along γ_v for each $i = 1, \dots, k$ and all $v \in V_x$. **Remark:** Lemma 31.5 is a special case of this problem.

Solution. Given $y \in U$, let $v \coloneqq \psi^{-1}(y)$. If the local frame $\{e_i\}$ exists, then $e_i \circ \gamma_v$ is parallel along γ_v and $e_i \circ \gamma_v(0) = p_i$. Since $\mathbb{P}_{\gamma_v}(p_i)$ has the same properties, by Proposition 29.7 we must have $e_i \circ \gamma_v = \mathbb{P}_{\gamma_v}(p_i)$. In particular, $e_i(y) = \widehat{\mathbb{P}}_{\gamma_v}(p_i)$. Hence, in order to show that the desired local frame exists, we define $e_i(y) \coloneqq \widehat{\mathbb{P}}_{\gamma_v}(p_i)$ (with $v \coloneqq \psi^{-1}(y)$).

The linear isomorphism axiom for $\mathbb{P}$ implies that $\{e_1(y), \dots, e_k(y)\}$ is a basis of E_y for all $y \in U$, so we are left to prove the smoothness of e_i.

We argue locally, invoking the smoothness axiom (in the way it is formulated in the footnote to Definition 29.8). Given $y_0 \in U$, let $v_0 \coloneqq \psi^{-1}(y_0)$ and pick a chart $\sigma\colon U' \to O$, with $x \in U' \subseteq U$. Then let $W \coloneqq D\sigma(x)[V_x]$, regarded as a subset of $\mathbb{R}^n$ (identifying $T_{\sigma(x)}\mathbb{R}^n$ with $\mathbb{R}^n$).

Next, let $\eta\colon \mathbb{R}^n \to \mathbb{R}$ be a smooth function compactly supported in W, with $0 \le \eta \le 1$ and $\eta = 1$ on a neighbourhood of the segment $S \coloneqq \{tD\sigma(x)[v_0] \mid t \in [0,1]\}$.[1] Since W is also starshaped about 0, we have $\eta(w)w \in W$ for all $w \in W$, and trivially also when $w \notin W$ (since in this case $\eta(w)w = 0$). Now we let

$$\Phi(v) \coloneqq D\sigma(\pi(v))[v] \in \mathbb{R}^n$$

for $v \in TM|_{U'}$ and finally we set

$$\Psi(v) \coloneqq \psi \circ D\sigma(x)^{-1}[\eta(\Phi(v))\Phi(v)].$$

The map $\Psi\colon TM|_{U'} \to M$ is smooth and well defined, since $D\sigma(x)^{-1}[\eta(\Phi(v))\Phi(v)] \in V_x$. It is immediate to check that $\Phi(tv) = D\sigma(x)[tv]$ and $\Psi(tv) = \gamma_v$, for $v \in T_xM$ close enough to v_0 and $t \in [0,1]$. Hence, by the smoothness axiom, $v \mapsto \widehat{\mathbb{P}}_{\gamma_v}(p_i)$ is smooth for $v \in V_x$ near v_0. Thus $e_i(y) = \widehat{\mathbb{P}}_{\gamma_{\psi^{-1}(y)}}(p_i)$ is smooth near y_0, as well.

(♣) Problem O.2. Let $\mathcal{H}$ denote the connection on TS^n from Problem N.3.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

[1] For instance we can apply Lemma 3.2 with $\mathbb{R}^n$, W and S' playing the roles of M, U, K respectively, where S' is a compact neighbourhood of S in W.

(i) Find an explicit formula for the connection map $\kappa\colon T(TS^n) \to TS^n$ and for the covariant derivative operator $\nabla\colon \mathfrak{X}(S^n) \times \mathfrak{X}(S^n) \to \mathfrak{X}(S^n)$.

(ii) Let x, y be two points in S^n such that $x \perp y$. Let $\gamma\colon [0, 2\pi] \to S^n$ denote the great circle $\gamma(t) = (\cos t)p + (\sin t)q$. Prove that the covariant derivative operator along γ satisfies $\nabla_T(\gamma') = 0$, where $T \in \mathfrak{X}([0, 2\pi])$ is the vector field $\frac{\partial}{\partial t}$.

Solution. (i). Recall from Problem N.3 that

$$\begin{aligned}
T_x S^n &= \{v \in \mathbb{R}^{n+1} : \langle x, v\rangle = 0\},\\
T_{(x,v)} TS^n &= \{(u, w) \in \mathbb{R}^{2n+2} : \langle u, v\rangle + \langle x, w\rangle = 0\},\\
\mathcal{H}_{(x,v)} &= \{(u, -\langle v, u\rangle x) \in \mathbb{R}^{2n+2} : \langle x, u\rangle = 0\},\\
\pi : TS^n \to S^n &\quad \text{is given by} \quad \pi((x, v)) = v,\\
D\pi(x, v) : T_{(x,v)} TS^n \to T_x S^n &\quad \text{is given by} \quad D\pi(x, v)[(u, w)] = u,
\end{aligned}$$

consequently, $\ker D\pi(x, v) = \{(0, w) : \langle x, w\rangle = 0\}$, and the map $\mathrm{pr}_2 : VTS^n \to TS^n$ is given by $\mathrm{pr}_2((x, v), (0, w)) = (x, w)$. For every element $\zeta = (u, w) \in T_{(x,v)} TS^n$, its splitting into vertical and horizontal part is then given by

$$\zeta = (u, w) = \underbrace{(0, w + \langle v, u\rangle x)}_{\zeta^{\mathtt{V}}} + \underbrace{(u, -\langle v, u\rangle x)}_{\zeta^{\mathtt{H}}}.$$

From Definition 31.1, the connection map $k : TTS^n \to TS^n$ acts on an element $\zeta = ((x, v), (u, w)) \in TTS^n$ as

$$k((x, v), (u, w)) = (x, w + \langle v, u\rangle x).$$

To compute $\nabla_X Y$ for any two vector fields $X, Y \in \mathfrak{X}(S^n)$, recall that Y can be represented as a smooth map $Y : S^n \to \mathbb{R}^{n+1}$, so that $\langle Y(x), x)\rangle = 0$ for every $x \in S^n$, so if $Z \in T_x S^n$, $DY(x)[Z]$ can be regarded as a vector in $\mathbb{R}^{n+1}$ (where the differential "D" is the one for maps $S^n \to \mathbb{R}^{n+1}$), and its equivalent in $T_{(x,Y(x))} TS^n$, which we denote in the same way, as $(Z, DY(x)[Z])$. Thanks to Theorem 31.10 with $\varphi = \mathrm{id}_{S^n}$ we conclude that

$$\nabla_X Y(x) = k(DY(x)[X(x)]) = DY(x)[X(x)] + \langle Y(x), X(x)\rangle x.$$

(ii). Thanks to Theorem 31.10 with $\varphi = \gamma$, we have that

$$\nabla_T(\gamma')(t) = k(D(\gamma')(t)[T(t)]).$$

Now, for any t we simply have $D(\gamma')(t)[T(t)] = \gamma''(t)$ so we get that $\nabla_T \gamma(t)$ is given by:

$$\nabla_T(\gamma')(t) = \gamma''(t) + |\gamma'(t)|^2\, \gamma(t).$$

Since

$$\begin{aligned}
\gamma'(t) &= -(\sin t)x + (\cos t)y,\\
\gamma''(t) &= -(\cos t)x - (\sin t)y,
\end{aligned}$$

It follows that $\nabla_T(\gamma')(t) = 0$.

Problem O.3. Let $\mathcal{H}$ be a parallel transport system in a vector bundle $\pi\colon E \to M$ with associated parallel transport system $\mathbb{P}$ and covariant derivative $\nabla\colon \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E)$.

(i) Define the **dual parallel transport system** in the dual bundle E^* by declaring that a section $\rho \in \Gamma_\gamma(E^*)$ is parallel if and only if $\rho(c)$ is constant for every parallel section $c \in \Gamma_\gamma(E)$. Prove *directly* that this defines a parallel transport system. (You may skip the verification of Axiom (iii) of Definition 29.8!).

(ii) Define the **dual covariant derivative operator** $\nabla^*\colon \mathfrak{X}(M) \times \Gamma(E^*) \to \Gamma(E^*)$ defined by
$$(\nabla^*_X \rho)(s) = X(\rho(s)) - \rho(\nabla_X(s)).$$
Prove *directly* that this is a covariant derivative operator in E^*.

(iii) The **dual connection** on E^* is the connection $\mathcal{H}^*$ whose associated parallel transport system is the dual parallel transport system from part (i) and whose associated covariant derivative operator is the dual covariant derivative operator from part (ii). How does one define $\mathcal{H}^*$ explicitly?

Solution. Ad (i): Let $\gamma\colon [a,b] \to M$ be any smooth curve, $\lambda \in E^*_x$. We will denote by
$$\mathbb{P}^*_\gamma(\lambda)\colon [a,b] \to E^*$$
the[2] parallel lift on $\pi\colon E^* \to M$ of γ starting at λ. We start by showing that such a lift is unique. Indeed, pick a parallel local frame $e_i\colon U \to E$ along γ around x, i.e.
$$e_i \circ \gamma\colon [a,b] \to E$$
is parallel and the $e_i(\gamma(t))$ form a basis of $E_{\gamma(t)}$ for any $t \in [a,b]$. Denote by
$$\epsilon^j\colon U \to E^*$$
the dual local frame on U along γ associated to e_i. By definition of parallel lifts on the dual bundle we have
$$\mathbb{P}^*_\gamma(\lambda)(t)[e_i(\gamma(t))] = \mathbb{P}^*_\gamma(\lambda)(a)[e_i(\gamma(a))] = \lambda(e_i(\gamma(a)) \in \mathbb{R}.$$
Since $e_i(\gamma(t))$ forms a basis the computation above shows that the parallel $\mathbb{P}^*_\gamma(t)$ is uniquely determined by its starting point λ. This computation already takes business of the first axiom, i.e. that
$$\widehat{\mathbb{P}}^*_\gamma\colon E^*_{\gamma(a)} \to E^*_{\gamma(b)}$$
is an isomorphism:
$$\begin{aligned}\widehat{\mathbb{P}}^*_\gamma(\epsilon^j(\gamma(a)))[e_i(\gamma(b))] &= \mathbb{P}^*_\gamma(\epsilon^j(\gamma(a)))(b)[e_i(\gamma(b))] \\ &= \epsilon^j(\gamma(a))[e_i(\gamma(a)] \\ &= \delta_{ij}.\end{aligned}$$

[2] Apriori we do not know that such a lift of γ is unique, but we will show that this is the case under our given assumptions.

This proves that with respect to the basis $\epsilon^j \circ \gamma$ we get $\widehat{\mathbb{P}}^*_\gamma = \mathrm{id}$, in particular that $\widehat{\mathbb{P}}^*_\gamma$ is an isomorphism.

The second axiom follows from another easy computation. In order to keep the notation in check, we will write

$$\epsilon^j_\gamma(t) = \epsilon^j(\gamma(t)),\ e^\gamma_i(t) = e_i(\gamma(t))$$

and so on. Now let

$$h\colon [a_1, b_1] \to [a, b]$$

be a diffeomorphism with $h(a_1) = a$, $h(b_1) = b$. Observe that

$$\begin{aligned}
\mathbb{P}^*_{\gamma\circ h}(\epsilon^j_{\gamma\circ h}(a_1))(t)[e^{\gamma\circ h}_i(t)] &\overset{(1)}{=} \mathbb{P}^*_{\gamma\circ h}(\epsilon^j_{\gamma\circ h}(a_1))(a_1)[e^{\gamma\circ h}_i(a_1)] \\
&\overset{(2)}{=} \delta_{ij} \\
&\overset{(3)}{=} \mathbb{P}^*_\gamma(\epsilon^j_\gamma(a))(a)[e^\gamma_i(a)] \\
&\overset{(4)}{=} \mathbb{P}^*_\gamma(\epsilon^j_\gamma(a))(h(t))[e^\gamma_i(h(t))] \\
&\overset{(5)}{=} \mathbb{P}^*_\gamma(\epsilon^j_{\gamma\circ h}(a_1))(h(t))[e^{\gamma\circ h}_i(t)],
\end{aligned}$$

where in (1) and (4) we used the definition of parallel lifts on E^*, in (2) add (3) we used the same reasoning as in the first axiom part and (5) uses $h(a_1) = a$. This proves

$$\mathbb{P}^*_{\gamma\circ h}(\lambda)(t) = \mathbb{P}^*_\gamma(\lambda)(h(t)).[3]$$

For the last axiom we pick yet another smooth curve $\delta\colon [a, b] \to M$ such that

$$\delta(a) = \gamma(a),\ \delta'(a) = \gamma'(a).$$

We want to prove that

$$\frac{d}{dt}\Big|_{t=a} \mathbb{P}^*_\gamma(\lambda)(t) = \frac{d}{dt}\Big|_{t=a} \mathbb{P}^*_\delta(\lambda)(t).$$

In analogy to γ and e_i, we pick another parallel local frame

$$f_i\colon U \to E$$

along δ with

$$f_i(\delta(a)) = f_i(\gamma(a)) = e_i(\gamma(a)) = p_i,$$

for all i.

First of all observe that

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=a} \left(\mathbb{P}^*_\gamma(\lambda)(t)[e^\gamma_i(t)]\right) &= 0, \\
\frac{d}{dt}\Big|_{t=a} \left(\mathbb{P}^*_\delta(\lambda)(t)[e^\delta_i(t)]\right) &= 0.
\end{aligned}$$

[3] Actually the equation above only proves this for $\lambda = \epsilon^j_{\gamma\circ h}(a_1)$, but after a moment of thought one sees that the same proof works for linear combinations, hence for any λ.

By shrinking U, if necessary, we can work locally on some Euclidean space on which $\mathbb{P}^*_\gamma(\lambda)(t)[e_i^\gamma(t)]$ corresponds to the scalar product

$$\left\langle \mathbb{P}^*_\gamma(\lambda)(t), e_i^\gamma(t) \right\rangle \in \mathbb{R}.$$

But then we have

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=a} \left\langle \mathbb{P}^*_\gamma(\lambda)(t), e_i^\gamma(t) \right\rangle \\ &= \left\langle \frac{d}{dt}\Big|_{t=a} \mathbb{P}^*_\gamma(\lambda)(t), e_i(\gamma(a)) \right\rangle + \left\langle \lambda, \frac{d}{dt}\Big|_{t=a} e_i(\gamma(t)) \right\rangle. \end{aligned}$$

We have an analogous formula for $\mathbb{P}^*_\delta$. Using these two and the initial uniqueness applied to $e_i \circ \gamma$ and $f_i \circ \delta$, i.e.

$$\frac{d}{dt}\Big|_{t=a} e_i(\gamma(t)) = \frac{d}{dt}\Big|_{t=a} f_i(\delta(t)),$$

we see that

$$\begin{aligned} \left\langle \frac{d}{dt}\Big|_{t=a} \mathbb{P}^*_\gamma(\lambda)(t), \underbrace{e_i(\gamma(a))}_{=p_i} \right\rangle &= -\left\langle \lambda, \frac{d}{dt}\Big|_{t=a} e_i(\gamma(t)) \right\rangle \\ &= -\left\langle \lambda, \frac{d}{dt}\Big|_{t=a} f_i(\gamma(t)) \right\rangle \\ &= \left\langle \frac{d}{dt}\Big|_{t=a} \mathbb{P}^*_\delta(\lambda)(t), \underbrace{f_i(\delta(a))}_{=p_i} \right\rangle, \end{aligned}$$

thus finishing the proof of the last axiom.

Ad (ii): We simply have to check the four covariant derivative axioms, so let us pick two vector fields $X, Y \in \mathfrak{X}(M)$, two sections $\rho\colon M \to E^*$, $s\colon M \to E$ and a smooth function $f\colon M \to \mathbb{R}$. Then we have

$$\begin{aligned} \left(\nabla^*_{X+Y}\rho\right)(s) &= X(\rho(s)) + Y(\rho(s)) - \rho(\underbrace{\nabla_{X+Y}(s)}_{=\nabla_X(s)+\nabla_Y(s)}) \\ &= \left(\nabla^*_X(\rho)\right)(s) + \left(\nabla^*_Y(\rho)\right)(s). \end{aligned}$$

Here we used the fact that ∇ is a covariant derivative on E. In an analogous fashion the linearity in the second entry of ∇^* follows from the one of ∇. Similarly, the $C^\infty(M)$-linearity in the first entry of ∇^* follows from the one of ∇. Let us do the computation for the Leibniz-rule, i.e. verify the last axiom:

$$\begin{aligned} \left(\nabla^*_X(f\cdot\rho)\right)(s) &= X((f\cdot\rho)(s)) - f\cdot\rho\left(\nabla_X(s)\right) \\ &= X(f\cdot\rho(s)) - f\cdot\rho\left(\nabla_X(s)\right) \\ &= \rho(s)\cdot X(f) + \underbrace{f\cdot X(\rho(s)) - f\cdot\rho\left(\nabla_X(s)\right)}_{=f\cdot\left(\nabla^*_X(\rho)\right)(s)} \\ &= \left(X(f)\cdot\rho\right)(s) + f\cdot\left(\nabla^*_X(\rho)\right)(s). \end{aligned}$$

Ad (iii): We want to define a distribution $\mathcal{H}^*$ on E^* that defines a connection and whose parallel transport system and covariant derivative coincide with those defined in (i) and (ii). Let again $\mathbb{P}^*$ denote the dual parallel transport system defined in (i) on E^* and consider its associated connection $\mathcal{H}^{\mathbb{P}^*}$ on E^* which is given by differentiating parallel lifts. More precisely

$$\mathcal{H}^{\mathbb{P}^*}(\lambda) := \left\{ \frac{d}{dt}\Big|_{t=0} \mathbb{P}^*_\gamma(\lambda)(t) \Big| \gamma\colon [0,1] \to M \text{ smooth} \right\}.$$

This is a connection, whose associated parallel transport system (defined via horizontal lifts of the connection) is simply $\mathbb{P}^*$ again, see second part of Theorem 30.1. Therefore $\mathcal{H}^* := \mathcal{H}^{\mathbb{P}^*}$ is a connection on E^* whose parallel transport system is given by $\mathbb{P}^*$.

Now we try to write down an explicit formula for $\mathcal{H}^*$ in terms of $\mathcal{H}$. For this we will need the direct sum bundle

$$E^* \oplus E = \coprod_{x \in M} E^*_x \oplus E_x \to M.$$

On this bundle we can define an evaluation map

$$e\colon E^* \oplus E \mapsto \mathbb{R},\ e(\lambda, p) = \lambda(p).$$

This map is obviously smooth, thus we can differentiate it to give us a map

$$De(\lambda, p)\colon T_\lambda E^* \oplus T_p E \to \mathbb{R},$$

where λ and p both lie in their respective fiber over the common point $x \in M$. The claim now is that

$$\mathcal{H}^*_\lambda = \bigcap_{p \in \pi^{-1}(x)} \ker De(\lambda, p)\,[\,\cdot\,, \mathcal{H}_p]\,.$$

First we show that the LHS is contained in the RHS, i.e. we pick $\rho\colon [0,1] \to E^*$ a parallel curve along $\gamma := \pi^* \circ \rho$ with $\rho(0) = \lambda$ and $c\colon [0,1] \to E$ any other parallel curve, also along γ satisfying $c(0) = p$ for some p over x. Observe that

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0} \rho(t)[c(t)] \\ &= \frac{d}{dt}\Big|_{t=0} e(\rho(t), c(t)) \\ &= De(\lambda, p)\Big[\underbrace{\frac{d}{dt}\Big|_{t=0}\rho(t)}_{\in \mathcal{H}^*_\lambda}, \underbrace{\frac{d}{dt}\Big|_{t=0} c(t)}_{\in \mathcal{H}_p}\Big], \end{aligned}$$

which proves the desired inclusion since p was an arbitrary element over x.

For the other inclusion we pick some $v \in T_\lambda E^*$ such that $De(\lambda, p)[v, \mathcal{H}_p] = 0$ for all p over x. Now we view v as the time-0 derivative of some curve $\rho\colon [0,1] \to E^*$ and show that this ρ is necessarily parallel which would then imply that $v \in \mathcal{H}^*_\lambda$. Indeed, picking any parallel lift $c\colon [0,1] \to E$ of $\gamma := \pi^* \circ \gamma$ starting at some p over x we can read the above computation backwards to obtain

$$\rho(t)[c(t)] = \text{const}.$$

Again, this holds for any parallel lift of γ since p is allowed to be any lift over x.

(♣) Problem O.4. Let $\pi_i \colon E_i \to M_i$ be two vector bundles, and let $\varphi \colon M_1 \to M_2$ denote a smooth map. Suppose $\Phi \colon E_1 \to E_2$ is a smooth map such that the following diagram commutes

$$\begin{tikzcd} E_1 \arrow[r, "\Phi"] \arrow[d, "\pi_1"'] & E_2 \arrow[d, "\pi_2"] \\ M_1 \arrow[r, "\varphi"'] & M_2 \end{tikzcd}$$

Note we are *not* assuming that Φ is linear on the fibres, and hence Φ need not be a vector bundle morphism along φ. For each $x \in M_1$, Φ defines a smooth map

$$\Phi_x := \Phi|_{E_1|_x} \colon E_1|_x \to E_2|_{\varphi(x)}.$$

This is a map between two vector spaces, so for any $p \in E_1|_x$ we can take its derivative

$$D\Phi_x(p) \colon T_p E_1|_x \to T_{\Phi(p)} E_2|_{\varphi(x)}.$$

Composing with the $\mathcal{J}$ maps from Problem B.3, we get a linear map from

$$\widehat{D}\Phi_x(p) := \mathcal{J}^{-1}_{\Phi(p)} \circ D\Phi_x(p) \circ \mathcal{J}_p \colon E_1|_x \to E_2|_{\varphi(x)}.$$

Consider the vector bundle $\tilde{\pi} \colon \mathrm{Hom}(E_1, \varphi^\star E_2) \to M_1$ over M_1. The fibre of this bundle over $x \in M_1$ is $\mathrm{L}(E_1|_x, E_2|_{\varphi(x)})$. Since Φ is smooth, the map $p \mapsto \widehat{D}\Phi_x(p)$ defines a smooth map $D^{\mathrm{fibre}}\Phi \colon E_1 \to \mathrm{Hom}(E_1, \varphi^\star E_2)$ which we call the **fibrewise derivative** of Φ

$$\begin{tikzcd} E_1 \arrow[rr, "D^{\mathrm{fibre}}\Phi"] \arrow[dr, "\pi_1"'] & & \mathrm{Hom}(E_1, \varphi^\star E_2) \arrow[dl, "\tilde{\pi}"] \\ & M_1 & \end{tikzcd}$$

Note again that $D^{\mathrm{fibre}}\Phi$ is *not* necessarily linear on the fibres (i.e. $\widehat{D}\Phi_x(p)$ does not have to depend linearly on p), and thus $D^{\mathrm{fibre}}\Phi$ is *not* necessarily a vector bundle morphism along φ.

(i) Show that the normal derivative $D\Phi \colon TE_1 \to TE_2$ of Φ restricts to define a map $D\Phi|_{VE_1} \colon VE_1 \to VE_2$. Prove that if $p, q \in E_1|_x$ then

$$D^{\mathrm{fibre}}\Phi(p)[q] = \mathrm{pr}_2^{E_2} \circ D\Phi|_{VE_1} \circ \mathcal{J}_p(q),$$

where $\mathrm{pr}_2^{E_2} \colon VE_2 \to E_2$ is the "projection onto the second factor" map (see (30.1) from Lecture 30 or Problem I.5).

(ii) Now take $M_1 = M_2 = M$ and φ to be the identity. Let $E_1 = TM$ denote the tangent bundle and let $E_2 = M \times \mathbb{R}$ denote the trivial bundle. Then a fibre preserving map $\Phi \colon E_1 \to E_2$ can be identified with a smooth function $f \colon TM \to \mathbb{R}$. Prove that the definition of $D^{\mathrm{fibre}} f$ given above is consistent with the fibrewise derivative $D^{\mathrm{fibre}} f \colon TM \to T^*M$ given in Problem C.3.

(iii) Now return to the general setup, and assume that both E_1 and E_2 are endowed with connections $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Let $\kappa_i \colon TE_i \to E_i$ denote the connection maps of $\mathcal{H}_i$. We define the **parallel derivative** of Φ to be the map

$$D^{\text{parallel}}\Phi \colon E_1 \to \operatorname{Hom}(TM_1, \varphi^\star E_2)$$

by

$$D^{\text{parallel}}\Phi(p)[v] := \kappa_2 \circ D\Phi(p)[\overline{v}], \qquad p \in E_1,\ v \in TM_1,$$

where $\overline{v}$ is the horizontal lift of v at p with respect to $\mathcal{H}_1$ (see Definition 28.7). Prove that for $x \in M_1$, $p \in E_1|_x$ and $\zeta \in T_pE_1$ that the following formula holds:

$$\kappa_2\big(D\Phi(p)[\zeta]\big) = D^{\text{fibre}}\Phi(p)[\kappa_1(\zeta)] + D^{\text{parallel}}\Phi(p)[D\pi_1(p)[\zeta].$$

(iv) Conclude that $D\Phi$ is entirely determined by $D\varphi$, $D^{\text{fibre}}\Phi$ and $D^{\text{parallel}}\Phi$. That is, under the vector bundle isomorphism $(D\pi_i, \kappa_i) \colon TE_i \to TM_i \oplus E_i$ along π_i given by Lemma 31.3, $D\Phi$ takes matrix form:

$$D\Phi = \begin{pmatrix} D\varphi & 0 \\ D^{\text{parallel}}\Phi & D^{\text{fibre}}\Phi \end{pmatrix}$$

This formula is often very useful in computations.

Solution. (i) First let us check that $D\Phi|_{VE_1}$ defines a map from VE_1 to VE_2. To do that it is sufficient to check that, given $z \in VE_1 \cap T_pE_1$, it holds that $D\Phi(p)[z] \in VE_2$. However, this is very easy to verify since

$$\begin{aligned} D\pi_2(\Phi(p))[D\Phi(p)[z]] &= D(\pi_2 \circ \Phi)(p)[z] = D(\varphi \circ \pi_1)(p)[z] \\ &= D\varphi(\pi_1(p))[D\pi_1(p)[z]] = 0, \end{aligned}$$

where we have used that $\pi_2 \circ \Phi = \varphi \circ \pi_1$. Therefore we have $D\Phi|_{VE_1}(p)[z] = D\Phi_x(p)[z]$ for every $z \in VE_1$.

Now recall that $\mathcal{J} \colon \pi^*E \to VE$ is a vector bundle isomorphism (see Problem I.5 and Lecture 30) and that $\operatorname{pr}_2^{E_2} \colon VE \to E$ is exactly defined as $\operatorname{pr}_2^{E_2}(\mathcal{J}_p(q)) = q$.

Putting together all these considerations, we easily obtain the thesis, that is

$$\mathcal{J}_{\Phi(p)}^{-1} \circ D\Phi_x(p) \circ \mathcal{J}_p = \operatorname{pr}_2^{E_2} \circ D\Phi|_{VE_1} \circ \mathcal{J}_p.$$

(ii) Let us consider the local coordinates (x^i, v^i) defined in Problem C.3. Then it is sufficient to check that the two definitions of $D^{\text{fibre}}f(x,v)$ coincides on $\frac{\partial}{\partial x^i}|_x$ for all $(x,v) \in TM$ and $i = 1, \dots, n$.

Following the definition of $D^{\text{fibre}}f$ given in Problem C.3, we have that

$$D^{\text{fibre}}f(x,v)\Big[\frac{\partial}{\partial x^i}\Big|_x\Big] = \sum_{i=1}^n \frac{\partial}{\partial v^i}\Big|_{(x,v)}(f).$$

On the other hand, let us explicit the definition of $D^{\text{fibre}} f$ given in this exercise. First recall that we can write Df as

$$Df(x,v) = \sum_{i=1}^{n} \frac{\partial}{\partial v^i}\Big|_{(x,v)}(f) dv^i|_{(x,v)} + \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\Big|_{(x,v)}(f) dx^i|_{(x,v)}$$

and consequently

$$Df|_{VTM}(x,v) = \sum_{i=1}^{n} \frac{\partial}{\partial v^i}\Big|_{(x,v)}(f) dv^i|_{(x,v)},$$

since $VTM = \bigcup_{(x,v)\in TM} \operatorname{span}\Big\{ \frac{\partial}{\partial v^i}\Big|_{(x,v)} \Big\}$.

Moreover notice that $\mathcal{J}_v\Big(\frac{\partial}{\partial x^i}\Big|_x\Big) = \frac{\partial}{\partial v^i}\Big|_{(x,v)}$ and that $\mathrm{pr}_2^{M\times\mathbb{R}}$ is just the projection on the second factor.

Therefore, following the definition given in this exercise, we have

$$\begin{aligned} D^{\text{fibre}} f(x,v)\Big[\frac{\partial}{\partial x^i}\Big|_x\Big] &= \mathrm{pr}_2^{E_2} \circ Df|_{VE_1} \circ \mathcal{J}_p\Big(\frac{\partial}{\partial x^i}\Big|_x\Big) = \mathrm{pr}_2^{E_2} \circ Df|_{VE_1}\Big[\frac{\partial}{\partial v^i}\Big|_{(x,v)}\Big] \\ &= \mathrm{pr}_2^{E_2}\Big(\sum_{i=1}^{n} \frac{\partial}{\partial v^i}\Big|_{(x,v)}(f)\Big) = \sum_{i=1}^{n} \frac{\partial}{\partial v^i}\Big|_{(x,v)}(f), \end{aligned}$$

which concludes the proof, since this coincides with the expression with the other definition of $D^{\text{fibre}} f$.

(iii) First notice that the horizontal lift of $D\pi_1(p)[\zeta]$ is exactly $\zeta^{\mathtt{H}}$ by definition. Therefore

$$D^{\text{parallel}}\Phi(p)[D\pi_1(p)[\zeta]] = \kappa_2(D\Phi(p)[\zeta^{\mathtt{H}}]).$$

On the other hand, we have that

$$\begin{aligned} D^{\text{fibre}}\Phi(p)[\kappa_1(\zeta)] &= \mathrm{pr}_2^{E_2} \circ D\Phi|_{VE_1} \circ \mathcal{J}_p(\kappa_1(\zeta)) = \mathrm{pr}_2^{E_2} \circ D\Phi(p)[\zeta^{\mathtt{V}}] \\ &= \kappa_2(D\Phi(p)[\zeta^{\mathtt{V}}]). \end{aligned}$$

And this concludes the proof using the linearity of $D\Phi(p)$ and κ_2, together with the fact that $\zeta = \zeta^{\mathtt{H}} + \zeta^{\mathtt{V}}$.

(iv) First notice that

$$D\pi_2(D\Phi(p)[\zeta]) = D(\pi_2 \circ \Phi)(p)[\zeta] = D(\varphi \circ \pi_1)(p)[\zeta] = D\varphi[D\pi_1(p)[\zeta]].$$

Moreover, thanks to the previous points, we have that

$$\kappa_2\big(D\Phi(p)[\zeta]\big) = D^{\text{parallel}}\Phi(p)[D\pi_1(p)[\zeta]] + D^{\text{fibre}}\Phi(p)[\kappa_1(\zeta)].$$

This two observations are patently sufficient to conclude the proof.

Problem Sheet P

Problem P.1. Let $\pi_i \colon E_i \to M$ be vector bundles with connections ∇^i for $i = 1, 2$.

(i) Prove that there is a unique connection on $E_1 \otimes E_2$ which on decomposable sections $s_1 \otimes s_2$ takes the form

$$\nabla^{\otimes}_X(s_1 \otimes s_2) := \nabla^1_X(s_1) \otimes s_2 + s_1 \otimes \nabla^2_X(s_2).$$

(ii) Prove that

$$\nabla^{\mathrm{Hom}}_X(\Phi)(s) := \nabla^2_X(\Phi(s)) - \Phi(\nabla^1_X(s))$$

is a connection on $\mathrm{Hom}(E_1, E_2)$. *Remark:* The connections in part (i) and part (ii) are consistent with the connection on the dual bundle from Problem (ii) under the isomorphism $\mathrm{Hom}(E_1, E_2) \cong E_1^* \otimes E_2$ from Corollary 15.13.

Problem P.2. Suppose ∇ is a connection on the tangent bundle $\pi \colon TM \to M$ of a manifold M. Show that for each $X \in \mathfrak{X}(M)$ there is a unique tensor derivation $\tilde{\nabla}_X \colon \mathcal{T}(M) \to \mathcal{T}(M)$ (cf. Definition 18.14) such that $\tilde{\nabla}_X(Y) = \nabla_X(Y)$ for all $Y \in \mathfrak{X}(M)$.

(♣) Problem P.3. Let $\pi \colon E \to M$ be a vector bundle over a connected manifold M, and let $\mathcal{H}$ denote a connection on E Let $\rho \colon \widetilde{M} \to M$ denote the universal covering of M. Prove that ∇ is flat if and only if $\rho^* E \to \widetilde{M}$ is the trivial bundle over $\widetilde{M}$ and the pullback connection $\rho^\star \mathcal{H}$ is the trivial connection.

(♣) Problem P.4. Consider TS^n equipped with the connection ∇ from Problem N.3.

(i) Prove that $\mathrm{Hol}^\nabla = \mathrm{SO}(n)$ (in the sense of Corollary 32.12).

(ii) Compute the curvature tensor R^∇.

Problem P.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$.

1. Suppose $\beta \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a bilinear map. Prove there exists a unique connection ∇^β on $TG \to G$ which satisfies the following condition: if $v, w \in \mathfrak{g}$ and X_v, X_w denote the corresponding left-invariant vector fields then

$$\nabla^\beta_{X_v}(X_w) = X_{\beta(v,w)}.$$

2. Prove that this connection is left-invariant in the sense that

$$(l_a)_\star(\nabla^\beta_X(Y)) = \nabla^\beta_{(l_s)_\star X}((l_a)_\star(Y)), \qquad \forall\, X, Y \in \mathfrak{X}(G),\ \forall\, a \in G.$$

Deduce that the parallel transport determined by this connection is left-invariant in the sense that if c is a parallel section along a curve γ then $Dl_a(\gamma) \circ c$ is a parallel section along $l_a \circ \gamma$.

3. Prove moreover that any such left-invariant connection ∇ determines such a bilinear map β via

$$\beta(v, w) \coloneqq \nabla_{X_v}(X_w)(e),$$

and hence that there is a bijective correspondence between bilinear maps $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and connections on TG.

Solutions to Problem Sheet P

Problem P.1. Let $\pi_i\colon E_i \to M$ be vector bundles with connections ∇^i for $i = 1, 2$.

(i) Prove that
$$\nabla_X(s_1 \otimes s_2) \coloneqq \nabla^1_X(s_1) \otimes s_2 + s_1 \otimes \nabla^2_X(s_2)$$
is a connection on $E_1 \otimes E_2$.

(ii) Prove that
$$\nabla_X(\Phi)(s) := \nabla^2_X(\Phi(s)) - \Phi(\nabla^1_X(s))$$
is a connection on $\mathrm{Hom}(E_1, E_2)$. *Remark:* The connections in part (i) and part (ii) are consistent with the connection on the dual bundle from Problem (ii) under the isomorphism $\mathrm{Hom}(E_1, E_2) \cong E_1^* \otimes E_2$ from Corollary 15.13.

Solution. Ad (i): Linearity in the second entry is given by definition as we define the connection $\nabla^\otimes$ on decomposable elements and then extend linearly. The linearity in the first variable follows from the linearity of ∇^i and properties of tensors: For two vector fields $X, Y \mathfrak{X}(M)$ we have

$$\begin{aligned}\nabla^\otimes_{X+Y}(s_1 \otimes s_2) &= \big(\nabla^1_X(s_1) + \nabla^1_Y(s_2)\big) \otimes s_2 + s_1 \otimes \big(\nabla^2_X(s_2) + \nabla^2_Y(s_2)\big)\\ &= \underbrace{\nabla^1_X(s_1) \otimes s_2 + s_1 \otimes \nabla^2_X(s_2)}_{=\nabla^\otimes_X(s_1\otimes s_2)} + \underbrace{\nabla^1_Y(s_1) \otimes s_2 + s_1 \otimes \nabla^2_Y(s_2)}_{=\nabla^\otimes_Y(s_1\otimes s_2)}.\end{aligned}$$

A similar argument proves that for any $f \in C^\infty(M)$ one has
$$\nabla^\otimes_{f\,X}(s_1 \otimes s_2) = f\nabla^\otimes_X(s_1 \otimes s_2),$$
so we are left to prove the Leibniz rule:

$$\begin{aligned}\nabla^\otimes_X\,(f \cdot (s_1 \otimes s_2)) &= \nabla^\otimes_X\,((f\, s_1) \otimes s_2)\\ &= \big(X(f) \cdot (s_1 \otimes s_2) + (f \cdot \nabla^1_X s_1) \otimes s_2\big) + (f\, s_1) \otimes \nabla^2_X s_2\\ &= X(f) \cdot (s_1 \otimes s_2) + f \cdot \Big(\underbrace{\nabla^1_X s_1 \otimes s_2 + s_1 \otimes \nabla^2_X s_2}_{=\nabla^\otimes_X(s_1\otimes s_2)}\Big).\end{aligned}$$

This concludes the proof of part (i).

Ad part (ii): The linearity in both entries, i.e. for $X + Y \in \mathfrak{X}(M)$ and $\Phi + \Psi \in \mathrm{Hom}(E_1, E_2)$ immediately follows from the linearity of both ∇^1 and ∇^2. The fact that $\nabla^{\mathrm{Hom}}_{f\,X}(\Phi)(s) = f \cdot \nabla^{\mathrm{Hom}}_X(\Phi)(s)$ follows again from properties of ∇^1, ∇^2 and $\Phi(f \cdot s) = f \cdot \Phi(s)$ for any section s on E_1. For the Leibniz bit we conclude with a straightforward computation:

$$\begin{aligned}\nabla^{\mathrm{Hom}}_X(f \cdot \Phi)(s) &= \nabla^2_X(f \cdot \Phi)(s) - f \cdot \Phi(\nabla^1_X(s))\\ &= X(f) \cdot \Phi(s) + \underbrace{f \cdot \nabla^2_X\Phi(s) - f \cdot \Phi(\nabla^1_X(s))}_{=f\cdot\nabla^{\mathrm{Hom}}_X\Phi(s)}.\end{aligned}$$

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

Problem P.2. Suppose ∇ is a connection on the tangent bundle $\pi\colon TM \to M$ of a manifold M. Show that for each $X \in \mathfrak{X}(M)$ there is a unique tensor derivation $\tilde{\nabla}_X\colon \mathcal{T}(M) \to \mathcal{T}(M)$ (cf. Definition 18.14) such that $\tilde{\nabla}_X(Y) = \nabla_X(Y)$ for all $Y \in \mathfrak{X}(M)$.

Solution. The result is a straightforward consequence of Proposition 18.17. Indeed, for each $X \in \mathfrak{X}(M)$, ∇_X is a sheaf morphism defined on smooth functions and vector fields which satisfies

$$\nabla_X(fg) = X(fg) = X(f)g + fX(g), \quad \nabla_X(fY) = X(f)Y + f\nabla_X Y$$

for all $f, g \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$. Therefore, applying the proposition, ∇ extends uniquely to a tensor derivation $\tilde{\nabla} : \mathcal{T}(M) \to \mathcal{T}(M)$.

Notice that, on a 1-form ω, $\tilde{\nabla}$ is defined as

$$\tilde{\nabla}_X \omega(Y) = X(\omega(Y)) - \omega(\nabla_X Y),$$

while, on a generic tensor $A \in \mathcal{T}^{r,s}(M)$, this tensor derivation is equal to

$$\begin{aligned}\tilde{\nabla}_X A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) = X(A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s))\\ - \sum_{i=1}^{r} A(\omega_1, \ldots, \tilde{\nabla}_X \omega_i, \ldots, \omega_r, X_1, \ldots, X_s)\\ - \sum_{i=1}^{s} A(\omega_1, \ldots, \omega_r, X_1, \ldots, \tilde{\nabla}_X X_i, \ldots, X_s).\end{aligned}$$

(♣) Problem P.3. Let $\pi\colon E \to M$ be a vector bundle over a connected manifold M, and let $\mathcal{H}$ denote a connection on E Let $\rho\colon \widetilde{M} \to M$ denote the universal covering of M. Prove that ∇ is flat if and only if $\rho^* E \to \widetilde{M}$ is the trivial bundle over $\widetilde{M}$ and the pullback connection $\rho^\star \mathcal{H}$ is the trivial connection.

Solution. We start the proof by first assuming that ∇ is flat. This means that the corresponding distribution $\mathcal{H}$ is integrable, i.e. for two vector fields $V, W\colon E \to \mathcal{H}$ we have

$$[V, W] \in \mathcal{H}.$$

First of all we claim that proving that $\rho^* \mathcal{H}$ is flat suffices[1]. Indeed, this is sufficient due to Corollary 33.5 as a universal cover is connected and simply connected by definition. For the 'if'-statement we are only left to show that the pullback distribution $\rho^* \mathcal{H}$ is integrable. Recall that

$$\rho^* \mathcal{H}_{(\tilde{x},p)} \left\{ (\tilde{v}, \zeta) \in T_{\tilde{x}} \widetilde{M} \times \mathcal{H}_p \,\middle|\, D\rho(\tilde{x})[\tilde{v}] = D\pi(p)[\zeta] \right\}.$$

Now pick two additional vector fields $\widetilde{X}, \widetilde{Y} \in \mathfrak{X}(\widetilde{M})$ such that

$$(\widetilde{X}, V),\ (\widetilde{Y}, W) \in \rho^* \mathcal{H},$$

or equivalently such that

$$D\rho(\widetilde{X}) = D\pi(V),\ D\rho(\widetilde{Y}) = D\pi(W).$$

[1] Here we actually mean the connection $\tilde{\nabla}$ associated to $\rho^* \mathcal{H}$.

But then we easily see that

$$\begin{aligned} D\rho\left[\widetilde{X},\widetilde{Y}\right] &= \left[D\rho(\widetilde{X}),D\rho(\widetilde{Y})\right] \\ &= [D\pi(V),D\pi(W)] \\ &= D\pi\underbrace{[V,W]}_{\in\mathcal{H}}, \end{aligned}$$

which proves that $[\widetilde{X},\widetilde{Y}]\in\rho^*\mathcal{H}$ and hence that the pullback is integrable.

Conversely, we assume that $\rho^*E\to\widetilde{M}$ and ρ^*H are trivial, i.e.

$$\rho^*E\cong\widetilde{M}\times\mathbb{R}^k,\ \rho^*\mathcal{H}_{(\tilde{x},p)}\cong T_{\tilde{x}}\widetilde{M}$$

with $\dim E=k$, where the second diffeomorphism is given by

$$D\mathrm{pr}_1(\tilde{x},p)[\tilde{v},\zeta]=\tilde{v}.$$

Using the description of the pullback distribution we then deduce that for $(\tilde{v},\zeta)$, $(\tilde{v},\eta)\in\rho^*\mathcal{H}_{(\tilde{x},p)}$ we must have $\zeta=\eta$. Adopting the notation from the first part of the proof then leads us to the conclusion

$$[V,W]=\overline{[V,W]}$$

for any two vector fields V,W on E taking values in $\mathcal{H}$. This proves that $\mathcal{H}$ is integrable, hence flat.

(♣) PROBLEM P.4. Consider TS^n equipped with the connection ∇ from Problem N.3.

(i) Prove that $\mathrm{Hol}^\nabla=\mathrm{SO}(n)$ (in the sense of Corollary 32.12).

(ii) Compute the curvature tensor R^∇.

SOLUTION. (i). Since for every $x\in S^n$ the groups $\mathrm{Hol}^\nabla(x)$ are all isomorphic, we will fix $x=x_N$ to be the north pole. Recall that, on the S^n, for any curve $\gamma:[a,b]\to S^n$ and a section of $\gamma^\star(TS^n)$, $V:[a,b]\to TS^n$, $V(t)=(\gamma(t),v(t))$, The condition of parallelism $\nabla_T V(t)=0$ is defined by the ODE

$$v'(t)+\langle v(t),\gamma'(t)\rangle\gamma(t)=0, \tag{P.1}$$

and the map $\mathbb{P}_\gamma:T_{\gamma(a)}S^n\to T_{\gamma(b)}S^n$ is given by $\mathbb{P}_\gamma(v(a))=v(b)$.

Let us first show that every element in Hol^∇ defines an element of $SO(n)$. For any continuous, piecewise smooth $\gamma:[a,b]\to S^n$ with $\gamma(a)=x_N=\gamma(b)$, we note that, because of (P.1) and since $v(t)\perp\gamma(t)$, we have (at all the points of differentiability of v)

$$\begin{aligned} \frac{d}{dt}|v(t)|^2=2\langle v(t),v'(t)\rangle &= \langle v(t),-\langle v(t),\gamma'(t)\rangle,\gamma(t)\rangle \\ &= -\langle v(t),\gamma'(t)\rangle\langle v(t),\gamma(t)\rangle \\ &= 0, \end{aligned}$$

hence the map $\mathbb{P}_\gamma : T_{x_N}S^n \to T_{x_N}S^n$ is length-preserving, and so necessarily an element of $O(n)$. To show that it is also orientation preserving, pick any basis $e_1, \dots, e_n$ of $T_{x_N}S^n$ which is positively oriented, a condition that can be expressed as $\det(x_N, e_1, \dots, e_n) > 0$. Since the parallel transport is an isomorphism, we have that, for every t, the vectors $e_i(t) = \mathbb{P}_\gamma(e_i)(t)$ form a basis of $T_{\gamma(t)}S^n$, hence, for every t, $\det(\gamma(t), e_1(t), \dots, e_n(t))$ is always nonzero. Since such function is continuous in $[a, b]$ and is it positive at $t = a$, it must be positive everywhere. Hence $\mathbb{P}_\gamma$ maps a (and so every) positive-oriented basis into a positive oriented basis and is consequently orientation preserving.

Vice versa, let us show that every element $A \in SO(n)$ can be represented as the parallel transport $\mathbb{P}_\gamma$ for some loop $\gamma : [a, b] \to S^n$. Let us suppose first that $n = 2$, so that A can be represented by a rotation matrix

$$A = R(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

of some angle $\beta \in [0, 2\pi)$. We claim that the loop γ can be obtained as the concatenation of the following three paths:

- $\gamma_1 : [0, \pi/2] \to S^2$, $\gamma_1(t) = (\sin t, 0, \cos t)$, namely the shortest arc of great circle from x_N to $(1, 0, 0)$
- $\gamma_2(t) : [0, 1] \to S^2$, $\gamma_2(t) = (\cos(\alpha t), \sin(\alpha t), 0)$, namely an arc of the equator from $(1, 0, 0)$ to $(\cos\alpha, \sin\alpha, 0)$;
- $\gamma_3 : [0, \pi/2] \to S^2$, $\gamma_3(t) = (\cos\alpha\cos t, \cos\alpha\cos t, \sin t)$, namely the arc of shortest great circle from $(\cos\alpha, \sin\alpha, 0)$ to x_N.

Where α will be chosen below.

From (P.1), the parallel transport of a vector $u_0 \in T_{x_N}S^2$ along γ_1 is defined by the value $u(\pi/2)$ of the solution to the problem

$$\begin{cases} \dot{u}^1(t) = \big(-u^1(t)\cos t + u^3(t)\sin t\big)\sin t, \\ \dot{u}^2(t) = 0, \\ \dot{u}^3(t) = \big(-u^1(t)\cos t + u^3(t)\sin t\big)\cos t, \\ u(0) = u_0, \end{cases}$$

whose solution can be computed with the help of change of variables

$$\begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} u^1(t) \\ u^3(t) \end{pmatrix}$$

and is given by

$$\begin{pmatrix} u^1(t) \\ u^2(t) \\ u^3(t) \end{pmatrix} = \begin{pmatrix} u_0^1\cos t + u_0^3\sin t \\ u_0^2 \\ -u_0^1\sin t + u_0^3\cos t \end{pmatrix}, \quad \text{consequently} \quad \mathbb{P}_{\gamma_1}(u_0) = \begin{pmatrix} u_0^3 \\ u_0^2 \\ -u_0^1 \end{pmatrix}.$$

The parallel transport of a vector $v_0 \in T_{(1,0,0)}S^2$ along γ_2 is defined by the value $v(1)$ of the solution to the problem

$$\begin{cases} \dot{v}^1(t) = \alpha\left(-v^1(t)\sin(\alpha t) + v^2\cos(\alpha t)\right)\cos(\alpha t), \\ \dot{v}^2(t) = \alpha\left(-v^1(t)\sin(\alpha t) + v^2\cos(\alpha t)\right)\sin(\alpha t), \\ \dot{v}^3(t) = 0, \\ v(0) = v_0, \end{cases}$$

whose solution similarly as before is computed to be

$$\begin{pmatrix} v^1(t) \\ v^2(t) \\ v^3(t) \end{pmatrix} = \begin{pmatrix} v_0^1\cos(\alpha t) - v_0^2\sin(\alpha t) \\ v_0^1\sin(\alpha t) + v_0^2\cos(\alpha t) \\ v_0^3 \end{pmatrix}, \quad \text{consequently} \quad \mathbb{P}_{\gamma_2}(v_0) = \begin{pmatrix} v_0^1\cos(\alpha) - v_0^2\sin\alpha \\ v_0^1\sin(\alpha) + v_0^2\cos\alpha \\ v_0^3 \end{pmatrix}.$$

Finally, the parallel transport of a vector $w_0 \in T_{(\cos\alpha,\sin\alpha,0)}S^2$ along γ_3 can be computed similarly to be

$$\mathbb{P}_{\gamma_3}(w_0) = \begin{pmatrix} -w_0^3 \\ -w_0^2\sin\alpha + w_0^2\cos\alpha \\ w_0^1\cos\alpha + w_0^2\sin\alpha \end{pmatrix}.$$

We conclude that, for $u_0 \in T_{x_N}S^2$ there holds

$$\mathbb{P}_\gamma(u_0) = \mathbb{P}_{\gamma_3}[\mathbb{P}_{\gamma_2}[\mathbb{P}_{\gamma_1}[u_0]]] = \begin{pmatrix} -u_0^1 \\ u_0^2\cos\alpha - u_0^3\sin\alpha \\ u_0^2\sin\alpha + u_0^3\cos\alpha \end{pmatrix}$$

which, being $u_0^1 = 0$, is precisely a counter-clockwise rotation of angle α. We may then choose $\beta = \alpha$ and conclude that Hol^∇, for S^2, is in fact equal to $SO(2)$.

Let us now come to the case of general n. From linear algebra, each $A \in SO(n)$ can be written as composition of m ($m \leq [n/2]$) number of rotations over pairwise orthogonal planes, i.e. $A = R_{\Pi_1}(\beta_1) \circ \cdots \circ R_{\Pi_m}(\beta_m)$, where Π_j's are the planes and β_j's are the respective angles. For each of these rotations, we may produce a loop γ_j at x_N, replacing the x-z plane with Π_j in the construction above, and finally setting $\gamma = \gamma_1 * \cdots * \gamma_m$. The endomorphism given by $\mathbb{P}_\gamma$ will coincide with the transformation A, and consequently, Hol^∇ coindices with $SO(n)$.

(ii). Recall from Problems N.3 and O.2 that:

- a vector field $X \in \mathfrak{X}(S^n)$ can be seen as a smooth function $X : S^n \to \mathbb{R}^{n+1}$ so that $\langle X(x), x\rangle = 0$ for every $x \in S^n$ and its horizontal lift is the vector field $\overline{X} \in \mathfrak{X}(T(TS^n))$

$$\overline{X}(x,v) = (X(x), -\langle X(x), v\rangle x);$$

- an element of $\mathfrak{X}(T(TS^n))$ can be represented by a smooth function $(u,w) : TS^n \to \mathbb{R}^{2(n+1)}$ so that, for any $(x,v) \in TS^n$,

$$\langle u(x,v), x\rangle = 0 \quad \text{and} \quad \langle u(x,v), v\rangle\rangle + \langle w(x,v), x\rangle = 0.$$

For any such $(u,v) \in \mathfrak{X}(T(TS^n))$, the differential of u is

$$\begin{aligned} Du(x,v): T_{(x,v)}TS^n &\to T_{u(x,v)}\mathbb{R}^{n+1} \simeq \mathbb{R}^{n+1}, \\ (U,W) &\mapsto D(u,v)(x,v)[(U,W)] = D_x u(x,v)[U] + D_v u(x,v)[W], \end{aligned}$$

where $D_x u$ and $D_v u$ denote partial differentiation with respect to x and v respectively, and an analogous expression holds for $Dw(x,v)$. In the case of a horizontal vector field, $u = X$ and we can compute

$$\begin{aligned} D\big(-\langle X(x), v\rangle x\big)[(U,W)] &= D_x\big(-\langle X(x), v\rangle x\big)[U] + D_v\big(-\langle X(x), v\rangle x\big)[W] \\ &= -\langle DX(x)[U], v\rangle x - \langle X(x), v\rangle U - \langle X(x), W\rangle x, \end{aligned}$$

so if $\overline{X}_i(x,v) = (X_i(x), -\langle X_i(x), v\rangle x)$ $(i = 1,2)$ are two horizontal vector fields, their Lie bracket will be:

$$\begin{aligned} &[\overline{X}_1, \overline{X}_2](x,v) \\ &= \Big(DX_1(x)[X_2] - DX_2(x)[X_1](x), \\ &\quad - \langle DX_1(x)[X_2], v\rangle x + \langle DX_2(x)[X_1], v\rangle x \\ &\quad - \langle X_1(x), v\rangle X_2(x) + \langle X_2(x), v\rangle X_1(x) \\ &\quad - \underbrace{\big\langle X_1(x), \big(-\langle X_2(x), v\rangle x\big\rangle}_{=0} + \underbrace{\big\langle X_2(x), \big(-\langle X_1(x), v\rangle x\big\rangle}_{=0}\big)\Big) \\ &= \Big([X_1, X_2](x), -\big\langle [X_1, X_2](x), v\big\rangle x - \langle X_1(x), v\rangle X_2(x) + \langle X_2(x), v\rangle X_1(x)\Big). \end{aligned}$$

On the other hand, we have

$$\overline{[X_1, X_2]}(x,v) = \big([X_1, X_2](x), -\langle [X_1, X_2](x), v\rangle x\big).$$

so thanks to Lemma 28.9 we deduce that

$$\begin{aligned} [\overline{X}_1, \overline{X}_2]^{\mathtt{V}}(x,v) &= [\overline{X}_1, \overline{X}_2](x,v) - [\overline{X}_1, \overline{X}_2]^{\mathtt{H}}(x,v) \\ &= [\overline{X}_1, \overline{X}_2](x,v) - \overline{[X_1, X_2]}(x,v) \\ &= \big(0, -\langle X_1(x), v\rangle X_2(x) + \langle X_2(x), v\rangle X_1(x)\big), \end{aligned}$$

and so we conclude that the curvature tensor is

$$\begin{aligned} R^{\nabla}(X_1, X_2)(V)(x) &= -\operatorname{pr}_2\left([\overline{X}_1, \overline{X}_2]^{\mathtt{V}}\right)(x) \\ &= \langle X_1(x), V(x)\rangle X_2(x) - \langle X_2(x), V(x)\rangle X_1(x), \end{aligned}$$

for every $X_1, X_2, V \in \mathfrak{X}(S^n)$.

Problem P.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$.

1. Suppose $\beta\colon \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ is a bilinear map. Prove there exists a unique connection ∇^β on $TG \to G$ which satisfies the following condition: if $v, w \in \mathfrak{g}$ and X_v, X_w denote the corresponding left-invariant vector fields then

$$\nabla^\beta_{X_v}(X_w) = X_{\beta(v,w)}.$$

2. Prove that this connection is left-invariant in the sense that

$$(l_a)_\star(\nabla^\beta_X(Y)) = \nabla^\beta_{(l_s)_\star X}((l_a)_\star(Y)), \qquad \forall\, X, Y \in \mathfrak{X}(G),\ \forall\, a \in G.$$

Deduce that the parallel transport determined by this connection is left-invariant in the sense that if c is a parallel section along a curve γ then $Dl_a(\gamma) \circ c$ is a parallel section along $l_a \circ \gamma$.

3. Prove moreover that any such left-invariant connection ∇ determines such a bilinear map β via

$$\beta(v, w) := \nabla_{X_v}(X_w)(e),$$

and hence that there is a bijective correspondence between bilinear maps $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and left-invariant connections on TG.

SOLUTION.

1. Let n be the dimension of G and fix a basis $\{v_1, \dots, v_n\}$ of $\mathfrak{g}$. Since l_a is a diffeomorphism, the linear map $Dl_a(e) : T_eG \to T_aG$ is an isomorphism for every $a \in G$. Hence

$$\{X_{v_1}(a) = Dl_a(e)[v_1], \dots, X_{v_n}(a) = Dl_a(e)[v_n]\}$$

is a basis of T_aG, which shows that $\{X_{v_1}, \dots, X_{v_n}\}$ is a global frame for the tangent bundle $TG \to G$. By Remark 16.9, given vector fields $X, Y \in \mathfrak{X}(G)$, we can write $X = a^i X_{v_i}$ and $Y = b^j X_{v_j}$ (remember the implicit sum convention). The covariant derivative operator ∇^β must satisfy

$$\begin{aligned} \nabla^\beta_X(Y) &= a^i \nabla^\beta_{X_{v_i}}(b^j X_{v_j}) \\ &= a^i X_{v_i}(b^j) X_{v_j} + a^i b^j \nabla^\beta_{X_{v_i}}(X_{v_j}) \\ &= a^i X_{v_i}(b^j) X_{v_j} + a^i b^j X_{\beta(v_i, v_j)}. \end{aligned} \tag{P.2}$$

This shows that ∇^β (hence also the corresponding connection) is uniquely determined, provided it exists. To prove existence, we define ∇^β using the formula (P.2) just found. We have

$$\begin{aligned} \nabla^\beta_X(fY) &= a^i X_{v_i}(f b^j) X_{v_j} + a^i f b^j X_{\beta(v_i, v_j)} \\ &= a^i X_{v_i}(f) b^j X_{v_j} + f(a^i X_{v_i}(b^j) X_{v_j} + a^i b^j X_{\beta(v_i, v_j)}) \\ &= X(f) Y + f \nabla^\beta_X(Y). \end{aligned}$$

The other axioms for a covariant derivative operator are clearly satisfied, as well as $\nabla^\beta_{X_{v_i}}(X_{v_j}) = X_{\beta(v_i, v_j)}$. By bilinearity, we also have $\nabla^\beta_{X_v}(X_w) = X_{\beta(v,w)}$ for all $v, w \in \mathfrak{g}$.

2. Let $a \in G$. Observe that, since l_a is a diffeomorphism, $(l_a)_* Z$ defines a vector field for all $Z \in \mathfrak{X}(G)$. Letting

$$\nabla^{\beta,a}_X(Y) := (l_{a^{-1}})_* \nabla^\beta_{(l_a)_* X}((l_a)_* Y), \tag{P.3}$$

we see that $\nabla^{\beta,a}$ is a covariant derivative operator: this is a consequence of

- $\phi_*(fZ) = (f \circ \phi^{-1})(l_a)_* Z$,
- $(\phi_* Z)(f \circ \phi^{-1}) = (Z(f)) \circ \phi^{-1}$,

for any diffeomorphism $\phi : G \to G$, $Z \in \mathfrak{X}(G)$ and $f \in C^\infty(G)$.[2] Since

$$\nabla^{\beta,a}_{X_v}(X_w) = (l_{a^{-1}})_* \nabla^{\beta}_{X_v}(X_w) = (l_{a^{-1}})_* X_{\beta(v,w)} = X_{\beta(v,w)},$$

by uniqueness we must have $\nabla^{\beta,a} = \nabla^{\beta}$, and thus (applying $(l_a)_*$ to both sides of (P.3)) we conclude that the connection is left-invariant. Alternatively, one can check left-invariance directly using (P.2). Now notice that, viewing $F := Dl_a$ as a map from TG to itself, we have $(l_a)_* s = F \circ s \circ l_a^{-1}$ for all $s \in \mathfrak{X}(G)$. So what we proved can be stated as

$$\nabla^{\beta}_{Dl_a(x)[v]}(F \circ s \circ l_a^{-1})(ax) = Dl_a(x)[\nabla^{\beta}_v(s)(x)]$$

for all $x \in G$ and $v \in T_x G$. Given $p \in T_x G$, using the fact that

$$\mathcal{J}_{F(p)}(Dl_a(x)[w]) = DF(p)[\mathcal{J}_p(w)]$$

for all $w \in T_x G$, from the proof of Theorem 32.1 (where we replace x with ax, p with $F(p)$, v with $Dl_a(x)[v]$ and s with the section $F \circ s \circ l_a^{-1}$ in the definition of $\mathcal{H}_{F(p)}$, where s is a section with $s(x) = p$) we conclude that

$$DF(p)[\mathcal{H}_p] \subseteq \mathcal{H}_{F(p)}. \tag{P.4}$$

Now c being parallel along γ means that $c'(t) \in \mathcal{H}_{c(t)}$, which by (P.4) implies $(F \circ c)'(t) = DF(c(t))[c'(t)] \in \mathcal{H}_{F \circ c(t)}$, as desired.

3. Given any such left-invariant connection[3] ∇, by the properties of a covariant derivative operator we have that $\beta(v,w) := \nabla_{X_v}(X_w)(e)$ is bilinear. Also, ∇ satisfies

$$\nabla_{X_v}(X_w)(a) = \nabla_{(l_a)_* X_v}((l_a)_* X_w)(l_a(e)) = Dl_a(e)[\nabla_{X_v}(X_w)(e)],$$

giving $\nabla_{X_v}(X_w) = X_{\beta(v,w)}$. Hence, by uniqueness, we get $\nabla = \nabla^{\beta}$ and thus any left-invariant connection arises from some bilinear map β. Moreover, the bilinear map associated to ∇^{β} (as in the statement of P.5.3) is precisely β, since $\nabla^{\beta}_{X_v}(X_w)(e) = X_{\beta(v,w)}(e) = \beta(v,w)$. This shows that ∇^{β} uniquely determines β and that $\beta \mapsto \nabla^{\beta}$ is a bijective correspondence between bilinear maps $\beta : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and left-invariant connections, with inverse given by the formula in the statement.

[2] These trivial facts have nothing to do with Lie groups and hold for manifolds in general.
[3] Or, more precisely, the corresponding covariant derivative operator.

Problem Sheet Q

(♣) Problem Q.1. Let $\pi \colon E \to M$ be a vector bundle with connection ∇. Let $E_0 \subset E$ be a vector subbundle such that ∇ is reducible to E_0. Prove that ∇ restricts to define a connection on E_0.

Problem Q.2. Let $\pi \colon E \to M$ denote a vector bundle, and let ∇_1 and ∇_2 denote two connections on E.

(i) Prove that $\nabla_1 - \nabla_2$ defines an element $A \in \mathcal{A}^1(M, E)$.

(ii) If $\mathcal{H}_1$ and $\mathcal{H}_2$ are the distributions on E corresponding to ∇_1 and ∇_2 respectively, prove that for all $p \in E$ one has

$$\mathcal{H}_2|_p = \big\{\zeta + \mathcal{J}_p\big(A(D\pi(p)[\zeta])\big) \mid \zeta \in \mathcal{H}_1|_p\big\}$$

where A is in the previous part.

(iii) Prove that

$$R^{\nabla_2} = R^{\nabla_1} - d^{\nabla_1}A + [A, A],$$

where $[A, A] \in \mathcal{A}^2(M, E)$ is defined by

$$[A, A](X, Y) = A(X)A(Y) - A(Y)A(X), \qquad X, Y \in \mathfrak{X}(M).$$

(iv) Conversely, prove that if ∇ is a connection on E and $A \in \mathcal{A}^1(M, E)$ then $\nabla_1 := \nabla + A$ is another connection. Deduce that the space of connections on E is (non-canonically) isomorphic to $\mathcal{A}^1(M, E)$.

Problem Q.3. Let $\pi \colon E \to M$ denote a vector bundle with connection ∇. Let ∇^{Hom} denote the induced connection on $\mathrm{Hom}(E, E)$ defined in part (ii) of Problem P.1, and let d^{∇} and $d^{\nabla^{\mathrm{Hom}}}$ denote the corresponding exterior covariant differentials. Prove that for $A \in \mathcal{A}^r_{M,E}$ and $\xi \in \Omega_{M,E}$ we have

$$d^{\nabla}(A \wedge \xi) = d^{\nabla^{\mathrm{Hom}}} A \wedge \xi + (-1)^r A \wedge d^{\nabla}\xi.$$

(♣) Problem Q.4. Let $\pi \colon E \to M$ be a vector bundle of rank k over a connected manifold M. Fix a Lie subgroup $G \subset \mathrm{GL}(k)$.

(i) Let us say that a connection ∇ on G is a **G-connection** if $\mathrm{Hol}^{\nabla}(x) \subset G$, up to conjugation (cf. Corollary 32.12). Prove that this is well-defined (i.e. independent of the choice of x).

(ii) Fix a G-connection ∇_1, and let ∇_2 denote any other connection. Suppose that the difference $\nabla_1 - \nabla_2$ actually lies in[1] $\Omega^1(M, \mathfrak{hol}^{\nabla_1}) \subset \mathcal{A}^1(M, E)$. Prove that ∇_2 is also a G-connection.

[1] Recall that $\mathfrak{hol}^{\nabla_1}$ is in particular a submanifold of $\mathrm{Hom}(E, E)$, so this assumption makes sense.

Solutions to Problem Sheet Q

(♣) Problem Q.1. Let $\pi\colon E \to M$ be a vector bundle with connection ∇. Let $E_0 \subset E$ be a vector subbundle such that ∇ is reducible to E_0. Prove that ∇ restricts to define a connection on E_0.

Solution. Let us prove that the parallel transport system $\mathbb{P}$ on E induces, simply by taking its restriction to E_0, a parallel transport system on E_0, which we denote by $\mathbb{P}_0$, by checking that the four axioms in Definition 29.8 are satisfied.

First of all, thanks to property (iv) of Definition 29.8, in the definition of reducibility of E_0 we can take the domain of γ in Definition 34.2 to be any interval.

As for property (i), since $\widehat{\mathbb{P}}_\gamma\colon E_{\gamma(a)} \to E_{\gamma(b)}$ is an isomorphism, so it is its restriction $\mathbb{P}_0$ to the subspace $(E_0)_{\gamma(a)}$ onto its image, which is by assumption contained, and hence equal (being vector subspace of the same dimension) to $(E_0)_{\gamma(b)}$. Moreover since for $p \in E_0$ the section $\mathbb{P}_\gamma(p) : [a,b] \to E$ has image contained in E_0, and since every embedding is a weak embedding (Definition 11.19) the map $\mathbb{P}_\gamma(p)\colon [a,b] \to E_0$ is also smooth and hence $\mathbb{P}_\gamma(p) \in \Gamma_\gamma(E_0)$. Properties (ii)-(iii)-(iv) for $\mathbb{P}_0$ are then inherited from $\mathbb{P}$.

Consequently, the parallel transport system $\mathbb{P}_0$ uniquely determines a connection $\mathcal{H}_0 \subset TE_0$, given by (cfr. Theorem 30.1)

$$\mathcal{H}_0 = \left\{ \zeta \in TE_0 : \zeta = \frac{d}{dt}\mathbb{P}_\gamma(p)(t)\bigg|_{t=0} \text{ for some } p \in E_0 \text{ and some } \gamma : [0,1] \to M \right\},$$

but this distribution is precisely $(\mathcal{H})|_{E_0}$, and the induced connection map (cfr. Definition 31.1) $\kappa_0\colon TE_0 \to E_0$ is given by the restriction of the connection map $k : TE \to E$ of $\mathcal{H}$ to TE_0. In particular, the induced covariant derivative ∇_0 (cfr. Theorem 31.10) is simply the restriction to E_0 of the covariant derivative ∇ defined on E.

Problem Q.2. Let $\pi\colon E \to M$ denote a vector bundle, and let ∇_1 and ∇_2 denote two connections on E.

(i) Prove that $\nabla_1 - \nabla_2$ defines an element $A \in \mathcal{A}^1(M,E)$.

(ii) If $\mathcal{H}_1$ and $\mathcal{H}_2$ are the distributions on E corresponding to ∇_1 and ∇_2 respectively, prove that for all $p \in E$ one has

$$\mathcal{H}_2|_p = \left\{ \zeta + \mathcal{J}_p\big(A(D\pi(p)[\zeta])\big) \mid \zeta \in \mathcal{H}_1|_p \right\}$$

where A is in the previous part.

(iii) Prove that

$$R^{\nabla_2} = R^{\nabla_1} - d^{\nabla_1}A + [A,A],$$

where $[A,A] \in \mathcal{A}^2(M,E)$ is defined by

$$[A,A](X,Y) = A(X)A(Y) - A(Y)A(X), \qquad X,Y \in \mathfrak{X}(M).$$

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

(iv) Conversely, prove that if ∇ is a connection on E and $A \in \mathcal{A}^1(M, E)$ then $\nabla_1 := \nabla + A$ is another connection. Deduce that the space of connections on E is (non-canonically) isomorphic to $\mathcal{A}^1(M, E)$.

Solution. Proof of (i): Thanks to the Hom-Γ Theorem 16.30 and the Vector-valued Differential Form Criterion (Theorem 26.3) it is sufficient to show that $\nabla_1 - \nabla_2$ seen as a map

$$\begin{aligned} A := \nabla_1 - \nabla_2 \colon \mathfrak{X}(M) \times \Gamma(E) &\to \Gamma(E) \\ (X, s) &\to (\nabla_1)_X s - (\nabla_2)_X s \end{aligned}$$

is $C^\infty(M)$-linear in both the variables.

Since every connection ∇ is $C^\infty(M)$-linear in the first variable (and so it is the difference of two connections), it is sufficient to check the $C^\infty(M)$-linearity on the second variable, that is $A(X, fs) = fA(X, s)$ for all $f \in C^\infty(M)$, $X \in \mathfrak{X}(M)$ and $s \in \Gamma(E)$. Therefore, let us compute

$$\begin{aligned} A(X, fs) &= (\nabla_1)_X(fs) - (\nabla_2)_X(fs) \\ &= X(f)s + f(\nabla_1)_X s - X(f)s - f(\nabla_2)_X s \\ &= f((\nabla_1)_X s - (\nabla_2)_X s) = fA(X, s), \end{aligned}$$

which is exactly what we wanted to prove.

Proof of (ii): By the proof of Theorem 32.1, we now that

$$\mathcal{H}_2|_p = \{Ds(x)[X] - \mathcal{J}_p((\nabla_2)_X(s)(x)) \mid s \in \Gamma(E), s(x) = p, X \in T_x M\}.$$

Therefore, using the definition of A, we obtain that

$$\begin{aligned} \mathcal{H}_2|_p = \{Ds(x)[X] - \mathcal{J}_p((\nabla_1)_X(s)(x)) + \mathcal{J}_p(A(X, s)(x)) \mid s \in \Gamma(E), \\ s(x) = p, X \in T_x M\}. \end{aligned}$$

However, notice that $Ds(x)[X] - \mathcal{J}_p((\nabla_1)_X(s)(x))$ is a generic element of $\mathcal{H}_1|_x$, thus we can call it ζ. In particular we have that $X = D\pi(p)[\zeta]$ and $s(x) = p$. Hence we have that

$$\mathcal{H}_2|_p = \{\zeta + \mathcal{J}_p(A(D\pi(p)[\zeta], p)) \mid s \in \Gamma(E), s(x) = p, X \in T_x M\}.$$

Observe that the we can write $A(X, s)(x) = A(D\pi(p)[\zeta], p)$, because we have proven in the previous point that A is a $C^\infty(M)$-linear operator and thus it is pointwise, namely $A(X, s)(x)$ depends only on the values of X and s in the point x.

Proof of (iii): Let us compute $R^{\nabla_2}(X, Y)$. We have that

$$\begin{aligned} R^{\nabla_2}(X, Y)s &= (\nabla_2)_X((\nabla_2)_Y s) - (\nabla_2)_Y((\nabla_2)_X s) - (\nabla_2)_{[X,Y]} s \\ &= ((\nabla_1)_X - A(X))((\nabla_1)_Y s - A(Y)s) \\ &\quad - ((\nabla_1)_Y - A(Y))((\nabla_1)_X s - A(X)s) - ((\nabla_1)_{[X,Y]} s - A([X, Y])s) \\ &= R^{\nabla_1}(X, Y)s - A(X)((\nabla_1)_Y s - A(Y)s) - (\nabla_1)_X(A(Y)s) \\ &\quad + A(Y)((\nabla_1)_X s - A(X)s) + (\nabla_1)_Y(A(X)s) + A([X, Y])s \\ &= R^{\nabla_1}(X, Y)s - A(X)((\nabla_1)_Y s) - (\nabla_1)_X A(Y)s \\ &\quad - A(Y)((\nabla_1)_X s) + A(Y)((\nabla_1)_X s) + (\nabla_1)_Y A(X)s + A(X)((\nabla_1)_Y s) \\ &\quad + A([X, Y])s + A(X)A(Y)s - A(Y)A(X)s \\ &= R^{\nabla_1}(X, Y)s - d^{\nabla_1}A(X, Y)s + [A, A](X, Y)s, \end{aligned}$$

where we have used the definition of $A([X,Y])$ and Theorem 35.5.

Notice that one should be careful of the difference between $(\nabla_1)_X A(Y)s$ and $(\nabla_1)_X(A(Y)s)$.

Proof of (iv): We need to check that ∇_1 satisfies all the four conditions in Definition 31.8:

(i) and (ii) $(\nabla_1)_X s$ is obviously C^∞-linear in the variable X because both ∇ and A are so,

(iii) $(\nabla_1)_X s$ is also linear in s because ∇ and A are linear in s,

(iv) $(\nabla_1)_X(fs) = \nabla_X(fs) + A(X, fs) = X(f)s + f\nabla_X s + fA(X,s) = X(f)s + f(\nabla_1)_X s$.

Together with (i) this implies that, fixing a connection ∇ on E, the map $A \mapsto \nabla + A$ is an isomorphism between $\mathcal{A}^1(M,E)$ and the space of connections. The isomorphism is non-canonical since we had to choose the "base connection" ∇.

Problem Q.3. Let $\pi\colon E \to M$ denote a vector bundle with connection ∇. Let ∇^{Hom} denote the induced connection on $\mathrm{Hom}(E,E)$ defined in part (ii) of Problem P.1, and let d^∇ and $d^{\nabla^{\mathrm{Hom}}}$ denote the corresponding exterior covariant differentials. Prove that for $A \in \mathcal{A}^r_{M,E}$ and $\xi \in \Omega_{M,E}$ we have

$$d^\nabla(A \wedge \xi) = d^{\nabla^{\mathrm{Hom}}} A \wedge \xi + (-1)^r A \wedge d^\nabla \xi.$$

Solution. Before we start the proof we recall some facts and relations that will be used throughout the computation: For any (honest) r-form $\omega \in \Omega^r_M$ and any bundle-valued k-form $\xi \in \Omega^k_{M,E}$ one has

$$\omega \wedge \xi = (-1)^{rk} \xi \wedge \omega \in \Omega^{r+k}_{M,E}.$$

For $k = 0$ one has $s = \xi \in \Omega^0_{M,E} = \Gamma(E)$, thus

$$\omega \wedge s = s \wedge \omega := \omega \otimes s.$$

Moreover, a decomposable element $\xi \in \Omega^r_{M,E}$ (resp. $A \in \mathcal{A}^k_{M,E}$) is of the form

$$\xi = \underbrace{\omega}_{\in \Omega^r(M)} \otimes \underbrace{s}_{\in \Gamma(E)}, \quad (\text{resp. } A = \underbrace{\vartheta}_{\in \Omega^k(M)} \otimes \underbrace{T}_{\in \Gamma(\mathrm{Hom}(E))}).$$

Also, the $\wedge$-operation between $\Omega^r_{M,E}$ and $\mathcal{A}^k_{(M,E)}$ gives back an element in $\Omega^{k+r}_{M,E}$. On decomposable elements we have

$$(\omega \otimes s) \wedge (\vartheta \otimes T) = (\omega \wedge \vartheta) \otimes T(s) = \omega \wedge \vartheta \wedge T(s),$$

where here we view T as C^∞-linear map between sections on E (cf. the Hom-Γ Theorem 16.30). Repeatedly invoking Theorem 35.4, we conclude the proof with the following painful computations on decomposable elements:

$$
\begin{aligned}
d^{\nabla}(A \wedge \xi) &= d^{\nabla}\big(\overbrace{(\vartheta \otimes T) \wedge (\omega \otimes s)}^{(\vartheta\wedge\omega)\wedge T(s)}\big) \\
&= d(\vartheta \wedge \omega) \wedge T(s) + (-1)^{k+r}(\vartheta \wedge \omega) \wedge \nabla(T(s)) \\
&= d\vartheta \wedge \omega \wedge T(s) + (-1)^{r}\vartheta \wedge d\omega \wedge T(s) + (-1)^{k+r}(\vartheta \wedge \omega) \wedge \nabla(T(s)) \\
&\overset{(1)}{=} d\vartheta \wedge \omega \wedge T(s) + (-1)^{r}\vartheta \wedge d\omega \wedge T(s) + (-1)^{r+k}\vartheta \wedge \omega \wedge \left(\nabla^{\mathrm{Hom}}T(s) + T(\nabla(s))\right).
\end{aligned}
$$

In (1) we used Problem (ii). At the same time we have

$$
\begin{aligned}
(-1)^{r}A \wedge d^{\nabla}\xi &= (-1)^{r}(\vartheta \otimes T) \wedge (d\omega \wedge s + (-1)^{k}\omega \wedge \nabla s) \\
&= (-1)^{r}\vartheta \wedge d\omega \wedge T(s) + (-1)^{r+k}(\vartheta \wedge \omega) \wedge T(\nabla(s)),
\end{aligned}
$$

and

$$
\begin{aligned}
d^{\mathrm{Hom}}A \wedge \xi &= d^{\mathrm{Hom}}(\vartheta \wedge T) \wedge (\omega \wedge s) \\
&= (d\vartheta \wedge T) \wedge (\omega \wedge s) + (-1)^{r}\vartheta \wedge \nabla^{\mathrm{Hom}}T \wedge \omega \wedge s \\
&= d\vartheta \wedge \omega \wedge T(s) + (-1)^{r+k\cdot 1}\vartheta \wedge \omega \wedge \nabla^{\mathrm{Hom}}T(s).
\end{aligned}
$$

It can be checked readily that the sum of the last two formulas equals the first expression which, by linearity, concludes the proof.

(♣) Problem Q.4. Let $\pi \colon E \to M$ be a vector bundle of rank k over a connected manifold M. Fix a Lie subgroup $G \subset \mathrm{GL}(k)$.

(i) Let us say that a connection ∇ on G is a **G-connection** if $\mathrm{Hol}^{\nabla}(x) \subset G$, up to conjugation (cf. Corollary 32.12). Prove that this is well-defined (i.e. independent of the choice of x).

(ii) Fix a G-connection ∇_1, and let ∇_2 denote any other connection. Suppose that the difference $\nabla_1 - \nabla_2$ actually lies in[1] $\Omega^1(M, \mathfrak{hol}^{\nabla_1}) \subset \mathcal{A}^1(M, E)$. Prove that ∇_2 is also a G-connection.

Solution.

(i) Essentially this was already shown in Corollary 32.12. Explicitly, let $x, y \in M$ and pick a linear isomorphism $B \colon \mathbb{R}^k \to E_x$. Assume that $\mathrm{Hol}^{\nabla}(x; B)$ is a subset of G, up to conjugation: this means that

$$
\mathrm{Hol}^{\nabla}(x; B) = B^{-1}\,\mathrm{Hol}^{\nabla}(x)B \subseteq CGC^{-1}
$$

for some $C \in \mathrm{GL}(k)$. We want to show that the same holds for y. We can assume that $C = I$ (just replace B with BC). Since M is connected, there exists a piecewise smooth curve γ joining x to y. By Lemma 32.11 we get

$$
\mathrm{Hol}^{\nabla}(y) = \widehat{\mathbb{P}}_{\gamma}\,\mathrm{Hol}^{\nabla}(x)\widehat{\mathbb{P}}_{\gamma}^{-1},
$$

hence letting $B' := \widehat{\mathbb{P}}_{\gamma}B$, which is a linear isomorphism from $\mathbb{R}^k$ to E_y, we get

$$
\mathrm{Hol}^{\nabla}(y; B') = B^{-1}\widehat{\mathbb{P}}_{\gamma}^{-1}\,\mathrm{Hol}^{\nabla}(y)\widehat{\mathbb{P}}_{\gamma}B = B^{-1}\,\mathrm{Hol}^{\nabla}(x)B \subseteq G.
$$

[1] Recall that $\mathfrak{hol}^{\nabla_1}$ is in particular a submanifold of $\mathrm{Hom}(E, E)$, so this assumption makes sense.

(ii) Write $\nabla_2 = \nabla_1 - A$ with $A \in \Omega^1(M, \mathfrak{hol}^{\nabla_1})$. Given a piecewise smooth loop $\gamma\colon [0,1] \to M$ based at x, as in the first part we can find a linear isomorphism $B\colon \mathbb{R}^k \to E_x$ such that

$$\mathrm{Hol}^{\nabla_1}(x; B) = B^{-1}\,\mathrm{Hol}^{\nabla_1}(x) B \subseteq G.$$

Let $v_i := B(e_i)$, so that $\{v_1, \dots, v_k\}$ is a basis of E_x, and let ℓ_i^1 be the ∇_1-parallel lift of γ starting at v_i. Notice that, calling $\widehat{\mathbb{P}}_t^{\nabla^1}$ the parallel transport (with respect to ∇_1) along the curve $\gamma\,|_{[0,t]}$, we have $\ell_i(t) = \widehat{\mathbb{P}}_t^{\nabla^1}(v_i)$. Similarly we define the ∇_2-parallel lift ℓ_i^2. We can write

$$\ell_j^2(t) = c_j^i(t)\ell_i^1(t)$$

for all $j = 1, \dots, k$, for suitable piecewise smooth coefficients $c_j^i(t)$.[2] We define $C(t)$ to be the matrix whose (i,j)-coefficient is $c_j^i(t)$, identified with an element of $\mathrm{GL}(k)$. Alternatively, $C(t)$ represents the linear transformation $R(t)\colon \ell_i(t) \mapsto \ell_i'(t)$ on $E_{\gamma(t)}$, with respect to the basis $\{\ell_i^1(t)\}_{i=1}^k$, hence

$$C(t) = \big(\widehat{\mathbb{P}}_t^{\nabla_1}B\big)^{-1}R(t)\big(\widehat{\mathbb{P}}_t^{\nabla_1}B\big) \tag{Q.1}$$

by virtue of the fact that $\big(\widehat{\mathbb{P}}_t^{\nabla_1}B\big)(e_i) = \widehat{\mathbb{P}}_t^{\nabla_1}(v_i) = \ell_i^1(t)$. Moreover, by (Q.1),

$$B^{-1}\widehat{\mathbb{P}}_\gamma^{\nabla_2}B = B^{-1}R(1)\widehat{\mathbb{P}}_\gamma^{\nabla_1}B = \big(B^{-1}\widehat{\mathbb{P}}_\gamma^{\nabla_1}B\big)C(1).$$

Since $B^{-1}\widehat{\mathbb{P}}_\gamma^{\nabla_1}B \in \mathrm{Hol}^{\nabla_1}(x;B) \subseteq G$, it suffices to show that also $C(1) \in G$.

Now we translate the fact that ℓ_i^2 is ∇_2-parallel into an ordinary differential equation for $C(t)$. Denoting the vector field $\frac{\partial}{\partial t}$ on $[0,1]$ by T, we have

$$\begin{aligned} 0 &= (\nabla_2)_{T(t)}(\ell_j^2)(t) \\ &= (\nabla_1)_{T(t)}(\ell_j^2)(t) - A(\gamma'(t))(\ell_j^2(t)) \\ &= (\nabla_1)_{T(t)}(c_j^i\ell_i^1)(t) - A(\gamma'(t))(c_j^i(t)\ell_i^1(t)) \\ &= \big(c_j^i\big)'(t)\ell_i^1(t) - c_j^i(t)A(\gamma'(t))(\ell_i^1(t)). \end{aligned} \tag{Q.2}$$

Let $A(t) := \big(\widehat{P}_t^{\nabla_1}B\big)^{-1}A(\gamma'(t))\big(\widehat{P}_t^{\nabla_1}B\big)$ be the matrix representing $A(\gamma'(t))$ with respect to the basis $\{\ell_i^1(t)\}_{i=1}^k$. We claim that $A(t) \in \mathfrak{g}$, the Lie algebra of G:[3] this holds because the isomorphism

$$\mathrm{GL}(E_{\gamma(t)}) \to \mathrm{GL}(k), \quad S \mapsto \big(\widehat{P}_t^{\nabla_1}B\big)^{-1}S\big(\widehat{P}_t^{\nabla_1}B\big)$$

sends the Lie subgroup $\mathrm{Hol}^{\nabla_1}(\gamma(t))$ to $\mathrm{Hol}^{\nabla_1}(x;B) \subseteq G$ (by Lemma 32.11, since the conjugation $S \mapsto \big(\widehat{P}_t^{\nabla_1}\big)^{-1}S\widehat{P}_t^{\nabla_1}$ sends $\mathrm{Hol}^{\nabla_1}(\gamma(t))$ to $\mathrm{Hol}^{\nabla_1}(x)$), hence its differential at the identity (given by the same formula) sends the Lie algebra $\mathfrak{hol}^{\nabla_1}(\gamma(t))$ into $\mathfrak{g}$.

[2] This holds as $\{\ell_i\}_{i=1}^k$ is a frame for the pullback bundle γ^*E. Strictly speaking, some care is needed as γ is a *piecewise* smooth curve; this is not harmful since one can get the smoothness of the coefficients $c_j^i(t)$ on the finitely many intervals where γ is genuinely smooth.

[3] The proof that we give is essentially the same as the one of Remark 34.5.

Now, using again $\ell_i^1(t) = \big(\widehat{P}_t^{\nabla_1} B\big)(e_i)$ and applying $\big(\widehat{P}_t^{\nabla_1} B\big)^{-1}$ to (Q.2), we get

$$\big(c_j^i\big)'(t)e_i - c_j^i(t)A(t)(e_i) = 0$$

for all $j = 1, \dots, k$, which can be rewritten as

$$\big(c_j^k\big)'(t)e_k = c_j^i(t)a_i^k(t)e_k,$$

or (using that e_k is a basis and equating the coefficients of e_k)

$$C'(t) = A(t)C(t).$$

We claim that, together with $C(0) = I$ and $A(t) \in \mathfrak{g}$, this implies $C(t) \in G$ for all $t \in [0,1]$. Unlike the scalar case, the solution is *not* given by the explicit formula $\exp\big(\int_0^t A(s)\,ds\big)$ in general, so we have to use a different argument. On the manifold $[0,1] \times G$ we define the (piecewise smooth) vector field

$$Y(t,g) \coloneqq \frac{d}{dt} + X_{A(t)}(g),$$

where $X_{A(t)}$ is the *right*-invariant vector field corresponding to $A(t) \in \mathfrak{g}$ on the G factor. The integral curve starting from $(0, I)$ has the form $s \mapsto (s, \widetilde{C}(s))$, with $\widetilde{C}'(s) = A(s)\widetilde{C}(s)$ (once we implicitly compose $\widetilde{C}$ with the inclusion $\iota\colon G \hookrightarrow \mathrm{GL}(k)$).[4] So by uniqueness $C \equiv \widetilde{C}$ (and $\widetilde{C}(s)$ is defined for all times $s \in [0,1]$). Since $\widetilde{C}(s) \in G$, we get $C(s) \in G$, as wanted.

[4]This is due to the trivial fact that, given a Lie group H and a Lie subgroup H', with inclusion map $\iota\colon H' \hookrightarrow H$, the right-invariant vector fields corresponding to $v \in \mathfrak{h}'$ in H' and H are ι-related (a similar statement holds for left-invariant vector fields), and to the fact that right-invariant vector fields in $\mathrm{GL}(k)$ are given by $X_M(T) = MT$, for all $M \in \mathfrak{gl}(k)$.

Problem Sheet R

Problem R.1. Let $\pi\colon E \to M$ be a vector bundle with Riemannian metric m, and suppose ∇ is a connection on E that is Riemannian with respect to m. Fix $x \in M$. Prove that the holonomy group $\mathrm{Hol}^{\nabla}(x)$ is a subgroup of the orthogonal group

$$\mathrm{O}(E_x, m_x) := \{T \in \mathrm{L}(E_x, E_x) \mid m_x(T(p), T(q)) = m_x(p,q), \ \forall\, p, q \in E_x\}.$$

Problem R.2. Let $\pi\colon E \to M$ be a vector bundle with Riemannian metric m. Given $p \in E$ define $p^\flat \in E^*$ by

$$p^\flat(q) := m_{\pi(p)}(p, q)$$

(i) Prove that $\flat\colon E \to E^*$ is a vector bundle isomorphism.

(ii) Let $\sharp\colon E^* \to E$ denote the inverse of E (written $\eta \mapsto \eta^\sharp$). Prove that

$$m^*(\eta, \mu) := m_x(\eta^\sharp, \mu^\sharp), \qquad \eta, \mu \in E_x^*$$

defines a Riemannian metric on E^*.

(iii) Prove that (E, m) and (E^*, m^*) are isometric vector bundles in the sense of Definition 36.10. *Remark:* The vector bundle isomorphisms $\flat$ and $\sharp$ are usually[1] called the **musical isomorphisms**.

(♣) Problem R.3. Let $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ be an invariant homogeneous polynomial of odd degree $2r+1$. Prove that $\mathrm{CW}_E(\phi) = 0$ for any vector bundle of rank k.

(♣) Problem R.4. Let $\pi\colon E \to M$ be a vector bundle of rank k. Prove that the Chern-Weil map

$$\mathrm{CW}_E\colon P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)) \to H^*_{\mathrm{dR}}(M)$$

is an algebra homomorphism (where the algebra structure on the left-hand side is just the pointwise product of functions, and on the right-hand side it is the wedge product, cf. Definition 37.20).

(♣) Problem R.5. Suppose that E_1 and E_2 are two vector bundles over a smooth manifold M. Prove the **Whitney product formula** for the Pontryagin classes[2]

$$p_r(E_1 \oplus E_2) = \sum_{i=0}^{r} p_i(E_1) \wedge p_{r-i}(E_2).$$

(♣) Problem R.6. Prove directly that $p_r(TS^n) = 0$ for all $r > 0$. (Don't just quote Proposition 37.22!) *Remark:* This shows that Pontryagin classes alone cannot determine a vector bundle up to isomorphism (since $TS^n \to S^n$ is not a trivial bundle).

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1] There is no relation between between the musical isomorphisms and the $\sharp$-$\flat$ correspondence in Theorem 26.17—this is purely a notational conincidence.

[2] For those of you who are familiar with Algebraic Topology: the statement would be more complicated if one worked with (singular) cohomology with coefficients in $\mathbb{Z}$, since then one would need to worry about 2-torsion elements.

Solutions to Problem Sheet R

Problem R.1. Let $\pi\colon E\to M$ be a vector bundle with Riemannian metric m, and suppose ∇ is a connection on E that is Riemannian with respect to m. Fix $x\in M$. Prove that the holonomy group $\mathrm{Hol}^{\nabla}(x)$ is a subgroup of the orthogonal group

$$\mathrm{O}(E_x, m_x) \coloneqq \{T\in \mathrm{L}(E_x,E_x) \mid m_x(T(p),T(q)) = m_x(p,q),\ \forall\, p,q\in E_x\}\,.$$

Solution. The thesis is equivalent to prove that for every piecewise smooth loop $\gamma\colon [a,b]\to M$ at x it holds that

$$m_x(\widehat{\mathbb{P}}_\gamma(p), \widehat{\mathbb{P}}_\gamma(q)) = m_x(p,q)$$

for all $p,q\in E_x$.

Actually we will prove something more, that is $\mathbb{P}_\gamma(\cdot)(t)$ is an isometry for all $t\in[a,b]$, which means that

$$m_{\gamma(t)}(\mathbb{P}_\gamma(p)(t), \mathbb{P}_\gamma(q)(t)) = m_x(p,q)$$

for all $p,q\in E_x$. Indeed, thanks to the solution of Problem P.2 we have that

$$\begin{aligned}\frac{d}{dt}[m_{\gamma(t)}(\mathbb{P}_\gamma(p)(t), \mathbb{P}_\gamma(q)(t))] &= (\nabla_{\gamma'(t)} m_{\gamma(t)})(\mathbb{P}_\gamma(p)(t), \mathbb{P}_\gamma(q)(t))\\ &\quad + m_{\gamma(t)}(\nabla_{\gamma'(t)}\mathbb{P}_\gamma(p)(t), \mathbb{P}_\gamma(q)(t))\\ &\quad + m_{\gamma(t)}(\mathbb{P}_\gamma(p)(t), \nabla_{\gamma'(t)}\mathbb{P}_\gamma(q)(t)).\end{aligned}$$

However notice that $\nabla_{\gamma'(t)} m_{\gamma(t)} = 0$ since the connection is Riemannian with respect to m and that $\nabla_{\gamma'(t)}\mathbb{P}_\gamma(p)(t) = \nabla_{\gamma'(t)}\mathbb{P}_\gamma(q)(t) = 0$ thanks to Proposition 32.3. Consequently $\frac{d}{dt}[m_{\gamma(t)}(\mathbb{P}_\gamma(p)(t), \mathbb{P}_\gamma(q)(t))] = 0$, which proves what we wanted.

Problem R.2. Let $\pi\colon E\to M$ be a vector bundle with Riemannian metric m. Given $p\in E$ define $p^\flat\in E^*$ by

$$p^\flat(q) \coloneqq m_{\pi(p)}(p,q)$$

(i) Prove that $\flat\colon E\to E^*$ is a vector bundle isomorphism.

(ii) Let $\sharp\colon E^*\to E$ denote the inverse of E (written $\eta\mapsto\eta^\sharp$). Prove that

$$m^*(\eta,\mu) \coloneqq m_x(\eta^\sharp,\mu^\sharp), \qquad \eta,\mu\in E_x^*$$

defines a Riemannian metric on E^*.

(iii) Prove that (E,m) and (E^*,m^*) are isometric vector bundles in the sense of Definition 36.10. *Remark:* The vector bundle isomorphisms $\flat$ and $\sharp$ are usuallyfootnote There is no relation between between the musical isomorphisms and the $\sharp$-$\flat$ correspondence in Theorem 26.17—this is purely a notational conincidence. called the **musical isomorphisms**. called the **musical isomorphisms**.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

Solution.

(i) First notice that, by definition, we have that $\pi_E(p) = \pi_{E^*}(p^\flat)$. Therefore the map $\flat$ is a vector bundle morphism, since it is also obviously linear. However $\flat$ is also patently injective since m_x is non-degenerate and consequently it is a vector bundle isomorphism.

(ii) First notice that m^* is an element of $\Gamma(E, E)$. Moreover m_x^* is an inner product, since $\sharp$ is an isomorphism between E_x^* and E_x.

(iii) We have that $\sharp\colon E^* \to E$ is a vector bundle isomorphism and the metric m^* has been defined exactly in order to satisfy the extra relation needed in Definition 36.10 to be a isometric vector bundle morphism.

(♣) Problem R.3. Let $\phi \in P_{\mathrm{GL}(k)}(\mathfrak{gl}(k))$ be an invariant homogeneous polynomial of odd degree $2r+1$. Prove that $\mathrm{CW}_E(\phi) = 0$ for any vector bundle of rank k.

Solution. Using Proposition 36.11 we endow E with a Riemannian metric m. Proposition 36.17 provides us with a Riemannian connection ∇ with respect to m. Given an endomorphism $T \in \mathrm{Hom}(E_x, E_x)$, we define its adjoint $T^* \in \mathrm{Hom}(E_x, E_x)$ to be the unique endomorphism satisfying

$$m_x(T(v), w) = m_x(v, T(w)) \qquad \text{for } v, w \in E_x.$$

Alternatively, identifying $\mathrm{Hom}(E_x, E_x)$ with $E_x^* \otimes E_x$ as in Corollary 15.13, this operation (which is an endomorphism of $\mathrm{Hom}(E_x, E_x)$) is defined on generators by $(\eta \otimes v)^* := v^\flat \otimes \eta^\sharp$.

If $\{e_1, \ldots, e_k\}$ is a local orthonormal frame (whose existence is guaranteed by Lemma 36.12), we can write a section $T \in \Gamma(\mathrm{Hom}(E, E))$ locally as

$$T = \sum_{i,j=1}^{k} t^{ij} e_j^\flat \otimes e_i$$

for suitable smooth functions t^{ij},[1] obtaining

$$T^* = \sum_{i,j=1}^{k} t^{ji} e_j^\flat \otimes e_i.$$

This shows that also T^* is a smooth section of $\mathrm{Hom}(E, E)$, and that $(t^{ij}(x))$, $(t^{ji}(x))$ are the matrices representing $T(x)$ and $T^*(x)$ with respect to the basis $\{e_i(x)\}$. In other words, $T \mapsto T^*$ corresponds to matrix transposition (with respect to an orthonormal basis). Observe that this operation extends canonically to sections in $\mathcal{A}^\ell(M, E)$: for a decomposable element $\eta \otimes T$ we let

$$(\omega \otimes T)^* := \omega \otimes T^*.$$

Moreover, according to Corollary 36.19, we have $(R^\nabla)^* = -R^\nabla$. Hence, calling Φ the parallel section of $(\mathrm{Hom}^r(E, E))^*$ induced by ϕ,

$$\begin{aligned} \Phi((R^\nabla)^* \otimes \cdots \otimes (R^\nabla)^*) &= \Phi((-R^\nabla) \otimes \cdots \otimes (-R^\nabla)) \\ &= (-1)^{2r+1} \Phi(R^\nabla \otimes \cdots \otimes R^\nabla) \end{aligned} \tag{R.1}$$

[1] Notice that $\{e_j^\flat(x)\}$ is a basis of E_x^*, and in fact is the dual basis to $\{e_i(x)\}$.

(being Φ fiberwise a linear functional). Now the key observation is that

$$\phi(A) = \phi(A^*) \tag{R.2}$$

for every element $A \in \mathfrak{gl}(k)$, where A^* denotes the transpose: this holds because, as we know from linear algebra, every matrix is conjugate to its transpose. An alternative argument is the following: viewing ϕ as a polynomial function in the entries of the matrix, we have $\phi(CAC^{-1}) = \phi(A)$ also when A and C have complex entries (and C is invertible).[2] If $A \in \mathfrak{gl}(k, \mathbb{C})$ is diagonalizable, then we can write $A = CDC^{-1}$ with D diagonal, hence

$$\phi(A) = \phi(D) = \phi(D^*) = \phi((C^*)^{-1}D^*(C^*)) = \phi(A^*)$$

(here A^* still denotes the transpose of A). Since diagonalizable matrices are dense in $\mathfrak{gl}(k, \mathbb{C})$, we deduce that our claim (R.2) holds for any $A \in \mathfrak{gl}(k, \mathbb{C})$ (and in particular for real matrices). Letting $S \coloneqq \operatorname{polar}(\phi)$, we also have

$$S(A_1^* \otimes \cdots \otimes A_{2r+1}^*) = S(A_1 \otimes \cdots \otimes A_{2r+1}) \qquad \text{for } A_1, \ldots, A_{2r+1} \in \mathfrak{gl}(k),$$

as the left-hand side defines a symmetric tensor S' with $\operatorname{polar}^{-1}(S') = \operatorname{polar}^{-1}(S)$ (whence $S' = S$). We infer that the same holds replacing the A_i's with elements of $\mathcal{A}^\ell(M, E)$,[3] so that

$$\Phi((R^\nabla)^* \otimes \cdots \otimes (R^\nabla)^*) = \Phi(R^\nabla \otimes \cdots \otimes R^\nabla). \tag{R.3}$$

From (R.1) and (R.3) we get

$$\Phi(R^\nabla \otimes \cdots \otimes R^\nabla) = 0$$

and so $\mathrm{CW}_E(\phi) = [\phi(\nabla)] = [\Phi(R^\nabla \otimes \cdots \otimes R^\nabla)] = 0$.

(♣) Problem R.4. Let $\pi\colon E \to M$ be a vector bundle of rank k. Prove that the Chern-Weil map

$$\mathrm{CW}_E\colon P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)) \to H^*_{\mathrm{dR}}(M)$$

is an algebra homomorphism (where the algebra structure on the left-hand side is just the pointwise product of functions, and on the right-hand side it is the wedge product, cf. Definition 37.20).

Solution. We are going to show that for any two homogeneous polynomials, say ϕ of degree r and ψ of degree s, one has

$$\mathrm{CW}_E(\phi \cdot \psi) = \mathrm{CW}_E(\phi) \wedge \mathrm{CW}_E(\psi). \tag{R.4}$$

[2] We have $\phi(AB) = \phi(A(BA)A^{-1}) = \phi(BA)$ when A is invertible, hence (being invertible matrices a dense subset of $\mathfrak{gl}(k)$) $\phi(AB) = \phi(BA)$ for all real $A, B \in \mathfrak{gl}(k)$. Both sides of this identity are polynomial functions in the entries of A and B, so it must hold also for $A, B \in \mathfrak{gl}(k, \mathbb{C})$. In particular, $\phi((CA)C^{-1}) = \phi(C^{-1}(CA)) = \phi(A)$ also for $A, C \in \mathfrak{gl}(k, \mathbb{C})$.

[3] This holds because, as already observed, for $T \in \operatorname{Hom}(E_x, E_x)$ the endomorphism $T^* \in \operatorname{Hom}(E_x, E_x)$ is given by transposition (with respect to an orthonormal basis): namely, if $F\colon \mathbb{R}^k \to E_x$ is a linear isometry, then $F^{-1} \circ T^* \circ F \in \mathfrak{gl}(k)$ is the transpose of $F^{-1} \circ T \circ F \in \mathfrak{gl}(k)$.

Recall that $\theta = \phi\cdot\psi$ is a homogeneous polynomial of degree $r+s$ defined by pointwise multiplication and that there is a section on the dual bundle $(\mathrm{Hom}^{r+s}(E,E))^*$ induced by the polarisation of θ, namely

$$\Theta_x(T_1 \otimes \cdots \otimes T_{r+s}) := \mathrm{polar}(\theta)(F^{-1} \circ T_1(x) \circ F, \ldots, F^{-1} \circ T_{r+s}(x) \circ F),$$

where the T_i's are sections on $\mathrm{Hom}(E,E)$ and $F\colon \mathbb{R}^k \to E_x$ is an element in $\mathrm{Fr}(E)_x$. For decomposable elements

$$U_i = \omega_i \otimes T_i \in \Omega(M, \mathrm{Hom}(E,E))$$

one extends Θ as follows:

$$\begin{aligned}\Theta(U_1 \otimes \cdots \otimes U_{r+s}) &\overset{(1)}{=} \Theta\left((\omega_1 \wedge \cdots \wedge \omega_{r+s}) \otimes T_1 \otimes \cdots \otimes T_{r+s}\right) \\ &= \underbrace{\Theta(T_1 \otimes \cdots \otimes T_{r+s})}_{\in C^\infty(M)} \cdot \underbrace{(\omega_1 \wedge \cdots \wedge \omega_{r+s})}_{\in \Omega^{r+s}(M)} \in \Omega^{r+s}(M).\end{aligned}$$

In (1) here we used Definition 37.12. Now that we understand how the induced sections Θ, Φ and Ψ operate on decomposable elements $\Omega(M, \mathrm{Hom}(E,E))$ we will prove the identity

$$\Theta(\overbrace{U \otimes \cdots \otimes U}^{(r+s)-\text{times}}) = \Phi(\overbrace{U \otimes \cdots \otimes U}^{r-\text{times}}) \wedge \Psi(\overbrace{U \otimes \cdots \otimes U}^{s-\text{times}}) \in \Omega(M). \tag{R.5}$$

By linearity it suffices to show (R.5) for a decomposable element $U = \omega \otimes T$ with $\omega \in \Omega(M)$ and $T \in \Gamma(\mathrm{Hom}(E,E))$. We compute

$$\begin{aligned}\Theta(U \otimes \cdots \otimes U) &= \Theta(T \otimes \cdots \otimes T) \cdot \overbrace{\omega \wedge \cdots \wedge \omega}^{(r+s)-\text{times}} \\ &= \mathrm{polar}(\theta)(F^{-1} \circ T \circ F, \ldots, F^{-1} \circ T \circ F) \cdot \omega \wedge \cdots \wedge \omega \\ &= \mathrm{polar}(\phi\cdot\psi)(F^{-1} \circ T \circ F, \ldots, F^{-1} \circ T \circ F) \cdot (\overbrace{\omega \wedge \cdots \wedge \omega}^{r-\text{times}}) \wedge (\overbrace{\omega \wedge \cdots \wedge \omega}^{s-\text{times}}),\end{aligned} \tag{R.6}$$

but using that polar is an algebra (iso)morphism with the corresponding multiplication[4] on the symmetric tensor (cf. Definition 37.4) we get

$$\begin{aligned}(\text{R.6}) &= \Phi(\overbrace{T \otimes \cdots \otimes T}^{r-\text{times}}) \cdot \Psi(\overbrace{T \otimes \cdots \otimes T}^{s-\text{times}}) \cdot (\overbrace{\omega \wedge \cdots \wedge \omega}^{r-\text{times}}) \wedge (\overbrace{\omega \wedge \cdots \wedge \omega}^{s-\text{times}}) \\ &= (\Phi(T \otimes \cdots \otimes T) \cdot (\omega \wedge \cdots \wedge \omega)) \wedge (\Psi(T \otimes \cdots \otimes T) \cdot (\omega \wedge \cdots \wedge \omega)) \\ &= \Phi(U \otimes \cdots \otimes U) \wedge \Psi(U \otimes \cdots \otimes U).\end{aligned}$$

This proves (R.5) for decomposable elements, whence for the general case. The claim, i.e. the identity (R.4), now follows from (R.5) by setting $U = R^\nabla$.

(♣) Problem R.5. Suppose that E_1 and E_2 are two vector bundles over a smooth manifold M. Prove the **Whitney product formula** for the Pontryagin classes[5]

$$p_r(E_1 \oplus E_2) = \sum_{i=0}^{r} p_i(E_1) \wedge p_{r-i}(E_2).$$

[4]Note how the product becomes very simply in our case as there are no permutations.

[5]For those of you who are familiar with Algebraic Topology: the statement would be more complicated if one worked with (singular) cohomology with coefficients in $\mathbb{Z}$, since then one would need to worry about 2-torsion elements.

Solution. In order to compute $p_r(E_1 \oplus E_2)$ one needs to choose a connection ∇ on the bundle $E_1 \oplus E_2$. Fixing ∇^i any connection on E_i for $i = 1, 2$, we define

$$\nabla \colon \mathfrak{X}(M) \times \Gamma(E_1 \oplus E_2) \to \Gamma(E_1 \oplus E_2)$$

as follows: For two sections $s_i \in \Gamma(E_i)$ $i = 1, 2$, and a vector field $X \in \mathfrak{X}(M)$ we define

$$\nabla_X(s_1 \oplus s_2) := \nabla^1_X s_1 \oplus \nabla^2_X s_2$$

and extend linearly. This defines a covariant derivative on $E_1 \oplus E_2$. Indeed, the linearity in both entries and the C^∞-homogeneous bit in the first entry are clear, thus we will only check the Leibniz rule: Pick a smooth function $f \in C^\infty(M)$ and observe

$$\begin{aligned}
\nabla_X(f \cdot (s_1 \oplus s_2)) &= \nabla_X((f\, s_1 \oplus f\, s_2)) \\
&= \nabla^1_X(f\, s_1) \oplus \nabla^2_X(f\, s_2) \\
&= \big(X(f) \cdot s_1 + f \cdot \nabla^1_X s_1\big) \oplus \big(X(f) \cdot s_2 + f \cdot \nabla^2_X s_2\big) \\
&= X(f) \cdot (s_1 \oplus s_2) + f \cdot \big(\nabla^1_X s_1 \oplus \nabla^2_X s_2\big) \\
&= X(f) \cdot (s_1 \oplus s_2) + f \cdot \nabla_X(s_1 \oplus s_2).
\end{aligned}$$

This choice of connection has the advantage that the corresponding curvature tensor R^∇ is of the form

$$R^\nabla = R^{\nabla^1} \oplus R^{\nabla^2},$$

or in matrix notation

$$R^\nabla = \begin{pmatrix} R^{\nabla^1} & 0 \\ 0 & R^{\nabla^2} \end{pmatrix}.$$

This is easy to see by invoking the formula $R^\nabla(X, Y) = \nabla_X \nabla_Y - \nabla_X \nabla_Y - \nabla_{[X,Y]}$[6].

There is a crucial relation between the coefficients of the characteristic polynomial $\phi^k_r(A)$ of a block diagonal matrix

$$A = A_1 \oplus A_2 \colon \mathbb{R}^{k_1} \oplus \mathbb{R}^{k_2} \to \oplus A_2 \colon \mathbb{R}^{k_1} \oplus \mathbb{R}^{k_2}$$

and the corresponding coefficients for A_1 and A_2: Denoting these coefficients by $\phi^{k_1}_i(A_1)$ and $\phi^{k_2}_j(A_2)$, we claim that

$$\phi^k_r(A) = \sum_{i+j=r} \phi^{k_1}_i(A_1) \cdot \phi^{k_2}_j(A_2) \tag{R.7}$$

holds. Indeed, the matrix $t\mathrm{I}_{k\times k} + A$ is again block diagonal of the form

$$(t\mathrm{I}_{k_1\times k_1} + A_1) \oplus (t\mathrm{I}_{k_2\times k_2} + A_2),$$

thus

$$\det(t\mathrm{I}_{k\times k} + A) = \det(t\mathrm{I}_{k_1\times k_1} + A_1) \cdot \det(t\mathrm{I}_{k_2\times k_2} + A_2).$$

Expanding the characteristic polynomials of A, A_1 and A_2, using the relation above and comparing the coefficients then gives us (R.7).

[6] I'm not lying! Try writing out $\nabla_X \nabla_Y(s_1 \oplus s_2)$ and you will see how it goes.

Similarly to the solutions of the previous problem, we will first show the equality

$$\Phi^k(\overbrace{V \otimes \cdots \otimes V}^{r-\text{times}}) = \sum_{i+j=r} \Phi_i^{k_1}(\overbrace{V_1 \otimes \cdots \otimes V_1}^{i-\text{times}}) \wedge \Phi^{k_j}(\overbrace{V_2 \otimes \cdots \otimes V_2}^{j-\text{times}}), \tag{R.8}$$

for $V = V_1 \oplus V_2 \in \Omega(M, \operatorname{Hom}(E,E))$ with $V_i \in \Omega(M, \operatorname{Hom}(E_i, E_i))$ for $i = 1, 2$. Once we have shown (R.8) the desired result follows by taking r even and setting $V_i = R^{\nabla^i}$ for $i = 1, 2$.

Let us prove (R.8) on decomposable elements of the form

$$V = \omega \otimes \overbrace{(T_1 \oplus T_2)}^{=T} = (\omega \otimes T_1) \oplus (\omega \otimes T_2).$$

The result for a general $V = V_1 \otimes V_2$ then follows by linearity. We start computing

$$\begin{aligned} \Phi^k(V \otimes \cdots \otimes V) &= \Phi^k(T \otimes \cdots \otimes T) \cdot (\omega \wedge \cdots \wedge \omega) \\ &= \operatorname{polar}(\phi_r^k)(F^{-1} \circ T \circ F, \ldots, F^{-1} \circ T \circ F) \cdot (\omega \wedge \cdots \wedge \omega). \end{aligned} \tag{R.9}$$

By the invariance of polar we are free to choose any $F \colon \mathbb{R}^k \to E_x$ in the fibre $\operatorname{Fr}(E)_x$, so we may choose an F that is block diagonal $F = F_1 \oplus F_2$ such that

$$F^{-1} \circ T \circ F = (F_1^{-1} \circ T_1 \circ F_1) \oplus (F_2^{-1} \circ T_2 \circ F_2).$$

Therefore

$$\begin{aligned} \text{(R.9)} &= \sum_{i+j=r} \operatorname{polar}(\phi_r^k)(F^{-1} \circ T \circ F, \ldots, F^{-1} \circ T \circ F) \cdot (\omega \wedge \cdots \wedge \omega) \\ &= \Bigg[\sum_{i+j=r} \operatorname{polar}(\phi_i^{k_1})(F_1^{-1} \circ T_1 \circ F_1, \ldots, F_1^{-1} \circ T_1 \circ F_1) \\ &\qquad \cdot \operatorname{polar}(\phi_j^{k_2})(F_2^{-1} \circ T_2 \circ F_2, \ldots, F_2^{-1} \circ T_2 \circ F_2) \Bigg] \cdot (\omega \wedge \cdots \wedge \omega) \\ &= \sum_{i+j=r} \Phi_i^{k_1}(V \otimes \cdots \otimes V) \wedge \Phi_j^{k_2}(V \otimes \cdots \otimes V). \end{aligned}$$

where the last two steps used a similar argument as in the solution to the previous problem, but one needs to be careful: Since $\phi_i^{k_1} \colon \mathfrak{gl}(k_1) \to \mathbb{R}$ and $\phi_j^{k_2} \colon \mathfrak{gl}(k_2) \to \mathbb{R}$ are defined on different vector spaces, they live in different spaces and one cannot invoke the algebra property of polar directly. However, one can view both as homogeneous polynomials on $\mathfrak{gl}(k) = \mathfrak{gl}(k_1 + k_2)$ in the obvious way. In particular, $\operatorname{polar}(\phi_i^{k_1})$ satisfies

$$\operatorname{polar}(\phi_i^{k_1})(F^{-1} \circ T \circ F, \ldots, F^{-1} \circ T \circ F) = \operatorname{polar}(\phi_i^{k_1})(F_1^{-1} \circ T_1 \circ F_1, \ldots, F_1^{-1} \circ T_1 \circ F_1).$$

A similar formula hods for $\phi_j^{k_2}$. Then using (R.7) one obtains

$$\phi_r^k = \sum_{i+j=r} \phi_i^{k_1} \cdot \phi_j^{k_2} \text{ in } P_{\mathrm{GL}(k)}(\mathfrak{gl}(k)),$$

whence justifying the equality above.

As already anticipated, taking $2r$ instead of r, setting $V = R^{\nabla^1} \oplus R^{\nabla^2}$ and taking the cohomology class on both sides shows the desired equality.

(♣) Problem R.6. Prove that $p_r(TS^n) = 0$ for all $r > 0$. *Remark:* This shows that Pontryagin classes alone cannot determine a vector bundle up to isomorphism (since $TS^n \to S^n$ is not a trivial bundle). Note that Proposition 37.22 is not contradicted, as S^n does embed in $\mathbb{R}^{n+1}$!

Solution. In Problem P.4 we computed that the curvature tensor of S^n is

$$R(X_1, X_2)(V) = \langle X_1, V\rangle X_2 - \langle X_2, V\rangle X_1, \quad \text{for every } X_1, X_2, V \in TS^n.$$

If $v_1, \dots, v_n$ is any local orthonormal frame on TS^n (i.e. $v_1(p), \dots v_n(p)$ are linearly independent elements of T_pS^n and $\langle v_j(p), v_k(p)\rangle = \delta_{jk}$) and $v^1, \dots, v^n$ is the corresponding dual frame, we see that

$$\begin{aligned} R(X_1, X_2)(v_j) &= \langle X_1^k v_k, v_j\rangle X_2^l v_l - \langle X_2^k v_k, v_j\rangle X_1^l v_l \\ &= (X_1^j X_2^l - X_2^j X_1^l) v_l \\ &= (v^j \wedge v^l)(X_1, X_2) v_l, \end{aligned}$$

that is, the expression of the endomorphism $R(X_1, X_2)\colon T_pS^n \to T_pS^n$ in the basis $v_1, \dots v_n$ is $R(X_1, X_2)_j^l = (v^j \wedge v^l)(X_1, X_2)$. The r-th Pontryagin class of S^n ($r > 0$) is then represented by the $4r$-form

$$\begin{aligned} p_r &= \frac{1}{(2\pi)^k} \sum_{\substack{1 \le j_1 < \dots < j_{2r} \le n \\ \sigma \in S_{j_1, \dots, j_{2r}}}} (\operatorname{sgn} \sigma) R_{j_1}^{\sigma(j_1)} \wedge \dots \wedge R_{j_{2r}}^{\sigma(j_{2r})} \\ &= \frac{1}{(2\pi)^k} \sum_{\substack{1 \le j_1 < \dots < j_{2r} \le n \\ \sigma \in S_{j_1, \dots, j_{2r}}}} (\operatorname{sgn} \sigma) v^{j_1} \wedge v^{\sigma(j_1)} \wedge \dots \wedge v^{j_{2r}} \wedge v^{\sigma(j_{2r})}, \end{aligned}$$

where $S_{j_1, \dots j_{2r}}$ denotes the set of permutations of $\{j_1, \dots, j_{2r}\}$. Now, for every term, every v^{j_k} appears two times in the wedge product and since since $\omega \wedge \omega \equiv 0$ for every 1-form, this implies that every term is zero. Thus $p_r \equiv 0$.

Problem Sheet S

Problem S.1. Let P be a manifold and $\mathfrak{g}$ a Lie algebra. Let $\vartheta \in \Omega^1(P, \mathfrak{g})$. Prove that the 3-form $[[\vartheta, \vartheta], \vartheta] \in \Omega^3(P, \mathfrak{g})$ (defined as in Example 26.7) is identically zero.

Problem S.2. Let $\pi \colon P \to M$ denote a principal G-bundle, and let ϖ denote a connection on P with curvature form Ω. Fix $X, Y \in \mathfrak{X}(M)$, and let $\overline{X}$ and $\overline{Y}$ denote their horizontal lifts. Prove that

$$\overline{[X,Y]}(p) - \big[\overline{X}, \overline{Y}\big](p) = D\eta_p(e)\big[\Omega_p(\overline{X}(p), \overline{Y}(p))\big],$$

where $\eta_p \colon G \to P$ is the map $a \mapsto p \cdot a$.

Problem S.3. Let $\pi \colon P \to M$ denote a principal G-bundle, and let $\rho \colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. Let $\lambda := D\rho(e)$, and suppose $f \colon P \to V$ is an equivariant smooth function. Prove that for any $v \in \mathfrak{g}$, one has

$$\xi_v(f) + \lambda(v)[f] = 0.$$

Problem S.4. Let $\pi \colon P \to M$ be a principal G-bundle. Let $\rho \colon G \to \mathrm{GL}(V)$ denote an effective representation. Let ϖ denote a connection on P and let ∇ denote the associated connection on $\rho(P)$. Fix $x \in M$. Then we can regard $\mathrm{Hol}^{\varpi}(x)$ and $\mathrm{Hol}^{\nabla}(x)$ as subgroups of G and $\mathrm{GL}(V)$ respectively, which are defined up to conjugation. Prove that (up to conjugation)

$$\rho\big(\mathrm{Hol}^{\varpi}(x)\big) = \mathrm{Hol}^{\nabla}(x).$$

(♣) Problem S.5. Let $\pi \colon E \to M$ be a vector bundle of rank k, and let $\mathrm{Fr}(E) \to M$ denote the principal $\mathrm{GL}(k)$-bundle. Then by Proposition 39.1 there is a bijective correspondence between connections on E and connections on $\mathrm{Fr}(E)$. Fix a Lie subgroup $G \subset \mathrm{GL}(k)$. Prove that a connection ∇ on E is a G-connection in the sense of Problem Q.4 if and only if the corresponding connection ϖ on $\mathrm{Fr}(E)$ is reducible to G in the sense of Definition 41.5.

(♣) Problem S.6. Use the principal bundle version of the Bianchi Identity (i.e. (39.6)) to prove the vector bundle version (Theorem 36.1).

(♣) Problem S.7. Use the principal bundle version of the Ambrose-Singer Holonomy Theorem (Theorem 41.7) to prove the vector bundle version (Theorem 34.8).

(♣) Problem S.8. Develop the theory of characteristic classes for principal bundles[1].

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

[1]Solutions will not be provided for this problem!

Solutions to Problem Sheet S

Problem S.1. Let P be a manifold and $\mathfrak{g}$ a Lie algebra. Let $\vartheta \in \Omega^1(P, \mathfrak{g})$. Prove that the 3-form $[[\vartheta, \vartheta], \vartheta] \in \Omega^3(P, \mathfrak{g})$ (defined as in Example 26.7) is identically zero.

Solution. For a fixed basis $e_1, \dots, e_n$ of $\mathfrak{g}$, we may represent ϑ as $\vartheta = \omega^i \otimes e_i$ for suitable 1-forms $\omega^1, \dots, \omega^n \in \Omega^1(P)$, hence

$$\begin{aligned} [\vartheta, \vartheta] &= [\omega^i \otimes e_i, \omega^j \otimes e_j] = \omega^i \wedge \omega^j \otimes [e_i, e_j], \\ [[\vartheta, \vartheta], \vartheta] &= [\omega^i \wedge \omega^j \otimes [e_i, e_j], \omega^k \otimes e_k] = \omega^i \wedge \omega^j \wedge \omega^k \otimes [[e_i, e_j], e_k]. \end{aligned}$$

From the Jacobi identity and the alternating property of the wedge product, we see that

$$\begin{aligned} [[\vartheta, \vartheta], \vartheta] &= \omega^i \wedge \omega^j \wedge \omega^k \otimes [[e_i, e_j], e_k] \\ &= -\omega^i \wedge \omega^j \wedge \omega^k \otimes [[e_j, e_k], e_i] - \omega^i \wedge \omega^j \wedge \omega^k \otimes [[e_k, e_i], e_j] \\ &= -\omega^j \wedge \omega^k \wedge \omega^i \otimes [[e_j, e_k], e_i] - \omega^k \wedge \omega^i \wedge \omega^j \otimes [[e_k, e_i], e_j] \\ &= -2\omega^i \wedge \omega^j \wedge \omega^k \otimes [[e_i, e_j], e_k] \\ &= -2[[\vartheta, \vartheta], \vartheta], \end{aligned}$$

where we relabelled the sums over repeated indexes in the last step. Necessarily then $[[\vartheta, \vartheta], \vartheta] = 0$.

Problem S.2. Let $\pi \colon P \to M$ denote a principal G-bundle, and let ϖ denote a connection on P with curvature form Ω. Fix $X, Y \in \mathfrak{X}(M)$, and let $\overline{X}$ and $\overline{Y}$ denote their horizontal lifts. Prove that

$$\overline{[X, Y]}(p) - \left[\overline{X}, \overline{Y}\right](p) = D\eta_p(e)\left[\Omega_p(\overline{X}(p), \overline{Y}(p))\right],$$

where $\eta_p \colon G \to P$ is the map $a \mapsto p \cdot a$.

Solution. First of all we observe that as a general fact for preconnections one has

$$[\overline{X}, \overline{Y}]^H = \overline{[X, Y]},$$

see Lemma 28.9. Thus the left-hand side of the desired equality is just minus the vertical part, i.e.

$$-[\overline{X}, \overline{Y}]^V.$$

Using

$$\varpi_p(\zeta) = D\eta_p(e)^{-1}[\zeta^V]$$

as the definition of the connection form we observe that

$$\begin{aligned} D\eta_p(e)\left[\Omega_p(\overline{X}(p), \overline{Y}(p))\right] &= -D\eta_p(e)\left[\varpi_p\left([\overline{X}, \overline{Y}](p)\right)\right] \\ &= -[\overline{X}, \overline{Y}]^V(p), \end{aligned}$$

which proves the desired equality.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

Problem S.3. Let $\pi\colon P \to M$ denote a principal G-bundle, and let $\rho\colon G \to \mathrm{GL}(V)$ denote a smooth effective representation of G. Let $\lambda := D\rho(e)$, and suppose $f\colon P \to V$ is an equivariant smooth function. Prove that for any $v \in \mathfrak{g}$, one has

$$\xi_v(f) + \lambda(v)[f] = 0.$$

Solution. First notice that

$$\xi_v(f)(p) = Df(p)[\xi_v(p)] = Df(p)[D\eta_p(e)[v]] = D(f \circ \eta_p)(e)[v],$$

for every $p \in P$ and $g \in \mathfrak{g}$. Moreover, using the equivariance of f, we have that

$$f \circ \eta_p(a) = f(p \cdot a) = \rho(a^{-1})(f(p)),$$

for every $p \in P$ and $a \in G$. If $i\colon G \to G$ denotes the inversion map then $Di(e)[v] = -v$ by the chain rule. Thus from the previous equation we have

$$\begin{aligned}\xi_v(f)(p) &= D(f \circ \eta_p)(e)[v] \\ &= D\rho \circ Di(e)[f(p)] \\ &= -D\rho(e)[v][f(p)] \\ &= -\lambda(v)[f(p)],\end{aligned}$$

which is what we wanted.

Problem S.4. Let $\pi\colon P \to M$ be a principal G-bundle. Let $\rho\colon G \to \mathrm{GL}(V)$ denote an effective representation. Let ϖ denote a connection on P and let ∇ denote the associated connection on $\rho(P)$. Fix $x \in M$. Then we can regard $\mathrm{Hol}^{\varpi}(x)$ and $\mathrm{Hol}^{\nabla}(x)$ as subgroups of G and $\mathrm{GL}(V)$ respectively, which are defined up to conjugation. Prove that (up to conjugation)

$$\rho\big(\mathrm{Hol}^{\varpi}(x)\big) = \mathrm{Hol}^{\nabla}(x),$$

Solution. Let $x \in M$ and fix a $p \in P_x$ and denote $E = \rho(P)$ the associated vector bundle. We will use superscripts ϖ and ∇ in order to distinguish the two parallel transport systems

$$\widehat{\mathbb{P}}^{\varpi}_{\gamma}\colon P_x \to P_x \text{ and } \widehat{\mathbb{P}}^{\nabla}_{\gamma}\colon E_x \to E_X$$

associated to the piecewise smooth loop $\gamma\colon [0,1] \to M$ based at x. Using the map $\phi_p\colon \mathrm{Hol}^{\varpi}(x) \to H^{\varpi}(p)$ from Proposition 41.2 we can identify $\widehat{\mathbb{P}}^{\varpi}_{\gamma}$ with the unique element

$$b = \phi_p(\widehat{\mathbb{P}}^{\varpi}_{\gamma}) \in G$$

such that

$$p \cdot b = \widehat{\mathbb{P}}^{\varpi}_{\gamma}(p).$$

The fibre $E_x = (P \times_G V)_x$ can be viewed as the set of equivalence classes $[q, v]$, where $q = p$ is fixed and v runs over V. Recall the isomorphism

$$L_p\colon V \to E_x, \qquad v \mapsto [p, v].$$

The task at hand is to show

$$\rho\left(\phi_p(\widehat{\mathbb{P}}^{\varpi}_{\gamma})\right) = L_p^{-1} \circ \widehat{\mathbb{P}}^{\nabla}_{\gamma} \circ L_p. \tag{S.1}$$

We prove (S.1) by feeding a vector $v \in V$ to both sides. Observe that

$$\rho\left(\phi_p(\widehat{\mathbb{P}}^{\varpi}_{\gamma})\right)(v) = \rho_b(v)$$

and

$$\begin{aligned}
L_p^{-1} \circ \widehat{\mathbb{P}}^{\nabla}_{\gamma} \circ L_p(v) &= L_p^{-1} \circ \widehat{\mathbb{P}}^{\nabla}_{\gamma}[p, v] \\
&\overset{(1)}{=} L_p^{-1}\left(\left[\widehat{P}^{\varpi}_{\gamma}(p), v\right]\right) \\
&= L_p^{-1}\left[p \cdot b, v\right] \\
&\overset{(2)}{=} L_p^{-1}[p, \rho_b(v)] \\
&= \rho_b(v).
\end{aligned}$$

In (1) we used the induced parallel transport system on E coming from P (cf. proof of Theorem 38.5) and in (2) we simply used the very definition of $\rho(P) = P \times_G V$. This proves (S.1) and thus completes the proof.

(♣) Problem S.5. Let $\pi\colon E \to M$ be a vector bundle of rank k, and let $\mathrm{Fr}(E) \to M$ denote the principal $\mathrm{GL}(k)$-bundle. Then by Proposition 39.1 there is a bijective correspondence between connections on E and connections on $\mathrm{Fr}(E)$. Fix a Lie subgroup $G \subset \mathrm{GL}(k)$. Prove that a connection ∇ on E is a G-connection in the sense of Problem Q.4 if and only if the corresponding connection ϖ on $\mathrm{Fr}(E)$ is reducible to G in the sense of Definition 41.5.

Solution. We show the following statement: a connection ∇ on E is a G-connection if and only if, for any $q \in \mathrm{Fr}(E)$, there exists a principal G'-subbundle $Q \ni q$ such that the corresponding connection ϖ is reducible to Q, for some subgroup $G' \subseteq \mathrm{GL}(k)$ conjugated to G.[1]

Assume ∇ is a G-connection. We fix $q \in \mathrm{Fr}(E)$ and observe that, by Problem S.4 (taking $V := \mathbb{R}^k$, $\rho := \iota$ and identifying $\rho(P)$ with E),

$$\mathrm{Hol}^{\varpi}(q) = \mathrm{Hol}^{\nabla}(x; q),$$

where $x := \pi(q)$. By hypothesis, we have $\mathrm{Hol}^{\nabla}(x;q) \subseteq G'$ for some subgroup $G' \subseteq \mathrm{GL}(k)$ conjugated to G. Now the strategy is essentially the same as the one used to prove Theorem 41.6. We call Q the set of points which can be joined, by means of a piecewise smooth horizontal path, to a point of the form $q \cdot g$, with $g \in G'$. In the sequel, we will often use the notation

$$q \cdot G' := \{q \cdot g \mid g \in G'\}.$$

The same argument used in the proof of Theorem 41.6 shows that $\pi|_Q$ is surjective and that local sections exist. We now show that the action of G' preserves Q: given $q' \in Q$, we have $q' = \widehat{\mathbb{P}}_{\gamma}(q \cdot g)$ for some $g \in G'$ and some curve γ joining x to $\pi(q')$. Then, by equivariance of parallel transport, we have

$$q' \cdot g' = \widehat{\mathbb{P}}_{\gamma}(q \cdot g) \cdot g' = \widehat{\mathbb{P}}_{\gamma}((q \cdot g) \cdot g') = \widehat{\mathbb{P}}_{\gamma}(q \cdot (gg')) \in Q,$$

[1] G' can depend on q.

for all $g' \in G'$. Finally, we show that the action of G' is transitive on each fibre: given $q_1, q_2 \in Q \cap P_y$ we can write $q_j = \widehat{\mathbb{P}}_{\gamma_j}(q \cdot g_j)$ for suitable $g_j \in G'$ and $\gamma_j \colon [0,1] \to M$ with $\gamma_j(0) = x$, $\gamma_j(1) = y$ (for $j = 1, 2$). The curve $\gamma_2 * \gamma_1^-$ is a loop based at x, so

$$\widehat{\mathbb{P}}_{\gamma_2 * \gamma_1^-}(q) = q \cdot b$$

for some $b \in \mathrm{Hol}^{\varpi}(q) \subseteq G'$. Thus,

$$\begin{aligned} q_2 &= \widehat{\mathbb{P}}_{\gamma_2 * \gamma_1^- * \gamma_1}(q \cdot g_2) \\ &= \widehat{\mathbb{P}}_{\gamma_2 * \gamma_1^- * \gamma_1}(q) \cdot g_2 \\ &= \widehat{\mathbb{P}}_{\gamma_1}(q \cdot b) \cdot g_2 \\ &= \widehat{\mathbb{P}}_{\gamma_1}(q \cdot g_1) \cdot (g_1^{-1} b g_2) \\ &= q_1 \cdot (g_1^{-1} b g_2) \end{aligned}$$

by equivariance. This shows that $q_2 \in q_1 \cdot G'$, as desired. So Proposition 24.20 applies, telling us that $Q \ni q$ is a principal G'-subbundle of $\mathrm{Fr}(E)$. The fact that ϖ is reducible to Q follows exactly as in the proof of Theorem 41.6.

Conversely, given $x \in M$, we pick a frame $q \in \mathrm{Fr}(E_x)$ and a principal G'-subbundle containing q. Given a piecewise smooth loop $\gamma \colon [0,1] \to M$ based at x, there exists a horizontal lift $c \colon [0,1] \to Q$ with $c(0) = q$. The endpoint $c(1)$ lies in $Q_x = q \cdot G'$. Also, observe that saying that c is horizontal with respect to the reduced connection on Q is equivalent to saying that c is ϖ-horizontal, so that $\widehat{\mathbb{P}}_{\gamma}(q) = c(1) \in q \cdot G'$. Hence,

$$\mathrm{Hol}^{\nabla}(x; q) = \mathrm{Hol}^{\varpi}(q) \subseteq G'.$$

(♣) Problem S.6. Use the principal bundle version of the Bianchi Identity (i.e. (39.6)) to prove the vector bundle version (Theorem 36.1).

Solution. Recall that any vector bundle $\pi \colon E \to M$ can be identified with the associated bundle $\rho(\mathrm{Fr}(E))$, where $\widehat{\pi} \colon \mathrm{Fr}(E) \to M$ is the frame bundle of E and $\rho = \mathrm{id}$ is the canonical representation. Specifically, the identification is given by the isomorphism

$$\rho(\mathrm{Fr}(E)) \to E, \qquad [A, v] \mapsto A(v)$$

where we view the frame A as a linear map $\mathbb{R}^k \to T_pM$ for $p := \widehat{\pi}(A)$. With the above identification in mind, fixing $A \in \mathrm{Fr}(E)$, Theorem 40.9 tells us that

$$R^{\nabla}(D\widehat{\pi}(A)[\zeta].D\widehat{\pi}(A)[\xi])(A(v)) = [A, \Omega_A(\zeta, \xi)(v)] = (\Omega_A(\zeta, \xi))^i_j v^j [A, e_i]. \tag{S.2}$$

Using Theorem 35.5 we get

$$\begin{aligned} d^{\nabla^{\mathrm{Hom}}}(R^{\nabla})(X, Y, Z) = \nabla^{\mathrm{Hom}}_X(R^{\nabla}(Y, Z)) - \nabla^{\mathrm{Hom}}_Y(R^{\nabla}(x, Z)) + \nabla^{\mathrm{Hom}}_Z(R^{\nabla}(X, Y)) \\ - R^{\nabla}([X, Y], Z) + R^{\nabla}([X, Z], Y) - R^{\nabla}([Y, Z], X). \end{aligned}$$

Observe that, by (S.2), for a fixed $v \in \mathbb{R}^k$ we have

$$R^{\nabla}([X, Y], Z)(A(v)) = [A, \Omega_A([\overline{X}, \overline{Y}], \overline{Z})(v)], \tag{S.3}$$

with $\overline{X}$ denoting the horizontal lift of X, and similarly for the last two terms. We claim that[2]

$$\nabla_X^{\mathrm{Hom}}(R^\nabla(Y,Z))(A(v)) = A(\overline{X}(\Omega(\overline{Y},\overline{Z})(v))). \tag{S.4}$$

Indeed, let $p := \widehat{\pi}(A)$ and let $\gamma\colon (-\epsilon,\epsilon) \to M$ be a curve with $\gamma'(0) = X(p)$. Let $A(t)$ be the parallel transport of A along γ (so that $A(0) = A$). Then

$$\begin{aligned}\nabla_X^{\mathrm{Hom}}(R^\nabla(Y,Z))(A(v)) &= \nabla_T(R^\nabla(Y,Z)(A(t)(v))) = \nabla_T[A(t),\Omega_{A(t)}(\overline{Y},\overline{Z})v] \\ &= \frac{d}{dt}(\Omega_{A(t)}(\overline{Y},\overline{Z}))^i_j v^j\big|_{t=0}[A,e_i]\end{aligned}$$

(where T is the vector $\frac{d}{dt}$ at 0), since $A(t)(v) = [A(t),v]$ and $[A(t),e_i]$ are parallel (see the first proof of Theorem 38.5). As $\frac{d}{dt}A(t)\big|_{t=0} = \overline{X}(A)$, (S.4) follows. Now (S.3) and (S.4) give

$$d^{\nabla^{\mathrm{Hom}}}(R^\nabla)(X,Y,Z)(A(v)) = A(d\Omega(\overline{X},\overline{Y},\overline{Z})(v)),$$

as $d\Omega(\overline{X},\overline{Y},\overline{Z})$ expands with an analogous formula with six terms.[3] The last term vanishes by the principal bundle version of the Bianchi identity, i.e. $d\Omega = [\Omega,\varpi]$, as ϖ vanishes on horizontal lifts. Being v and A arbitrary, the statement follows.

(♣) Problem S.7. Use the principal bundle version of the Ambrose-Singer Holonomy Theorem (Theorem 41.7) to prove the vector bundle version (Theorem 34.8).

Solution. Consider a vector bundle $\pi\colon E \to M$ and fix a point $x \in M$, we want to prove that the holonomy algebra $\mathfrak{hol}^\nabla(x)$ is the subalgebra of $\mathfrak{gl}(E_x)$ spanned by all the elements of the form

$$\hat{\mathbb{P}}^\nabla_\gamma \circ R^\nabla(v_1,v_2) \circ (\hat{\mathbb{P}}^\nabla_\gamma)^{-1}$$

with $y \in M$, $v_1, v_2 \in T_yM$ and γ a piecewise smooth curve from y to x.

Consider the principal $\mathrm{GL}(k)$-bundle $\pi\colon \mathrm{Fr}(E) \to M$ associated to E and denote $P = \mathrm{Fr}(E)$. Notice that the correspondence between E and $\mathrm{Fr}(E)$ is given by the trivial representation $\rho = \mathrm{id}\colon \mathrm{GL}(k) \to \mathrm{GL}(k)$. In particular the vector space considered is $V = \mathbb{R}^k$. The key idea of the translation from the principal bundle setting to the vector bundle setting is that, for every $x \in M$ and $p \in P_x$, the conjugation with $L_p\colon \mathbb{R}^k \to E_x$ gives an isomorphism from $\mathrm{Hom}(E_x,E_x)$ and $\mathfrak{gl}(k)$. Moreover this isomorphism well-behaves with respect to the parallel transport.

With this idea in mind we can prove our problem. Fix $p \in P_x$; then, by Problem S.4 (look in particular at the solution), we have that

$$\mathrm{Hol}^\nabla(x) = L_p \circ H^\varpi(p) \circ L_p^{-1}.$$

Thanks to the principal bundle version of the Ambrose-Singer Holonomy Theorem (Theorem 41.7), we know that the holonomy algebra $\mathfrak{h}$ of $H^\varpi(p)$ is the subalgebra of $\mathfrak{gl}(k)$ spanned by all the elements of the form $\Omega_q(\xi_1,\xi_2)$, for $q \in P$ that

[2] The right-hand side is A applied to the $\mathbb{R}^k$-valued function $\overline{X}(\Omega(\overline{Y},\overline{Z})(v))$ evaluated at A.
[3] Note that, according to our identification, $[A,\Omega_A([\overline{X},\overline{Y}],\overline{Z})(v)] = A(\Omega([\overline{X},\overline{Y}],\overline{Z})(v))$.

can be reached from p via a piecewise smooth horizontal path and for $\xi_1, \xi_2 \in T_qP$ horizontal.

Thus, $\mathfrak{hol}^\nabla(x)$ is spanned by all the elements of the form $L_p \circ \Omega_q(\xi_1, \xi_2) \circ L_p^{-1}$ for q, ξ_1, x_2 as above.

Now, given q as above, consider a curve γ in M connecting $y = \pi(q)$ and x such that $\hat{\mathbb{P}}^\varpi_\gamma(q) = p$ (which exists by definition of q). Then notice that, for all $v \in \mathbb{R}^k$, it holds

$$L_p^{-1} \circ \hat{\mathbb{P}}^\nabla_\gamma \circ L_q(v) = L_p^{-1}(\hat{\mathbb{P}}^\nabla_\gamma([q, v])) = L_p^{-1}([\hat{\mathbb{P}}^\varpi_\gamma(q), v]) = L_p^{-1}([p, v]) = v.$$

Consequently we have that

$$L_p \circ \Omega_q(\xi_1, \xi_2) \circ L_p^{-1} = \hat{\mathbb{P}}^\nabla_\gamma \circ L_q \circ \Omega_q(\xi_1, \xi_2) \circ L_q^{-1} \circ (\hat{\mathbb{P}}^\nabla_\gamma)^{-1}.$$

Observe that, when we consider the correspondence between $\pi \colon E \to M$ and $\pi \colon P \to M$, Theorem 40.9 can be written more simply as

$$R^\nabla(v_1, v_2) = L_q \circ \Omega_q(\xi_1, \xi_2) \circ L_q^{-1}$$

for all $q \in P$, $\xi_1, \xi_2 \in T_q\mathrm{Fr}(E)$ horizontal and $v_1 = D\pi(q)[\xi_1], v_2 = D\pi(q)[\xi_2]$.

Therefore we obtain that

$$L_p \circ \Omega_q(\xi_1, \xi_2) \circ L_p^{-1} = \hat{\mathbb{P}}^\nabla_\gamma \circ R^\nabla(v_1, v_2) \circ (\hat{\mathbb{P}}^\nabla_\gamma)^{-1}, \tag{S.5}$$

where $v_i = D\pi(q)[\xi_i]$ for $i = 1, 2$. Moreover notice that for every choice of $y \in M$, $v_1, v_2 \in T_yM$ and γ piecewise smooth curve from y to x, we can find $q \in P$ and $\xi_1, \xi_2 \in T_qP$ horizontal such that (S.5) holds (just choosing $q = \hat{\mathbb{P}}^\varpi(p)$ and x_1, x_2 the horizontal liftings of v_1, v_2).

Thus $\mathfrak{hol}^\nabla(x)$ is spanned by all the elements of the form $\hat{\mathbb{P}}^\nabla_\gamma \circ R^\nabla(v_1, v_2) \circ (\hat{\mathbb{P}}^\nabla_\gamma)^{-1}$ for $y \in M$, $v_1, v_2 \in T_yM$ and γ piecewise smooth curve from y to x, as we wanted.

(♣) Problem S.8. Develop the theory of characteristic classes for principal bundles[4].

Solution. Enjoy!

[4]Solutions will not be provided for this problem!

Problem Sheet T

(♣) PROBLEM T.1. Let ∇ be a connection on M. Let $\sigma\colon U \to O$ and $\tau\colon V \to \Omega$ denote two charts on M with local coordinates (x^i) and (y^i) respectively. Assume that $U \cap V \neq \emptyset$. Let Γ^k_{ij} denote the Christoffel symbols of σ and $\tilde{\Gamma}^k_{ij}$ denote the Christoffel symbols of τ, so that

$$\nabla_{\frac{\partial}{\partial x^i}}\left(\frac{\partial}{\partial x^j}\right) = \Gamma^k_{ij}\frac{\partial}{\partial x^k}, \qquad \nabla_{\frac{\partial}{\partial y^i}}\left(\frac{\partial}{\partial y^j}\right) = \tilde{\Gamma}^k_{ij}\frac{\partial}{\partial y^k}.$$

Investigate the relationship between

$$\Gamma^k_{ij}\big|_{U\cap V} \qquad \text{and} \qquad \tilde{\Gamma}^k_{ij}\big|_{U\cap V}.$$

PROBLEM T.2. Let ∇ denote a connection on M, and let d^∇ denote the associated exterior covariant differential. Prove that

$$T^\nabla = d^\nabla(\mathrm{id}).$$

PROBLEM T.3. Let ∇ be a torsion-free connection on M with curvature tensor R^∇. Prove that for all $X, Y, Z \in \mathfrak{X}(M)$, one has

$$\big(\nabla_X R^\nabla\big)(Y,Z) + \big(\nabla_Y R^\nabla\big)(Z,X) + \big(\nabla_Z R^\nabla\big)(X,Y) = 0.$$

PROBLEM T.4. Let G denote a Lie group, and let $\mathfrak{g}$ denote the Lie algebra of G. A Riemannian metric m on G is **left-invariant** if

$$l_a^\star(m) = m, \qquad \forall\, a \in G$$

and **right-invariant** if

$$r_a^\star(m) = m, \qquad \forall\, a \in G.$$

A Riemannian metric is **bi-invariant** if it is both left and right-invariant.

(i) Suppose $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is an inner product on $\mathfrak{g}$. Prove that $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ induces a left-invariant Riemannian metric m on G by

$$m_a(X_v(a), X_w(a)) \coloneqq \langle v, w\rangle_{\mathfrak{g}}, \qquad \forall\, v, w \in \mathfrak{g},\ a \in G,$$

where X_v is the left-invariant vector field on G with $X_v(e) = v$. Prove moreover that every left-invariant Riemannian metric on G is of this form.

(ii) Prove that the Riemannian metric m associated to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is right-invariant (and hence bi-invariant) if and only if

$$\langle \mathrm{Ad}_a(v), \mathrm{Ad}_a(w)\rangle_{\mathfrak{g}} = \langle v, w\rangle_{\mathfrak{g}}, \qquad \forall\, v, w \in \mathfrak{g},\ a \in G.$$

(iii) Assume now that G is connected. Prove that the Riemannian metric m associated to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is bi-invariant if and only if ad_v is skew-symmetric with respect to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ for all $v \in \mathfrak{g}$.

(♣) Problem T.5. Let G denote a Lie group, and let $\mathfrak{g}$ denote the Lie algebra of G. Let ∇^c denote the connection[1] on G defined by

$$\nabla^c_{X_e}(X_w) = c[X_v, X_w], \qquad \forall\, v, w \in \mathfrak{g}.$$

Let m denote a bi-invariant Riemannian metric on G.

(i) Prove that ∇^c is complete for any $c \in \mathbb{R}$.

(ii) Prove that ∇^c is Riemannian with respect to m for all $c \in \mathbb{R}$.

(iii) Prove that $\nabla^{\frac{1}{2}}$ is torsion-free. *Remark:* This shows that $\nabla^{\frac{1}{2}}$ is the Levi-Civita connection of (G, m).

(iv) Prove that $\nabla^{\frac{1}{2}}$ is right-invariant[2] in the sense that

$$(r_a)_\star(\nabla^{\frac{1}{2}}_X(Y)) = \nabla^{\frac{1}{2}}_{(r_a)_\star X}((r_a)_\star(Y)), \qquad \forall\, X, Y \in \mathfrak{X}(G),\ \forall\, a \in G.$$

(v) Compute the curvature tensor $R^{\nabla^{\frac{1}{2}}}$ of $\nabla^{\frac{1}{2}}$.

[1] This is the connection on G given by taking $\beta = c[\cdot,\cdot]$ in Problem P.5.
[2] We already know from Problem P.5 that $\nabla^{\frac{1}{2}}$ is left-invariant.

Solutions to Problem Sheet T

(♣) Problem T.1. Let ∇ be a connection on M. Let $\sigma\colon U \to O$ and $\tau\colon V \to \Omega$ denote two charts on M with local coordinates (x^i) and (y^i) respectively. Assume that $U \cap V \neq \emptyset$. Let Γ^k_{ij} denote the Christoffel symbols of σ and $\tilde{\Gamma}^k_{ij}$ denote the Christoffel symbols of τ, so that

$$\nabla_{\frac{\partial}{\partial x^i}}\left(\frac{\partial}{\partial x^j}\right) = \Gamma^k_{ij}\frac{\partial}{\partial x^k}, \qquad \nabla_{\frac{\partial}{\partial y^i}}\left(\frac{\partial}{\partial y^j}\right) = \tilde{\Gamma}^k_{ij}\frac{\partial}{\partial y^k}.$$

Investigate the relationship between

$$\Gamma^k_{ij}\big|_{U\cap V} \qquad \text{and} \qquad \tilde{\Gamma}^k_{ij}\big|_{U\cap V}.$$

Solution. For greater clarity, we will denote with Greek letters the indices pertaining to the x-coordinates and with Latin letter those pertaining to the y-coordinates.

Let $A\colon \sigma(U\cap V) \to \mathrm{GL}(n,\mathbb{R})$, $A(x) = D(\tau\circ\sigma^{-1})(x)$ be the Jacobian matrix of the coordinate change, and let B be its inverse. Recall that, when changing coordinates, for $x \in \sigma(U \cap V)$, the local representation of a vector field

$$X(x) = X^\mu(x)\frac{\partial}{\partial x^\mu}\bigg|_x \quad \text{transforms as} \quad Y((\tau\circ\sigma^{-1})(x)) = A^i_\mu(x)X^\mu(x)\frac{\partial}{\partial y^i}\bigg|_{(\sigma\circ\tau^{-1})(x)}.$$

Consequently, using the properties of the covariant derivative we can compute:

$$\begin{aligned}
\Gamma^\xi_{\mu\nu}\frac{\partial}{\partial x^\xi} &= \nabla_{\frac{\partial}{\partial x^\mu}}\left(\frac{\partial}{\partial x^\nu}\right)\\
&= \nabla_{\frac{\partial}{\partial x^\mu}}\left(A^i_\nu\frac{\partial}{\partial y^i}\right)\\
&= \frac{\partial A^i_\nu}{\partial x^\mu}\frac{\partial}{\partial y^i} + A^i_\nu\nabla_{A^j_\mu\frac{\partial}{\partial y^j}}\left(\frac{\partial}{\partial y^i}\right)\\
&= \frac{\partial A^i_\nu}{\partial x^\mu}\frac{\partial}{\partial y^i} + A^i_\nu A^j_\mu\widetilde{\Gamma}^k_{ji}\frac{\partial}{\partial y^k},
\end{aligned}$$

hence, using again the transformation law, that

$$\Gamma^\xi_{\mu\nu}A^k_\xi\frac{\partial}{\partial y^k} = \left(\frac{\partial A^k_\nu}{\partial x^\mu} + A^i_\nu A^j_\mu\widetilde{\Gamma}^k_{ji}\right)\frac{\partial}{\partial y^k}.$$

Inverting the matrix A (an operation that reads as: $A^k_\xi B^\eta_k = \delta^\eta_\xi$) we deduce that the transformation law for the Christoffel symbols is

$$\Gamma^\eta_{\mu\nu} = \frac{\partial A^k_\nu}{\partial x^\mu}B^\eta_k + A^i_\nu A^j_\mu B^\eta_k\widetilde{\Gamma}^k_{ji} \quad \text{for } \eta,\mu,\nu = 1,\dots,m = \dim(M).$$

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

Using the classical notation for the matrices

$$A^i_\mu = \frac{\partial y^i}{\partial x^\mu} \quad \text{and} \quad B^\nu_j = \frac{\partial x^\nu}{\partial y^j},$$

the above identity can be rewritten as

$$\Gamma^\eta_{\mu\nu} = \frac{\partial^2 y^k}{\partial x^\nu \partial x^\mu}\frac{\partial x^\mu}{\partial y^k} + \frac{\partial y^i}{\partial x^\nu}\frac{\partial y^j}{\partial x^\mu}\frac{\partial x^\eta}{\partial y^k}\widetilde{\Gamma}^k_{ji} \quad \text{for } \eta, \mu, \nu = 1, \dots, m = \dim(M).$$

Problem T.2. Let ∇ denote a connection on M, and let d^∇ denote the associated exterior covariant differential. Prove that

$$T^\nabla = d^\nabla(\mathrm{id}).$$

Solution. We view $\mathrm{id} \in \Omega^1_{M,TM}$, with $\mathrm{id}(v) = v$ for all $v \in T_pM$ and all $p \in M$. Let $X, Y \in \mathfrak{X}(M)$. Theorem 35.5 gives

$$\begin{aligned} d^\nabla(\mathrm{id})(X,Y) &= \nabla_X(\mathrm{id}(Y)) - \nabla_Y(\mathrm{id}(X)) - \mathrm{id}([X,Y]) \\ &= \nabla_X(Y) - \nabla_Y(X) - [X,Y], \end{aligned}$$

which is precisely the torsion $T^\nabla(X,Y)$.

Problem T.3. Let ∇ be a torsion-free connection on M with curvature tensor R^∇. Prove that for all $X, Y, Z \in \mathfrak{X}(M)$, one has

$$\big(\nabla_X R^\nabla\big)(Y,Z) + \big(\nabla_Y R^\nabla\big)(Z,X) + \big(\nabla_Z R^\nabla\big)(X,Y) = 0.$$

Solution. We show that, for all $X, Y, Z, W \in \mathfrak{X}(M)$,

$$(\nabla_X R)(Y,Z)(W) + (\nabla_Y R)(Z,X)(W) + (\nabla_Z R)(X,Y)(W) = 0.$$

Fix $x \in M$. Since the left-hand side involves point operators, in order to show that it vanishes at x we can replace X, Y, Z, W with other vector fields, as long as the vectors $X(x), Y(x), Z(x), W(x)$ do not change. Working on a suitable neighbourhood of x, we can thus assume that

$$\nabla_v(X) = \nabla_v(Y) = \nabla_v(Z) = \nabla_v(W) = 0$$

for all $v \in T_xM$, and that $[A,B] = 0$ for all possible choices $A, B \in \{X, Y, Z, W\}$: this can be achieved using Proposition 44.11 (replacing X with $X^i(x)\partial_i$, and similarly for Y, Z, W).[1] In particular, the vector field $\nabla_A(B)$ vanishes at p for all possible choices $A, B \in \{X, Y, Z, W\}$. Now

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

[1] These simplifying assumptions are not really necessary, but they will shorten the computations considerably.

so $(\nabla_X R)(Y,Z)(W) = \nabla_X(R(Y,Z)(W)) - R(Y,Z)(\nabla_X W)$ at p.[2] Thus, it suffices to show

$$\nabla_X(R(Y,Z)(W)) - R(Y,Z)(\nabla_X W) + \cdots = 0, \tag{T.1}$$

where the dots account for the similar terms obtained by *cyclically* permuting X, Y, Z (namely, replacing (X,Y,Z) with (Y,Z,X) and (Z,X,Y)). Expanding the curvature tensor and recalling that all Lie brackets vanish identically, we see that

$$\begin{aligned}\nabla_X(R(Y,Z)(W)) - R(Y,Z)(\nabla_X W) = \nabla_X\nabla_Y\nabla_Z W - \nabla_X\nabla_Z\nabla_Y W \\ - \nabla_Y\nabla_Z\nabla_X W + \nabla_Z\nabla_Y\nabla_X W,\end{aligned}$$

where we omitted the parentheses in order to lighten the notation. In order to conclude, note that the first term cancels out the third one when we sum over cyclic permutations, and similarly the second one cancels out the fourth one.

Alternatively, observe that ∇_X, ∇_Y, ∇_Z are linear operators on the vector space $\mathfrak{X}(M)$. The space of linear operators on a vector space always forms a Lie algebra, with $[S,T] \coloneqq S \circ T - T \circ S$. In particular, the Jacobi identity holds and so we get

$$[\nabla_X, [\nabla_Y, \nabla_Z]](W) + [\nabla_Y, [\nabla_Z, \nabla_X]](W) + [\nabla_Z, [\nabla_X, \nabla_Y]](W) = 0,$$

which is just a restatement of (T.1).

Problem T.4. Let G denote a Lie group, and let $\mathfrak{g}$ denote the Lie algebra of G. A Riemannian metric m on G is **left-invariant** if

$$l_a^\star(m) = m, \qquad \forall\, a \in G$$

and **right-invariant** if

$$r_a^\star(m) = m, \qquad \forall\, a \in G.$$

A Riemannian metric is **bi-invariant** if it is both left and right-invariant.

(i) Suppose $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is an inner product on $\mathfrak{g}$. Prove that $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ induces a left-invariant Riemannian metric m on G by

$$m_a(X_v(a), X_w(a)) \coloneqq \langle v, w\rangle_{\mathfrak{g}}, \qquad \forall\, v, w \in \mathfrak{g},\ a \in G, \tag{T.2}$$

where X_v is the left-invariant vector field on G with $X_v(e) = v$. Prove moreover that every left-invariant Riemannian metric on G is of this form.

(ii) Prove that the Riemannian metric m associated to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is right-invariant (and hence bi-invariant) if and only if

$$\langle \mathrm{Ad}_a(v), \mathrm{Ad}_a(w)\rangle_{\mathfrak{g}} = \langle v, w\rangle_{\mathfrak{g}}, \qquad \forall\, v, w \in \mathfrak{g},\ a \in G. \tag{T.3}$$

(iii) Assume now that G is connected. Prove that the Riemannian metric m associated to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ is bi-invariant if and only if ad_v is skew-symmetric with respect to $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ for all $v \in \mathfrak{g}$.

[2] One could also drop the term $R(Y,Z)(\nabla_X W)$. However, it is not convenient to do so.

Solution. Before starting the proof notice that a Riemannian metric m is left-invariant if and only if

$$m_a = l^*_{a^{-1}}(m_e) = (l^*_a)^{-1}(m_e) \tag{T.4}$$

for every $a \in G$, that is it is sufficient to check the left-invariance at the identity. Indeed, if we take an other $b \in G$, we then have that

$$l^*_a(m_b) = l^*_a(l^*_{b^{-1}} m_e) = l^*_{ab^{-1}}(m_e) = m_{ba^{-1}},$$

which proves the left-invariance. Obviously the same statement holds also substituting "left-invariance" with "right-invariance".

We can now start with the proof of the three points:

(i) Recall that $X_v(a) \coloneqq Dl_a(e)[v]$, thus for every $Y, Z \in T_aG$ we have that

$$m_a(Y, Z) = \langle (Dl_a(e))^{-1}[Y], (Dl_a(e))^{-1}[Z] \rangle_{\mathfrak{g}}$$
$$\implies m_a = (l^*_a)^{-1} \langle \cdot, \cdot \rangle_{\mathfrak{g}}.$$

Therefore it is obvious that m is a Riemannian metric on G and now we need only to prove that m is left-invariant if and only if (T.2) holds.

However notice that (T.2) is equivalent to

$$\begin{aligned} m_e(v, w) = \langle v, w \rangle_{\mathfrak{g}} &= m_a(X_v(a), X_w(a)) = m_a(Dl_a(e)[v], Dl_a(e)[w]) \\ &= (l^*_a)^{-1}(m_a)(v, w), \end{aligned}$$

which is exactly (T.4), that we have already proved to be equivalent to the left-invariance.

(ii) Recall that $\mathrm{Ad}_a(v) = D\mu_a(e)[v] = D(r_{a^{-1}} \circ l_a)(e)[v]$. Therefore (T.3) is equivalent to

$$\langle \cdot, \cdot \rangle_{\mathfrak{g}} = r^*_{a^{-1}} \circ l^*_a \langle \cdot, \cdot \rangle_{\mathfrak{g}} = r^*_{a^{-1}}(m_a),$$

using the definition of m. However, for what we said at the beginning of the solution, being right-invariant is equivalent to $m_e = r^*_{a^{-1}}(m_a)$ for all $a \in G$, which concludes the proof.

(iii) Let us first prove that, if m is bi-invariant, then ad_v is skew-symmetric for all $v \in \mathfrak{g}$. Let $\gamma \colon (-\varepsilon, \varepsilon) \to G$ be a curve with $\gamma(0) = e$ and $\gamma'(0) = u \in \mathfrak{g}$, then we have that

$$\begin{aligned} 0 = \left.\frac{d}{dt}\right|_{t=0} \left\langle \mathrm{Ad}_{\gamma(t)}(v), \mathrm{Ad}_{\gamma(t)}(w) \right\rangle_{\mathfrak{g}} &= \langle D\,\mathrm{Ad}(e)[u](v), w \rangle_{\mathfrak{g}} + \langle v, D\,\mathrm{Ad}(e)[u](w) \rangle_{\mathfrak{g}} \\ &= \langle \mathrm{ad}_u(v), w \rangle_{\mathfrak{g}} + \langle v, \mathrm{ad}_u(w) \rangle_{\mathfrak{g}}, \end{aligned}$$

which proves that ad_v is skew-symmetric for all $v \in \mathfrak{g}$.

On the other hand, let us now assume that ad_v is skew-symmetric for all $v \in \mathfrak{g}$. First of all notice that

$$\mathrm{Ad}_{ab} = D\mu_{ab}(e) = D(\mu_a \circ \mu_b)(e) = D\mu_a(e) \circ D\mu_b(e) = \mathrm{Ad}_a \circ \mathrm{Ad}_b$$

for all $a, b \in G$.

Now consider $a \in G$ and take a curve $\gamma\colon [0,1] \to G$ with $\gamma(0) = e$ and $\gamma(1) = a$. To prove that m is bi-invariant it is sufficient to prove that $\frac{d}{dt}\left\langle \mathrm{Ad}_{\gamma(t)}(v), \mathrm{Ad}_{\gamma(t)}(w)\right\rangle_{\mathfrak{g}} = 0$[3]. Hence let us compute it using the skew-symmetry of ad_v, obtaining

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=s} \left\langle \mathrm{Ad}_{\gamma(t)}(v), \mathrm{Ad}_{\gamma(t)}(w)\right\rangle_{\mathfrak{g}} &= \frac{d}{dt}\bigg|_{t=s} \left\langle \mathrm{Ad}_{\gamma(t)\gamma(s)^{-1}\gamma(s)}(v), \mathrm{Ad}_{\gamma(t)\gamma(s)^{-1}\gamma(s)}(w)\right\rangle_{\mathfrak{g}} \\
&= \frac{d}{dt}\bigg|_{t=s} \left\langle \mathrm{Ad}_{\gamma(t)\gamma(s)^{-1}}(\mathrm{Ad}_{\gamma(s)}(v)), \mathrm{Ad}_{\gamma(t)\gamma(s)^{-1}}(\mathrm{Ad}_{\gamma(s)}(w))\right\rangle_{\mathfrak{g}} \\
&= \left\langle \mathrm{ad}_{\gamma'(t)\gamma(s)^{-1}}(\mathrm{Ad}_{\gamma(s)}(v)), \mathrm{Ad}_{\gamma(s)}(w)\right\rangle_{\mathfrak{g}} + \left\langle \mathrm{Ad}_{\gamma(s)}(v), \mathrm{ad}_{\gamma'(t)\gamma(s)^{-1}}(\mathrm{Ad}_{\gamma(s)}(w))\right\rangle_{\mathfrak{g}} \\
&= 0,
\end{aligned}$$

which is what we wanted.

(♣) Problem T.5. Let G denote a Lie group, and let $\mathfrak{g}$ denote the Lie algebra of G. Let ∇^c denote the connection[4] on G defined by

$$\nabla^c_{X_e}(X_w) = c[X_v, X_w], \qquad \forall\, v, w \in \mathfrak{g}.$$

Let m denote a bi-invariant Riemannian metric on M.

(i) Prove that ∇^c is complete for any $c \in \mathbb{R}$.

(ii) Prove that ∇^c is Riemannian with respect to m for all $c \in \mathbb{R}$.

(iii) Prove that $\nabla^{\frac{1}{2}}$ is torsion-free. *Remark:* This shows that $\nabla^{\frac{1}{2}}$ is the Levi-Civita connection of (G, m).

(iv) Prove that $\nabla^{\frac{1}{2}}$ is right-invariant[5] in the sense that

$$(r_a)_\star(\nabla^{\frac{1}{2}}_X(Y)) = \nabla^{\frac{1}{2}}_{(r_a)_\star X}((r_a)_\star(Y)), \qquad \forall\, X, Y \in \mathfrak{X}(G),\ \forall\, a \in G.$$

(v) Compute the curvature tensor $R^{\nabla^{\frac{1}{2}}}$ of $\nabla^{\frac{1}{2}}$.

Solution. For (i) we first need to find the geodesics. We make an educated guess and try do show that the exponential map

$$\exp\colon \mathfrak{g} \to G$$

gives us the geodesics with respect to the chosen connection ∇^c. Indeed, if we define

$$\gamma(t) = \exp(t \cdot v),$$

then by definition of the exponential map we know that γ is an integral curve of the left-invariant vector field X_v on G. With this we compute

[3] Notice that a priori ad_v is skew-symmetric if and only if the derivative is 0 only for $t = 0$

[4] This is the connection on G given by taking $\beta = c[\cdot, \cdot]$ in Problem P.5.

[5] We already know from Problem P.5 that $\nabla^{\frac{1}{2}}$ is left-invariant.

$$
\begin{aligned}
\nabla^c_{\frac{\partial}{\partial t}}(\gamma')(t) &= \nabla^c_{\frac{\partial}{\partial t}}(X_v \circ \gamma)(t) \\
&\overset{(1)}{=} \nabla^c_{\frac{\partial}{\partial t}\gamma}(X_v)(\gamma(t)) \\
&= \nabla^c_{\gamma'(t)}(X_v)(\gamma(t)) \\
&= \nabla^c_{X_v(\gamma(t))}(X_v)(\gamma(t)) \\
&\overset{(2)}{=} \nabla^c_{X_v}(X_v)(\gamma(t)) \\
&= c[X_v, X_v]_{\gamma(t)} \\
&= 0,
\end{aligned}
$$

where in (1) we used the chain rule for covariant derivatives and in (2) we used the fact that any covariant derivative is a point operator in the first entry. This shows that γ' is a parallel curve along γ, i.e. γ is a geodesic. But $\gamma(t) = \exp(t \cdot v)$ is defined for all $t \in \mathbb{R}$. Now observe that for any given $a \in G$ and $w \in T_a G$ we can define $v := Dl_{a^{-1}}(a)[w]$ and that then

$$\gamma(t) = a \exp(t \cdot v)$$

is still a geodesic satisfying

$$\gamma(0) = a,\ \gamma'(0) = w.$$

By uniqueness of geodesics we thus have found all the geodesic and they are all defined on the whole $\mathbb{R}$ which proves completeness of the connection ∇^c.

For part (ii) we first show that for left-invariant vector fields, say X_u, X_v and X_w with $u, v, w \in \mathfrak{g}$ the Ricci equation is satisfied, i.e.

$$X_u\left(\langle X_v, X_w\rangle\right) = \langle \nabla^c_{X_u} X_v, X_w\rangle + \langle X_v, \nabla^c_{X_u} X_w\rangle,$$

cf. Proposition 36.15. Indeed, an quick computation grants

$$
\begin{aligned}
\langle \nabla^c_{X_u} X_v, X_w\rangle &= c \cdot \langle [X_u, X_v], X_w\rangle \\
&= c \cdot \langle X_{[u,v]}, X_w\rangle \\
&\overset{(1)}{=} c \cdot \langle [u,v], w\rangle_{\mathfrak{g}} \\
&\overset{(2)}{=} c \cdot \langle \mathrm{ad}_u(v), w\rangle_{\mathfrak{g}} \\
&\overset{(3)}{=} -c \cdot \langle v, \mathrm{ad}_u(w)\rangle_{\mathfrak{g}}.
\end{aligned}
$$

Step (1) and (3) are consequences of the previous problem, more precisely (i) and (iii). Step (2) is just the general fact that $\mathrm{ad}_u(v) = [u, v]$ which we have seen last semester. Similarly we see

$$\langle X_v, \nabla^c_{X_u} X_w\rangle = c \cdot \langle v, \mathrm{ad}_u(w)\rangle_{\mathfrak{g}}.$$

Moreover, by the bi-invariance and the previous problem again we obtain that

$$\langle X_v, X_w\rangle = \langle v, w\rangle_{\mathfrak{g}}$$

is a constant smooth function, thus

$$X_u(\langle X_v, X_w\rangle) = 0.$$

In particular this proves the Ricci identity for X_u, X_v, X_w.

The general Ricci equation now follows easily[6]: Any vector field on a Lie group can be written as linear combination of left-invariant vector fields, thus is suffices to consider X_u, $a^i\, X_{v_i}$ and $b^j\, X_{w_j}$ as ∇^c is linear in the first component. Using the Leibniz-rule of ∇^c and the first part of the proof we obtain

$$\begin{aligned}
\langle \nabla^c_{X_u}(a^i\, X_{v_i}), b^j\, X_{w_j}\rangle &= b^j \cdot \langle X_u(a^i)\, X_{v_i} + a^i\, \nabla^c_{X_u}(X_{v_i}), X_{w_j}\rangle && \text{(T.5)}\\
&= (b^j\, X_u(a^i)) \cdot \langle X_{v_i}, X_{w_j}\rangle + (b^j\, a^i) \cdot \langle \nabla^c_{X_u} X_{v_i}, X_{w_j}\rangle && \text{(T.6)}\\
&= (b^j\, X_u(a^i)) \cdot \langle X_{v_i}, X_{w_j}\rangle + (b^j\, a^i) \cdot \langle \nabla^c_{X_u} X_{v_i}, X_{w_j}\rangle && \text{(T.7)}\\
&= (b^j\, X_u(a^i)) \cdot \langle X_{v_i}, X_{w_j}\rangle + (b^j\, a^i) \cdot \langle \mathrm{ad}_u(v_i), w_j\rangle_{\mathfrak{g}} && \text{(T.8)}
\end{aligned}$$

and analogously

$$\langle X_{v_i}, \nabla^c_{x_u}(b^j\, X_{w_j})\rangle = (a^i X_u(b^j))\langle X_{v_i}, X_{w_j}\rangle + (b^j\, a^i) \cdot \langle v_i, \mathrm{ad}_u(w_j))\rangle_{\mathfrak{g}}. \qquad \text{(T.9)}$$

Now adding (T.8) and (T.9) (plus part (iii) in the previous problems) gives

$$(b^j\, X_u(a^i) + a^i\, X_u(b^j)) \cdot \langle X_{v_i}, X_{w_j}\rangle. \qquad \text{(T.10)}$$

Applying the derivation property and invoking the first part of the proof grants

$$\begin{aligned}
X_u\left(\langle a^i\, X_{v_i}, b^j\, X_{w_j}\rangle\right) &= X_u(a^i\, b^j) \cdot \langle X_{v_i}, X_{w_j}\rangle + (a^i\, b^j) \cdot \underbrace{X_u\langle X_{v_i}, X_{w_j}\rangle}_{=0}\\
&= \text{(T.10)},
\end{aligned}$$

which proves the Ricci identity in the general case.

Torsion-freeness (i.e. part (iii)) for $c = \frac{1}{2}$ is trivial as it directly follows from the definition of the connection $\nabla^{\frac{1}{2}}$.

In part (iv) we will mimic the idea from P.5, i.e. we define a new covariant derivative

$$\nabla^{a,\frac{1}{2}}_X(Y) := (r_a)_\star^{-1}\left(\nabla^{\frac{1}{2}}_{(r_a)_\star X}((r_a)_\star(Y)\right)$$

and show that it is the Levi-Civita connection on G with respect to $\langle\cdot,\cdot\rangle$ which then tells us that

$$\nabla^{\frac{1}{2}} = \nabla^{a,\frac{1}{2}}$$

by the Fundamental Theorem of Riemannian Geometry (cf. Theorem 45.1). We first recall two general facts about push-forwards: For any diffeomorphism φ, vector field X and smooth function f we have

- $\varphi_\star Z(f) = Z(f \circ \varphi) \circ \varphi^{-1}$ and equivalently $\varphi_\star Z(f) \circ \varphi = Z(f \circ \varphi)$,
- $(\varphi_\star)^{-1} Z = (\varphi^{-1})_\star Z$.

[6] The amount of TeX-symbols is inversely proportional to the difficulty.

We start by torsion-freeness, i.e.

$$T^{\nabla^{a,\frac{1}{2}}} = 0. \tag{T.11}$$

Indeed, we have

$$\begin{aligned}
T^{\nabla^{a,\frac{1}{2}}}(X,Y) &= \nabla^{a,\frac{1}{2}}_X(Y) - \nabla^{a,\frac{1}{2}}_Y(X) - [X,Y] \\
&= (r_{a^{-1}})_\star \left(\nabla^{\frac{1}{2}}_{(r_a)_\star X}((r_a)_\star(Y)) - \nabla^{\frac{1}{2}}_{(r_a)_\star Y}((r_a)_\star(x))\right) - [X,Y] \\
&= (r_{a^{-1}})_\star \Bigg(\underbrace{[(r_a)_\star X, (r_a)_\star Y]}_{=(r_a)_\star[X,Y]} \Bigg) - [X,Y] \\
&= 0.
\end{aligned}$$

Turning our attention to the Ricci identity we first compute

$$\begin{aligned}
\langle \nabla^{a,\frac{1}{2}}_Z(X), Y\rangle &= \left\langle (r_a)_\star^{-1}\left(\nabla^{\frac{1}{2}}_{(r_a)_\star Z}((r_a)_\star X)\right), (r_a)_\star^{-1}((r_a)_\star Y)\right\rangle && \text{(T.12)} \\
&\overset{(1)}{=} \left\langle \nabla^{\frac{1}{2}}_{(r_a)_\star Z}((r_a)_\star X)\circ r_a, (r_a)_\star Y\circ r_a\right\rangle && \text{(T.13)}
\end{aligned}$$

where in (1) we used the left-invariance of the Riemannian metric. Similarly

$$\langle X, \nabla^{a,\frac{1}{2}}_Z(Y)\rangle = \left\langle (r_a)_\star X\circ r_a, \nabla^{\frac{1}{2}}_{(r_a)_\star Z}((r_a)_\star Y)\circ r_a\right\rangle. \tag{T.14}$$

Adding (T.13) and (T.14) and using the Ricci identity for $\nabla^{\frac{1}{2}}$ gives

$$\begin{aligned}
(r_a)_\star Z\left(\left\langle (r_a)_\star X, (r_a)_\star Y\right\rangle\right)\circ r_a = &\left\langle \nabla^{\frac{1}{2}}_{(r_a)_\star Z}((r_a)_\star X), (r_a)_\star Y\right\rangle\circ r_a \\
&+ \left\langle (r_a)_\star X, \nabla^{\frac{1}{2}}_{(r_a)_\star Z}((r_a)_\star Y)\right\rangle\circ r_a,
\end{aligned}$$

but using the first general formula for push-forwards above plus the bi-invariance of the metric:

$$\begin{aligned}
(r_a)_\star Z\left(\left\langle (r_a)_\star X, (r_a)_\star Y\right\rangle\right)\circ r_a &= Z\left(\left\langle (r_a)_\star X, (r_a)_\star Y\right\rangle\circ r_a\right) \\
&= Z\Big(\left\langle X, Y\right\rangle\circ \underbrace{r_{a^{-1}}\circ r_a}_{=\mathrm{id}}\Big) \\
&= Z\left(\langle X, Y\rangle\right).
\end{aligned}$$

All in all we have shown that

$$Z\left(\langle X,Y\rangle\right) = \langle \nabla^{a,\frac{1}{2}}_Z(X), Y\rangle + \langle X, \nabla^{a,\frac{1}{2}}_Z(Y)\rangle,$$

which as discussed above implies $\nabla^{a,\frac{1}{2}} = \nabla^{\frac{1}{2}}$ and thus concludes the proof of the right-invariance.

For the last part of the exercise (i.e.(v)) we note that it suffices to compute $R^{\nabla^{\frac{1}{2}}}$ for left-invariant vector fields as they span the whole tangent bundle and because $R^{\nabla^{\frac{1}{2}}}$ is a point operator in all entries (see Lecture 33). We have

$$\begin{aligned}
R^{\nabla^{\frac{1}{2}}}(X_v, X_w)X_u &= \nabla^{\frac{1}{2}}_{X_v}\nabla^{\frac{1}{2}}_{X_w}(X_u) - \nabla^{\frac{1}{2}}_{X_w}\nabla^{\frac{1}{2}}_{X_v}(X_u) - \nabla^{\frac{1}{2}}_{[X_v,X_w]}(X_u) \\
&= \frac{1}{4}[X_v,[X_w,X_u]] - \frac{1}{4}[X_w,[X_v,X_u]] - \frac{1}{2}[[X_v,X_w],X_u] \\
&= \frac{1}{4}\left([X_v,[X_w,X_u]] + [X_w,[X_u,X_v]] + 2\cdot[X_u,[X_v,X_w]]\right) \\
&\overset{(1)}{=} \frac{1}{4}\left(-[X_u,[X_v,X_w]] + 2\cdot[X_u,[X_v,X_w]]\right) \\
&= \frac{1}{4}[X_u,[X_v,X_w]],
\end{aligned}$$

where in (1) we used the Jacobi identity for Lie brackets. Now invoking the C^∞-linearity of the curvature tensor we can deduce that for general vector fields X, Y and Z on G one has

$$R^{\nabla^{\frac{1}{2}}}(X,Y)Z = \frac{1}{4}[Z,[X,Y]].$$

Problem Sheet U

PROBLEM U.1. Consider S^n equipped with the metric $m = \imath^\star(m_{\mathrm{Eucl}})$, where $\imath\colon S^n \to \mathbb{R}^{n+1}$ is the inclusion. Prove that the Levi-Civita connection of m is the connection introduced in Problem N.3.

PROBLEM U.2. Let m be a Riemannian metric on M, and let ∇ denote the Levi-Civita connection of m.

(i) Prove that for all $X, Y, Z \in \mathfrak{X}(M)$,

$$\mathcal{L}_X(m)(Y,Z) = \mathcal{L}_X(m)(Y,Z) = \langle \nabla_Y(X), Z\rangle + \langle Y, \nabla_Z(X)\rangle .$$

(ii) We say that a vector field X is a **Killing field** if $\mathcal{L}_X(m) = 0$. Prove that a vector field is a killing field if and only if its maximal flow consists of local isometries.

PROBLEM U.3. Let $\varphi\colon M \to N$ be an isometric map between Riemannian manifolds. Prove that for $x \in M$ the restriction of $(\cdot)^\top$ to $T_{\varphi(x)}N$ is the orthogonal projection onto $D\varphi(x)[T_xM]$.

(♣) PROBLEM U.4. Let $\varphi\colon M \to N$ be a smooth normal covering map and m is a Riemannian metric on M which is invariant under all deck transformations. Prove there is a unique Riemannian metric on N such that φ is a Riemannian covering.

PROBLEM U.5. Let M be a smooth manifold and suppose $\mu\colon G \times M \to M$ is a smooth transitive left action of a Lie group G on M. Fix $x \in M$ and let H denote the isotropy group at x, so that $M \cong G/H$ is a homogeneous space (cf. Theorem 12.11). Let also $\rho\colon H \to \mathrm{GL}(T_xM)$ denote the linear isotropy representation of H (cf. Definition 12.10), so that

$$\rho_a(v) = D\mu_a(e)[v], \qquad a \in H,\ v \in T_xM.$$

Let us say that a Riemannian metric m on M is **invariant** if $\mu_a\colon M \to M$ is an isometry for every $a \in G$. Prove that there is a bijective correspondence between invariant Riemannian metrics on M and inner products on T_xM that are invariant under ρ_a for each $a \in H$.

(♣) PROBLEM U.6. Let M^n be a connected manifold and suppose ∇ is a torsion-free connection on M. Prove that ∇ is the Levi-Civita connection of some Riemannian metric m on M if and only if Hol^∇ is conjugate in $\mathrm{GL}(n)$ to a subgroup of $\mathrm{O}(n)$.

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.

Solutions to Problem Sheet U

Problem U.1. Consider S^n equipped with the metric $m = \imath^\star(m_{\mathrm{Eucl}})$, where $\imath\colon S^n \to \mathbb{R}^{n+1}$ is the inclusion. Prove that the Levi-Civita connection of m is the connection introduced in Problem N.3.

Solution. Let us denote here by $m_{\mathrm{Eucl}} = \langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{R}^{n+1}$ and by $\nabla^{\mathbb{R}^{n+1}}$ the associated Levi-Civita connection (which is just the plain derivative). If ∇ denotes the Levi-Civita connection induced on S^n by $m = \iota^\star m_{\mathrm{Eucl}}$, if X, Y are two vector fields on S^n then by proposition 46.18, there holds

$$\nabla_X Y(p) = (\nabla_X^{\mathbb{R}^{n+1}} Y(p))^\top,$$

where $\nabla^{\mathbb{R}^{n+1}}$ denotes the flat connection on $\mathbb{R}^{n+1}$ and "$\top$" denotes the orthogonal projection from $T_p\mathbb{R}^{n+1}$ onto T_pS^n.

Recall first of all that for any (non-necessarily tangential) vector field along S^n, $Z \in \Gamma_\iota(T\mathbb{R}^{n+1})$, and every tangent vector $X \in T_pS^n$, the expression $\nabla_X^{\mathbb{R}^{n+1}} Y(p)$ is well-defined at $p \in S^n$ and can be computed by taking any extension of X and Z to elements of $\mathfrak{X}(\mathbb{R}^{n+1})$.

Next, recall from Problem N.3 and Problem O.2 we have

$$T_pS^n = \{v \in \mathbb{R}^{n+1}, \langle v, p \rangle = 0\},$$

hence for $w \in T_p\mathbb{R}^{n+1}$ and hence we have $w^\top = w - \langle v, x \rangle x$.

Since the inclusion $\iota\colon S^n \to \mathbb{R}^{n+1}$ can be interpreted as a vector field along S^n, and the identity $\mathrm{id}_{\mathbb{R}^{n+1}}$ is one extension of it to $\mathbb{R}^{n+1}$, we can compute

$$\begin{aligned}
\nabla_X Y(p) &= (\nabla_X^{\mathbb{R}^{n+1}} Y(p))^\top \\
&= \nabla_X^{\mathbb{R}^{n+1}} Y(p) - \langle \nabla_X^{\mathbb{R}^{n+1}} Y(p), p \rangle p \\
&= \nabla_X^{\mathbb{R}^{n+1}} Y(p) - \langle \nabla_X^{\mathbb{R}^{n+1}} Y(p), \mathrm{id}_{\mathbb{R}^{n+1}}(p) \rangle p \\
&= \nabla_X^{\mathbb{R}^{n+1}} Y(p) + \langle Y(p), \nabla_X^{\mathbb{R}^{n+1}}(\mathrm{id}_{\mathbb{R}^{n+1}})(p) \rangle p \qquad \text{(U.1)} \\
&= \nabla_X^{\mathbb{R}^{n+1}} Y(p) + \langle Y(p), X(p) \rangle p,
\end{aligned}$$

where in (U.1) we used the fact that since Y is always orthogonal to $\mathrm{id}_{\mathbb{R}^{n+1}}$ along S^n we have, for every $X \in T_pS^n$,

$$0 \equiv D\langle Y, \mathrm{id}_{\mathbb{R}^{n+1}} \rangle(p)[X] = \langle \nabla_X^{\mathbb{R}^{n+1}} Y(p), p \rangle + \langle Y(p), \nabla_X^{\mathbb{R}^{n+1}}(\mathrm{id}_{\mathbb{R}^{n+1}})(p) \rangle.$$

But, as we proved in Problem O.2 the above expression for ∇_Y is exactly the one for the covariant derivative associated to the connection $\mathcal{H}$ defined in Problem N.3.

Problem U.2. Let m be a Riemannian metric on M, and let ∇ denote the Levi-Civita connection of m.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

(i) Prove that for all $X, Y, Z \in \mathfrak{X}(M)$,

$$\mathcal{L}_X(m)(Y, Z) = \mathcal{L}_X(m)(Y, Z) = \langle \nabla_Y(X), Z\rangle + \langle Y, \nabla_Z(X)\rangle\,.$$

(ii) We say that a vector field X is a **Killing field** if $\mathcal{L}_X(m) = 0$. Prove that a vector field is a killing field if and only if its maximal flow consists of local isometries.

Solution. First of all we observe that the Lie derivative of a Riemannian metric is again a $0, 2$-tensor and by relations derived in Lecture 18 we have

- $\mathcal{L}_X(m)(Y, Z) = X(m(Y, Z)) - m(\mathcal{L}_X(Y), Z) - m(Y, \mathcal{L}_X Z)$,
- $\mathcal{L}_X(Y) = [X, Y]$
- $\mathcal{L}_X(m)(Y, Z) = \lim_{t\to 0} \frac{\varphi_t^*(m)(Y,Z) - m(Y,Z)}{t}$,

where $\varphi\colon D \subseteq \mathbb{R} \times M \to M$ is the maximal flow of X. The first identity together with the Ricci identity gives

$$\mathcal{L}_X(m)(Y, Z) = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle - \langle [X, Y], Z\rangle - \langle Y, [X, Z]\rangle.$$

Thanks to the torsion-freeness of the Levi-Civita connection we also have

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

Now we plug in these two equations into the equation above to obtain

$$\begin{aligned} \mathcal{L}_X(m)(Y, Z) &= \langle [X, Y] + \nabla_Y X, Z\rangle + \langle Y, [X, Z] + \nabla_Z X\rangle - \langle [X, Y], Z\rangle - \langle Y, [X, Z]\rangle \\ &= \langle \nabla_Y X, Z\rangle + \langle Y, \nabla_Z X\rangle. \end{aligned}$$

This proves the desired equality.

For the second part we observe that

$$\mathcal{L}_X(m) = 0 \iff \forall Y, Z \in \mathfrak{X}(M)\colon \lim_{t\to 0} \frac{\varphi_t^*(m)(Y, Z) - m(Y, Z)}{t} = 0,$$

by the third identity stated above. For each $x \in M$ there exists a maximal time interval (t_x^-, t_x^+) on which the integral curve γ_x of X through x is defined, i.e.

$$\gamma_x\colon (t_x^-, t_x^+) \to M,\ \gamma_x(t) = \varphi_t(x),$$

is defined. For $t_x \in (t_x^-, t_x^+)$ near 0 we have

$$\forall Y, Z \in \mathfrak{X}(M)\colon\ (\varphi_t^* m)_x(Y, Z) = m_x(Y, Z)$$

by the formula above and $\varphi_0 = \mathrm{id}_M$. But then due to the group property we actually obtain that

$$(\varphi_t^* m)_x = m_x$$

for all $t \in (t_x^-, t_x^+)$. It is another general fact that φ_t is a local diffeomorphism around x for $t \in (t_x^-, t_x^+)$ (exercise: why?). This proves that the maximal flow of X consist of local isometries.

The converse also holds. Indeed, the proof above can be read backwards.

Problem U.3. Let $\varphi\colon M \to N$ be an isometric map between Riemannian manifolds. Prove that for $x \in M$ the restriction of $(\cdot)^\top$ to $T_{\varphi(x)}N$ is the orthogonal projection onto $D\varphi(x)[T_xM]$.

Solution. Consider $W \in T_{\varphi(x)}N$, then we want to prove that $W^\top$ is the orthogonal projection of W on $D\varphi(x)[T_xM]$. Recall that $W^\top$ is defined as

$$W^\top \coloneqq D\varphi[(\varphi^*(W^\flat))^\sharp].$$

It obviously holds that $W^\top \in D\varphi(x)[T_xM]$ and therefore, to our purpose, it is sufficient to show that $\langle W^\top, V\rangle = \langle W, V\rangle$ for all $V \in D\varphi(x)[T_xM]$.

However this follows easily from the definition. Indeed, given $V \in D\varphi(x)[T_xM]$, let $X \in T_xM$ be such that $V = D\varphi[X]$; then we have that

$$\begin{aligned}\langle W^\top, V\rangle &= \langle D\varphi[(\varphi^*(W^\flat))^\sharp], D\varphi[X]\rangle = \langle (\varphi^*(W^\flat))^\sharp, X\rangle = \varphi^*(W^\flat)(X)\\ &= W^\flat(D\varphi[X]) = \langle W, D\varphi[X]\rangle = \langle W, V\rangle,\end{aligned}$$

which concludes the proof.

(♣) Problem U.4. Let $\varphi\colon M \to N$ be a smooth normal covering map and m is a Riemannian metric on M which is invariant under all deck transformations. Prove there is a unique Riemannian metric on N such that φ is a Riemannian covering.

Solution. Let us denote by $\langle\cdot,\cdot\rangle_M$ the Riemannian metric on M, i.e. m. For any two vector fields $X, Y \in \mathfrak{X}(N)$ we pick the unique φ-related vector fields $\widetilde{X}, \widetilde{Y} \in \mathfrak{X}(M)$, i.e.

$$D\varphi \circ \widetilde{X} = X \circ \varphi \text{ and } D\varphi \circ \widetilde{Y} = Y \circ \varphi.$$

The fact that $\widetilde{X}, \widetilde{Y}$ are unique is an immediate consequence of φ being a covering map and the Inverse Function Theorem. If we assume that there exists *some* Riemannian metric $\langle\cdot,\cdot\rangle_N$ on N that pulls back to m on M under φ, then we must have

$$\langle X(p), Y(p)\rangle_N = \langle \widetilde{X}(\tilde{p}), \widetilde{Y}(\tilde{p})\rangle_M,$$

for any $\tilde{p}$ in the fibre $\varphi^{-1}(p)$. This already proves uniqueness. Now let us take the above equation as a definition of $\langle\cdot,\cdot\rangle_N$ and prove its well-definedness.

We need to check that this definition is independent of the choice of $\tilde{p}$, i.e. we need to verify that for another element $\tilde{q} \in \varphi^{-1}(p)$ we have

$$\langle \widetilde{X}(\tilde{p}), \widetilde{Y}(\tilde{p})\rangle_M = \langle \widetilde{X}(\tilde{q}), \widetilde{Y}(\tilde{q})\rangle_M.$$

For this we observe that there exists a deck transformation $A \in \mathrm{Deck}(\widetilde{M}, \varphi)$ such that

$$A\,\tilde{q} = \tilde{p}^{1}.$$

By deck-transformation invariance of the metric m we have

$$\langle \widetilde{X}(\tilde{q}), \widetilde{Y}(\tilde{q})\rangle_M = \left\langle DA(\tilde{q})\left[\widetilde{X}(\tilde{q})\right], DA(\tilde{q})\left[\widetilde{Y}(\tilde{q})\right]\right\rangle_M.$$

[1]This holds because the covering is normal.

At the same time we can use the deck-transformation property, i.e. $\varphi \circ A = \varphi$ to see that

$$D\varphi(\tilde{p}) \circ DA(\tilde{q}) \left[\widetilde{X}(\tilde{q})\right] = D(\varphi \circ A)(\tilde{q}) \left[\widetilde{X}(\tilde{q})\right] = D\varphi(\tilde{q}) \left[\widetilde{X}(\tilde{q})\right] = X(p),$$

which then proves

$$\widetilde{X}(\tilde{p}) = DA(\tilde{q}) \left[\widetilde{X}(\tilde{q})\right]$$

by uniqueness of $\widetilde{X}$. The same argument works for $\widetilde{Y}$ and thus we obtain

$$\left\langle DA(\tilde{q}) \left[\widetilde{X}(\tilde{q})\right], DA(\tilde{q}) \left[\widetilde{Y}(\tilde{q})\right]\right\rangle_M = \left\langle \widetilde{X}(\tilde{p}), \widetilde{Y}(\tilde{p})\right\rangle_M .$$

This shows that $\langle \cdot, \cdot \rangle_N$ is well defined. Smoothness of $\langle \cdot, \cdot \rangle_N$ is clear. Since $D\varphi$ is an isomorphism at every point (remember, a covering map is a local diffeomorphism around every point!) it is also clear that $\langle \cdot, \cdot \rangle_N$ defines an inner product at every point in N. This concludes the proof.

PROBLEM U.5. Let M be a smooth manifold and suppose $\mu \colon G \times M \to M$ is a smooth transitive left action of a Lie group G on M. Fix $x \in M$ and let H denote the isotropy group at x, so that $M \cong G/H$ is a homogeneous space (cf. Theorem 12.11). Let also $\rho \colon H \to \mathrm{GL}(T_xM)$ denote the linear isotropy representation of H (cf. Definition 12.10), so that

$$\rho_a(v) = D\mu_a(e)[v], \qquad a \in H,\ v \in T_xM.$$

Let us say that a Riemannian metric m on M is **invariant** if $\mu_a \colon M \to M$ is an isometry for every $a \in G$. Prove that there is a bijective correspondence between invariant Riemannian metrics on M and inner products on T_xM that are invariant under ρ_a for each $a \in H$.

SOLUTION. Given an invariant Riemannian metric m on M, its restriction to T_xM is an inner product invariant under ρ_a for each $a \in H$. Indeed, for all $a \in H$ and $v, w \in T_xM$, it holds

$$m_x(\rho_a(v), \rho_a(w)) = m_x(D\mu_a(e)[v], D\mu_a(e)[w]) = m_x(v, w),$$

where we have used that μ_a is an isometry with respect to the metric m.

We want to prove that this map $\mathcal{F} \colon m \mapsto m_x$ from the invariant Riemannian metrics on M to the inner products on T_xM invariant under ρ_a for all $a \in H$ is a bijective correspondence.

Let $\langle \cdot, \cdot \rangle$ be an inner product on T_xM invariant under ρ_a for all $a \in H$. Given $y \in M$, consider $b \in G$ such that $\mu_b(y) = x$ and define m_y as

$$m_y = \mu_b^* \langle \cdot, \cdot \rangle. \tag{U.2}$$

Notice that the definition does not depend on the choice of b. Indeed, if we consider another $c \in G$ with $\mu_c(y) = x$ we have that

$$\mu_c^* \langle \cdot, \cdot \rangle = \mu_{bb^{-1}c}^* \langle \cdot, \cdot \rangle = \mu_b^* \circ \mu_{b^{-1}c}^* \langle \cdot, \cdot \rangle = \mu_b^* \langle \cdot, \cdot \rangle,$$

where we used that $\langle\cdot,\cdot\rangle$ is invariant under $\rho_{b^{-1}c}$, since $b^{-1}c \in H$.

Moreover observe that m is easily a Riemannian metric on M and $m_x = \langle\cdot,\cdot\rangle$. Thus we have only to prove that m is invariant (i.e. $\mathcal{F}$ is surjective) and it is the only invariant metric that coincides with $\langle\cdot,\cdot\rangle$ in x (i.e. $\mathcal{F}$ is injective).

The second assertion is obvious since every invariant metric must satisfy (U.2), thus let us prove the first one. Consider $c \in G$, $y \in M$ and $b \in G$ such that $\mu_b(y) = x$, then we have that

$$\mu_c^* m_y = \mu_c^* \circ \mu_b^* \langle\cdot,\cdot\rangle = \mu_{cb}^* \langle\cdot,\cdot\rangle = m_{\mu_c(y)},$$

which proves μ_c is an isometry, as we wanted.

(♣) Problem U.6. Let M^n be a connected manifold and suppose ∇ is a torsion-free connection on M. Prove that ∇ is the Levi-Civita connection of some Riemannian metric m on M if and only if Hol^∇ is conjugate in $\mathrm{GL}(n)$ to a subgroup of $\mathrm{O}(n)$.

Solution. Fix $x \in M$. If ∇ is the Levi-Civita connection of a Riemannian metric m, then we claim that the parallel transport along any loop γ based at x gives a linear isometry of T_xM. Indeed, the pullback connection on γ^*TM is still Riemannian and thus satisfies the Ricci identity, which gives

$$\frac{d}{dt}\langle \mathbb{P}_\gamma(v), \mathbb{P}_\gamma(w)\rangle = 0$$

whenever $v, w \in T_xM$, implying that $\langle v, w\rangle = \langle \widehat{\mathbb{P}}_\gamma(v), \widehat{\mathbb{P}}_\gamma(w)\rangle$, as claimed. Letting $A\colon \mathbb{R}^n \to T_xM$ be an orthonormal frame at x, this means that $A^{-1} \circ \widehat{\mathbb{P}}_\gamma \circ A$ belongs to $\mathrm{O}(n)$. It follows that $\mathrm{Hol}^\nabla(x; A) \subseteq \mathrm{O}(n)$. Since M is connected, the claim follows from Corollary 32.12.

Conversely, we can find (and fix in the sequel) a frame A at x such that $\mathrm{Hol}^\nabla(x; A) \subseteq \mathrm{O}(n)$. Given $y \in M$ and a piecewise smooth curve γ from x to y, we define a (positive definite) inner product m_y on T_yM by declaring that $\{\widehat{\mathbb{P}}_\gamma \circ A(e_i)\}$ is an orthonormal basis at y. The definition of m_y does not depend on γ: if δ is another such curve, then

$$\widehat{\mathbb{P}}_\delta \circ A = \widehat{\mathbb{P}}_\gamma \circ A \circ (A^{-1} \circ \widehat{\mathbb{P}}_{\delta * \gamma^-} \circ A)$$

and $\delta * \gamma^-$ is a loop based at x; thus,

$$\widehat{\mathbb{P}}_\delta \circ A(e_j) = O^i_j\, \widehat{\mathbb{P}}_\gamma \circ A(e_i),$$

with $O := A^{-1} \circ \widehat{\mathbb{P}}_{\delta * \gamma^-} \circ A \in \mathrm{Hol}^\nabla(x; A) \subseteq \mathrm{O}(n)$. Since O is an orthogonal transformation, this shows that m_y is well defined.

The smoothness of $y \mapsto m_y$ follows as in Problem O.1. Indeed, given $y \in M$, we can find a diffeomorphism $\psi\colon V \to U$ with $V \subseteq \mathbb{R}^n$ open and starshaped about 0 and U an open neighbourhood of y. Given $v \in V$, we let $\gamma_v(t) := \psi(tv)$ (for $t \in [0, 1]$, so that γ_v joins y to $\psi(v)$). For $p \in U$ we set

$$e_i(p) := \widehat{\mathbb{P}}_{\alpha * \gamma_{\psi^{-1}(p)}} \circ A(e_i),$$

where α is a curve from x to y chosen in advance. By definition of m_p, $\{e_i(\cdot)\}$ defines a local orthonormal frame on U, which is smooth by Problem O.1. It follows that also the metric m is smooth on U, hence everywhere (as y was arbitrary).

We claim that the parallel transport along any curve $\delta\colon [0,1] \to M$, joining $y = \delta(0)$ to $z = \delta(1)$, gives a linear isometry from T_yM to T_zM: indeed, let γ be a curve from x to y and note that $\widehat{\mathbb{P}}_\gamma$ and $\widehat{\mathbb{P}}_{\gamma*\delta}$ are both linear isometries by construction of m (from T_xM to T_yM and T_zM, respectively). Since $\widehat{\mathbb{P}}_{\gamma*\delta} = \widehat{\mathbb{P}}_\delta \circ \widehat{\mathbb{P}}_\gamma$, we deduce that $\widehat{\mathbb{P}}_\delta\colon T_yM \to T_zM$ is a linear isometry, as well.

Since ∇ is torsion-free, in order to conclude it suffices to show that ∇ is Riemannian. Given vector fields $X, Y, Z \in \mathfrak{X}(M)$ and a point $p \in M$, take a curve γ with $\gamma'(0) = X(p)$ and fix an orthonormal frame $\{e_i(t)\}$ parallel along γ. Writing $Y(\gamma(t)) = Y^i(t)e_i(t)$ (and similarly for Z), we have

$$\begin{aligned} X(\langle Y, Z\rangle)(p) &= \frac{d}{dt}\langle Y(\gamma(t)), Z(\gamma(t))\rangle\Big|_{t=0} \\ &= \frac{d}{dt}(Y^i(t)Z^i(t))\Big|_{t=0} \\ &= (Y^i)'(0)Z^i(0) + Y^i(0)(Z^i)'(0). \end{aligned}$$

On the other hand, calling T the vector $\frac{\partial}{\partial t}$ at 0, the chain rule (31.7) gives

$$\begin{aligned} \langle \nabla_X Y, Z\rangle(p) &= \langle \nabla_T(Y^i e_i), Z^j(0)e_j(0)\rangle \\ &= (Y^i)'(0)Z_i(0) + \langle Y^i(0)\nabla_T e_i, Z^j(0)e_j(0)\rangle \\ &= (Y^i)'(0)Z_i(0), \end{aligned}$$

being $e_i(t)$ parallel along γ. Similarly, $\langle \nabla_X Y, Z\rangle(p) = Y^i(0)(Z^i)'(0)$. We deduce that the Ricci identity holds, which concludes the proof in view of Proposition 36.15.

Problem Sheet V

Problem V.1. Let (M, m) be an oriented Riemannian manifold with Riemannian volume form μ. Prove that any orientation-preserving isometry $\varphi\colon (M, m) \to (M, m)$ is volume-preserving in the sense of Definition 48.5, i.e. that $\varphi^\star(\mu_m) = \mu_m$. Give an example to show that the converse is false.

Problem V.2. Let (M, m) be a compact oriented Riemannian manifold with boundary. Let μ_m denote the Riemannian volume form. Let $N \in \Gamma(TM|_{\partial M})$ denote an outward pointing vector field[1] such that $|N(x)| = 1$ for all $x \in \partial M$. The metric m restricts to a metric on ∂M, and $i_N(\mu_m)$ is the Riemannian volume form of $(\partial M, m)$ when ∂M is given the induced orientation (Definition 21.21). Prove that

$$\int_{M,m} \operatorname{div}_m(X) = \int_{\partial M,m} \langle X, N \rangle, \qquad \forall\, X \in \mathfrak{X}(M).$$

Remark: This is the generalisation of the Divergence Theorem 48.10 to Riemannian manifolds with boundary.

Problem V.3. Let (M^3, m) be an oriented Riemannian manifold of dimension three with Riemannian volume form μ_m. Define

$$\chi\colon \mathfrak{X}(M) \to \Omega^2(M), \qquad \chi(X) := i_X(\mu_m).$$

We define the **curl** operator as

$$\operatorname{curl}_m\colon \mathfrak{X}(M) \to \mathfrak{X}(M), \qquad \operatorname{curl}_m(X) := \chi^{-1} \circ dX^\flat,$$

where $X^\flat \in \Omega^1(M)$ is the one-form obtained from X via the musical isomorphism.

(i) Let $\tau\colon C^\infty(M) \to \Omega^3(M)$ denote the map $f \mapsto f\,\mu_m$. Prove that the following diagram commutes:

$$\begin{tikzcd}
C^\infty(M) \arrow[r, "\operatorname{grad}_m"] \arrow[d, "\mathrm{id}"'] & \mathfrak{X}(M) \arrow[r, "\operatorname{curl}_m"] \arrow[d, "\flat"] & \mathfrak{X}(M) \arrow[r, "\operatorname{div}_m"] \arrow[d, "\chi"] & C^\infty(M) \arrow[d, "\tau"] \\
C^\infty(M) \arrow[r, "d"'] & \Omega^1(M) \arrow[r, "d"'] & \Omega^2(M) \arrow[r, "d"'] & \Omega^3(M)
\end{tikzcd}$$

(ii) Deduce that

$$\operatorname{div}_m(\operatorname{curl}_m(X)) = 0, \qquad \forall\, X \in \mathfrak{X}(M),$$

a formula you no doubt remember from Analysis II.

Problem V.4. Let M be a manifold of dimension two or three.

[1]Such a vector field N exists by normalising the vector field found in Lemma 21.20.

(i) Prove that the curvature tensor $\mathcal{R}^{\nabla}_m$ is completely determined by the Ricci tensor Ric_m.

(ii) Prove that a Riemannian metric m on M is Einstein if and only if m has constant curvature.

(♣) Problem V.5. Let (M, m) be a Riemannian manifold. Let $\kappa \colon TTM \to TM$ denote the connection map of the Levi-Civita connection of m, and let Θ_t denote the geodesic flow of m (i.e. the flow of the geodesic spray of the Levi-Civita connection of M). Using Lemma 31.3 we can regard $D\Theta_t$ as a vector bundle isomorphism

$$\widetilde{D\Theta_t} := (D\pi, \kappa) \circ D\Theta_t \circ (D\pi, \kappa)^{-1} \colon TM \oplus TM \to TM \oplus TM$$

along $\pi \colon TM \to M$. Fix $x \in M$ and $u, v, w \in T_x M$. Let γ denote the unique geodesic with $\gamma(0) = x$ and $\gamma'(0) = u$, and let $c \in \mathrm{Jac}(\gamma)$ denote the unique Jacobi field along γ with

$$c(0) = v, \qquad \nabla_T(c)(0) = w.$$

Prove that

$$\widetilde{D\Theta_t}(x, u)(v, w) = (c(t), c'(t)).$$

(♣) Problem V.6. Let $\pi \colon E \to M$ be a vector bundle, and assume we are given a connections ∇^E on E and a connection ∇^M on M. Let $\nabla^{E,M}$ denote the induced connection on the bundle $E \otimes T^*M$ over M. If $s \in \Gamma(E)$ then $\nabla^E(s) \in \Gamma(E \otimes T^*M)$, and hence $\nabla^{E,M}(\nabla(s)) \in \Gamma(E \otimes T^*M \otimes T^*M)$. We abbreviate

$$\nabla^{(2)}(s) := \nabla^{E,M}(\nabla^E(s))$$

and call the operator $\nabla^{(2)} \colon \Gamma(E) \to \Gamma(E \otimes T^*M \otimes T^*M)$ a **higher covariant derivative**.

(i) Prove that

$$\nabla^{E,M}(\nabla^E(s))(X, Y) = \nabla^E_X(\nabla^E_Y(s)) - \nabla^E_{\nabla^M_X(Y)}(s),$$

(ii) Suppose that ∇^M is torsion-free. Prove that

$$\nabla^{(2)}(s)(X, Y) - \nabla^{(2)}(s)(Y, X) = R^{\nabla^E}(X, Y)(s).$$

(iii) We can repeat this process inductively to define covariant derivatives

$$\nabla^{(k)} \colon \Gamma(E) \to \Gamma(E \otimes T^*M^{\otimes k})$$

Find an explicit formula for $\nabla^{(k)}$ in terms of ∇^E and ∇^M.

(iv) Now suppose (M, m_1) and (N, m_2) are two Riemannian manifolds. Let ∇^M and ∇^N denote the Levi-Civita connections. Suppose $\varphi \colon M \to N$ is a smooth map. Then $D\varphi$ can be thought of as a section of the pullback bundle $\varphi^\star TN \to TM$. By taking $E = \varphi^\star TN$ and applying the construction above, we can define higher order covariant derivatives of $D\varphi$. Give an explicit formula in terms of local coordinates on M and N (and the Christoffel symbols of ∇^M and ∇^N) for these higher order covariant derivatives.

(v) Now take $(N, m_2) = (\mathbb{R}, m_{\mathrm{Eucl}})$, so that φ is a simply a smooth function on M. How does the second order covariant derivative compare to the Hessian?

Solutions to Problem Sheet V

Problem V.1. Let (M, m) be an oriented Riemannian manifold with Riemannian volume form μ. Prove that any orientation-preserving isometry $\varphi\colon (M, m) \to (M, m)$ is volume-preserving in the sense of Definition 48.5, i.e. that $\varphi^\star(\mu_m) = \mu_m$. Give an example to show that the converse is false.

Solution. Let us assume that M is connected. Then we know that the pullback $\varphi^*(\mu_m)$ is a non-zero multiple of the volume form μ_m at every point. But for any $x \in M$ and any positive orthonormal basis $(v_1, \dots, v_n)$ of $(T_x M, m_x)$ we have, by assumption, that $(D\varphi(x)[v_1], \dots, D\varphi(x)[v_n])$ is a positive orthonormal basis of $(T_{\varphi(x)}M, m_{\varphi(x)})$, thus

$$\varphi^*(\mu_m)(x)(v_1, \dots, v_n) = \mu_m(\varphi(x))(D\varphi(x)[v_1], \dots, D\varphi(x)[v_n]) = 1,$$

which proves $\varphi^*(\mu_m) = \mu_m$.

To show that the converse does not hold we pick $M = \mathbb{R}^2$ with its flat metric m, i.e. $m = m_{\mathrm{Eucl}}$ is just the Euclidean scalar product at each point. We observe that a linear transformation

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2$$

is volume-preserving if and only if

$$\det(T) = 1.$$

Indeed, we have $\mu_{m_{\mathrm{Eucl}}}(\cdot,\cdot) = e^1 \wedge e^2$[1] by the normalisation condition and therefore

$$T^*(\mu_{\mathrm{Eucl}})(e_1, e_2) = \det(T).$$

Similarly, it is easy to see that T preserves the metric m_{Eucl} if and only if T is orthogonal with respect to the Euclidean inner product. Hence, the linear transformation

$$T = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

defines a volume-preserving diffeomorphism on $(\mathbb{R}^2, m_{\mathrm{Eucl}})$ which is not isometric.

Problem V.2. Let (M, m) be a compact oriented Riemannian manifold with boundary. Let μ_m denote the Riemannian volume form. Let $N \in \Gamma(TM|_{\partial M})$ denote an outward pointing vector field[2] such that $|N(x)| = 1$ for all $x \in \partial M$. The metric m restricts to a metric on ∂M, and $i_N(\mu_m)$ is the Riemannian volume form of $(\partial M, m)$ when ∂M is given the induced orientation (Definition 21.21). Prove that

$$\int_{M,m} \mathrm{div}_m(X) = \int_{\partial M, m} \langle X, N \rangle, \qquad \forall\, X \in \mathfrak{X}(M).$$

Remark: This is the generalisation of the Divergence Theorem 48.10 to Riemannian manifolds with boundary.

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

[1] Here we are suppressing the $\mathcal{J}$-maps.

[2] Such a vector field N exists by normalising the vector field found in Lemma 21.20.

Solution. As in the proof of Theorem 48.10, we apply Stokes' Theorem:

$$\int_{M,m} \operatorname{div}_m(X) = \int_M d(i_X(\mu_m)) = \int_{\partial M} i_X(\mu_m).$$

Since the right-hand side of the desired identity equals $\int_{\partial M} \langle X, N \rangle \, i_N(\mu_m)$, it suffices to show that

$$i_X(\mu_m) = \langle X, N \rangle \, i_N(\mu_m)$$

pointwise on ∂M. To see this, write $X = \langle X, N \rangle N + Y$, so that Y is orthogonal to N and thus $Y(p) \in T_p \partial M$, for all $p \in \partial M$. In particular Y, which is a section of $\iota^*(TM)$ ($\iota \colon \partial M \hookrightarrow M$ being the inclusion), can be seen as a vector field on ∂M. We have

$$i_X(\mu_m) = \langle X, N \rangle \, i_N(\mu_m) + i_Y(\mu_m).$$

Finally, $i_Y(\mu_m)$ vanishes as a differential form on ∂M: indeed,

$$\iota^*(i_Y(\mu_m)) = i_Y(\iota^*(\mu_m)) = i_Y(0) = 0,$$

since an n-form necessarily vanishes on an $(n-1)$-dimensional manifold.

Problem V.3. Let (M^3, m) be an oriented Riemannian manifold of dimension three with Riemannian volume form μ_m. Define

$$\chi \colon \mathfrak{X}(M) \to \Omega^2(M), \qquad \chi(X) \coloneqq i_X(\mu_m).$$

We define the **curl** operator as

$$\operatorname{curl}_m \colon \mathfrak{X}(M) \to \mathfrak{X}(M), \qquad \operatorname{curl}_m(X) \coloneqq \chi^{-1} \circ dX^\flat,$$

where $X^\flat \in \Omega^1(M)$ is the one-form obtained from X via the musical isomorphism.

(i) Let $\tau \colon C^\infty(M) \to \Omega^3(M)$ denote the map $f \mapsto f \, \mu_m$. Prove that the following diagram commutes:

$$\begin{tikzcd}
C^\infty(M) \arrow[r, "\operatorname{grad}_m"] \arrow[d, "\mathrm{id}"'] & \mathfrak{X}(M) \arrow[r, "\operatorname{curl}_m"] \arrow[d, "\flat"] & \mathfrak{X}(M) \arrow[r, "\operatorname{div}_m"] \arrow[d, "\chi"] & C^\infty(M) \arrow[d, "\tau"] \\
C^\infty(M) \arrow[r, "d"'] & \Omega^1(M) \arrow[r, "d"'] & \Omega^2(M) \arrow[r, "d"'] & \Omega^3(M)
\end{tikzcd}$$

(ii) Deduce that

$$\operatorname{div}_m(\operatorname{curl}_m(X)) = 0, \qquad \forall\, X \in \mathfrak{X}(M),$$

a formula you no doubt remember from Analysis II.

Solution.

(i) It is sufficient to prove that the three square diagrams commute. The first one commutes if and only if $df = (\operatorname{grad}_{\mathrm{m}} f)^\flat$ for all $f \in C^\infty(M)$, which follows directly from the definition of gradient (Definition 48.12).

The commutation of the second square is equivalent to $d(X^\flat) = \chi(\mathrm{curl}_m(X))$ for all $X \in \mathfrak{X}(M)$, which is exactly the definition of curl_m given in the text of the problem.

Finally the third square commutes if and only if for all $X \in \mathfrak{X}(M)$ it holds

$$\mathrm{div}_m(X)\mu_m = \tau(\mathrm{div}_m(X)) = d(\chi(X)) = d(i_X\mu_m),$$

which is implied by Cartan's Magic Formula (Theorem 20.6) as in the proof of Theorem 48.10.

Therefore we have proven the commutation of the diagram.

(ii) Given $X \in \mathfrak{X}(M)$, by the commutation of the diagram in the first part of the problem, we know that

$$\mathrm{div}_m(\mathrm{curl}_m(X))\mu_m = \tau(\mathrm{div}_m(\mathrm{curl}_m(X))) = dd(X^\flat) = 0,$$

which implies that $\mathrm{div}_m(\mathrm{curl}_m(X)) = 0$ as we wanted.

Problem V.4. Let M be a manifold of dimension two or three.

(i) Prove that the curvature tensor $\mathcal{R}_m^\nabla$ is completely determined by the Ricci tensor Ric_m.

(ii) Prove that a Riemannian metric m on M is Einstein if and only if m has constant curvature.

Solution.

(i) We wish to invoke Corollary 47.10; hence, we have to show that the Ricci tensor determines all the sectional curvatures (in dimension two or three). In general, pick $x \in M$ and an orthonormal basis $\{e_1, \dots, e_n\}$ of T_xM. For the Levi-Civita connection ∇, the formula derived below Definition 49.8 gives

$$\mathrm{Ric}_m(e_j, e_j) = \sum_{i=1}^n R_m^\nabla(e_i, e_j, e_i, e_j) = \sum_{i|i\neq j} \mathrm{sect}_m(x; \mathrm{span}\{e_i, e_j\}) \tag{V.1}$$

for all $j = 1, \dots, n$ (as $R_m^\nabla(e_j, e_j, e_j, e_j) = 0$ by the symmetries of the Riemann tensor). In particular, if the dimension is two we get

$$\mathrm{sect}_m(x; T_xM) = \mathrm{sect}_m(x; \mathrm{span}\{e_1, e_2\}) = \mathrm{Ric}_m(e_2, e_2).$$

In dimension three, summing (V.1) over j, we get

$$\begin{aligned}\sum_{j=1}^3 \mathrm{Ric}_m(e_j, e_j) &= \sum_{(i,j)|i\neq j} \mathrm{sect}_m(x; \mathrm{span}\{e_i, e_j\}) \\ &= 2\,\mathrm{sect}_m(x; \mathrm{span}\{e_1, e_2\}) + 2\,\mathrm{sect}_m(x; \mathrm{span}\{e_1, e_3\}) \\ &\quad + 2\,\mathrm{sect}_m(x; \mathrm{span}\{e_2, e_3\}).\end{aligned}$$

But then

$$\sum_{j=1}^{3} \operatorname{Ric}_m(e_j, e_j) - 2\operatorname{Ric}_m(e_3, e_3) = 2\operatorname{sect}_m(x; \operatorname{span}\{e_1, e_2\}),$$

so the sectional curvature of $\operatorname{span}\{e_1, e_2\}$ is determined by the Ricci tensor. As the orthonormal basis was arbitrary, it follows that all sectional curvatures are determined by the Ricci tensor.

(ii) Assume that $\operatorname{Ric}_m = \lambda m$. The argument used for the first part shows actually the following:[3] given a (real) vector space V of dimension $n \in \{2, 3\}$ with an inner product, fix an orthonormal basis $\{e_1, \ldots, e_n\}$ and consider two quadrilinear maps R_1, R_2 as in Lemma 47.9; if

$$\sum_{i=1}^{n} R_1(e_i, v, e_i, v) = \sum_{i=1}^{n} R_2(e_i, v, e_i, v) \tag{V.2}$$

for all $v \in V$, then $R_1 = R_2$. Given $x \in M$, we now apply this fact with $V := T_xM$ and $R_1 := R^{\nabla}_m|_x$, $R_2 := \frac{\lambda}{n-1}\mathcal{S}_m|_x$ (as in the proof of Corollary 47.12). Observe that

$$\sum_{i=1}^{n} \mathcal{S}_m|_x(e_i, v, e_i, v) = n\langle v, v\rangle - \sum_{i=1}^{n} \langle v, e_i\rangle^2 = (n-1)\langle v, v\rangle,$$

while

$$\sum_{i=1}^{n} R_1(e_i, v, e_i, v) = \operatorname{Ric}_m(v, v) = \lambda\langle v, v\rangle$$

by hypothesis. Thus (V.2) holds and we get $R_1 = R_2$, which evidently implies that m has constant curvature $\frac{\lambda}{n-1}$.

(♣) PROBLEM V.5. Let (M, m) be a Riemannian manifold. Let $\kappa\colon TTM \to TM$ denote the connection map of the Levi-Civita connection of m, and let Θ_t denote the geodesic flow of m (i.e. the flow of the geodesic spray of the Levi-Civita connection of M). Using Lemma 31.3 we can regard $D\Theta_t$ as a vector bundle isomorphism

$$\widetilde{D\Theta_t} := (D\pi, \kappa) \circ D\Theta_t \circ (D\pi, \kappa)^{-1}\colon TM \oplus TM \to TM \oplus TM$$

along $\pi\colon TM \to M$. Fix $x \in M$ and $u, v, w \in T_xM$. Let γ denote the unique geodesic with $\gamma(0) = x$ and $\gamma'(0) = u$, and let $c \in \operatorname{Jac}(\gamma)$ denote the unique Jacobi field along γ with

$$c(0) = v, \qquad \nabla_T(c)(0) = w.$$

Prove that

$$\widetilde{D\Theta_t}(x, u)(v, w) = (c(t), \nabla_T(c)(t)).$$

[3] The proof in this situation uses Lemma 47.9 in place of Corollary 47.10.

Solution. Set
$$\xi = (D\pi, \kappa)^{-1}(v, w) \in T_u TM$$
and define the vector field
$$\tilde{c}(t) = D\pi \circ D\Theta_t[\xi] \in T_{\pi(\Theta_t(x,u))}M.$$
Recall that as a function of t, the projection $\pi(\Theta_t(x,u)) \in M$ is the unique geodesic with starting point x and derivative u. By uniqueness of geodesics this means that $\pi(\Theta_t(x,u))$ is precisely γ and $\tilde{c}$ is therefore a vector field along γ.

We will shortly show that $\tilde{c}$ is a Jacobi field along γ with initial data
$$\tilde{c}(0) = v \text{ and } \nabla_T(\tilde{c})(0) = w.$$
This will then imply by Proposition 50.8 that $\tilde{c} = c$ and consequently prove half of the statement.

In order to see why $\tilde{c}$ is a Jacobi field we will write it as a geodesic variation of γ (cf. Proposition 50.13). For this we pick a smooth curve $\eta\colon (-\epsilon, \epsilon) \to TM$ such that
$$\eta(0) = u,\ \eta'(0) = \xi.$$
By the chain rule we thus have
$$\tilde{c}(t) = \frac{\partial}{\partial s}\Big|_{s=0} \underbrace{\exp_{\pi(\eta(s))}(t \cdot \eta(s))}_{=:\Gamma(s,t)}.$$
It is clear that $\Gamma(0,t) = \gamma(t)$, hence $\tilde{c}$ is a Jacobi field along γ. For the initial data we compute

- $\tilde{c}(0) = \frac{\partial}{\partial s}\big|_{s=0}(\pi(\eta(s)) = D\pi(\xi) = v,$
- $\frac{\partial}{\partial t}\big|_{t=0}\tilde{c}(t) = \frac{\partial}{\partial s}\big|_{s=0}\eta(s) = \xi$

where in the second bullet point we first of all used the fact that we can swap T and $S = \frac{\partial}{\partial s}$ just as in the proof of Proposition 50.13, and
$$D(\exp_y)(0_y)[\mathcal{J}_{0_Y}(\eta(s))]) = \mathrm{id}_{T_yM},$$
where $y = \pi(\eta(s))$ (cf. Theorem 43.3). This proves
$$D\pi\left(D\theta_t(x,u)[\xi]\right) = c(t).$$
We are only left to show that
$$\kappa(D\Theta_t(x,u)[\xi]) = \nabla_T(c)(t).$$
Using again the fact that ∇ is torsion-free, $[S,T] = 0$ and $\Theta_t(\eta(s)) = \frac{\partial}{\partial t}\exp_{\pi(\eta(s))}(t \cdot \eta(s))$ (see end of Lecture 42) one gets
$$\begin{aligned}
D\Theta_t(x,u)[\xi] &= \frac{\partial}{\partial s}\Big|_{s=0}(\theta_t(\eta(s))) \\
&= \frac{\partial}{\partial t}\left(\frac{\partial}{\partial s}\Big|_{s=0}\exp_{\pi(\eta(s))}(t \cdot \eta(s))\right) \\
&= \frac{\partial}{\partial t}\tilde{c}(t) \\
&= D\tilde{c}(t)[T].
\end{aligned}$$

And finally we obtain

$$\kappa\left(D\Theta_t(x,u)[\xi]\right) = \kappa(D\tilde{c}(t)[T]) = \nabla_T(\tilde{c})(t).$$

Here we used the very first definition of a covariant derivative via the connection map κ (see Theorem 31.10). This finishes the proof.

(♣) PROBLEM V.6. Let $\pi\colon E \to M$ be a vector bundle, and assume we are given a connections ∇^E on E and a connection ∇^M on M. Let $\nabla^{E,M}$ denote the induced connection on the bundle $E \otimes T^*M$ over M. If $s \in \Gamma(E)$ then $\nabla^E(s) \in \Gamma(E \otimes T^*M)$, and hence $\nabla^{E,M}(\nabla(s)) \in \Gamma(E \otimes T^*M \otimes T^*M)$. We abbreviate

$$\nabla^{(2)}(s) := \nabla^{E,M}(\nabla^E(s))$$

and call the operator $\nabla^{(2)}\colon \Gamma(E) \to \Gamma(E \otimes T^*M \otimes T^*M)$ a **higher covariant derivative**.

(i) Prove that

$$\nabla^{E,M}(\nabla^E(s))(X,Y) = \nabla^E_X(\nabla^E_Y(s)) - \nabla^E_{\nabla^M_X(Y)}(s),$$

(ii) Suppose that ∇^M is torsion-free. Prove that

$$\nabla^{(2)}(s)(X,Y) - \nabla^{(2)}(s)(Y,X) = R^{\nabla^E}(X,Y)(s).$$

(iii) We can repeat this process inductively to define covariant derivatives

$$\nabla^{(k)}\colon \Gamma(E) \to \Gamma(E \otimes T^*M^{\otimes k})$$

Find an explicit formula for $\nabla^{(k)}$ in terms of ∇^E and ∇^M.

(iv) Now suppose (M, m_1) and (N, m_2) are two Riemannian manifolds. Let ∇^M and ∇^N denote the Levi-Civita connections. Suppose $\varphi\colon M \to N$ is a smooth map. Then $D\varphi$ can be thought of as a section of the bundle $\varphi^\star TN \otimes T^*M$. By taking $E = \varphi^\star TN$ and applying the construction above, we can define higher order covariant derivatives of $D\varphi$ via:

$$\nabla^{(2)}(\varphi) := \nabla^{E,M}(D\varphi),$$

and inductively $\nabla^{(k)}(\varphi)$ is a section of $\varphi^\star(TN) \otimes T^*M^{\otimes k}$. Give an explicit formula for $\nabla^{(k)}(\varphi)$. Find a local expression for $\nabla^{(2)}(\varphi)$ in terms of local coordinates on M and N (and the Christoffel symbols of ∇^M and ∇^N) for $\nabla^{(2)}(\varphi)$.

(v) Now take $(N, m_2) = (\mathbb{R}, m_{\mathrm{Eucl}})$, so that φ is a simply a smooth function on M. How does the second order covariant derivative compare to the Hessian?

SOLUTION.

(i) First, recall that $\nabla^{E,M}$ acts on a section $e \otimes \omega \in \Gamma(E \otimes T^*M)$ as

$$\nabla_X^{E,M}(e \otimes \omega) = \nabla_X^E e \otimes \omega + e \otimes \nabla_X^M \omega,$$

where the covariant differentiation on forms is, as we know,

$$\nabla_X^M \omega(Y) = X(\omega(Y)) - \omega(\nabla_X^M Y).$$

Hence, if $e_1, \dots e_k$ is a fixed local frame for E defined on some open subset $U \subset M$ and s is a section of the bundle E, there exist k uniquely determined 1-forms $\omega_1, \dots, \omega_k$ so that

$$\nabla^E s = e_i \otimes \omega^i$$

on U (if we write $s = s^i e_i$, the explicit expression for such forms is $\omega^i(X) = X(s^i) + s^j \omega_j^i(X)$ where ω_j^i are the connection 1-forms). We may then compute:

$$\begin{aligned}
\nabla^{E,M}(\nabla^E s)(X_1, X_2) &= \nabla_{X_1}^{E,M}(\nabla^E s)(X_2) \\
&= \nabla_{X_1}^{E,M}(e_i \otimes \omega^i)(X_2) \\
&= \nabla_{X_1}^E e_i \, \omega^i(X_2) + e_i \, (\nabla_{X_1}^M \omega_i) X_2 \\
&= \nabla_{X_1}^E e_i \, \omega^i(X_2) + e_i \big(X_1(\omega^i(X_2)) - \omega^i(\nabla_{X_1}^M X_2)\big) \\
&= \nabla_{X_1}^E(\nabla_{X_1}^E s) - \nabla_{\nabla_{X_1}^M X_2}^E s,
\end{aligned}$$

where we used, for the last inequality, the fact that covariant derivatives satisfy the Leibniz rule.

(ii) Using (i) and adding and subtracting the term $\nabla_{[X_1,X_2]}^E s$, we see that

$$\begin{aligned}
\nabla^{(2)} s(X_1, X_2) - \nabla^{(2)} s(X_2, X_1) &= \nabla_{X_1}^E(\nabla_{X_2}^E s) - \nabla_{X_2}^E(\nabla_{X_1}^E s) - \Big(\nabla_{\nabla_{X_1}^M X_2}^E s - \nabla_{\nabla_{X_2}^M X_1}^E s\Big) \\
&= \nabla_{X_1}^E(\nabla_{X_2}^E s) - \nabla_{X_2}^E(\nabla_{X_1}^E s) - \nabla_{[X_1,X_2]}^E s - \nabla_{T(X_1,X_2)}^E s \\
&= R^{\nabla^E}(X_1, X_2) s,
\end{aligned}$$

where T is the torsion tensor of ∇^M which is identically zero by assumption.

(iii) As in (i), recall first that, for $k \in \mathbb{N}$, the induced connection on $E \otimes T^*M^{\otimes k}$ is

$$\nabla_X^{E,M}(e \otimes \eta) = \nabla_X^E e \otimes \eta + e \otimes \nabla_X^M \eta,$$

where the covariant differentiation on k-forms is

$$\nabla_X^M \eta(Y_1, \dots, Y_k) = X(\eta(Y_1, \dots, Y_k)) - \sum_{j=1}^k \eta(X_1, \dots, \nabla_X^M Y_j, \dots, Y_k).$$

If $e_1, \dots e_k$ is a fixed local frame for E and s is a section of E, we may locally write

$$\nabla^{(k)} s = e_i \otimes \eta^i,$$

where $\eta_1, \dots \eta_k$ are k-forms uniquely determined by the frame. We then compute, for $k \geq 1$:

$$\begin{aligned}
\nabla^{(k+1)} s(X_1, \dots, X_{k+1}) &= \nabla^{E,M}_{X_1}(\nabla^k s)(X_2, \dots, X_{k+1}) \\
&= \nabla^{E,M}_{X_1}(e_i \otimes \eta^i)(X_2, \dots X_{k+1}) \\
&= \nabla^E_{X_1} e_i \, \eta^i(X_2, \dots, X_{k+1}) + e_i \, (\nabla^M_{X_1} \eta_i)(X_2, \dots, X_{k+1}) \\
&= \nabla^E_{X_1} e_i \, \eta^i(X_2, \dots, X_{k+1}) \\
&\quad + e_i \Big(X_1(\eta^i(X_2, \dots, X_{k+1})) - \sum_{j=2}^{k+1} \eta^i(X_1, \dots, \nabla^M_X Y_j, \dots, Y_k) \Big),
\end{aligned}$$

So we conclude that a (recursive) formula for higher-order covariant differentiation is given by

$$\begin{aligned}
\nabla^{(k+1)} s(X_1, \dots, X_{k+1}) &= \nabla^E_{X_1}\big(\nabla^{(k)} s(X_2, \dots, X_{k+1})\big) \\
&\quad - \sum_{j=2}^{k+1} \nabla^{(k)} s(X_2, \dots, \nabla^M_{X_1} X_j, \dots, X_{k+1}).
\end{aligned}$$

(iv) Here, the bundle is $E = \varphi^\star(TN)$ with covariant derivative $\nabla^E_X = \nabla^N_{D\varphi[X]}$ and we regard $D\varphi$ as a section of the bundle $\varphi^\star(TN) \otimes T^*M$, and we put

$$\nabla^{(2)} \varphi(X_1, X_2) \coloneqq \nabla^{E,M}(D\varphi)(X_1, X_2) = \nabla^{E,M}_{X_1}(D\varphi)(X_2).$$

Similarly as in (i) we deduce that

$$\begin{aligned}
\nabla^{(2)} \varphi(X_1, X_2) &= \nabla^E_{X_1}(D\varphi[X_2]) - D\varphi[\nabla^M_{X_1} X_2] \\
&= \nabla^N_{D\varphi[X_1]}(D\varphi[X_2]) - D\varphi[\nabla^M_{X_1} X_2]
\end{aligned}$$

The higher-order derivatives $\nabla^{(k)} \varphi$ are defined inductively as in (iii), and as in that case we have the formula, for $k \geq 1$,

$$\begin{aligned}
\nabla^{(k)} \varphi(X_1, \dots, X_k) &= \nabla^E_{X_1}(\nabla^{(k-1)} \varphi(X_2, \dots, X_k)) \\
&\quad - \sum_{j=2}^{k} \nabla^{(k-1)} \varphi(X_2, \dots, \nabla^M_{X_1} X_j, \dots, X_k)
\end{aligned}$$

We find the local expression for $\nabla^{(2)} \varphi$. If $x^1, \dots, x^m$ are local coordinates of M and $y^1, \dots, y^n$ are local coordinates on N, and we denote with $\Gamma^\sigma_{\mu\nu}$ and γ^k_{ij} the correspondent Christoffel symbols on M and N respectively, by the formula we have

$$\begin{aligned}
\nabla^{(2)} \varphi(\partial_\mu, \partial_\nu) &= \nabla^N_{\partial_\mu \varphi}(\partial_\nu \varphi) - D\varphi[\nabla^M_{\partial_\mu} \partial_\nu] \\
&= \nabla^N_{\partial_\mu \varphi}(\partial_\nu \varphi^i \partial_i) - D\varphi[\Gamma^\sigma_{\mu\nu} \partial_\sigma] \\
&= \partial^2_{\mu\nu} \varphi^i \partial_i|_\varphi + \partial_\mu \varphi^j \partial_\nu \varphi^i \gamma^k_{ji}(\varphi) \partial_k|_\varphi - \Gamma^\sigma_{\mu\nu} \partial_\sigma \varphi^i \partial_i|_\varphi.
\end{aligned}$$

This means that the local components of $\nabla^{(2)} \varphi$ are

$$(\nabla^{(2)} \varphi)^k_{\mu\nu}(x) = \partial^2_{\mu\nu} \varphi^k + \gamma^k_{ij}(\varphi(x)) \partial_\mu \varphi^i(x) \partial_\nu \varphi^j(x) - \Gamma^\sigma_{\mu\nu}(x) \partial_\sigma \varphi^k(x).$$

(v) In the particular case where $N = \mathbb{R}$ with its standard metric, we see from the expressions in (iv) that we recover precisely the notion of Hessian for a real-valued map given in Definition 48.20.

Problem Sheet W

(♣) Problem W.1. Prove that the converse to Proposition 53.10 does not hold: Find a Riemannian manifold (M, m) with a geodesic γ that does not have conjugate points yet is not minimal.

(♣) Problem W.2. Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds of the same dimension. Suppose (M, m_1) is complete. Prove that φ is a Riemannian covering.

Problem W.3. Let (M, m) be a Riemannian manifold. Let $\gamma\colon [0, b] \to M$ be a regular geodesic. A point $t_0 \in (0, b]$ is a conjugate point of γ if and only if $\exp_{\gamma(0)}$ does not have maximal rank at $t_0\gamma'(0)$. In fact,

$$\dim \ker D\exp_{\gamma(0)}(t_0\gamma'(0)) = \dim \operatorname{null}(\gamma|_{[0,t_0]}).$$

Problem W.4. Let (M, m) be a Riemannian manifold. Let $\gamma \in \mathcal{C}_{xy}([a, b])$ be a geodesic and let $c_1, c_2 \in T_\gamma\mathcal{C}_{xy}([a, b])$. Prove that

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1, c_2) = -\int_a^b \big\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1, \gamma')(\gamma'), c_2 \big\rangle\, dt.$$

If instead we only require $c_1, c_2 \in T_\gamma\mathcal{P}_{xy}([a, b])$, prove that

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1, c_2) = &-\int_a^b \big\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1, \gamma')(\gamma'), c_2 \big\rangle\, dt \\ &+ \sum_{i=1}^{k-1} \big\langle \nabla_T(c_1)|_{[a_{i-1},a_i]}(a_i) - \nabla_T(c_1)|_{[a_i,a_{i+1}]}(a_i), c_2(a_i) \big\rangle,\end{aligned}$$

where $a = a_0 < a_1 < \cdots < a_k = b$ is any subdivision of $[a, b]$ such that $c_1|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \dots, k$.

(♣) Problem W.5. Let (M, m) and $(\tilde{M}, \tilde{m})$ be two Riemannian manifolds of the same dimension n with associated exponential maps $\exp$ and $\widetilde{\exp}$. Fix $x \in M$ and $\tilde{x} \in \tilde{M}$, and let $v \in T_xM$ and $\tilde{v} \in T_{\tilde{x}}\tilde{M}$ be two vectors of unit norm. Let $\gamma(t) := \exp_x(tv)$ and $\tilde{\gamma}(t) := \widetilde{\exp}_{\tilde{x}}(t\tilde{v})$. Let $b > 0$ be such that both γ and $\tilde{\gamma}$ are defined on $[0, b]$, and let $y = \gamma(b)$, $\tilde{y} = \tilde{\gamma}(b)$. Suppose that for all $t \in [0, b]$ and for all 2-planes $\Pi \subset T_{\gamma(t)}M$ one has

$$\operatorname{sect}_m(\gamma(t); \Pi) \le \operatorname{sect}_{\tilde{m}}\big(\tilde{\gamma}(t); T_t[\Pi]\big).$$

Then for all $c \in T_\gamma\mathcal{P}_{xy}([0, b])$ one has

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c, c) \ge \operatorname{Hess}(\mathbb{E}_{\tilde{m}})(\tilde{\gamma})(\tau(c), \tau(c)),$$

where $\tau\colon T_\gamma\mathcal{P}_{xy}([0, b]) \to T_{\tilde{\gamma}}\mathcal{P}_{\tilde{x}\tilde{y}}([0, b])$ was defined in (53.8).

Solutions to Problem Sheet W

(♣) Problem W.1. Prove that the converse to Proposition 53.10 does not hold: Find a Riemannian manifold (M, m) with a geodesic γ that does not have conjugate points yet is not minimal.

Solution. Consider the flat torus (T^2, m_{flat}), i.e. the torus equipped with its unique Riemannian metric such that the universal covering

$$\pi\colon (\mathbb{R}^2, m_{\text{Eucl}}) \to (T^2, m_{\text{flat}})$$

defines a Riemannian covering (see Example 46.10). We know that $\mathbb{R}^2$ equipped with its Levi-Civita connection is flat, in particular its sectional curvature is 0. It immediately follows from $\pi\colon \mathbb{R}^2 \to T^2$ being a Riemannian covering that (T^2, m_{flat}) with its Levi-Civita connection is also flat. Thus there are no conjugate points since (T^2, m_{flat}) has zero constant curvature (see Remark 50.10). Also, the projection is an isometric map and therefore sends geodesics to geodesics. In particular, the projection of straight lines on the flat torus are geodesics.

We consider the (closed) geodesic

$$\gamma\colon [0,3] \to T^2,\ \gamma(t) = \left(t, \frac{1}{3}t\right).$$

The distance between the two points

$$x = \gamma\left(\frac{1}{2}\right) = \left(\frac{1}{2}, \frac{1}{6}\right),\ y = \gamma\left(\frac{5}{2}\right) = \left(\frac{1}{2}, \frac{5}{6}\right),$$

is given by

$$d(x,y) = \sqrt{0^2 + \left(\frac{2}{3}\right)^2} = \frac{2}{3}.$$

The length of γ restricted to $\left[\frac{1}{2}, \frac{5}{2}\right]$ however is

$$\int_{\frac{1}{2}}^{\frac{5}{2}} |\dot\gamma(t)| dt = \int_{\frac{1}{2}}^{\frac{5}{2}} \sqrt{1 + \frac{1}{9}} dt = \frac{2\sqrt{10}}{3},$$

which is strictly greater than the distance $d(x,y)$, hence γ is *not* minimal. This concludes the proof [1].

Solutions written by Giada Franz, Francesco Palmurella, Alessio Pellegrini, and Alessandro Pigati.
Last modified: June 28, 2019.

[1] Note that this does not contradict Corollary 53.9 and Corollary 53.6: The geodesic we have chosen just happens not to be a minimal geodesic between x and y.

(♣) PROBLEM W.2. Let $\varphi\colon (M, m_1) \to (N, m_2)$ be an isometric map between Riemannian manifolds of the same dimension. Suppose (M, m_1) is complete. Prove that φ is a Riemannian covering.

SOLUTION. Let us proceed by steps.

(i). *Preliminary: existence and uniqueness of the "geodesic lift".* We show that, for every $q \in \varphi(M)$, every $p \in \varphi^{-1}(q)$ and every geodesic σ in N starting at q, there exists a unique geodesic $\tilde{\sigma}$ in M starting at p so that $\sigma = \varphi \circ \tilde{\sigma}$. Indeed, since φ is a local isometry, it maps geodesics to geodesics, hence if we consider the unique geodesic $\tilde{\sigma}$ in M starting at p and with initial velocity $w = D\varphi(p)^{-1}[\sigma'(0)]$, this must coincide, by the existence and uniqueness theorem for geodesics,s with σ.

(ii). *N is complete.* Indeed, pick a point $q \in \varphi(M)$ and any tangent vector $w \in T_qN$ and consider the unique geodesic σ starting at q with initial velocity w. By (i), there exists a geodesic $\tilde{\sigma}$ in M so that $\sigma = \varphi \circ \tilde{\sigma}$. Since M is complete, $\tilde{\sigma}$ is defined on the whole $\mathbb{R}$, and consequently so is σ. Since v was arbitrary, then N is complete by the Hopf-Rinow Theorem 53.7.

(iii). *φ is surjective.* Indeed, let q be any point in N, and let $q_0 = \varphi(p_0)$ be a point in $\varphi(M)$. Since N is complete, there exists a length-minimising geodesic σ with $\sigma(0) = q_0$ and $\sigma(1) = q$. Then by (i) we have $\sigma = \varphi \circ \tilde{\sigma}$ for some geodesic $\tilde{\sigma}$ in M, and since both are complete, necessarily than $q = \varphi(\tilde{\sigma}(1))$ and so $q \in \varphi(M)$.

(iv). *φ is a covering map.* We further divide this step in three sub-steps.
(iv.a). Let $q_0 \in N$ be fixed, and let $U = B^N_\varepsilon(q_0)$ be the geodesic ball in N of radius ε centred at q_0. Set $\varphi^{-1}(q_0) = \{q_\alpha\}_{\alpha \in A}$ and let us verify that, for every $\alpha \neq \beta$, there holds $B^M_\varepsilon(p_\alpha) \cap B^M_\varepsilon(p_\beta) = \varnothing$. Indeed, if $\tilde{\sigma}$ is a length-minimising geodesic from p_α to p_β, then we consider, by (i), the geodesic (in N) $\sigma = \varphi \circ \tilde{\sigma}$ which is a loop at q_0. Since all geodesics in $B^N_\varepsilon(q_0)$ are radial by Corollary 52.8, σ must leave, and then return, the ball $B^N_\varepsilon(p_0)$, and so it must have length bigger than 2ε. Now since φ is a local isometry it preserves the length of curves, and thus the length of $\tilde{\sigma}$ in M must be bigger than 2ε. By the triangle inequality applied to d^M is then follows that $U_\alpha \cap U_\beta = \varnothing$.
(iv.b). Let us prove that $\varphi^{-1}(U) = \bigcup_\alpha U_\alpha$. For the inclusion "$\subseteq$", if $p \in \varphi^{-1}(U)$ then $\varphi(p) \in U = B^N_\varepsilon(p_0)$, hence a length-minimising geodesic σ from $\varphi(p)$ to q_0 lies in U and has length $\delta < \varepsilon$. The corresponding geodesic $\tilde{\sigma}$ in M given by (i) joins p with p_α for some α and has length $\delta < \varepsilon$, hence $p \in U_\alpha$. hence $p \in U_\alpha$. For the inclusion "$\supseteq$", again since φ is an isometry, it preserves the length of curves, and so (since the Riemannian distance is an infimum over all curves joining the two points) it contracts the distances: $d^N(\varphi(p_1), \varphi(p_2)) \leq d^M(p_1, p_2)$, and in particular, since $\varphi(p_\alpha) = q_0$ and each U_α is a ball of radius ε, its image via φ will be contained in U.

(v). *$\varphi|_{U_\alpha}\colon U_\alpha \to U$ is a diffeomorphism for every α.* Since φ sends the radial geodesic from p_α tangent to $v \in T_{p_\alpha}M$ into the radial geodesic from q_0 with initial velocity $D\varphi(p_\alpha)[v] \in T_{q_0}N$, it follows that $\varphi|_{U_\alpha} = \exp_{q_0} \circ D\varphi_{p_\alpha} \circ (\exp_{p_\alpha}|_{B_\varepsilon(0)})^{-1}$, and that such map is invertible with $(\varphi|_{U_\alpha})^{-1} = \exp_{p_\alpha} \circ (D\varphi_{p_\alpha})^{-1} \circ (\exp_{p_\alpha}|_{B_\varepsilon(0)})^{-1}$.

PROBLEM W.3. Let (M, m) be a Riemannian manifold. Let $\gamma\colon [0, b] \to M$ be a regular geodesic. A point $t_0 \in (0, b]$ is a conjugate point of γ if and only if $\exp_{\gamma(0)}$ does not have maximal rank at $t_0\gamma'(0)$. In fact,

$$\dim \ker D\exp_{\gamma(0)}(t_0\gamma'(0)) = \dim \operatorname{null}(\gamma|_{[0,t_0]}).$$

Solution. Let $x := \gamma(0)$. Given a vector $v \in T_xM$, the vector $\mathcal{J}_{t_0\gamma'(0)}(v) \in T_{t_0\gamma'(0)}(T_xM)$ belongs to the kernel of $D\exp_x(t_0\gamma'(0))$ if and only if the same holds for $t_0\mathcal{J}_{t_0\gamma'(0)}(v)$, as $t_0 > 0$.

Call c_v the Jacobi field along γ with initial conditions $c_v(0) = 0_x$, $\nabla_T(c_v)(0) = v$. By Corollary 50.15 we have

$$c_v(t_0) = D\exp_x(t_0\gamma'(0))[t_0\mathcal{J}_{t_0\gamma'(0)}(v)],$$

so $t_0\mathcal{J}_{t_0\gamma'(0)}(v)$ belongs to the kernel of $D\exp_x(t_0\gamma'(0))$ if and only if $c_v(t_0) = 0$. Since the assignment $v \mapsto c_v$ is a linear bijection between T_xM and Jacobi fields (along γ) vanishing at the initial time 0, the claim follows.

Problem W.4. Let (M, m) be a Riemannian manifold. Let $\gamma \in \mathcal{C}_{xy}([a, b])$ be a geodesic and let $c_1, c_2 \in T_\gamma\mathcal{C}_{xy}([a, b])$. Prove that

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1, c_2) = -\int_a^b \left\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1, \gamma')(\gamma'), c_2 \right\rangle dt.$$

If instead we only require $c_1, c_2 \in T_\gamma\mathcal{P}_{xy}([a, b])$, prove that

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1, c_2) = &-\int_a^b \left\langle \nabla_T(\nabla_T(c_1)) + R^\nabla(c_1, \gamma')(\gamma'), c_2 \right\rangle dt \\ &+ \sum_{i=1}^{k-1} \left\langle \nabla_T(c_1)|_{[a_{i-1},a_i]}(a_i) - \nabla_T(c_1)|_{[a_i,a_{i+1}]}(a_i), c_2(a_i) \right\rangle,\end{aligned}$$

where $a = a_0 < a_1 < \cdots < a_k = b$ is any subdivision of $[a, b]$ such that $c_1|_{[a_{i-1},a_i]}$ is smooth for each $i = 1, \ldots, k$.

Solution. Consider the variation

$$\Gamma(r, s, t) = \exp_{\gamma(t)}(r \cdot c_1(t) + s \cdot c_2(t)).$$

For the first identity we first of all observe that by definition of the Hessian (cf. Definition 51.15) and the first bit of the proof of Proposition 51.13 one has

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c_1, c_2) = -\int_a^b R\left\langle \partial_s\Gamma, \nabla_T(\partial_t\Gamma)\right\rangle dt,$$

where here R, S and T denote the obvious vector fields on $\mathbb{R}$. Now we start massaging the integrand above:

$$R\left\langle \partial_s\Gamma, \nabla_T(\partial_t\Gamma))\right\rangle = \underbrace{\left\langle \nabla_R(\partial_s\Gamma), \nabla_T(\partial_t\Gamma)\right\rangle}_{(A)} + \underbrace{\left\langle \partial_s\Gamma, \nabla_R\nabla_T(\partial_t\Gamma)\right\rangle}_{(B)}.$$

By choice of our variation we have

$$\partial_r\Gamma(0, 0, t) = c_1(t),\ \partial_s\Gamma(0, 0, t) = c_2(t),\ \partial_t\Gamma(0, 0, t) = \gamma'(t).$$

Therefore (A) vanishes as γ is a geodesic by assumption, thus $\nabla_T(\partial_t\Gamma) = 0$. For (B) on the other hand we have

$$(B) = \langle \partial_s\Gamma, R^\nabla(R, T)(\partial t\Gamma) + \nabla_T\nabla_R(\partial_t\Gamma) + \nabla_{[T,R]}(\partial_t\Gamma)\rangle.$$

The last term is 0 since $[T, R] = 0$[2]. Also note that above we have been implicitly using the pullback connection $\gamma^*\nabla$ which we still called ∇ by abuse of notation. Following this abuse of notation, the chain rule for covariant derivatives and plugging in $s = r = 0$ finally grants

$$\begin{aligned}(B) &= \langle c_2, R^\nabla(c_1, \gamma')(\gamma') + \nabla_T(\nabla_R(\partial_t\Gamma))\rangle \\ &= \langle c_2, R^\nabla(c_1, \gamma')(\gamma') + \nabla_T(\nabla_T(c_1))\rangle.\end{aligned}$$

In the last line we used torsion-freeness of the Levi-Civita connection and $[R, T] = 0$. This proves the first identity.

Now for the second identity we observe again by referring to the proof of the first variation formula (cf. proof of Proposition 51.13) that the Hessian has an additonal error term

$$(C) := \int_a^b R\left(T\langle\partial_s\Gamma, \partial_t\Gamma\rangle\right) dt.$$

We can swap R and T, apply the Ricci identity and obtain

$$(C) = \underbrace{\int_a^b T\langle\nabla_R(\partial_s\Gamma), \gamma'\rangle dt}_{(C')} + \underbrace{\int_a^b T\langle c_2, \nabla_R(\gamma')\rangle dt}_{(C'')}.$$

For (C'') we observe that

$$\nabla_R(\gamma'(t))(0) = \nabla_T(\partial_r\Gamma(0, 0, \cdot))(t) = \nabla_T(c_1)(t),$$

which is by assumption a discontinuous expression at the points $a = a_0 < a_1 < \cdots < a_k = b$, and therefore we end up with

$$(C'') = \int_{i=1}^{k-1} \left\langle \nabla_T(c_1)\big|_{[a_{i-1},a_i]}(a_i) - \nabla_T(c_1)\big|_{[a_i,a_{i+1}]}(a_i), c_2(a_i)\right\rangle.$$

For (C') on the other hand we observe that

$$\nabla_R(\partial_s\Gamma(\cdot, 0, t))(0) \text{ and } \partial_s\Gamma(r, 0, t) = D\exp_{\gamma(t)}(r \cdot c_1(t))[c_2(t)].$$

The second term is a smooth function in r, and so is the first then. Therefore (C'') is given by the difference of $\langle\nabla_R(\partial_s\Gamma), \gamma'\rangle$ evaluated at $t = a, b$[3], but by properness of c_2 this is just 0. This proves the second identity.

(♣) Problem W.5. Let (M, m) and $(\tilde{M}, \tilde{m})$ be two Riemannian manifolds of the same dimension n with associated exponential maps $\exp$ and $\widetilde{\exp}$. Fix $x \in M$ and $\tilde{x} \in \tilde{M}$, and let $v \in T_xM$ and $\tilde{v} \in T_{\tilde{x}}\tilde{M}$ be two vectors of unit norm. Let $\gamma(t) := \exp_x(tv)$ and $\tilde{\gamma}(t) := \widetilde{\exp}_{\tilde{x}}(t\tilde{v})$. Let $b > 0$ be such that both γ and $\tilde{\gamma}$ are defined on $[0, b]$, and let $y = \gamma(b)$, $\tilde{y} = \tilde{\gamma}(b)$. Suppose that for all $t \in [0, b]$ and for all 2-planes $\Pi \subset T_{\gamma(t)}M$ one has

$$\mathrm{sect}_m(\gamma(t); \Pi) \le \mathrm{sect}_{\tilde{m}}\left(\tilde{\gamma}(t); T_t[\Pi]\right).$$

Then for all $c \in T_\gamma\mathcal{P}_{xy}([0, b])$ one has

$$\mathrm{Hess}(\mathbb{E}_m)(\gamma)(c, c) \ge \mathrm{Hess}(\mathbb{E}_{\tilde{m}})(\tilde{\gamma})(\tau(c), \tau(c)),$$

where $\tau\colon T_\gamma\mathcal{P}_{xy}([0, b]) \to T_{\tilde{\gamma}}\mathcal{P}_{\tilde{x}\tilde{y}}([0, b])$ was defined in (53.8).

[2] Here we cannot use the fact that γ is a geodesic. Exercise: Why?

[3] Note that γ' is smooth as we are assuming $\gamma \in \mathcal{C}_{xy}([a, b])$ even though $c_1, c_2 \in T_\gamma\mathcal{P}_{xy}([a, b])$.

Solution. Given $c \in T_\gamma \mathcal{P}_{xy}([0,b])$, we can write it as $c = f^i e_i$ where $\{\gamma' = e_1, e_2, \dots, e_n\}$ is a parallel orthonormal frame along γ and $f^i \coloneqq \langle c, e_i\rangle \colon [0,b] \to \mathbb{R}$ are piecewise smooth functions. Calling $0 = a_0 < a_1 < \cdots < a_k = b$ a subdivision of $[0,b]$ such that $c|_{[a_{j-1},a_j]}$ (and thus $f^i|_{[a_{j-1},a_j]}$ for $i = 1, \dots, n$) is smooth for each $j = 1, \dots, k$, by Proposition 51.16 (proven in Problem W.4) we have that

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c,c) = &- \int_0^b \left\langle \nabla_T(\nabla_T(c)) + R^\nabla(c,\gamma')(\gamma'), c\right\rangle dt \\ &+ \sum_{j=1}^{k-1} \left\langle \nabla_T(c)|_{[a_{j-1},a_j]}(a_j) - \nabla_T(c)|_{[a_j,a_{j+1}]}(a_j), c(a_j)\right\rangle .\end{aligned}$$

Now observe that $\nabla_T(c) = \nabla_T(f^i e_i) = (f^i)' e_i$ and analogously $\nabla_T(\nabla_T(c)) = (f^i)'' e_i$, since e_i are parallel. In a point a_j of discontinuity of c, let us denote $(f^i)'_-(a_j)$ the left derivative (computed taking $t \nearrow a_i$) and $(f^i)'_+$ right derivative (computed taking $t \searrow a_i$). Then we obtain that

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c,c) &= - \int_0^b \left\langle (f^i)'' e_i + f^i R^\nabla(e_i,\gamma')(\gamma'), f^i e_i\right\rangle dt \\ &\quad + \sum_{i=1}^{k-1} \left\langle (f^i)'_-(a_i) e_i - (f^i)'_+(a_i) e_i, f^i(a_i) e_i\right\rangle \\ &= - \int_0^b \sum_{i=1}^n \left[(f^i)'' + (f^i)^2 \sec_m(\gamma(t); \operatorname{span}(\gamma', e_i))\right] dt \\ &\quad + \sum_{j=1}^{k-1} \left[\sum_{i=1}^n (f^i)'_-(a_j) f^i(a_j) - (f^i)'_+(a_j) f^i(a_j)\right].\end{aligned} \tag{W.1}$$

However, defining $\tilde{e}_i(t) \coloneqq T_t(e_i(t))$, $\{\tilde{\gamma}' = \tilde{e}_1, \dots, \tilde{e}_n\}$ is a parallel orthonormal frame along $\tilde{\gamma}$ and $\tau(c) = f^i \tilde{e}_i$. Therefore, with the same computations, we have that

$$\begin{aligned}\operatorname{Hess}(\mathbb{E}_{\tilde{m}})(\tilde{\gamma})(\tau(c),\tau(c)) = &- \int_0^b \sum_{i=1}^n \left[(f^i)'' + (f^i)^2 \sec_{\tilde{m}}(\tilde{\gamma}(t); \operatorname{span}(\tilde{\gamma}', \tilde{e}_i))\right] dt \\ &+ \sum_{j=1}^{k-1} \left[\sum_{i=1}^n (f^i)'_-(a_j) f^i(a_j) - (f^i)'_+(a_j) f^i(a_j)\right].\end{aligned} \tag{W.2}$$

Observe that, by hypothesis, it holds

$$\sec_m(\gamma(t); \operatorname{span}(\gamma', e_i)) \le \sec_{\tilde{m}}(\tilde{\gamma}(t); \operatorname{span}(\tilde{\gamma}', \tilde{e}_i)),$$

thus comparing (W.1) and (W.2) gives exactly the sought inequality

$$\operatorname{Hess}(\mathbb{E}_m)(\gamma)(c,c) \ge \operatorname{Hess}(\mathbb{E}_{\tilde{m}})(\tilde{\gamma})(\tau(c),\tau(c)).$$

Problem Sheet XYZ

Problem XYZ.1. Something easy to get started:

(i) Let M be a simply connected compact topological manifold of dimension three. Prove that M is homeomorphic to S^3. *Hint:* If you get stuck, try here.

(ii) Now let M be a simply connected compact smooth manifold of dimension four. Is M necessarily diffeomorphic to S^4?

Problem XYZ.2. Infinite-dimensional manifolds are fun. For each result proved in Lectures 1–53, decide whether the statement holds for infinite-dimensional manifolds. If so, prove it (adding additional hypotheses where needed). If not, construct a counterexample.

Problem XYZ.3. Prove that there does not exist a manifold M with $\dim M = \pi$.

Problem XYZ.4. Go surfing.

Problem XYZ.5. Study hard for your Differential Geometry II exam, and enjoy your summer. (In that order.)

Will J. Merry, Diff. Geometry II, Spring 2019, ETH Zürich. Last modified: June 28, 2019.