# Differential Geometry

by Will Merry

Lecture notes based on the Differential Geometry course lectured by Prof. P.M.H. Wilson in Michaelmas term 2007 for Part III of the Cambridge Mathematical Tripos.

## Contents


| | | |
|---|---|---|
| 1 | Smooth manifolds | 1 |
| 2 | Vector fields | 9 |
| 3 | Submanifolds | 14 |
| 4 | Lie groups | 18 |
| 5 | Vector bundles and sheaves | 24 |
| 6 | Differential forms and cohomology | 35 |
| 7 | Integration on manifolds and Lie derivatives | 44 |
| 8 | Connections on vector bundles | 49 |
| 9 | Koszul connections | 57 |
| 10 | Elementary Riemannian geometry | 66 |


Please let me know on **w.merry@dpmms.cam.ac.uk** if you find any errors - I'm sure there are many!

## 1 Smooth manifolds

### 1.1 Definitions

A **topological manifold** $M$ of dimension $n$ is a second countable Hausdorff topological space that is locally homeomorphic to $\mathbb{R}^n$; that is, for any $p \in M$ there exists an open neighborhood $U$ of $p$ and a homeomorphism $h : U \to \mathcal{O} \subseteq \mathbb{R}^n$, where $\mathcal{O}$ is open in $\mathbb{R}^n$. We call the homeomorphism $h : U \to \mathcal{O}$ a **chart**, and we call $U$ a **coordinate neighborhood** of $p$. We write $M^n$ to signify that the dimension of $M$ is $n$.

### 1.2 Charts

Let $r^i : \mathbb{R}^n \to \mathbb{R}$ denote projection onto the $i$th coordinate, and given a chart $h : U \to \mathcal{O}$, let $x^i = r^i \circ h : U \to \mathbb{R}$. The functions $x^i$ are the **coordinates** of $h$ on $U$. Given two charts $h : U \to \mathcal{O}$ and $k : V \to \Omega$ on a topological manifold $M$ such that $U \cap V \neq \emptyset$, the function

$$k \circ h^{-1} : h(U \cap V) \to k(U \cap V)$$

is a homeomorphism between open sets of $\mathbb{R}^n$. We call $k \circ h^{-1}$ the **transition function** between the charts $h$ and $k$.

### 1.3 Definitions

An **atlas** $\mathcal{A}$ on a topological manifold $M^n$ is a collection $\{(U_\alpha, h_\alpha) \mid \alpha \in A\}$ of charts on $M$ such that:

1. $\{U_\alpha \mid \alpha \in A\}$ is an open cover of $M$,
2. For any $\alpha, \beta \in A$ such that $U_\alpha \cap U_\beta \neq \emptyset$, the transition function $h_{\alpha\beta} := h_\alpha \circ h_\beta^{-1}$ is smooth (posesses continuous partial deriviatives of all orders).

In general, we say two charts satisfying the conditions of (2) above are **compatible**. Note that this implies that the transition function between the two charts is a diffeomorphism.

A **smooth structure** on $M$ is an atlas $\mathcal{A}$ that is maximal with respect to property (2) above.

## 1.4 Lemma

Any atlas determines a unique smooth structure.

◀ Let $\mathcal{A} = \{(U_\alpha, h_\alpha) | \alpha \in A\}$ be an atlas on $M$, and consider the collection of charts $\mathcal{A}^*$ consisting of all charts on $M$ that are compatible with all the charts in $\mathcal{A}$. Clearly $\mathcal{A} \subseteq \mathcal{A}^*$. We now check $\mathcal{A}^*$ is an atlas. Indeed, property (1) is immediate, and to check property (2), let $h : U \to \mathcal{O}$ and $k : V \to \Omega$ be two charts in $\mathcal{A}^*$ such that $U \cap V \neq \emptyset$. We must show that $k \circ h^{-1}$ is smooth, and it is enough to check this locally. Given $p \in U \cap V$, choose a chart $h_\alpha : U_\alpha \to \mathcal{O}_\alpha$ in $\mathcal{A}$ such that $p \in U_\alpha$. Then $W = U \cap V \cap U_\alpha$ is an open neighborhood of $p$ and

$$k \circ h^{-1}\big|_{h(W)} = k \circ h^{-1}\big|_{\phi(W)} = (k \circ h_\alpha^{-1}) \circ (h_\alpha \circ h^{-1}) : h(W) \to k(W)$$

is smooth by assumption. Thus $\mathcal{A}^*$ is an atlas, and clearly $\mathcal{A}^*$ is maximal amongst atlases containing $\mathcal{A}$. ▶

Therefore it is enough to specify an atlas when defining a smooth structure on a topological manifold, and we will do this without further comment.

## 1.5 Definition

A **smooth manifold of dimension** $n$ is an $n$-dimensional topological manifold $M$ equipped with a smooth structure.

## 1.6 Examples

1. Any open subset $V$ of a smooth manifold $M^n$ is itself a smooth manifold; second countability, the Hausdorff condition and the locally Euclidean property are inherited by subsets, and if $\{(U_\alpha, h_\alpha) | \alpha \in A\}$ is an atlas on $M$, then $\{(U_\alpha \cap V, h_\alpha|_{U_\alpha \cap V}) | \alpha \in A\}$ forms an atlas on $V$.

2. $\mathbb{R}^n$ is Hausdorff and second countable (for a countable base, take all balls of rational radii with rational centres), and certainly locally Euclidean. The identity map determines an atlas on $\mathbb{R}^n$ (there are no transition functions to worry about) and smooth structure determined by this atlas makes $\mathbb{R}^n$ into a smooth $n$-manifold. We call this the **standard** smooth structure on $\mathbb{R}^n$.

3. The $n$-dimensional sphere $S^n \subseteq \mathbb{R}^{n+1}$ is a $n$-dimensional smooth manifold. Indeed, for $i = 1 \ldots n+1$, let $U_i^+ := \{y \in S^n \mid y^i > 0\}$ and similarly let $U_i^-$ denote the set of points in $S^n$ whose $i$th coordinate is negative. Note that $\{U_i^\pm\}$ is an open cover of $S^n$. Now define $h_i : U_i^\pm \to \mathbb{R}^n$ to be the map that forgets the $i$th coordinate; $h_i(y^1, \ldots, y^{n+1}) = (y^1, \ldots, \hat{y^i}, \ldots, y^{n+1})$. Suppose $i < j$; then the transition function

$$h_{ij} : \left(u^1, \ldots, u^n\right) \mapsto \left(u^1, \ldots, \hat{u^i}, \ldots, \pm\sqrt{1 - \|u\|^2}, \ldots, u^n\right),$$

which is smooth (note $\|u\| \leq 1$). A similar formula holds for $i > j$, and for $i = j$ the transition function is just the identity map. Thus all the transition functions are diffeomorphisms (in particular homeomorphisms), and we have a smooth structure on $S^n$.

4. $\mathbb{R}P^n := S^n / \{\pm 1\}$ is a smooth manifold of dimension $n$. Let $U_i^+ := \{x \in S^n | x^i > 0\}$ and similarly define $U_i^-$; note that $\{U_i^\pm \mid i = 1, \ldots, n+1\}$ define an open cover of $S^n$. Let $\pi : S^n \to S^n/\{\pm 1\} =: \mathbb{R}P^n$ denote the canonical projection. Observe that $\pi(U_i^+) = \pi(U_i^-)$, and that $\pi$ restricted to $U_i^\pm$ is a homeomorphism ($\mathbb{R}P^n$ is given the quotient topology induced by $\pi$). Hence the composition

$$\pi(U_i^+) \xrightarrow{\pi^{-1}} U_i^+ \xrightarrow{h_i} \mathbb{R}^n$$

is a well defined homeomorphism, call it $k_i$. Moreover we have

$$k_j \circ k_i^{-1} = (h_j \circ \pi^{-1}) \circ (h_i \circ \pi^{-1})^{-1} = h_j \circ h_i^{-1}$$

which is smooth. It follows that $(\pi(U_i^+), k_i)$ defines an atlas on $\mathbb{R}P^n$.

5. The two-dimensional torus $T$. Define an equivalence relation on $\mathbb{R}^2$ by $x \sim y$ if and only if $x^i - y^i \in \mathbb{Z}$ for $i = 1, 2$. Let $T$ denote $\mathbb{R}^2/\sim$. Any unit square $Q \subseteq \mathbb{R}^2$ with vertices $(a, b), (a+1, b), (a, b+1)$ and $(a+1, b+1)$ determines a homeomorphism $\mathrm{int}(Q) \to U(Q) \subseteq T$, where $U(Q)$ is an open subset of $T$ (in fact, $U(Q)$ is all of $T$ apart from two circles). The inverse map $h_Q : U(Q) \to \mathrm{int}(Q)$ then is a homeomorphism from an open set of $T$ into an open set of $\mathbb{R}^2$. Clearly the $\{U(Q)\}$ cover $T$, and if $Q_1$ and $Q_2$ are two squares with a non-empty intersection then one easily sees that the coordinate transformation $h_{Q_2} \circ h_{Q_1}^{-1}$ is given locally by translations; that is, each component of $h_{Q_1}(Q_1 \cap Q_2)$ (of which in general there will be seven - draw a picture!) is mapped by a translation onto a corresponding component of $h_{Q_2}(Q_1 \cap Q_2)$ by $h_{Q_2} \circ h_{Q_1}^{-1}$. Thus we have a smooth structure, and $T$ is a smooth 2-manifold.

## 1.7 Definition

Let $M^m$ and $N^n$ be smooth manifolds. A continuous map $\Phi : M \to N$ is called **smooth** if for each $p \in M$ for some (and hence all) charts $h : U \to \mathcal{O} \subseteq \mathbb{R}^m$ and $k : V \to \Omega \subseteq \mathbb{R}^n$ on $M$ and $N$ respectively with $p \in U$ and $\Phi(p) \in V$ such that the map the composite map (called the **local expression** of $\Phi$)

$$k \circ \Phi \circ h^{-1} : h(U \cap \Phi^{-1}(V)) \to k(\Phi(U) \cap V)$$

is smooth.

If $\Phi$ is a homeomorphism and its inverse $\Phi^{-1} : N \to M$ is also smooth then we say $\Phi$ is a **diffeomorphism**. We say that $\Phi$ is a **local diffeomorphism** if given any $p \in M$ we can find a neighborhood $U$ of $p$ such that $\Phi|_U : U \to \Phi(U)$ is a diffeomorphism.

A **smooth function** on an open subset $U \subseteq M$ is a smooth map $f : U \to \mathbb{R}$ where $\mathbb{R}$ is given the standard smooth structure (cf. Example 1.6.2).

Observe that when we give $\mathbb{R}^n$ the standard smooth structure, the charts on manifolds become diffeomorphisms (their local expression is the identity).

## 1.8 Germs

Let $M$ be a manifold and $p \in M$. Functions $f, g$ defined on open subsets $U, V$ respectively containing $p$ are said to have the same **germ** at $p$ if there exists a neighborhood $W$ of $p$ contained in $U \cap V$ such that $f|_W \equiv g|_W$. More precisely, define an equivalence relation on the space of smooth functions defined in a neighborhood of $p$, by $(U, f) \sim (V, g)$ if and only if there exists a neighborhood $W$ of $p$ contained in $U \cap V$ such that $f|_W \equiv g|_W$. A **germ** is an equivalence class under this relation. Notationally, we will not differentiate between a germ $f$ at $p$ and a representative $(U, f)$ of $f$. This will hopefully not be confusing.

Let $C_p^\infty$ denote the set of germs of smooth functions at $p$ (occasionally we write $C_{M,p}^\infty$ when there is more than one manifold under consideration). Observe that $C_p^\infty$ is a ring; given germs $f$ and $g$ with representatives $(U, f)$ and $(V, g)$ respectively we define $f + g$ to be germ containing $(U \cap V, f + g)$ and $f \cdot g$ to be the germ containing $(U \cap V, f \cdot g)$ (where $f \cdot g(p) := f(p) \cdot g(p)$). Moreover we have a natural inclusion of the constant germs into $C_p^\infty$, which induces a natural map $\mathbb{R} \hookrightarrow C_p^\infty$ making $C_p^\infty$ into an $\mathbb{R}$-algebra.

A germ $f$ has a well defined value at $p$ (although nowhere else though), and this defines a surjective ring homomorphism $\mathrm{eval} : C_p^\infty \to \mathbb{R}$ sending $f \mapsto f(p)$. If we let $\mathcal{F}_p$ denote the kernel of eval, then $\mathcal{F}_p$ is an ideal of $C_p^\infty$. In fact, since eval is surjective, $\mathcal{F}_p$ is a maximal ideal, and in fact is the unique maximal ideal, since if $f(p) \neq 0$ then if $(U, f)$ is any representative there exists a neighborhood $V \subseteq U$ of $p$ such that $f$ is never zero on $V$. Then the germ containing $(V, 1/f)$ is an inverse for $f$. Hence any germ in $C_p^\infty \backslash \mathcal{F}_p$ is invertible; equivalently $\mathcal{F}_p$ is the unique maximal ideal and $C_p^\infty$ is thus a local ring.

## 1.9 Definitions

A **tangent vector** $v$ at $p \in M$ is a **linear derivation** of the $\mathbb{R}$-algebra $C_p^\infty$, that is, a linear map $C_p^\infty \to \mathbb{R}$ such that $v(f \cdot g) = f(p)v(g) + v(f)g(p)$. The tangent vectors form a real vector space in the obvious way; this space is denoted $T_p(M)$ and is called the **tangent space** to $M$ at $p$. Note that if $c$ is the constant germ $c(p) = c \in \mathbb{R}$ then if $v$ is any tangent vector we have $v(c) = 0$. Indeed, $v(c) = cv(1)$ and $v(1) = v(1 \cdot 1) = v(1) + v(1)$ implies $v(1) = 0$.

## 1.10 Charts and tangent vectors

Let $C_0^\infty$ denote $C_{\mathbb{R}^n,0}^\infty$ and let $M$ be a smooth $n$-manifold, $p \in M$ and $h : U \to \mathcal{O} \subseteq \mathbb{R}^n$ a chart on $M$ about $p$ such that $h(p) = 0$ (in general we say such a chart is **centred** about $p$). Then $h$ defines an isomorphism $h^*$ of $\mathbb{R}$-algebras $h^*$ defined by, for $f \in C_0^\infty$,

$$h^*(f) = f \circ h.$$

$h^*$ is an isomorphism precisely because $\phi$ is a diffeomorphism.

Then given $v \in T_p(M)$ we can associate a tangent vector $h_*(v) \in T_0(\mathbb{R}^n)$ where $h_*(v)$ is the derivation of $C_0^\infty$ defined by

$$h_*(v)(f) = v(h^*(h)).$$

As $h^*$ is an isomorphism, so is $h_*$ and thus we have defined an isomorphism $h_* : T_p(M) \to T_0(\mathbb{R}^n)$.

## 1.11 A basis for the tangent space

Let $(r^1, \ldots, r^n)$ denote the standard coordinates on $\mathbb{R}^n$. Consider the operator $\frac{\partial}{\partial r^i}\big|_0$ defined by $\frac{\partial}{\partial r^i}\big|_0(f) = \frac{\partial f}{\partial r^i}(0)$. Then $\frac{\partial}{\partial r^i}\big|_0$ is an linear derviation of $C_0^\infty$ and hence an element of $T_0(\mathbb{R}^n)$. Observe that $\frac{\partial}{\partial r^i}\big|_0(r^j) = \delta_i^j$. In fact, we claim that $\left\{\frac{\partial}{\partial r^i}\big|_0 | i = 1, \ldots, n\right\}$ forms a basis for $T_0(\mathbb{R}^n)$.

To prove this we need the following result from calculus.

## 1.12 Calculus lemma

Let $f : U \to \mathbb{R}$ be smooth, where $U \subseteq \mathbb{R}^n$ is open and convex. Then there exist smooth functions $g_{ij} : U \to \mathbb{R}$ $(i, j = 1, \ldots, n)$ such that for any $y \in U$ we have

$$f(y) = f(0) + \frac{\partial f}{\partial r^i}(0) r^i(y) + r^i(y) r^j(y) g_{ij}(y), \tag{1}$$

where as in the rest of these notes we use the **summation convention** that we sum over indices in an expression that appear in the top and the bottom - we note that the index in $\frac{\partial}{\partial r^i}\big|_0$ is considered to be on the bottom (it's on the bottom of a fraction).

## 1.13 Proposition

$\left\{\frac{\partial}{\partial r^i}\big|_0 | i = 1, \ldots, n\right\}$ is a basis of $T_0(\mathbb{R}^n)$.

◀ Suppose $v = a^i \partial_i|_0$ is the zero derivation (here and elsewhere, where possible we will abbreviate an expression of the form $\frac{\partial}{\partial x^i}\big|_0$ to $\partial_i|_0$ - this of course won't work if we are working with several different coordinate systems at once). Then $v(r^i) = a^i = 0$, and so $v = 0$. Now let $v \in T_0(\mathbb{R}^n)$, and set $a^i := v(r^i)$. Consider $v_0 := a^i \partial_i|_0 \in T_0(\mathbb{R}^n)$. Given a germ $f \in C_0^\infty$, pick a representative $(U, f)$, where we may assume $U$ is convex. Write $f$ as in (1) and compute:

$$v(f) = v(f(0)) + \frac{\partial f}{\partial r^i}(0) v(r^i) + v(r^i \cdot r^j \cdot g_{ij}).$$

Now the first term disappears, as $v$ is zero on constants, and since $r^i(0) = 0$, the derivation property kills the last term. Thus

$$v(f) = \frac{\partial f}{\partial r^i}(0) v(r^i) = a^i \partial_i|_0(f) = v_0(f).$$

Thus $v = v_0$ and the we have a basis. ▶

## 1.14 Observation

Observe we can canonically identify $T_0(\mathbb{R}^n)$ with $\mathbb{R}^n$ via $a^i\partial_i|_0 \leftrightarrow a^i e_i$, where $\{e_i\}$ is the standard basis of $\mathbb{R}^n$ (so $r^i(e_j) = \delta^i_j$). In fact, given any $p \in \mathbb{R}^n$, the same proof shows $\{\partial_i|_p\}$ forms a basis of $T_p(\mathbb{R}^n)$ and thus the identification $a^i\partial_i|_p \leftrightarrow a^i e_i$ allows us to identify $T_p(\mathbb{R}^n)$ with $\mathbb{R}^n$ for any $p \in \mathbb{R}^n$.

## 1.15 A basis for $T_p(M)$

Now let $M^n$ be a smooth manifold and $p \in M$. Pick a chart $h : U \to \mathcal{O} \subseteq \mathbb{R}^n$ centered about $p$. Let $(x^1, \dots, x^n)$ be the coordinates of $h$. Recall we have a map $(h^{-1})_* : T_0(\mathbb{R}^n) \to T_p(M)$ that is an isomorphism. Define the tangent vector $\frac{\partial}{\partial x^i}\big|_p$ by,

$$\frac{\partial}{\partial x^i}\Big|_p(f) := \left(h^{-1}\right)_* \left(\frac{\partial}{\partial r^i}\Big|_0\right)(f) = \frac{\partial}{\partial r^i}\left(f \circ h^{-1}\right)(0).$$

Since $(h^{-1})_*$ is an isomorphism, we have the following immediate corollary.

## 1.16 Corollary

Let $M^n$ be a smooth manifold and $p$ in $M$. If $(U, h)$ is a chart centred at $p$ with coordinates $(x^1, \dots, x^n)$ then $T_p(M)$ is a $n$-dimensional real vector space with basis $\left\{\frac{\partial}{\partial x^i}\big|_p | i = 1, \dots, n\right\}$. Moreover if $v \in T_p(M)$ then $v = a^i \frac{\partial}{\partial x^i}\big|_p$ where $a^i := v(x^i)$.

## 1.17 Jacobians

Suppose now $(U, h)$ and $(V, k)$ are both charts centred at $p$. Write $(x^1, \dots, x^n)$ for the coordinates of $h$ and $(y^1, \dots, y^n)$ for the coordinates of $k$. Let $(r^1, \dots r^n)$ denote the coordinates on $h(U)$ and $(s^1, \dots, s^n)$ the coordinates on $k(V)$.

Observe by the previous Corollary we have

$$\frac{\partial}{\partial y^j}\Big|_p = \frac{\partial x^i}{\partial y^j}(p)\frac{\partial}{\partial x^i}\Big|_p. \tag{2}$$

Let $F$ be the coordinate transformation $h \circ k^{-1}$. We can write $F = (F^1, \dots, F^n)$ where $F^i = r^i \circ F$. Now note

$$\begin{aligned}
\frac{\partial x^i}{\partial y^j}(p) &= \frac{\partial}{\partial y^j}\Big|_p(x^i) \\
&= \frac{\partial}{\partial s^j}(x^i \circ k^{-1})(0) \\
&= \frac{\partial}{\partial s^j}(r^i \circ h \circ k^{-1})(0) \\
&= \frac{\partial}{\partial s^j}(r^i \circ F)(0) \\
&= \frac{\partial F^i}{\partial s^j}(0),
\end{aligned}$$

and hence

$$\frac{\partial x^i}{\partial y^j}(p) = JF(0), \tag{3}$$

where $JF(0)$ is the **Jacobian** of $F$ at 0. Thus we have

$$\frac{\partial}{\partial y_j}\Big|_p = JF(0)^i_j \frac{\partial}{\partial x^i}\Big|_p.$$

## 1.18 General charts

More generally, given any chart $(U, h)$ on $M^n$ with coordinates $(x^1, \ldots, x^n)$, we can define $\frac{\partial}{\partial x^i}\big|_p$ for any $p \in U$ by the formula

$$\frac{\partial}{\partial x^i}\Big|_p(f) := \frac{\partial}{\partial r^i}\left(f \circ h^{-1}\right)(h(p)),$$

where $(r^1, \ldots, r^n)$ are the coordinates on $h(U)$. If, $h(p) = c = (c^1, \ldots, c^n)$, say if we consider the linear coordinate transformation $k = (y^1, \ldots, y^n)$ where $y^i := x^i - c^i$, then if $(s^1, \ldots, s^n)$ denote the coordinates on $k(U)$ we have

$$\begin{aligned}\frac{\partial}{\partial y^i}\Big|_p(f) &= \frac{\partial}{\partial s^i}(f \circ k^{-1})(0)\\ &= \frac{\partial}{\partial r^i}\left(f \circ h^{-1}\right)(c)\\ &= \frac{\partial}{\partial x^i}\Big|_p(f).\end{aligned}$$

Hence for any $p \in U$, $\left\{\frac{\partial}{\partial x^i}\big|_p\right\}$ is a basis of $T_p(M)$.

## 1.19 Changing coordinates

If $v \in T_p(M)$ and $(U, h)$ and $(V, k)$ are charts about $p$ with coordinates $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ respectively then writing

$$a^i \frac{\partial}{\partial x^i}\Big|_p = v = b^j \frac{\partial}{\partial y^j}\Big|_p$$

and using (2)we see that

$$a^i = v(x^i) = b^j \frac{\partial}{\partial y^j}\Big|_p(x^j) = b^j \frac{\partial x^i}{\partial y^j}(p). \tag{4}$$

## 1.20 Definition

Let $\Phi : M \to N$ be a smooth map between smooth manifolds. Let $p \in M$. We have a map $\Phi^* : C^\infty_{N,\Phi(p)} \to C^\infty_{M,p}$ defined by $\Phi^*(f) = f \circ \Phi$. Now define the **derivative** of $\Phi$ at $p$ to be the map

$$d\Phi(p) : T_p(M) \to T_{\Phi(p)}(N)$$

defined by

$$d\Phi(p)(v)(f) = v(\Phi^*(f)).$$

Thus $d\Phi(p)$ is a linear map between the tangent spaces. Where possible we will omit the '$p$' from the notation and just write $d\Phi$.

The chain rule is tautologous: if $\Psi : N \to P$ is a smooth map of smooth manifolds such that $\Psi \circ \Phi : M \to P$ is defined then

$$d(\Psi \circ \Phi) = d\Psi \circ d\Phi.$$

Indeed, if $v \in T_p(M)$ and $f \in \mathcal{A}_{P,\Psi\Phi(p)}$ then

$$\begin{aligned}d(\Psi \circ \Phi)(v)(f) &= v((\Psi \circ \Phi)^*(f))\\ &= v(\Psi \circ \Phi \circ f)\\ &= v(\Phi^*(f \circ \Psi))\\ & \quad d\Phi(v)(f \circ \Psi)\\ & \quad d\Phi(v)(\Psi^*(f))\\ & \quad d\Psi(d\Phi(v))(f).\end{aligned}$$

If $\left\{\frac{\partial}{\partial x^i}\big|_p\right\}$ is a basis of $T_p(M)$ and $\left\{\frac{\partial}{\partial y^j}\big|_{\Phi(p)}\right\}$ is a basis of $T_{\Phi(p)}(N)$ then

$$d\Phi\left(\frac{\partial}{\partial x^j}\Big|_p\right)(y^i) = \frac{\partial}{\partial x^j}(y^i \circ \Phi) = \frac{\partial \Phi^i}{\partial x^j}(p),$$

where $\Phi^i = y^i \circ \Phi$ and thus

$$d\Phi\left(\frac{\partial}{\partial x^j}\Big|_p\right) = \frac{\partial \Phi^i}{\partial x^j}(p)\frac{\partial}{\partial y^i}\Big|_{\Phi(p)}.$$

Hence if

$$v = a^j \frac{\partial}{\partial x^j}\Big|_p, \qquad d\Phi(v) = b^i \frac{\partial}{\partial y^i}\Big|_{\Phi(p)}$$

then

$$d\Phi(v) = a^j d\Phi\left(\frac{\partial}{\partial x^j}\Big|_p\right) = a^j \frac{\partial \Phi^i}{\partial x^j}(p)\frac{\partial}{\partial y^i}\Big|_{\Phi(p)},$$

and thus

$$b^i = a^j \frac{\partial \Phi^i}{\partial x^j}(p). \tag{5}$$

## 1.21 Example

As a special case, if $\Phi : \mathbb{R}^m \to \mathbb{R}^n$ is smooth and $p \in \mathbb{R}^m$, with $(r^1, \ldots, r^m)$ the standard coordinates on $\mathbb{R}^m$ and $(s^1, \ldots, s^n)$ the standard coordinates of $\mathbb{R}^n$ we have

$$d\Phi\left(\frac{\partial}{\partial r^i}\Big|_p\right) = J\Phi(p)^j_i \frac{\partial}{\partial s^j}\Big|_{\Phi(p)}. \tag{6}$$

In particular, under the identification of $T_p(\mathbb{R}^n)$ with $\mathbb{R}^d$ as in Section 1.14, the derivative $d\Phi(p)$ is just the linear map determined by the Jacobian $J\Phi(p)$. Given this, one might ask why the chain rule was so easy to prove (as it is not so simple to prove in standard multivariate calculus). The answer is in our use of Lemma 1.12, which in turn used the standard chain rule in multivariate calculus.

## 1.22 Definition

A **smooth curve** on a manifold $M$ is a smooth map $c : (a, b) \to M$, where $(a, b) \subseteq \mathbb{R}$ is an interval and $\mathbb{R}$ is given the standard smooth structure. For $t \in (a, b)$ the **tangent vector to** $c$ **at** $t$ is

$$dc\left(\frac{d}{dr}\Big|_t\right) \in T_{c(t)}(M),$$

(we write $d/dr$ instead of $\partial/\partial r$ when $n = 1$), and is denoted $\dot{c}(t)$.

If $c : (a, b) \to M$ is a smooth curve, and $(x^1, \ldots, x^n)$ local coordinates about $c(t_0)$ then for $t$ close to $t_0$ we set $c^i(t) := x^i \circ c(t)$. Then

$$\dot{c}(t)(x^i) = dc\left(\frac{d}{dr}\Big|_t\right)(x^i) = \frac{d}{dr}(x^i \circ c) = \frac{dc^i}{dr}(t),$$

and hence

$$\dot{c}(t) = \frac{dc^i}{dr}(t)\partial_i|_{c(t)}, \tag{7}$$

that is,

$$\dot{c} = \frac{dc^i}{dr} \cdot \partial_i \circ c.$$

In particular if $M = \mathbb{R}^n$ and we use the identification of $T_{c(t)}(\mathbb{R}^n)$ with $\mathbb{R}^n$ given in Section 1.12 we have $\dot{c}(t) = \frac{dc^i}{dr}(t)e_i$, which recovers the standard definition from multivariate calculus for the derivative of a smooth curve $c : (a, b) \to \mathbb{R}^n$.

## 1.23 An alternative definition of tangent vectors

We now focus our attention on smooth curves in $M$ defined on a neighborhood of 0. A smooth curve $c:(-\epsilon,\epsilon)\to M$ with $c(0)=p$ defines a tangent vector $\dot{c}(0)\in T_p(M)$. Note that smooth curves $c$ and $\gamma$ define the same tangent vector if and only if, for some (and hence every) chart $h$ centred at $p$ we have $h\circ c$ and $h\circ\gamma$ defining the same tangent vector in $T_0(\mathbb{R}^n)$. By the previous section, this is if and only if

$$\frac{d}{dr}(h\circ c)(0)=\frac{d}{dr}(h\circ\gamma)(0). \tag{8}$$

Conversely suppose $v\in T_p(M)$ is any tangent vector. By making a linear change of coordinates in the vector space $T_p(M)$ we may assume that we have local coordinates $h=(x^1,\dots,x^n)$ about $p$ such that $v=\partial_1|_p$. Define $c(t)=h^{-1}(t,0,\dots,0)$. Then for $f\in C_p^\infty$ we have

$$\begin{aligned}\dot{c}(0)(f) &= dc\left(\frac{d}{dr}\Big|_0\right)(f)\\ &= \frac{d}{dr}(f\circ c)(0)\\ &= \frac{\partial}{\partial r^1}(f\circ h)(0)\\ &= \frac{\partial}{\partial x^1}\Big|_p(f)\\ &= v(f).\end{aligned}$$

Thus any tangent vector $v\in T_p(M)$ can be written as $\dot{c}(0)$ for some smooth curve $c:(-\epsilon,\epsilon)\to M$. Thus we can make the following alternative definition of $T_p(M)$: a tangent vector at $p\in M$ is an equivalence class of smooth curves $c:(-\epsilon,\epsilon)\to M$ such that $c(0)=p$, where $c\sim\gamma$ if and only if for some chart $h$ centred at $p$, (8) holds.

## 1.24 Definition

Let $M$ be a smooth manifold. The **tangent bundle** of $M$ is the disjoint union of the tangent spaces;

$$T(M):=\coprod_{p\in M}T_p(M).$$

We have a natural projection $\pi:T(M)\to M$ sending $v\in T_p(M)\mapsto p$. When referring to an element of $T(M)$, we will often write $(p,v)$ to indicate that $v\in T_p(M)$.

## 1.25 Theorem

Let $M$ be a smooth $n$-manifold. Then $T(M)$ is naturally a smooth $2n$-manifold such that $\pi$ is smooth.

◀ Let $\{(U_\alpha,h_\alpha)\}$ be an atlas on $M$, where $h_\alpha:U_\alpha\to\mathcal{O}_\alpha$ has coordinates $(x_\alpha^1,\dots,x_\alpha^n)$. Define a **local trivialisation** $t_\alpha:T(U_\alpha):=\coprod_{p\in U_\alpha}T_p(M)\to U_\alpha\times\mathbb{R}^n$ by

$$t_\alpha(p,v)=\left(p,v\left(x_\alpha^1\right),\dots,v\left(x_\alpha^n\right)\right).$$

If $U_\alpha\cap U_\beta\neq\emptyset$ then the map $t_\alpha\circ t_\beta^{-1}:(U_\alpha\cap U_\beta)\times\mathbb{R}^n\to(U_\alpha\cap U_\beta)\times\mathbb{R}^n$ sends

$$\left(p,a^1,\dots,a^n\right)\mapsto\left(p,a^i\frac{\partial x_\beta^1}{\partial x_\alpha^i}(p),\dots,a^i\frac{\partial x_\beta^n}{\partial x_\alpha^i}(p)\right)$$

(cf. (4)). In particular, the map $\psi_{\alpha\beta}(p):=(t_\beta\circ t_\alpha^{-1})(p,\cdot):\mathbb{R}^n\to\mathbb{R}^n$ is a linear isomorphism, and moreover the map

$$\psi_{\alpha\beta}:U_\alpha\cap U_\beta\to GL(n,\mathbb{R}),\quad p\mapsto\psi_{\alpha\beta}(p)$$

is smooth.

Now define $\bar{h}_\alpha : T(U_\alpha) \to \mathcal{O}_\alpha \times \mathbb{R}^n$ by $\bar{h}_\alpha = (h_\alpha \times \mathrm{id}) \circ t_\alpha$. Then observe for $\alpha, \beta$ such that $U_\alpha \cap U_\beta \neq \emptyset$ the transition function $\bar{h}_\alpha \circ \bar{h}_\beta^{-1} : h_\beta(U_\alpha \cap U_\beta) \times \mathbb{R}^n \to h_\alpha(U_\alpha \times U_\beta) \times \mathbb{R}^n$ sends

$$(q, a^1, \ldots, a^n) \mapsto \left(h_{\alpha\beta}(q), \psi_{\alpha\beta}\left(h_\beta^{-1}(q)\right)(a^1, \ldots, a^n)\right),$$

which is smooth. Moreover the collection $\left\{\left(T(U_\alpha), \bar{h}_\alpha\right)\right\}$ is a collection of bijections such that $T(M) = \bigcup_\alpha T(U_\alpha)$. Now define a topology on $T(M)$ by declaring the $\bar{h}_\alpha$ to be homeomorphisms. This will make $T(M)$ into a smooth $2n$-dimensional manifold with atlas $\left\{\left(T(U_\alpha), \bar{h}_\alpha\right)\right\}$ as soon as we know that is Hausdorff and second countable under this topology. If $(p,v) \neq (q,w)$ then either $p \neq q$ and we can use the Hausdorff property of $M$ or $p = q$ and $v \neq w$ and we can use the Hausdorff property of $\mathbb{R}^n$ to separate $(p,v)$ and $(q,w)$. To see second countability, we may assume that $\{U_\alpha\}$ is a countable cover of $M$; then each $U_\alpha \times \mathbb{R}^n$ is second countable, and hence so are the $T(U_\alpha)$ and since$\{T(U_\alpha)\}$ is a countable cover of $T(M)$ it follows $T(M)$ is second countable.

It is immediate that $\pi$ is smooth, as its local expression with respect to charts $(U_\alpha, h_\alpha)$ and $\left(T(U_\alpha), \bar{h}_\alpha\right)$ is the map $\mathrm{proj}_1 : U_\alpha \times \mathbb{R}^n \to U_\alpha$. This completes the proof. ▶

In fact this proof actually shows that $T(M)$ is a smooth **vector bundle** of rank $n$ over $M$ (vector bundles will be defined in Chapter 5).

## 1.26 Bundle maps between tangent bundles

Let $\Phi : M^m \to N^n$ be smooth. Then $\Phi$ induces a **bundle morphism** $d\Phi : T(M) \to T(N)$ defined by

$$d\Phi(p,v) = d\Phi(p)(v).$$

Moreover, $d\Phi$ is smooth. Indeed, if $(p,v) \in T(M)$ and $(U,h)$ is a chart about $p$, and $(V,k)$ a chart on $N$ about $\Phi(p)$ then the local expression $\bar{k} \circ d\Phi \circ \bar{h}^{-1}$ is the map

$$\left(h(p), a^1, \ldots, a^m)\right) \mapsto \left(k \circ \Phi \circ h^{-1}(h(p)), a^j \frac{\partial \Phi^1}{\partial x^j}(p), \ldots, a^j \frac{\partial \Phi^n}{\partial x^j}(p)\right),$$

(cf. (5)) which is smooth.

# 2 Vector fields

## 2.1 Definition

A **vector field** on $M$ is a smooth **section** of $\pi : T(M) \to M$, that is, a smooth map $X : M \to T(M)$ such that $\pi \circ X = \mathrm{id}_M$. Thus if $X$ is a vector field, $X(p)$, which we will often write as $X_p$, lies in $T_p(M)$ for all $p \in M$. The assertion that $X$ is smooth is equivalent to the following. Let $(U,h)$ be a chart on $M$ with coordinates $(x^1, \ldots, x^n)$. Then for $p \in U$ we can write $X_p = X^i(p)\partial_i|_p$ for some functions $X^i : U \to \mathbb{R}$, and to say $X : M \to T(M)$ is smooth is equivalent to saying that the $X^i$ are smooth functions on $U$. Indeed, with respect to the chart $(T(U), \bar{h})$ on $T(M)$, $X$ has local expression

$$p \mapsto \left(x^1(p), \ldots, x^n(p), X^1(p), \ldots, X^n(p)\right)$$

and thus $X$ is smooth if and only if all the $X^i$ are smooth.

Let $\mathcal{X}(M)$ denote the set of all smooth vector fields on $M$. For $f \in C^\infty(M)$ and $X \in \mathcal{X}(M)$, we can define a new vector field $fX : M \to T(M)$ by $(fX)(p) = f(p)X_p \in T_p(M)$. Similarly we can also define for $f, g \in C^\infty(M)$ and $X, Y \in \mathcal{X}(M)$ a vector field in the obvious way, and so $\mathcal{X}(M)$ becomes a module over the ring $C^\infty(M)$.

We can also define **local vector fields** to be smooth sections of $\pi$ defined only on some open set $U \subseteq M$. We denote the local vector fields over $U \subseteq M$ by $\mathcal{X}(U)$. In particular if $(x^1, \ldots, x^n)$ are local coordinates on $U \subseteq M$ then $\partial_i \in \mathcal{X}(U)$.

## 2.2 Proposition

Let $M^n$ be a smooth manifold on which there exist $n$ **independent** vector fields $X_1, \ldots, X_n$, that is, for all $p \in M$, $\{X_1(p), \ldots, X_n(p)\}$ is a basis of $T_p(M)$. Then $T(M)$ is isomorphic as a vector bundle to $M \times \mathbb{R}^n$. By this we mean (this will be explained in more detail in Chapter 5 - see Section 5.4 in particular) that there exists a diffeomorphism $F : T(M) \to M \times \mathbb{R}^n$ such that $F(T_p(M)) \subseteq \{p\} \times \mathbb{R}^n$ and the restriction of $F$ to $T_p(M)$, $F_p : T_p(M) \to \{p\} \times \mathbb{R}^n$ is a linear isomorphism.

◀ Define $F : T(M) \to M \times \mathbb{R}^n$ by $F(p, v) = (p, c^1, \ldots, c^n)$ where $v = c^i X_i(p)$ in the basis $\{X_i(p)\}$ of $T_p(M)$. Then $F$ is a bijection by assumption, and restricts as required. We thus need only show $F$ and $F^{-1}$ are smooth. Given a chart $(U, h)$ with local coordinates $(x^1, \ldots, x^n)$ on $M$, the local expression of $F$ in the charts $(T(U), \hat{h})$ and $(U \times \mathbb{R}^n, h \times \mathrm{id})$ is

$$(p, a^1, \ldots, a^n) \mapsto (p, c^1, \ldots, c^n)$$

where $a^i \partial_i|_p = v = c^i X_i(p)$.

Now we can write $X_i(p) = b_i^j(p)\partial_j|_p$ for some smooth functions $b_i^j : U \to \mathbb{R}$. Then

$$a^j \partial_j|_p = v = c^i X_i(p) = c^i b_i^j(p)\partial_j|_p,$$

and thus $a^j = c_i b_i^j(p)$. This shows that $F^{-1}$ is smooth, as the $b_i^j$ are smooth. Moreover, since matrix inversion is smooth, and $p \mapsto \left[b_i^j(p)\right]$ is smooth, if $\left[d_j^i(p)\right]$ denotes the inverse matrix to $\left[b_i^j(p)\right]$ then $p \mapsto \left[d_j^i(p)\right]$ is also smooth. Then as $c^i = d_j^i(p)a^j$, we see that $F$ is also smooth. ▶

## 2.3 Definition

Given a smooth vector field $X$ and $f \in C^\infty(M)$, we can define a function $Xf : M \to \mathbb{R}$ by $Xf(p) = X_p(f)$. If $(U, h)$ is a chart about $p$ with coordinates $(x^1, \ldots, x^n)$, then we can locally write $X_p = X^i(p)\partial_i|_p$, and thus

$$Xf(p) = X^i(p)\frac{\partial f}{\partial x^i}(p).$$

In particular, $Xf$ is smooth. Hence we may also view $X$ as a **derivation** of $C^\infty(M)$; since $X_p$ is a derivation of $C_p^\infty$ we immediately have

$$X(fg)(p) = Xf(p) \cdot g(p) + f(p) \cdot Xg(p).$$

In fact, more is true.

## 2.4 Proposition

A map $\mathfrak{X} : C^\infty(M) \to C^\infty(M)$ is a deriviation if and only if there exists $X \in \mathcal{X}(M)$ such that for all $f \in C^\infty(M)$, $\mathfrak{X}(f) = Xf$.

◀ We have just shown given $X \in \mathcal{X}(M)$, defining $\mathfrak{X} : C^\infty(M) \to C^\infty(M)$ by $\mathfrak{X}(f) = Xf$ does indeed define a derivation. Conversely, suppose $\mathfrak{X}$ is a derivation of $C^\infty(M)$, $p \in M$ and define a tangent vector $X_p \in T_p(M)$ by, for $f \in C_p^\infty$,

$$X_p(f) := \mathfrak{X}\left(\tilde{f}\right)(p),$$

where $\tilde{f}$ is any smooth function defined on all of $M$ such that the image of $\tilde{f}$ in $C_p^\infty$ is $f$.

There are three things to check in order to conclude this is well defined. Firstly we need to know that there exists a smooth extension $\tilde{f}$ of $f$ to a function defined on all of $M$ (a priori, we only know that $f$ defines a function on a neighborhood of $p$). Secondly we need to know that if $\tilde{f}$ and $\tilde{f}'$ are two such extensions then $\mathfrak{X}(\tilde{f})(p) = \mathfrak{X}(\tilde{f}')(p)$. Then we need to check $X_p$ is indeed a derivation of $C_p^\infty$, and thus does indeed define an element of $T_p(M)$.

The key fact we need is the following: given any $p \in M$ and any neighborhoods $U \subset V$ of $p$ there exists a smooth function $\psi : M \to \mathbb{R}$ such that $\psi|_U \equiv 1$ and $\psi|_{M \setminus V} \equiv 0$. Such a function is called a **bump function**. That such functions exist is not obvious, and depends on the existence of **partitions of unity**. We will define these in Chapter 5, Theorem 5.17, and prove the existence of bump functions in Corollary 5.18. For now however we will just accept such functions $\psi$ exist.

It is then immediate that such an extension $\tilde{f}$ exists. In order to check $X_p(f)$ is independent of the choice of extension, it is enough to observe that if $\tilde{f} \in C^\infty(M)$ vanishes in a neighborhood of $p$ then $\mathfrak{X}(\tilde{f})(p) = 0$. If $\tilde{f}|_V \equiv 0$ for some neighborhood of $p$, then choose a neighborhood $U \subset V$ of $p$, pick a bump function $\psi$ such that $\psi|_U \equiv 1$ and $\psi|_{M \setminus V} \equiv 0$, and then consider the smooth function $\psi' : M \to \mathbb{R}$, $\psi' := 1 - \psi$. Then $\psi'|_U \equiv 0$ and $\psi'|_{M \setminus V} \equiv 1$, and so as functions on $M$, $\tilde{f} = \tilde{f}\psi'$. The derivation property of $\mathfrak{X}$ ensures $\mathfrak{X}(\tilde{f}\psi')(p) = 0$, and so also $\mathfrak{X}(\tilde{f})(p) = 0$ as required.

It is clear now that $X_p$ is a derivation of $C_p^\infty$, since $\mathfrak{X}$ acts as a derivation on the chosen extensions. To complete the proof, we show that $p \mapsto X_p$ is smooth, and thus this construction defines us a vector field $X$ on $M$. For this it is enough to check that if $\left(x^1, \dots, x^n\right)$ are any local coordinates on a neighborhood $U$ of $p$ then $Xx^i \in C^\infty(U)$, regarding $x^i : U \to \mathbb{R}$ as smooth functions on $U$. But this is clear, since if $f \in C^\infty(U)$ is any function then using a bump function we extend $f$ to $\tilde{f} \in C^\infty(M)$, and then observe that $\mathfrak{X}(\tilde{f}) \in C^\infty(M)$. ►

## 2.5 $XY$ is not a derivation

Given $X, Y \in \mathcal{X}(M)$ and $f \in C^\infty(M)$, we can define $(XY) f = X(Yf)$. However this does **not** define a vector field, as this is not a derivation of $C^\infty(M)$. Indeed,

$$(XY)(fg) = X\{(Yf)g + fYg\} = (XY)f \cdot g + Yf \cdot Xg + Xf \cdot Yg + f \cdot (XY)g \neq (XY)f \cdot g + f \cdot (XY)g.$$

However, this shows that $[X, Y] := XY - YX$ is a derivation of $C^\infty(M)$. It thus follows that $[X, Y] := XY - YX$ defines a vector field on $M$. We call $[X, Y]$ the **Lie bracket** of $X$ and $Y$.

## 2.6 Properties of the Lie bracket

1. The map $[\cdot, \cdot] : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$ is bilinear over $\mathbb{R}$ and skew-symmetric.
2. If $f, g \in C^\infty(M)$ and $X, Y \in \mathcal{X}(M)$ then $[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$.
3. If $(U, h)$ is a chart on $M$ with coordinates $(x^1, \dots, x^n)$ and in this chart $X(p) = X^i(p)\partial_i|_p$ and $Y(p) = Y^i(p)\partial_i|_p$ then
$$[X, Y](p) = \left( X^i(p) \frac{\partial Y^j}{\partial x_i}(p) - Y^i(p) \frac{\partial X^j}{\partial x_i}(p) \right) \partial_j|_p. \tag{9}$$
4. The coordinate local vector fields $\partial_i, \partial_j$ always satisfy $[\partial_i, \partial_j] = 0$.
5. The **Jacobi identity**: for $X, Y, Z \in \mathcal{X}(M)$,
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

◄ (1) is clear. To prove (2), we first compute $X(gY)$. Let $\ell \in C^\infty(M)$. Then

$$(X(gY))\,\ell = X(g \cdot Y\ell) = Xg \cdot Y\ell + g \cdot (XY)\ell,$$

and hence $X(gY) = (Xg) \cdot Y + g \cdot XY$. Then

$$\begin{aligned} [fX, gY] &= fX(gY) - gY(fX) \\ & f(Xg)Y + fgXY - g(Yf)X + gfYX \\ & fg[X, Y] + f(Xg)Y - g(Yf)X. \end{aligned}$$

Next, to prove (3) we compute

$$X(Yf) = X\left(Y^j \frac{\partial f}{\partial x^j}\right) = X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

Symmetry of the mixed partial derivatives then proves (3). Finally both (4) and (5) follow from (3). ►

## 2.7 Corollary

$\mathcal{X}(M)$ is an infinite dimensional Lie algebra under $[\cdot,\cdot]$.

## 2.8 Definition

Let $\Phi : M \to N$ be a diffeomorphism and $X \in \mathcal{X}(M)$. We can define the **pushforward** of $X$ to $N$ under $\Phi$ to be $\Phi_* X$ defined by

$$(\Phi_* X)(q) = d\Phi\left(X_{\Phi^{-1}(q)}\right) \in T_q(N). \tag{10}$$

Thus if $f \in C^\infty(N)$ we have

$$(\Phi_* X) f = X(f \circ \Phi) \circ \Phi^{-1}, \tag{11}$$

so

$$d\Phi \circ X = \Phi_* X \circ \Phi.$$

## 2.9 Lemma

Let $\Phi : M \to N$ be a diffeomorphism and $X, Y \in \mathcal{X}(M)$. Then

$$\Phi_*[X,Y] = [\Phi_* X, \Phi_* Y].$$

In other words, $\Phi$ defines a Lie algebra isomorphism $\Phi_* : \mathcal{X}(M) \to \mathcal{X}(N)$.

◀ Let $q \in N$ and set $p = \Phi^{-1}(q)$. Suppose $f \in C^\infty(N)$. Then we compute:

$$\begin{aligned}
(\Phi_*[X,Y])_q f &= d\Phi\left([X,Y]_p\right) f \\
&= [X,Y]_p(f \circ \Phi) \\
&= X_p(Y(f\circ\Phi)) - Y_p(X(f\circ\Phi)) \\
&= X_p((d\Phi \circ Y)(f)) - Y_p((d\Phi \circ X)(f)) \\
&= X_p((\Phi_* Y)(f) \circ \Phi) - Y_p((\Phi_* X)(f)\circ\Phi) \\
&= d\Phi\left(X_p\right)((\Phi_* Y)(f)) - d\Phi\left(Y_p\right)((\Phi_* X)(f)) \\
&= (\Phi_* X)_q((\Phi_* Y)(f)) - (\Phi_* Y)_q((\Phi_* X)(f)) \\
&= [\Phi_* X, \Phi_* Y]_q f.
\end{aligned}$$

An unpleasant calculation. Note that we used both characterisations (10) and (11) at different stages in the proof. ▶

## 2.10 Definition

Let $X \in \mathcal{X}(M)$ and $p \in M$. A smooth curve $c : (a,b) \to M$ (with $a < 0$ and $b > 0$) is called an **integral curve of $X$ at $p$** if

$$c(0) = p, \quad X(c(t)) = \dot{c}(t) \text{ for all } t \in (a,b).$$

## 2.11 Theorem (existence and uniqueness of integral curves)

Let $M$ be a smooth manifold, $p \in M$ and $X \in \mathcal{X}(M)$. Then there exists an open interval $I_p$ containing 0 and an integral curve $c_p : I_p \to M$ of $X$ such that $c_p(0) = p$. Moreover, if $\gamma : J \to M$ is another integral curve of $X$ such that $\gamma(0) = p$, with $J$ an open interval containing 0 then $J \subseteq I_p$ and $c_p|_J \equiv \gamma$.

◀ If $c : (a,b) \to M$ is a smooth curve, and $(x^1, \ldots, x^n)$ local coordinates about $c(t_0)$ then for $t$ close to $t_0$, recall by (7)we have

$$\dot{c}(t) = \frac{dc^i}{dr}(t)\partial_i|_{c(t)},$$

where $c^i = x^i \circ c$. Suppose in this chart $X(q) = X^i(q)\partial_i|_q$. Then the assertion that $c$ is an integral curve becomes

$$\frac{dc^i}{dr}(t)\partial_i|_{c(t)} = X^i(c(t))\partial_i|_{c(t)},$$

and thus the assertion comes down to solving the system of ODE's

$$\frac{dc^i}{dr}(t) = X^i(c(t)), \quad c^i(0) = X^i(p). \tag{12}$$

So choose a chart $(U, \phi)$ about $p$. Let $I_p$ denote the union of all the open intervals containing the origin which are the domains of integral curves of $X$ satisfying the initial condition that the origin maps to $p$. Applying the standard theorem on existence to solutions of ODE's to the system of equations (12) we see that $I_p \neq \emptyset$. Suppose now $c$ and $\gamma$ are integral curves of $X$ defined on open intervals $A, B$ respectively with $A \cap B \neq \emptyset$ . Then if there exists $t_0 \in A \cap B$ such that $c(t_0) = \gamma(t_0)$, then by the standard theorem on uniqueness to solutions of ODE's, the subset of $A \cap B$ on which $c$ and $\gamma$ is open and non-empty. By continuity it is also closed, and hence by connectedness it is equal to $A \cap B$. It follows there exists an integral curve $c_p$ of $X$ defined on all of $I_p$, which completes the proof. ▶

## 2.12 Definitions

Let $M$ be a smooth manifold, and $X \in \mathcal{X}(M)$. In the notation of the proof above, for $t \in \mathbb{R}$ define $U_t := \{p \in M \mid t \in I_p\}$ and define $\phi_t : M \to M$ by $\phi_t(p) = c_p(t)$.

## 2.13 Theorem (flow theorem)

For each $t$ it holds that:

1. $U_t$ is open, and $M = \bigcup_{t>0} U_t$.
2. If $t \in I_p$ then $I_{\phi_t(p)} = \{s - t | s \in I_p\}$.
3. $\phi_t : U_t \to U_{-t}$ is a diffeomorphism with inverse $\phi_{-t}$.
4. If $s, t \in \mathbb{R}$ then the domain of $\phi_s \circ \phi_t$ is contained in (but not generally equal to) $U_{s+t}$. If $s, t$ have the same sign however then we have equality. In any case, on the domain of $\phi_s \circ \phi_t$ we have $\phi_s \circ \phi_t = \phi_{s+t}$.
5. Given $p \in M$ there exists a maximal open neighborhood $V$ of $p$ and maximal $\epsilon > 0$ such that the map
$$\phi : (-\epsilon, \epsilon) \times V, \quad (t, q) \mapsto \phi_t(q) = \gamma_q(t)$$
is a well defined smooth map. Note by (3) we have for $s, t \in \mathbb{R}$ such that all of $|s|, |t|$ and $|s + t| < \epsilon$ this implies $\phi(s + t, p) = \phi(s, \phi_t(p))$. We say that $\phi$ is the **local flow** of $X$ at $p$.

We won't prove the Flow Theorem; it essentially follows from the ODE theorem on smooth the dependance on initial conditions.

## 2.14 Definitions

A smooth vector field $X$ is called **complete** if $U_t = M$ for all $t \in \mathbb{R}$, in other words, for all $p \in M$,the domain of $c_p$ is all of $\mathbb{R}$. In this case the transformations $\{\phi_t\}_{t\in\mathbb{R}}$ form a group of transformations of $M$ parametrized by the real numbers. It is called the **one-parameter group** *of* $X$. The associated map $\phi$ is then a (global) **flow**. Note that if $X$ is not complete, $\{\phi_t\}_{t\in\mathbb{R}}$ do not form a group, since their domains depend on $t$. In this case we refer to $\{\phi_t\}_t \in \mathbb{R}$ as the **local** one-parameter group of $X$.

## 2.15 Proposition

Let $X$ be a smooth vector field on a smooth manifold $M$. If $c$ is an integral curve of $M$ whose maximal domain is not all of $\mathbb{R}$, then the image of $c$ cannot lie in any compact subset of $M$.

◀ Let $(a, b)$ denote the maximal domain of $c$, and $p := c(0)$. Suppose $b < \infty$ but that the image of $c$ lies in a compact set $K$ of $M$ (the case $a > -\infty$ is similar). If $\{t_i\}$ is any sequence of times approaching $b$ from below, then $\{c(t_i)\}$ is a sequence of points in $K$, and thus passing to a subsequence if necessary we may assume $c(t_i) \to q \in K$. Choose a neighborhood $U$ of $q$ and $\epsilon > 0$ such that that the local flow $\phi$ of $X$ is defined on $(-\epsilon, \epsilon) \times U$ (so for $t \in (a, b) \cap (-\epsilon, \epsilon)$, $\phi(t, p) = c(t)$). Pick $i$ large enough such that $c(t_i) \in U$ and $t_i + \epsilon > b$, and define $\gamma : (a, t_i + \epsilon) \to M$ by

$$\gamma(t) = \begin{cases} c(t) & a < t < b \\ \phi(t - t_i, c(t_i)) & t_i - \epsilon < t < t_i + \epsilon. \end{cases}$$

By assumption these definitions agree on the overlap, since

$$\phi(t - t_i, c(t_i)) = \varphi(t - t_i, \phi_{t_i}(p)) = \varphi(t - t_i + t_i, p) = c(t).$$

and hence $\gamma$ is an integral curve of $X$ satisfying $\gamma(0) = p$ and defined on a larger interval then $(a, b)$. Contradiction. ▶

## 2.16 Corollary

If $M$ is a compact smooth manifold then every smooth vector field on $M$ is complete.

# 3 Submanifolds

## 3.1 Definition

Let $\Phi : M \to N$ be a smooth map between manifolds. We say that $\Phi$ is an **immersion** if $d\Phi(p)$ is injective for each $p \in M$.

## 3.2 Definitions

Let $\Phi : M \to N$ an injective immersion. Then the pair $(M, \Phi)$ is an **immersed submanifold** of $N$. If in addition $\Phi$ is a topological embedding, that is, $\Phi$ is a homeomorphism onto its image (with the subspace topology) then the pair $(M, \Phi)$ is an **embedded submanifold**. In this case we often suppress $\Phi$ and identify $M$ with its image $\Phi(M) \subseteq N$ and thus regard $\Phi$ as the inclusion $M \hookrightarrow N$.

Let $\Phi : M \to N$ be an embedding. Then for $p \in M$, the map $d\Phi(p) : T_p(M) \to T_{\Phi(p)}(N)$ identifies the tangent space $T_p(M)$ with a subspace of $T_{\Phi(p)}(N)$. Then $d\Phi : T(M) \to T(N)$ is a smooth embedding, so $T(M)$ is an embedded submanifold of $T(N)$.

## 3.3 Definitions

Let $\Phi : M^m \to N^n$ be smooth. A point $p \in M$ is called a **regular point** of $\Phi$ if $d\Phi(p)$ is surjective. If $p \in M$ is not a regular point then it is a **critical point**. A point $q \in N$ is a **regular value** of $\Phi$ if for any $p \in \Phi^{-1}(q)$, $p$ is a regular point. Note that if $\Phi^{-1}(q) = \emptyset$ then this condition is vacuously true. If $q \in N$ is not a regular value then it is a **critical value**. Note that if $m < n$ then any point $q \in N$ such that $f^{-1}(q)$ is non-empty is a critical value.

## 3.4 Theorem (implicit function theorem)

Let $U$ be a neighborhood of 0 in $\mathbb{R}^m$ and $\Phi : U \to \mathbb{R}^n$ a smooth map such that $\Phi(0) = 0$. Then:

1. If $m \leq n$, let $i : \mathbb{R}^m \to \mathbb{R}^n$ denote the canonical inclusion. If $\Phi$ has maximal rank $m$ at 0 then there exists a chart $h$ on $\mathbb{R}^n$ and a neighborhood $W$ of $0 \in \mathbb{R}^m$ such that $h \circ \Phi|_W = i|_W$.

2. If $m \geq n$, let $\pi : \mathbb{R}^m \to \mathbb{R}^n$ denote projection onto the first $n$ coordinates. If $\Phi$ has maximal rank $n$ at 0 then there exists a chart $k$ on $\mathbb{R}^m$ and a neighborhood $V$ of $0 \in \mathbb{R}^m$ such that $\Phi \circ k|_V = \pi|_V$.

◀ First we prove (1). The hypotheses imply that the $m \times n$ matrix $\left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]$ has rank $m$. Hence by rearranging the component functions $\Phi^i$ of $\Phi$ if necessary (which amounts to composing $\Phi$ with an invertible transformation of $\mathbb{R}^m$, which is a diffeomorphism) we may assume that the $m \times m$ minor $\left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{1 \leq i,j \leq m}$ is invertible. Define $F : U \times \mathbb{R}^{n-m} \to \mathbb{R}^n$ by

$$F(x^1, \ldots, x^m, x^{m+1}, \ldots, x^n) := \Phi(x^1, \ldots, x^m) + (0, \ldots, 0, x^{m+1}, \ldots, x^n).$$

Then $F \circ i = \Phi$, and the Jacobian matrix of $F$ at 0 is

$$\begin{bmatrix} \left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{1 \leq i \leq m} & 0 \\ \left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{m+1 \leq i \leq n} & I_{n-m} \end{bmatrix}$$

Thus by the inverse function theorem $F$ has a local inverse $h$, and $h \circ \Phi = h \circ F \circ i = i$. This proves (1).

Similarly in (2), we may assume that the $n \times n$ minor $\left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{1 \leq i,j \leq n}$ is invertible, and hence defining $G : U \times \mathbb{R}^{m-n} \to \mathbb{R}^m$ by

$$G(x^1, \ldots, x^m) := \left(\Phi(x^1, \ldots, x^m), x^{n+1}, \ldots, x^m\right).$$

Then $\Phi = \pi \circ G$, and the Jacobian matrix of $G$ at 0 is

$$\begin{bmatrix} \left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{1 \leq j \leq n} & \left[\frac{\partial \Phi^i}{\partial x^j}(0)\right]_{n+1 \leq j \leq m} \\ 0 & I_{m-n} \end{bmatrix}$$

Thus by the inverse function theorem $G$ has a local inverse $k$, and $\Phi \circ k = \pi \circ G \circ k = \pi$. ▶

## 3.5 Definition

Let $M^m$ be a submanifold of $N^n$. A chart $(U, h)$ on $N$ with local coordinates $(x^1, \ldots, x^n)$ is called a **slice chart for $M$ in $N$** if

$$M \cap U = \{p \in U \mid x^{m+1}(p) = \cdots = x^n(p) = 0\}.$$

## 3.6 Proposition (submanifold criterion)

Let $\Phi : M^m \to N^n$ be an immersion. Then for any $p \in M$, there exists a neighborhood $U$ of $p$ and a coordinate map $(V, k)$ defined on some neighborhood $V$ of $\Phi(p)$, with local coordinates $(y^1, \ldots, y^n)$ such that:

1. A point $q$ belongs to $\Phi(U) \cap V$ if and only if $y^{m+1}(q) = \cdots = y^n(q) = 0$, so

$$k\left(\Phi(U) \cap V\right) = (\mathbb{R}^m \times \{0\}) \cap k(V).$$

2. $\Phi|_U$ is an embedding.

If $\Phi$ was an embedding, then we may take $U = M$, and hence

$$\Phi(M) \cap V = \{q \in V \mid y^{m+1}(q) = \cdots = y^n(q) = 0\}.$$

In other words, an immersed submanifold $M^m \subseteq N^n$ is one such that every point $p \in M$ has a chart $(V, k)$ about $\Phi(p)$ and a neighborhood $U$ such that $k$ is a slice chart for $U$ in $N$, and an embedded submanifold is one such that every point $p \in M$ has a chart $(V, k)$ about $\Phi(p)$ such that

$k$ is a slice chart for $M$ in $N$.

◀ Let $i : \mathbb{R}^m \hookrightarrow \mathbb{R}^n$ denote the inclusion. Let $h$ be a chart centred about $p$, and $\ell$ a chart centred about $\Phi(p)$. Since $\ell \circ \Phi \circ h^{-1}$ has maximal rank $m$ at 0, by the implicit function theorem there exists a chart $\ell'$ of $\mathbb{R}^n$ and a neighborhood $W$ of $0 \in \mathbb{R}^m$ such that

$$\ell' \circ \ell \circ \Phi \circ h^{-1}|_W = i|_W.$$

Then set $U = h^{-1}(W)$ and $k = \ell' \circ \ell$. Then by restricting the domain of $\ell'$ if necessary, (1) clearly holds. Furthermore, since $\Phi|_U = k^{-1} \circ i \circ h$ is the composition of embeddings, so is $\Phi|_U$.

Finally, if $\Phi$ was an embedding then $\Phi(U) = \Phi(M) \cap V'$ for some open set $V' \subseteq N$, and hence $\Phi(U) \cap V = \Phi(M) \cap (V \cap V')$, and so replacing $V$ by $V \cap V'$, the last statement follows. ▶

## 3.7 Proposition

Every regular level set of a smooth map is an embedded submanifold.

◀ Suppose $\Phi : M^m \to N^n$ is smooth. Suppose initially that $\Phi$ is a submersion. If $q \in N$, and $p \in P := \Phi^{-1}(q)$, then by the Theorem 3.4 there are charts $(U, h)$ of $M$ about centred about $p$ and $(V, k)$ on $N$ centered about $q$ such that $\Phi$ has local expression $(y^1, \dots, y^m) \mapsto (y^1, \dots, y^n, 0, \dots, 0)$ in these two charts. Thus $P \cap U$ is the slice $\{(x^1, \dots, x^m) \in U \mid x^{n+1} = \cdots = x^m = 0\}$, and so $P$ is an embedded submanifold of $M$.

Now we consider the general case, and drop the assumption that $\Phi$ is a submersion. If $q$ is any regular value of of $\Phi$ such that $P := \Phi^{-1}(q) \neq \emptyset$, then for each $p \in P$, $d\Phi(p)$ has rank $n$. Let $U := \{p \in M \mid d\Phi(p) \text{ has rank } n\}$. Then $P \subseteq U$, and we will show that $U$ is open. Given this, $\Phi|_U : U \to N$ is a submersion, and we can apply the above to conclude that $P$ is an embedded submanifold of $U$, and hence also an embedded submanifold of $M$.

If $p \in U$, then the determinant of some $n \times n$ minor of the $n \times m$ matrix representing $d\Phi(p)$ in some smooth local coordinates is non-zero. Since the determinant is continuous, there is a neighborhood $V$ of $p$ such that this minor has non-zero determinant, and thus $V \subseteq U$ and $U$ is open. ▶

## 3.8 Theorem (Sard)

If $\Phi : M \to N$ is smooth then the set of critical values of $\Phi$ has measure zero in $N$.

We will not prove this theorem in this course.

## 3.9 Theorem

Let $\Phi : M^m \to N^n$ be a smooth map, with $m \geq n$is smooth. If $q \in N$ is a regular value of $\Phi$ such that $P := \Phi^{-1}(q) \neq \emptyset$, then $P$ is a topological manifold of dimension $m - n$. Moreover there exists a unique smooth for which $(P, i)$ (where $i : P \hookrightarrow M$ is inclusion) becomes an embedded submanifold of $M$.

◀ Let $k : V \to \mathbb{R}^n$ be a chart on $N$ centred about $q$; given $p \in P$, let $h : U \to \mathbb{R}^m$ be a chart on $M$ centred about $p$. Decompose $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n}$, and let $\pi_1 : \mathbb{R}^m \to \mathbb{R}^n$ and $\pi_2 : \mathbb{R}^m \to \mathbb{R}^{n-m}$ denote the projections onto the two factors. Let $i_2 : \mathbb{R}^{m-n} \hookrightarrow \mathbb{R}^n$ denote the inclusion

$$\left(a^1, \dots, a^{m-n}\right) \mapsto \left(0, \dots, 0, a^1, \dots, a^{m-n}\right).$$

Since $k \circ \Phi \circ h^{-1}$ has maximum rank at $0 \in \mathbb{R}^m$ by Theorem 3.4 there exists a chart $(W, \ell)$ around $0 \in \mathbb{R}^m$ such that $k \circ \Phi \circ h^{-1} \circ \ell = \pi_1|_W$. Let $W' := \pi_2(W)$, so $W'$ is open in $\mathbb{R}^{m-n}$ and $k \circ \Phi \circ h^{-1} \circ \ell \circ i_2|_{W'} = \pi_1 \circ i_2|_{W'} = 0$. Thus if $j := h^{-1} \circ \ell \circ i_2|_{W'}$then $j(W') \subseteq P$. In fact, we claim that

$$j(W') = P \cap \left(h^{-1} \circ \ell\right)(W),$$

so that $j$ maps $W'$ homeomorphically onto a neighborhood of $p \in P$ in the subspace topology. Indeed, clearly $j(W') \subseteq P \cap \left(h^{-1} \circ \ell\right)(W)$, since

$$j(W') = \left(h^{-1} \circ \ell \circ i_2\right)(W) = h^{-1} \circ \ell\left(W \cap \left(0 \times \mathbb{R}^{m-n}\right)\right).$$

Conversely, if $p' \in P \cap \left(h^{-1} \circ \ell\right)(W)$, then $p' = h^{-1} \circ \ell(u)$ for some unique $u \in W$, and since

$$0 = k \circ \Phi(p') = \left(k \circ \Phi \circ h^{-1} \circ \ell\right)(u) = \pi_1(u),$$

we have

$$u = (0, a) \in 0 \times W'$$

for some $a \in W'$, and thus $p' = j(a) \in j(W')$. It follows that the inclusion $i : P \hookrightarrow M$ is a topological embedding.

Finally, we endow $P$ with the smooth structure induced by the charts $\left\{\left(j(W'), j^{-1}\right)\right\}$ as $p$ ranges over $P$. Then $i : P \hookrightarrow M$ is smooth, since $h \circ i \circ \left(j^{-1}\right)^{-1} = \ell \circ i_2$. ▶

## 3.10 Corollary

Let $\Phi : M^m \to N^n$ be smooth, $q \in N$ a regular value and $P = \Phi^{-1}(q) \neq \emptyset$. Then for $p \in P$, we have

$$di\left(T_p(P)\right) = \ker d\Phi(p).$$

◀ Since both subspaces have common dimension $m - n$, it suffices to check that $di(T_p(P)) \subseteq \ker d\Phi(p)$. Let $v \in T_p(P)$. Then for $f \in C^\infty_{N,q}$ we have

$$d\Phi(di(v))(f) = d(\Phi \circ i)(v)(f) = v(f \circ \Phi \circ i).$$

But $\Phi \circ i \equiv q$, and hence $f \circ \Phi \circ i$ is a constant function, and thus $v(f \circ \Phi \circ i) = 0$. ▶

## 3.11 Example

$GL(n, \mathbb{R})$ is an open subset of $\mathrm{Mat}(n, \mathbb{R}) = \mathbb{R}^{n^2}$. The set $\mathrm{Sym}(n, \mathbb{R})$ of symmetric matrices may be identified with $\mathbb{R}^{\frac{n(n+1)}{2}}$. Now define $\Phi : GL(n, \mathbb{R}) \to \mathrm{Sym}(n, \mathbb{R})$ by $\Phi(A) = AA^t$. Observe that $\Phi^{-1}(I) = O(n)$, the real orthogonal $n \times n$ matrices. We claim that $I$ is a regular value of $\Phi$.

First for $A \in GL(n, \mathbb{R})$ we can define a diffeomorphism $R_A : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ by right multiplication: $R_A(X) = XA$. Now observe that if $A \in O(n)$ then $\Phi \circ R_A = \Phi$. Thus by the chain rule, we have $d\Phi(A) \circ dR_A(I) = d\Phi(I)$. Since $R_A$ is a diffeomorphism this shows that $\mathrm{rank}(d\Phi(A)) = \mathrm{rank}(d\Phi(I))$. It is therefore enough to check that $d\Phi(I)$ is surjective.

Now observe that $T_I\left(GL(n, \mathbb{R})\right) \cong T_I\left(\mathrm{Mat}(n, \mathbb{R})\right) \cong \mathrm{Mat}(n, \mathbb{R})$. Specifically, if we take global coordinates $(x^i_j)$ on $GL(n, \mathbb{R})$, where

$$x^i_j(A) = a^i_j, \quad A = \left[a^i_j\right],$$

we send

$$a^i_j \frac{\partial}{\partial x^i_j}\Big|_I \leftrightarrow \left[a^i_j\right] \in \mathrm{Mat}(n, \mathbb{R}).$$

We can represent an arbitary element $A \in T_I\left(GL(n, \mathbb{R})\right)$ by the curve $c_A : t \mapsto I + tA$ (note that $c(t) \in GL(n, \mathbb{R})$ for small enough $t$). Now since

$$\Phi(I + tA) = I + t(A + A^t) + t^2 AA^t$$

we have

$$d\Phi(I)(A) = \frac{d}{dt}\Big|_{t=0}(\Phi(I + tA)) = A + A^t.$$

Now finally let $S$ by an arbitrary symmetric matrix, then if $A = S/2$ we have $A + A^t = S$. Thus $d\Phi(I)$ is surjective, and thus $O(n)$ is an embedded submanifold of $GL(n, \mathbb{R})$ of dimension $\frac{n(n-1)}{2}$.

## 3.12 Distributions

Let $M$ be a smooth manifold. A $r$-dimensional **distribution** $\mathcal{D}$ on $M$ is a choice of an $r$-dimensional subspace $\mathcal{D}(p) \subseteq T_p(M)$ for each $p \in M$. A distribution $\mathcal{D}$ is smooth if $\mathcal{D}(M) := \coprod_{p\in M} \mathcal{D}(p)$ is a smooth subbundle of $T(M)$ (see Chapter 6 for the precise definition of subbundles). Equivalently, $\mathcal{D}$ is smooth if every $p \in M$ has a neighborhood $U$ and smooth vector fields $X_1, \ldots, X_r \in \mathcal{X}(U)$ such that $\{X_1(q), \ldots, X_r(q)\}$ is a basis for $\mathcal{D}(q)$ for each $q \in U$.

We say a smooth vector field $X \in \mathcal{X}(U)$ (where $U \subseteq M$ is an open subset) **belongs** to $\mathcal{D}$ if $X_p \in \mathcal{D}(p)$ for all $p \in U$. A distribution $\mathcal{D}$ is called **involutive** if given $X, Y \in \mathcal{X}(U)$ belonging to $\mathcal{D}$ we also have $[X, Y] \in \Theta(U)$ belonging to $\mathcal{D}$. Note that a one-dimensional distribution is just a vector field.

If $(M, f)$ is an immersed submanifold of $N$ and $\mathcal{D}$ is a distribution on $N$, we call $(M, f)$ an **integral submanifold** of $\mathcal{D}$ if $d\Phi(T_p(M)) = \mathcal{D}(\Phi(p))$ for all $p \in M$. A distribution $\mathcal{D}$ on $N$ is called **completely integrable** if each point is contained in an integral submanifold of $N$. A **locally integrable distribution** is one such that every point $p \in N$ is contained in an integral submanifold of an open subset $U \subseteq N$. Note that an integral manifold of a one-dimensional distribution is just (the image of) a curve.

## 3.13 Theorem (Frobenius)

A distribution is involutive if and only if it is completely integrable.

## 3.14 Theorem (Whitney)

Any smooth manifold $M^n$ may be embedded in $\mathbb{R}^{2n}$.

We will not prove either of these theorems in this course. The version of Theorem 3.14 stated is a truly difficult result. An earlier (and much easier) result also due to Whitney states that we can embed $M^n$ in $\mathbb{R}^{2n+1}$.

# 4 Lie groups

## 4.1 Definitions

A **Lie group** $G$ is a smooth manifold endowed with a group structure such that the multiplication map $m : G \times G \to G, (p, q) \mapsto pq$ and the inversion map $i : G \to G, p \mapsto p^{-1}$ are smooth maps.

A **Lie subgroup** $H$ of $G$ is the image of an immersed submanifold $(H', \Phi)$ of $G$ such that $H'$ is a Lie group, and $\Phi : H' \to G$ is a homomorphism of the (abstract) groups $H'$ and $G$. Thus if $H \leq G$ is a Lie subgroup then the inclusion $i : H \hookrightarrow G$ is an immersion. It is a non-trivial fact that a closed subgroup $H \leq G$ is a Lie group with respect to the subspace topology.

## 4.2 Examples

1. From the previous lecture, $GL(n, \mathbb{R})$, $O(n)$ and $SO(n)$ are all Lie groups.

2. The $n$-torus $T^n = \mathbb{R}^n/\mathbb{Z}$ is an abelian Lie group (the group structure is induced by addition on $\mathbb{R}^n$). In fact, any compact abelian Lie group is a torus.

3. A given manifold can carry multiple Lie group structures. In addition to the standard one, we can make $\mathbb{R}^3$ into a Lie group by defining $m(a, b) = (a_1 + b_1, a_2 + b_2 + a_1 b_3, a_3 b_3)$ (where $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3)$). This corresponds to identifying $\mathbb{R}^3$ with the subgroup $H$ of $GL(3, \mathbb{R})$ consisting of matrices of the form

$$\begin{pmatrix} 1 & a_1 & a_2 \\ 0 & 1 & a_3 \\ 0 & 0 & 1 \end{pmatrix}.$$

The fact that this is a subgroup confirms that $m$ does indeed endow $\mathbb{R}^3$ with a group structure.

## 4.3 Definition

Let $G$ be a Lie group. Set $\mathfrak{g} := T_e(G)$ (where $e$ is the identity element of $G$). $\mathfrak{g}$ is called the **Lie algebra** of $G$. Our initial aim is to show that $\mathfrak{g}$ is indeed a Lie algebra (and thus the definition is not completely inane).

## 4.4 Computing the Lie algebra $\mathfrak{o}_n$ of $O(n)$.

Recall from Example 3.11 that $O(n)$ is an embedded submanifold of $GL(n,\mathbb{R})$ of dimension $\frac{n(n-1)}{2}$. Thus $\mathfrak{o}_n := T_I(O(n))$ is naturally a subspace of $\mathfrak{gl}_n := T_I(GL(n,\mathbb{R})) = \mathrm{Mat}(n,\mathbb{R})$. We can no longer represent a element of $\mathfrak{o}_n$ by a curve of the form $c(t) = I + tA$ for some $A \in \mathrm{Mat}(n,\mathbb{R})$, as even for small $t$, $I + tA$ has no reason to lie in $O(n)$. However by elementary calculus, we can write

$$c(t) = I + tA + O(t^2).$$

Then if we require $c(t) \in O(n)$ then we need

$$(I + tA + O(t^2))(I + tA + O(t^2))^t = I,$$

or equivalently $A + A^t = 0$. This gives $\mathfrak{o}_n \subseteq \{A \in \mathfrak{gl}_n \mid A + A^t = 0\}$ and then counting dimensions gives equality. Alternatively one could proceed as in Example 3.11, with the map $d\Phi : T_I(GL(n,\mathbb{R})) \to T_I(\mathrm{Sym}(n,\mathbb{R}))$, and then use Corollary 3.10 to conclude that $T_I(O(n))$ was the kernel of $d\Phi(I)$. Since $d\Phi(I)(A) = A + A^t$, we recover the same result.

## 4.5 Definition

Let $G$ be a Lie group. Let $\ell_p : G \to G$ denote the diffeomorphism $q \mapsto pq$. Let $X \in \mathcal{X}(G)$. We say that $X$ is a **left-invariant vector field** if $\ell_{p*}X = X$ for all $p \in G$. We let $\mathcal{X}_\ell(G)$ denote the set of left-invariant vector fields on $G$. In more detail, this means that for any $q \in G$ we require for $f \in C^\infty_{pq}$,

$$X_{pq}f = (\ell_{p*}X)_{pq} f = d\ell_p(X_q)(f) = X_q(f \circ \ell_p).$$

Suppose $X \in \mathcal{X}_\ell(G)$, and let $\xi = X_e$. Let $f \in C^\infty_p$. Then

$$X_p f = X_e(f \circ \ell_p) = d\ell_p(X_e)(f) = d\ell_p(\xi)(f).$$

It follows that if we define a section $X_\xi : G \to T(G)$ by, for $p \in G$ and $f \in C^\infty_p$,

$$X_\xi(p)f = d\ell_p(\xi)(f),$$

then $X_\xi = X$ and thus $X_\xi$ is a left-invariant vector field. Hence a left-invariant vector field $X$ is determined by $\xi = X_e$. In fact, if $\xi$ is an arbitrary element of $\mathfrak{g}$, then defining $X_\xi$ as above yields a left-invariant vector field.

## 4.6 Proposition

Let $G$ be a Lie group, and $\xi \in \mathfrak{g} = T_eG$. Define a section $X_\xi : G \to T(G)$ by

$$X_\xi(p) = d\ell_p(\xi) \in T_p(G). \tag{13}$$

Then $X_\xi$ is a vector field on $G$; moreover $X_\xi$ is left-invariant.

◀ We need only verify that $X_\xi$ is a smooth section of $\pi : T(G) \to G$ and left-invariant. Let $h_e : U'_e \to \mathcal{O}'$ be a chart centred about $e$, with coordinates $(x^1, \dots, x^n)$ $(n = \dim G)$. Fix a point $p \in G$. Observe that if $U'_p := \ell_p(U'_e)$ and $h_p := h_e \circ \ell_{p^{-1}}$ then $h_p : U'_p \to \mathcal{O}'$ is a chart centered about $p$. Let $(y^1, \dots, y^n)$ denote the coordinates of this chart. Now choose $U_e \subseteq U'_e$ such that if $U_p = \ell_p(U_e)$ then $m(U_p \times U_e) \subseteq U'_p$. We then have the following commutative diagram

$$\begin{array}{ccc} U_p \times U_e & \xrightarrow{m} & U'_p \\ \downarrow & & \downarrow \\ \mathcal{O} \times \mathcal{O} & \xrightarrow{F} & \mathcal{O}' \end{array}$$

where $F$ is defined to be the smooth composite at the bottom. Given $q \in U_p$, wee can restrict this to

$$\begin{array}{ccc} \{q\} \times U_e & \overset{\ell_q}{\to} & U'_p \\ \downarrow & & \downarrow \\ \{h_p(q)\} \times \mathcal{O} & \overset{F(h_p(q),\cdot)}{\to} & \mathcal{O}' \end{array}$$

Then the map $d\ell_q : T_e(G) \to T_q(G)$ is given, with respect to the bases $\left\{ \frac{\partial}{\partial x^i}\big|_e \right\}$ of $T_e(G)$ and $\left\{ \frac{\partial}{\partial y^j}\big|_q \right\}$ of $T_q(G)$ by

$$d\ell_q \left( \frac{\partial}{\partial x^i}\big|_e \right) = JF\left(h_p(q), 0\right)^j_i \frac{\partial}{\partial y^j}\big|_q,$$

where $JF(h_p(q), 0)$ is the Jacobian matrix of $F$ at the point $(h_p(q), 0) \subseteq \mathcal{O} \times \mathcal{O}$ (cf. (6)). Since the entries of $JF(h_p(q), 0)$ are smooth functions of $q \in U_g$, we have shown that for any fixed $\xi \in T_e(G)$, the images $d\ell_q(\xi)$ depend smoothly on $q$. Thus $X_\xi$ is smooth.

Finally, we check $X_\xi$ is left-invariant. Indeed,

$$(\ell_{p*} X_\xi)(pq) = d\ell_p(X_\xi(q)) = d\ell_p \circ d\ell_q(\xi) = d\ell_{pq}(\xi) = X_\xi(pq).$$

This completes the proof. ▶

## 4.7 Corollary

We have a linear isomorphism between $\mathfrak{g}$ and $\mathcal{X}_\ell(G)$ given by $\xi \mapsto X_\xi$.

## 4.8 Corollary

Let $G^n$ be a Lie group, and $\mathfrak{g}$ its Lie algebra. Then $\mathfrak{g}$ is a $n$-dimensional Lie algebra (!) under the bracket induced $\mathfrak{g}$ from the bracket on $\mathcal{X}_\ell(G)$ inherited from $\mathcal{X}(G)$ .

◀ We need only show that if $X_\xi$ and $X_\eta$ are left invariant then so is $[X_\xi, X_\eta]$. But by Lemma 2.9, for any $p \in G$, since $\ell_p$ is a diffeomorphism we have

$$\ell_{p*}[X_\xi, X_\eta] = [\ell_{p*}X_\xi, \ell_{p*}X_\eta] = [X_\xi, X_\eta],$$

so $[X_\xi, X_\eta]$ is left-invariant. Thus we can **define** a Lie bracket on $\mathfrak{g}$ by setting $[\xi, \eta] = \zeta$ where $[X_\xi, X_\eta] = X_\zeta$. ▶

## 4.9 The commutator bracket

Suppose now that $G$ is a **matrix Lie group**, that is, $G$ is a Lie group and a **closed** subgroup of $GL(n, \mathbb{R})$. Then as we have seen, $T_e(G)$ can be identified with a subspace of $\mathrm{Mat}(n, \mathbb{R})$. We already have a bracket on $\mathrm{Mat}(n, \mathbb{R})$, namely the **commutator** $[A, B] := AB - BA$. In fact, these two brackets coincide.

◀ In this section we'll prove the case when $G = GL(n, \mathbb{R})$. Using the matrix entries $\left(x^i_j\right)$ as global coordinates on $GL(n, \mathbb{R})$, the natural isomorphism $T_I(GL(n, \mathbb{R})) \leftrightarrow \mathfrak{gl}_n$ takes the form

$$a^i_j \frac{\partial}{\partial x^i_j}\big|_I \leftrightarrow \left(a^i_j\right).$$

Any matrix $A = \left(a^i_j\right)$ defines a left-invariant vector field $X_A$, defined by (for $P \in GL(n, \mathbb{R})$)

$$X_A(P) = d\ell_P \left( a^i_j \frac{\partial}{\partial x^i_j}\big|_I \right).$$

Under the above identification the map $d\ell_P(I)$ is just left multiplication by $P = \left(p^i_j\right)$, and thus in coordinates

$$X_A(P) = p^i_j a^j_k \frac{\partial}{\partial x^i_k}\big|_P. \tag{14}$$

Given two matrices $A, B$, using (9), the Lie bracket of the corresponding left-invariant vector fields is given by

$$\begin{aligned}[X_A, X_B](P) &= \left\{p^i_j a^j_k \frac{\partial}{\partial x^i_k}\left(p^\ell_m b^m_h\right) - p^i_j b^j_k \frac{\partial}{\partial x^i_k}\left(p^\ell_m a^m_h\right)\right\} \frac{\partial}{\partial x^\ell_h}|_P \\ & \left\{p^i_j a^j_k \delta^\ell_i \delta^k_m b^m_h - p^i_j b^j_k \delta^\ell_i \delta^k_m a^m_h\right\} \frac{\partial}{\partial x^\ell_h}|_P \\ & \left\{p^i_j a^j_k b^k_h - p^i_j b^j_k a^k_h\right\} \frac{\partial}{\partial x^i_h}|_P.\end{aligned}$$

In particular, if $P = I$ then we obtain

$$[X_A, X_B](I) = \left\{a^i_k b^k_h - b^i_k a^k_h\right\} \frac{\partial}{\partial x^i_h}|_P = [A, B]^i_h \frac{\partial}{\partial x^i_h}|_I = X_{[A,B]}(I).$$

Since a left-invariant vector field is determined by its value at the identity, this shows that $[X_A, X_B] = X_{[A,B]}$, which is precisely what we wanted to show. ▶

To deal with the case when $G$ is merely a subgroup of $GL(n, \mathbb{R})$ we consider the following more general situation.

## 4.10 Restriction to subgroups

Let $H \leq G$ be a Lie subgroup and let $i : H \hookrightarrow G$ denote inclusion. Then $i$ is an immersion, and so $di\,(T_e(H)) = di(\mathfrak{h})$ is a subspace of $\mathfrak{g}$. If $\xi \in \mathfrak{h}$ then letting $\bar{\xi} = di(\xi) \in di(\mathfrak{h})$ we have left-invariant vector fields $X_\xi \in \mathcal{X}_\ell(H)$ and $X_{\bar{\xi}} \in \mathcal{X}_\ell(G)$. If $\ell_{G,p}$ and $\ell_{H,p}$ denote the respective left multiplication diffeomorphisms for $p \in H$ then $\ell_{G,p} \circ i = i \circ \ell_{H,p}$ and thus

$$di\,(X_\xi(p)) = di \circ d\ell_{H,p}(\xi) = d\ell_{G,p} \circ di(\xi) = d\ell_{G,p}\left(\bar{\xi}\right) = X_{\bar{\xi}}(p).$$

We have previously only defined $\Phi_* X$ for $\Phi$ a diffeomorphism: if $\Phi : M \to N$ is smooth and $X \in \mathcal{X}(M)$ then it may still be the case that there exists a well defined vector field $Y \in \mathcal{X}(N)$ satisfying $Y_q = d\Phi\left(X_{\Phi^{-1}(q)}\right)$ - in this case we write $Y = \Phi_* X$ and say $X$ and $Y$ are **$\Phi_*$-related**.The difference is that if $\Phi$ is not a diffeomorphism there is no guarantee that such a vector field exists. However if it does, the proofs of our previous results (eg. Lemma 2.9) still go through.

Thus the computation above shows that $i_* X_\xi = X_{\bar{\xi}}$, and hence by Lemma 2.9 we conclude that $di\,(\mathfrak{h})$ is a Lie subalgebra of $\mathfrak{g}$, and $di : \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra homomorphism, that is

$$[X_\xi, X_\eta]_H = \left[X_{\bar{\xi}}, X_{\bar{\eta}}\right]_G.$$

In particular the bracket on $H$ is induced from that of $G$. Applying to this to $H \leq GL(n, \mathbb{R})$ we conclude that the bracket is again given by matrix commutation, as required.

## 4.11 Lemma

Let $G$ be a Lie group, and $X_\xi \in \mathcal{X}_\ell(G)$ be a left-invariant vector field associated to some $\xi \in \mathfrak{g}$. Let $\theta : (-\epsilon, \epsilon) \to G$ be the integral curve for $X_\xi$ such that $\theta(0) = e$. Then for $|s|, |t| < \epsilon/2$ we have $\theta(s + t) = \theta(s) \cdot \theta(t)$, where '$\cdot$' denotes group multiplication in $G$, that is $\theta(s + t) = \ell_{\theta(s)} \circ \theta(t)$.

◀ For fixed $s$, we show that the curves $t \mapsto \theta(s+t)$ and $t \mapsto \theta(s) \cdot \theta(t)$ are both integral curves of $X_\xi$ defined on $(-\epsilon/2, \epsilon/2)$ though $e$. Uniqueness of integral curves then forces equality.

Define $c(t) : (-\epsilon/2, \epsilon/2) \to G$ by $c(t) = \theta(s + t)$. Then certainly $c(0) = \theta(s)$. Moreover

$$\begin{aligned}\dot{c}(t) &= dc\left(\frac{d}{dr}|_t\right) \\ &= d\theta\left(\frac{d}{dr}|_{s+t}\right). \\ &= \dot{\theta}(s + t) \\ &= X_\xi\left(\theta(s + t)\right) \\ &= X_\xi(c(t)).\end{aligned}$$

Suppose now $p = \theta(s) \in G$. Define $\gamma(t) = \ell_p \circ \theta(t)$. Then $\gamma(0) = p \cdot e = p = \theta(s)$ and

$$\begin{aligned} \dot{\gamma}(t) &= d\left(\ell_p \circ \theta\right)\left(\frac{d}{dr}\Big|_t\right) \\ &= d\ell_p \circ d\theta\left(\frac{d}{dr}\Big|_t\right) \\ &= d\ell_p\left(\dot{\theta}(t)\right) \\ &= d\ell_p\left(X_\xi(\theta(t))\right) \\ &= X_\xi\left(\ell_p \circ \theta(t)\right) \\ &= X_\xi\left(\gamma(t)\right), \end{aligned}$$

where for the last but one equality we used left-invariance of $X_\xi$. Thus both $c$ and $\gamma$ are indeed integral curves of $X_\xi$ through $e$ defined on the same interval as claimed. ▶

## 4.12 Corollary

Any left-invariant vector field on a Lie group $G$ is complete.

◀ Let $X_\xi$ be an arbitrary left-invariant vector field on $G$, corresponding to $\xi \in \mathfrak{g}$.Let $\theta : (-\epsilon, \epsilon) \to G$ be the integral curve through $e$ for $X_\xi$. Define $c : \mathbb{R} \to G$ as follows. Given $t \in \mathbb{R}$, choose $N \in \mathbb{N}$ such that $t/N \in (-\epsilon/2, \epsilon/2)$. Define $c(t) = \theta(t/N)^N$. First, let use check $\theta$ is well defined. Suppose $N'$ was another such integer. Then since $\theta(t/NN')^{NN'} = \theta(t/NN')$ by Corollary 4.11, we have $\theta(t/N)^N = \theta(t/NN')^{NN'} = \theta(t/N')^{N'}$. Certainly $c(0) = e$, and given $t \in \mathbb{R}$, letting $p = \theta(t/N)^{N-1}$ as before we have

$$\begin{aligned} \dot{c}(t) &= d\left(\ell_p \circ \theta\right)\left(\frac{d}{dr}\Big|_t\right) \\ &= d\ell_p \circ d\theta\left(\frac{d}{dr}\Big|_t\right) \\ &= d\ell_p\left(\dot{\theta}(t)\right) \\ &= d\ell_p\left(X_\xi(\theta(t))\right) \\ &= X_\xi\left(\ell_p \circ \theta(t)\right) \\ &= X_\xi\left(c(t)\right). \end{aligned}$$

Thus $c$ is an integral curve of $X_\xi$ at $e$ defined on all of $\mathbb{R}$. By maximality, $\theta$ must also be defined on all of $\mathbb{R}$.

To complete the proof, we must show that for any $p \in G$, if $c_p(t)$ is the integral curve for $X_\xi$ with $c_p(0) = p$, then $c_p$ is also defined on all of $\mathbb{R}$. But given $p \in G$, define $\gamma : \mathbb{R} \to G$ by $t \mapsto p \cdot \theta(t) = \ell_p \circ \theta(t)$. Then certainly $\gamma(0) = p$, and an identical calculation to the above shows that $\gamma$ is an integral curve of $X_\xi$. Maximality then implies that $c_p$ is defined on all of $\mathbb{R}$. ▶

## 4.13 Definition

A **one-parameter subgroup** of a Lie group $G$ is a homomorphism of Lie groups $\theta : \mathbb{R} \to G$ (where $\mathbb{R}$ is given the additive group structure). If $\xi \in \mathfrak{g}$, the one-parameter subgroup **generated** by $\xi$ is the one-parameter subgroup $\theta(t)$ that is the integral curve through $e$ of $X_\xi$.

## 4.14 Proposition

There is a one-to-one correspondence between one-parameter subgroups of $G$ and left-invariant vector fields (and hence also with $\mathfrak{g}$).

◀ We already know that a left-invariant vector field gives rise to a one-parameter subgroup. Conversely, suppose $\theta : \mathbb{R} \to G$ is a one-parameter subgroup. Define $\xi = d\theta\left(\frac{d}{dr}\big|_0\right) \in T_e(G)$. Then we have a left-invariant vector field $X_\xi$, and to complete the proof it is enough to show that $\theta$ is an integral curve for $X$. We compute:

$$\begin{aligned} X_\xi\left(\theta(t)\right) &= d\ell_{\theta(t)}(\xi) \\ &= d\ell_{\theta(t)} \circ d\theta\left(\frac{d}{dr}\Big|_0\right) \\ &= d(\ell_{\theta(t)} \circ \theta)\left(\frac{d}{dr}\Big|_0\right) \\ &= d\theta\left(\frac{d}{dr}\Big|_t\right) \\ &= \dot\theta(t). \end{aligned}$$

This completes the proof. ▶

## 4.15 Definition

For any $A \in \mathrm{Mat}(n,\mathbb{R})$, define

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

## 4.16 Proposition

For any $A \in \mathrm{Mat}(n,\mathbb{R})$, $\exp(A) \in GL(n,\mathbb{R})$. Moreover, the one-parameter subgroup of $GL_n(\mathbb{R})$ generated by $A \in \mathfrak{gl}_n$ is $\theta(t) = \exp(tA)$.

◀ First let us check that $\exp(A)$ converges. We have $|AB| \le |A||B|$, where $|\cdot|$ is the norm induced from $\mathbb{R}^{n^2}$, and hence by induction $|A^k| \le |A|^k$. The Weierstrass M-test shows then shows that $\exp(A)$ converges uniformly on any bounded subset of $M_n(\mathbb{R})$ (by comparison with the series expansion for $e^{|A|}$).

Fix $A \in \mathfrak{gl}_n$. The one-parameter subgroup generated by $A$ is the integral curve $\theta(t)$ satisfying $\theta(0) = I$ and $\theta'(t) = X^{(A)}_{\theta(t)}$. Using (14), the condition for $\theta(t) = \left[\theta^i_j(t)\right]$ to be an integral curve is

$$\dot\theta^i_k(t) = \theta^i_j(t) a^j_k, \tag{15}$$

where $A = [a_{ij}]$. We claim that $\theta(t) := \exp(tA)$ satisfies this equation. Since $\theta(0) = I$, by uniqueness this proves that $\theta$ is the desired one-parameter subgroup.

First however we should check that $\theta(t)$ is a smooth $GL(n,\mathbb{R})$-valued curve. To check smoothness, we note that differentiating the series formally term by term gives

$$\dot\theta(t) = \sum_{k=1}^{\infty} \frac{k}{k!} t^{k-1} A^{k-1} = \left(\sum_{k=1}^{\infty} \frac{1}{(k-1)!}(tA)^{k-1} A\right) = \theta(t)A. \tag{16}$$

Since the differentiated series converges uniformly on bounded sets (because apart from the additional factor of $A$, it is the same series), this term by term differentiation is justified. A similar computation shows $\dot\theta(t) = A\theta(t)$. By smooth dependence of solutions of ODE's, $\theta$ is a smooth curve.

Finally we show that $\theta(t)$ is invertible for all $t$, so that $\theta$ actually takes its values in $GL_n(\mathbb{R})$. If $c(t) := \theta(t)\theta(-t)$, then $c$ is a smooth curve, and using the previous computation and the product rule,

$$c'(t) = (\theta(t)A)\theta(-t) - \theta(t)(A\theta(-t)) = 0,$$

and thus $\sigma$ is constant. Since $\sigma(0) = I$, we obtain $\theta(t)\theta(-t) = I$. Similarly, $\theta(-t)\theta(t) = I$.

Finally, (16) shows that $\theta(t)$ satisfies (15), and this completes the proof. ▶

# 5 Vector bundles and sheaves

## 5.1 Definition

Let $M$ be a smooth manifold. A smooth manifold $E$ together with a surjective smooth map $\pi : E \to M$ is called a **vector bundle of rank $m$ over $M$** if:

1. For each $p \in M$, $\pi^{-1}(p) =: E_p$ admits the structure of an $m$-dimensional (real) vector space.
2. Any $p \in M$ has an open neighborhood $U$ and a diffeomorphism $t : \pi^{-1}(U) \to U \times \mathbb{R}^m$ such that $\pi|_U = \mathrm{proj}_1 \circ t$ and such that for each $q \in U$, $t^q := \mathrm{proj}_1 \circ t|_{E_q} : E_q \to \mathbb{R}^m$ is a vector space isomorphism. The pair $(U, t)$ is called a **local trivialisation** of $E$.

$E$ is called the **total space**, and $M$ is called the **base space**. We often refer to $E$ as 'the' vector bundle. If $m = 1$, we call $E$ a **line bundle**. What we have actually defined are **real** vector bundles; there is a similar concept of complex vector bundles. In this course however we shall mainly be interested in real vector bundles. We say that a vector bundle $\pi : E \to M$ of rank $m$ is **trivial** if we can find a local trivialisation defined on all of $M$, that is, a trivialisation $t : E \to M \times \mathbb{R}^m$.

## 5.2 Proposition

Let $\pi : E \to M$ be a smooth vector bundle of rank $m$, and $(U, t)$ and $(V, \tau)$ two local trivialisations of $E$, such that $U \cap V \neq \emptyset$. Then there exists a smooth map $\psi : U \cap V \to GL(m, \mathbb{R})$ such that $\tau \circ t^{-1} : (U \cap V) \times \mathbb{R}^m \to (U \cap V) \times \mathbb{R}^m$ has the form $(p, v) \mapsto (p, \psi(p)(v))$, where $\psi(p)(v)$ denotes the usual action of the $r \times r$ matrix $\psi(p)$ on the vector $v \in \mathbb{R}^m$.

◀ It is clear that $\tau \circ t^{-1}(p, v) = (p, \sigma(p.v))$ for some smooth map $\sigma : (U \cap V) \times F \to F$. Moreover, for each fixed $p \in U \cap V$, the map $v \mapsto \sigma(p, v)$ is a linear isomorphism of $F$, so there is exists a map $\psi(p) \in GL(m, \mathbb{R})$ such that $\sigma(p, v) = \psi(p)(v)$. It remains to show that the map $\psi : U \cap V \to GL(m, \mathbb{R})$ is smooth.

To see this, pick a basis $\{e_i\}$ of $F$, so that we may identify $F$ with $\mathbb{R}^m$ and write $v = v^j e_j$,

$$\psi(p)(v) = \psi(p)^i_j v^j e_i.$$

Then $\psi(p)^i_j = r^i(\sigma(p, e_j))$ where $r^i : \mathbb{R}^m \to \mathbb{R}$ is projection onto the $i$th coordinate. This is smooth by composition. Since the matrix entries are smooth (global) coordinates on $GL(m, \mathbb{R})$, this shows that $\psi$ is smooth. ▶

## 5.3 Definitions

We call the smooth map $\psi$ the **transition function** between the local trivialisations $t$ and $\tau$. More generally, if $\pi : E \to M$ is an $m$-dimensional vector bundle then we say $E$ has **cocycle** $\{U_\alpha, \psi_{\alpha\beta}\}$ if $\{U_\alpha\}$ is an open cover of $M$, such that there exists local trivialisations $t_\alpha : \pi^{-1}(U_\alpha) \to U \times \mathbb{R}^m$ with transition functions $\psi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m, \mathbb{R})$ for $\alpha, \beta$ such that $U_\alpha \cap U_\beta \neq \emptyset$.

Observe that if $E$ has cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$ then the $\psi_{\alpha\beta}$ satisify:

$$\begin{cases} \psi_{\alpha\alpha}(p) = \mathrm{id} & \text{for all } \alpha \text{ and all } p \in U_\alpha, \\ \psi_{\alpha\beta}(p)\psi_{\beta\alpha}(p) = \mathrm{id} & \text{for all } \alpha, \beta \text{ with } U_\alpha \cap U_\beta \neq \emptyset \text{ and all } p \in U_\alpha \cap U_\beta, \\ \psi_{\alpha\beta}(p)\psi_{\beta\gamma}(p)\psi_{\gamma\alpha}(p) = \mathrm{id} & \text{for all } \alpha, \beta \text{ with } U_\alpha \cap U_\beta \cap U_\alpha \neq \emptyset \text{ and all } p \in U_\alpha \cap U_\beta \cap U_\gamma. \end{cases} \quad (17)$$

These are called the **cocycle conditions**. More generally, if $M$ is a smooth manifold and $\{U_\alpha, \psi_{\alpha\beta}\}$ a collection such that $\{U_\alpha\}$ is an open cover of $M$ and for $\alpha, \beta$ such that $U_\alpha \cap U_\beta \neq \emptyset$ the $\psi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m, \mathbb{R})$ are smooth matrix-valued functions satisfying the cocycle conditions (17) then we will still call $\{U_\alpha, \psi_{\alpha\beta}\}$ a cocycle.

## 5.4 Examples

1. We have so far met one example of a vector bundle in this course; namely the tangent bundle $\pi : T(M) \to M$. Examining the proof of Theorem 1.25 it is immediate that the tangent bundle $T(M)$ of an $n$-dimensional smooth manifold $M$ is indeed a vector bundle of rank $n$ over $M$. It is clear from the proof of Theorem 1.25 that if $\{U_\alpha, h_\alpha\}$ is an atlas for $M$ then $T(M)$ has cocycle $\left\{U_\alpha, J(h_\alpha \circ h_\beta^{-1}) \circ h_\beta\right\}$ where $J(h_\alpha \circ h_\beta^{-1})$ is the Jacobian matrix of $h_\alpha \circ h_\beta^{-1}$.

2. Here is an explicit example of a vector bundle: the **Hopf line bundle** (sometimes called the **tautological bundle**) over $\mathbb{C}P^n$. First, $\mathbb{C}P^n$ is an $n$-dimensional **complex manifold**: that is, the charts are maps from open sets in $\mathbb{C}P^n$to open sets of $\mathbb{C}^n$, and the transition functions are **holomorphic**. Indeed, $\mathbb{C}P^n$ has an open cover $\{U_i\}_{i=0}^n$ where

$$U_i := \left\{ \left(z^0 : \cdots : z^n\right) \in \mathbb{C}P^n \mid z^i \neq 0 \right\},$$

and the chart $h_i : U_i \to \mathbb{C}^n$ carries

$$h_i : \left(z^0 : \cdots : z^n\right) \mapsto \left( \frac{z^0}{z^i}, \ldots, \widehat{\frac{z^i}{z^i}}, \ldots, \frac{z^n}{z^i} \right) \in \mathbb{C}^n.$$

Moreover the transition function

$$h_i \circ h_j^{-1} : h_j \left(U_i \cap U_j\right) \to h_i \left(U_i \cap U_j\right)$$

for $i < j$ is the map

$$\left(w^1, \ldots, w^n\right) \mapsto \left( \frac{w^1}{w^i}, \ldots, \frac{w^{i-1}}{w^i}, \frac{w^{i+1}}{w^i}, \ldots, \frac{w^{j-1}}{w^i}, \frac{1}{w^i}, \frac{w^j}{w^i}, \ldots, \frac{w^n}{w^i} \right),$$

which is evidentally holomorphic on $h_j(U_i \cap U_j)$. Thus $\mathbb{C}P^n$ is an $n$-dimensional complex manifold as claimed (contrast this with Example 4 in Section 1.6).
We define the Hopf bundle $E$ to be

$$E = \coprod_{p \in \mathbb{C}P^n} E_p,$$

where $E_p$ is the line in $\mathbb{C}^{n+1}$ that represents the point $p \in \mathbb{C}P^n$, and we let $\pi$ be the map carrying $E_p$ onto $p$. For notational simplicity, in what follows we shall assume that $n = 1$ (the general case is similar, the only difference is essentially harder notation).
We have the following obvious trivialisation. On $U_0$, we may write any point as $(1 : z)$ for some $z \in \mathbb{C}$, and

$$E_{(1:z)} = \{(w, wz) | w \in \mathbb{C}\}.$$

Define $t_0 : \pi^{-1}(U_0) \to U_0 \times \mathbb{C}$ by

$$(w, wz) \mapsto ((1 : z), w).$$

Similarly, a point in $U_1$ may be written as $(\zeta : 1)$ for some $\zeta \in \mathbb{C}$, and

$$E_{(\zeta:1)} = \{(v\zeta, v) | v \in \mathbb{C}\},$$

and we define $t_1 : \pi^{-1}(U_1) \to U_1 \times \mathbb{C}$ by

$$(v\zeta, v) \mapsto ((\zeta; 1), v).$$

On $U_0 \cap U_1$ we have $(1 : z) = (\zeta : 1)$ if and only if $\zeta = {}^1\!/\!_z$, and if this is the case then $(w, wz) \in E_{(1:z)} = (v\zeta, v) \in E_{(\zeta:1)}$ if and only if $v = wz$. Thus $\psi_{01} : U_0 \cap U_1 \to GL_1(\mathbb{C}) = \mathbb{C}^*$ is the map defined by

$$\psi_{10}(z)(w) = wz.$$

Similarly $\psi_{01}(\zeta)(v) = v\zeta$. This the Hopf bundle $E$ is a complex line bundle over $\mathbb{C}P^n$.

## 5.5 Definitions

Given two bundles $\pi : E \to M$ and $\pi' : E' \to M'$, a **bundle morphism** from $E$ to $E'$ is a pair of smooth maps $F : E \to E', \Phi : M \to M'$ such that $\Phi \circ \pi = \pi' \circ F$ and such that the restriction to each fibre, $F|_{E_p} : E_p \to E'_{F(p)}$ is a linear map. If both maps are a diffeomorphism and the restrictions to fibre are all linear isomorphisms, then $F$ is called a **bundle isomorphism**. We say that $F$ **covers** $\Phi$.

Given two vector bundles $E, E'$ over the same base space $M$, we generally use a slightly more restrictive definition of a bundle morphism. Namely, a bundle morphism $F : E \to E'$ is a smooth map such that $\pi = \pi' \circ F$ and such that the restriction to each fibre, $F|_{E_p} : E_p \to E'_{F(p)}$ is a linear map (in other words we require $F$ to cover the identity map on $M$). If $F$ is a diffeomorphism and the restrictions to fibre are all linear isomorphisms, then $F$ is again called a bundle isomorphism. In this language, a vector bundle of rank $m$ is trivial if and only if it is isomorphic to the vector bundle $M \times \mathbb{R}^m$.

## 5.6 Lemma

Let $E$ and $E'$ be vector bundles over $M$ with the same cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$. Then $E$ and $E'$ are isomorphic as vector bundles.

◀ Suppose $E$ and $E'$ have (necessarily common) rank $m$. Let $\{t_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^m\}$ and $\{t'_\alpha : \pi'^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^m\}$ be the local trivialisations of $E$ and $E'$ respectively. Given $\alpha$, define a map $F_\alpha : \pi^{-1}(U_\alpha) \to \pi'^{-1}(U_\alpha)$ by

$$F_\alpha(z) = t'^{-1}_\alpha \circ t_\alpha(z).$$

Clearly $F_\alpha$ is a linear isomorphism on each fibre, and $\tau_i$ is a diffeomorphism. We show that on $\pi^{-1}(U_\alpha \cap U_\beta)$, the maps $F_\alpha$ and $F_\beta$ are equal. Indeed, if $z \in E_p$ so $t_\alpha(z) = (p, v)$ and $t_\beta(z) = (p, w)$ for some $v, w \in \mathbb{R}^m$, then $\psi_{\alpha\beta}(p)(w) = v$ and we compute:

$$\begin{aligned}
F_\alpha(z) &= t'^{-1}_\alpha \circ t_\alpha(z) \\
&= t'^{-1}_\alpha(p, v) \\
&= t'^{-1}_\alpha \left(p, \psi_{\alpha\beta}(p)(w)\right) \\
&= t'^{-1}_\beta \circ \left(t'_\beta \circ t'^{-1}_\alpha\right) \left(p, \psi_{\alpha\beta}(p)(w)\right) \\
&= t'^{-1}_\beta \left(p, \psi_{\beta\alpha}(p)\psi_{\alpha\beta}(p)(w)\right) \\
&= t'^{-1}_\beta(p, w) \\
&= t'^{-1}_\beta \circ t_\beta(z) \\
&= F_\beta(z).
\end{aligned}$$

Hence the maps $F_\alpha$ patch together to give us a well defined diffeomorphism $F : E \to E'$ that is a linear isomorphism on the fibres. ▶

Thus we have shown that a cocycle of a vector bundle determines the vector bundle up to isomorphism. In fact more is true - given a smooth manifold $M$ and a cocycle on $M$, we can always find a vector bundle $E$ with this cocycle. Thus giving a cocycle on $M$ is equivalent to specifying an isomorphism class of vector bundles over $M$. This is the content of the following theorem.

## 5.7 Theorem (the vector bundle construction theorem)

Let $M^n$ be a smooth manifold, and $\{U_\alpha | \alpha \in A\}$ an open cover of $M$ such that for all $\alpha, \beta$ such that $U_\alpha \cap U_\beta \neq \emptyset$ we have smooth maps $\psi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m, \mathbb{R})$ satisifying the cocycle conditions (17). Then there exists a vector bundle $\pi : E \to M$ of rank $m$ with cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$, where $E$ has dimension $n + m$. Moreover, $E$ is unique up to vector bundle isomorphism.

◀ Define

$$E = \{(p, \alpha, v) \in M \times A \times \mathbb{R}^m \mid p \in U_\alpha\} / \sim,$$

where $(p, \alpha, v) \sim (q, \beta, w)$ if and only if $p = q \in U_\alpha \cap U_\beta$ and $w = \psi_{\beta\alpha}(p)(v)$ (this is an equivalence relation due to the cocycle conditions). For $p \in M$, define

$$E_p = \pi^{-1}(p) = \{[p, \alpha, v] \mid p \in U_\alpha, v \in \mathbb{R}^m\}$$

where $[p, \alpha, v]$ denotes the equivalence class of $(p, \alpha, v)$ in $E$. Observe that $\pi : E \to M$ is surjective. Introduce a real vector space structure on $E_p$ by defining

$$[p, \alpha, v] + [p, \beta, w] = [p, \alpha, v + \psi_{\alpha\beta}(p)(w)],$$

$$\lambda[p, \alpha, v] = [p, \alpha, \lambda v].$$

This structure does note depend on the specific $\alpha \in A$ chosen, as each $\psi_{\alpha\beta}$ is a vector space isomorphism.

Now define $t_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^m$ by $t_\alpha([p, \beta, v]) = (p, v)$. Then $t_\alpha$ is a bijection onto an open subset of $M \times \mathbb{R}^m$: it is bijective as any $[p, \beta, v] \in \pi^{-1}(U)$ can be expressed uniquely as $[p, \alpha, \psi_{\alpha\beta}(p)(w)]$ by non-singularity of $\psi_{\alpha\beta}(p)$. Now

$$t_\alpha \circ t_\beta^{-1} : (U_\alpha \cap U_\beta) \times \mathbb{R}^m \to (U_\alpha \cap U_\beta) \times \mathbb{R}^m$$

is the map $(p, v) \mapsto (p, \psi_{\alpha\beta}(p)(v))$, which is smooth with smooth inverse $(p, w) \mapsto (p, \psi_{\beta\alpha}(p)(w))$. Thus $t_\alpha \circ t_\beta^{-1}$ is a diffeomorphism.

Now let $\{(V_\gamma, h_\gamma) | \gamma \in G\}$ be any atlas for $M$. Then

$$\left\{ (\pi^{-1}(V_\gamma \cap U_\alpha), (h_\gamma \times \mathrm{id})) \circ t_\alpha \mid \alpha \in A, \gamma \in G \right\}$$

is an atlas for $E$, making $E$ into a $n + m$ dimensional manifold. With this smooth structure, $\{U_\alpha\}$ becomes a trivialising cover for $M$, and the transition functions are clearly the $\{\psi_{\alpha\beta}\}$. $\pi : E \to B$ is smooth, as it is the composition $\mathrm{proj}_1 \circ t_\alpha : [p, \alpha, v] \mapsto (p, v) \mapsto p$. Thus $E$ is a rank $m$ vector bundle over $M$.

Uniqueness up to vector bundle isomorphism is immediate from Lemma 5.6. ▶

## 5.8 An alternative version of the vector bundle construction theorem

It is often more convenient to apply the following version of the vector bundle construction theorem, whose proof is just one stage of the proof of the full vector bundle construction theorem.

Let $M$ be a smooth manifold, and suppose for each $p \in M$ we have an $m$-dimensional real-vector space $E_p$. Let $E = \coprod_{p \in B} E_p$ and $\pi$ the map $E \to M$ such that $\pi^{-1}(p) = E_p$. Suppose $\{U_\alpha\}$ is an open cover of $M$ such that for each $\alpha$, there exists a map

$$t_\alpha : \pi^{-1}(U_i) \to U_i \times \mathbb{R}^m$$

which is a bijection, and such that for each $\alpha, \beta$ such that $U_\alpha \cap U_\beta \neq \emptyset$, the map

$$t_\alpha \circ t_\beta^{-1} : (U_\alpha \cap U_\beta) \times \mathbb{R}^m \to (U_\alpha \cap U_\beta) \times \mathbb{R}^m$$

is a map of the form

$$(p, v) \mapsto (p, \psi_{\alpha\beta}(p)(v)),$$

where $\{\psi_{\alpha\beta}\}$ are a collection of smooth functions $U_\alpha \cap U_\beta \to GL(m, \mathbb{R})$ satisfying the cocycle conditions.

Then there exists a unique topological and smooth structure on $E$ making $E$ into a smooth manifold and $\pi : E \to M$ into a smooth vector bundle of rank $m$ with cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$.

The advantage of using this version rather than the one stated in Theorem 5.7 is that it allows us to explicitly construct the desired vector bundle, rather than just assert the existence of such a vector bundle up to isomorphism. In fact, as we shall see later, we often want to distinguish between isomorphic vector bundles (for instance, the tangent bundle and the cotangent bundle - see Chapter 6), and then Theorem 5.7 is not much use.

## 5.9 Definition

If $\pi : E \to M$ is a vector bundle, a **subbundle** of $E$ is a subset $E' \subseteq E$ such that $E'$ is an embedded submanifold of $E$, and for each $p \in M$, the fibre $E'_p := E' \cap \pi^{-1}(p)$ is a linear subspace of $E_p$, and that with the vector space structure on $E'_p$ inherited from $E_p$, $\pi|_{E'} : E' \to M$ is a vector bundle. In other words, we require the inclusion map $i : E' \to E$ to be a bundle morphism.

## 5.10 Definitions

Let $\pi : E \to M$ be a vector bundle. A smooth map $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$ is called a **section** of $\pi$. The set of sections of $\pi$ is written $\Gamma(E)$, although it often has other notations depending on the bundle, eg. $\Gamma(T(M)) = \mathcal{X}(M)$. Sections need not be defined on all of $M$; if $U$ is an open subset of $M$ we write $\Gamma(U, E)$ for the smooth **local sections** $s : U \to E$ such that $\pi \circ s = \mathrm{id}_U$.

A **local frame** of a vector bundle $\pi : E \to M$ of rank $m$ is a family $\mathbf{e} = \{e_1, \ldots, e_k\}$ of smooth sections in $\Gamma(U, E)$ (where $U \subseteq M$ is open) such that for all $p \in U$, $\{e_1(p), \ldots, e_k(p)\}$ is a basis of $E_p$. A **global frame e** is a frame defined on all of $M$. We say a manifold is **parallelisable** if it admits a global frame.

## 5.11 Lemma

There is a bijective correspondence between local trivialisations of $E$ and local frames of $E$.

◀ If $t : \pi^{-1}(U) \to U \times \mathbb{R}^m$ is a local trivialisation, define a local frame $\mathbf{e} = \{e_1, \ldots, e_m\}$ over $U$ by

$$e_i(p) = t^{-1}(p, e_i),$$

where the $\{e_i\}$ on the right hand side of the above equation denote the standard basis of $\mathbb{R}^m$. It is clear the $\mathbf{e}$ is a local frame.

Conversely if $\mathbf{e} = \{e_1, \ldots, e_m\}$ is a local frame over $U$, define a map $\tau : U \times \mathbb{R}^m \to \pi^{-1}(U)$ by

$$\tau\left(p, a^1, \ldots, a^n\right) = a^i e_i(p).$$

Since $\mathbf{e}$ is a local frame $\tau$ is bijective, and to complete the proof we need only show that $\tau$ is a diffeomorphism (it is clear that these two operations are mutually inverse). Since $\tau$ is bijective is suffices to check $\tau$ is a local diffeomorphism, and hence it is enough to show that if $p \in U$, and $t : \pi^{-1}(V) \to V \times \mathbb{R}^m$ is a local trivialisation over $V$, where $V \subseteq U$ is a neighborhood of $p$ then $t \circ \tau : V \times \mathbb{R}^m \to V \times \mathbb{R}^m$ is a diffeomorphism. For each $i$, the map $t \circ e_i : V \to V \times \mathbb{R}^m$is smooth, and hence there are smooth functions $f_i^j$ such that

$$t \circ e_i(p) = \left(p, f_i^1(p), \ldots, f_i^m(p)\right).$$

Thus

$$t \circ \tau\left(p, a^1, \ldots, a^m\right) = \left(p, a^i f_i^1(p), \ldots, a^i f_i^m(p)\right)$$

is smooth, Finally $(t \circ \tau)^{-1}$ is smooth since matrix inversion is smooth - if $\left[d_i^j(p)\right]$ denotes the matrix inverse to $\left[f_j^i(p)\right]$ then $p \mapsto \left[d_i^j(p)\right]$ is smooth and

$$t \circ \tau\left(p, b^1, \ldots, b^m\right) = \left(p, b^i d_i^1(p), \ldots, b^i d_i^m(p)\right)$$

which is also smooth. ▶

## 5.12 Corollary

A vector bundle $\pi : E \to M$ is trivial if and only if it admits a global frame.

◀ Apply the previous result with $U = M$. ▶

## 5.13 Definitions

Let $\pi : E \to M$ be a rank $m$ vector bundle. If all the transition functions $\{\psi_{\alpha\beta}\}$ of $E$ take their values in some subgroup $G \subseteq GL(m,\mathbb{R})$ then we say that **the structure group of $E$ can be reduced to $G$**. We say that a vector bundle is **orientable** if its structure group may be reduced to $GL^+(m,\mathbb{R}) := \{A \in GL(m,\mathbb{R}) \mid \det(A) > 0\}$. We say $E$ admits an **orthogonal structure** if we can reduce the structure group of $E$ to $O(m)$. Similarly a complex vector bundle of rank $m$ admits a **unitary structure** if we can reduce its structure group to $U(m)$.

## 5.14 The Hopf bundle admits a unitary structure.

By choosing a different trivialisation, we give the Hopf bundle $E$ (Example 2 of Section 5.4) a unitary structure. Define

$$t_0 : (w, wz) \in E_{(1:z)} \mapsto \left((1:z), w\sqrt{1+|z|^2}\right) \in U_0 \times \mathbb{C},$$

$$t_1 : (v\zeta, v) \in E_{(\zeta:1)} \mapsto \left((\zeta:1), v\sqrt{1+|\zeta|^2}\right) \in U_1 \times \mathbb{C}.$$

Writing $x = w\sqrt{1+|z|^2}$ and $y = v\sqrt{1+|\zeta|^2}$, we see that

$$t_1 \circ t_0^{-1}((1:z), x) = t_1\left(\frac{x}{\sqrt{1+|z|^2}}, \frac{xz}{\sqrt{1+|z|^2}}\right)$$

$$= t_1\left(\frac{x|\zeta|}{\sqrt{1+|\zeta|^2}}, \frac{x|\zeta|}{\zeta\sqrt{1+|\zeta|^2}}\right) = \left((\zeta:1), \frac{|\zeta|x}{\zeta}\right),$$

and thus the transition maps $\psi_{10}$ is defined by $\psi_{10}(z)(x) = \frac{zx}{|z|}$, and similarly $\psi_{01}(\zeta)(y) = \frac{\zeta y}{|\zeta|}$.

Thus both $\psi_{10}$ and $\psi_{01}$ are maps from $U_0 \cap U_1 \to U(1) \subseteq GL_1(\mathbb{C})$.

## 5.15 Metrics

A **metric** on a vector bundle $\pi : E \to M$ of rank $m$ is an assignment $p \mapsto \langle\cdot,\cdot\rangle_p$ where $\langle\cdot,\cdot\rangle_p$ is an inner product on $E_p$ varying smoothly with $p$. Slightly more precisely, and using terminology that will become clearer shortly, a metric is a smooth section of the bundle $E^* \otimes E^*$.

If we can reduce the structure group $G$ of $E$ to $O(m)$ then we can obtain a metric on $E$ as follows: define $\langle v, w\rangle_p$ for $v, w \in E_p$ to be $\langle t^p(v), t^p(w)\rangle_{\mathbb{R}^m}$, where $t$ is some local trivialization $t^p := \mathrm{proj}_1 \circ t|_{E_p} : E_p \to \mathbb{R}^m$, and $\langle\cdot,\cdot\rangle_{\mathbb{R}^m}$ denotes the Eucliden dot product on $\mathbb{R}^m$. This is well defined as if $\tau$ is another local trivialization, with transition function $\psi$, so

$$\tau^p(v) = \psi(p)\left(t^p(v)\right),$$

then

$$\langle \tau^p(v), \tau^p(w)\rangle_{\mathbb{R}^m} = \langle \psi(p)\left(t^p(v)\right), \psi(p)\left(t^p(w)\right)\rangle_{\mathbb{R}^m} = \langle t^p(v), t^p(w)\rangle_{\mathbb{R}^m},$$

since $\psi(p)$ is an orthogonal matrix. By construction, this inner product varies smoothly over the fibres. We shall see below in Lemma 5.19 below that every (real) vector bundle of rank $m$ admits an orthogonal structure, and thus we can in fact always define a metric on a vector bundle.

Similarly we can define a **Hermitian metric** on a complex vector bundle $\pi : E \to M$ of rank $m$ is an assignment $p \mapsto \langle\cdot,\cdot\rangle_p$ where $\langle\cdot,\cdot\rangle_p$ is a Hermitian inner product on $E_p$ varying smoothly with $p$. In exactly the same way, if $E$ admits a unitary structure then we can define a Hermitian metric on $E$. We have just shown in Section 5.14 that the Hopf bundle admits such a unitary structure; we define a Hermitian inner product on it by

$$\langle (w_1, w_1 z), (w_2, w_2 z)\rangle_{(1:z)} = \left\langle w_1\sqrt{1+|z|^2}, w_2\sqrt{1+|z|^2}\right\rangle = w_1\bar{w}_2(1+|z|)^2.$$

In particular, the associated norm gives

$$\|(w, wz)\|_p^2 = \langle (w, wz), (w, wz)\rangle_p = |w|^2(1+|z|)^2 = |w|^2 + |wz|^2,$$

which therefore corresponds to the standard length of a vector $(w, wz) \in \mathbb{C}^2$. Note again that to say a bundle can have its structure group reduced to $G$ is to say that in **some** trivialisation we can force all the transition functions to take their values in $G$, and not that in every trivialisation the transition functions take their values in $G$; indeed the original trivialisation of the Hopf bundle in Example 2 of Section 5.4 did not have its transition functions taking their values in $U(1)$.

## 5.16 The Gram-Schmidt process

Conversely, given a metric $\langle \cdot, \cdot \rangle$ on $E$, we may reduce the structure group to $O(m)$ as follows: let $\{(U_\alpha, t_\alpha) | \alpha \in A\}$ denote an open covering of $M$ by trivialising neighborhoods for $E$. Let $\{e_1, \dots e_k\}$ denote the canonical basis of $\mathbb{R}^m$, and for each $\alpha$, let $e_1^\alpha, \dots, e_k^\alpha$ denote the smooth sections of $\pi$ on $U_\alpha$ defined by $e_i^\alpha(p) = t_\alpha^{-1}(p, e_i)$. Then define new sections $s_i^\alpha$ inductively by

$$s_i^\alpha(p) = \frac{e_i^\alpha(p) - \sum_{j=1}^{i-1} \langle e_j^\alpha(p), s_j^\alpha(p) \rangle s_j^\alpha(p)}{\left\| e_i^\alpha(p) - \sum_{j=1}^{i-1} \langle e_j^\alpha(p), s_j^\alpha(p) \rangle s_j^\alpha(p) \right\|};$$

(this is just the Gram-Schmidt orthogonalization process); then the $s_i^\alpha$ then form a smooth frame that are orthonormal with respect to $\langle \ , \ \rangle$.

Next, define new local trivializations $\tau_\alpha$ by

$$\tau_\alpha \left( a^i s_i^\alpha(p) \right) = \left( p, a^1, \dots, a^m \right);$$

note that the $\tau_\alpha$ are local trivializations as $\{s_i^\alpha\}$ is a smooth frame. Claim now that the transition functions $\psi_{\alpha\beta}$ with respect to $\{U_\alpha, \tau_\alpha\}$ take their values in $O(m)$. Indeed, the matrix $\psi_{\alpha\beta}(p)$ is just the change of basis matrix from the orthonormal basis $\{s_i^\beta(p)\}$ of $E_p$ to the orthonormal basis $\{s_i^\alpha(p)\}$, and thus $\psi_{\alpha\beta}(p)$ is an orthogonal matrix.

The next result is a key result used throughout differential geometry, and is the reason second countability was included in the definition of a topological manifold (Section 1.1). We shall use it several times throughout this course - we have already used the existence of bump functions in the proof of Proposition 2.4. Below in Lemma 5.19 we shall use the existence of a partition of unity to show that every real vector bundle admits an orthogonal structure.

## 5.17 Theorem (existence of partition of unity)

Let $M$ be a smooth manifold and $\{U_\alpha \mid \alpha \in A\}$ an open cover of $M$. Then there exists a countable collection $\{\lambda_i \mid i \in \mathbb{N}\}$ of smooth functions $\lambda_i \in C^\infty(M)$ such that:

- for any $i$, there exists $\alpha(i)$ such that $\mathrm{supp}(\lambda_i) := \overline{\{p \in M \mid \lambda_i(p) \neq 0\}}$ is contained in $U_{\alpha(i)}$ and is compact,
- for all $p \in M$, there exists a neighborhood $V$ of $p$ such that only finitely many of the $\lambda_i$ are not identically zero in $V$ (in other words, the collection $\{\lambda_i\}$ is **locally finite**),
- each $\lambda_i$ is non-negative, and for all $p \in M$, we have $\sum_{i=1}^\infty \lambda_i(p) = 1$ (note this sum only has finitely many non-zero terms).

The collection $\{\lambda_i\}$ is called a **partition of unity subordinate to the open cover** $\{U_\alpha\}$.

We will not prove this theorem; its proof is essentially general (point-set) topology and thus falls out of the remit for this course.

## 5.18 Lemma (bump functions)

Let $M$ be a smooth manifold, $p \in M$ and $U \subset V$ any open neighborhoods of $p$ (with $U$ strictly contained in $V$). Then there exists a smooth bump function $\psi : M \to \mathbb{R}$ such that $0 \leq \psi \leq 1$ on $M$, $\psi|_{\bar{U}} \equiv 1$ and $\psi|_{M \setminus V} \equiv 0$.

◄ Let $\{\psi, \psi'\}$ be a partition of unity subordinate to the open cover $\{V, M \setminus \bar{U}\}$ of $M$, so $\mathrm{supp}(\psi) \subseteq V$ and $\mathrm{supp}(\psi') \subseteq M \setminus \bar{U}$. Then $\psi$ is the desired function. ►

## 5.19 Lemma

Any vector bundle admits a metric.

◄ Let $E$ be a vector bundle of rank $m$ over $M^n$. Let $\{U_\alpha, t_\alpha\}$ be a trivialising cover of $M$ for $E$, and let $\{\lambda_i\}$ be a partition of unity subordinate to $\{U_\alpha\}$. The Euclidean dot product on $\mathbb{R}^m$ induces an inner product on $U_\alpha \times \mathbb{R}^m$, and hence, via $t_\alpha^{-1}$, an inner product on $\pi^{-1}(U_\alpha)$. We will write this inner product as $\langle\ ,\ \rangle_\alpha$.

Now define $\langle\cdot,\cdot\rangle$ on $E$ by

$$\langle v, w\rangle = \sum_{i=1}^{\infty} \lambda_i(\pi(v))\, \langle v, w\rangle_{\alpha(i)}\,,$$

where $\mathrm{supp}(\lambda_i) \subseteq U_{\alpha(i)}$, and $\lambda_i(\cdot)\,\langle\cdot,\cdot\rangle_{\alpha(i)}$ is defined to be zero if $\pi(v) \notin U_{\alpha(i)}$. Since $\{\lambda_i\}$ is locally finite, for any $p$ and any, $\langle\cdot,\cdot\rangle_p := \langle\cdot,\cdot\rangle|_{E_p\times E_p}$ is a finite sum of inner products varying smoothly with $p$, and hence $p \mapsto \langle\cdot,\cdot\rangle_p$ is smooth. Finally, since the properties of being symmetric and positive definite are convex, that is, if $A$ and $B$ are symmetric and positive definite then so is $tA + (1-t)B$ for all $t \in [0,1]$, and at least one $\lambda_i$ is strictly positive at each point $p$, it follows $\langle\cdot,\cdot\rangle$ is indeed a metric on $E$. ►

We conclude our discussion of vector bundles by constructing a few more standard bundles.

## 5.20 The Whitney sum

Let $\pi : E \to M$ and $\pi' : E' \to M$ be two vector bundles over $M$ of rank $m$ and $m'$ respectively. We first claim we can find a common trivialising cover $\mathcal{U}$ of $M$ for $E$ and $E'$. Indeed, if $\{U_\alpha|\alpha \in A\}$ and $\{U'_\beta|\beta \in B\}$ are trivialising covers for $E$ and $E'$ respectively, then we simply consider $\mathcal{U} = \{U_\alpha \cap U'_\beta|\alpha \in A, \beta \in B\}$. Now define the **Whitney sum** $E \oplus E'$ to be the bundle with total space

$$E \oplus E' := \{(v, w) \in E \times E'|\pi(v) = \pi'(w)\}$$

and projection map $\tilde{\pi} : E \oplus E' \to M$ defined by $\tilde{\pi}(v, w) = \pi(v) = \pi'(w)$. Note that $(E \oplus E')_p = E_p \oplus E'_p$. We define local trivialisations $\tau$ for $U \in \mathcal{U}$ by $\tau : \tilde{\pi}^{-1}(U) \to U \times \mathbb{R}^m \oplus \mathbb{R}^{m'}$ fibrewise by

$$\tau^p(v, w) = (t^p(v), t'^p(w))\,.$$

If $E$ and $E'$ have transition functions $\{\psi_{\alpha\beta}\}$ and $\{\psi'_{\alpha\beta}\}$ with respect to $\mathcal{U}$, then $E\oplus E'$ has transition functions

$$\psi_{\alpha\beta} \oplus \psi'_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(m + m', \mathbb{R}),$$

which have matrix representation

$$\begin{bmatrix} [\psi_{\alpha\beta}] & 0 \\ 0 & \left[\psi'_{\alpha\beta}\right] \end{bmatrix},$$

where $[\psi_{\alpha\beta}]$ and $\left[\psi'_{\alpha\beta}\right]$ denote the matrices of $\psi_{\alpha\beta}$ and $\psi'_{\alpha\beta}$ respectively.

Thus $E \oplus E'$ is given by the cocycle $\{U_\alpha, \psi_{\alpha\beta} \oplus \psi'_{\alpha\beta}\}$. It is clear that this satisfies the cocycle condition, and thus Theorem 5.8 guarantees that this is indeed a well defined vector bundle of rank $m + m'$.

This illustrates the merits of Theorem 5.7 versus Theorem 5.8. Given two bundles $E, E'$ with cocycles $\{U_\alpha, \psi_{\alpha\beta}\}$ and $\{U_\alpha, \psi'_{\alpha\beta}\}$, it is clear that $\{U_\alpha, \psi_{\alpha\beta} \oplus \psi'_{\alpha\beta}\}$ defines a cocycle and thus gives us a bundle $E \oplus E'$. However we have no way of telling a priori that the fibres of $E \oplus E'$ can be given by $(E \oplus E')_p = E_p \oplus E'_p$ (although in this case it is rather obvious). By applying the alternative version, Theorem 5.8 we were not only able to obtain the existence of the bundle $E \oplus E'$ but also to obtain explicit information about how it is constructed.

## 5.21 The Dual Bundle

Let $\pi : E \to M$ be a vector bundle of rank $m$. Let $E_p^* := \mathrm{Hom}(E_p, \mathbb{R}^m)$ and $E^* = \coprod_{p \in M} E_p^*$. Define $\pi^* : E^* \to B$ mapping $E_p^* \to p$. Let $\{U_\alpha, t_\alpha\}$ be a trivializing cover of $M$ for $E$. Then if $p \in U_\alpha$, we have $t_\alpha^p : E_p \to \mathbb{R}^m$ an isomorphism, and so if we define

$$\tau_\alpha^p := (t_\alpha^p)^*,$$

then

$$\tau_\alpha^p : E_p^* \to (\mathbb{R}^m)^* \cong \mathbb{R}^m,$$

and then $\tau_\alpha : (\pi^*)^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^m$ is a bijective map such that its restriction to each fibre $E_p^*$ is a linear isomorphism. If $U_\alpha \cap U_\beta \neq \emptyset$, then

$$\tau_\beta \circ \tau_\alpha(p, v) = \left(p, \left(\psi_{\beta\alpha}(p)^{-1}\right)^t (v)\right),$$

and thus the associated transition functions $\psi_{\alpha\beta}^*$ are transposed inverses of the original ones,

$$\psi_{\beta\alpha}^*(p) := \left(\psi_{\alpha\beta}(p)\right)^t.$$

The $\{\psi_{\alpha\beta}^*\}$ satisfy the cocycle conditions as the $\{\psi_{\alpha\beta}\}$ do. Thus by Theorem 5.8, this does indeed define a vector bundle of rank $m$ over $M$, which we call the **dual bundle**. Again, we could proceed using Theorem 5.7 and simply define $E^*$ to be the bundle with cocycle

$$\left\{U_\alpha, \left(\psi_{\alpha\beta}^{-1}\right)^t\right\},$$

where $E$ has cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$, but then as before we would not a priori know that a concrete representation of this has fibres equal to the dual space of the original fibres.

## 5.22 The tensor bundle

Now let $\pi : E \to M$ and $\pi : E' \to M$ be two vector bundles of rank $m$ and $m'$ respectively. Let $(E \otimes E')_p := E_p \otimes E'_p$, and $E \otimes E' = \coprod_{p \in M}(E \otimes E')_p$, with $\bar{\pi} : E \otimes E' \to M$ the map such that $\bar{\pi}^{-1}(p) = (E \otimes E')_p$. If $\{U_\alpha\}$ is a common trivialising cover for $E$ and $E'$, with associated trivialisations $t_\alpha$ and $t'_\alpha$, then define

$$\tau_\alpha : (E \otimes E')_p \to \mathbb{R}^m \otimes \mathbb{R}^{m'} \cong \mathbb{R}^{mm'}$$

to be the map defined on decomposable element $v \otimes w$ of $E_p \otimes E'_p$ by

$$v \otimes w \mapsto t_\alpha^p(v) \otimes t_\alpha'^p(w),$$

and then extending by linearity. Then $\tau_\alpha : \bar{\pi}^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^{mm'}$ is a bijective map such that its restriction to each fibre $(E \otimes E')_p$ is a linear isomorphism. If $U_\alpha \cap U_\beta \neq \emptyset$, then

$$\tau_\alpha \circ \tau_\beta^{-1}(p, v \otimes w) = \left(p, \psi_{\alpha\beta}(p)(v) \otimes \psi'_{\alpha\beta}(p)(w))\right),$$

and thus $E \otimes E'$ has cocycle $\{U_\alpha, \psi_{\alpha\beta} \otimes \psi'_{\alpha\beta}\}$. We call $E \otimes E'$ the **tensor product bundle** of $E$ and $E'$.

## 5.23 The Hom bundle

Let $\pi : E \to M$ and $\pi' : E' \to M$ be vector bundles over $M$ of ranks $m$ and $m'$, with cocycles $\{U_\alpha, \psi_{\alpha\beta}\}$ and $\{U_\alpha, \psi'_{\alpha\beta}\}$ respectively. Let $\mathrm{Hom}\left(E_p, E'_p\right)$ denote the set of linear maps $E_p \to E'_p$, so $\mathrm{Hom}\left(E_p, E'_p\right) \cong \mathrm{Hom}\left(\mathbb{R}^m, \mathbb{R}^{m'}\right)$. Define the **homomorphism bundle of $E$ and $E'$** to be the bundle $\mathrm{Hom}\left(E, E'\right) \to M$ with have fibre

$$\mathrm{Hom}\left(E, E'\right)_p = \mathrm{Hom}\left(E_p, E'_p\right).$$

In order to check that this is indeed a vector bundle of rank $mm'$ over $M$, we use the natural isomorphism

$$\mathrm{Hom}(V, W) \cong V^* \otimes W$$

given by $(v^*, w) \mapsto [v \mapsto v^*(v)w]$, to regard this construction as a special case of Sections 5.22 and 5.23; namely, $\mathrm{Hom}(E, E') \cong E^* \otimes E'$.

In the special case $E' = E$ we write $\mathrm{End}(E)$ for $\mathrm{Hom}(E, E)$.

## 5.24 The pullback bundle

Let $\pi : E \to N$ be a vector bundle of rank $m$ over a smooth manifold $N$, and suppose $\Phi : M \to N$ is smooth. We define the **pullback bundle** $\Phi^*E$ over $M$ to be the bundle with total space

$$\Phi^*E := \{(p, v) \in M \times E \mid \Phi(p) = \pi(v)\},$$

and projection $\pi_\Phi : \Phi^*E \to M$ defined by $\pi_\Phi(p, v) = p$. We also define $F : \Phi^*E \to E$ by $F(p, v) = v$, so that $\Phi \circ \pi_\Phi = \pi \circ F$. We define a vector space structure on $(\Phi^*E)_p$ by

$$\lambda(p, v) + \mu(p, w) := (p, \lambda v + \mu w),$$

and so $F$ maps $(\Phi^*E)_p$ onto $E_p$ isomorphically. To see the local triviality of $\Phi^*E$ if $(U, t)$ is local trivialisation of $E$ then we define

$$\tau : \pi_\Phi^{-1}\left(\Phi^{-1}(U)\right) \to \Phi^{-1}(U) \times \mathbb{R}^m$$

by

$$\tau(p, v) = (p, t(F(v))).$$

Thus $\Phi^*E \to M$ is a bundle of rank $m$ over $M$. If $E$ has cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$ then $\Phi^*E$ has cocycle $\{\Phi^{-1}(U_\alpha\}, \psi_{\alpha\beta} \circ \Phi\}$.

## 5.25 Sheaves

In this course we will introduce sheaves very superficially. Our use of them will solely be as an aid in notation and as a way of expressing the definition of a connection (see Chapter 8) more concisely.

A **sheaf** $\mathcal{E}$ over a topological space $T$ is an assignment to each nonempty open set $U \subseteq T$ a group $\mathcal{E}(U)$, called the **sections** of $\mathcal{E}$ over $U$, and to each pair $U \subseteq V$ of open sets a map $r_U^V : \mathcal{E}(V) \to \mathcal{E}(U)$ called the **restriction map** satisfying:

1. For any triple $U \subseteq V \subseteq W$ of open sets

$$r_U^W = r_U^V \circ r_V^W.$$

   Because of this relation for $s \in \mathcal{E}(V)$ we may write $s|_U$ for $r_U^V(s)$ without losing any information.

2. For any pair of open sets $U, V \subseteq T$ and sections $s \in \mathcal{E}(U)$ and $\sigma \in \mathcal{E}(V)$ such that

$$s|_{U\cap V} = \sigma|_{U\cap V}$$

   there exists a section $\varepsilon \in \mathcal{E}(U \cap V)$ such that

$$\varepsilon|_U = s, \quad \varepsilon|_V = \sigma.$$

3. If $s \in \mathcal{E}(U \cap V)$ and

$$s|_U = s|_V = 0$$

   then $s = 0$.

## 5.26 Examples of sheaves

1. Here is the simplest example relevant to us. Let $M$ be a smooth manifold, and define a sheaf $\mathcal{C}^\infty$on $M$ by letting
$$\mathcal{C}^\infty(U) = C^\infty(U).$$
The restriction maps are just the standard restrictions of smooth functions.

2. A slight modification gives the sheaf $\mathcal{C}^*$ on $M$ given by
$$\mathcal{C}^*(U) = \{\text{multiplicative group of nonzero } C^\infty \text{ functions on } U\}.$$

3. The main example we consider is the following. Suppose $E \to M$ is a smooth vector bundle. Define the **sheaf of sections** of $E$ to be
$$\mathcal{E}(U) := \Gamma(U, E),$$
the smooth sections of $E$ over $U \subseteq M$. Again the restriction maps are the obvious ones.

4. A particular example of this is the sheaf $\mathcal{X}$ over $M$ consisting of the smooth sections of $T(M)$, that is, the vector fields.

5. In the next chapter we will meet the sheaf $\Omega^r$ on $M$ of **differential** $r$**-forms**, given by
$$\Omega^r(U) = \Gamma\left(U, \Lambda^r\left(T^*(M)\right)\right)$$
(this notation will make more sense later).

6. In Chapter 8 we will meet the sheaf $\mathcal{A}^r$ of $E$**-valued differential** $r$**-forms** given by
$$\mathcal{A}^r(U) = \Gamma\left(U, \Lambda^r\left(T^*(M)\right) \otimes E\right),$$
as well as variations on this theme:
$$\mathcal{A}^r_{\mathrm{End}(E)}(U) = \Gamma\left(U, \Lambda^r\left(T^*(M)\right) \otimes \mathrm{End}\left(E\right)\right).$$

7. If $\mathcal{E}$ is any sheaf over a topological space $T$, and $U \subseteq T$ is any open subset we can define the **restriction sheaf** $\mathcal{E}|_U$ in the obvious way; namely for $V \subseteq U$ open we set
$$\mathcal{E}|_U(V) = \mathcal{E}(V).$$

## 5.27 Definition

A **sheaf morphism** $\alpha : \mathcal{E} \to \mathcal{F}$ of sheaves over a topological space $T$ is given by a collection of homomorphisms $\alpha_U : \mathcal{E}(U) \to \mathcal{F}(U)$ such that given $U \subseteq V$, the maps $\alpha_U$ and $\alpha_V$ commute with the restriction maps, that is
$$\alpha_V \circ r^V_U = \rho^V_U \circ \alpha_U,$$
where $r^V_U$ is the restriction map of $\mathcal{E}$ and $\rho^V_U$ is the restriction map of $\mathcal{F}$.

The following is essentially the only sheaf-theoretic result we will prove in the entire course. It will be important in Chapter 8, when we come to define the curvature of a connection.

## 5.28 Proposition

Let $\pi : E \to M$ and $\pi' : E' \to M$ be two vector bundles over $M$ with sheaves of sections $\mathcal{E}$ and $\mathcal{E}'$ respectively. Then there is a natural bijective correspondence between vector bundle morphisms $F : E \to E'$ and sheaf morphisms $\alpha : \mathcal{E} \to \mathcal{E}'$ that are **linear over the sheaf** $\mathcal{C}^\infty$on $M$. By this we mean that if $s \in \mathcal{E}(U)$ and $f \in C^\infty(U)$ then $\alpha_U(fs) = f\alpha_U(s)$.

◀ Given a vector bundle morphism $F : E \to E'$ and $U \subseteq M$ open, define $\alpha_U : \mathcal{E}(U) \to \mathcal{E}'(U)$ by

$$\alpha_U(s) = F|_{\pi^{-1}(U)} \circ s.$$

It is clear that this is compatible with the restrictions and so defines a sheaf morphism $\alpha : \mathcal{E} \to \mathcal{E}'$. Moreover it is clear that $\alpha$ is linear over $\mathcal{C}^\infty$.

Conversely suppose that $\alpha : \mathcal{E} \to \mathcal{E}'$ is a sheaf morphism that is linear over $\mathcal{C}^\infty$. Given $p \in M$, select a local frame $\mathbf{e} = \{e_1, \ldots, e_m\}$ ($m = \mathrm{rank}(E)$) over a neighborhood $U$ of $p$. Then given $v \in E_p$ we can uniquely write

$$v = a^i e_i(p).$$

Consider now the $a^i : U \to \mathbb{R}$ as smooth (constant) functions on $U$. Then

$$\alpha_U\left(a^i e_i\right) = a^i \alpha_U\left(e_i\right),$$

where the $\{\alpha_U(e_i)\}$ are smooth sections of $E'$ over $U$, and we define

$$F_p(v) = a^i \alpha_U\left(e_i\right)(p).$$

Then we define $F : E \to E'$ by $F|_{E_p} = F_p$. $F$ is well defined as $\alpha$ is compatible with restrictions, and it is clear that $F$ is a vector bundle morphism. Moreover these operations are obviously mutually inverse. ▶

# 6 Differential forms and cohomology

## 6.1 Differentials

Given a smooth manifold $M^n$, and $U \subseteq M$ open, a function $f \in C^\infty(U)$ gives a map $df(p) : T_p(M) \to T_{f(p)}(\mathbb{R})$ for $p \in U$. Under the identification $T_{f(p)}(\mathbb{R}) \cong \mathbb{R}$, given by $v \mapsto v(r)$ (where $r : \mathbb{R} \to \mathbb{R}$ is the coordinate on $\mathbb{R}$), we map think of $df(p)$ as a map $T_p(M) \to \mathbb{R}$, in other words, $df(p) \in (T_p(M))^* =: T_p^*(U)$. In this case we will normally write $df_p$ instead and call $df_p$ the **differential** of $f$. We call $T_p^*(M)$ the **cotangent space** to $M$ at $p$. If $(U, h)$ is a chart, with coordinates $(x^1, \ldots, x^n)$, then observe

$$df(p)\left(\partial_j|_p\right)(r) = \frac{\partial}{\partial x^j}(r \circ f)(p) = \frac{\partial f}{\partial x_j}(p).$$

In particular, consider $dx^i|_p \in T_p^*(M)$ . Then $dx^i|_p\left(\partial_j|_p\right) = \delta^i_j$. Thus

$$df_p = \frac{\partial f}{\partial x^i}(p) dx^i|_p,$$

and $\{dx^i|_p \mid i = 1, \ldots, d\}$ is the basis of $T_p^*(M)$ dual to the basis $\{\partial_i|_p \mid i = 1, \ldots, d\}$ of $T_p(M)$.

## 6.2 Change of coordinates

Let $\alpha \in T_p^*(M)$ and $(U, h)$ and $(V, k)$ be charts about $p$, with coordinates $\left(x^1, \ldots, x^n\right)$ and $\left(y^1, \ldots, y^n\right)$ respectively. Then we can write

$$\alpha = a_i dx^i|_p = \alpha = b_j dy^j|_p$$

for some $a_i, b_j \in \mathbb{R}$. Then

$$b_j = \alpha\left(\frac{\partial}{\partial y^j}\Big|_p\right) = \alpha\left(\frac{\partial x^i}{\partial y^j}(p)\frac{\partial}{\partial x^i}\Big|_p\right) = a_i \frac{\partial x^i}{\partial y^j}(p). \tag{18}$$

Since

$$\frac{\partial x^i}{\partial y^j}\frac{\partial y^j}{\partial x^k} = \delta_{ik},$$

using equation the transition matrix is the transposed inverse to the one occuring for the tangent bundle.

## 6.3 Definition

Let $M^n$ be a smooth manifold. The **cotangent bundle** of $M$ is the disjoint union of the cotangent spaces;

$$T^*(M) := \coprod_{p \in M} T_p^*(M).$$

We have a natural projection $\pi^* : T^*(M) \to M$ sending $\alpha \in T_p^*(M) \mapsto p$. When referring to an element $\alpha$ of $T^*(M)$, we will often write $\alpha = (p, \alpha)$ to indicate that $\alpha \in T_p^*(M)$. $T^*(M)$ is the dual bundle to $T(M)$ and hence $T^*(M)$ is a $2n$-dimensional manifold; the above section shows if $M$ has atlas $\{U_\alpha, h_\alpha\}$ then $T^*(M)$ has cocycle

$$\left\{ U_\alpha \left( \left( J \left( h_\alpha \circ h_\beta^{-1} \right) \circ h_\beta \right)^{-1} \right)^t \right\},$$

which agrees with the computations in Section 5.21.

## 6.4 Definitions

A **differential** 1**-form** is a smooth section of $\pi^* : T^*(M) \to M$; that is, a map $\omega : M \to T^*(M)$ such that $\omega(p) =: \omega_p \in T_p^*(M)$, and if $(U, h)$ is a chart about $p$ with coordinates $(x^1, \ldots, x^n)$ then on $U$ we can write

$$\omega_q = f_i(q) dx^i|_q,$$

for $q \in U$, where $f_i \in C^\infty(U)$. We let $\Omega^1$ denote the sheaf of differential 1-forms;

$$\Omega^1(U) = \Gamma(U, T^*(M)).$$

Note that the differential $df$ of a smooth function $f \in C^\infty(U)$ as defined in Section 6.1 is a differential 1-form, $df \in \Omega^1(U)$.

## 6.5 Definition

Now let $M^n$ be a smooth manifold and consider the tensor bundle

$$T^{(k,\ell)}(M) := \overbrace{T(M) \otimes \cdots \otimes T(M)}^{k} \otimes \overbrace{T^*(M) \otimes \cdots \otimes T^*(M)}^{\ell}.$$

We let $\mathcal{T}^{(k,\ell)}$ denote the sheaf of smooth sections of $T^{(k,\ell)}(M)$ and call its sections **(mixed) tensors of type** $(k, \ell)$. We define the **tensor algebra** of $M$ to be

$$\mathcal{T}(M) = \bigoplus_{k,\ell \geq 0} \mathcal{T}^{(k,\ell)}(M).$$

If we have a chart $(U, h)$ with local coordinates $(x^1, \ldots, x^n)$ then a tensor $T$ can be written locally on $U$ as

$$T(p) = T^{i_1 \cdots i_r}_{j_1 \cdots j_s}(p) \partial_{i_1}|_p \otimes \cdots \otimes \partial_{i_k}|_p \otimes dx^{j_1}|_p \otimes \cdots \otimes dx^{j_k}|_p,$$

where the functions $T^{i_1, \ldots, i_r}_{j_1, \ldots, j_s} : U \to \mathbb{R}$ are smooth.

A **contraction** $C$ of $\mathcal{T}(M)$ is a map $C^i_j : \mathcal{T}^{(k,\ell)}(M) \to \mathcal{T}^{(k-1,\ell-1)}(M)$ given on decomposable elements $X_1 \otimes \cdots \otimes X_k \otimes \omega^1 \otimes \cdots \otimes \omega^\ell$ of $\mathcal{T}^{(k,\ell)}(M)$ by

$$C^i_j \left( X_1 \otimes \cdots \otimes X_k \otimes \omega^1 \otimes \cdots \otimes \omega^\ell \right) = \omega^i(X_j) \cdot X_1 \otimes \cdots \otimes \widehat{X_j} \otimes \cdots \otimes X_k \otimes \omega^1 \otimes \cdots \otimes \widehat{\omega^i} \otimes \cdots \otimes \omega^\ell$$

and then extended by linearity.

## 6.6 Multilinear algebra

To progress further we need to study some multilinear algebra. Recall that given vector space $V_1, \ldots, V_r$, the tensor product is the universal multilinear object. Thus any multilinear form $\alpha : V_1 \times \cdots \times V_r \to \mathbb{R}$ induces a unique linear map $\beta : V_1 \otimes \cdots \otimes V_r \to \mathbb{R}$, and in this way we can naturally identify the vector space of multilinear forms $V_1 \times \cdots \times V_r \to \mathbb{R}$ with $\mathrm{Hom}(V_1 \otimes \cdots \otimes V_r, \mathbb{R}) = (V_1 \otimes \cdots \otimes V_r)^*$.

A **perfect pairing** between finite dimensional vector spaces $V$ and $W$ is a bilinear map $(\cdot,\cdot) : V \times W \to \mathbb{R}$ such that if $v \neq 0 \in V$ then there exists some $w \in W$ such that $(v, w) \neq 0$, and similarly if $w \neq 0 \in W$ then there exists some $v \in V$ such that $(v, w) \neq 0$. Such a perfect pairing induces isomorphisms

$$V \cong W^*, W \cong V^*$$

given by $v \mapsto (w \mapsto (v, w))$ and $w \mapsto (v \mapsto (v, w))$, and similarly an isomorphism $\varphi : V \to W^*$ gives a perfect pairing $(v, w) := \varphi(v)(w)$, or an isomorphism $\psi : W \to V^*$ gives a perfect pairing $(v, w) := \psi(w)(v)$.

We have a natural perfect pairing

$$(\cdot,\cdot) : (V_1^* \otimes \cdots \otimes V_r^*) \times (V_1 \otimes \cdots \otimes V_r) \to \mathbb{R}$$

given by

$$((v_1^*, \ldots, v_r^*), (v_1, \ldots, v_r)) = v_1^*(v_1) \ldots v_r^*(v_r).$$

This gives us a natural isomorphism $V_1^* \otimes \cdots \otimes V_r^* \cong (V_1 \otimes \cdots \otimes V_r)^*$ and thus allows us to make the identification with the space of multilinear maps $V_1 \times \cdots \times V_r \to \mathbb{R}$ with $V_1^* \otimes \cdots \otimes V_r^*$.

Thus for a fixed vector space $V$, we may identify $\mathrm{Mult}^r(V)$, the vector space of multilinear maps $V^r \to \mathbb{R}$ with $(V^*)^{\otimes r}$.

## 6.7 The exterior algebra

Let $V$ be a fixed vector space. The **$r$th exterior algebra** $\Lambda^r(V)$ is the universal object for **alternating** multilinear map, that is, maps

$$\alpha : V^r \to \mathbb{R}$$

such that for any $v_1, \ldots, v_r \in V$ and any permutation $\pi \in S_r$ , we have $\alpha\left(v_{\pi(1)}, \ldots, v_{\pi(r)}\right) = \mathrm{sgn}(\pi)\alpha(v_1, \ldots, v_r)$. In other words, given any alternating multilinear map $\alpha : V^r \to \mathbb{R}$, $\alpha$ factors uniquely to give a linear map $\beta : \Lambda^r(V) \to \mathbb{R}$ .

To constuct $\Lambda^r(V)$, we let $S^r(V)$ be the ideal of $V^{\otimes r}$ generated by the elements $v \otimes \cdots \otimes v$, and then we let $\Lambda^r(V) := T^r(V)/S^r(V)$. Let $v_1 \wedge \cdots \wedge v_r$ be the image of $v_1 \otimes \cdots \otimes v_r$.

## 6.8 The 'natural' convention

We can identify the space of altenating multilinear maps on $V^r \to \mathbb{R}$, written $\mathrm{Alt}^r(V)$ with $(\Lambda^r(V))^*$. Now we a natural map

$$\varphi : \mathrm{Alt}^r(V) \hookrightarrow \mathrm{Mult}^r(V) \cong (V^*)^{\otimes r} \xrightarrow{q} \Lambda^r(V^*),$$

where $q$ is the quotient map, which is an isomorphism, as its inverse is given by

$$\psi : f_1 \wedge \cdots \wedge f_r \mapsto \frac{1}{r!} \sum_{\pi \in S_r} \mathrm{sgn}(\pi) f_{\pi(1)} \otimes \cdots \otimes f_{\pi(r)}.$$

## 6.9 The 'usual' convention

Unfortunately most books on differential geometry do not use this convention (there are actually compelling reasons not to use this convention, but we will not go into them) and instead define an isomorphism $\psi' : \Lambda^r(V^*) \to \mathrm{Alt}^r(V)$ given by

$$f_1 \wedge \cdots \wedge f_r \mapsto \sum_{\pi \in S_r} \varepsilon(\pi) f_{\pi(1)} \otimes \cdots \otimes f_{\pi(r)}.$$

This makes the formulas prettier, but has the unfortunate side effect that its composition with the natural map $\varphi$ of in the previous section is **not** the identity, and is fact multiplication by $r!$.

From now on, we will identify the alternating tensor $\sum_{\pi \in S_r} \varepsilon(\pi) f_{\pi(1)} \otimes \cdots \otimes f_{\pi(r)}$ with the wedge $f_1 \wedge \cdots \wedge f_r$ without further comment. Under this convention, we have for decomposable elements $f_1 \wedge \cdots \wedge f_r$ of $\Lambda^r(V^*)$,

$$(f_1 \wedge \cdots \wedge f_r)(v_1, \ldots, v_r) = \sum_{\pi \in S_r} f_{\pi(1)}(v_1) \ldots f_{\pi(r)}(v_r) = \det[f_i(v_j)];$$

in fact this association also gives a perfect pairing, which proves that we are in fact defining an isomorphism (under the more natural convention we would have $f_1 \wedge \cdots \wedge f_r)(v_1, \ldots, v_r) = \frac{1}{r!}\det[f_i(v_j)]$).

## 6.10 The wedge product

Under the identification we have $\mathrm{Alt}^r(V)$ with $\Lambda^r(V^*)$, the natural map

$$\Lambda^p(V) \times \Lambda^q(V) \xrightarrow{\wedge} \Lambda^{p+q}(V)$$

induces a **wedge product** on $\wedge : \mathrm{Alt}^p(V) \times \mathrm{Alt}^q(V) \to \mathrm{Alt}^{p+q}(V)$, defined by $(f, g) \mapsto f \wedge g$, with

$$(f \wedge g)(v_1, \ldots, v_{p+q}) = \frac{1}{p!q!} \sum_{\pi \in S_{p+q}} \mathrm{sgn}(\pi) f\left(v_{\pi(1)}, \ldots, v_{\pi(p)}\right) g\left(v_{\pi(p+1)}, \ldots, v_{\pi(p+q)}\right).$$

## 6.11 The algebra of alternating forms

In this way we form the **algebra of alternating forms** on an $n$-dimensional vector space $V$,

$$\mathrm{Alt}(V) = \bigoplus_{r=1}^{n} \mathrm{Alt}^r(V).$$

Note that $\dim(\mathrm{Alt}^r(V)) = \binom{n}{r}$ (where $\dim(V) = n$), since if $\{v_1, \ldots, v_r\}$ is a basis of $V$ then

$$\{v_{i_1} \wedge \cdots \wedge v_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq n\}$$

is a basis of $\Lambda^r(V^*)$.

## 6.12 The exterior bundle

We construct one more bundle. If $\pi : E \to M$ is a vector bundle of rank $m$, we can construct in the same way as the $r$**th exterior bundle** $\Lambda^r(E) \to M$, whose fibres $\Lambda^r(E)_p$ are defined to be $\Lambda^r(E_p)$. This gives a bundle of rank $\binom{m}{r}$. In particular, the line bundle $\Lambda^m(E)$ is called the **determinant line bundle of** $E$ and is written $\det(E)$. Note that if $E$ has cocycle $\{U_\alpha, \psi_{\alpha\beta}\}$ then $\det(E)$ has cocycle $\{U_\alpha, \det \psi_{\alpha\beta}\}$ (hence the name).

## 6.13 Definition

A **differential** $r$**-form** on a smooth manifold $M^n$ is a smooth section of the bundle $\Lambda^r(T^*(M)) \to M$, where $0 \leq r \leq n$, and by convention $\Lambda^0(T^*(M))$ is the trivial bundle $T^*(M) \times \mathbb{R}$. Under the identification $\Lambda^r\left(T_p^*(M)\right)$ with the space of alternating forms on $T_p(M)$, we may identify a differential $r$-form $\omega$ with a map such for each $p \in M$, $\omega(p) =: \omega_p$ is an alternating multilinear map $T_p(M)^r \to \mathbb{R}$. We let $\Omega^r$ denote the sheaf of differential $r$-forms.

If $(U, h)$ is a chart on $M$ with local coordinates $\left(x^1, \ldots, x^n\right)$, we can locally write $\omega_p$ for $p \in U$ as

$$\omega_p = \sum_{1 \leq i_1 < \cdots < i_r \leq d} f_{i_1 \ldots i_r}(p) dx^{i_1}|_p \wedge \cdots \wedge dx^{i_r}|_p,$$

where we are using the fact that

$$\left\{dx^{i_1}|_p \wedge \cdots \wedge dx^{i_r}|_p \mid 1 \leq i_1 < \cdots < i_r \leq d\right\}.$$

is a basis of $\Lambda^r\left(T_p^*(M)\right)$.The assertion that $\omega$ is smooth is equivalent to the functions $f_{i_1 \ldots, i_r} : U \to \mathbb{R}$ being smooth.

## 6.14 Theorem (orientations)

Let $M^n$ be a smooth manifold. The following are equivalent:

1. There exists a nowhere vanishing smooth $n$-form $\omega$ on $M$ (such a form is called a **volume form**).

2. The vector bundle $\det(T^*(M))$ is trivial.

3. There exists an atlas $\mathcal{A} = (U_\alpha, h_\alpha)$ such that the Jacobian matrices of all the transition functions have strictly positive determinant.

◀ The equivalenve of (1) and (2) is clear from Corollary 5.12.

To progress further we first need observe the following: given charts $(U, h)$ and $(V, k)$ with local coordinates $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ respectively, and $U \cap V \neq \emptyset$, the $n$-forms $dx^1 \wedge \cdots \wedge dx^n$ and $dy^1 \wedge \cdots \wedge dy^n$ in $\Omega^n(U \cap V)$ are related by

$$dx^1 \wedge \cdots \wedge dx^n = \det \left[\frac{\partial x^i}{\partial y^j}\right] dy^1 \wedge \cdots \wedge dy^n. \tag{19}$$

This can either we deduced from the cocycle of $\det(T^*(M))$ or directly, as follows: since $\Lambda^n(T^*(M))$ is one dimensional, we know that $dx^1 \wedge \cdots \wedge dx^n = f dy^1 \wedge \cdots \wedge dy^n$ for some smooth function $f$, and to determine $f$ we simply evaluate

$$\begin{aligned} dx^1|_p \wedge \cdots \wedge dx^n|_p \left(\frac{\partial}{\partial y^1}\Big|_p, \ldots, \frac{\partial}{\partial y^n}\Big|_p\right) &= \det\left[dx^i|_p\left(\frac{\partial}{\partial y^j}\Big|_p\right)\right] \\ &= \det\left[dx^i|_p\left(\frac{\partial x^k}{\partial y^k}(p)\frac{\partial}{\partial x^k}\Big|_p\right)\right] \\ &= \det\left[\frac{\partial x^i}{\partial y^j}(p)\right]. \end{aligned}$$

But on the other hand,

$$f(p) dy^1|_p \wedge \cdots \wedge dy^n|_p \left(\frac{\partial}{\partial y^1}\Big|_p, \ldots, \frac{\partial}{\partial y^n}\Big|_p\right) = f(p),$$

and thus $f = \det\left[\frac{\partial x^i}{\partial y^j}\right]$ as claimed.

Suppose $\omega$ is a non-vanishing $n$-form. Let $\mathcal{A}$ denote the collection of all the charts $(U, h)$ such if $(x^1, \ldots, x^n)$ are the associated local coordinates, we have

$$\omega\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right) > 0 \text{ on } U.$$

$\mathcal{A}$ is an atlas, since we can always reorder the local coordinates of a given chart. Moreover if $(U, h)$ and $(V, k)$ are in $\mathcal{A}$ with coordinates $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ respectively, and $U \cap V \neq \emptyset$, then we can write $\omega|_U = g dx^1 \wedge \cdots \wedge dx^n$ and $\omega|_U = h dy^1 \wedge \cdots \wedge dy^n$ for some positive smooth functions $g$ and $h$. Then on $U \cap V$, we have $dx^1 \wedge \cdots \wedge dx^n = f dy_1 \wedge \cdots \wedge dy_d$ where $f = g/h > 0$. But by the previous computation, $f = \det\left[\frac{\partial x^i}{\partial y^j}\right]$, and thus the atlas $\mathcal{A}$ satisifies condition (3). Thus (1) $\Rightarrow$ (3).

Finally we prove that (3) $\Rightarrow$ (1). Given such an atlas $\mathcal{A}$, let $\{\lambda_i\}$ be a partition of unity subordinate to the open cover $\{U_\alpha\}$ of $M$. For each $i$, define the $n$-form $\omega_i$ on $M$ to by

$$\omega_i(p) = \lambda_i(p) dx^1_{\alpha(i)}|_p \wedge \cdots \wedge dx^n_{\alpha(i)}|_p \text{ for } p \in U_{\alpha(i)}, \quad \omega_i(p) = 0 \text{ for } p \neq U_{\alpha(i)}.$$

Then if

$$\omega := \sum_{i=1}^{\infty} \omega_i,$$

we see that $\omega$ is a nowhere vanishing $n$-form, since for any $p \in M$, if $I$ denotes the finite non-empty set of $i \in \mathbb{N}$ such that $\lambda_i(p) \neq 0$, then if $i_0 \in I$ and $\beta = \alpha(i_0)$ we have

$$\begin{aligned}\omega(p)\left(\frac{\partial}{\partial x_\beta^1}\Big|_p, \dots, \frac{\partial}{\partial x_\beta^n}\Big|_p\right) &= \omega_{i_0}(p)\left(\frac{\partial}{\partial x_\beta^1}\Big|_p, \dots, \frac{\partial}{\partial x_\beta^n}\Big|_p\right) \\ &+ \sum_{j \neq i_0 \in I} \omega_j(p)\left(\frac{\partial}{\partial x_\beta^1}\Big|_p, \dots, \frac{\partial}{\partial x_\beta^n}\Big|_p\right),\end{aligned}$$

and the first term is equal to $\lambda_{i_0}(p) > 0$, and all the others are non-negative by equation (19) above and assumption on the atlas $\mathcal{A}$. The theorem is proved. ▶

## 6.15 Definitions

A smooth $n$-manifold is called **orientable** if it satisfies any the three equivalent conditions of Theorem 6.14. If $M$ is orientable, there exist precisely two **orientations**, by which we mean choices of equivalence classes of nowhere vanishing $n$-forms, under the relation $\omega \sim \omega'$ if ${}^{\omega(p)}/_{\omega'(p)} > 0$ for some (and hence every) $p \in M$. By an **orientated manifold** we mean an orientable manifold equipped with a choice of orientation.

## 6.16 Theorem (exterior differentiation)

Let $M$ be a smooth manifold. There exists a unique linear sheaf morphism $d : \Omega^r \to \Omega^{r+1}$ such that

1. If $f \in \Omega^0(M)$ then $df$ is the differential $df$.
2. If $\omega \in \Omega^r(M)$ and $\eta$ is any smooth form we have $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^r \omega \wedge d\eta$.
3. $d^2 = 0$, that is, $d(d\omega) = 0$ for all forms $\omega$.

$d$ is called the **exterior differentiation** operator.

We will first define $d$ locally in terms of charts, and then show that the definition is independent of the choice of chart. Given $p \in M$ and a chart $(U, h)$ around $p$, any $r$ form $\omega$ defined on a neighborhood of $p$ may be locally written as

$$\omega(p) = \sum_{1 \leq i_1 < \dots < i_r \leq d} f_{i_1 \dots i_r}(p) dx^{i_1}|_p \wedge \dots \wedge dx^{i_r}|_p.$$

We shall use the shorthand

$$\omega(p) = \sum_I f_I(p) dx^I|_p,$$

where $I = (i_1, \dots, i_r)$ is a strictly increasing multiindex, and $dx^I|_p := dx^{i_1}|_p \wedge \dots \wedge dx^{i_r}|_p$. Define

$$d\omega(p) := \sum_I (df_I)_p \, dx^I|_p.$$

We will show that $d$ enjoys the following properties.

- $d\omega(p) \in \Lambda^{r+1}\left(T_p^*(M)\right)$ for any local $r$-form $\omega$.
- If two $r$-forms $\omega, \eta$ agree on a neighborhood of $p$ then $d\omega(p) = d\eta(p)$.
- $d(a\omega + b\eta)(p) = ad\omega(p) + bd\eta(p)$ for $a, b \in \mathbb{R}$.
- $d(\omega \wedge \eta)(p) = (d\omega \wedge \eta)(p) + (-1)^r(\omega \wedge d\eta)(p)$ for any form $\eta$.
- $d(df)(p) = 0$ for $f$ a smooth function on a neighborhood of $p$.

The first three are immediate. To check the fourth, by linearity we reduce to the case $\omega = f dx^I$ and $\eta = g dx^J$. The left-hand side is

$$\begin{aligned} d(\omega \wedge \eta)(p) &= d(fg dx^I \wedge dx^J)(p) \\ &= (df_p \cdot g(p) + f(p) \cdot dg_p) \wedge dx^I|_p \wedge (dx^J|_p \\ &= \left(df_p \wedge dx^I|_p\right) \wedge \left(g(p) \cdot dx^J|_p\right) + (-1)^r \left(f(p) \cdot dx^I|_p\right) \wedge \left(dg_p \wedge dx^J|_p\right) \\ &= d\omega(p) \wedge \eta(p) + (-1)^p \omega(p) \wedge d\eta(p), \end{aligned}$$

which is the right-hand side. Now we check that for a local smooth function $f$, $d(df)(p) = 0$. Write $df_p$ locally as $\frac{\partial f}{\partial x^i}(p) dx^i|_p$ - note that by definition $d(f)(p) = df_p$. Then

$$\begin{aligned} d(df)(p) &= d\left(\frac{\partial f}{\partial x^i}(p) dx^i|_p\right)(p) \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j}(p) \cdot dx^j|_p \wedge dx^i|_p \\ & \sum_{i<j} \left(\frac{\partial^2 f}{\partial x^i \partial x^j}(p) - \frac{\partial^2 f}{\partial x^j \partial x^i}(p)\right) dx^i|_p \wedge dx^j|_p, \end{aligned}$$

which is zero by equality of mixed partial deriviatives.

Now we will show that $d$ is well defined, that is, independent of the choice of chart. Suppose $d'$ is defined in the same way relative to some other chart around $p$. Then $d' f(p) = df_p = df(p)$. Furthermore, the five properties we have just shown that $d$ has also apply to $d'$ and hence,

$$\begin{aligned} d'\omega(p) &= \sum_I d'(f_I dx^I)(p) \\ &= \sum_I \left((df_I)_p \wedge dx^I|_p + f_I(p) d'\left(dx^I|_p\right)\right), \end{aligned}$$

and thus it is enough to show that $d'\left(dx^I|_p\right) = 0$. But this follows immediately by applying the third property to $d'\left(dx^{i_1}|_p \wedge \cdots \wedge dx^{i_r}|_p\right)$, together with the fact that $d'\left(dx^i|_p\right) = 0$, since $d'\left(x^i\right) = dx^i = d\left(x^i\right)$ (thinking of $x^i$ as a local smooth function), and thus $d'\left(dx^i|_p\right) = d'(d'\left(x^i\right))(p) = 0$.

Thus $d = d'$ and so $d$ is well defined. Hence we have a well defined operator $d : \Omega^r(M) \to \Omega^{r+1}(M)$ that satisfies all the required conditions which is trivially a sheaf morphism. It remains to show that $d$ is unique. Suppose $d''$ is any sheaf morphism satisfying the conditions of the theorem. By the previous reasoning, it is enough to show that $d''$ satisfies the five properties above. This time, all but the second are immediate. To show this, it is enough to show that if $\omega$ is a form that is zero on a neighborhood of $V$ of $p$ then $d''\omega(p) = 0$. To see this, let $U \subseteq V$ be a smaller neighborhood of $p$ and let $\psi$be a smooth function that is identically 0 on $U$ and identically 1 on $M \backslash V$ (i.e. $\psi = 1 - \psi'$ for some bump function $\psi'$). Then $\psi\omega = \omega$ on all of $M$, so

$$d''\omega(p) = d''(\psi\omega)(p) = d\psi_p \wedge \omega(p) + \psi(p) d''\omega(p) = 0.$$

This establishes uniqueness, and completes the proof. ►

## 6.17 Lemma

Let $\omega$be any 1-form and $X, Y$ any two vector fields. Then

$$d\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]). \tag{20}$$

◄ We may reduce to the case $\omega = f dg$, as the given equation is clearly additive in $\omega$, and any 1-form can be expressed locally as a sum of terms of this form. Then $d\omega = df \wedge dg$ and we have

$$d\omega(X, Y) = df(X) dg(Y) - dg(X) df(Y) = Xf \cdot Yg - Xg \cdot Yf.$$

The right hand side is

$$\begin{aligned} X(f dg(Y)) - Y(f dg(X)) - f dg([X, Y]) &= X(f \cdot Yg) - Y(f \cdot Xg) - f[X, Y]g \\ &= Xf \cdot Yg + fXYg - Yf \cdot Xg - fYXg - f(XYg - YXg), \end{aligned}$$

and everything then cancels to give the desired equality. ►

## 6.18 Definitions

A form $\omega$ is **closed** if it is in the kernel of $d$. It is **exact** if it is in the image of $d$. Since $d^2 = 0$, it follows that $\Omega^*(M)$ forms a **cochain complex** under $d$; we define **de Rham** cohomology $H^*_{\mathrm{dR}}(M)$ to be the associated cohomology of this chain complex. Note that unlike most other cohomology theories, the groups $H^*_{\mathrm{dR}}(M)$ are in fact vector spaces.

## 6.19 Definition

Let $\Phi : M \to N$ be a smooth map between smooth manifolds. Let $\Phi^*$ denote the map dual to $d\Phi$, that is, $\Phi^* : T^*(N) \to T^*(M)$ maps the fibre $T^*_{\Phi(p)}(N)$ to the fibre $T^*_p(M)$; if $\alpha \in T^*_{\Phi(p)}N$,

$$\Phi^*(\alpha)(v) := \alpha(d\Phi(v)).$$

More generally, if $f$ is a local smooth function on $N$, define $\Phi^* f$ to be the local smooth function $f \circ \Phi$ on $M$, and if $\omega$ is a local $r$-form on $N$, define $\Phi^*\omega$ to be the map defined by

$$(\Phi^*\omega)_p (v_1, \dots, v_r) := \omega_{\Phi(p)} (d\Phi(v_1), \dots, d\Phi(v_r)).$$

We shall see below that $\Phi^*\omega$ is smooth and hence a local $r$-form on $M$. The map $\Phi^*$ is called the **pullback map** of $\Phi$. If $\Phi$ is a diffeomorphism, $\omega$ a local 1-form on $N$ and $X$ a local vector field on $M$ then the pushforward $\Phi_*$ and the pullback $\Phi^*$are related by

$$(\Phi^*\omega)(X) = \omega(\Phi_* X) \circ \Phi. \tag{21}$$

## 6.20 Lemma (properties of $\Phi^*$)

Let $\Phi : M \to N$ be smooth. Then:

1. If $\omega$ is a local $r$-form on $N$ then $\Phi^*\omega$ is a smooth and hence a local $r$-form on $M$.
2. $\Phi^* : \Omega^r(N) \to \Omega^r(N)$ is an algebra homomorphism.
3. $\Phi^*$ is a **cochain map** between chain complexes, that is, $d\Phi^* = \Phi^* d$.
4. $\Phi^*$ induces a linear map $H^*_{\mathrm{dR}}(N) \to H^*_{\mathrm{dR}}(M)$.

◄ We will not prove this in the order listed. First we will prove that if $\omega_1, \dots, \omega_r$ are 1-forms then

$$\Phi^*(\omega_1 \wedge \cdots \wedge \omega_r) = \Phi^*\omega_1 \wedge \cdots \wedge \Phi^*\omega_r.$$

Indeed, if $v_1, \dots, v_r \in T_p(M)$,

$$\begin{aligned}
\Phi^*(\omega_1 \wedge \cdots \wedge \omega_r)_p (v_1, \dots, v_r) &= (\omega_1 \wedge \cdots \wedge \omega_r)_{\Phi(p)} (d\Phi(v_1), \dots, d\Phi(v_r)) \\
&= \det\left[(\omega_j)_{\Phi(p)} (d\Phi(v_i))\right] \\
&= \det\left[(\Phi^*\omega_j)_p (v_i)\right] \\
&= (\Phi^*\omega_1 \wedge \cdots \wedge \Phi^*\omega_r)_p (v_1, \dots, v_r).
\end{aligned}$$

Now we check (3) on functions; if $f$ is a local smooth function then we show $\Phi^*(df) = d(\Phi^* f)$. Let $v \in T_p(M)$. Then

$$\begin{aligned}
\Phi^*(df)(v) &= df_p(d\Phi(v)) \\
&= d\Phi(v)(f) \\
&= v(f \circ \Phi) \\
&= v(\Phi^* f) \\
&= d(\Phi^* f)_p (v).
\end{aligned}$$

Now we prove (1) and (2). Let $(U, h)$ be a chart on $N$, with local coordinates $(x^1, \ldots, x^n)$. Since $\Phi^*$ is clearly additively linear, it is enough to show that $\Phi^*\left(f dx^{i_1} \wedge \cdots \wedge dx^{i_r}\right)$ is smooth to prove (1), and note as soon as we have done this we have also proved (2), from the first calculation.

Indeed, from what we have already shown,

$$\begin{aligned}\Phi^*\left(f dx^{i_1} \wedge \cdots \wedge dx^{i_r}\right) &= \left(\Phi^* f\right) \Phi^*\left(dx^{i_1}\right) \wedge \cdots \wedge \Phi^*(dx^{i_r}) \\ &= (f \circ \Phi) d\left(x^{i_1} \circ \Phi\right) \wedge \cdots \wedge d\left(x^{i_r} \circ \Phi\right),\end{aligned}$$

and the latter is smooth, since $x^i \circ \Phi$ and $f \circ \Phi$ are smooth local functions on $M$. This proves (1) and (2), and since

$$\begin{aligned}d\left\{\Phi^*\left(f dx^{i_1} \wedge \cdots \wedge dx^{i_r}\right)\right\} &= d\left\{(f \circ \Phi) d\left(x^{i_1} \circ \Phi\right) \wedge \cdots \wedge d\left(x^{i_r} \circ \Phi\right)\right\} \\ & d(f \circ \Phi) \wedge d\left(x^{i_1} \circ \Phi\right) \wedge \cdots \wedge d\left(x^{i_r} \circ \Phi\right) \\ & \Phi^*(df) \wedge d\left(x^{i_1} \circ \Phi\right) \wedge \cdots \wedge d\left(x^{i_r} \circ \Phi\right) \\ & \Phi^*\left\{d\left(f dx^{i_1} \wedge \cdots \wedge dx^{i_r}\right)\right\},\end{aligned}$$

by linearity we have also proved (3). Finally, (4) is an immediate consequence of (2) and (3). ▶

## 6.21 Corollary

The assignment $\Phi \mapsto \Phi^*$ defines a contravariant functor from the category of smooth manifolds to the category of $\mathbb{Z}$-graded abelian groups.

## 6.22 Extending $\Phi^*$

In the case when $\Phi : M \to M$ is a diffeomorphism it is convenient to extend the definition of $\Phi^*$ to an operator on the tensor algebra $\mathcal{T}(M)$ as follows.

If $X \in \mathcal{X}(M)$ define $\Phi^* X := \left(\Phi^{-1}\right)_* X$. Then extend $\Phi^* : \mathcal{T}^{(k,\ell)}(M) \to \mathcal{T}^{(k,\ell)}(M)$ by setting

$$\Phi^*\left(X_1 \otimes \cdots \otimes X_k \otimes \omega^1 \otimes \cdots \otimes \omega^\ell\right) = \Phi^* X_1 \otimes \cdots \otimes \Phi^* X_k \otimes \Phi^* \omega^1 \otimes \cdots \otimes \Phi^* \omega^\ell.$$

## 6.23 Lemma

$H^0_{\mathrm{dR}}(M) = \mathbb{R}^k$, where $k$ is the number of components (equivalently, path components) of $M$.

◀ A closed 0-from is a smooth real-valued function such that $df = 0$, and this happens if and only if $f$ is constant on each component of $M$. Since there are no $(-1)$-forms, this proves the result. ▶

We conclude our brief discussion of de Rham cohomology by stating the following two important theorems.

## 6.24 Theorem (homotopy invariance)

$H^*_{\mathrm{dR}}(M)$ is a homotopy invariant of the smooth manifold $M$. In particular, it is independent of the smooth structure on $M$.

## 6.25 Theorem (Poincaré lemma)

1. Let $U$ be a star-shaped open subset of $\mathbb{R}^n$. Then $H^n_{\mathrm{dR}}(U) = 0$ for all $r \geq 1$.
2. Let $M$ be a smooth manifold. Then $H^*(M \times \mathbb{R}) \cong H^*(M)$.

# 7 Integration on manifolds and Lie derivatives

## 7.1 Integration on manifolds - the simple case

Let $M^n$ be a smooth orientated manifold, and $\omega$ a compactly supported $n$-form on $M$. Suppose there exists a positively orientated chart $(U, h)$ on $M$ such that $\text{supp}(\omega) \subseteq U$ and $h(U) \subseteq \mathbb{R}^n$ is bounded and measurable. Write $\omega = f dx^1 \wedge \cdots \wedge dx^n$ on $U$, where $(x^1, \ldots, x^n)$ are the local coordinates of $h$ and $f \in C^\infty(U)$ is smooth. Define the **integral of** $\omega$ **over** $M$ to be

$$\int_M \omega = \int_U \omega := \int_{h(U)} \left(h^{-1}\right)^* \omega = \int_{h(u)} f \circ h^{-1} dr^1 \ldots dr^n.$$

## 7.2 Lemma

The integral $\int_M \omega$ is well defined; it does not depend on the choice of chart $(U, h)$.

◀ Suppose $(V, k)$ is another positively orientated chart such that $\text{supp}(\omega) \subseteq V$, with $k(V)$ bounded and measurable, and let $(y^1, \ldots, y^n)$ be the local coordinates associated to $k$. Let $F = h \circ k^{-1} : k(V) \to h(U)$ be the coordinate transformation. Then by (3)

$$\frac{\partial x^i}{\partial y^j}(p) = JF(k(p))^i_j.$$

Thus by equation (19) we have

$$dx^1 \wedge \cdots \wedge dx^n = \det\left(JF \circ k\right) dy^1 \wedge \cdots \wedge dy^n.$$

Thus with respect to coordinates $(y^1, \ldots, y^n)$,

$$\omega_p = f(p) \det\left(JF \circ k\right) dy^1 \wedge \cdots \wedge dy^n.$$

Set $W = k(V)$ . Then by the change of variable formula for multiple integrals, and using the fact that $\det(JF \circ k) > 0$ as both $(U, h)$ and $(V, k)$ are positively orientated (this is why we required the manifold to be orientable) we have

$$\begin{aligned}
\int_{h(U)} \left(h^{-1}\right)^* \omega \quad = \quad & \int_{F(W)} f \circ h^{-1} dr^1 \ldots dr^n \\
& \int_W f \circ h^{-1} \circ F |\det\left(JF \circ k\right)| ds^1 \ldots ds^n \\
& \int_{F(W)} f \circ k^{-1} \det\left(JF \circ k\right) ds^1 \ldots ds^n \\
& \int_{k(V)} \left(k^{-1}\right)^* \omega.
\end{aligned}$$

This completes the proof of the lemma. ▶

## 7.3 Integration on manifolds - the general case

Let $M^n$ be a smooth orientated manifold, and $\omega$ a compactly supported $n$-form on $M$. Let $\mathcal{A}$ be a positively orientated atlas on $M$, such that for each chart $(U, \phi h \in \mathcal{A}$, $h(U)$ is a bounded measurable subset of $\mathbb{R}^n$. There exist finitely many charts $(U_1, h_1), \ldots, (U_r, h_r) \in \mathcal{A}$ such that $\text{supp}(\omega) \subseteq \bigcup_{i=1}^r U_i$.

Set

$$A_{i_1 \ldots i_\ell} := \int_{U_{i_1} \cap \cdots \cap U_{i_\ell}} \omega,$$

where we may integrate with any of the charts $(U_{i_m}, h_{i_m})$ $(m \leq \ell)$ to determine $A_{i_1 \ldots i_\ell}$ by the previous lemma.

Then set

$$\int_M \omega := \sum_{i=1}^r A_i - \sum_{i<j} A_{ij} + \sum_{i<j<k} A_{ijk} - \cdots + (-1)^{r+1} A_{12\ldots r} = \sum_{\ell=1}^r (-1)^{\ell+1} \sum_{i_1<\cdots<i_\ell} A_{i_1\ldots i_\ell}$$

(i.e. we are using the **inclusion-exclusion prinicple**). To justify this, we need.

## 7.4 Lemma

This is well defined: if $\{(V_j, k_j) \mid j = 1, \ldots, s\}$ is another choice of charts from $\mathcal{A}$ such that $\mathrm{supp}(\omega) \subseteq \bigcup_{j=1}^s V_j$ and

$$B_{j_1\ldots j_k} := \int_{V_{j_1}\cap\cdots\cap V_{j_k}} \omega,$$

then

$$\sum_{\ell=1}^r (-1)^{\ell+1} \sum_{i_1<\cdots<i_\ell} A_{i_1\ldots i_\ell} = \sum_{k=1}^s (-1)^{k+1} \sum_{j_1<\cdots<j_k} B_{j_1\ldots j_k}. \tag{22}$$

We could consider the cover $\{(U_i, h_i), (V_j, k_j) \mid i = 1, \ldots, r, j = 1, \ldots, s\}$ and set

$$C_{i_1\ldots i_\ell}^{j_1\ldots j_k} := \int_{U_{i_1}\cap\cdots\cap U_{i_\ell}\cap V_{j_1}\cap\cdots\cap V_{j_k}} \omega,$$

and by the previous lemma it does not matter on which of the maps $h_{i_p}$ or $k_{j_q}$ we use to compute $C_{i_1\ldots i_\ell}^{j_1\ldots j_k}$.

It follows from the definition that for any $m \le r$

$$A_{i_1\ldots i_m} = \sum_{k=1}^s (-1)^{k+1} \sum_{j_1<\cdots<j_k} C_{i_1\ldots i_m}^{j_1\ldots j_k},$$

and similarly for

$$B_{j_1\ldots j_m} = \sum_{\ell=1}^r (-1)^{\ell+1} \sum_{i_1<\cdots<i_\ell} C_{i_1\ldots i_\ell}^{j_1\ldots j_m}.$$

Thus by rearranging the order of summation, both sides of (22) are equal to

$$= \sum_{i=1}^r (-1)^{\ell+1} \sum_{k=1}^s (-1)^{k+1} \sum_{i_1<\cdots<i_\ell} \sum_{j_1<\cdots<j_k} C_{i_1\ldots i_\ell}^{j_1\ldots j_k}.$$

This completes the proof. ▶

## 7.5 Theorem (Stoke's theorem - without boundary)

Let $M^n$ be a smooth orientated manifold and $\omega \in \Omega^{n-1}(M)$. Then

$$\int_M d\omega = 0.$$

We will not prove Stoke's theorem in this course.

## 7.6 Corollary (integration by parts)

Let $M^n$ be a smooth orientated manifold and $\alpha \in \Omega^p(M), \beta \in \Omega^q(M)$ where $p + q = n - 1$. Then

$$\int_M \alpha \wedge d\beta = (-1)^{p+1} \int_M d\alpha \wedge \beta.$$

◀ Apply Stoke's theorem to $\omega = \alpha \wedge \beta$. ▶

## 7.7 Corollary

If $M^n$ is a smooth compact orientable manifold then $H^n_{\text{dR}}(M) \neq 0$.

◀ Choose a volume form $\omega$. Then $\int_M \omega$ is a positive real number. $\omega$ is clearly closed, but not exact by Stoke's theorem, as if $\omega = d\eta$ then $\int_M \omega = \int_M d\eta = 0$. Thus $H^n_{\text{dR}}(M) \neq 0$, as it contains the non-zero class $[\omega]$. ▶

## 7.8 Lie Derivatives

Let $M$ be a smooth manifold and $X \in \mathcal{X}(M)$ and $\phi_t$ the local flow of $X$. Define the **Lie derivative** of $X$, written $L_X$ to be the operator on tensors defined by

$$L_X T(p) = \lim_{t \to 0} \frac{(\phi_t^* T)(p) - T(p)}{t} = \frac{d}{dt}\Big|_{t=0} (\phi_t^* T)(p).$$

We will investigate what this definition means on successively more complicated objects, starting with functions. Recall (see Section 2.12) the notation $\phi_t(p) = c_p(t)$ with $c_p$ the unique maximal integral curve of $X$ through $p$.

We will need the following result to prove the key result on the Lie derivative, Theorem 7.10.

## 7.9 Lemma

Let $\Phi : M \to M$ be a diffeomorphism. Then there exists a unique operator $\alpha : \mathcal{T}(M) \to \mathcal{T}(M)$ such that

1. $\alpha$ preserves the type of tensors.
2. For $f \in C^\infty(M)$, $\alpha(f) = f \circ \Phi$.
3. For $X \in \mathcal{X}(M)$, $\Phi^* X = \left(\Phi^{-1}\right)_* X$.

◀ Suppose that $\alpha$ satisfies the conditions of the Lemma. We will show that if $\omega \in \Omega^*(M)$ then $\alpha(\omega) = \Phi^* \omega$, which thus uniquely determines $\alpha$. For this it is enough to check this for $\omega \in \Omega^1(M)$. Let $C$ be the contraction $X \otimes \omega \mapsto \omega(X)$. Then $\alpha \circ C = C \circ \alpha$ implies that

$$\begin{aligned}
\omega(X) \circ \Phi &= \alpha(\omega(X)) \\
&= \alpha(C\{X \otimes \omega\}) \\
&= C\{\alpha(X \otimes \omega)\} \\
&= C\left\{\left(\Phi^{-1}\right)_* X \otimes \alpha(\omega)\right\} \\
&= \alpha(\omega)\left(\left(\Phi^{-1}\right)_* X\right),
\end{aligned}$$

and hence $\alpha(\omega)$ is the 1-form such that

$$\alpha(\omega)\left(\left(\Phi^{-1}\right)_* X\right) = \omega(X) \circ \Phi,$$

and then comparing to (21) completes the proof. ▶

## 7.10 Theorem (properties of the Lie derivative)

Let $M^n$ be a smooth manifold and $X \in \mathcal{X}(M)$. Then the Lie derivative is the unique operator $\mathcal{T}(M) \to \mathcal{T}(M)$ which maps $\mathcal{X}(M)$ to itself and satisfies:

1. $L_X f = Xf$ for $f \in C^\infty(M)$.
2. $L_X Y = [X, Y]$ for $Y \in \mathcal{X}(M)$.
3. $L_X : \mathcal{T}(M) \to \mathcal{T}(M)$ is a derivation of $\mathcal{T}(M)$ which preserves the type of tensors.

4. $L_X$ commutes with all contractions.

◄ First we verify that $L_X$ as defined satisfies properties (1) to (4). Let $f \in C^\infty(M)$. Then

$$\begin{aligned} Xf(p) &= X_p(f) \\ &= c_p'(0)(f) \\ &= dc_p\left(\frac{d}{dr}\Big|_0\right)(f) \\ &= \frac{d}{dr}(f \circ c_p)(0) \\ &= \frac{d}{dt}\Big|_{t=0}(\phi_t^* f)(p) \\ &= L_X f(p). \end{aligned}$$

Now take $Y \in \mathcal{X}(M)$.

$$L_X Y(p) = \lim_{t\to 0} \frac{(\phi_t^* Y)(p) - Y(p)}{t} = \lim_{t\to 0} \frac{(\phi_{-t})_*(Y)(\phi_t(p)) - Y(p)}{t}.$$

Take a smooth function $f$ defined on a neighborhood of $p$. Then for $t$ small,

$$(\phi_{-t})_*(Y)(\phi_t(p))(f) - Y(p)(f) == (Y(\phi_t(p))(f \circ \phi_{-t}) - Y(\varphi_t(p))(f)) + (Y(\phi_t(p))(f) - Y(p)(f)),$$

and since

$$\begin{aligned} \lim_{t\to 0} \frac{Y(\phi_t(p))(f\circ\phi_{-t}) - Y(\phi_t(p))(f)}{t} &= \lim_{t\to 0} Y(\phi_t(p))\left(\frac{f\circ\phi_{-t} - f}{t}\right) \\ &= Y(\phi_0(p))(-L_X f) \\ &= -Y_p(Xf) \end{aligned}$$

and

$$\lim_{t\to 0} \frac{Y(\phi_t(p))(f) - Y(p)(f)}{t} = L_X(Yf)(p) = X_p(Yf),$$

we have

$$(L_X Y)(p)(f) = X_p(Yf) - Y_p(Xf) = [X,Y]_p(f).$$

Now leet $T$ and $S$ be two smooth tensors on $M$. Then

$$L_X(T\otimes S) = L_X T \otimes S + T \otimes L_X S.$$

Indeed,

$$\begin{aligned} L_X(T\otimes S)(p) &= \lim_{t\to 0} \frac{(\phi_t^*(T\otimes S))(\phi_t(p)) - (T\otimes S)(p)}{t} \\ &= \lim_{t\to 0} \frac{(\phi_t^* T)(\phi_t(p)) \otimes (\phi_t^* S)(\phi_t(p)) - (T\otimes S)(p)}{t} \\ &= \lim_{t\to 0} \frac{(\phi_t^* T)(\phi_t(p)) \otimes (\phi_t^* S)(\phi_t(p)) - (\phi_t^* T)(\phi_t(p)) \otimes S(p)}{t} \\ &\quad + \lim_{t\to 0} \frac{(\phi_t^* T)(\phi_t(p)) \otimes S(p) - T(p)\otimes S(p)}{t} \\ &= \lim_{t\to 0}\left((\phi_t^* T)(\phi_t(p)) \otimes \frac{(\phi_t^* S)(\phi_t(p)) - S(p)}{t}\right) + \lim_{t\to 0}\left(\frac{(\phi_t^* T)(\phi_t(p)) - T(p)}{t} \otimes S(p)\right) \\ &= T(p)\otimes L_X S(p) + L_X S(p) \otimes T(p). \end{aligned}$$

Thus $L_X$ satisfies property (3) and Lemma 7.9 shows that $L_X$ satisfies property (4).

Finally we show uniqueness. Suppose such an operator $\alpha_X : \mathcal{T}(M) \to \mathcal{T}(M)$ exists. Let $C : \mathcal{T}^{(1,1)}(M) \to C^\infty(M)$ be the contraction operator. Then

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

Thus we define $\alpha_X$ on $\Omega^1(M)$ by this relation and observe that

$$(\alpha_X \omega)(fY) = f(\alpha_X \omega)(Y),$$

and so $\alpha_X\left(\Omega^1(M)\right) \subseteq \Omega^1(M)$. If $U$ is a coordinate neighborhood with local coordinates $(x^1, \ldots, x^n)$ then $\alpha_X$ induces endomorphisms of $C^\infty(U)$, $\mathcal{T}^{(0,1)}(M)$ and $\mathcal{X}(U) = \mathcal{T}^{(1,0)}(U)$. Any $T \in \mathcal{T}^{(k,\ell)}(U)$ can be written as

$$T = T^{i_1 \ldots i_k}_{j_1 \ldots j_\ell} \partial_{i_1} \otimes \cdots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_\ell},$$

and we can uniquely extend $\alpha_X|_U$ to $\mathcal{T}(U)$ satisfying properties (1), (2) and (4). Property (3) is verified by induction on $k$ and $\ell$. Finally, $\alpha_X$ is defined on $\mathcal{T}(M)$ by the requirement

$$(\alpha_X T)|_U = (\alpha_X|_U)(T|_U),$$

which is enforced by the requirement that $\alpha_X$ should be a derivation, as in the proof of Theorem 6.16. ▶

## 7.11 Lemma (properties of the Lie derivative of a smooth form)

Let $M$ be a smooth manifold, $\omega, \eta$ smooth forms on $M$ and $f \in C^\infty(M)$. Then:

1. If $\omega \in \Omega^r(M)$ then

$$L_X \omega(Y_1, \ldots, Y_r) = X(\omega(Y_1, \ldots, Y_r)) - \sum_i \omega(Y_1, \ldots, [X, Y_i], \ldots, Y_r).$$

2. $L_X(\omega \wedge \eta) = L_X\omega \wedge \eta + \omega \wedge L_X \eta$.
3. $L_X(d\omega) = d(L_X \omega)$.

(1) is immediate from the previous result, and (2) follows from (1). To see (3), observe that $L_X \circ d - d \circ L_X$ is a skew-derivation of $\mathcal{T}(M)$ and vanishes on $f$ and $df$ for $f \in C^\infty(M)$, and hence an argument similar to the previous result shows it vanishes identically. ▶

We should note here that $L_X Y(p)$ does **not** just depend on the value of $X_p$, it also depends on a $X$ on a neighborhood of $p$. In the next chapter we will define a **connection**, $D_X Y(p)$ which will depend only on $X_p$, and thus give us a way to 'differentiate' vector fields.

## 7.12 The interior product

Let $X$ be a vector field on a smooth manifold $M$. Let $\omega$ be a smooth form of positive degree $r$. Define the **interior product** of $\omega$ with $X$, the **contraction** of $\omega$ with $X$ to be $i_X w$, an $(r-1)$-form defined by

$$i_X(\omega)(Y_1, \ldots, Y_{r-1}) := \omega(X, Y_1, \ldots, Y_{r-1}).$$

Define $i_X f = 0$ for a form of degree 0.

## 7.13 Theorem (properties of $i_X$)

$i_X$ is unique linear mapping $\Omega^*(M) \to \Omega^*(M)$ satisfying

1. $i_X f = 0$ for $f \in C^\infty(M)$.
2. $i_X \omega = \omega(X)$ for $\omega \in \Omega^1(M)$.
3. $i_X(\omega \wedge \omega') = i_X \omega \wedge \eta + (-1)^{\deg(\omega)} \omega \wedge i_X \omega'$.

Moreover $i_X \circ i_X = 0$ and we have **Cartan's formula**

$$L_X\omega = i_X(d\omega) + d(i_X\omega).$$

◀ Uniqueness follows by a similar argument to Theorem 7.10. To prove (3) we first show that

$$i_X\left(\omega^1 \wedge \cdots \wedge \omega^k\right) = \sum_{i=1}^{k}(-1)^{i=1}\omega^i(X)\omega^1 \wedge \cdots \wedge \widehat{\omega^i} \wedge \cdots \wedge \omega^k.$$

To check this, write $X_1 = X$ and we evaluate both sides on vector fields $X_2, \ldots, X_k$. We need to show

$$\left(\omega^1 \wedge \cdots \wedge \omega^k\right)(X_1, \ldots, X_k) = \sum_{i=1}^{k}(-1)^{i=1}\omega^i(X_1)\left(\omega^1 \wedge \cdots \wedge \widehat{\omega^i} \wedge \cdots \wedge \omega^k\right)(X_2, \ldots, X_k). \quad (23)$$

Let $\mathfrak{X}$ denote the matrix $\mathfrak{X} = \left[\omega^i(X_j)\right]$. Then the left-hand side of (23) is $\det \mathfrak{X}$. If $\mathfrak{X}_{(i,j)}$ denotes the $(k-1) \times (k-1)$ minor of $\mathfrak{X}$ obtained by deleting row $i$ and column $j$ then the right-hand side of (23) is

$$\sum_{i=1}^{k}(-1)^{i-1}\omega^i(X_1)\det \mathfrak{X}_{(i,j)}.$$

Equality follows as the above is the cofactor expansion of $\det \mathfrak{X}$ in the first column.

It then follows that (3) holds for $\omega = f dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ and $\eta = g dx^{j_1} \wedge \cdots \wedge dx^{j_q}$, and the general case follows from this by linearity.

The fact the $i_X \circ i_X = 0$ is clear since by (3), $i_X \circ i_X$ is a derivation that vanishes on $C^\infty(M)$ and $\mathcal{T}^{(0,1)}(M)$, and thus vanishes identically. Finally Cartan's formula follows as both sides are derivations that coincide on $C^\infty(M)$ and $\Omega^1(M)$. ▶

# 8 Connections on vector bundles

## 8.1 Vector bundle valued forms

Suppose $\pi : E \to M$ is a smooth vector bundle of rank $m$ over a smooth manifold $M$. Consider the bundle $\Lambda^r(T^*(M)) \otimes E$. We write $\mathcal{A}^r$ for the sheaf of sections

$$\mathcal{A}^r(U) = \Gamma\left(U, \Lambda^r\left(T^*(M)\right)\right).$$

We call elements of $\mathcal{A}^r$ **$E$-valued $r$-forms**. By definition we set $\mathcal{A}^0 = \mathcal{E}$, that is,

$$\mathcal{A}^0(U) = \mathcal{E}(U) = \Gamma(U, E).$$

Under the standard identification $\Lambda^r(T^*(M)) \cong (\Lambda^r(T(M))^*$ that we have been making, we have

$$\Lambda^r(T^*(M)) \otimes E \cong (\Lambda^r(T(M))^* \otimes E \cong \operatorname{Hom}(\Lambda^r(T(M)), E).$$

In other words, we can think of the fibre $(\Lambda^r(T^*(M)) \otimes E)_p$ to be the set of alternating $r$-multilinear maps $T_p(M) \times \cdots \times T_p(M) \to E_p$.

## 8.2 Trivialising $\Lambda^r(T^*(M)) \otimes E$

Suppose first that $U \subseteq M$ is a trivialising neighborhood for $E$, and $\mathbf{e} = \{e_1, \ldots, e_m\}$ is a local frame. Given $p \in U$, let $\{\omega^1, \ldots, \omega^\ell\}$ be a local frame for $\Lambda^r(T^*(M))$. Then given $s \in \mathcal{A}^r(U)$ as

$$s = f^i_j\omega^j \otimes e_i,$$

where the $f^i_j$ are uniquely determined $C^\infty$ functions near $p$. Let $\xi^i := f^i_j\omega^j$. One easily checks that the differential $r$-form $\xi^i$ is independent of the choice of local frame $\{\omega^1, \ldots, \omega^\ell\}$. Since $p$ was arbitrary we conclude that $\xi^i \in \Omega^r(U)$, and hence we can write

$$\xi = \xi^i \otimes e_i$$

on $U$. We will write $\underline{\xi}$ for the **column** vector

$$\underline{\xi} = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \\ \vdots \\ \xi^m \end{pmatrix},$$

and thus with $\mathbf{e}$ denoting the **row** vector $(e_1, \dots, e_m)$ we can write in matrix notation

$$\xi = \mathbf{e} \cdot \underline{\xi}$$

where the '$\cdot$' denotes matrix multiplication.

## 8.3 The $\Lambda^r(T^*(M)) \otimes$ End$(E)$ bundle

The next bundle we need to work with is $\Lambda^r(T^*(M)) \otimes \mathrm{End}(E)$; its sheaf of sections is denoted $\mathcal{A}^r_{\mathrm{End}(E)}$. If $U \subseteq M$ trivialises $E$ with local frame $\mathbf{e}$ then similarly to the above we can write a section $a \in \mathcal{A}^r_{\mathrm{End}(E)}(U)$ as

$$A = a^i_j \otimes \varepsilon^j \otimes e_i,$$

and we will let $A$ denote the matrix $\left[a^i_j\right]$, $a^i_j \in \Omega^r(U)$, where $\{\varepsilon^1, \dots, \varepsilon^m\}$ is the dual coframe to $\mathbf{e}$.

An element $a \in \mathcal{A}^p_{\mathrm{End}(E)}(U)$ can act on an element $\xi \in \mathcal{A}^q(U)$ in the obvious way

$$\left(a^i_j \otimes \varepsilon^j \otimes e_i\right) \wedge \left(\xi^k \otimes e_k\right) := a^i_k \wedge \xi^k \otimes e_k,$$

which we write in matrix notation as

$$a \wedge \xi = \mathbf{e} \cdot \left(A \cdot \underline{\xi}\right).$$

## 8.4 Definition

A **(linear) connection** of a vector bundle $E$ over $M$ is $\mathbb{R}$-linear sheaf morphism $D : \mathcal{A}^0 \to \mathcal{A}^1$ satisfying the Leibniz property

$$D(fs) = df \otimes s + fDs \tag{24}$$

for a local section $s$ of $E$ and a local smooth function $f$.

If $U \subseteq M$ is a trivialising neighborhood for $E$ with local frame $\mathbf{e}$ then we may associate a **connection matrix** $\theta$ of $D$ to $\mathbf{e}$, given as follows. Since $De_i \in \mathcal{A}^1(U)$ we can write

$$De_i = \theta^i_j \otimes e_j$$

for $\theta^i_j \in \Omega^1(U)$, and we let $\theta = \left[\theta^i_j\right]$. An element $s \in \mathcal{A}^0(U)$ can be written as $s = s^i e_i$ for $s^i \in C^\infty(U)$, or in matrix notation, $s = \mathbf{e} \cdot \underline{s}$, where as above $\underline{s}$ denotes the **column** vector. Then we have

$$\begin{aligned} Ds &= D\left(s^i e_i\right) \\ &= ds^i \otimes e_i + s^i De_i \\ &= \left(ds^j + s^i \theta^j_i\right) \otimes e_j, \end{aligned}$$

or in matrix notation,

$$D(\mathbf{e} \cdot \underline{s}) = \mathbf{e} \cdot (d\underline{s} + \theta \cdot \underline{s}),$$

where $d\underline{s}$ denotes the column vector

$$d\underline{s} = \begin{pmatrix} ds^1 \\ ds^2 \\ \vdots \\ \\ \vdots \\ ds^m \end{pmatrix}.$$

We will write

$$D\underline{s} = d\underline{s} + \theta \cdot \underline{s} \tag{25}$$

as a shorthand for the previous equation, and say that $D$ **acts locally** as $D = d + \theta$.

We say a section $s$ is **parallel** with respect to $D$ if $Ds = 0$.

## 8.5 Changing the frame

We now investigate how the connection matrix changes when we change the frame. Suppose $\mathbf{e}' = \{e'_1, \ldots, e'_m\}$ is another local frame. Then there exists a map $\psi : U \to GL(m, \mathbb{R})$ such that

$$e'_j = \psi^j_i e_i,$$

where $\psi = \left[\psi^j_i\right]$. In matrix notation,

$$\mathbf{e}' = \mathbf{e} \cdot \psi.$$

The corresponding column vector $\underline{s}$ changes to $\underline{s}'$ where

$$\underline{s}' = \psi \cdot \underline{s},$$

since $s'^j e'_j = s^i e_i$ implies $s^i = \psi^i_j s'^j$. Then if $\theta'$ is the matrix with respect to $\mathbf{e}'$ then

$$\begin{aligned} D\mathbf{e}' &= D(\mathbf{e} \cdot \psi) \\ &= \mathbf{e} \cdot d\psi + \mathbf{e} \cdot \theta \\ &= \mathbf{e}' \cdot \psi^{-1} \cdot d\psi + \mathbf{e}' \cdot \psi^{-1} \cdot \theta \cdot \psi, \end{aligned}$$

where $d\psi = \left[d\psi^j_i\right]$ and so

$$\theta' = \psi^{-1} \cdot d\psi + \psi^{-1} \cdot \theta \cdot \psi. \tag{26}$$

## 8.6 Example

This is the 'standard' connection on $T(\mathbb{R}^n)$. Define $D$ by $D\left(\frac{\partial}{\partial r^i}\right) = 0$. This clearly satisfies the required properties of a connection. Thus if $X = X^i \frac{\partial}{\partial r^i}$ is a smooth vector field, we have

$$DX = D\left(X^i \frac{\partial}{\partial r^i}\right) = \sum_{i=1}^{n} dX^i \frac{\partial}{\partial r^i},$$

and thus $X$ is parallel with respect to $D$ if and only if $dX^i = 0$ for all $i$, that is, each of the $X^i$ are constant.

## 8.7 Lemma

Let $E$ be a rank $m$ vector bundle over $M^n$. Then there exists a connection on $E$.

◀ Let $\{U_\alpha\}$ be a open covering of $M$ by trivialising neighborhoods of $E$. Let $\mathbf{e}^\alpha = \{e^\alpha_1, \ldots, e^\alpha_m\}$ be a local frame associated to the local trivialisation $\pi^{-1}(U_\alpha) \cong U_\alpha \times \mathbb{R}^m$. Define a connection $D_\alpha$ on $\pi^{-1}(U_\alpha)$ by $D_\alpha(e^\alpha_k) = 0$ for $k = 1, \ldots, m$. Now let $\{\lambda_i\}$ be a partition of unity subordinate to $\{U_\alpha\}$, and define

$$D = \sum_{i=1}^{\infty} \lambda_i D_{\alpha(i)},$$

where $\text{supp}(\lambda_i) \subseteq U_{\alpha(i)}$. This is again clearly a connection. ▶

## 8.8 Lemma

Suppose $E$ is a smooth vector bundle over $M$, and $D_1$ and $D_2$ are two connections on $E$. Then $D_1 - D_2$ determines an element of $\mathcal{A}^1_{\mathrm{End}(E)}(M)$, that is, a global section of $\Lambda^1(T^*(M)) \otimes \mathrm{End}(E)$.

◀ If $s$ is a local section of $E$ and $f$ a local smooth function then

$$\begin{aligned}(D_1 - D_2)(fs) &= (df \otimes s + fD_1 s) - (df \otimes s + fD_2 s)\\ &= f(D_1 - D_2)s.\end{aligned}$$

Thus $D_1 - D_2$ corresponds to a vector bundle morphism $E \to \Lambda^1(T^*(M)) \otimes E$ by Proposition 5.28 and thus also a global section of $\Lambda^1(T^*(M)) \otimes \mathrm{End}(E)$. ▶

## 8.9 Corollary

The space of connections on a bundle are an (infinite dimensional) affine space over the vector space $\mathcal{A}^1_{\mathrm{End}(E)}(M)$.

## 8.10 Covariant derivatives

Let $E$ be a vector bundle of rank $m$ over a smooth manifold $M^n$. Let $D$ be a connection on $E$. Then $D$ has a natural extension to an $\mathbb{R}$-linear sheaf morphism $d^E : \mathcal{A}^r \to \mathcal{A}^{r+1}$ defines as follows: if $\omega$ is a local $r$-form on $M$ and $s$ is a local section of $E$ then we set

$$d^E(\omega \otimes s) = d\omega \otimes s + (-1)^r \omega \wedge Ds.$$

Observe that for $r = 0$ this is just the Leibniz rule (24), which is also ensures that $d^E$ is well defined, that is,

$$d^E(\omega \otimes fs) = d^E(f\omega \otimes s).$$

Moreover a generalised Leibniz rule also holds; namely if $\xi \in \mathcal{A}^p$ and $\omega \in \Omega^q$ then

$$d^E(\omega \wedge \xi) = d\omega \wedge \xi + (-1)^q \omega \wedge d^E \xi.$$

Indeed, it is enough to verify this for $\xi = \eta \otimes s$ where $\eta$ is a local $p$-form and $s$ a local section of $E$. Then

$$\begin{aligned}d^E(\omega \wedge \xi) &= d^E((\omega \wedge \eta) \otimes s)\\ &= d(\omega \wedge \eta) \otimes s + (-1)^{p+q}(\omega \wedge \eta) \otimes Ds\\ &= d\omega \wedge \xi + (-1)^q \{(\omega \wedge d\eta) \otimes s + (-1)^p(\omega \wedge \eta) \otimes Ds\}\\ &= d\omega \wedge \xi + (-1)^q \omega \wedge d^E \xi.\end{aligned}$$

## 8.11 Definition

Let $\pi : E \to M$ be a vector bundle and $D$ a connection on $E$. Let $d^E$ denote the corresponding covariant derivative, and consider

$$\mathcal{R} := d^E \circ d^E : \mathcal{A}^0 \to \mathcal{A}^2.$$

We call $\mathcal{R}$ the **curvature** of $D$; it is the obstruction to $\{\mathcal{A}^*, d^E\}$ being a complex. Unlike $D$, $\mathcal{R}$ is linear over the smooth functions, as one easily checks:

$$\begin{aligned}\mathcal{R}(fs) &= d^E(df \otimes s + fDs)\\ &= -df \wedge Ds + df \wedge Ds + f\mathcal{R}s\\ &= f\mathcal{R}s.\end{aligned}$$

Thus by Proposition 5.28, $\mathcal{R}$ corresponds to a vector bundle morphism $E \to \Lambda^2(T^*(M)) \otimes E$ and hence a global section $R$ of the bundle $\Lambda^2(T^*(M)) \otimes \mathrm{End}(E)$. We will also call $R$ the curvature of $D$. Observe we can also think of $R$ as an element of $\mathcal{T}^{(1,3)}(M)$.

## 8.12 Relating $\mathcal{R}$ and $R$

$\mathcal{R}$ and $R$ are related by

$$\mathcal{R}s = R \wedge s.$$

In fact the following is true: for all $r \geq 0$ we have

$$d^E \circ d^E : \xi \mapsto R \wedge \xi.$$

Indeed, it is enough to verify this for $\xi = \omega \otimes s$ where $\omega$ is a local $r$-form and $s$ a local section of $E$. Then

$$\begin{aligned} d^E\left(d^E \xi\right) &= d^E\left(d\omega \otimes s + (-1)^r \omega \wedge Ds\right) \\ &= (-1)^{r+1} d\omega \wedge Ds + (-1)^r d\omega \wedge Ds + (-1)^{2r} \omega \wedge \mathcal{R}s \\ &= \omega \wedge \mathcal{R}s \\ &= \omega \wedge (R \wedge s) \\ &= R \wedge (\omega \otimes s) \\ &= R \wedge \xi. \end{aligned}$$

## 8.13 The curvature with respect to a local frame

Let $E$ be a smooth rank $m$ vector bundle over $M^n$, $D$ a connection on $E$ and $\mathcal{R}$ the curvature of $D$. Suppose $U \subseteq M$ is a trivialising neighborhood for $E$, and $\mathbf{e} = \{e_1, \dots, e_m\}$ a local frame. Let $\{\varepsilon^1, \dots, \varepsilon^m\}$ denote the corresponding coframe. There exist 2-forms $\Theta_i^j \in \Omega^2(U)$ such that

$$\mathcal{R}e_i = \Theta_i^j \otimes e_j,$$

we let $\Theta = \left[\Theta_i^j\right]$ denote the matrix, called the **curvature matrix** of $\mathcal{R}$ with respect to $\mathbf{e}$. Thus

$$R = \Theta_k^j \otimes \varepsilon^k \otimes e_j.$$

Observe that

$$\begin{aligned} \mathcal{R}e_i &= d^E\left(De_i\right) \\ &= d^E\left(\theta_i^j \otimes e_j\right) \\ &= d\theta_i^j \otimes e_j - \theta_i^j \wedge \theta_j^k e_k \\ &= \left(d\theta_i^k + \theta_j^k \wedge \theta_i^j\right) \otimes e_k, \end{aligned}$$

and thus

$$\Theta_i^k = d\theta_i^k + \theta_j^k \wedge \theta_i^j, \tag{27}$$

or in matrix notation

$$\Theta = d\theta + \theta \wedge \theta. \tag{28}$$

This time changing the frame is relatively painless. If $\mathbf{e}' = \mathbf{e} \cdot \psi$ is another frame then since $\mathcal{R}$ is linear over smooth functions,

$$\begin{aligned} \mathcal{R}\mathbf{e}' &= \mathcal{R}\left(\mathbf{e} \cdot \psi\right) \\ &= \mathcal{R}\mathbf{e} \cdot \psi \\ &= \mathbf{e} \cdot \Theta \cdot \psi \\ &= \mathbf{e}' \cdot \psi^{-1} \cdot \Theta \cdot \psi, \end{aligned}$$

and hence the curvature matrix $\Theta'$ of $\mathcal{R}$ with respect to $\mathbf{e}'$ is related to $\Theta$ by

$$\Theta' = \psi^{-1} \cdot \Theta \cdot \psi. \tag{29}$$

## 8.14 Proposition (the general Bianchi identity)

Let $\pi : E \to M$ be a vector bundle and $D$ a connection on $E$. Then in any local trivialisation the following relationship holds between the connection and curvature matrices:

$$d\Theta = \Theta \wedge \theta - \theta \wedge \Theta. \tag{30}$$

◀ We simply apply the exterior derivative to both sides of (28) to obtain:

$$\begin{aligned} d\Theta &= d\theta \wedge \theta - \theta \wedge d\theta \\ &= (\Theta - \theta \wedge \theta) \wedge \theta - \theta \wedge (\Theta - \theta \wedge \theta) \\ &= \Theta \wedge \theta - \theta \wedge \Theta. \end{aligned}$$ ▶

## 8.15 The dual connection

Let $\pi : E \to M$ be a vector bundle of rank $m$, and $D$ a connection on $E$. We wish to define a connection $D^*$ on the dual bundle $E^*$. We do so by the recipe

$$d(\sigma(s)) = (D^*\sigma)(s) - \sigma(Ds) \tag{31}$$

for $\sigma$ a local section of $E^*$ and $s$ a local section of $E$. It is easily checked that (31) does indeed define a connection. Suppose $\mathbf{e} = \{e_1, \ldots, e_m\}$ is a local frame for $E$ over $U \subseteq M$, so the dual coframe $\mathbf{e}^* = \{e_1^*, \ldots, e_m^*\}$ is a local frame for $E^*$ (here we are using $e_i^*$ for the element normally denoted $\varepsilon^i$, i.e. $e_i^*(e_j) = \delta_{ij}$ - this is order to make the indices work). Thus (31) yields

$$0 = d(e_i^*(e_j)) = (D^* e_i^*)(e_j) - e_i^*(De_j),$$

and hence if $\theta^*$ denotes the connection matrix of $D^*$ with respect to $\mathbf{e}^*$, that is,

$$D^* e_i^* = \theta^{*j}_{\ i} e_j^*$$

then we have

$$\begin{aligned} 0 &= \theta^{*k}_{\ i} e_k^*(e_j) - e_i^*\left(\theta^k_j e_j\right) \\ &= \theta^{*j}_{\ i} - \theta^i_j, \end{aligned}$$

and hence

$$\theta^* = -\theta^t.$$

## 8.16 Further new connections

Suppose $\pi : E \to M$ and $\pi' : E' \to M$ are vector bundles over $M$ with connections $D$ and $D'$ respectively. We define a connection $D \otimes D'$ on the bundle $E \otimes E'$ by

$$(D \otimes D')(s \otimes s') = Ds \otimes s' + s \otimes D's',$$

and then extending by linearity to all local sections of $E \otimes E'$. It is clear that this defines a connection. In particular we are interested in using this formula to define a connection on $E^* \otimes E = \mathrm{End}(E)$. Write $\tilde{D}$ for the connection $D^* \otimes D$ on $E^* \otimes E$. Then if $D$ has connection matrix $\theta$ and $D^*$ has connection matrix $\theta^*$ with respect to local frames $\mathbf{e}, \mathbf{e}^*$ over $U \subseteq M$ respectively then if $a = a^i_j e_i^* \otimes e_j$ is an arbitrary element of $\mathcal{A}^0_{\mathrm{End}(E)}(U)$ then

$$\tilde{D}a = da^i_j e_i^* \otimes e_j + a^i_j \theta^{*k}_{\ i} e_k^* \otimes e_j + a^i_j e_i^* \otimes \theta^k_j e_j,$$

and so

$$\tilde{D}a = \left(da^k_j + \theta^k_i a^i_j - a^i_j \theta^j_k\right) e_k^* \otimes e_j,$$

and so if $A$ denotes the matrix $\left[a^i_j\right]$ then (with a similar shorthand notation to (25))

$$\tilde{D}A = dA + \theta A - A\theta,$$

that is, $\tilde{D}$ acts locally by

$$\tilde{D} = d + [\theta, \cdot].$$

If we write $d^{\mathrm{End}(E)}$ for the corresponding covariant derivative then $d^{\mathrm{End}(E)}$ acts locally in the same way as $\tilde{D}$, that is, if now $a \in \mathcal{A}^r_{\mathrm{End}(E)}(U)$, so $a^i_j \in \Omega^r(U)$ then we still have

$$d^{\mathrm{End}(E)} A = dA + A \wedge \theta - A \wedge \theta. \tag{32}$$

## 8.17 The coordinate-free version of the general Bianchi identity

In particular, we can apply this to $\Theta \in \mathcal{A}^2_{\mathrm{End}(E)}(U)$ to obtain

$$d^{\mathrm{End}(E)}\Theta = d\Theta + \Theta \wedge \theta - \theta \wedge \Theta = 0,$$

by (30). In fact it is clear that the statement

$$d^{\mathrm{End}(E)}\Theta = 0$$

is equivalent to (30), and thus is a coordinate-free way of stating the general Bianchi identity.

## 8.18 Definition

Let $\pi : E \to M$ be a vector bundle and $D$ a connection on $E$. Let $X \in \mathcal{X}(U)$ for some open set $U \subseteq M$. Then we define an $\mathbb{R}$-linear sheaf morphism $D_X : \mathcal{A}^0|_U \to \mathcal{A}^0|_U$ (recall the notation $\mathcal{A}^0|_U$ denotes the sheaf $\mathcal{A}^0$ restricted to $U$; see Example 7 of Section 5.26) by

$$D_X s = Ds(X).$$

Thus for $s$ a local section and $f$ a local smooth function we have

$$D_X(fs) = df(X) + D_X s = Xf + D_X s.$$

Similarly given $X, Y$ local vector fields defined on some open set $U \subseteq M$ we can define an $\mathbb{R}$-linear sheaf morphism $R(X,Y) : \mathcal{A}^0|_U \to \mathcal{A}^0|_U$ by

$$R(X,Y)s = (\mathcal{R}s)(X,Y).$$

Thus for $s$ a local section and $f$ a local smooth function we have

$$R(X,Y)(fs) = fR(X,Y)s.$$

This notation will become more helpful in the following chapter, where we will be investigating connections on the tangent bundle of a smooth manifold, $M$ and will primarily view them as operators $\mathcal{X}(M) \to \mathcal{X}(M)$ (recall that for $E = T(M)$, $\mathcal{X}(M) = \mathcal{A}^0(M)$). For now however we will use it to prove a very useful formula for the curvature $R$.

## 8.19 Lemma

Let $E$ be a smooth rank $m$ bundle over $M$ and $D$ a connection on $M$. Let $X, Y$ be local vector fields defined on $U \subseteq M$. Then

$$R(X,Y)s = D_X D_Y s - D_Y D_X s - D_{[X,Y]} s. \tag{33}$$

◀ The result will follow almost immediately from the following more general statement. Let $V$ be open in $U$. Let $\mu \in \mathcal{A}^1(V)$. Then we claim

$$d^E(\mu)(X,Y) = D_X(\mu(Y)) - D_Y(\mu(X)) - \mu([X,Y]). \tag{34}$$

Without loss of generality, we may assume $\mu = \omega \otimes s$ where $\omega \in \Omega^1(V)$ and $s \in \mathcal{A}^0(V)$. Then

$$d^E(\mu) = d\omega \otimes s - \omega \wedge Ds,$$

and hence

$$\begin{aligned} d^E(\mu)(X,Y) &= d\omega \otimes s(X,Y) - (\omega \wedge Ds)(X,Y) \\ &= d\omega(X,Y) \otimes s - \omega(X) D_Y s + \omega(Y) D_X s. \end{aligned}$$

Now by (23)

$$d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]),$$

and thus

$$d^E(\mu)(X,Y) = \{X\omega(Y) \otimes s + \omega(Y) D_X s\} - \{Y\omega(X) \otimes s + \omega(X) D_Y s\} - \{\omega([X,Y]) \otimes s\}.$$

Then since $\mu([X,Y]) = \omega([X,Y]) \otimes s$ and

$$\begin{aligned} D_X(\mu(Y)) &= D(\omega(Y)s)(X) \\ &= (d(\omega(Y)) \otimes s)(X) + \omega(Y) D(s)(X) \\ &= X\omega(Y) \otimes s + \omega(Y) D_X(s), \end{aligned}$$

and similarly

$$D_Y(\mu(X)) = Y\omega(X) \otimes s + \omega(X) D_Y s,$$

(34) follows. Then to complete the proof we have

$$\begin{aligned} \mathcal{R}s(X,Y) &= d^E(Ds)(X,Y) \\ &= D_X(Ds(Y)) - D_Y(Ds(X)) - Ds([X,Y]) \\ & D_X D_Y s - D_Y D_X s - D_{[X,Y]} s, \end{aligned}$$

and thus $R(X,Y)s = D_X D_Y s - D_Y D_X s - D_{[X,Y]} s$ as claimed. ▶

## 8.20 Definition

Let $E$ be a smooth vector bundle of rank $r$ over $M^n$, $\langle \cdot, \cdot \rangle$ a metric on $E$ and $D$ a connection on $E$. We say that $D$ is **orthogonal** if

$$d\langle s_1, s_2 \rangle = \langle Ds_1, s_2 \rangle + \langle s_1, Ds_2 \rangle$$

for any sections local sections $s_1, s_2$ of $E$. Equivalently, for any smooth vector field $X$ we require

$$X\langle s_1, s_2 \rangle = \langle D_X s_1, s_2 \rangle + \langle s_1, D_X s_2 \rangle. \tag{35}$$

## 8.21 Proposition

An orthogonal connection $D$ with respect to a metric $\langle \cdot, \cdot \rangle$ has skew-symmetric connection and curvature matrices with respect to any orthonormal local frame $\mathbf{e} = \{e_1, \ldots, e_m\}$.

◀ Let $D$ be orthogonal connection and $\mathbf{e} = \{e_1, \ldots, e_m\}$ an orthonormal frame and $\theta$ the curvature matrix of $D$ with respect to $\mathbf{e}$. Then

$$\begin{aligned} 0 &= d\langle e_i, e_j \rangle = \langle De_i, e_j \rangle + \langle e_i, De_j \rangle \\ &= \langle \theta_i^k e_k, e_j \rangle + \langle e_i, \theta_j^\ell e_\ell \rangle \\ &= \theta_i^j + \theta_j^i, \end{aligned}$$

which shows that $\theta$ is skew-symmetric. Then by (27),

$$\begin{aligned} \Theta_i^k &= d\theta_i^k + \theta_j^k \wedge \theta_i^j \\ &= -d\theta_k^i - \theta_i^j \wedge \theta_j^k \\ &= -d\theta_k^i - (-1)^2 \theta_j^i \wedge \theta_k^j \\ &= -\Theta_k^i, \end{aligned}$$

we see that $\Theta$ is also skew-symmetric. ▶

# 9 Koszul connections

## 9.1 Koszul connections

We now specialise to the case $E = T(M)$, and thus consider connections on the tangent bundle. These are called **Koszul connections**. We often abuse language and refer to such a connection as a connection on $M$. Notationally we will use the symbol '$\nabla$' instead of '$D$' to distinguish between Koszul connections and arbitary connections.

Let $\nabla$ be a Koszul connection on $M^n$. Suppose $(x^1, \dots, x^n)$ are local coordinates on $U$. Then $U$ trivialises $T(M)$ and $\Omega^1(M)$, so if $\theta$ is the connection matrix for $\nabla$ over $U$ then we can write

$$\theta_j^k = \Gamma_{ij}^k dx^i$$

for some smooth functions $\Gamma_{ij}^k : U \to \mathbb{R}$.We call the $n^3$ functions $\Gamma_{ij}^k$ the **Christoffel symbols** of the connection with respect to the local frame $\{\partial_1, \dots, \partial_n\}$.

Thus

$$\nabla \partial_j = \Gamma_{ij}^k dx^i \otimes \partial_k,$$

and hence

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

Similarly we can decompose the curvature matrix $\Theta$ of $\nabla$ over $U$. We define

$$R_{jk\ell}^i := \Theta_j^i (\partial_k, \partial_\ell),$$

so that

$$\Theta_j^i = \frac{1}{2} R_{jk\ell}^i dx^k \wedge dx^\ell. \tag{36}$$

Thus

$$R(\partial_k, \partial_\ell)(\partial_j) = R_{jk\ell}^i \partial_i,$$

and thus

$$R = R_{jk\ell}^i dx^k \otimes dx^\ell \otimes dx^j \otimes \partial_i$$

Since $\Theta_j^i$ is alternating, we have

$$R_{jk\ell}^i = -R_{j\ell k}^i \quad \text{for all } 1 \leq i, j, k, \ell \leq n. \tag{37}$$

More generally, for local vector fields $X, Y$ and $Z$ the local vector fields $R(X,Y)Z$ and $R(Y,X)Z$ satisfy

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

## 9.2 Definition

Let $c : [a,b] \to M^n$ be a smooth curve in $M$. A **vector field along** $c$ (note that this is **not** normally an actual vector field) is a smooth function $V : [a,b] \to T(M)$ such that

$$V(t) \in T_{c(t)}(M) \text{ for all } t \in [a,b].$$

If $(x^1, \dots, x^n)$ are local coordinates on $M$, we may write a vector field $V$ along $\gamma$ as

$$V(t) = \sum_{i=1}^n V^i(t) \partial_i|_{c(t)},$$

and the assertion that $V$ is smooth is that is equivalent to the $V^i(t)$ being smooth functions $c^{-1}(U) \to M$.

Let $\mathrm{Vect}(c)$ denote the set of all vector fields along $c$. Note that in particular, $\dot{c}(t)$ is a vector field along $c$, and more generally if $t_0 \in (a,b)$ and $(x^1, \dots, x^n)$ are local coordinates about $t_0$ then $\partial_i \circ c$ is a vector field along $c$ for $t$ sufficiently small, where

$$(\partial_i \circ c)(t) := \partial_i|_{c(t)}.$$

## 9.3 Definition

Let $M$ be a smooth manifold, and $\nabla$ a Koszul connection on $M$, and $c : [a,b] \to M$ a smooth curve. Suppose $X$ is a vector field defined in a neighborhood of $c([a,b])$. Then we can define a vector field along $c$, written $\frac{DX}{dr}$ by

$$\frac{DX}{dr}(t) = \nabla_{\dot{c}(t)} X = \nabla X(\dot{c}(t)) \in T_{c(t)}(M).$$

We call $\frac{DX}{dr}$ the **covariant derivative** of $X$ along $c$ with respect to $\nabla$. We wish to generalise this, so that if $V$ is any vector field along $c$ we can define another covariant derivative $\frac{DV}{dr}$ along $c$. The notation $\frac{D}{dr}$ is somewhat confusing - it depends on $c$ and $\nabla$as well.

## 9.4 Proposition

Let $M^n$ be a smooth manifold, $\nabla$ a Koszul connection on $M$ and $c : [a,b] \to M$ a smooth curve on $M$. Then there exists a unique operation

$$\frac{D}{dr} : \mathrm{Vect}(c) \to \mathrm{Vect}(c)$$

such that:

1. Given $V, W \in \mathrm{Vect}(c)$, we have
$$\frac{D(V+W)}{dr} = \frac{DV}{dr} + \frac{DW}{dr}.$$
2. Given a smooth function $f : [a,b] \to \mathbb{R}$ and $V \in \mathrm{Vect}(c)$,
$$\frac{D(fV)}{dr}(t) = \frac{df}{dr}(t)V(t) + f(t)\frac{DV}{dr}(t).$$
3. If $X$ is a smooth vector field defined in a neighborhood of $c(t_0)$ then $\frac{DX}{dr}$ is equal to the construction given above, that is, for $t$ sufficiently close to $t_0$,
$$\frac{DX}{dr}(t) = \nabla_{\dot{c}(t)} X.$$

◀ Let $t_0 \in (a,b)$ and set $p := c(t_0)$. Let $(x^1, \ldots, x^n)$ be local coordinates on $U \subseteq M$, where $U$ is some neighborhood of $p$, and select $\epsilon > 0$ such that $I := (t_0 - \epsilon, t_0 + \epsilon) \subseteq [a,b]$ and $c(I) \subseteq U$. Then if $V$ is a vector field along $c$ we can write for $t \in I$

$$V(t) = V^j(t)\partial_j|_{c(t)},$$

that is,

$$V = V^j \cdot \partial_j \circ c,$$

with the $V^i : I \to \mathbb{R}$ smooth functions. We show that if we have a function $\frac{D}{dr} : \mathrm{Vect}(c) \to \mathrm{Vect}(c)$ satisfying (1), (2) and (3) then $\frac{D}{dr}$ is uniquely determined in $U$.

Then by (1) and (2),

$$\frac{DV}{dt} = \frac{dV^j}{dr} \cdot \partial_j \circ c + V^j \frac{D}{dr}(\partial_j \circ c).$$

But since $\partial_i \circ c$ is a smooth vector field defined in a neighborhood of $t_0$, by (3)

$$\frac{D}{dr}(\partial_i \circ c) = \nabla_{\dot{c}(t)}(\partial_i \circ c),$$

and since by (7) we gave

$$\dot{c}(t) = \frac{dc^i}{dr}(t)\partial_i|_{c(t)},$$

it follows that

$$\nabla_{\dot{c}(t)}(\partial_j \circ c) = \frac{dc^i}{dr} \cdot \Gamma^k_{ij} \circ c \cdot \partial_k \circ c,$$

and hence putting this all together we obtain

$$\frac{DV}{dr}(t) = \left(\frac{dV^k}{dr}(t) + V^j(t)\frac{dc^i}{dr}(t)\Gamma^k_{ij}(c(t))\right)\partial_k|_{c(t)}. \tag{38}$$

Thus $\frac{DV}{dt}$ is uniquely determined by conditions (1), (2) and (3). Moreover, the argument above reverses to show that defining $\frac{DV}{dr}$ locally as above does indeed satisfy (1), (2) and (3). Since we already know that quantities in the formula are well defined when we change coordinates, it follows that defining $\frac{D}{dr}$ locally as this does indeed yield a well defined map $\mathrm{Vect}(c) \to \mathrm{Vect}(c)$, and this completes the proof. ▶

## 9.5 Definition

We say that a vector field $V$ along $c$ is **parallel** along $c$ if $\frac{DV}{dr} \equiv 0$. A **parallel frame** along $c$ is a set $\{V_1, \dots, V_n\}$ of vector fields along $V$ such that each $V_i$ is parallel along $c$, and that for all $t \in [a,b]$, $\{V_1(t), \dots, V_n(t)\}$ is a basis of $T_{c(t)}(M)$.

## 9.6 Lemma

Given a smooth curve $c : [a,b] \to M$ and a tangent vector $v_a \in T_{c(a)}(M)$, there exists a unique vector field $V$ along $c$ such that $V$ is parallel along $c$ and $V(a) = v_a$.

◀ Choose local coordinates $(x^1, \dots, x^n)$ on a coordinate neighborhood $U$ containing $c(a)$. Write $v_a = v_a^k \partial_k|_{c(a)}$ in this coordinate system.

By (38) finding such a $V$ on $U$ is equivalent to solving the system of ODE's

$$\frac{dV^k}{dr} + V^j . \frac{dc^i}{dr} \cdot \Gamma^k_{ij} \circ c = 0 \text{ for } k = 1, \dots, n, \tag{39}$$

subject to to the inital conditions

$$V^k(a) = v_a^k.$$

By standard ODE theory, we obtain unique smooth solutions $V^k(t)$ defined on $c^{-1}(U \cap c([a,b]))$. Then in exactly the same way we can repeatedly solve the (finitely many) initial value problems to define the $V^k(t)$ on all of $[a,b]$ (the solutions must agree on overlaps, due to the uniqueness clause of the ODE theory). Thus glueing the solutions together we obtain the desired vector field $V \in \mathrm{Vect}(c)$. ▶

## 9.7 Definition

Let $M^n$ be a smooth manifold, and $p, q$ in the same path component of $M$. Let $c : [a,b] \to M$ be a smooth curve with $c(a) = p, c(b) = q$. We define a map

$$\tau^c_{p,q} : T_p(M) \to T_q(M)$$

sending $v \in T_p(M)$ to the vector $V(b) \in T_q(M)$, where $V$ is the unique parallel vector field along $c$ with $V(a) = v$. From the properties of $\frac{D}{dr}$ it is clear that $\tau^c_{p,q}$ is linear. Moreover, if $\bar{c}$ denotes the path from $q$ to $p$ obtained by traversing $c$ backwards, it is clear that

$$\tau^{\bar{c}}_{q,p} = \left(\tau^c_{p,q}\right)^{-1},$$

and hence $\tau^c_{p,q}$ is a linear isomorphism between $T_p(M)$ and $T_q(M)$; this gives a way of 'connecting' two different tangent spaces together, and is the origin of the word 'connection'. We call $\tau^c_{p,q}(v) \in T_q(M)$ the vector obtained from $v \in T_p(M)$ by **parallel translation** along $c$.

We have thus shown that if $M$ carries a Koszul connection $\nabla$ then we can obtain a **system of parallel transport** on $M$; namely isomorphisms $\tau^c_{p,q} : T_p(M) \to T_q(M)$ for any points $p, q \in M$ that are in the same path component. In fact, as we shall now show, we can recover the connection for the system of parallel transport.

## 9.8 Proposition

Let $c : (-\epsilon, \epsilon) \to M$ be a curve in $M$ with $c(0) = p$ and $\dot{c}(0) = v \in T_p(M)$. Write

$$\tau_t := \tau^c_{p,c(t)} : T_p(M) \to T_{c(t)}(M).$$

Then for any local vector field $Y$, it holds that

$$\nabla_v Y = \lim_{t\to 0} \frac{\tau_t^{-1}(Y(c(t))) - Y(p)}{t}.$$

◀ Fix a basis $\{v_1, \dots, v_n\}$ of $T_p(M)$, and let $\{V_1, \dots, V_n\}$ be the correponding parallel vector fields along $c$. Then $\{V_1, \dots, V_n\}$ is a parallel frame along $c$, since if $a^i V_i(t) = 0$ for some $t \in (-\epsilon, \epsilon)$ and $a^i \in \mathbb{R}$ then applying $\tau_{-t}$ to both sides we obtain $a^i v_i = 0$, whence $a^i = 0$ for all $1 \le i \le n$.

Set

$$Y(c(t)) = Y^i(t) V_i(t),$$

where the $Y^i : (-\epsilon, \epsilon) \to \mathbb{R}$ are smooth. Then

$$\begin{aligned} \lim_{t\to 0} \frac{\tau_t^{-1}(Y(c(t)) - Y(p)}{t} &= \lim_{t\to 0} \sum_{i=1}^n \frac{Y^i(t)\tau_t^{-1}(V_i(t)) - Y^i(0)V_i(0)}{t} \\ &= \sum_{i=1}^n \lim_{t\to 0} \frac{Y^i(t) - Y^i(0)}{t} V_i(0) \\ &= \frac{dY^i}{dr}(0) V_i(0) \\ &= \frac{D}{dr}\left(Y^i V_i\right)(0) \\ &= \nabla_v Y. \quad ▶ \end{aligned} \tag{40}$$

In particular, this shows that if $X$ is a smooth vector field, we can compute $(\nabla_X Y)(p) = \nabla_{X_p} Y$ with knowledge only of $X_p$ and of the values $Y$ takes on a $c$ curve in $M$ with $c(0) = p$ and $\dot{c}(0) = X_p$. This should be contrasted with the Lie derivative (see the remark after the proof of Lemma 7.11), where we needed to know $X$ in a neighborhood of $p$ in order to compute $L_X Y(p)$.

## 9.9 Definition

Suppose now $T \in \mathcal{T}^{(k,\ell)}(M)$ with $k + \ell > 0$. We want to define $\tau_t(T)$. We do this as follows: for $\omega \in \Omega^1(M)$ we define $\tau_t^*(\omega)$ by

$$\tau_t^*(\omega)(Y) := \omega(\tau_t(Y)),$$

and then for $T \in \mathcal{T}^{(k,\ell)}(M)$ we define $\tau_t(T)$ by

$$\tau_t(T)\left(Y_1, \dots, Y_k, \omega^1, \dots, \omega^\ell\right) := T\left(\tau_t(Y_1), \dots, \tau_t(Y_k), \tau_t^*\left(\omega^1\right), \dots, \tau_t^*\left(\omega^\ell\right)\right).$$

## 9.10 Theorem (extension of $\nabla$ to $\mathcal{T}(M)$)

Let $M^n$ be a smooth manifold and $\nabla$ a Koszul connection on $M$. Given $p \in M$ and $T \in \mathcal{T}^{(k,\ell)}(M)$ with $k + \ell > 0$, define

$$\nabla_X T(p) := \begin{cases} \lim_{t\to 0} \frac{\tau_t^{-1}(T(c(p))) - T(p)}{t} & X_p \neq 0 \\ 0 & X_p = 0, \end{cases}$$

where $c$ is any curve such that $c(0) = p$ and $\dot{c}(0) = X_p$. Similarly given $f \in C^\infty(M)$, define

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

Then $\nabla_X$ is an $\mathbb{R}$-linear derivation of the full tensor algebra $\mathcal{T}(M)$ which preserves the type of tensors and commutes with all contractions.

Moreover $\nabla_X$ is the unqiue $\mathbb{R}$-linear derivation $\alpha$ of the full tensor algebra $\mathcal{T}(M)$ which preserves the type of tensors and commutes with all contractions such that

$$\alpha(Y) = \nabla_X Y, \quad \alpha(f) = Xf.$$

Finally $X \mapsto \nabla_X$ is linear over the smooth functions and $\nabla_{X+Y} = \nabla_X + \nabla_Y$.

◀ Uniqueness (which we will prove last) will ensure that $\nabla_X$ is well defined (i.e. independent of the choice of curve $c$). It is clear that $\nabla_X T$ is $\mathbb{R}$-linear in the $T$-variable, and since

$$\begin{aligned}
\nabla_X (S \otimes T)(p) &= \lim_{t\to 0} \frac{\tau_t^{-1}(S \otimes T)(c(t)) - (S \otimes T)(p)}{t} \\
&= \lim_{t\to 0} \frac{\tau_t^{-1}(S(c(t))) \otimes \tau_t^{-1}(T(c(t))) - S(p) \otimes T(p)}{t} \\
&= \lim_{t\to 0} \frac{\tau_t^{-1}(S(c(t))) - S(p)}{t} \otimes T(p) \\
&\quad + \lim_{t\to 0} \tau_t^{-1}(S(c(t))) \otimes \frac{\tau_t^{-1}(T(c(t))) - T(p)}{t} \\
&= \nabla_X S(p) \otimes T(p) + S(p) \otimes \nabla_X T(p)
\end{aligned}$$

(this calculation is formally identical if either $S, T$ or both are functions) we see that $\nabla_X$ is indeed a derivation of $\mathcal{T}(M)$. It is clear that $\nabla_X$ preserves the type of $T$, and to show $\nabla_X$ commutes with all the contractions, we will start with the special case $T = \omega \otimes Y$ with $\omega \in \mathcal{T}^{(0,1)}(M)$ so $CT = \omega(Y)$. Then

$$\nabla_X CT = X\omega(Y),$$

and

$$\begin{aligned}
C(\nabla_X T) &= C(\nabla_X \omega \otimes Y + \omega \otimes \nabla_X Y) \\
&= (\nabla_X \omega)(Y) + \omega(\nabla_X Y).
\end{aligned}$$

Thus we need to show

$$(\nabla_X \omega)(Y) = X\omega(Y) - \omega(\nabla_X Y). \tag{41}$$

To prove (41), let $p \in M$ and $\{v_1, \ldots, v_n\}$ a basis of $T_p(M)$ and $\{\ell^1, \ldots, \ell^n\}$ the corresponding dual basis of $T_p^*(M)$. Let $V_j(t) := \tau_t(v_j)$ so $\{V_1, \ldots, V_n\}$ is a parallel frame along $c$, and define

$$L^i(t) := \tau_{-t}^*(\ell^i).$$

Then

$$L^i(t)(V_j(t)) = \tau_{-t}^*(\ell^i)(\tau_t(v_j)) = \ell^i(v_j) = \delta_j^i,$$

and thus $\{L^1, \ldots, L^n\}$ is the dual coframe to $\{V_1, \ldots, V_n\}$. If $\omega$ is a 1-form defined on a neighborhood of $p$ then we can write

$$\omega(c(t)) = \omega_i(t) L^i(t)$$

for some smooth functions $\omega_i$. If $Y$ is a vector field defined on a neighborhood of $p$ we can write

$$Y(c(t)) = Y^j(t) V_j(t),$$

and so

$$\begin{aligned}
\omega(Y)(c(t)) &= \omega_i(t) L^i(t) \left\{ Y^j(t) V_j(t) \right\} \\
&= \omega_i(t) Y^i(t),
\end{aligned}$$

and so

$$\begin{aligned}
\nabla_X (\omega(Y))(p) &= X\omega(Y)(p) \\
&= \lim_{t\to 0} \frac{\omega_i(t) Y^i(t) - \omega_i(0) Y^i(0)}{t} \\
&= \frac{d\omega_i}{dr}(0) Y^i(0) + \omega_i(0) \frac{dY^i}{dr}(0).
\end{aligned} \tag{42}$$

Next,

$$\begin{aligned}
\nabla_X\omega(p) &= \lim_{t\to 0}\frac{\tau_t^{-1}\left(\omega_i(t)L^i(t)\right)-\omega_i(0)L^i(0)}{t}\\
&= \lim_{t\to 0}\frac{\omega_i(t)\tau_t^{-1}\left(L^i(t)\right)-\omega_i(0)L^i(0)}{t}\\
&= \lim_{t\to 0}\frac{\omega_i(t)-\omega_i(0)}{t}L^i(0)\\
&= \frac{d\omega_i}{dr}(0)L^i(0). \qquad (43)
\end{aligned}$$

We laread know that

$$\nabla_X Y(p)=\frac{dY^i}{dr}(0)V_j(0),$$

by (40), and hence

$$\omega\left(\nabla_X Y\right)(p)=\omega_i(0)\frac{dY^i}{dr}(0). \qquad (44)$$

Putting (42),(43) and (44) together proves (41).

Now we can prove the general case that $\nabla_X$ commutes with all contractions. Without loss of generality, suppose

$$T=\omega^1\otimes\cdots\otimes\omega^k\otimes Y_1\otimes\cdots\otimes Y_\ell$$

and $C=C^i_j$. Thus

$$CT=\omega^i\left(Y_j\right)\omega^1\otimes\cdots\otimes\widehat{\omega^i}\otimes\cdots\otimes\omega^k\otimes Y_1\otimes\cdots\otimes\widehat{Y_j}\otimes\cdots\otimes Y_\ell=:\omega^i\left(Y_j\right)\cdot T'.$$

Thus

$$\begin{aligned}
\nabla_X CT\quad(p) = &\quad \nabla_X\left(\omega^i\left(Y_j\right)\right)(p)\otimes T'(p)+\omega^i\left(Y_j\right)(p)\otimes\nabla_X T'(p)\\
= &\quad X\omega^i\left(Y_j\right)(p)T'(p)+\omega^i\left(Y_j\right)(p)\nabla_X T'(p).
\end{aligned}$$

Next,

$$\begin{aligned}
C\left(\nabla_X T\right)(p) &= C\left(\sum\omega^1\otimes\cdots\otimes\nabla_X(\cdot)\otimes\cdots\otimes Y_\ell\right)(p)\\
&= \omega^i\left(Y_j\right)(p)\nabla_X T'(p)+\omega^i\left(\nabla_X Y_j\right)(p)T'(p)+\left(\nabla_X\omega^i\right)\left(Y_j\right)(p)T'(p),
\end{aligned}$$

and then (41) shows that

$$\nabla_X CT(p)=C\left(\nabla_X T\right)(p)$$

as required.

Now suppose $\alpha:\mathcal{T}(M)\to\mathcal{T}(M)$ satisfies all of these properties. Then

$$\alpha\left(\omega\otimes Y\right)=\alpha\left(\omega\right)\otimes Y+\omega\otimes\nabla_X Y,$$

and hence we must have

$$\alpha(\omega)(Y)=X\omega(Y)-\nabla_X Y,$$

since

$$\begin{aligned}
\alpha\left(\omega(Y)\right) &= \alpha\left(\text{contraction of }\omega\otimes Y\right)\\
&= \text{contraction of }\alpha\left(\omega\otimes Y\right)\\
&= \alpha(\omega)(Y)+\omega\left(\nabla_X Y\right).
\end{aligned}$$

Then the derivation property uniquely determines $\alpha$ on tensors of the form $f\omega^1\otimes\cdots\otimes\omega^k\otimes Y_1\otimes\cdots\otimes Y_\ell$. Since any tensor is a finite sum of terms of this form, this shows that $\alpha$ is uniquely determined by these conditions.

Finally, if reamins to see that $X\mapsto\nabla_X$ is linear over the smooth functions. Since we already know $X\mapsto\nabla_X Y$ and $X\mapsto Xf$ are $C^\infty(M)$-linear, we see that (41) shows that $X\mapsto\nabla_X\omega$ is also $C^\infty(M)$-linear, and hence $X\mapsto\nabla_X T$ is $C^\infty(M)$-linear for any tensor $T$. A similar argument shows that $\nabla_{X+Y}=\nabla_X+\nabla_Y$. This completes the proof. ▶

## 9.11 Definition

Given a Koszul connection $\nabla$ on $M$, define for vector fields $X, Y \in \mathcal{X}(M)$ a vector field $T(X,Y) \in \mathcal{X}(M)$ by

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

Observe that $T$ is clearly bilinear, and for $f \in C^\infty(M)$,

$$\begin{aligned} T(fX,Y) &= \nabla_{fX} Y - \nabla_Y fX - [fX,Y] \\ &= f\nabla_X Y - Yf \cdot X - f\nabla_Y X - f[X,Y] + Yf \cdot X \\ &= fT(X,Y), \end{aligned}$$

and similarly $T(X,fY) = fT(X,Y)$. Thus by Proposition 5.28, $T$ determines a global section (also called) $T$ of the bundle $T^{(1,2)}(M)$, that is, $T \in \mathcal{T}^{(1,2)}(M)$. Thus $T$ is a tensor of type $(1,2)$ which is known as the **torsion tensor** of $\nabla$.

We write the components of $T$ in local coordinates $(x^1, \dots, x^n)$ as $T^k_{ij}$, that is,

$$T(\partial_i, \partial_j) =: T^k_{ij} \partial_k,$$

i.e.

$$T = T^k_{ij} dx^i \otimes dx^j \otimes \partial_k.$$

If $\nabla$ has Christoffel symbols $\{\Gamma^k_{ij}\}$ with respect to the coordinates $(x^i)$ then as $[\partial_i, \partial_j] = 0$ by Proposition 2.6.4 we have

$$T(\partial_i, \partial_j) = \left(\Gamma^k_{ij} - \Gamma^k_{ji}\right) \partial_k$$

and hence

$$T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}. \tag{45}$$

We say that a Koszul connection $\nabla$ on $M$ is **symmetric** or **torsion-free** if its torsion tensor $T$ is identically zero, that is, $T^k_{ij} \equiv 0$ for every coordinate system, or $\Gamma^k_{ij} = \Gamma^k_{ji}$ in every coordinate system (hence the name 'symmetric'). Conversely if $\Gamma^k_{ij} = \Gamma^k_{ji}$ in a set of coordinate systems that cover $M$ then $\nabla$ is symmetric, as the following result shows.

## 9.12 Proposition

Let $M^n$ be a smooth manifold and $\nabla$a Koszul connection on $M$. Let $p \in M$. The torsion tensor $T$ of a connection $\nabla$ satisfies $T(p) = 0$ if and only if there exists a coordinate system $(x^1, \dots, x^n)$ centred at $p$ such that $\Gamma^k_{ij}(p) = 0$ for all $1 \leq i,j,k \leq n$ .

◀ For this proof we will suspend our use of the summation convention - this is because this is unfortunately one of those rare occasions where it is not possible to make the indices sum correctly.

Clearly if we have a coordinate system on which all the $\Gamma^k_{ij}$ vanish at $p$ in that system then certainly $T(p) = 0$. The converse is where the work lies.

Suppose that $(x^1, \dots, x^n)$ are local coordinates defined on a neighborhood $U$ of $p$, such that the $\left(x^i\right)$ are centred at $p$, and such that

$$\Gamma^k_{ij}(p) = \Gamma^k_{ji}(p) \quad \text{for all } 1 \leq i,j,k \leq n \tag{46}$$

(so $T(p) = 0$). Set

$$y^k := x^k + \frac{1}{2} \sum_{i,j} \Gamma^k_{ij}(p) x^i x^j.$$

Then due to (46) we have for $q \in U$,

$$\frac{\partial y^k}{\partial x^\ell}(q) = \delta^k_\ell + \sum_i \Gamma^k_{i\ell}(p) x^i(q), \tag{47}$$

and so in particular

$$\frac{\partial y^k}{\partial x^\ell} = \delta^k_\ell.$$

Thus by the inverse function theorem there exists a neighborhood $V \subseteq U$ of $p$ such that $(y^1, \ldots, y^n)$ forms a coordinate system on $V$, which is clearly centered at $p$. Let $\bar{\Gamma}^k_{ij}$ denote the Christoffel symbols of $\nabla$ with respect to the $(y^i)$. Then

$$\begin{aligned}
\sum_k \bar{\Gamma}^k_{ij} \frac{\partial}{\partial y^k} &= \nabla_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^j} \\
&= \nabla_{\frac{\partial}{\partial y^i}} \left( \sum_\ell \frac{\partial x^\ell}{\partial y^j} \frac{\partial}{\partial x^\ell} \right) \\
&= \sum_\ell \frac{\partial^2 x^\ell}{\partial y^i \partial y^j} \frac{\partial}{\partial x^\ell} + \frac{\partial x^\ell}{\partial y^j} \nabla_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial x^\ell} \\
&= \sum_\ell \frac{\partial^2 x^\ell}{\partial y^i \partial y^j} \frac{\partial}{\partial x^\ell} + \frac{\partial x^\ell}{\partial y^j} \left\{ \sum_m \frac{\partial x^m}{\partial y^i} \nabla_{\frac{\partial}{\partial x^m}} \frac{\partial}{\partial x^\ell} \right\} \\
&= \sum_\ell \frac{\partial^2 x^\ell}{\partial y^i \partial y^j} \frac{\partial}{\partial x^\ell} + \frac{\partial x^\ell}{\partial y^j} \left\{ \sum_{m,h} \frac{\partial x^m}{\partial y^i} \Gamma^h_{m\ell} \frac{\partial}{\partial x^h} \right\} \\
&= \sum_{k,\ell,m,h} \left\{ \frac{\partial^2 x^h}{\partial y^i \partial y^j} + \frac{\partial x^\ell}{\partial y^j} \frac{\partial x^m}{\partial y^i} \Gamma^h_{m\ell} \right\} \frac{\partial y^k}{\partial x^h} \frac{\partial}{\partial y^k},
\end{aligned}$$

and thus for $q \in V$,

$$\bar{\Gamma}^k_{ij}(q) = \sum_{\ell,m,h} \left\{ \frac{\partial^2 x^h}{\partial y^i \partial y^j}(q) + \frac{\partial x^\ell}{\partial y^j}(q) \frac{\partial x^m}{\partial y^i}(q) \Gamma^h_{m\ell}(q) \right\} \frac{\partial y^k}{\partial x^h}(q).$$

In particular evaluating at $p$ gives

$$\bar{\Gamma}^k_{ij}(p) = \frac{\partial^2 x^k}{\partial y^i \partial y^j}(p) + \Gamma^k_{ij}(p). \tag{48}$$

Next, starting from

$$\sum_k \frac{\partial y^\ell}{\partial x^k} \frac{\partial y^k}{\partial y^j} = \delta^\ell_j$$

and differentiating with respect to $y^i$ gives

$$\begin{aligned}
\frac{\partial y^\ell}{\partial x^k} \frac{\partial^2 x^k}{\partial y^i \partial y^j} &= -\frac{\partial}{\partial y^i} \left( \frac{\partial y^\ell}{\partial x^k} \right) \frac{\partial x^k}{\partial y^j} \\
&= -\sum_m \frac{\partial x^m}{\partial y^i} \frac{\partial^2 y^\ell}{\partial x^m \partial x^k} \frac{\partial x^k}{\partial y^j},
\end{aligned}$$

and thus evaluating at $p$ gives

$$\frac{\partial^2 x^k}{\partial y^i \partial y^j}(p) = -\frac{\partial^2 y^k}{\partial x^i \partial x^j}(p).$$

Now from (47),

$$\frac{\partial^2 y^k}{\partial x^i \partial x^j}(p) = \Gamma^k_{ij}(p),$$

and so substituting this into (48) we obtain

$$\bar{\Gamma}^k_{ij}(p) = -\Gamma^k_{ij}(p) + \Gamma^k_{ij}(p) = 0,$$

and thus $(y^1, \ldots, y^n)$ is the a coordinate system centred at $p$ satisfying the requirements of the proposition. ▶

## 9.13 Extending $\nabla$ to $T^{(0,2)}(M) \otimes \mathbf{End}(T(M))$

We have already shown in Theorem 9.10 how to extend a Koszul connection $\nabla$ to an $\mathbb{R}$-linear derivation of the full tensor algebra $\mathcal{T}(M)$. In particular we defined a connection $\nabla$ on $T^{(0,2)}(M)$ as follows: if $A \in \mathcal{T}^{(0,2)}(M)$ and $X, Y, Z \in \mathcal{X}(M)$ we set

$$(\nabla_X A)(Y, Z) = \nabla_X (A(Y, Z)) - A(\nabla_X Y, Z) - A(Y, \nabla_X Z). \tag{49}$$

Suppose now $A \in \Gamma\left(T^{(0,2)}(M) \otimes \mathrm{End}(T(M))\right)$, so $A(Y, Z) \in \mathrm{End}(T(M))$. Then we extend we can still use (49) to define a connection on $T^{(0,2)}(M) \otimes \mathrm{End}(T(M))$, only now both sides are to read as endomorphisms of $T(M)$. In particular, we can consider $\nabla_X R$, where $R$ is the curvature of $\nabla$.

## 9.14 Proposition (Bianchi's identities for symmetric connections)

Let $\nabla$ be a torsion-free connection on $T(M)$, and $R$ its curvature tensor. Then for all $X, Y, Z \in \mathcal{X}(M)$:

1. **(Bianchi's first identity)**

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

In particular, in any coordinate system we have

$$R^i_{jk\ell} + R^i_{k\ell j} + R^i_{\ell jk} = 0 \text{ for all } 1 \le i, j, k, \ell \le n.$$

2. **(Bianchi's second identity)**

$$(\nabla_X R)(Y, Z) + (\nabla_Y R)(Z, X) + (\nabla_Z R)(Y, X) = 0.$$

◄ To prove (1), we first note it is sufficient to verify this for the coordinate vector fields, as both sides are $C^\infty(M)$ linear, and since $[\partial_k, \partial_\ell] = 0$ by Proposition 2.6.4 we use Lemma 8.19 to obtain

$$R(\partial_k, \partial_\ell)(\partial_j) = \left(\nabla_{\partial_k}\nabla_{\partial_\ell} - \nabla_{\partial_\ell}\nabla_{\partial_k}\right)(\partial_j),$$

and since $\nabla$ is symmetric we have $\nabla_{\partial_k}\partial_j = \nabla_{\partial_j}\partial_k$ at thus when we take the cyclic sum of the above equation it vanishes.

To prove (2), since both sides are again $C^\infty(M)$-linear it is enough to verify this pointwise for the coordinate vector fields. So let $p \in M$, and take coordinates $(x^1, \dots, x^n)$ centred about $p$ such that the Christoffel symbols all vanish at $p$ (possible by Proposition 9.12).

Then since the Christoffel symbols vanish at $p$, we have

$$\begin{aligned}
\left\{(\nabla_{\partial_i} R)(\partial_j, \partial_k)\right\}\partial_\ell\big|_p &= \left(\nabla_{\partial_i} R(\partial_j, \partial_k)\right)\partial_\ell\big|_p - R(\nabla_{\partial_i}\partial_j, \partial_k)\partial_\ell\big|_p - R(\partial_j, \nabla_{\partial_i}\partial_k)\partial_\ell\big|_p \\
&= \left(\nabla_{\partial_i} R(\partial_j, \partial_k)\right)\partial_\ell\big|_p + 0 \\
&= \nabla_{\partial_i}\left(R(\partial_j, \partial_k)\partial_\ell\right)\big|_p - R(\partial_j, \partial_k)\nabla_{\partial_i}\partial_\ell\big|_p \\
&= \nabla_{\partial_i}\left(R(\partial_j, \partial_k)\partial_\ell\right)\big|_p + 0 \\
&= \nabla_{\partial_i}\left(R^m_{\ell jk}\partial_m\right)\big|_p \\
&= \left\{\frac{\partial}{\partial x^i}\left(R^m_{\ell jk}\right)(p)\partial_m + R^m_{\ell jk}\left(\nabla_{\partial_i}\partial_m\right)\right\}\big|_p \\
&= \frac{\partial}{\partial x^i}\left(R^m_{\ell jk}\right)(p)\partial_m|_p.
\end{aligned}$$

Thus we have reduced the proof to showing that **in these coordinates** we have

$$\frac{\partial}{\partial x^i}\left(R^m_{\ell jk}\right)(p) + \frac{\partial}{\partial x^j}\left(R^m_{\ell ki}\right)(p) + \frac{\partial}{\partial x^k}\left(R^m_{\ell ij}\right)(p) = 0. \tag{50}$$

To see this last statement, we use the general Bianchi identity (30) proved in Proposition 8.14, that is

$$d\Theta^m_\ell = \Theta^m_j \wedge \theta^j_\ell - \theta^m_j \wedge \Theta^j_\ell,$$

where $\theta$ is the connection matrix of $\nabla$ and $\Theta$ is the curvature matrix of $\nabla$. By assumption $\theta_i^j = \Gamma_{ki}^j dx^k$ is zero at $p$, and hence this identity implies that $d\Theta_\ell^m(p) = 0$ for all $1 \le m, \ell \le n$. Now by (36),

$$\Theta_\ell^m = \frac{1}{2} R_{\ell jk}^m dx^j \wedge dx^k,$$

and hence

$$d\Theta_\ell^m = \frac{1}{2} \frac{\partial}{\partial x_i} \left( R_{\ell jk}^m \right) dx^i \wedge dx^j \wedge dx^k.$$

Equating the coefficients of $dx^i \wedge dx^j \wedge dx^k$ in $d\Theta_\ell^m(p)$ (6 terms in total) we obtain precisely the left-hand side of (50), which thus completes the proof. ▶

# 10 Elementary Riemannian geometry

## 10.1 Definition

A **Riemannian manifold** $(M, g)$ is a smooth manifold $M$ together with a **Riemannian metric** $g$ on $M$, where by definition a Riemannian metric on $M$ is just a metric on the vector bundle $T(M)$. Thus $g \in \mathcal{T}^{(0,2)}(M)$ is symmetric, and if $(x^1, \dots, x^n)$ are local coordinates on $M$ we can write $g$ as

$$g = g_{ij} dx^i \otimes dx^j,$$

where the $g_{ij} : U \to \mathbb{R}$ are smooth, and for any $p \in U$ the matrix $[g_{ij}(p)]$ is positive definite.

We will denote the metric both by $g$ and $\langle \cdot, \cdot \rangle$. Note by Lemma 5.19 any smooth manifold admits a Riemannian metric.

## 10.2 Definition

Let $(M, g)$ be a Riemannian manifold. Given a Koszul connection $\nabla$ on $M$, we have an induced connection $\nabla$ on $T^{(0,2)}(M)$. We say that $\nabla$ is a **metric connection** or $\nabla$ is **compatible** with $g$ if $\nabla_X g \equiv 0$ for all $X \in \mathcal{X}(M)$, that is, if

$$(\nabla_X g)(Y, Z) = \nabla_X(g(X, Y)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0$$

for all $X, Y, Z \in \mathcal{X}(M)$. Equivalently $\nabla$ is a metric connection if and only if for all $X, Y, Z \in \mathcal{X}(M)$ we have

$$X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.$$

## 10.3 Lemma

Let $(M, g)$ be a Riemannian manifold and $\nabla$ a Koszul connection on $M$. Let $c : [a, b] \to M$ be a smooth curve. Then the following are equivalent:

1. For any two vector fields $V, W \in \mathrm{Vect}(c)$

$$\frac{d}{dr} \langle V, W \rangle = \left\langle \frac{DV}{dr}, W \right\rangle + \left\langle V, \frac{DW}{dr} \right\rangle.$$

2. For any $t \in [a, b]$ the parallel transport map $\tau_t : T_{c(a)}(M) \to T_{c(t)}(M)$ is an isometry with respect to $g$.

◀ Suppose (1) holds. Then if $V$ is parallel along $c$ we have

$$\frac{d}{dr} \langle V, V \rangle = 2 \left\langle \frac{DV}{dr}, V \right\rangle = 0,$$

and thus $\langle V, V \rangle$ is constant along $c$. Thus each $\tau_t$ is norm preserving, and hence an isometry.

Conversely suppose $\tau_t$ is always an isometry. Choose parallel vector fields $U_1, \dots, U_n \in \mathrm{Vect}(c)$ that are orthonormal with respect to $g$ at one point of $c$, and hence at every point of $c$. Then given $V, W \in \mathrm{Vect}(c)$ we can write

$$V(t) = V^i(t)U_i(t), \quad W(t) = W^j(t)U_j(t)$$

for some smooth functions $V^i, W^j : [a.b] \to \mathbb{R}$. Then

$$\langle V, W \rangle = \sum_i V^i W^j,$$

and so

$$\frac{d}{dr}\langle V, W \rangle = \sum_i \frac{dV^i}{dr} W^i + V^i \frac{DW^i}{dr}.$$

Next,

$$\frac{DV}{dr} = \frac{dV^i}{dr} U_i, \quad \frac{DW}{dr} = \frac{dW^j}{dr} U_j,$$

by (38), since $\frac{DU_i}{dr} \equiv 0$ as the $U_i$ are parallel, and thus also

$$\left\langle \frac{dV}{dr}, W \right\rangle + \left\langle V, \frac{DW}{dr} \right\rangle = \sum_i \frac{dV^i}{dr} W^i + V^i \frac{DW^i}{dr}. \quad \blacktriangleright$$

## 10.4 Corollary

Let $(M, g)$ be a Riemannian manifold and $\nabla$ a connection on $M$. Then $\nabla$ is a metric connection if and only for any smooth curve $c : [a, b] \to M$ the parallel transport isomorphisms $\tau_t : T_{c(a)}M \to T_{c(t)}M$ are isometries with respect to $g$.

◀ By definition, $\nabla$ is a metric connection if and only if for all vector fields $X, Y, Z \in \mathcal{X}(M)$ and $p \in M$ we have

$$X_p \langle Y, Z \rangle = \left\langle \nabla_{X_p} Y, Z \right\rangle + \left\langle Y, \nabla_{X_p} Z \right\rangle.$$

Now simply apply the previous lemma to a curve $c$ with $\dot{c}(0) = X_p$. ▶

## 10.5 Definition

Given a connection Koszul $\nabla$ on a Riemannian manifold $(M, g)$ we can form an element $R \in \mathcal{T}^{(0,4)}(M)$ defined by

$$R(W, Z, X, Y) := \langle R(X, Y)Z, W \rangle$$

for $X, Y, Z, W \in \mathcal{X}(M)$. It will hopefully not prove confusing that we now use the symbol '$R$' to refer to two different tensors; the curvature tensor $R \in \mathcal{T}^{(1,3)}(M)$ and now a different $R$ residing in $\mathcal{T}^{(0,4)}(M)$. It should be clear from the context to which we are referring to.

In local coordinates $(x^1, \dots, x^n)$ on $M$, we can write

$$R = R_{ijk\ell} dx^k \otimes dx^\ell \otimes dx^j \otimes dx^i,$$

where

$$R_{ijk\ell} := \langle R(\partial_k, \partial_\ell)(\partial_j), \partial_i \rangle = \left\langle R^m_{jk\ell} \partial_m, \partial_i \right\rangle = g_{mi} R^m_{jk\ell}.$$

## 10.6 Theorem (symmetries of $R$)

Let $(M, g)$ be a Riemannian manifold, $\nabla$ a Koszul connection on $M$ and $R$ the $(0, 4)$-curvature tensor. Then for any $X, Y, Z, W \in \mathcal{X}(M)$:

1. $R(W, Z, Y, X) = -R(W, Z, X, Y)$.
2. If $\nabla$ is **metric** then $R(Z, W, X, Y) = -R(W, Z, X, Y)$.

3. If $\nabla$ is **symmetric** then $R(W,Z,X,Y)+R(W,X,Y,Z)+R(W,Y,Z,X)=0$.

4. If $\nabla$ is **both metric and symmetric** then $R(W,Z,X,Y)=R(X,Y,W,Z)$.

◀ (1) is clear since $\Theta$ is alternating and so $R(X,Y)=-R(Y,X)$.
To prove (2) it suffices to show that

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

for all $X,Y,Z\in\mathcal{X}(M)$.

It suffices to check this for the coordinate vector fields in arbitrary local coordinates; then since

$$\begin{aligned} R\left(\partial_i,\partial_i,\partial_j,\partial_k\right) &= \left\langle R\left(\partial_j,\partial_k\right)\partial_i,\partial_i\right\rangle \\ &= \left\langle\left(\nabla_{\partial_j}\nabla_{\partial_k}-\nabla_{\partial_k}\nabla_{\partial_j}\right)\partial_i,\partial_i\right\rangle \end{aligned}$$

by Lemma 8.19 and the fact that $[\partial_j,\partial_k]=0$ by Proposition 2.6.4. But now

$$\begin{aligned} \partial_j\partial_k\left\langle\partial_i,\partial_i\right\rangle &= 2\partial_j\left\langle\nabla_{\partial_k}\partial_i,\partial_i\right\rangle \\ &= 2\left\langle\nabla_{\partial_j}\nabla_{\partial_k}\partial_i,\partial_i\right\rangle+2\left\langle\nabla_{\partial_k}\partial_i,\nabla_{\partial_j}\partial_i\right\rangle, \end{aligned}$$

and hence

$$0=[\partial_j,\partial_k]=2\left\langle\left(\nabla_{\partial_j}\nabla_{\partial_k}-\nabla_{\partial_k}\nabla_{\partial_j}\right)\partial_i,\partial_i\right\rangle.$$

This proves (2). (3) is immediate from Bianchi's first identity for symmetric connections (Proposition 9.14.1).

Property (4) is an algebraic consequence of properties (1), (2) and (3). Indeed,

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

and also

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

and so

$$2R(W,Z,X,Y)=R(W,Y,X,Z)+R(W,X,Z,Y)+R(Z,X,Y,W)+R(Z,Y,W,X).$$

Similarly

$$2R(X,Y,W,Z)=R(X,Z,W,Y)+R(X,W,Y,Z)+R(Y,W,Z,X)+R(Y,Z,X,W).$$

But then

$$\begin{aligned} R(X,Z,W,Y)&=(-1)^2R(Z,X,Y,W),\\ R(X,W,Y,Z)&=(-1)^2R(W,X,Z,Y),\\ R(Y,W,Z,X)&=(-1)^2R(W,Y,X,Z),\\ R(Y,Z,X,W)&=(-1)^2R(Z,Y,W,X). \end{aligned}$$

Thus

$$2R(X,Y,W,Z)=2R(W,Z,X,Y),$$

and this completes the proof. ▶

We have shown that in some sense, metric symmetric Koszul connections are the 'best' type of Koszul connection, in the sense that then the Riemannian curvature tensor posseses the most symmetries. But do such metric symmetric connections exist?

## 10.7 Theorem (the fundamental lemma of Riemannian geometry)

Let $(M, g)$ be a Riemannian manifold. Then there exists a unique metric symmetric Koszul connection $\nabla$ on $M$, called the **Levi-Civita** connection.

◀ First deal with uniqueness. Suppose that $\nabla$ is a symmetric metric connection. Let $X, Y, Z$ be arbitrary smooth vector fields. Compatibility with the metric implies that

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle .$$

That the connection is symmetric implies

$$\langle \nabla_X Y, Z\rangle - \langle \nabla_Y X, Z\rangle = \langle [X, Y], Z\rangle ,$$

and hence we have

$$\begin{aligned} X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle X,Y\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle + \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle - \langle \nabla_Z X, Y\rangle - \langle X, \nabla_Z Y\rangle \\ &= 2\langle \nabla_X Y, Z\rangle - \langle [X,Y], Z\rangle + \langle [X,Z], Y\rangle + \langle [Y,Z], X\rangle , \end{aligned}$$

and hence

$$\langle \nabla_X Y, Z\rangle = \frac{1}{2}\left\{ X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle [Y,Z],X\rangle + \langle [Z,X],Y\rangle + \langle [X,Y],Z\rangle \right\} . \quad (51)$$

This establishes uniqueness.

For existence, we need to show that defining $\nabla_X$ by equation (51) does indeed define a symmetric metric connection. We must verify:

- $\nabla_{fX} Y = f\nabla_X Y$,
- $\nabla_X (fY) = Xf \cdot Y + f\nabla_X Y$,
- $\nabla_X Y - \nabla_Y X = [X, Y]$,
- $\langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle = X\langle Y, Z\rangle$,

the remaining conditions being trivial.

For the first point, observe that

$$\begin{aligned} 2\langle \nabla_{fX} Y, Z\rangle &= fX\langle Y,Z\rangle + Y\langle Z,fX\rangle - Z\langle fX,Y\rangle - \langle [Y,Z],fX\rangle + \langle [Z,fX],Y\rangle + \langle [fX,Y],Z\rangle \\ &= f\left\{ X\langle Y,Z\rangle - Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle [Y,Z],X\rangle + \langle [Z,X],Y\rangle + \langle [X,Y],Z\rangle \right\} \\ &\quad + Yf\cdot\langle Z,X\rangle - Zf\cdot\langle X,Y\rangle + Zf\cdot\langle X,Y\rangle - Yf\cdot\langle X,Z\rangle \\ &= 2f\langle \nabla_X Y, Z\rangle . \end{aligned}$$

To prove the second point, we see that

$$\begin{aligned} 2\langle \nabla_X (fY), Z\rangle &= X\langle fY,Z\rangle + fY\langle Z,X\rangle - Z\langle X,fY\rangle - \langle [fY,Z],X\rangle + \langle [Z,X],fY\rangle + \langle [X,fY],Z\rangle \\ &= f\left\{ X\langle Y,Z\rangle - Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle [Y,Z],X\rangle + \langle [Z,X],Y\rangle + \langle [X,Y],Z\rangle \right\} \\ &\quad + Xf\cdot\langle Y,Z\rangle - Zf\cdot\langle X,Y\rangle + Zf\cdot\langle Y,X\rangle + Xf\cdot\langle Y,Z\rangle \\ &= 2f\langle \nabla_X Y, Z\rangle + 2Xf\cdot\langle Y,Z\rangle . \end{aligned}$$

To prove that $\nabla$ is symmetric we compute

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle - 2\langle \nabla_Y X, Z\rangle &= X\langle Y,Z\rangle - Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle [Y,Z],X\rangle + \langle [Z,X],Y\rangle + \langle [X,Y],Z\rangle \\ &\quad - Y\langle X,Z\rangle + X\langle Z,Y\rangle + Z\langle Y,X\rangle + \langle [X,Z],Y\rangle - \langle [Z,Y],X\rangle - \langle [Y,X],Z\rangle \\ &= -\langle [Y,Z],X\rangle - \langle [X,Z],Y\rangle + \langle [X,Y],Z\rangle + \langle [X,Z],Y\rangle + \langle [Y,Z],X\rangle + \langle [X,Y],Z\rangle \\ &= 2\langle [X,Y],Z\rangle , \end{aligned}$$

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

Finally, to prove that $\nabla$ is compatible with the metric we compute

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle + 2\langle Y, \nabla_X Z\rangle &= X\langle Y,Z\rangle - Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle [Y,Z],X\rangle + \langle [Z,X],Y\rangle + \langle [X,Y],Z\rangle \\ &\quad + X\langle Z,Y\rangle - Z\langle Y,X\rangle - Y\langle X,Z\rangle - \langle [Z,Y],X\rangle + \langle [Y,X],Z\rangle + \langle [X,Z],Y\rangle \\ &= 2X\langle Y,Z\rangle . \end{aligned}$$

This completes the proof of existence. ▶

## 10.8 The Levi-Civita connection in terms of the Christoffel symbols

Let $(x^1, \dots, x^n)$ be local coordinates on a Riemannian manifold $(M, g)$, and $\nabla$ the Levi-Civita connection on $M$. Firstly we compute:

$$2 \langle \nabla_{\partial_i} \partial_j, \partial_\ell \rangle = 2 \left\langle \Gamma^k_{ij} \partial_k, \partial_\ell \right\rangle = 2\Gamma^k_{ij} g_{k\ell}.$$

But then we also have by equation (51) that

$$\begin{aligned} 2 \langle \nabla_{\partial_i} \partial_j, \partial_\ell \rangle &= \partial_i \langle \partial_j, \partial_\ell \rangle + \partial_j \langle \partial_\ell, \partial_i \rangle - \partial_\ell \langle \partial_i, \partial_j \rangle \\ &= \frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{\ell i}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell}, \end{aligned}$$

and thus we see that the Christoffel symbols for the Levi-Civita connection satisfy

$$\Gamma^k_{ij} = \frac{1}{2} g^{k\ell} \left( \frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{\ell i}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right), \tag{52}$$

where $\left[g^{ij}\right] := \left[g_{ij}\right]^{-1}$ is the inverse matrix, so

$$g^{i\ell} g_{\ell j} = \delta^i_j.$$

## 10.9 Definition

If $(M, g)$ is a Riemannian manifold, the **Riemannian curvature tensor** on $M$ is the $(0, 4)$-curvature tensor $R$ with respect to the Levi-Civita connection on $M$.

## 10.10 Definitions

Let $(M, g)$ be a Riemannian manifold and $p \in M$. Given 2 linearly independent tangent vectors $v_1, v_2 \in T_pM$ we define the **sectional curvature** of the 2-plane $\Pi = \operatorname{span}\{v_1, v_2\} \subseteq T_p(M)$ to be

$$K_p(\Pi) := \frac{R(v_1, v_2, v_1, v_2)}{\langle v_1, v_1 \rangle \langle v_2, v_2 \rangle - \langle v_1, v_2 \rangle^2}.$$

Note that this depends only on the 2-plane $\Pi$ and not the choice of basis $\{v_1, v_2\}$, since both $R$ and $g$ are linear and the top and bottom are homogeneous of degree 2. In particular, if $\{e_1, e_2\}$ is are orthonormal, and $\Pi = \operatorname{span}\{e_1, e_2\}$ then $K_p(\Pi) = R(e_1, e_2, e_1, e_2)$. If there exists $K \in \mathbb{R}$ such that $K_p(\Pi) \equiv K$ for all $p \in M$ and $\Pi \subseteq T_p(M)$ then we say $M$ has **constant curvature**.

If $M$ is two-dimensional, then we only have one sectional curvature at a point $p$, and in this case it is written $K(p)$ and called the **Gaussian curvature** of $M$ at $p$.

In fact, the sectional curvature determines the full Riemannian curvature tensor. In order to prove this, we need the following algebraic lemma.

## 10.11 Lemma

Let $V$ be a vector space and $R_1, R_2 : V \times V \times V \times V \to \mathbb{R}$ two quadrilinear maps such that for all $w, x, y, z \in V$ and $i = 1, 2$:

1. $R_i(w, z, y, x) = -R_i(w, z, x, y)$,
2. $R_i(z, w, x, y) = -R_i(w, z, x, y)$,
3. $R_i(w, z, x, y) + R_i(w, x, y, z) + R_i(w, y, z, x) = 0$.
4. $R_i(w, z, x, y) = R_i(x, y, w, z)$.

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 $R_1 \equiv R_2$.

◀ It is clearly sufficient to show that if a multilinear map $R$ satisfying the (1), (2), (3) and (4) also 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

$$\begin{aligned} 0 &= R(w,z,x,y) + R(w,x,y,z) + R(w,y,z,x) \\ &= R(w,z,x,y) - R(w,z,y,x) - R(w,y,x,z) \\ &= 3R(w,z,x,y). \quad ▶ \end{aligned}$$

## 10.12 Corollary

The sectional curvatures determine the full Riemannian curvature tensor.

## 10.13 Corollary

Suppose that $(M, g)$ is a Riemannian manifold and $\nabla$ the Levi-Civita connection on $M$. Suppose there exists $p \in M$ such that

$$K_p(\Pi) \equiv K(p)$$

is independent of the choice of 2-plance $\Pi \subseteq T_p(M)$ (this is necessarily the case for all $p \in M$ if $\dim M = 2$). Then if $R$ is the Riemannian curvature tensor we have for all $W, X, Y, Z \in \mathcal{X}(M)$ that

$$R(W_p, Z_p, X_p, Y_p) = K(p)\left\{\langle W_p, X_p\rangle \langle Z_p, Y_p\rangle - \langle W_p, Z_p\rangle \langle X_p, Y_p\rangle\right\}. \tag{53}$$

◀ Let $R_1(W,X,Y,Z) := K\{\langle W,X\rangle\langle Z,Y\rangle - \langle W,Z\rangle\langle X,Y\rangle\}$. Then it easily seen that $R_1$ is quadrilinear and satisfies properties (1), (2), (3) and (4) of Lemma 10.11. Moreover at $p$ we have $R(X_p, Y_p, X_p, Y_p) = R_1(X_p, Y_p, X_p, Y_p)$ for all $X, Y \in \mathcal{X}(M)$. Hence by Lemma 10.11, $R|_p = R_1|_p$, as required. ▶

In fact, a stronger result holds: if $\dim M \geq 3$ and $K_p(\Pi) \equiv K(p)$ **for all** $p \in M$ then actually $K(p) \equiv K$ is constant, that is, $M$ has constant curvature. This is **Schur's Theorem** and we will prove this shortly; firstly however we will need to deduce the existence of a special coordinate system about any point in $M$.

## 10.14 Definition

Given a metric $g$ on $M$, we say that a local coordinate system $(x^1, \ldots, x^n)$ centred about $p \in M$ is **normal** at $p$ if

$$g_{ij}(p) = \delta_{ij} \quad \text{and} \quad \frac{\partial g_{jk}}{\partial x^i}(p) = 0 \quad \text{for all } 1 \leq i,j,k \leq n.$$

Note that by (52) in normal coordinates at $p$ we also have

$$\Gamma^k_{ij}(p) = 0 \quad \text{for all } 1 \leq i,j,k \leq n,$$

and thus (this will be useful later) (50) holds.

## 10.15 Proposition (normal coordinates)

Let $(M, g)$ be a Riemannian manifold and $p \in M$. Then there exists a neighborhood $U$ of $p$ and coordinates $\left(x^1, \ldots, x^n\right)$ that are normal at $p$.

◀ This is the second time in the course where we will be unable to make the summation convention work, and thus as in the proof of Proposition 9.12 we will explicitly write in the summation signs in this proof. By the Gram-Schmidt process (see Section 5.16) we may assume we have local coordinates $(x^1, \ldots, x^n)$ centred at $p$ that satisfy

$$g_{ij}(p) = \delta_{ij} \quad \text{for all } 1 \le i, j \le n.$$

Now define

$$a_{ijk} := \frac{\partial g_{ij}}{\partial x_k}(p),$$

so

$$a_{ijk} = a_{jik}.$$

Then set

$$b_{ijk} := \frac{1}{2}\left(a_{ijk} + a_{kij} - a_{jki}\right),$$

so

$$b_{ijk} = b_{ikj} \tag{54}$$

and

$$b_{ijk} + b_{jik} = a_{ijk}. \tag{55}$$

Define functions $y_1, \ldots, y_n$ by

$$y^k := x^k + \frac{1}{2}\sum_{h,m} b_{khm} x^h x^m.$$

Then observe that by (54),

$$\frac{\partial y^k}{\partial x^\ell} = \delta^k_\ell + \sum_m b_{k\ell mr} x^m,$$

and in particular

$$\frac{\partial y^k}{\partial x^\ell}(p) = \delta^k_\ell,$$

and hence by the inverse function theorem there exists a neighborhood $V \subseteq U$ of $p$ such that $(y^1, \ldots, y^n)$ form local coordinates on $V$. Set

$$\bar{g}_{ij} := \left\langle \frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j} \right\rangle.$$

Now we perform some messy calculations. First, note that

$$\frac{\partial}{\partial x^\ell} = \sum_k \frac{\partial y^k}{\partial x^\ell}\frac{\partial}{\partial y^k} = \sum_{k,m}\left(\delta^k_\ell + b_{k\ell m}x^m\right)\frac{\partial}{\partial y^k}.$$

Thus

$$\begin{aligned} g_{\ell m} &= \left\langle \frac{\partial}{\partial x_\ell}, \frac{\partial}{\partial x_m} \right\rangle \\ &= \left\langle \sum_{i,k}\left(\delta^i_\ell + b_{i\ell k}x^k\right)\frac{\partial}{\partial y^i}, \sum_{j,h}\left(\delta^j_m + b_{jmh}x^h\right)\frac{\partial}{\partial y^j} \right\rangle \\ &= \sum_{i,j,h,k} \bar{g}_{ij}\left(\delta^i_\ell + b_{i\ell k}x^k\right)\left(\delta^j_m + b_{jmh}x^h\right), \end{aligned}$$

and hence evaluating at $p$ gives

$$\delta_{\ell m} = g_{\ell m}(p) = \sum_{i,j} \bar{g}_{ij}(p)\delta^i_\ell \delta^j_m,$$

and hence

$$\bar{g}_{ij}(p) = \delta_{ij}.$$

Next,

$$\frac{\partial g_{\ell m}}{\partial x^r} = \sum_s \frac{\partial y^s}{\partial x^r}\frac{\partial}{\partial y^s}\left\{\sum_{i,j,h,k} \bar{g}_{ij}\left(\delta^i_\ell + b_{i\ell k}x^k\right)\left(\delta^j_m + b_{jmh}x^h\right)\right\}.$$

Now note that

$$\frac{\partial}{\partial y^s}\left(\delta^i_\ell + b_{i\ell k}x^k\right) = b_{i\ell k}\frac{\partial x^k}{\partial y^s},$$

and thus evaluating at $p$ gives

$$\frac{\partial}{\partial y^s}\left(\delta^i_\ell + b_{i\ell k}x^k\right)(p) = b_{i\ell s}.$$

Hence when we evaluate $\frac{\partial g_{\ell m}}{\partial x^r}$ at $p$ we obtain

$$\begin{aligned} a_{\ell m r} &= \sum_{i,j,s} \delta^s_r \left(\frac{\partial \bar{g}_{ij}}{\partial y^s}(p)\delta^i_\ell\delta^j_m + \delta_{ij}b_{i\ell s}\delta^j_m + \delta_{ij}\delta^i_\ell b_{jms}\right) \\ &= \frac{\partial \bar{g}_{\ell m}}{\partial y^r} + b_{m\ell r} + b_{\ell m r}, \end{aligned}$$

and thus

$$\frac{\partial \bar{g}_{\ell m}}{\partial y^r} = a_{\ell m r} - b_{m\ell r} - b_{\ell m r} = 0,$$

by (55). This completes the proof. ▶

## 10.16 The second Bianchi identity for the Riemannian curvature tensor $R$

Observe that if $(x^1, \ldots, x^n)$ are normal coordinates at $p \in M$ then

$$\begin{aligned} \frac{\partial}{\partial x^i}\left(R_{m\ell jk}\right)(p) &= \frac{\partial}{\partial x^i}\left(g_{mh}R^h_{\ell jk}\right)(p) \\ &= \frac{\partial g_{mh}}{\partial x^i}(p)R^h_{\ell jk}(p) + g_{mh}(p)\frac{\partial}{\partial x^i}\left(R^h_{\ell jk}\right)(p) \\ &= \frac{\partial}{\partial x_i}\left(R^m_{\ell jk}\right)(p). \end{aligned}$$

Thus by (50) we deduce that in **normal coordinates** at $p$ we have

$$\frac{\partial}{\partial x^i}\left(R_{m\ell jk}\right)(p) + \frac{\partial}{\partial x^j}\left(R_{m\ell ki}\right)(p) + \frac{\partial}{\partial x^k}\left(R_{m\ell ij}\right)(p) = 0. \quad (56)$$

## 10.17 Theorem (Schur)

Let $(M, g)$ be a connected Riemannian manifold of dimension $n \geq 3$. Then if the sectional curvatures of $M$ are pointwise constant, that is, $K_p(\Pi) = f(p)$ for all 2-planes $\Pi \subseteq T_p(M)$, where $f : M \to \mathbb{R}$, then $f$ is constant.

◀ Let $p \in M$, and $(x^1, \ldots, x^n)$ be normal coordinates on a neighborhood $U$ of $p$. Then by (53) we can write

$$R(W_p, Z_p, X_p, Y_p) = K(p)\left\{\langle W_p, X_p\rangle \langle Z_p, Y_p\rangle - \langle W_p, Z_p\rangle \langle X_p, Y_p\rangle\right\}$$

for any smooth vector fields $W, X, Y, Z$ on $U$, and hence on $U$ we have

$$R_{ijk\ell}(p) = f(p)\left(g_{ik}g_{j\ell} - g_{i\ell}g_{jk}\right).$$

But by (56),

$$\frac{\partial}{\partial x^h}(R_{ijk\ell})(p) + \frac{\partial}{\partial x^k}(R_{ij\ell h})(p) + \frac{\partial}{\partial x^\ell}(R_{ijhk})(p) = 0,$$

and hence

$$\frac{\partial f}{\partial x^h}(p)\left(\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}\right) + \frac{\partial f}{\partial x^k}(p)\left(\delta_{i\ell}\delta_{jh} - \delta_{ih}\delta_{j\ell}\right) + \frac{\partial f}{\partial x^\ell}(p)\left(\delta_{ih}\delta_{jk} - \delta_{ik}\delta_{jh}\right) = 0.$$

Since $n \geq 3$, given $h$ we can find $i, j$ such that $i, j, h$ are all distinct. Setting $k = i, \ell = j$ it then follows from the above that $\frac{\partial f}{\partial x^h}(p) = 0$. Since $h$ was arbitrary, it follows $df_p = 0$. Thus $f$ is locally constant. Since $M$ is connected, $f$ is constant. ▶

## 10.18 Example

We claim that $S^n$ has constant sectional curvature, when equipped with the metric induced from the ambient space $\mathbb{R}^{n+1}$ (that is, for given $p \in S^n$, the metric on $T_p(S^n) \subseteq T_p\left(\mathbb{R}^{n+1}\right) \cong \mathbb{R}^{n+1}$ is just the restriction of the dot product on $\mathbb{R}^{n+1}$).

First, claim that the group of orientation preserving isometries of $S^n$, $SO(n+1)$ operates transitively on the set of 2-planes in $T(S^n)$. To see this, it is enough to show that given $p \in S^n$, if $H$ is the stabiliser of $p$ in $SO(n+1)$ then we can take any 2-plane in $T_p(S^n)$ to any other. To check this, since $SO(n+1)$ is certainly transitive on $S^n$, we may take $p = (1, 0, \ldots, 0)$. The stabiliser of $p$ is then

$$H = \left\{ \left( \begin{array}{cc} 1 & 0 \\ 0 & A \end{array} \right) \middle| A \in SO(n) \right\}.$$

With respect to the Euclidean metric, $T_p(S^n)$ is orthogonal to $p$, and $H$ operates on $T_p(S^n)$ by $v \mapsto Av$. This action is certainly transitive on the 2-planes in $T_p(S^n)$.

It is easy to see that sectional curvatures are preserved by isometries, whence it follows $S^n$ has constant sectional curvature as claimed.

We conclude the our discussion on sectional curvature with a theorem we won't prove.

## 10.19 Theorem

Any simply connected complete Riemannian manifold $(M, g)$ with constant sectional curvature $\kappa$ is diffeomorphic to one of the following three manifolds, where $|\kappa| = \frac{1}{r^2}$:

1. if $\kappa = 0$, $M \cong \mathbb{R}^n$,
2. if $\kappa > 0$, $M \cong S^n(r) = \left\{x \in \mathbb{R}^{n+1} \mid \|x\| = r\right\}$,
3. if $\kappa < 0$, $M \cong H^n(r)$, where $H^n(r)$ denotes the **hyperbolic space**.

## 10.20 Definition

Let $(M, g)$ be a Riemannian manifold. Define a $(0, 2)$-tensor $\mathrm{Ric}(X, Y)$ defined by

$$\mathrm{Ric}(X, Y) := \mathrm{tr}(v \mapsto R(v, X)(Y)).$$

We call Ric the **Ricci tensor** of $g$.

Take an orthonormal basis $\{e_1, \ldots, e_n\}$ of $T_p(M)$. Then

$$\begin{aligned} \mathrm{Ric}(X_p, Y_p) &= \mathrm{tr}\left(v \mapsto R(v, X_p)\, Y_p\right) \\ &= \sum_{i=1}^n \langle R(e_i, Y_p)\, e_i, X_p \rangle = \sum_{i=1}^n R(e_i, Y_p, e_i, X_p) \\ &= \sum_{i=1}^n R(e_i, X_p, e_i, Y_p). \end{aligned}$$

In particular if $(x^1, \ldots, x^n)$ are local normal coordinates about $p \in M$ then if we work with the orthonormal basis $\{\partial_i|_p\}$ of $T_p(M)$ and

$$r_{kl} := \mathrm{Ric}\,(\partial_k, \partial_\ell)$$

then

$$\mathrm{Ric}\,(X, Y) = \sum_i \{\partial_i\text{-coefficient of } (\partial_i \mapsto \mathrm{Ric}\,(\partial_i, X)\, Y)\},$$

and so

$$r_{k\ell} = R^i_{\ell ik} = g^{ij} R_{j\ell ik}.$$

Ric is symmetric since

$$r_{k\ell} = g^{ij} R_{j\ell ik} = g^{ij} R_{ikj\ell} = g^{ji} R_{jki\ell} = r_{\ell k}.$$

Both $g$ and Ric are elements of $\mathcal{T}^{(0,2)}(M)$. A metric $g$ is called **Einstein** if $r = \lambda g$ for some constant $\lambda$.

## 10.21 Definition

For any non-zero $v \in T_p\,(M)$ the **Ricci curvature in the direction** $v$ is defined by

$$\mathrm{Ric}(v) := \frac{\mathrm{Ric}(v, v)}{\|v\|^2}.$$

If $\|v\| = 1$ then $r(v) = r(v, v)$. Moreover if $\|v\| = 1$ we may extend $\{v\}$ to an orthonormal basis $\{e_1 = v, e_2, \ldots, e_n\}$ of $T_p\,(M)$. Then

$$r(v) = \sum_{i=1}^n R(e_i, v, e_i, v) = \sum_{i=2}^n R(e_i, v, e_i, v),$$

since $R(e_1, e_1, e_1, e_1) = 0$, and thus $\frac{\mathrm{Ric}(v)}{n-1}$ is an average of sectional curvatures $K_p\,(\Pi_i)$ where

$$\Pi_i = \mathrm{span}\,\{v, e_i\}\,, \quad i \geq 2.$$

## 10.22 Lemma

The Ricci curvatures at $p$ are all equal to a constant (say $\lambda$) if and only if Ric $= \lambda g$ at $p$ ($g$ is Einstein 'at $p$').

◀ One way is clear. For the converse, we simply note that the Ricci curvatures all being equal to $\lambda$ imply that for any non-zero $v \in T_p\,(M)$ we have $\mathrm{Ric}(v, v) = \lambda \langle v, v\rangle$. Since $\mathrm{Ric}(\cdot, \cdot)$ is a symmetric bilinear form the **polarization identity** gives

$$\begin{aligned} 2\mathrm{Ric}(v, w) &= r(v + w, v + w) - r(v, v) - r(w, w) \\ &= \lambda \langle v + w, v + w\rangle - \lambda \langle v, w\rangle - \lambda \langle w, w\rangle \\ &= \lambda \langle v, w\rangle\,, \end{aligned}$$

since the polarisation identity also applies to the symmetric bilinear form $\langle \cdot, \cdot \rangle$. ▶

## 10.23 Definition

The Ricci tensor $r$ and the metric together determine another endomorphism $\theta : T_p\,(M) \to T_p\,(M)$ defined

$$\mathrm{Ric}(v, \cdot) = \langle \theta(v), \cdot \rangle\,.$$

The **scalar curvature** $\mathrm{scal}(p)$ at $p$ is the trace of $\theta$. If $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $T_p\,(M)$ then we have

$$\mathrm{Ric}(e_i) = \mathrm{Ric}(e_i, e_i) = \langle \theta(e_i), e_i\rangle = \theta_{ii},$$

and so

$$\begin{aligned}\mathrm{tr}(\theta) &= \sum_{i=1}^{n} \theta_i^i \\ &= \sum_{i=1}^{n} \langle \theta(e_i), e_i \rangle \\ &= \sum_{i=1}^{n} \mathrm{Ric}(e_i).\end{aligned}$$

Thus $\frac{\mathrm{scal}(p)}{n}$ is an average of Ricci curvatures.

If $(x^1, \ldots, x^n)$ are normal coordinates about $p$ then writing

$$\theta(\partial_i) = \theta_i^j \partial_j,$$

we have

$$g_{j\ell}\theta_i^j = \langle \theta(\partial_i), \partial_\ell \rangle = \mathrm{Ric}(\partial_i, \partial_\ell) = r_{i\ell},$$

and hence

$$\theta_i^j = g^{j\ell} r_{i\ell}$$

and so

$$\begin{aligned}\mathrm{scal}(p) &= \sum_i \theta_i^i(p) \\ &= g^{i\ell}(p) r_{i\ell}(p) \\ &= g^{i\ell}(p) g^{jk}(p) R_{k\ell ji}(p).\end{aligned}$$

The following theorem is in a similar vein to Schur's Theorem 10.17.

## 10.24 Theorem

Let $(M, g)$ be a connected Riemannian manifold of dimension $n \geq 3$. Then if the Ricci curvatures of $M$ are pointwise constant, that is, $\mathrm{Ric}(v) = \lambda(p)$ for all $v \neq 0 \in T_p(M)$, where $\lambda : M \to \mathbb{R}$, then $\lambda$ is constant, and so $M$ is Einstein.

◀ By Lemma 10.22 and the assumption if $(x^1, \ldots, x^n)$ are normal coordinates about $p$ then

$$r_{ij}(p) = \lambda(p) g_{ij}(p).$$

In what follows, everything is to be evaluated at $p$; 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 $m \in \{1, \ldots, n\}$. Then by (56),

$$\frac{\partial}{\partial x^m}(R_{hjhi}) + \frac{\partial}{\partial x^i}(R_{hjmh}) + \frac{\partial}{\partial x^h}(R_{hjim}) = 0. \tag{57}$$

Using $r_{ij} = \lambda g_{ij}$ we obtain

$$\begin{aligned}\delta_{ij}\frac{\partial \lambda}{\partial x^m} &= \frac{\partial r_{ij}}{\partial x^m} \\ &= \frac{\partial}{\partial x^m}\left(\sum_{k,\ell} g^{k\ell} R_{kj\ell i}\right) \\ &= \sum_{k,\ell} \delta^{k\ell} \frac{\partial}{\partial x^m}(R_{kj\ell i}),\end{aligned}$$

and hence for **any** $1 \le i \le n$,

$$\frac{\partial \lambda}{\partial x^m} = \sum_h \frac{\partial}{\partial x^m}\left(R_{hihi}\right). \tag{58}$$

Thus setting $i = j$ in (57)and substituting we have

$$\frac{\partial \lambda}{\partial x^m} + \frac{\partial}{\partial x^i}\left(R_{himh}\right) + \frac{\partial}{\partial x^h}\left(R_{hiim}\right) = 0,$$

and so summing both sides over $i$,

$$n\frac{\partial \lambda}{\partial x^m} = \sum_h \sum_{i \neq m} \frac{\partial}{\partial x^i}\left(R_{hihn}\right) + \sum_h \frac{\partial}{\partial x^m}\left(R_{hmhm}\right) + \sum_{h \neq m} \sum_i \frac{\partial}{\partial x^h}\left(R_{ihmi}\right) + \sum_i \frac{\partial}{\partial x^m}\left(R_{imim}\right). \tag{59}$$

Now

$$\sum_h \sum_{i \neq m} \frac{\partial}{\partial x^i}\left(R_{hihn}\right) = -\sum_{h \neq m} \sum_i \frac{\partial}{\partial x^h}\left(R_{ihmi}\right),$$

and since (58) holds for any $i$, (59) becomes

$$n\frac{\partial \lambda}{\partial x^m} = 2\frac{\partial \lambda}{\partial x^m},$$

and hence

$$(n-2)\frac{\partial \lambda}{\partial x^m} = 0.$$

Since $m$ was arbitrary we conclude $d\lambda_p = 0$; thus $\lambda$ is locally constant. Since $M$ is connected, $\lambda$ is constant. ►

We will conclude the course with a short discussion of how to define a metric on a Riemannian manifold.

## 10.25 Definitions

Let $(M, g)$ be a connected Riemannian manifold. Given a piecewise smooth curve $c : [a, b] \to M$, we define the **length** of $c$ to be

$$\ell(c) := \int_a^b \langle \dot{c}(t), \dot{c}(t) \rangle \, dt.$$

If the image of $c$ is contained in a single chart $(U, h)$ with coordinates $(x^1, \dots, x^n)$, and $g_{ij}$ are the coefficients of $g$ with respect to this chart then we can write

$$\ell(c) := \int_a^b \langle \dot{c}(t), \dot{c}(t) \rangle \, dt = \int_a^b \sqrt{g_{ij}(h(c(t))\frac{d}{dt}\left\{x^i(c(t) \cdot x^j(c(t)\right\}} dt.$$

Given $p, q \in M$, we define the **distance** between $p$ and $q$ to be

$$d(p, q) := \inf\left\{\ell(c) \mid c : [a, b] \to M \text{ piecewise smooth, } c(a) = p, c(b) = q\right\}. \tag{60}$$

We shall shortly prove that this defines a metric on $M$. First however we check that $d(p, q)$ is always finite.

## 10.26 Lemma

Let $(M, g)$ be a connected Riemannian manifold. Then given any two $p, q \in M$, there exists a piecewise smooth curve from $c$ from $p$ to $q$.

◄ Fix $p \in M$. Set

$$U_p := \{q \in M \mid \text{there exists a piecewise smooth curve } c \text{ from } p \text{ to } q\}.$$

Using local coordinates, it is easy to see that $U_p$ is open. Similarly $M \backslash U_p$ is open. Since $U_p \neq \emptyset$ as $p \in M$ and $M$ is connected, it follows $U_p = M$ as claimed. ►

## 10.27 Proposition

Let $(M, g)$ be a connected Riemannian manifold. Then the function $d : M \times M \to \mathbb{R}$ defined in (60) is a metric on $M$.

◀ The only property of a metric that it is not immediate $d$ satisfies is $d(p, q) > 0$ for $p \neq q$. To check this, let $p \neq q \in M$ and $(U, h)$ a chart about $p$. Then there exists $\epsilon > 0$ such that

$$q \notin V := h^{-1}\left(\overline{B_\epsilon(h(p))}\right).$$

Let $(x^1, \ldots, x^n)$ be the local coordinates of $h$, and $g_{ij}$ the coordinates of $g$ with respect to the $x^i$. Since $[g_{ij} \circ h(x)]$ is a positive definite smooth (in $x$) matrix and $\overline{B_\epsilon(h(p))} \subseteq \mathbb{R}^n$ is compact, there exists $\lambda > 0$ such that for any $\xi \in \mathbb{R}^n$ and any $x \in V$ we have

$$g_{ij}(h(x))\xi^i\xi^j \geq \lambda|\xi|^2.$$

Thus for any curve piecewise smooth curve $c : [a, b] \to M$ with $c(a) = p, c(b) = q$, we have

$$\ell(c) \geq \ell\left(c|_{c^{-1}(V \cap c([a,b]))}\right) \geq \lambda\epsilon,$$

since as $q \notin V$, there exists $y \in \partial B_\epsilon(h(p))$ such that $h^{-1}(y) = c(d)$ for some $d \in (a, b]$, and then

$$\ell(c) = \int_a^b \langle \dot{c}(t), \dot{c}(t)\rangle\, dt \geq \int_a^d \sqrt{g_{ij}(h(c(t))\frac{d}{dt}\{x^i(c(t) \cdot x^j(c(t)\}}dt. \geq \lambda\epsilon.$$

This completes the proof. ▶

## 10.28 Definition

Let $(M, g)$ be a connected Riemannian manifold. We say that $M$ is **complete** if $M$ is complete as a metric space under the metric $d$ of Proposition 10.27.