MANIFOLD GEOMETRY AND ANALYSIS

NOTES BY YOUSUF SOLIMAN

ABSTRACT. These notes were compiled for me to get a better grasp of geometry and analysis on manifolds. I decided to work on these notes to explore and expand on the material taught in *Differential Geometry (21-759)* by Dejan Slepčev at Carnegie Mellon University. The first chapter roughly covers most of the material taught in 21-759 at Carnegie Mellon University (with more emphasis on topics I was interested in). Many of the figures here come from Keenan Crane's discrete differential geometry notes. I make no claim that any of the material here is original. If you find errors please let me know!

CONTENTS

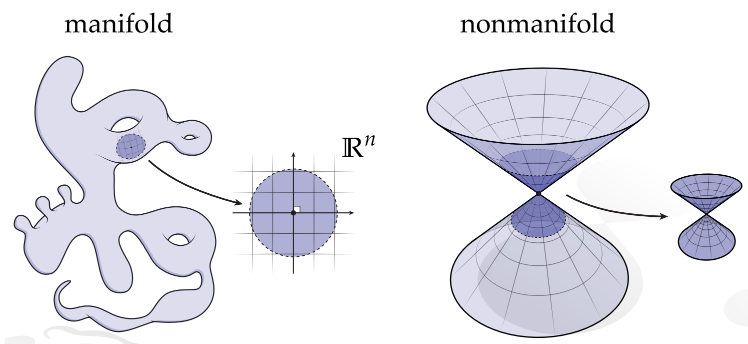
E-mail address: `ysoliman@cmu.edu`.
Date: November 1, 2017.

CHAPTER 1

Basic Riemannian Geometry

1. Manifolds.

Intuitively, an n-dimensional manifold is a topological space that locally looks like $\mathbf{R}^n$.



The formal definition is the following:

Definition 1.1. An n-dimensional **manifold** is a second countable Hausdorff topological space M endowed with a maximal atlas $\mathscr{A}$. $\diamond$

In the next sections I will explain what this definition means, why this is a sensible definition, and provide some examples.

1.1. Charts and atlases.

Let M be a topological space.

Definition 1.2. A **chart** is a homeomorphism $\varphi : U \to \widetilde{U}$ from an open subset $U \subseteq M$ to an open subset $\widetilde{U} \subseteq \mathbf{R}^n$. $\diamond$

Remark 1.3. Note that the components $x_1, \ldots, x_n$ of φ define a **local coordinate system** on M. $\diamond$

Definition 1.4. Two charts $\varphi_1 : U_1 \to \widetilde{U}_1$ and $\varphi_2 : U_2 \to \widetilde{U}_2$ are called **compatible** if the **transition maps**

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2) \quad \text{and}$$
$$\varphi_1 \circ \varphi_2^{-1} : \varphi_2(U_1 \cap U_2) \to \varphi_1(U_1 \cap U_2)$$

are smooth. $\diamond$

Definition 1.5. An **atlas** $\mathscr{A}$ on M is a collection

$$\mathscr{A} = \{\varphi_i : U_i \to \widetilde{U}_i \ : \ i \in I\}$$

of pairwise compatible charts such that

$$M = \bigcup_{i \in I} U_i.$$

$\diamond$

Definition 1.6. An atlas $\mathscr{A}$ is called **maximal** if for every atlas $\mathscr{B}$ with $\mathscr{A} \subset \mathscr{B}$ it holds that $\mathscr{A} = \mathscr{B}$. $\diamond$

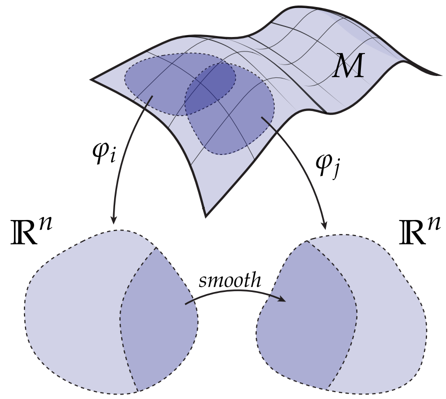


FIGURE 1. The key idea that will show up time and time again is that we can leverage the notion of differentiability in $\mathbf{R}^n$ to talk about differentiability on manifolds.

The reason why we would like to only work with maximal atlases is that we would like to rule out the possibility of creating a different manifold by (trivially) enlarging the atlas. Often a maximal atlas is referred to as a smooth structure on M. *A smooth structure dictates precisely how the locally Euclidean geometry of the smooth manifold behaves!*

Theorem 1.7. *Every atlas $\mathscr{A}$ is contained in a unique maximal atlas $\widetilde{\mathscr{A}}$.*

Proof. Consider the family

$$\mathscr{F} := \{\mathscr{B} - \text{an atlas} : \mathscr{A} \subseteq \mathscr{B}\}.$$

Note that $\mathscr{F}$ is partially ordered by inclusion. We claim that $\mathscr{F}$ is a chain-complete poset. Consider any chain $\{\mathscr{A}_\alpha\}_{\alpha\in\Lambda}$, and note that

$$\bigcup_{\alpha\in\Lambda} \mathscr{A}_\alpha$$

is also an atlas. So we have that the hypothesis of Zorn's lemma are satisfied and so there is some maximal atlas $\widetilde{\mathscr{A}}$ containing $\mathscr{A}$.

Now we prove uniqueness. Suppose that $\mathscr{A}^*$ is another maximal atlas that contains $\mathscr{A}$. So we have that $\mathscr{A}^* \subseteq \mathscr{A}^* \cup \widetilde{\mathscr{A}}$ which implies that $\mathscr{A}^* = \mathscr{A}^* \cup \widetilde{\mathscr{A}}$ by the maximality of $\mathscr{A}^*$ and so we have that $\widetilde{\mathscr{A}} \subseteq \mathscr{A}^*$ and so we have that $\widetilde{\mathscr{A}} = \mathscr{A}^*$ by the maximality of $\widetilde{\mathscr{A}}$. □

Given the result of this theorem we can specify a manifold simply as a pair $(M, \mathscr{A})$ where M is a second countable Hausdorff topological space and $\mathscr{A}$ is an atlas on M. In this representation of a manifold the atlas $\mathscr{A}$ represents the unique maximal atlas containing $\mathscr{A}$.

Example 1.8. We will show that we can endow $\mathbf{R}$ with two distinct smooth structures. Consider the two atlases:

$$\mathscr{A} := \{\mathrm{id} : \mathbf{R} \to \mathbf{R}\} \quad \text{and}$$
$$\mathscr{B} := \{x \mapsto x^3\}.$$

Now we will show that $\mathscr{A}$ and $\mathscr{B}$ define two different manifolds on $\mathbf{R}$. Namely, we will show that $\mathscr{A}$ and $\mathscr{B}$ are not compatible. Let $\varphi(x) = x$ and $\psi(y) = y^3$, notice that $\varphi \in \mathscr{A}$ and $\psi \in \mathscr{B}$ and we have that

$$\varphi \circ \psi^{-1}(x) = \sqrt[3]{x}.$$

So we have that the transition maps are *not* smooth at 0. Hence, we have shown that $\mathbf{R}$ has two distinct smooth structures. Note that $(\mathbf{R}, \mathscr{A})$ is diffeomorphic to $(\mathbf{R}, \mathscr{B})$ – just not via the identity! ⋄

1.2. **Topological assumptions on M.** The topological assumptions we make in **??** are used to rule out certain undesirable pathologies. To understand why exactly these are undesirable it is useful to take the standpoint that a manifold should be determined by its smooth functions.

Example 1.9 (**R** with a double origin). Consider

$$M := \mathbf{R} \times \{0\} \cup \mathbf{R} \times \{1\} / \sim$$

with $(x, 0) \sim (x, 1)$ whenever $x \neq 0$. This is a real line with two origins $o = [(0, 0)]$ and $o^* = [(0, 1)]$ and there is no smooth (or even continuous) function which can distinguish between o and o^*. ⋄

Example 1.10 (Long line). The long line L is obtained by sticking together uncountably many copies of $[0, 1)$.

Formally, the construction is as follows: Let Ω denote the smallest uncountable well-ordered set and set $L := \Omega \times [0, 1)$. The lexicographical order $\prec$ makes L into a totally ordered set; and hence, a topological space via the order topology. The order topology is the topology generated by the basis of open intervals

$$(x, z) := \{y \in L \,:\, x \prec y \prec z\}.$$

Now we can endow L with an atlas with charts $\varphi : \{\omega\} \times [0, 1) \cup \{\omega'\} \times (0, 1) \to (-1, 1)$ defined via

$$\varphi(\omega, t) = t - 1 \quad \text{and} \quad \varphi(\omega', t) = 1.$$

In the above ω' is the successor of ω.

L has a number of undesirable properties. The most undesirable of them all is that it does not admit a Riemannian metric. ⋄

These topological assumptions not only rule out undesirable pathologies, but also provide us with some of the technical machinery necessary to study the geometry of manifolds (such as allowing us to construct partitions of unity). For example, the Hausdorff assumption ensures that limits of sequences are unique and that compact sets are closed. Now we provide a few technical results that will be key in our analysis.

Theorem 1.11. *A smooth manifold is paracompact.*

Proof. Let M be a smooth manifold as defined in 4. First we show that M admits a compact exhaustion, that is an increasing sequence of compact sets $\{K_i\}_{i\in\mathbf{N}}$ such that

$$M = \bigcup_{i\in\mathbf{N}} K_i$$

and

$$K_i \subset \operatorname{int} K_{i+1}$$

for all $i \in \mathbf{N}$.

Since M is second countable we can find a countable base $\mathscr{B} = \{V_i\}_{i\in\mathbf{N}}$ of the topology on M. Note that without loss of generality we may assume that each V_i has compact closure since M is locally homeomorphic to Euclidean space, and therefore locally compact. So we may restrict to those V_i that have compact closure. Now define

$$K_1 = \overline{V_1}.$$

We will construct the sequence $\{K_n\}$ inductively. Suppose that K_n has been defined. Let i_n be the smallest integer such that

$$K_n \subset V_1 \cup \cdots \cup V_{i_n}.$$

Now define

$$K_{n+1} := \overline{V_1} \cup \cdots \cup \overline{V_{i_n}}.$$

It is easy to see that the sequence $\{K_i\}_{i\in\mathbf{N}}$ is a compact exhaustion of M.

Now we show that M is paracompact. Consider a compact exhaustion $\{K_i\}_{i\in\mathbf{N}}$ of M. For all $i \in \mathbf{N}$ define $U_i := \operatorname{int} K_i$. Note that $\overline{U_i} \subseteq K_i$ is compact for every $i \in \mathbf{N}$ and the U_i form an open cover of M with the property that $\overline{U_i} \subset U_{i+1}$. Now consider any open cover $\{V_\alpha\}_{\alpha\in\Lambda}$ of M. Note that $\overline{U_i} \setminus U_{i-1}$ is compact for all $i \in \mathbf{N}$. For all $i \geq 3$ we have that $\overline{U_i} \setminus U_{i-1}$ is covered by the open sets

$$V_{\alpha,i} := V_\alpha \cap U_{i+1} \cap \left(\overline{U}_{i-2}\right)^{\mathsf{c}}.$$

Now we can find a finite subcover $\mathscr{V}_i$ since the set is compact. Similarly, for $i = 1, 2$ we can let

$$V_{\alpha,i} := V_\alpha \cap U_3.$$

So again we can find a finite subcover $\mathscr{V}_i$.

Now we claim that

$$\bigcup_{i \in \mathbf{N}} \mathscr{V}_i$$

is a locally finite refinement of $\{V_\alpha\}$. It is clearly a refinement since $V_{\alpha,i} \subset V_\alpha$ for all $\alpha \in \Lambda$ and $i \in \mathbf{N}$. To see that it is locally finite consider any point $p \in M$. Note that there is a smallest $i \in \mathbf{N}$ such that $p \in U_i$ and $p \notin U_{i-1}$. So we have that $p \in U_i \setminus U_{i-1}$ for some $i \in \mathbf{N}$. Now we can consider the open set $O := U_i \setminus \overline{U_{i-2}}$. By construction we have that the only sets that intersect O are elements of $\mathscr{V}_{i+1}, \mathscr{V}_i$, and $\mathscr{V}_{i-1}$, which are all finite. Lastly, we have that the collection is a cover by construction.

Since $\{V_\alpha\}$ was an arbitrary cover we have that any cover admits a locally finite, countable refinement, and so we have that M is paracompact as desired. ☐

Now we come to one of the fundamental lemmas needed – a smooth version of Urysohn's lemma.

Lemma 1.12 (Smooth Urysohn's Lemma)**.** *If M is a smooth manifold and $C_0, C_1 \subset M$ are disjoint closed sets then there exists a smooth function $f : M \to [0,1]$ such that $C_0 = f^{-1}(0)$ and $C_1 = f^{-1}(1)$.*

Proof. The proof is almost identical to the standard proof of Urysohn's lemma, but since we have the language of smooth manifolds it can easily be modified to show that the functions are smooth. The proof is left as a simple exercise to the reader. ☐

Now we come to the construction of partitions of unity, which we will need.

Lemma 1.13. *Let M be a smooth manifold. Any countable locally finite covering $\{U_i : i \in I\}$, where I is an (at most countable) index set, of open sets has a partition of unity subordinate to this covering.*

Proof. For every $i \in I$ we can use the smooth Urysohn lemma to find functions $\lambda_i : M \to [0,1]$ such that $\lambda_i^{-1}(0) = M \setminus U_i$. As the cover is locally finite, we have that the sum $\sum_{i \in I} \lambda_i$ is well defined. Furthermore, it is always positive as $\{U_i : i \in I\}$ cover M. Now we can define

$$\varphi_i = \frac{\lambda_i}{\sum_{i \in I} \lambda_i},$$

which is a partition of unity subordinate to the covering $\{U_i : i \in I\}$. ☐

An interesting consequence of these topological assumptions is the following theorem due to Milnor.

Theorem 1.14. *If M is a non-empty connected 1-dimensional manifold, then it is diffeomorphic to either $\mathbf{S}^1$ or $\mathbf{R}$.*

This result has a very similar flavor to the famous uniformization theorem for Riemann surfaces. We will prove this result later on – an original (sketch of the) proof can be found in the appendix of []. Although this in some sense coincides with our intuition, this isn't a the best justification of the topological assumptions on M, in and of itself, since there are many spaces that can be endowed with (non-diffeomorphic) *exotic* smooth structures.

In light of the fact that paracompact spaces are normal and Urysohn's metrization theorem we have that all smooth manifolds are metrizable.

1.3. **Orientability.** We present a global notion of orientation on manifolds. This will be key in understanding integration on manifolds in the next section. It will also allow us to better understand the examples that will arise time and time again throughout these notes.

Definition 1.15. Let M be a smooth manifold. An **orientation** on M is a subatlas $\mathscr{A}$ such that all of the transition functions have $\det(d\tau) > 0$. $\diamond$

Definition 1.16. A M is **orientable** if it admits an orientation. If M does not admit an orientation, we say that M is **non-orientable**. $\diamond$

Definition 1.17. Two distinct orientations **determine the same orientation** if their union is an orientation. $\diamond$

Proposition 1.18. *Let M be a smooth orientable and connected manifold. Then there exist exactly two distinct orientations on M.*

We delay this proof until we introduce differential forms, since we obtain a less brute force approach to analyzing orientability.

1.4. **Examples.** Now we present a few examples of manifolds.

Example 1.19 ($\emptyset$). The empty set is a manifold; in fact, it is an n-dimensional manifold for every $n \in \mathbf{N} \cup \{0\}$ $\diamond$

Example 1.20 ($\mathbf{R}^n$). $\mathbf{R}^n$ is a manifold with atlas $\mathscr{A} = \{\mathrm{id} : \mathbf{R}^n \to \mathbf{R}^n\}$. $\diamond$

Example 1.21 ($\mathbf{S}^n$). Consider the n-sphere

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

Let

$$\begin{aligned} U^+ &:= \{\boldsymbol{x} \in \mathbf{S}^n \setminus (0, \cdots, 0, +1)\} \quad \text{and} \\ U^- &:= \{\boldsymbol{x} \in \mathbf{S}^n \setminus (0, \cdots, 0, -1)\}. \end{aligned}$$

The maps $\varphi_\pm : U^\pm \to \mathbf{R}^n$ defined by

$$\varphi_\pm(x) := \frac{(x_1, \ldots, x_n)}{1 \mp x_{n+1}}$$

are charts on $\mathbf{S}^n$. To see that $\mathscr{A} = \{\varphi_+, \varphi_-\}$ defines an atlas on $\mathbf{S}^n$, we need to compute the transition maps. Note that

$$\begin{aligned} |\varphi_\pm(x)|^2 &= \frac{1 - x_{n+1}^2}{(1 \mp x_{n+1})^2} \\ &= \frac{(1 - x_{n+1})(1 + x_{n+1})}{(1 \mp x_{n+1})^2} \\ &= \frac{1 \pm x_{n+1}}{1 \mp x_{n+1}}. \end{aligned}$$

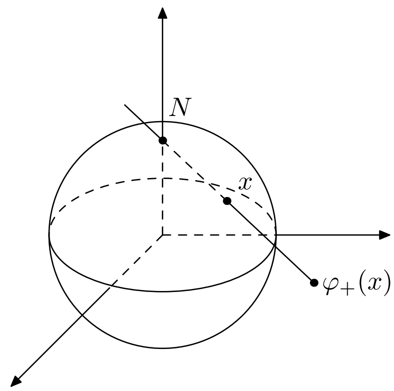
So we see that

$$\frac{\varphi_\pm(x)}{|\varphi_\pm(x)|^2} = \frac{(x_1, \ldots, x_n)}{1 \mp x_{n+1}} \cdot \frac{1 \mp x_{n+1}}{1 \pm x_{n+1}} = \frac{(x_1, \ldots, x_n)}{1 \pm x_{n+1}} = \varphi_\mp(x)$$

for $x \in U^+ \cap U^-$. Now we can compute the transition maps trivially:

$$(\varphi_\pm \circ \varphi_\mp^{-1})(x) = \frac{\varphi_\mp(\varphi_\mp^{-1}(x))}{|\varphi_\mp(\varphi_\mp^{-1}(x))|^2} = \frac{x}{|x|^2},$$

which is clearly smooth on the domain of the transition maps since $\varphi_\pm \circ \varphi_\mp^{-1} : \mathbf{R}^n \setminus \{0\} \to \mathbf{R}^n \setminus \{0\}$. So we have that $\mathscr{A}$ is indeed an atlas on $\mathbf{S}^n$.



◇

Example 1.22 (Normed vector space). Let $(V, \|\cdot\|)$ be a normed n-dimensional vector space. This norm induces a topology (which is independent of the chosen norm) on V. Pick a basis $\{e_i\}_{i=1}^n$ of V. Define $\varphi : V \to \mathbf{R}^n$ via

$$\sum_{i=1}^n \lambda_i e_i \mapsto (\lambda_1, \ldots, \lambda_n).$$

Note that φ is a vector space isomorphism, as well as homeomorphism. So we have that $\mathscr{A} = \{\varphi : V \to \mathbf{R}^n\}$ is a smooth atlas / global chart on V. Note that since a linear coordinate change with matrix $P \in \mathbf{GL}(V)$ is a smooth diffeomorphism the **standard smooth structure** on V determined above is independent of choice of basis. ◇

Example 1.23. If M is a smooth manifold and $\Omega \subseteq M$ is open, then Ω is also a smooth manifold when endowed with the subspace topology. More precisely, given a smooth atlas $\mathscr{A} = \{\varphi_i : U_i \to \widetilde{U}_i \ : \ i \in I\}$ on M we have that $\mathscr{B} = \{\varphi_i|_\Omega : U_i \cap \Omega \to \widetilde{U}_i \ : \ i \in I\}$ is a smooth atlas on Ω. ◇

Example 1.24. Consider the **real projective space** $\mathbf{R}P^n$ of 1-dimensional subspaces of $\mathbf{R}^{n+1}$, or equivalently $\mathbf{R}^{n+1} \setminus \{0\}/\mathbf{R}^*$. Given $(x_0, \ldots, x_n) \in \mathbf{R}^{n+1} \setminus \{0\}$, we write $[x_0 : \cdots : x_n]$ to represent the corresponding element of $\mathbf{R}P^n$. These are called **homogenous coordinates**. To obtain an atlas, note that $\mathbf{R}P^n$ is covered by

$$U_i := \{[x_0 : \ldots, : x_{i-1} : 1 : x_{i+1} : \cdots : x_n]\},$$

for $i = 0, \ldots, n$ and $\varphi : U_i \to \mathbf{R}^n$ defined by

$$\varphi_i([x_0 : \ldots, x_{i-1} : 1 : x_{i+1} : \ldots, x_n]) := (x_0, \ldots, \widehat{x}_i, \ldots, x_n)$$

is a chart, where $\widehat{x}_i$ means that x_i is omitted.

Note that

$$\varphi_i \circ \varphi_j^{-1}(x_0, \ldots, \widehat{x}_j, \ldots, x_n) = \frac{1}{x_i}(x_0, \ldots, \widehat{x}_i, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_n),$$

and so therefore $\mathscr{A} = \{\varphi_i : U_i \to \mathbf{R}^n\}_{i=0}^n$ defines an atlas on $\mathbf{R}P^n$. ◇

Example 1.25. In a completely analogous way one can define the **complex projective space** $\mathbb{C}P^n$ of 1-dimensional complex subspaces of $\mathbb{C}^{n+1}$. $\mathbb{C}P^n$ has complex homogenous coordinates and those induce a natural atlas on $\mathbb{C}P^n$. In fact, this atlas endows $\mathbb{C}P^n$ the structure of a complex manifold. ◇

Example 1.26 (Grassmanian). Let $\mathbf{Gr}_k(\mathbf{R}^n)$ denote the **Grassmannian**, that is, the space of k-dimensional subspaces of $\mathbf{R}^n$. Here we show that $\mathbf{Gr}_k(\mathbf{R}^n)$ is a smooth compact manifold. Later we will show that a general (possibly infinite dimensional) version of the Grassmanian is a Banach manifold, from which the fact that $\mathbf{Gr}_k(\mathbf{R}^n)$ is a smooth compact manifold follows.

Consider the map

$$\begin{aligned}\{T \in \hom(\mathbf{R}^k, \mathbf{R}^n) \ : \ \operatorname{rank}(T) = k\}/\mathbf{GL}(\mathbf{R}^k) &\to \mathbf{Gr}_k(\mathbf{R}^n) \\ [T] &\mapsto \operatorname{im}(T).\end{aligned}$$

This map is a bijection, and the action of $\mathbf{GL}(\mathbf{R}^k)$ is free. We equip $\mathbf{Gr}_k(\mathbf{R}^n)$ with the quotient topology, which is Hausdorff and second countable. Now consider the set of indices

$$\mathbf{I} := \{(i_1, \dots, i_k) \,:\, 1 \leq i_1 < \cdots < i_k \leq n\}.$$

For $I \in \mathbf{I}$ set

$$U_I := \{T \in \hom(\mathbf{R}^k, \mathbf{R}^n) \,:\, \det(T_{ij})_{i \in I, j \in [k]\} \neq 0} / \mathbf{GL}(\mathbf{R}^k).$$

Now recall from basic linear algebra that a rank k matrix must have at least one non-vanishing $k \times k$ minor, so we deduce that U_I form an open cover of $\mathbf{Gr}_k(\mathbf{R}^n)$. Now define charts $\phi_I : U_I \to \hom(\mathbf{R}^k, \mathbf{R}^{n-k})$ as follows: without loss of generality (by reindexing) we can assume that $I = (1, 2, \dots, k)$. Given $[T] \in U_I$, we can use Gaussian elimination to find a unique $g \in GL(\mathbf{R}^k)$ such that

$$Tg = \begin{pmatrix} \mathbb{1}_{k \times k} \\ B \end{pmatrix},$$

where $\mathbb{1}_{k \times k}$ is the $k \times k$ identity matrix and $B \in \hom(\mathbf{R}^{n-k}, \mathbf{R}^k)$. We now define $\phi_I([T]) := B$. This is clearly well defined, in the sense that it does not depend on the representative. Now we show that $\mathscr{A} := \{\phi_I : U_I \to \hom(\mathbf{R}^{n-k}, \mathbf{R}^k) \cong \mathbf{R}^{(n-k)k}\}$ is an atlas.

⋄

Proposition 1.27. *Let $(M, \mathscr{A})$ and $(N, \mathscr{B})$ be manifolds of dimension m and n respectively. There is a unique maximal atlas on $M \times N$ such that the projection maps $M \times N \to M$ and $M \times N \to N$ are smooth.*

Proof. Suppose there is a maximal atlas $\mathscr{C}$ with respect to which the projection maps are smooth. If $\varphi : U \to \widetilde{U}$ and $\psi : V \to \widetilde{V}$ are charts on M and N respectively, then $\varphi \times \psi : U \times V \to \widetilde{U} \times \widetilde{V}$ is a chart in $\mathscr{C}$.

Conversely, the set of all such charts defines an atlas with respect to which the projection maps are smooth. Indeed, we have that

$$\bigcup_{\varphi_\alpha \in \mathscr{A}} \bigcup_{\psi_\beta \in \mathscr{B}} U_\alpha \times V_\beta = M \times N,$$

and if

$$(U_\alpha \times V_\beta) \cap (U_\gamma \times V_\delta) \neq \emptyset$$

then we have that

$$(\varphi_\gamma \times \psi_\delta) \circ (\varphi_\alpha \times \psi_\delta)^{-1} = (\varphi_\gamma \circ \varphi_\alpha^{-1}) \times (\psi_\delta \circ \psi_\beta^1)$$

is clearly smooth since M and N are smooth manifolds. So we see that this defines a smooth structure on the Cartesian product $M \times N$. With respect to this atlas we clearly have that the projection maps are smooth. □

1.5. **Smooth maps between manifolds.** We now present the natural notion of a smooth map between manifolds. Since we don't have a natural way to talk about differentiation on arbitrary topological spaces, we pull back into local coordinates using the atlas and check that the smoothness holds between subsets of $\mathbf{R}^n$.

Definition 1.28. Let $(M, \mathscr{A})$ and $(N, \mathscr{B})$ be two smooth manifolds. A continuous map $f : M \to N$ is called **smooth** if for all charts $\varphi : U \to \widetilde{U}$ in $\mathscr{A}$ and $\psi : V \to \widetilde{V}$ in $\mathscr{B}$ we have that

$$\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(V)) \to \psi(V)$$

is smooth. The space of all smooth maps $f : M \to N$ is denoted by $\mathscr{C}^\infty(M, N)$. ⋄

Theorem 1.29. *If $f : M \to N$ and $g : N \to O$ both are smooth maps between manifolds, then so is $g \circ f : M \to O$.*

Proof. □

We reserve the term **smooth function** for smooth maps into $\mathbf{R}$. The space of all smooth functions on M is denoted by $\mathscr{C}^\infty(M)$.

Now we present an alternative perspective for thinking about smooth functions between manifolds. Note that for any function $F : M \to N$ we have a natural dual (or **pullback**) that takes functions defined on subsets of N to functions defined on subsets of M. Namely, for any function $f : A \subset N \to \mathbf{R}$ we define

$$F^*(f) = f \circ F : F^{-1}(A) \subset M \to \mathbf{R}.$$

Now we present the following lemma which motivates this alternative notion of smoothness.

Lemma 1.30. *A function $F : M \to N$ is continuous if and only if $F^*(f)$ is continuous for all continuous functions $f : A \subset N \to \mathbf{R}$ defined on open subsets of N.*

Proof.

($\implies$) First suppose that $F : M \to N$ is continuous. Now consider any arbitrary $f : A \subset N \to \mathbf{R}$ where A is an open set. Now it is immediate that $F^*(f) = f \circ F$ is continuous since it is the composition of two continuous functions.

($\impliedby$) Now suppose that $F^*(f)$ is continuous for all continuous functions $f : A \subset N \to \mathbf{R}$ defined on open subsets of N.

Fix an open set $U \subset N$. Now choose a continuous function $\lambda : N \to [0, \infty)$ such that $\lambda^{-1}((0,\infty)) = U$. Now we have that

$$(F^*(\lambda))^{-1}((0,\infty)) = (\lambda \circ F)^{-1}((0,\infty)) = F^{-1}(\lambda^{-1}((0,\infty))) = F^{-1}(U).$$

Note that this is open since we have that $F^*(\lambda)$ is open by assumption, and so we have that F is continuous since U was arbitrary.

□

So we see that $F : M \to N$ is continuous if and only if $F^*(\mathscr{C}^0(N)) \subset \mathscr{C}^0(M)$. This suggests that a function $F : M \to N$ is smooth if and only if $F^*(\mathscr{C}^\infty(N)) \subset \mathscr{C}^\infty(M)$. The following proposition shows several equivalent notions of smoothness:

Proposition 1.31. *Let $f : M \to N$ be continuous. Then the following conditions are equivalent:*

(1) If $\mathscr{B}$ is an atlas on N (a differentiable structure), then $F^*(\mathscr{B}) \subset \mathscr{C}^\infty(M)$.
(2) $F^*(\mathscr{C}^\infty(N)) \subset \mathscr{C}^\infty(M)$.
(3) For all charts $\varphi : U \to \widetilde{U}$ of M and $\psi : V \to \widetilde{V}$ of N we have that

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

is smooth.

Definition 1.32. A smooth map $f : M \to N$ is called a **diffeomorphism** if it is bijective and $f^{-1} : N \to M$ is smooth. The space of diffeomorphisms $f : M \to M$ is denoted by $\mathrm{Diff}(M)$. ⋄

Definition 1.33. Two manifolds M and N are called **diffeomorphic** if there exists a diffeomorphism $M \to N$. ⋄

Remark 1.34.

- In dimensions $n \leq 3$ there is only a single smooth structure that can be put on a topological n-manifold – if two manifolds are homeomorphic then they have the same smooth structure. As a result in these cases one generally does not specify the smooth structure.
- There is a unique smooth structure compatible with the Euclidean topology on $\mathbf{R}^n$, except in the case of $n = 4$ where there are infinitely many *exotic* smooth structures.
- The seven-dimensional sphere $\mathbf{S}^7$ admits 28 *distinct* smooth structures which form an abelian monad with respect to the connected sum.

⋄

1.6. **Tangent spaces.** Throughout this section we will consider an n-dimensional manifold M unless otherwise specified.

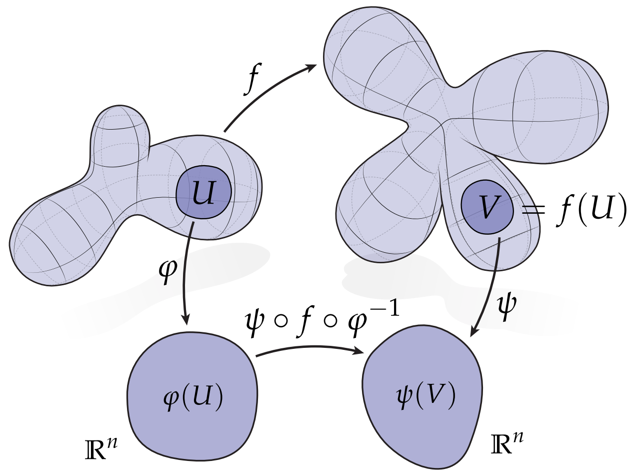


FIGURE 2. If two manifolds M and N are diffeomorphic then we say that they have the same smooth structure since a smooth function on M is also a smooth function on N after being pushed-forward by the diffeomorphism.

1.6.1. *Motivation.* Tangent vectors in Euclidean space are thought of as tangents or velocities to curves. When considering smooth manifolds, we would like tangent vectors to retain this meaning. However, tangent vectors have no such place to live unless we know the manifold is embedded in Euclidean space. When we think about tangent vectors in $\mathbf{R}^n$ we really are thinking about a separate copy of $\mathbf{R}^n$ at each point of the surface – that is the tangent space at a point $p \in \mathbf{R}^n$ is really a separate copy of $\mathbf{R}^n$ translated such that the origin is at p.

Now we present an example of what the tangent bundle (the set of all tangent planes at each point of the surface) of a sphere is that will motivate the definition of the tangent space that follows.

Example 1.35. We will consider the manifold $\mathbf{S}^n$ embedded in $\mathbf{R}^{n+1}$. Now we would like to consider the set of vectors tangent to a sphere. By tangent to the sphere we mean that they are velocity vectors for curves in the sphere. If $c : I \to \mathbf{S}^n$ is a curve then we have that $\|c\|^2 = 1$ and consequently we have that

$$\dot{c}c = 0.$$

This tells us that the velocity is always perpendicular to the base vector. So the tangent bundle of $\mathbf{S}^n$ is the following:

$$T\mathbf{S}^n \cong \left\{(p, v) \in \mathbf{R}^{n+1} \times \mathbf{R}^{n+1} \ : \ \|p\| = 1 \text{ and } \langle p, v\rangle = 0\right\}.$$

Notice that for any $(p, v) \in T\mathbf{S}^n$ we have that v is the velocity of

$$c(t) = p\cos t + v\sin t,$$

at the base point $p \in \mathbf{S}^n$ at time $t = 0$.

This is a very geometric way to describe the tangent space of $\mathbf{S}^n$, and does not generalize well to arbitrary topological spaces endowed with a manifold structure. Since $\mathbf{S}^n$ is a smooth manifold, we have that for every point $p \in \mathbf{S}^n$ that we can write

$$p_{n+1} = F(p_1, \ldots, p_n)$$

for some smooth function F – this is just saying that locally we can describe the manifold as the graph of a smooth function. Now if $c : I \to \mathbf{S}^n$ is a curve, then we have that

$$c_{n+1}(t) = F(c_1(t), \ldots, c_n(t)).$$

Hence

$$\dot{c}_{n+1}(t) = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\dot{c}_i(t).$$

So we see that for the tangent vectors $v \in T_p\mathbf{S}^n$ that we can write

$$v_{n+1} = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} v_i.$$

That is that we can write v_{n+1} as a smooth function of our chosen coordinates given that p_{n+1} is a smooth function of p_n. This last argument is sufficiently general to describe the tangent space of an n-dimensional submanifold of $\mathbf{R}^m$. ⋄

Now that we have worked through this example we go on to formally define the tangent space of a smooth manifold. Here we present two equivalent notions of the tangent space that each illustrate various important ways of thinking about tangent vectors.

1.6.2. *Tangent vectors as curves.* Since we want our tangent vectors to describe the velocities of curves on manifolds it is a natural idea to let the curves describe the tangent space of the manifold. Let M be a smooth manifold and fix $p \in M$. Now we will consider the space of curves that go through p at time zero:

$$\mathscr{C}_p(M) := \left\{\gamma : (-\varepsilon_\gamma, \varepsilon_\gamma) \to M \ : \ \varepsilon_\gamma > 0,\ \gamma(0) = p \text{ and } \gamma \text{ is smooth}\right\}.$$

Now consider a chart $\varphi : U \to \widetilde{U}$ containing p. Then we have that $\varphi \circ \gamma$ is a smooth curve in Euclidean space, and so we may consider it's differential in the classical sense.

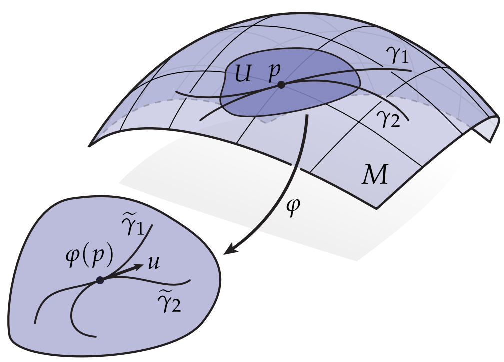


FIGURE 3. By using the smooth structure of M we are able to pullback smooth curves on the manifold to Euclidean realizations where we know how to talk about tangent vectors.

Definition 1.36. We say that two curves $\gamma_1 : (-\varepsilon_1, \varepsilon_1) \to M$ and $\gamma_2 : (-\varepsilon_2, \varepsilon_2) \to M$ are **equivalent curves** if for all charts $\varphi : U \to \widetilde{U}$ containing p we have that

$$\mathrm{d}(\varphi \circ \gamma_1)(0) = \mathrm{d}(\varphi \circ \gamma_2)(0).$$

So we have that the differentials agree at the image of p. We write $\gamma_1 \sim \gamma_2$. ⋄

Proposition 1.37. *Consider $\gamma_1 : (-\varepsilon_1, \varepsilon_1) \to M$ and $\gamma_2 : (-\varepsilon_2, \varepsilon_2) \to M$. We have that $\gamma_1 \sim \gamma_2$ if for any chart $\varphi : U \to \widetilde{U}$ containing p we have that*

$$\mathrm{d}(\varphi \circ \gamma_1)(0) = \mathrm{d}(\varphi \circ \gamma_2)(0).$$

That is if two curves are equivalent in one chart containing p, then they are equivalent in all charts containing p.

Proof. Suppose that for some chart $\varphi \in \mathscr{A}$ containing p we have that

$$\mathrm{d}(\varphi \circ \gamma_1)(0) = \mathrm{d}(\varphi \circ \gamma_2)(0).$$

Now consider another chart $\psi \in \mathscr{A}$ containing p. We then have that

$$\psi \circ \gamma_1 = (\psi \circ \varphi^{-1}) \circ (\varphi \circ \gamma_1).$$

Note that the right hand side is defined on the intersection of the domains of ψ and φ, which is still an open set, and so we can talk about derivatives. So we have that

$$\mathrm{d}(\psi \circ \gamma_1)(0) = \mathrm{d}(\psi \circ \varphi^{-1}) \circ \mathrm{d}(\varphi \circ \gamma_1)(0) = \mathrm{d}(\psi \circ \varphi^{-1}) \circ \mathrm{d}(\psi \circ \gamma_2)(0) = \mathrm{d}(\psi \circ \gamma_2)(0).$$

The last step follows since $\mathrm{d}(\psi \circ \varphi^{-1})$ is invertible and therefore injective. $\square$

Note that it is clear that $\sim$ defines an equivalence relation on $\mathscr{C}_p(M)$.

Definition 1.38. The **tangent space of** M **at** p is defined as

$$T_p^{(1)}M := \mathscr{C}_p(M)/\sim .$$

$\diamond$

We use the superscript right now to differentiate the two definition of the tangent space, but will eventually just denote the tangent space as T_pM once we show the equivalence between spaces.

Theorem 1.39. *The map* $F_\varphi : T_p^{(1)}M \to \mathbf{R}^n$ *defined via*

$$[\gamma] \mapsto \mathrm{d}(\varphi \circ \gamma)(0)$$

for any chart φ *containing* p *is a bijection.*

Proof. Let $\varphi : U \to \widetilde{U}$ be a chart containing p. Consider $v \in \mathbf{R}^n$, and consider the curve $\gamma_v : I_v \to M$ defined as

$$\gamma_v(t) := \varphi^{-1}(\varphi(p) + tv),$$

where I_v is an open interval such that for all $t \in I_v$ we have that $\varphi(p) + tv \in \widetilde{U}$. Now we have that

$$\mathrm{d}(\varphi \circ \gamma_v)(0) = \mathrm{d}(\varphi(p) + tv) = v.$$

Since $v \in \mathbf{R}^n$ was arbitrary we have that F_φ is surjective.

Now to see that φ is injective note that if $F_\varphi[\gamma_1]) = F_\varphi([\gamma_2])$ then we have that the image of the differentials coincide at zero, and by definition we have that $\gamma_1 \sim \gamma_2$ and hence $[\gamma_1] = [\gamma_2]$. So we see that F_φ is injective, and therefore bijective. $\square$

Remark 1.40. We would like to endow $T_p^{(1)}M$ with a canonical vector space structure. That is we would like $T_p^{(1)}M$ to be isomorphic to $\mathbf{R}^n$. This vector space structure is not entirely clear, but one can use the map F_φ to pullback the vector space structure of $\mathbf{R}^n$ into $T_p^{(1)}M$. This induces the following vector space operations for $[\gamma_1], [\gamma_2] \in T_pM$

$$[\gamma_1] + [\gamma_2] := [\varphi^{-1} \circ (\varphi \circ \gamma_1 + \varphi \circ \gamma_2)]$$

and

$$c[\gamma_1] := [\varphi^{-1} \circ (c\varphi \circ \gamma_1)].$$

Although this vector space structure depends on the choice of coordinates, one can show that this definition is indeed coordinate invariant.

The main advantage of the next definition of T_pM is that there is a clear canonical vector space structure. This definition using equivalence classes of curves is very geometric in nature and shows us that our definition of tangent vectors really represent what we would like them to. $\diamond$

1.6.3. *Sheaf-theoretic tangent vectors.* Here we present a very different approach to the construction of the tangent space of a manifold that uses very elementary sheaf-theoretic ideas.

Fix $p \in M$. We will consider the subset $\mathscr{C}_p(M) \subset \mathscr{C}^\infty(M)$ of smooth functions whose domain contains p. Now we have a natural vector space structure, where we add functions on the intersection of their domains – which is clearly an open set. To simplify some of the results at the cost of an extra layer of abstraction we will modify $\mathscr{C}_p(M)$ to $\mathfrak{G}_p(M)$ by taking the quotient space with the relation being that two functions that coincide on some neighborhood of p are considered equivalent. The elements of $\mathfrak{G}_p(M)$ are called the **germs** of functions at p. This idea is very similar the definition of L^p spaces where we consider the quotient space of functions that are identical almost everywhere. The reason why will use this modification will be clear in light of Proposition 1.45.

Now consider a curve $c : I \to M$ with $c(0) = p$. We would like to make sense of the velocity of c at 0. Now for any coordinate function x_i then we have that $x \circ c$ measures this coordinates component of c in some chart of M. Similarly, we have that $\frac{\mathrm{d}}{\mathrm{d}t}(x \circ c)$ measures the change in velocity with respect to f. However, since there is no canonical way to choose a differentiable framework, or choice of coordinates, we will instead consider all of the time derivatives of $f \circ c$ for all $f \in \mathscr{C}_p(M)$.

Definition 1.41. The **velocity of the curve** c **at** 0 is the map $\dot{c}(0)$

$$\begin{aligned}\mathscr{C}_p(M) &\to \mathbf{R}\\ f &\mapsto \frac{\mathrm{d}}{\mathrm{d}t}(f\circ c)(0).\end{aligned}$$

⋄

This definition comes from the idea that we can completely determine the velocity of a curve c by the specifying the directional derivatives with respect to $\dot{c}(0)$ for all smooth functions defined on a neighborhood of $p = c(0)$:

$$D_{\dot{c}(0)}f = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ c)(0).$$

Definition 1.42. A **derivation at** p is a linear map $D : \mathscr{C}_p(M) \to \mathbf{R}$ that also satisfies the Leibniz product rule for differentiation at p:

$$D(fg) = D(f)g(p) + f(p)D(g)$$

for all $f, g \in \mathscr{C}_p(M)$. ⋄

We now prove two important properties of derivations:

Proposition 1.43. *Consider some point $p \in M$.*

(1) Derivations vanish on constant functions. That is if D is a derivation and c is a constant function we have that $D(c) = 0$.
(2) If for two functions $f, g \in \mathscr{C}_p(M)$ such that $[f] = [g]$ in $\mathfrak{G}_p(M)$ then $D(f) = D(g)$.

Proof. Both of these properties are straightforward.

(1) This is straightforward since

$$D(c) = cD(1) = cD(1\cdot 1) = c(D(1)+D(1)) \implies D(c) = 0.$$

(2) Consider a function f that vanishes on a neighborhood of p. Then we can find $\lambda : M \to \mathbf{R}$ that is 1 on a neighborhood of p and $\lambda = 0$ on the complement where f vanishes. So we have that $\lambda f = 0$ on M. It follows that

$$0 = D(\lambda f) = D(\lambda)f(p) + \lambda(p)D(f) = D(f).$$

So we see that if two functions f, g agree on a neighborhood of p, then $f - g$ vanishes on a neighborhood of p and so we have that $D(f) = D(g)$.

□

Since functions that coincide on any neighborhood of p have the same derivation, we only consider derivations defined on the germs of functions at p – we can compute the derivation by choosing any representative.

Definition 1.44. The **tangent space of** M **at** p is defined as the vector space of derivations (on $\mathfrak{G}_p(M)$) at p, and is denoted by $T_p^{(2)}M$. ⋄

Proposition 1.45. *If $p \in U \subset M$, where U is open, then $T_p^{(2)}U = T_p^{(2)}M$.*

Proof. This follows directly from 4. Since the $\mathfrak{G}_p(M)$ is a purely local definition we have that $\mathfrak{G}_p(M) = \mathfrak{G}_p(U)$. The equivalence follows directly from the definition of $T_p^{(2)}M$. □

Now we show that the velocity of a curve is indeed a derivation at p.

Proposition 1.46. *For any smooth curve $c : I \to M$ such that $c(0) = p$ we have that the map $f \mapsto \frac{\mathrm{d}}{\mathrm{d}t}(f\circ c)(0)$ is a derivation at p.*

Proof. This map is clearly linear in f from the fact that differentiation is linear. Note that the Leibniz product rule also follows from the product rule for differentiation. Indeed, let $f, g \in \mathscr{C}_p(M)$ and then we have that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}((fg)\circ c)(0) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}(f\circ c)(0)\right)(g\circ c)(0) + (f\circ c)(0)\left(\frac{\mathrm{d}}{\mathrm{d}t}(g\circ c)(0)\right)\\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}(f\circ c)(0)\right)g(p) + f(p)\left(\frac{\mathrm{d}}{\mathrm{d}t}(g\circ c)(0)\right).\end{aligned}$$

□

Here we can see that restricting derivations to equivalence classes in $\mathfrak{G}_p(M)$ we are doing something analogous to considering equivalence classes of curves in the definition of $T_p^{(1)}M$. Notice that in the case of equivalence classes of curves we are in some sense considering an analog to the space of derivations that arise from curves.

Lemma 1.47. *The map $\mathbf{R}^n \to T_p^{(2)}M$ defined via*

$$v \mapsto D_v(f) := \sum_{i=1}^n v^i \frac{\partial}{\partial x_i}(f\circ\varphi^{-1})(\varphi(p))$$

is an isomorphism for any chart φ containing p.

Proof. Fix any chart $\varphi \in \mathscr{A}$ containing p. This mapping is clearly linear. Note that since

$$D_v(x_i \circ \varphi) = v^i$$

we have that the kernel of the mapping is trivial. So it remains to show that the map is surjective.

Note that for any smooth function $f \in \mathscr{C}_p(M)$ we have that $f\circ\varphi^{-1}$ is a *smooth* map from $\mathbf{R}^n$ to $\mathbf{R}$. For simplicity define $g := f\circ\varphi^{-1}$. Now we have that since g is a smooth function we have a Taylor expansion formula:

$$g(x) = g(\varphi(p)) + \sum_{i=1}^n (x_i - \varphi(p))g_i(x)$$

where the g_i are smooth and $g_i(\varphi(p)) = \frac{\partial g}{\partial x_i}(\varphi(p))$. Namely, we define

$$g_i(x) = \int_0^1 \frac{\partial g}{\partial x_i}(tx)\,\mathrm{d}t,$$

and so we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}g(tx) = \sum_{i=1}^n x_i \frac{\partial g}{\partial x_i}(tx).$$

Now consider an arbitrary derivation $D \in T_p^{(2)}M$ and note that

$$\begin{aligned}D(f) = D(g\circ\varphi) &= D(g(\varphi(p))) + \sum_{i=1}^n [D(x_i\circ\varphi)g_i(\varphi(p)) + x_i(\varphi(p))D(g_i\circ\varphi)]\\ &= D(g(\varphi(p))) + \sum_{i=1}^n D(x_i\circ\varphi)g_i(\varphi(p))\\ &= \sum_{i=1}^n D(x_i\circ\varphi)\frac{\partial g}{\partial x_i}(\varphi(p))\\ &= \sum_{i=1}^n D(x_i\circ\varphi)\frac{\partial}{\partial x_i}(f\circ\varphi^{-1})(\varphi(p)).\end{aligned}$$

So now consider the vector

$$v := (D(x_1\circ\varphi), \ldots, D(x_n\circ\varphi)),$$

and we see that

$$D(f) = D_v(f).$$

So we conclude that the mapping is surjective, hence bijective, and thus an isomorphism of vector spaces. □

It now immediately follows that these two definitions of the tangent space coincide.

Theorem 1.48. *The mapping $T_p^{(1)}M \to T_p^{(2)}M$ defined via*

$$[\gamma] \mapsto D_\gamma(f) := \mathrm{d}(f \circ \gamma)(0),$$

is an isomorphism of vector spaces.

Proof. Note that the mapping is clearly linear. Since we have bijections between both $T_p^{(1)}M$ and $T_p^{(2)}M$ with $\mathbf{R}^n$ we have that

$$\dim T_p^{(1)}M = \dim T_p^{(2)}M = n.$$

So we have that they mapping is surjective and hence an isomorphism. □

From now on we will denote the tangent space as T_pM.

1.6.4. *Local basis of the tangent space.* Since T_pM has a natural vector space structure it will be useful to describe a natural choice of basis tangent vectors.

Definition 1.49. Let $(M, \mathscr{A})$ be a smooth n-dimensional manifold. Consider $p \in M$ and let $\varphi \in \mathscr{A}$ be a chart containing p. The **associated basis of** φ **of** T_pM is

$$\left\{ \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right\}$$

where $\frac{\partial}{\partial x_i} : \mathscr{C}_p(M) \to \mathbf{R}$ is the derivation defined by

$$\frac{\partial}{\partial x_i}(f) = \frac{\partial (f \circ \varphi^{-1})}{\partial x_i}(\varphi(p)).$$

⋄

It is easy to check that the associated basis of any chart is indeed a basis of the tangent space. This follows from the proof of 4 that

$$D(f) = \sum_{i=1}^n D(x_i \circ \varphi^{-1}) \frac{\partial (f \circ \varphi^{-1})}{\partial x_i}(\varphi(p))$$

That is

$$D = \sum_{i=1}^n v^i \frac{\partial}{\partial x_i}$$

and the v^i are uniquely determined. One can define a similar associated basis for the tangent space described by equivalence classes of curves.

Definition 1.50. The **cotangent space** T_pM^* to M at $p \in M$ is the vector space of linear functionals on T_pM. ⋄

Using local coordinates we obtain a natural dual basis $\mathrm{d}x_i$ satisfying

$$\mathrm{d}x_i \left(\frac{\partial}{\partial x_j} \right) = \delta^i_j.$$

This tells us that

$$\mathrm{d}x_i(v) = \mathrm{d}x_i \left(\sum_{j=1}^n v^j \frac{\partial}{\partial x_j} \right) = v^i,$$

which is that $\mathrm{d}x_i$ calculates the i^{th} coordinate of a vector.

Now we give a change of basis formula for the tangent and cotangent spaces.

Lemma 1.51. *Let $(M, \mathscr{A})$ be a smooth manifold, and fix $p \in M$. Suppose that $\{x_i : U \to \mathbf{R}\}_{i=1}^n$ and $\{y^i : V \to \mathbf{R}\}_{i=1}^n$ form two coordinate systems around $p \in M$ in the charts φ and ψ respectively. Then*

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^n \frac{\partial y_j}{\partial x_i} \frac{\partial}{\partial y_j}$$

and

$$\mathrm{d}x_i = \sum_{j=1}^n \frac{\partial x_i}{\partial y_j} \mathrm{d}y_j,$$

where $\left\{\frac{\partial}{\partial x_i}\right\}$ form a basis of T_pM and $\{\mathrm{d}x_i\}$ form a the corresponding basis of T_pM^*.

Proof. For all $1 \leq i, j \leq n$ we have that

$$\frac{\partial}{\partial x_i}(y_j) = \frac{\partial(y_j \circ \varphi^{-1})}{\partial x_i}(\varphi(p)) = \frac{\partial(y_j \circ (\psi \circ \varphi^{-1}))}{\partial x_i}(\varphi(p)) = \frac{\partial y_j}{\partial x_i}(\varphi(p)).$$

The last equality follows from the fact that $\psi \circ \varphi^{-1}$ the transition map $y \circ x$ and we are taking the i^{th} derivative of the j^{th} component.

Now consider any smooth function $f \in \mathscr{C}_p(M)$. We have that

$$\left(\sum_{j=1}^n \frac{\partial y_j}{\partial x_i} \frac{\partial}{\partial y_j}\right)(f) = \sum_{j=1}^n \frac{\partial y_j}{\partial x_i} \frac{\partial(f \circ \psi^{-1})}{\partial y_j}(\psi(p))$$

and

$$\frac{\partial}{\partial x_i}(f) = \frac{\partial(f \circ \varphi^{-1})}{\partial x_i}(\varphi(p)) = \sum_{j=1}^n \frac{\partial}{\partial x_i}(y_j) \frac{\partial(f \circ \psi^{-1})}{\partial y_j}(\psi(p)) = \sum_{j=1}^n \frac{\partial y_j}{\partial x_i} \frac{\partial(f \circ \psi^{-1})}{\partial y_j}(\psi(p)).$$

The second equality follows from a simple change of variables. Now since these two derivations act the same on $\mathscr{C}_p(M)$ we conclude that they are the same vector in T_pM.

Now we compute

$$\mathrm{d}x_i\left(\frac{\partial}{\partial y_j}\right) = \left(\frac{\partial}{\partial y_j}\right)[x_i] = \frac{\partial x_i}{\partial y_j}.$$

So we see that for any tangent vector $v \in T_pM$ we can write

$$v = \sum_{j=1}^n v^j \frac{\partial}{\partial y_j}$$

and so

$$\left(\sum_{j=1}^n \frac{\partial x_i}{\partial y_j} \mathrm{d}y_j\right)(v) = \sum_{j=1}^n v^j \frac{\partial x_i}{\partial y_j}.$$

Similarly, we have that

$$\mathrm{d}x_i(v) = \sum_{j=1}^n v^j \frac{\partial x_i}{\partial y_j}.$$

So we conclude that $\mathrm{d}x_i$ and $\sum \frac{\partial x_i}{\partial y_j} \mathrm{d}y_j$ are the same covector in T_pM^* since they act on tangent vectors in the same way. □

Note that the matrices $\left[\frac{\partial y_j}{\partial x_i}\right]$ and $\left[\frac{\partial x_i}{\partial y_j}\right]$ have entries that are functions on the intersection of the domains of φ and ψ. Furthermore, we have that these two matrices are inverses of each other.

1.6.5. *Differential of a smooth mapping.* Now suppose we want to make sense of the derivative of a smooth map $f : M \to N$. We can always pull back into coordinates in the following way: Fix $p \in M$, $\varphi \in \mathscr{A}$, the atlas of M, and $\psi \in \mathscr{B}$, the atlas of N where p is in the domain of φ and $f(p)$ is the domain of ψ. Now define

$$\widehat{f} := \psi \circ f \circ \varphi^{-1}$$

at $\varphi(p)$ and we would want $d\widehat{f}_{\varphi(p)} : \mathbf{R}^m \to \mathbf{R}^n$ to represent the derivative of f at p. However, this definition depends on the choice of charts. For example if we choose other charts φ^* and ψ^* and set

$$f^* := \psi^* \circ f \circ (\varphi^*)^{-1},$$

then

$$df^*_{\varphi^*(p)} = d(\psi^* \circ \psi^{-1})_{\psi(f(p))} \circ d\widehat{f}_{\varphi(p)} \circ d(\varphi \circ (\varphi^*)^{-1})_{\varphi^*(p)}.$$

However, now that we have a notion of tangent space we can talk about the differential of a mapping in a coordinate invariant manner. The key observation is that the differential of a mapping in Euclidean space maps tangent vectors to tangent vectors.

Definition 1.52. The **differential** of $f : M \to N$ at $p \in M$ is the mapping

$$df_p : T_pM \to T_{f(p)}N$$

define as follows for all $v \in T_pM$ we define

$$df_p(v)[g] = v[g \circ f]$$

for all $g \in \mathscr{C}_{f(p)}(N)$. ◇

It is easy to see that $df_p(v) \in T_{f(p)}N$ and that the differential is linear in f.

Example 1.53 (Smooth functions). As a special case we would like to explore how the differential acts when we have a smooth function $f : M \to \mathbf{R}$. Note that $T_q\mathbf{R} \cong \mathbf{R}$ and so we can consider the coordinate chart given by the natural identity mapping $\mathrm{id}_d : \mathbf{R} \to \mathbf{R}$. Using this global chart we can consider any vector $w \in T_q\mathbf{R}$ can be written as

$$w = c\frac{\partial}{\partial y}.$$

We now have an identification between $T_q\mathbf{R}$ and $\mathbf{R}$ given by

$$c\frac{\partial}{\partial y} \mapsto c.$$

Using this identification we have that the differential df_p becomes a linear mapping

$$df_p : T_pM \to \mathbf{R}.$$

That is that $df_p \in T_pM^*$. Now we claim that for any smooth $f : M \to \mathbf{R}$ and any $v \in T_pM$ that

$$df(v) = v[f].$$

Now that if $df(v)$ is identified with $c \in \mathbf{R}$ then for $g : \mathbf{R} \to \mathbf{R}$ we have

$$df(v)[g] = cg'(q).$$

So it remains to show that $df(v)[g] = v[f]g'(q)$. Consider any smooth curve $\gamma : (-\varepsilon, \varepsilon) \to M$ such that $\gamma(0) = p$ and $\gamma'(0) = v$. That is γ is a curve that represents $v \in T_pM$. Then a direct application of the chain rule shows that

$$df(v)[g] = v[g \circ f] = \frac{\mathrm{d}}{\mathrm{d}t}(g \circ f \circ \gamma)(0) = \frac{\partial g}{\partial y}(f(\gamma(0)))\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma)(0) = g'(q)v[f].$$

So we see that our abstract notion of the differential coincides with what we expect that is since $v[f] = D_vf$ we have that

$$df(v) = D_vf.$$

◇

Proposition 1.54 (Chain rule of the differential)**.** *Let $f : M \to N$ and $g : N \to P$ be smooth maps. Then*

$$d(g \circ f)_p = dg_{f(p)} \circ df_p.$$

Proof. Fix any vector $v \in T_pM$ and a smooth function $h : P \to \mathbf{R}$. Then

$$\begin{aligned}\big(d(g\circ f)_p[v]\big)(h) &= v\,[h\circ(g\circ f)]\\ &= v\,[(h\circ g)\circ f]\\ &= (df_p[v])(h\circ g)\\ &= dg_{f(p)}\big(df_p[v]\big)(h)\\ &= \big(\big(dg_{f(p)}\circ df_p\big)(v)\big)(h).\end{aligned}$$

Since v and f were arbitrary we deduce the desired result. □

1.7. **Fiber bundles and vector fields.** As we vary the point $p \in M$ the tangent space T_pM vary in a smooth fashion. To formalize this notion we define fiber bundles (and maybe more general fibrations). We will be able to use fiber bundles and fibrations to encode topological and geometric information about the spaces over which they are defined. Namely, we will be able to discuss and study many constructions, such as the tangent bundle of a manifold.

Definition 1.55. Let $\pi : E \to B$ be a mapping of sets. Then the **fiber** of π over $p \in B$ is

$$\pi^{-1}(p) \subset E.$$

⋄

Definition 1.56. A **bundle** is a triple (E, B, π) where $\pi : E \to B$ is a continuous map. The space B is called the **base space**. The space E is called the **total space**, and π is called the **projection** of the bundle. ⋄

Note that if π is surjective then each fiber is nonempty, and the map π partitions E as

$$E = \bigsqcup_{p\in B} \pi^{-1}(p).$$

Definition 1.57. Let F, E, and B be topological spaces, and $\pi : E \to B$. The 4-tuple (E, B, π, F) is a **fiber bundle** if π is a surjection and for every $x \in E$ there exists a neighborhood Ω of $\pi(x)$ in B and a homeomorphism $\varphi : \pi^{-1}(\Omega) \to \Omega \times F$ such that the following diagram commutes:

$$\begin{tikzcd} \pi^{-1}(\Omega) \arrow[r, "\varphi"] \arrow[d, "\pi"] & \Omega \times F \arrow[dl, "\mathrm{proj}_\Omega"] \\ \Omega & \end{tikzcd}$$

where proj_Ω is the projection onto Ω. We call F the **fiber**, and φ a **trivialization**. ⋄

Notation 1.58. We use E_x to denote the fiber, $\pi^{-1}(x)$, over x. Fiber bundles are often called locally trivial fibrations, and we will often use the projection to represent the fibration, i.e. $\pi : E \to B$ will represent (E, B, π, F) where F is not specified.

Definition 1.59. A fiber bundle is **smooth** if E and B are smooth manifolds, π is smooth, and the local trivializations φ can be chosen to be a diffeomorphism. ⋄

In the context of Riemannian geometry we will generally only consider smooth fiber bundles.

Definition 1.60. Let (E, M, π, F) be a fiber bundle where E and M are manifolds. A (smooth) **section** is a (smooth) map $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$. The space of sections is denoted by $\Gamma(M, E)$, or $\Gamma(E)$ if the manifold M is clear from the context. ⋄

Example 1.61. The projection map $X \times F \to X$ is the *trivial* fibration over X with fiber F. ⋄

Example 1.62. Let $\mathbf{S}^1 \subseteq \mathbb{C}$ be the unit circle with basepoint $1 \in \mathbf{S}^1$. Consider the map $f_n : \mathbf{S}^1 \to \mathbf{S}^1$ given by $f_n(z) = z^n$. Then $f_n : \mathbf{S}^1 \to \mathbf{S}^1$ is a locally trivial fibration with fiber a set of n distinct points (the n^{th} roots of unity in $\mathbf{S}^1$). ⋄

Example 1.63. Consider the Mobïus band $M = [0,1]\times[0,1]/\sim$ where $(t,0)\sim(1-t,1)$. Let C be the *center circle* $C = \{(1/2,s)\in M\}$ and consider the projection

$$\begin{aligned}\pi : M &\to C\\ (t,s) &\mapsto (1/2,s).\end{aligned}$$

This map is a locally trivial fibration with fiber $[0,1]$. ⋄

A very import class of fiber bundles that will be important to our study of Riemannian geometry are vector bundles.

Definition 1.64. A **vector bundle** is a fiber bundle where the space F is a vector space, and with the additional requirement of the local trivialization map φ induces a linear transformation on each fiber. That is for every $x\in\Omega$ we have that

$$\varphi : \pi^{-1}(x) \to \{x\}\times F$$

is an isomorphism of vector spaces. The **rank** of the vector bundle is the dimension of F. ⋄

Definition 1.65. Let (E_1, M_1, π_1, F_1) and (E_2, M_2, π_2, F_2) be two vector bundles, and let $f : M_1 \to M_2$ be a smooth map between manifolds. A **vector bundle map** over f is a smooth map $g : E_1 \to E_2$ such that

- $\pi_2 \circ g = f \circ \pi_1$, and
- for every $x\in M$, the induced map $\pi_1^{-1}(x) \to \pi_2^{-1}(f(x))$ is linear.

⋄

If $M_1 = M_2$ and we do not specify f, they we take $f = \mathrm{id}_{M_1}$.

Definition 1.66. A **vector bundle isomorphism** between (E_1, M_1, π_1, V_1) and (E_2, M_2, π_2, V_2) is a vector bundle map $g : E_1 \to E_2$ which has an inverse vector bundle map $g^{-1} : E_2 \to E_1$. ⋄

Although I probably will not spend much time talking about gauge transformations, we have the necessary notation to define them – gauge transformations arise in semi-riemannain geometry and general relativity.

Definition 1.67. A vector bundle isomorphism between E and itself is called a **gauge transformation**. The **gauge group** $\mathscr{G}(E)$ is the group of gauge transformations of E. ⋄

Definition 1.68. A vector bundle from $E\to M$ is called **trivial** if it is isomorphic to a product bundle $V\times M\to M$ for some vector field V. ⋄

Example 1.69. Note that the Grassmanian $\mathbf{Gr}_k(\mathbf{R}^n)$ comes with a **tautological vector bundle**

$$\gamma_k := \{(\Pi, v)\in \mathbf{Gr}_k(\mathbf{R}^n)\times\mathbf{R}^n \;:\; v\in\Pi\}$$

It is easy to check that this is indeed a vector bundle. ⋄

1.7.1. *Transition Functions for Vector Bundles.* One way that one can think about fiber bundles is that they are locally product spaces. This is a very similar idea to that of manifolds.

Given a vector bundle (E, M, π, F) we can find a cover of M by open neighborhoods U_α and local trivializations $\varphi_\alpha : \pi^{-1} \to U_\alpha\times F$. On the overlaps $U_{\alpha\beta} := U_\alpha\cap U_\beta$ we can form the **transition functions** $g^\alpha_\beta \in \mathscr{C}^\infty(U_{\alpha,\beta}, \mathbf{GL}(F))$ defined by

$$g^\alpha_\beta(x) := \varphi_{\beta,x}\circ\varphi^{-1}_{\alpha,x}.$$

These transition functions satisfy the **co-cycle condition** that $g^\alpha_\alpha = \mathrm{id}_F$ and on $U_{\alpha\beta\gamma} := U_\alpha\cap U_\beta\cap U_\gamma$

$$g^\gamma_\alpha g^\beta_\gamma g^\alpha_\beta = \mathrm{id}_F\,.$$

Now suppose that $\{U_\alpha\}_{\alpha\in\Lambda}$ is an open cover of M and $g^\alpha_\beta\in\mathscr{C}^\infty(U_{\alpha\beta}, \mathbf{GL}(F))$ satisfy the co-cycle conditions. Now define

$$E := \left(\bigsqcup_\alpha U_\alpha\times V\right)/\sim$$

with $(x,v)\sim(x',v')$ if and only if $x\in U_\alpha$, $x'\in U_\beta$, $x=x'\in M$ and $v'=g^\alpha_\beta v$. We claim that E has the structure of a vector bundle over M. Indeed, let

$$\pi_E:\bigsqcup_\alpha U_\alpha\times V\to E\quad\text{and}\quad \pi_M:E\to M$$

be the canonical quotient and projection maps. Now consider some index α and O a subset of $U_\alpha\times V$. Then we see that

$$\pi_E^{-1}(\pi_E(O))=\bigsqcup_\beta h^\alpha_\beta(O),$$

where

$$\begin{aligned} h^\alpha_\beta : U_{\alpha\beta}\times V &\to U_{\alpha\beta}\times V\\ (x,v)&\mapsto (x,g^\alpha_\beta(x)v). \end{aligned}$$

So we see that if O is open then $\pi_E^{-1}(\pi_E(O))$ is an open subset of $\bigsqcup_\alpha U_\alpha\times V$. So we have that π_E is a continuous mapping. Since the restriction to the sets $U_\alpha\times V$ is injective we have that

$$\pi_E^\alpha=\pi_E|^{-1}_{U_\alpha\times V}:\pi_E(U_\alpha\times V)\to U_\alpha\times V$$

form a chart for E. The overlap maps are h^α_β. So we see that E has a smooth structure. Furthermore, π_M is smooth with respect to this smooth structure, since it induces projection maps on all of the charts. Now let

$$\pi_\alpha:\pi_M\circ\pi_E^\alpha:U_\alpha\times V\to U_\alpha.$$

These diffeomorphisms on each U_α induce a vector space structure on each fiber when restricted to a point. Since the restriction of the overlap maps h^α_β induce a vector space isomorphism we have that the vector space structures induced by π^α_E and π^β_E are the same. So we have that the fiber bundle diagram commutes and hence we have shown that E is indeed a vector bundle over M.

A very similar idea can be done for arbitrary fiber bundles by taking transition functions from the overlaps to $\mathrm{End}(X)$ for any object X.

Proposition 1.70. *If (E,M,π,V) is a vector bundle and (g^α_β) is a co-cycle obtained from local trivializations of E then the vector bundle obtained from (g^α_β) is isomorphic to E.*

Proof. Consider (g^α_β) the transition data arising from the local trivializations of E. Let E' be the vector bundle obtained from (g^α_β). Define $f:E\to E'$ by

$$f(v)=[\alpha,g_\alpha(v)]\quad\text{if }\pi(v)\in U_\alpha.$$

If $\pi(v)\in U_\alpha\cap U_\beta$ then

$$[\beta,g_\beta(v)]=[\alpha,g^\alpha_\beta(g_\beta(v))]=[\alpha,g_\alpha(v)]\in E'.$$

This shows us that the map f is well-defined since it does not depend on α. We clearly see that $\pi_M\circ f=\pi$.

Since the map

$$\left(\pi_E^\alpha\right)^{-1}\circ f\circ g_\alpha^{-1}:U_\alpha\times V\to U_\alpha\times V$$

is the identity and this smooth, we have that f is a smooth map. Since the restrictions of g_α and π^E_α are vector space isomorphisms restricted to each fiber, we have that so is f. So we see that f is indeed a vector space isomorphism. □

1.7.2. *Vector bundle constructions.*

Definition 1.71. If $f:M\to N$ is a smooth map and (E,N,π,F) is a vector bundle, then the **pullback** of E via f is the vector bundle

$$f^*E:=\{(x,v)\in M\times F\ :\ f(x)=\pi(v)\}.$$

⋄

Definition 1.72. If $E \to M$ and $F \to M$ are two vector bundles over M, then $\mathrm{Hom}(E,F)$ is the vector bundle over M defined by

$$\mathrm{Hom}(E,F) := \{(x,\lambda) \,:\, x \in M, \lambda \in \mathrm{Hom}(E_x, F_x)\}.$$

The **dual vector bundle** of E is the bundle $E^* := \mathrm{Hom}(E, \mathbf{R})$ (or $\mathrm{Hom}(E,\mathbb{C})$ if E is complex). ⋄

Proposition 1.73. *$\mathrm{Hom}(E,F)$, $E \otimes F$, $\bigwedge^i E$, $E^{\otimes k}$ are indeed smooth vector bundles.*

Proof. We will show how to construct the tensor product and exterior product in a separate manner later as well. Here we do not provide an explanation of these spaces, but they will be introduced later. Anyone not familiar with this material already can freely skip it.

- If $\pi_E : E \to M$ and $\pi_F : F \to M$ are a smooth vector bundles over M we can form $\mathrm{Hom}(E,F)$ such that
$$\mathrm{Hom}(E,F)_m = \mathrm{Hom}(E_m, F_m) \quad \text{for all } m \in M.$$
The topology and smooth structure are determined by requiring that if $s_1 \in \Gamma(E)$ and $s_2 \in \Gamma(\mathrm{Hom}(E,F))$ then $s_2 \circ s_1 \in \Gamma(F)$.
- If $\pi_E : E \to M$ and $\pi_F : F \to M$ are smooth vector bundles over M, we can form their tensor product, $E \otimes F$, such that
$$(E \otimes F)_m = E_m \otimes F_m \quad \text{for all } m \in M.$$
The topology and smooth structure on $E \otimes F$ are determined from those of E and F by requiring that if $s \in \Gamma(E)$ and $s' \in \Gamma(F)$ then $s \otimes s' \in \Gamma(E \otimes F)$.
- If $\pi : E \to M$ is a smooth vector bundle, we can form the k^{th} exterior power $\bigwedge^k E$ so that
$$(\textstyle\bigwedge^k E)_m = \bigwedge^k E_m \quad \text{for all } m \in M.$$
The topology and smooth structure on $\bigwedge^k E$ are determined by requiring that if $s_1, \ldots, s_k \in \Gamma(E)$ then $s_1 \wedge \cdots \wedge s_k \in \Gamma(\bigwedge^k E)$.
- Follows immediately from the second bullet point.

□

1.7.3. *Tangent bundle.*

Definition 1.74. Let $(M, \mathscr{A})$ be a manifold with $\mathscr{A} = \{\varphi_\alpha : U_\alpha \to \mathbf{R}^n\}$, then we define the **tangent bundle** as

$$TM := \bigsqcup_{p \in M} \{p\} \times T_p M.$$

⋄

Now we will show that TM is indeed a vector bundle over M.

Lemma 1.75. *TM is a smooth $2n$-dimensional manifold.*

Proof. Consider a differentiable manifold M of dimension n with atlas $\mathscr{A}_M = \{(U_i, \varphi_i)\}_{i \in I}$, where I is some index set. For each $i \in I$ define the mapping $\Phi_i : U_i \times \mathbf{R}^n \to TM$ as follows:

$$\Phi_i(\vec{x}, a^1, \ldots, a^n) = \left(\phi_i(\vec{x}), a^i \frac{\partial}{\partial x_i}\right),$$

where $\frac{\partial}{\partial x_i}$ are the basis elements of the tangent plane using coordinate chart (U_i, ϕ_i). Now we want to show that $\mathscr{A}_{TM} = \{(U_i \times \mathbf{R}^n, \Phi_i)\}_{i \in I}$ is indeed an atlas for TM.

(1) First we show that $\mathscr{A}_{TM}$ covers the tangent bundle. Notice that
$$\bigcup_{i \in I} \Phi_i(U_i \times \mathbf{R}^n) = \{(p,v) : p \in M, v \in T_p M\} = TM,$$
as desired.

(2) Next we must show that the transition maps are differentiable. Consider a tangent vector at the point $p \in M$ that are represented in two separate coordinate charts (U, φ) and (V, ψ) as $a^i \frac{\partial}{\partial x_i}$ and $b^j \frac{\partial}{\partial y_j}$, respectively. Let Φ and Ψ be the corresponding coordinate charts into the tangent bundle. Now can consider the transition map

$$\Psi^{-1} \circ \Phi(x_1, \ldots, x_n, a^1, \ldots, a^n) = \left(\psi^{-1} \circ \phi(x_1, \ldots, x_n), \sum_{i=1}^n a^i \frac{\partial y_1}{\partial x_i}, \ldots, \sum_{i=1}^n a^n \frac{\partial y_n}{\partial x_i}\right).$$

Note that in the second n coordinates we are just pulling the coefficient from the vector that we get from the change of basis. Note that the first n coordinates are differentiable since M is a differentiable manifold, and the second n coordinates are also differentiable since they are each sums of smooth functions. So we have that $\Psi^{-1} \circ \Phi$ is a differentiable mapping. Since the choice of coordinates was arbitrary, we have that all of the transition maps are differentiable mappings.

Since both of these properties are satisfied, we have shown that TM does indeed have a differentiable manifold structure. □

Theorem 1.76. *The tangent bundle*

$$\mathbf{T}M := (TM, M, \pi, T_pM),$$

where $\pi : TM \to M$ is defined via and (by abusing notation) we write T_pM instead of $\mathbf{R}^n$ to emphasize that the fiber is indeed the tangent space, is a smooth vector bundle.

Definition 1.77. We define the **cotangent bundle** as

$$TM^* := \mathrm{Hom}(TM, \mathbf{R}) = \bigsqcup_{p \in M} \{p\} \times T_pM^*.$$

⋄

Remark 1.78. We will often just use the notation TM or TM^* to denote all of the data specified in the tangent bundle $\mathbf{T}M$ or cotangent bundle $\mathbf{T}M^*$ respectively. ⋄

Definition 1.79. If $f : M \to N$ is a smooth map, then $\mathrm{d}f : TM \to TN$ defined by $\mathrm{d}f := \mathrm{d}f_p$ on T_pM is called the **derivative** of f. ⋄

Proposition 1.80. *The derivative* $\mathrm{d}f$ *is a vector bundle map over* f*.*

Proof. Since f is differentiable, by definition we have that there are atlases $\{(U_\alpha, \varphi_\alpha)\}$ and $\{(V_\beta, \psi_\beta)\}$ for M and N, respectively, such that the following compositions are smooth:

$$\mathbf{R}^m \supseteq \varphi_\alpha(f^{-1}(V_\beta) \cap U_\alpha) \xrightarrow[\cong]{\varphi_\alpha^{-1}} f^{-1}(V_\beta) \cap U_\alpha \xrightarrow{f} V_\beta \xrightarrow[\cong]{\psi_\beta} \psi_\beta(V_\beta) \subseteq \mathbf{R}^n.$$

So we obtain a commutative diagram

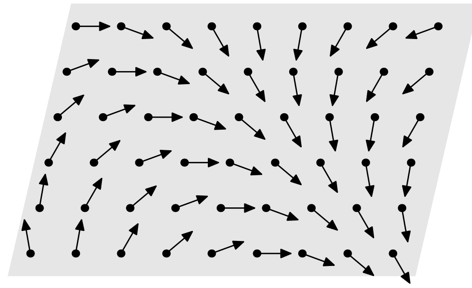
$$\begin{tikzcd}
\mathrm{d}f^{-1}(T(V_\beta)) \cap T(U_\alpha) = T(f^{-1}(V_\beta) \cap U_\alpha) \arrow[r, "T_{\varphi_\alpha}"] \arrow[d, "\mathrm{d}f"] & T(\varphi_\alpha(f^{-1}(V_\beta) \cap U_\alpha)) \cong \varphi_\alpha(f^{-1}(V_\beta) \cap U_\alpha) \times \mathbf{R}^m \arrow[d, "\Phi"] \\
T(V_\beta) \arrow[r, "T_{\psi_\beta}"] & T(\psi_\beta(V_\beta)) \cong \psi_\beta(V_\beta) \times \mathbf{R}^n
\end{tikzcd}$$

where

$$\Psi(p, \boldsymbol{v}) := (\psi_\beta \circ f \circ \varphi_\alpha^{-1}(p), D(\psi_\beta \circ f \circ \varphi_\alpha^{-1})|_p(\boldsymbol{v})),$$

and T_{φ_α} and T_{ψ_β} are the charts of TM and TN, respectively. So we see that $\mathrm{d}f$ is differentiable. Since $\mathrm{d}f$ restricted to each fiber is a linear transformation we see that $\mathrm{d}f$ is indeed a morphism of tangent bundles over f, as desired. □

Definition 1.81. A **vector field** on M is a smooth section of the tangent bundle $X \in \Gamma(TM)$. We denote $\Gamma(TM)$ as $\mathfrak{X}(M)$.



⋄

Intuitively, we can just think of this as a vector attached to each point of the manifold, which varies smoothly.

Lemma 1.82. *The space $\mathfrak{X}(M)$ is canonically isomorphic to the space of derivations on the $\mathscr{C}^\infty(M)$ algebra. That is the space of all linear operators $D : \mathscr{C}^\infty(M) \to \mathscr{C}^\infty(M)$ such that*

$$D(fg) = D(f)g + fD(g).$$

Proof. Consider some vector field $X \in \mathfrak{X}(M)$. We then clearly have that X defines a derivation $X : \mathscr{C}^\infty(M) \to \mathscr{C}^\infty(M)$ defined via

$$X(f)(p) := X(p)(f) = df(X(x)).$$

Now suppose we have an arbitrary derivation $D : \mathscr{C}^\infty(M) \to \mathscr{C}^\infty(M)$. Now for all $p \in M$ define $D_p : \mathscr{C}^\infty(M) \to \mathbf{R}$ defined by

$$D_p(f) = D(f)(p).$$

So we have that $D_p \in T_pM$. In this way we obtain a section X, defined by the mapping $p \mapsto (p, D_p)$, of TM. The fact that X is smooth follows directly from the fact that the derivations are over the algebra $\mathscr{C}^\infty(M)$. Indeed the vector mapping is smooth. □

Remark 1.83. The above result is quite subtle. Rather than needing to show any results about vector fields pointwise we can consider them globally as a derivation over $\mathscr{C}^\infty(M)$. This will be a useful tool in elegantly constructing new vector fields given existing ones. We will also use X to represent the derivation corresponding to the vector field $X \in \mathfrak{X}(M)$. ⋄

Notation 1.84. We will use the notation $X(p)$ and $X|_p$ interchangably to represent a vector field at a point p. The second notation is sometimes nicer when we want to treat our tangent vectors as derivations.

Definition 1.85. Let $f : M \to N$ be a smooth function between manifolds. Let $X \in \mathfrak{X}(M)$ and $Y \in \mathfrak{X}(M)$. We say that Y is f**-related** to X if

$$W|_{f(p)} = df \circ V|_p.$$

⋄

Lemma 1.86. *Let $X, Y \in \mathfrak{X}(M)$. There is a a unique vector field $[X, Y] \in \mathfrak{X}(M)$ such that $[X, Y](f) = X(Y(f)) - Y(X(f))$ for all $f \in \mathscr{C}^\infty(M)$.*

Proof. Note that it is clear that $[X, Y]$ is linear. It remains to check that it satisfies the Leibniz rule:

$$\begin{aligned}
[X, Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X(gY(f) + fY(g)) - Y(gX(f) + fX(g)) \\
&= X(g)Y(f) + X(f)Y(g) + gX(Y(f)) + fX(Y(g)) \\
&\quad - Y(g)X(f) - Y(f)X(g) - gY(X(f)) - fY(X(g)) \\
&= g[X, Y](f) + f[X, Y](g).
\end{aligned}$$

So by the identification in the previous lemma we have that $[X, Y] \in \mathfrak{X}(M)$. □

Definition 1.87. Given two vector fields $X, Y \in \mathfrak{X}(M)$ we define the **Lie Bracket** $[X,Y]$ implicitly as the derivation:

$$[X,Y](f) := X(Y(f)) - Y(X(f))$$

$\diamond$

Note that since in local coordinates $X(Y(f))$ involves derivatives of order higher than 1, we do not necessarily have that $X(Y(f))$ is a derivation over $\mathscr{C}^\infty(M)$, i.e. there is no reason to believe that $X(Y(f))$ is a vector field at all. So it was necessary to show that the Lie Bracket is indeed a vector field.

Theorem 1.88. *Suppose that $X, Y, Z \in \mathfrak{X}(M)$, $a, b \in \mathbf{R}$, and $f, g \in \mathscr{C}^\infty(M)$. Then:*

(a) $[X,Y] = -[Y,X]$
(b) $[aX + bY, Z] = a[X,Z] + b[Y,Z]$
(c) $[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$
(d) $[fX, gY] = fg[X,Y] + fX(g)Y - gY(f)X$.

Proof.

(a) This is immediate from the definition.
(b) This is immediate from the definition.
(c) Note that

$$[[X,Y],Z] = [XY - YX, Z] = X \circ Y \circ Z - Y \circ X \circ Z - Z \circ X \circ Y + Z \circ Y \circ X.$$

On the other hand, we have that

$$\begin{aligned}[X,[Y,Z]] + [Y,[X,Z]] = X \circ Y \circ Z - X \circ Z \circ Y - Y \circ Z \circ X + Z \circ Y \circ X \\ + Y \circ Z \circ X - Y \circ X \circ Z - Z \circ X \circ Y + X \circ Z \circ Y.\end{aligned}$$

So we have that

$$[[X,Y],Z] = [X,[Y,Z]] + [Y,[Z,X]].$$

Now by using (a) we have the desired Jacobi identity

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0.$$

(d) We can compute

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

$\square$

Proposition 1.89. *Let $f : M \to N$ be a smooth map between manifolds. If $X, Y \in \mathfrak{X}(M)$ and $W, Z \in \mathfrak{X}(N)$ are such that W is f-related to X and Z is f-related to Y then for every $p \in M$ we have*

$$[W,Z]_{f(p)} = df_p[X,Y]|_p.$$

Proof. Consider any smooth function $g : N \to \mathbf{R}$ and fix $p \in M$. Let $q = f(p) \in N$. Then

$$\begin{aligned}[W,Z]|_q(g) &= W|_y(Zg) - Z|_y(Wg) \\ &= df_p(X|_p)(Zg) - df_p(Y|_p)(Wg) \\ &= X|_p((Zg) \circ f) - Y|_p((Wg) \circ f) \\ &= X|_p(Y(g \circ f)) - Y|_p(X(g \circ f)) \qquad = [X,Y]|_p(g \circ f) \\ &= df_p([X,Y]|_p)f.\end{aligned}$$

Here we simply used the f-relatedness of the vector fields to deduce that

$$(Zg) \circ f|_p = (Zg)|_{f(p)} = df_p(Y|_p)g = Y|_p(g \circ f),$$

and similarly that $(Wg) \circ f = X(g \circ f)$. □

1.7.4. *Flows and integral curves.* Since tangent vectors are also velocities to curves, we would hope that our vector fields had a similar interpretation. A curve $c(t)$ such that

$$\dot{c}(t) = X(c(t))$$

is called an **integral curve** for X. Given an initial value $p \in M$, there is in fact a unique integral curve $c(t)$ such that $c(0) = p$ and it is defined on some maximal interval I that contains 0 as an interior point. Indeed, in local coordinates we just have a system of ordinary differential equations. So by the theorem of Cauchy-Lipschitz-Picard-Lindelöf we have local existence and uniqueness (uniqueness follows from the fact that X is a *smooth* vector field on M). To obtain a maximal interval we need to use the local uniqueness of solutions and patch them together through a covering of coordinate charts.

Now we will use the general notation that $\Phi_t^X(p) = c_p(t)$ is the flow corresponding to the vector field X; that is to say

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^X = X \circ \Phi_t^X.$$

We will also write the flow as $\Phi^X(t, p)$ indicate this same flow

Many of the results that hold for ODEs in Euclidean space hold for ODEs on manifolds since we can always pullback the differential equation into charts, solve it there, and push it back onto the manifold. We now state and prove some of the most important ones.

Theorem 1.90. *Let $X \in \mathfrak{X}(M)$. The associated flow map $\Phi^X : \mathrm{Dom}(X) \to M$ is smooth where*

$$\mathrm{Dom}(X) := \bigcup_{p \in M} I_p \times \{p\},$$

where I_p is the maximal interval of existence for the flow originating from p. Furthermore, we have that

$$\Phi^X(t + s, p) = \Phi^X(t, \Phi^X(s, p))$$

in the sense that whenever the right hand side of the equation exists, so does the left and equality holds.

Proof. Smoothness of the flow map is straightforward from ordinary differential equation theory.

Now suppose that the right hand side of the above equation exists. Consider the following ordinary differential equation:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\left(\Phi^X(t+s,p)\right) = X(\Phi^X(t+s,p)) & \text{for } t > 0, \\ \Phi^X(t+s,p)|_{t=0} = \Phi^X(s,p). \end{cases}$$

Note that $\Phi^X(t, \Phi^X(s, p))$ is the *unique* solution of the above ODE and so we have that the left hand side exists and the equality holds. □

Definition 1.91. Consider a smooth vector field $X \in \mathfrak{X}(M)$. The flow map Φ^X is called **complete** or **global** if its domain of definition $\mathrm{Dom}(X)$ is $\mathbf{R} \times M$. In this case we say that X is a **complete vector field**. ⋄

Lemma 1.92. *A vector field with compact support on M is complete.*

Proof. Consider a smooth compactly supported vector field $X \in \mathfrak{X}(M)$ such that $\mathrm{supp}(X)$ is compact. Note that the flow map starting from points in $\mathrm{supp}(X)$ is defined for some open interval around 0 since X is compactly supported. So we can find some $\varepsilon > 0$ such that

$$[-\varepsilon, \varepsilon] \times \mathrm{supp}(X) \subset \mathrm{Dom}(X)$$

Now consider some point $p \notin \mathrm{supp}(X)$. Since $X(p) = 0$ we have that $\Phi^X(t, p) = p$ for all $t \in \mathbf{R}$ and so we have that $\mathbf{R} \times \{p\} \subset \mathrm{Dom}(X)$. Putting these two notions together we have that

$$[-\varepsilon, \varepsilon] \times M \subset \mathrm{Dom}(X).$$

By the previous lemma we have that

$$\Phi^X(t + \varepsilon, p) = \Phi^X(t, \Phi^X(\varepsilon, x))$$

for all $|t| \leq \varepsilon$, and so
$$[-2\varepsilon, 2\varepsilon] \times M \subset \mathrm{Dom}(X).$$
By repeating this process inductively we obtain that
$$\mathbf{R} \times M = \mathrm{Dom}(X).$$
□

In particular, we have that every vector field on a compact manifold is complete.

We are now in a position to define the *Lie derivative* of a vector field.

Definition 1.93. Let X and Y be vector fields on M, and $p \in M$. Let Φ_t be the flow of the vector field X. The **Lie derivative** of Y with respect to X is
$$\mathscr{L}_X Y|_p = \lim_{t \to 0} \frac{\mathrm{d}\Phi_{-t} Y|_{\Phi_t(p)} - Y|_p}{t} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d}\Phi_{-t} Y|_{\Phi_t(p)}.$$
⋄

Proposition 1.94. $\mathscr{L}_X Y = [X, Y]$.

Proof. Recall that for a function $f : M \to \mathbf{R}$ we have that
$$Xf(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\Phi_t(p)).$$
Now let Ψ_s be the flow of the vector field Y, and note that
$$\mathscr{L}_X Y|_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} Y|_{\Phi_t(p)}(f \circ \Phi_{-t}) = \left.\frac{\partial^2 \mathscr{H}}{\partial t \partial s}\right|_{(t,s)=(0,0)},$$
where
$$H(t,s) := f(\Phi_{-t}(\Psi_s(\Phi_t(p)))).$$
This is clear since
$$\left.\frac{\partial H}{\partial s}\right|_{(t,0)} = Y|_{\Phi_t(p)}(f \circ \Phi_{-t}).$$
Now define $K(r,s,t) = f(\Phi_r(\Psi_s(\Phi_t(p))))$. Then $H(t,s) = K(-t,s,t)$, and so
$$\left.\frac{\partial^2 H}{\partial t \partial s}\right|_{(0,0)} = -\left.\frac{\partial^2 K}{\partial r \partial s}\right|_{(0,0,0)} + \left.\frac{\partial^2 K}{\partial t \partial s}\right|_{(0,0,0)}.$$
Now we compute
$$\begin{aligned} \left.\frac{\partial K}{\partial r}\right|_{(0,s,0)} &= Xf(\Psi_s(p)), \\ \left.\frac{\partial^2 K}{\partial r \partial s}\right|_{(0,0,0)} &= YXf(p), \\ \left.\frac{\partial K}{\partial s}\right|_{(0,0,t)} &= Yf(\Phi_t(p)), \\ \left.\frac{\partial^2 K}{\partial t \partial s}\right|_{(0,0,0)} &= XYf(p). \end{aligned}$$
Now the result follows since
$$\left.\frac{\partial^2 H}{\partial t \partial s}\right|_{(0,0)} = [X,Y]f(p).$$
□

1.8. **Immersions, submersions, and embeddings.** There are many different notions of submanifolds.

Definition 1.95. A smooth map $f : M \to N$ is called an **immersion** if for every $p \in M$ we have that $df_p : T_pM \to T_{f(p)}N$ is injective. In this case we call M an immersed submanifold of N. ⋄

Since a linear map, L, is injective if and only if $\ker L = \{0\}$ the main takeaway is that an immersion does not squash tangent vectors – that is to say non-zero tangent vectors get pushed forward to non-zero tangent vectors.

Example 1.96. Consider the map $f : \mathbf{S}^1 \to \mathbf{R}^2$ given in polar coordinates by

$$f(\theta) = (r(\theta), \theta), \quad \text{where} \quad r(\theta) = \cos 2\theta.$$

It is clear that $f(\mathbf{S}^1)$ is an immersion, but not a submanifold of $\mathbf{R}^2$. This is intuitively clear since any neighborhood of $\mathbf{0}$ in $\mathbf{R}^2$ intersects $f(\mathbf{S}^1)$ is a set *with corners*, which is clearly not diffeomorphic to any open interval.

In this case we say that f is an **immersion with self-intersections**. ⋄

Definition 1.97. A smooth map $f : M \to N$ is called an **embedding** if it is an injective immersion and the topology of M coincides with the induced topology of $f(M) \subset N$. In this case we call M an **embedded submanifold** of N. ⋄

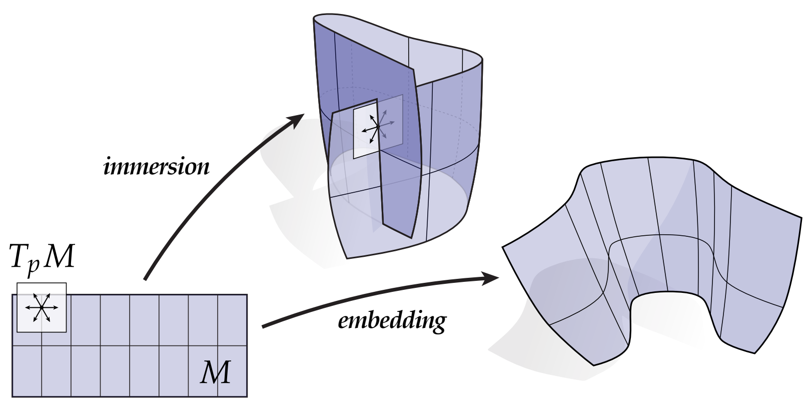


FIGURE 4. An immersion is still a manifold of the same dimension, while an embedding has the additional requirement of being a homeomorphism onto its image.

Definition 1.98. A subset M of a manifold N is called a **submanifold** of dimension m if for each $p \in M$ there exists a chart $\varphi : U \to \widetilde{U}$ of N around p such that $\varphi(M \cap U)$ is an open subset of an m-dimensional linear subspace of $\mathbf{R}^n$. ⋄

Definition 1.99. A smooth map $f : M \to N$ is a **submersion** if for all $p \in M$ we have that $df_p : T_pM \to T_pN$ is surjective. ⋄

1.9. **Lie groups and Lie algebras.** Lie groups are simply manifolds which also have a group structure. In particular, they are internal groups in the category of smooth manifolds.

Definition 1.100. A **Lie group** is a group $(G, \cdot)$ such that G is a smooth manifold, $m : G \times G \to G$ given by $m(g, h) = g \cdot h$ is a smooth map, and such that $i : G \to G$ given by $i(g) = g^{-1}$ is smooth. ⋄

Exercise 1.101. Show that if $m : G \times G \to G$ is smooth then $i : G \to G$ is also smooth.

Definition 1.102. Let G and H be Lie groups, a **Lie group homomorphism** from G to H is a smooth group homomorphism $\varphi : G \to H$. ⋄

Example 1.103. $\mathbf{S}^1 = \mathbf{R}/\mathbb{Z}$ is a Lie group. $\mathbf{S}^1 \subseteq \mathbf{C} \setminus \{0\}$ is also a Lie group (where $\mathbf{C} \setminus \{0\}$ is endowed with complex multiplication). ⋄

Example 1.104. We provide some examples of matrix Lie groups.

- The **general linear group**: $\mathbf{GL}(n,\mathbf{R}) := \{A \in \mathbf{R}^{n\times n} \,:\, \det A \neq 0\}$.

  Note that $\mathbf{GL}(n,\mathbf{R})$ is a group since the determinant is multiplicative, and so if $A, B \in \mathbf{GL}(n,\mathbf{R})$ we see that $\det(AB) = \det(A)\det(B) \neq 0$; hence $AB \in \mathbf{GL}(n,\mathbf{R})$. Similarly, we obtain that $\det(A^{-1}) = \frac{1}{\det A} \neq 0$, and so $\mathbf{GL}(n,\mathbf{R})$ is closed under inverses.

  Note that we can view $\mathbf{GL}(n,\mathbf{R})$ embedded in $\mathbf{R}^{n\times n}$. Note that since $\det : \mathbf{GL}_n(\mathbf{R}) \to \mathbf{R}^\times$ is a polynomial in the n^2 entries of the matrices, we have that det is continuous. In particular, we see that
  $$\mathbf{GL}(n,\mathbf{R}) = \det{}^{-1}(\mathbf{R} \setminus \{0\}),$$
  and so $\mathbf{GL}(n,\mathbf{R})$ is an open subset of $\mathbf{R}^{n\times n}$. Hence $\mathbf{GL}(n,\mathbf{R})$ is a smooth manifold. Furthermore, note that since $(A,B) \to AB$ is a polynomial in all of the coordinates we deduce that the multiplication is smooth, and hence $\mathbf{GL}(n,\mathbf{R})$ is a Lie group. It is not hard to see that $\dim \mathbf{GL}(n,\mathbf{R}) = n^2$.

- The **special linear group**: $\mathbf{SL}(n,\mathbf{R}) := \{A \in \mathbf{GL}(n,\mathbf{R}) \,:\, \det A = 1\}$.

  Clearly $\mathbf{SL}(n,\mathbf{R})$ is a subgroup of $\mathbf{GL}(n,\mathbf{R})$. By identical reasoning we see that $\mathbf{SL}(n,\mathbf{R})$ is a Lie group. The rest of the matrix subgroups of $\mathbf{GL}(n,\mathbf{R})$ are Lie groups for exactly the same reason. Note that $\mathbf{SL}(n,\mathbf{R})$ is closed, as it is the preimage of $\{1\}$ under det.

- The **orthogonal group**: $\mathsf{O}(n) := \{A \in \mathbf{GL}(n,\mathbf{R}) \,:\, A^\top A = \mathbb{1}_{n\times n}\}$.
- The **special orthogonal group**: $\mathbf{SO}(n) := \{A \in \mathsf{O}(n) \,:\, \det A = 1\}$.
- The **unitary group**: $\mathsf{U}(n) := \{A \in \mathbf{GL}(n,\mathbf{C}) \,:\, TT^* = \mathbb{1}_{n\times n}\}$

⋄

Definition 1.105. Let M be a manifold and G a Lie group. A **Lie group action** of G on M is a smooth action $\rho : G \times M \to M$. ⋄

Note that if ρ is a Lie group action of G on M, then for any $g \in M$ the map $\rho_g := \rho(g,\cdot) : M \to M$ is a diffeomorphism. In this way we can think of ρ as a representation of a subgroup of $\mathrm{Diff}(M)$.

Example 1.106. If G is a Lie group, then G acts on itself by left multiplication, right multiplication, and conjugation:
$$\begin{aligned} L &: G \times G \to G, & (g,h) &\mapsto g\cdot h,\\ R &: G \times G \to G, & (g,h) &\mapsto h\cdot g^{-1},\\ C &: G \times G \to G, & (g,h) &\mapsto g\cdot h\cdot g^{-1}. \end{aligned}$$

⋄

Definition 1.107. A **Lie algebra** is a $\mathbb{F}$-vector space, $\mathfrak{g}$, endowed with an operation $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which is

(i) *bilinear*: for all $a,b,c \in \mathbb{F}$ and $x,y,z \in \mathfrak{g}$
$$[ax+by,z] = a[x,z]+b[y,z] \qquad \text{and} \qquad [x,by+cz] = b[x,y]+c[x,z].$$
(ii) *alternates on* $\mathfrak{g}$: for all $x \in \mathfrak{g}$
$$[x,x] = 0.$$
(iii) *satisfies the Jacobi identity*: for all $x,y,z \in \mathfrak{g}$
$$[[x,y],z]+[[y,z],x]+[[z,x],y] = 0.$$

⋄

Note that bilinearity and the "alternating on $\mathfrak{g}$" properties imply that the Lie bracket is anticommutative.

Example 1.108. The following are Lie algebras over $\mathbf{R}$.

- The set of vector fields on M, $\mathfrak{X}(M)$, endowed with the Lie bracket $[X,Y] = XY - YX$.
- $\mathbf{R}^3$ endowed with the Lie bracket $[x,y] = x \times y$.
- $\mathbf{R}^{n\times n}$, endowed with the Lie bracket $[A,B] = AB - BA$.

- $\mathfrak{sl}(n) := \{A \in \mathbf{R}^{n\times n} : \operatorname{tr}(A) = 0\}$. The Lie bracket is that of $\mathbf{R}^{n\times n}$.
- $\mathfrak{so}(n) := \{A \in \mathbf{R}^{n\times n} : A^\top = -A\}$. The Lie bracket is the commutator (i.e. the Lie bracket of $\mathbf{R}^{n\times n}$)

◇

Every Lie group has an associated Lie algebra that in some sense describes the infinitesimal changes in the Lie group.

Definition 1.109. Let G be a Lie group, and define the **associated Lie algebra**

$$\mathfrak{g} := T_e G,$$

where $e \in G$ is the identity element. ◇

We still need to endow $\mathfrak{g}$ with a Lie bracket and prove that this makes $\mathfrak{g}$ a Lie algebra.

Example 1.110. If $G = \mathbf{GL}(n, \mathbf{R})$ then $g = \mathfrak{gl}(n) := \operatorname{End}(\mathbf{R}^n)$. ◇

Definition 1.111. A vector field $X \in \mathfrak{X}(G)$ is **left-invariant** if

$$dL_g(X|_h) = X|_{gh}$$

for all $g, h \in G$. The space of left-invariant vector spaces is denoted $\operatorname{Lie}(G)$. ◇

Proposition 1.112. *Let G be a Lie group. $\mathfrak{g}$ is canonically isomorphic to* $\operatorname{Lie}(G)$*;* $\operatorname{Lie}(G)$ *is closed under the Lie bracket, and makes it a Lie algebra of* $\dim G$*.*

Proof. Fix $v \in \mathfrak{g}$. Now define $X \in \mathfrak{X}(G)$ pointwise as

$$X|_g = dL_g(v).$$

Conversely, every left invariant vector field satisfies the above equality. So we see that $\mathfrak{g}$ is canonically isomorphic to $\operatorname{Lie}(G)$. Hence $\dim \operatorname{Lie}(G) = \dim G$. Now consider two left invariant vector fields $X, Y \in \operatorname{Lie}(G)$, and note that by definition X is dL_g-related to X and Y is dL_g-related to Y for all $g \in G$. In light of Proposition 1.89 we deduce that

$$dL_g[X, Y] = [dL_g X, dL_g Y],$$

and so the Lie bracket $[X, Y]$ is left-invariant. Finally, we see that $\operatorname{Lie}(G)$ is a Lie algebra by Theorem 1.88. □

Note that since $\mathfrak{g}$ is canonically isomorphic to $\operatorname{Lie}(G)$, we have that $\mathfrak{g}$ is a Lie algbera. Explicitly we have the Lie bracket on $\mathfrak{g}$ given as follows: for $v, w \in \mathfrak{g}$ let X and Y be the associated left-invariant vector fields on G. Then we define

$$[\![v, w]\!] = [X, Y]|_e.$$

Example 1.113. We compute the Lie Algebra of $\mathbf{GL}(n, \mathbf{R})$. Since $\mathbf{GL}(n, \mathbf{R})$ is an open submanifold of $\mathbf{R}^{n\times n}$, we know that the tangent space of $\mathbf{GL}(n, \mathbf{R})$ at any matrix A is the same as its tangent space in $\mathbf{R}^{n\times n}$. Furthermore, since $\mathbf{R}^{n\times n} \cong \mathbf{R}^{n^2}$ we can identify $\mathbf{R}^{n\times n} \cong T_A \mathbf{R}^{n\times n}$. In particular, we can identify $\mathfrak{gl}(n, \mathbf{R}) \cong \mathbf{R}^{n\times n}$ as *vector spaces*.

Since $\mathbf{R}^{n\times n}$ is an associative algebra, we know that it automatically has a bracket given by the commutator. We claim that this bracket is the same as that in $\mathfrak{gl}(n, \mathbf{R})$ under the natural isomorphism stated above. Fix $A \in \mathbf{R}^{n\times n}$ and let $X(I) \cong A \in T_I \mathbf{GL}(n, \mathbf{R})$. Now let X be the associated left-invariant vector field by setting

$$X(g) = dL_g X(I) \cong dL_G A \cong gA.$$

Note that under this isomorphism the left-translation is within the larger manifold $\mathbf{R}^{n\times n}$ instead of $\mathbf{GL}(n, \mathbf{R})$. Now repeat this process for another matrix $B \in \mathbf{R}^{n\times n}$ to obtain another left-invariant vector field Y. Now we compute the Lie bracket $[X, Y]$. Let $x^{ij} : \mathbf{GL}(n, \mathbf{R}) \to \mathbf{R}$ be the function sending a matrix to it's (i, j) element. Now we compute

$$Yx^{ij}(g) = Y|_g x^{ij} \cong (gB)x^{ij} = x^{ij}(gB).$$

So we see that $Yx^{ij} = x^{ij} \circ R_B$. Now we apply this function to the vector $X(I) = A$. We associate to A the curve $\gamma(t) = I + tA$, and we find that

$$X|_I Y x^{ij} = \dot{\gamma}(0)(x^{ij} \circ R_B) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left(x^{ij}(B + tAB)\right) = (AB)_{i,j}.$$

Similarly, we see that $Y|_I X x^{ij} = (BA)_{i,j}$. So we conclude that

$$[X,Y]|_I x^{ij} = (AB-BA)_{i,j} =\cong (AB-BA)(x^{ij}).$$

Now we see that for any $M \in \mathbf{R}^{n\times n}$ that we can write in coordinates (by considering this as a tangent space)

$$Q = \sum_{i,j} Q(x^{i,j}) \frac{\partial}{\partial x^{ij}}.$$

Since we have that $[A,B](x^{i,j}) = (AB-BA)x^{i,j}$ we see that these two vectors have all the same components, and hence are equal. Finally, we conclude that $\mathfrak{gl}(n,\mathbf{R}) \cong \mathbf{R}^{n\times n} \cong \mathrm{End}(n,\mathbf{R})$ as Lie algebras. ⋄

Definition 1.114. The **adjoint representation** $\mathrm{Ad} : G \to \mathbf{GL}(\mathfrak{g})$ is defined via

$$\mathrm{Ad}(g)v = (dC_g)v,$$

for $v \in T_eG$, and where C_g is the action of G on itself by conjugation. ⋄

Before moving forward, we try and give some intuition about the adjoint representation. Let $G = \mathbf{GL}(n,\mathbf{R})$. Now for any $X \in \mathfrak{g} = \mathfrak{gl}(n) = \mathbf{R}^{n\times n}$, a parametric curve in G through the identity with velocity vector X at $t = 0$ is given by $\gamma(t) := \exp(tX)$, where exp is the usual matrix exponential. So we see that the adjoint representation for some $g \in G$ is given by

$$\mathrm{Ad}(g)X = (dC_g)X = (dC_g)\dot\gamma(0),$$

and so it is the velocity at $t = 0$ of the parametric curve $(C_g \circ \gamma)(t) = g\exp(tX)g^{-1}$. Note that

$$(C_g \circ \gamma)(t) = 1 + gtXg^{-1} + \sum_{i\geq 2} \frac{t^i}{i!} gX^i g^{-1}.$$

Now clearly $(C_g \circ \gamma)'(0) = gXg^{-1}$, and so

$$\mathrm{Ad}(g)X = gXg^{-1} = (C_g)X.$$

This is clearly smooth in g. We can use a similar technique to show compute the differential $d\,\mathrm{Ad}$. Again, let $X \in \mathfrak{gl}(n)$ and let $\gamma(t) = \exp(tX)$ be as above. Now we see that $d\,\mathrm{Ad}_e : \mathfrak{g} \to T_e(\mathbf{GL}(\mathfrak{g}))$ applied to X is given by the velocity of the parametric curve $t \mapsto \mathrm{Ad}(\exp(tX)) \in \mathbf{GL}(\mathfrak{g})$ at $t = 0$. In light of the above, we see that this is simply the derivative of $C_{\exp(tX)} \in \mathbf{GL}(\mathfrak{g})$. Now for $Y \in \mathfrak{g}$ we have that

$$\begin{aligned}
\mathrm{Ad}(\exp(tX))Y &= C_{\exp(tX)}Y \\
&= \exp(tX)\cdot Y \cdot \exp(-tX) \\
&= (1 + tX + o(t))Y(1 - tX + o(t)) \\
&= (Y + tXY + o(t))(1 - tX + o(t)) \\
&= Y + t(XY - YX) + o(t).
\end{aligned}$$

Now we see the $\mathrm{End}(\mathfrak{g})$ valued velocity at $t = 0$ is simply the usual commutator $[X,Y] = XY - YX$.

Lemma 1.115. *Let G and H be Lie groups. A group homomorphism $\varphi : G \to H$ is continuous if and only if it is continuous at the identity. Similarly, φ is smooth if and only if it is smooth in some open set containing the identity.*

Proof. Note that the action G induces on itself of left-translation $L_g : G \to G$ is a homeomorphism taking e to g. Note that similarly, $L_{\varphi(g)} : H \to H$ is also a homeomorphism. Since φ is a homomorphism we have

$$\varphi \circ L_g = L_{\varphi(g)} \circ \varphi.$$

Now it is clear that continuous of φ at g is equivalent to continuity of φ at e. Similarly, since the left translations are in fact diffeomorphisms and $L_g(U)$ is an open neighborhood of g for any open set U containing e we see that if φ is smooth in some U around e, then it is smooth on some neighborhood around every $g \in G$. Since smoothness is a local property, we are done. □

Lemma 1.116. Ad *is a Lie group homomorphism.*

Proof. Note that for $g, h, k \in G$ we have that

$$(C_h \circ C_g)(k) = h \cdot g \cdot k \cdot g^{-1} \cdot h^{-1} = C_{h\cdot g}(k),$$

that is $C_h \circ C_g = C_{h\cdot g}$. Furthemore, note that $C_g(e) = e$. Now we see that by the chain rule that

$$\operatorname{Ad}(h \cdot g) = dC_{h\cdot g} = dC_h \circ dC_g = \operatorname{Ad}(h) \circ \operatorname{Ad}(g).$$

So we see that Ad is a group homomorphism. Now it remains to show that Ad is smooth. Fix $v \in \mathfrak{g}$ and $g \in G$. Now consider a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to G$ satisfying $\gamma(0) = e$ and $\gamma'(0) = v$. Let X be the associated left-invariant vector. Recall that

$$\operatorname{Ad}(g)v = dC_g v = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} C_g(\gamma(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} C(g, \gamma(t)) = dC(\mathbf{0}_g, v),$$

where $\mathbf{0}_g$ is the zero vector in T_gG. Here we consider $(\mathbf{0}_g, v) \in T_{(g,e)}(G \times G) \cong T_gG \oplus T_eG$. Note that since $dC : T(G \times G) \to TG$ is a smooth vector bundle map, we have that $\operatorname{Ad}(g)v$ varies smoothly with respect to g. We can obtain smooth coordinates for $\mathbf{GL}(\mathfrak{g})$ by taking a basis (e_i) for $\mathfrak{g}$ and using $\dim G \times \dim G$ matrix entries with respect to this basis. If α^j is the dual basis then the matrix entries of $\operatorname{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ are given by $(\operatorname{Ad}(\mathfrak{g}))_{i,j} = \alpha^j(\operatorname{Ad}(\mathfrak{g})e_i)$. Now we see that this is clearly a smooth function of g, which establishes that $\operatorname{Ad} : \mathfrak{g} \to \mathbf{GL}(\mathfrak{g})$ is a Lie group homomorphism. $\square$

Lemma 1.117. *For $v, w \in \mathfrak{g}$ we have that*

$$[\![v, w]\!] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\operatorname{Ad}(\gamma_v(t))w),$$

where $\gamma_v : (-\varepsilon, \varepsilon) \to G$ is the smooth curve satisfying $\gamma_v'(0) = v$.

Proof. Let γ_v be as stated above, and let X and Y be the left-invariant vector fields associated to v and w, respectively. Note that since X is left invariant we have that $\Phi_t(h) = h \cdot \gamma_v(t) = h \cdot \Phi_t(e)$, where Φ_t is the flow map associated to X. Putting this all together gives

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \operatorname{Ad}(\gamma_X(t))w &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} dC_{\gamma(t)}w \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} dR_{\gamma(t)}\left(dL_{\gamma(t)}w\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} dR_{\gamma(t)}\left(Y|_{\gamma(t)}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} d\Phi_{-t}Y|_{\Phi_t(e)} \\
&= \mathscr{L}_X Y \\
&= [X, Y] \\
&= [\![v, w]\!].
\end{aligned}$$

$\square$

Now we explicitly compute the differential of Ad.

Proposition 1.118. $d\operatorname{Ad}_e(v) = [\![v, w]\!]$

Proof. Since $\operatorname{Ad} : G \to \mathbf{GL}(\mathfrak{g})$ we have that $d\operatorname{Ad} : \mathfrak{g} \to T_{\mathbb{1}}\mathbf{GL}(\mathfrak{g}) \cong \mathbf{GL}(\mathfrak{g})$. Fix $v \in \mathfrak{g}$ and let $\gamma_v : (-\varepsilon, \varepsilon) to G$ be a smooth map such that $\gamma_v'(0) = v$ and $\gamma_v(0) = e$. Now $d\operatorname{Ad}_e(v)$ is the velocity of $\operatorname{Ad}(\gamma_v(t))$ at $t = 0$, where the velocity takes values in the open subset $\mathbf{GL}(\mathfrak{g})$ of $\operatorname{End}(\mathfrak{g})$. The above lemma tells us that $d\operatorname{Ad}_e(v)w = [\![v, w]\!]$. That is to say $d\operatorname{Ad}_e(v) = [\![v, \cdot]\!]$. $\square$

Note that we could have taken this to be the definition of the Lie bracket on $\mathfrak{g}$.

Definition 1.119. The **adjoint representation** of the Lie algebra $\mathfrak{g}$ is $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ is defined by

$$\mathrm{ad}(v)w = (d\,\mathrm{Ad})(v)w.$$

◇

Theorem 1.120. *Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. There exists a map $\exp : \mathfrak{g} \to G$ which is a local diffeomorphism in a neighborhood of the origin in $\mathfrak{g}$ such that for any $v \in \mathfrak{g}$, the map $t \mapsto \exp(tv)$ is an integral curve for the left-invariant vector field associated to v. Furthermore, $\exp((s+t)v) = \exp(sv) \cdot \exp(tv)$.*

Proof. Fix $v \in \mathfrak{g}$ and let X be the associated left-invariant vector field. By Theorem 1.90 we have the existence of a flow map $\Phi : G \times I_G \to G$. Fix $g \in G$, and let

$$\Psi(h,t) = gh^{-1}\Phi(g,t)$$

for all $h \in G$ and $t \in I_g$, the maximal interval of existence of the flow starting from g. Now we see that

$$\frac{\mathrm{d}\Psi}{\mathrm{d}t} = dL_{g\cdot h^{-1}}\frac{\mathrm{d}\Phi}{\partial t} = X|_{\Psi(h,t)}$$

since X is left-invariant. Now it follows that $I_g \subseteq I_h$, and so by symmetry it follows that $I_h = I_g$ for all $g, h \in G$.

Now let $t \in I_e$. By the above, we can define $\Phi(\Phi(e,t),s)$ for all $s \in I_e$. Now since for any such s the maps $t \mapsto \Phi(\Phi(e,t),s)$ and $t \mapsto \Phi(e,t+s)$ satisfy the same differential equation with the same initial value that they are the same map. Hence, we have that $\Phi(e,t+s)$ is well defined for all $t,s \in I_e$. Hence $I_e + I_e \subseteq I_e$. So clearly I_e cannot be finite, and so $I_e = \mathbf{R}$.

Now define the exponential map as follows: for $v \in \mathfrak{g}$ let X be the associated left-invariant vector field. Then

$$\exp(v) := \Phi(e,1).$$

Note that we immediately have that $\exp(tv) = \Phi(e,t)$, and that $\exp(tv)\cdot g = \Phi(g,t)$ for all $g \in G$ and $t \in \mathbf{R}$. This in turn shows that $\exp((s+t)v) = \exp(sv)\exp(tv)$.

Now it remains to show that the exponential map is a local diffeomorphism. Observe that $d\exp_0 : T_0\mathfrak{g} \to T_eG$. Now for $w \in T_0\mathfrak{g} \cong \mathfrak{g}$ we associate it to the curve $0 + tw$, and we compute

$$d\exp_0(w) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(tw) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi^{tw}(e,1) = w \cdot e = w.$$

Hence $d\exp_0 = \mathrm{id}_{\mathfrak{g}}$, and is in particular invertible. Now it immediately follows from the inverse function theorem that exp is a local diffeomorphism around $0 \in \mathfrak{g}$. □

To conclude this section we compute a couple of examples.

Example 1.121. Consider the Lie group $\mathbf{SL}(n)$ and the associated Lie algebra $\mathfrak{sl}(n)$. Let $I \in \mathbf{SL}(n)$ be the $n \times n$ identity matrix, which is also the identity of the group. Consider any arbitrary $X \in \mathfrak{sl}(n)$ and note that the left invariant vector field is given by

$$\mathbf{X}|_A = AX \in \mathfrak{sl}(n).$$

This is clear since if we consider a left translation $L_B : SL(n) \to \mathbf{SL}(n)$ defined as

$$L_B(A) = BA,$$

then the differential dL_B is given by

$$dL_B(\mathbf{X}|_A) = B\left(\mathbf{X}|_A\right) = BAX = \mathbf{X}|_{BA}.$$

Now if we consider the curve $\gamma : \mathbf{R} \to SL(n)$, we claim that $\gamma(t) = e^{Xt}$ is a solution to the initial value problem

$$\gamma'(t) = \mathbf{X}\gamma(t), \qquad \gamma(0) = I.$$

Notice that

$$\gamma'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{Xt}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{n=0}^{\infty} X^n \frac{t^n}{n!} = \sum_{n=1}^{\infty} X^n \frac{t^{n-1}}{(n-1)!} = \left(\sum_{n=0}^{\infty} X^n \frac{t^n}{n!}\right) X = \left(e^{Xt}\right) X = \mathbf{X}\gamma(t),$$

and
$$\gamma(0) = e^{X\cdot 0} = e^{\mathbf{0}} = I + \sum_{n=1}^{\infty} \frac{\mathbf{0}^n}{n!} = I.$$
So we have that γ solves this initial value problem and we have that $\gamma(1) = e^X$. So we have that the exponential map $\exp : \mathfrak{sl}(n) \to SL(n)$ is defined as
$$\exp(X) = e^X.$$
$\diamond$

Theorem 1.122. *Every connected Abelian Lie group G is of the form $\mathbf{T}^k \times \mathbf{R}^{n-k}$ where $\mathbf{T} = \mathbf{R}/\mathbb{Z} := \{e^{i\theta} : \theta \in [0, 2\pi)\}$.*

Proof. Note that
$$G_0 := \{\exp(X_1)\cdots\exp(X_k) : X_i \in \mathfrak{g}\}$$
is the connected subgroup of G containing the identity.

Now we see that since G is Abelian that for $X, Y \in \mathfrak{g}$ that the map $\psi(t) = \exp(tX)\cdot\exp(tY)$ satisfies
$$d\psi(t) = X|_{\psi(t)} + Y|_{\psi(t)},$$
where we associate $\mathfrak{g} \cong \mathrm{Lie}(G)$. In particular, we deduce that
$$\exp(X)\exp(Y) = \exp(X+Y).$$
So the map $E : \mathfrak{g} \to G$ given by $X \mapsto e^X$ is surjective since G_0 is the connected subgroup containing the identity. Furthermore, we have that E is a group homomorphism (where $\mathfrak{g}$ inherits the group structure from its vector space structure) and diffeomorphism. So $G \cong \mathfrak{g}/\ker(E)$ as groups.

Since $\ker(E)$ is a subgroup of $\mathfrak{g}$, and since exp is a local diffeomorphism around the origin in $\mathfrak{g}$ we can find a neighborhood U of $0 \in \mathfrak{g}$ such that $U \cap \ker(E) = \{0\}$. Now by standard group theory we see that such a discrete subgroup of $\mathfrak{g}$ has a basis $e_1, \ldots, e_n$ such that $\ker E = \{\alpha_1 e_1 + \cdots + \alpha_n e_m : \alpha_i \in \mathbb{Z}\} \cong \mathbb{Z}^m$. So we see that $\mathfrak{g}/\ker(E) \cong \mathbf{T}^k \times \mathbf{R}^{n-k}$, as desired. $\square$

1.10. **Sard's theorem.**

Definition 1.123. We say that a set $\Omega \subseteq M$ is **null** (or of **measure zero**) if it's image in $\mathbf{R}^n$ under *every* chart has $\mathscr{L}^n$ measure zero. $\diamond$

Definition 1.124. We say that $x \in M$ is a **critical point** of $f : M \to N$ if
$$df_x(T_xM) \neq T_{f(x)}N.$$
Note that this is clearly equivalent to $\operatorname{rank} df_x < n$ where $\dim N = n$. $\diamond$

The following theorem is found in Milnor's *Topology from the differentiable viewpoint*.

Theorem 1.125. *Let M be an m-dimensional manifold, N be an n-dimensional manifold, and $f : M \to N$ be a smooth map. The set of critical values of f in N has measure zero.*

Proof. We can assume that $f : U \subseteq \mathbf{R}^m \to \mathbf{R}^n$. The proof proceeds by induction on m. Note that the assertion trivially holds for $m = 0$.

Now set
$$C_0 := \{x \in U : \operatorname{rank} df_x < n\}$$
and
$$C_k := \{x \in U : \text{all partial derivatives of } f \text{ less than or equal to order } k \text{ vanish}\}.$$
Note that each C_k is the finite intersection of sets which are the preimage of zero under a continuous map into $\mathbf{R}$. So each C_k is closed, and we have a decreasing sequence of closed sets
$$C_0 \supseteq C_1 \supseteq C_2 \supseteq \cdots.$$
We claim that $f(C_k \setminus C_{k+1})$ has measure zero for all $k \in \mathbf{N}_0$ and that $f(C_k)$ has measure zero for large enough k.

Step 1. $f(C_0 \setminus C_1)$ **has measure zero.**

Note that if $n = 1$ then $C_0 = C_1$, and hence this step trivially holds. So suppose that $n \geq 2$.

For every $\boldsymbol{x} \in C_0 \setminus C_1$ we will find an open neighborhood V of $\boldsymbol{x}$ such that $f(V \cap C)$ has measure zero. Since $\mathbf{R}^n$ is second countable we have that $C_0 \setminus C_1$ is covered by countably many such sets, the claim will follow from the countable additivity of the Lebesgue measure.

If $\boldsymbol{x} \in C_0 \setminus C_1$ then some partial derivative of f does not vanish at $\boldsymbol{x}$. Without loss of generality we can assume that $\partial_1 f_1(\boldsymbol{x}) \neq 0$. Now define $h : U \to \mathbf{R}^m$ by

$$h(\boldsymbol{x}) = (f_1(\boldsymbol{x}), x_2, \ldots, x_m).$$

We immediately see that

$$dh_{\boldsymbol{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}) & * \\ 0 & \mathbb{1} \end{bmatrix}$$

Since $dh_{\boldsymbol{x}}$ is non-singular it is invertible, and we can use the inverse function theorem to find an open neighborhood V of x such that $h : V \to \widetilde{V} := \operatorname{im}(V)$ is a diffeomorphism. Now set $g := f \circ h^{-1} : \widetilde{V} \to \mathbf{R}^n$. Observe that the set $\widetilde{C}$ of critical points of g is precisely $h(V \cap C)$:

$$\begin{aligned} \widetilde{C} &:= \left\{\boldsymbol{x} \in \widetilde{V} \,:\, \operatorname{rank} d(f \circ h^{-1})_{\boldsymbol{x}} < n\right\} \\ &= \left\{\boldsymbol{x} \in \widetilde{V} \,:\, \operatorname{rank}\left(df_{h^{-1}(\boldsymbol{x})} \circ dh^{-1}_{\boldsymbol{x}}\right) < n\right\} \\ &= \{\boldsymbol{x} \in \widetilde{V} \,:\, \operatorname{rank} df_{h^{-1}(\boldsymbol{x})} < n\} \\ &= \{h(\boldsymbol{x}) \in \widetilde{V} \,:\, x \in V,\ \operatorname{rank} df_{\boldsymbol{x}} < n\} \\ &= h(V \cap C). \end{aligned}$$

So we see that

$$g(\widetilde{C}) = (f \circ h^{-1})(\widetilde{C}) = (f \circ h^{-1})(h(V \cap C)) = f(V \cap C).$$

So now we show that the set of critical values of g restricted to V is null. Note that since $h_1(\boldsymbol{x}) = f_1(\boldsymbol{x})$, we see that $g_1(\boldsymbol{x}) = x_1$ for all $\boldsymbol{x} \in \widetilde{V}$. That is to say g fixes the first component of its argument, and so for fixed $t \in \mathbf{R}$ we write

$$g_t : (\{t\} \times \mathbf{R}^{m-1}) \cap \widetilde{V} \to \{t\} \times \mathbf{R}^{n-1},$$

as the restriction of g to the slice of $\widetilde{V}$ with first component t. Thinking of g_t as a map $g_t : \widetilde{V}_t \subset \mathbf{R}^{m-1} \to \mathbf{R}^{n-1}$ we can consider the differential

$$\mathrm{d}g(t, \boldsymbol{x}') = \begin{bmatrix} 1 & 0 \\ * & \mathrm{d}g_t(\boldsymbol{x}') \end{bmatrix}$$

for elements of $\widetilde{V}_t$. One should observe that dg_t is lower triangular, which follows immediately since dh is upper triangular. So we see that (t, y) is a critical value of g if and only if y is a critical value of g_t. By induction over the dimension of the space we obtain for each t this set, which we will denote as V_t, has measure zero. Now the set of critical points of g is closed, and so the set of critical values of g, which is V, is Borel. So we have that χ_V is measurable and by Tonelli's theorem we see that

$$\mathscr{L}^n(V) = \int_{\mathbf{R}^n} \chi_V(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{-\infty}^{\infty} \int_{\mathbf{R}^{n-1}} \chi_{V_t}(\boldsymbol{x}', t)\, \mathrm{d}\boldsymbol{x}'\, \mathrm{d}t = \int_{\mathbf{R}^{n-1}} 0\, \mathrm{d}\boldsymbol{x}' = 0.$$

So V is a null set, as desired.

Step 2. $f(C_k \setminus C_{k+1})$ **has measure zero for** $k \in \{1, 2, \ldots\}$**.**

Let $x \in C_k \setminus C_{k+1}$. So we can find a multi-index α with $|\alpha| = k$ such that

$$w := \partial^\alpha f_1$$

satisfies

$$w(x) = 0 \quad \text{and} \quad \partial_1 w(x) \neq 0.$$

Now define $h : U \to \mathbf{R}^m$ by

$$h(x) := (w(x), x_2, \ldots, x_m).$$

For some neighborhood V of x, the map $h : V \to \widetilde{V} := h(V)$ is a diffeomorphism. Now set $g := f \circ h^{-1} : \widetilde{V} \to \mathbf{R}^n$.

We have that $h(C_k \cap V) \subset \{0\} \times \mathbf{R}^{n-1}$, and, in fact, is contained in the set of critical points of g restricted to this domain. So we have that $f(C_k \cap V)$ is contained in the set of critical values of $g|_{\widetilde{V}\cap(\{0\}\times\mathbf{R}^n)}$, which is zero by induction.

Step 3. For all $k > m/n - 1$, $f(C_k)$ has measure zero.

Let $I \subset U$ be a closed cube of side length δ. By Taylor's theorem and compactness of I there exists a $c > 0$ such that, for all $\boldsymbol{x} \in C_k \cap I$ and $\boldsymbol{x} + \boldsymbol{h} \in I$,

$$|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})| \le c\|\boldsymbol{h}\|^{k+1}.$$

We need an inequality of order $k+1$ since the derivatives of lower order will vanish at $\boldsymbol{x}$. Now divide I into ℓ^m cubes of side length δ/ℓ. If $\widetilde{I}$ is such a cube and $\widetilde{I} \cap C_k \neq 0$, then fix $\boldsymbol{x} \in \widetilde{I} \cap C_k$ and for each point of $\widetilde{I}$ can be written as $\boldsymbol{x} + \boldsymbol{h}$ where

$$\|\boldsymbol{h}\| \le \frac{\sqrt{m}\delta}{\ell}.$$

However, the first inequality above tells us that

$$\|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})\| \le c\left(\frac{\sqrt{m}\delta}{\ell}\right)^{k+1},$$

and so $f(\widetilde{I})$ is contained in a cube of side length $2c(\sqrt{m}\delta/\ell)^{k+1}$. Since this bound holds for all such sub-cubes we have that $f(I \cap C_k)$ is contained in a union of at most ℓ^m cubes, each with volume at most

$$\left(\frac{2c(\sqrt{m}\delta)^{k+1}}{\ell^{k+1}}\right)^n.$$

So we have that

$$\mathscr{L}^N(f(I\cap C_k)) \le \ell^m \left(\frac{2c(\sqrt{m}\delta)^{k+1}}{\ell^{k+1}}\right)^n = \left(2c(\sqrt{m}\delta)^{k+1}\right)^n \ell^{m-n(k+1)}.$$

Now we see that if $k > m/n - 1$, then this goes to zero as $\ell \to \infty$. Hence, we have obtained that $f(I \cap C_k)$ is a $\mathscr{L}^n$ measure zero set.

Now we finish the proof. We can write

$$\begin{aligned} f(C_0) &= f(C_0 \setminus C_1 \cup C_1 \setminus C_1 \cup \cdots \cup C_{k-1} \setminus C_k \cup C_k) \\ &= f(C_0 \setminus C_1) \cup f(C_1 \setminus C_2) \cup \cdots \cup f(C_{k-1} \setminus C_k) \cup f(C_k). \end{aligned}$$

Since each of these is a set of measure zero, we see that $f(C_0)$ is a set of measure zero and we are done. □

1.11. **Whitney embedding theorem.** One might ask whether every abstract n-dimensional topological manifolds can be realized as a submanifold of some Euclidean space $\mathbf{R}^m$ for some $m \in \mathbf{N}$. The main result of this section, Whitney's embedding theorem, provides a positive answer to this question. Intuitively, we should not be able to embed an n-dimensional manifold in $\mathbf{R}^n$ as witnessed by the existence of nonplanar graphs or the fact that $\mathbf{S}^n$ is naturally embedded in $\mathbf{R}^{n+1}$ – the fact that $\mathbf{S}^n$ does not embed into $\mathbf{R}^n$ can be shown using some basic algebraic topology. Every topological (read not necessarily smooth) manifold can be embedded in $\mathbf{R}^{2n+1}$ and every smooth manifold can be embedded in $\mathbf{R}^{2n}$.

Theorem 1.126 (Whitney's embedding theorem – $\mathbf{R}^{2n+1}$ version)**.** *Every compact smooth n-dimensional manifold M can be realized as a submanifold of $\mathbf{R}^{2n+1}$.*

Proof. Let M be a smooth compact n-dimensional manifold.

Step 1. M **embeds into** $\mathbf{R}^m$ **for some** $m \in \mathbf{N}$.

Let $\mathscr{U} = \{U_i\}_{i=1}^k$ be a finite open cover of M such that each U_i is contained in the domain of some chart $\varphi_i : V_i \to \mathbf{R}^n$. Let $\{\rho_i\}_{i=1}^k$ be a smooth partition of unit subordinated to $\mathscr{U}$. For all $i = 1, \ldots, k$ let $\widetilde{\varphi_i} = \rho_i \varphi_i : M \to \mathbf{R}^n$ (where we define this product to be zero whenever ρ_i is). Now define

$$\begin{aligned} \iota : M &\to (\mathbf{R}^n \times \mathbf{R})^k \\ p &\mapsto ((\widetilde{\varphi_1}(p), \rho_1(p)), \ldots, (\widetilde{\varphi_k}(p), \rho_k(p))). \end{aligned}$$

Now we show that ι is injective. Let $p, q \in M$ be such that $\iota(x) = \iota(y)$. Then we see that for some $i = 1, \ldots, k$ that $p, q \in U_i$ and hence $\varphi_k(p) = \varphi_k(q)$. Since φ_k is a homeomorphism onto it's image we see that this implies $p = q$, and so ι is injective.

Now we show that ι is an immersion, i.e. that $d\iota$ is injective everywhere. Fix $p \in M$, $v \in T_pM$, and we see that for $i = 1, \ldots, k$

$$(d\iota_p(v))_i = (\rho_i(p)(d\varphi_i)_p(v) + (d\rho_i)_p(v)\varphi_i(p), (d\rho_i)_p(v)).$$

Note that the above computation follows immediately since tangent vectors are derivations that satisfy the Leibniz product rule. Now we see that if $d\iota_p(v) = 0$ for some $v \in T_pM$ then $(d\varphi_i)_p(v) = 0$ for all $i = 1, \ldots, k$; hence $v = 0$ since $\varphi_i : V_i \to \mathbf{R}^n$ is an immersion for all i (here $V_i \subseteq M$ is viewed as a submanifold of M).

Finally, we show that $\iota : M \to \iota(M)$ is an embedding. This follows immediately since M is compact, $\iota(M)$ endowed with the subspace topology of $(\mathbf{R}^n \times \mathbf{R})^k$ is Hausdorff, and ι is a bijective, continuous map – this is a classical topological that any bijective, continuous map from a compact space into a Hausdorff space is a homeomorphism. We conclude this step by noting that $(\mathbf{R}^n \times \mathbf{R})^k \cong \mathbf{R}^{k(n+1)}$.

Step 2. If $m > 2n+1$ **then the set of hyperplanes** $\Pi \in (\mathbf{RP}^{m-1})^\star$ **such that the composition** $p_\Pi \circ \iota$ **is injective from** $M \to \Pi$ **is full measure, where** $p_\Pi : \mathbf{R}^m \to \Pi$ **is the orthogonal projection.**

Note that $p_\Pi \circ \iota$ is injective if and only if every line orthogonal to Π intersects the manifold at exactly one point. Intuitively, that is to say that most directions are not directions which can joint distinct pairs of points in the manifold. Formally, let $\Delta = \{(p,p) \in M \times M \ : \ p \in M\}$ be the diagonal set. Now define

$$\delta : M \times M \setminus \Delta \to^{m-1}$$

via $\delta(p,q)$ is the line through the origin parallel to the line joining $\iota(p)$ and $\iota(q)$ in $\mathbf{R}^m$. Note that $\dim(\mathbf{RP}^{m-1}) = m-1$ and $\dim(M \times M \setminus \Delta) = 2n$. Since we assume that $m > 2n+1$ we see that $\dim(\mathbf{RP}^{m-1}) \geq 2n+1$. Now by Sard's theorem applied to the map δ we conclude that the measure of $\delta(M \times M \setminus \Delta)$ is zero. Hence, the complement of the image of Δ is full measure, and since $p_\Pi \circ \iota$ is injective for all Π in the complement of the image of δ we are done.

Step 3. If $m > 2n+1$ **then the set of hyperplanes** $\Pi \in (\mathbf{RP}^{m-1})^\star$ **where** $p_\Pi \circ \iota$ **is an immersion is full measure.**

Note that dp_Π is not injective whenever there exists some non-zero $v \in T_pM \subseteq \mathbf{R}^m$ such that $p_\Pi(v) = 0$. Consider the projectivization of the tangent bundle, $\mathbf{P}(TM)$. This is a fiber bundle over M with fiber $\mathbf{RP}^{m-1}$. So the dimension of the total space is $2n-1$. Now consider the map

$$\gamma : \mathbf{P}(TM) \to \mathbf{RP}^{m-1}$$

defined as $\gamma(v) \mapsto d\iota_p(v) \in \mathbf{R}^m$ for $v \in T_pM$. Then since the dimension of $\mathbf{P}(TM)$ is at least $2n+1$ by Sard's theorem again we conclude that the image of γ is measure zero. Hence, the collection of hyperplanes Π where dp_Π is injective iz full measure.

We conclude by noting that $d(p_\Pi \circ \iota) = dp_\Pi \circ d\iota$.

Step 4. Every n**-dimensional smooth manifold embeds into** $\mathbf{R}^{2n+1}$**.**

Now by taking any Π such that the above two conditions hold, we obtain that $p_\Pi \circ \iota : M \to \Pi \cong \mathbf{R}^{m-1}$ is an embedding. By repeatedly applying the previous two steps to this new embedding we obtain an embedding of M into $\mathbf{R}^{2n+1}$.

□

Corollary 1.127. *Every compact n-dimensional manifold immerses in* $\mathbf{R}^{2n}$.

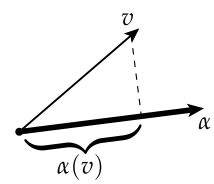
Proof. Since the conditions on Step 3 of the above proof are more relaxed than those of Step 2, we can iterate one more time to show that every n-dimensional compact manifold immerses into $\mathbf{R}^{2n}$. □

2. Differential Forms.

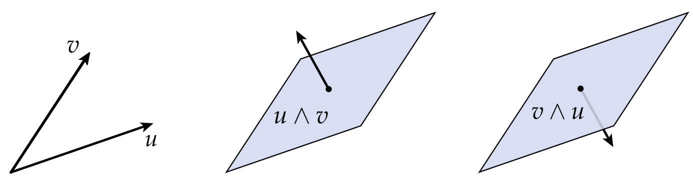
Before defining tensors, differential forms, and the exterior algebra in the context of manifolds, we give some intuition by reviewing the geometric aspects of the exterior algebra of vector spaces. Note that the treatment in this introduction is not meant to be fully rigorous.

Let V be an n-dimensional real vector space. Given a vector $v \in V$ we can inspect v by considering the *length* of v along a given direction α.

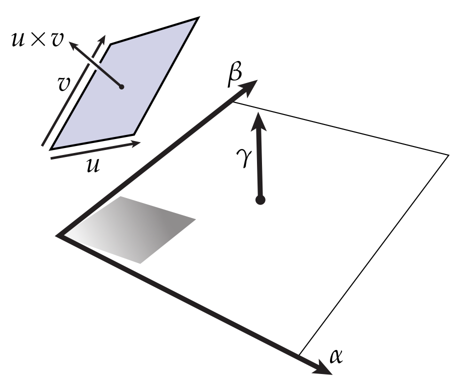
The result is simply a real number, which we can call $\alpha(v)$. This notation is meant to emphasize that α is a function $V \to \mathbf{R}$: in particular, it is a *linear* function which takes in a vector and produces a scalar. Any such function is called a *covector* (or 1-form).



Now if we have a collection of vectors, we can measure a similar quantity by projecting the parallelogram spanned by these vectors onto the space spanned by two other vectors α and β. First we need to introduce an operation called the wedge product of two vectors u and v, which will give us the parallelogram spanned by u and v. This is most easily described visually:



The object $u \wedge v$ is called a *2-vector*. The wedge produce is antisymmetric in the sense that $u \wedge v = -v \wedge u$. Now we introduce an operation $\alpha \wedge \beta$ which returns the area of the projection of $u \wedge v$ onto span$\{\alpha, \beta\}$.



This projection operation is given the name of $\alpha \wedge \beta$ and takes in two vectors u, v and returns a real number $\alpha \wedge \beta(u \wedge v)$. This *wedge* operation is also antisymmetric in the sense that $\alpha \wedge \beta = -\beta \wedge \alpha$. The object $\alpha \wedge \beta$ is called a 2-vector. More generally, a k-vector is the wedge product of k vectors (this isn't exactly right, but it gives some good intuition).

The exterior algebra roughly allows us to talk about *signed volumes* spanned by vectors and get complementary vectors (i.e. the Hodge dual). Differential forms will be smoothly varying (k-)covectors and will be the technical foundation for calculus and analysis on manifolds. It is always a good idea to keep in mind that although the construction of these objects is very algebraic they have a very concrete geometric interpretation.

The rest of this section closely follows the first chapter of Federer's Geometric Measure Theory.

2.1. **Tensor products.** The language of tensors and multilinear algebra is extremely useful and prevalent in differential geometry and manifold analysis. We give the basic constructions and prove some basic propositions that will become useful to us later.

Definition 2.1. Let V and W be vector spaces over $\mathbf{R}$. The **tensor product** is a vector space $V \otimes W$ and a bilinear map, called the multiplication, $\mu : V \times W \to V \otimes W$ such that for any bilinar map $f : V \times W \to Z$ for any other vector space Z we obtain a *unique* linear map $F : V \otimes W \to Z$ such that the following diagram commutes:

$$\begin{tikzcd} V \otimes W \arrow[dr, dashed, "F"] & \\ V \times W \arrow[u, "\mu"] \arrow[r, "f"'] & Z \end{tikzcd}$$

This is known as the **universal mapping property** of the tensor product. ⋄

Proposition 2.2. *Let V, W be vector spaces. The tensor product $V \otimes W$ exists.*

Proof. Let $F(V \times W)$ be the free vector space over $V \times W$, i.e.

$$F(V \times W) := \left\{ \sum_{i=1}^{k} \alpha_i (v_i, w_i) \ : \ \alpha_i \in \mathbf{R},\ v_i \in V,\ w_i \in W \right\}.$$

Note that these are formal sums in this product space where every element of $V \times W$ is a basis element. Now define $\widetilde{\mu} : V \times W \to F(V \times W)$ as follows:

$$\widetilde{\mu}((v, w)) = (v, w).$$

Now let $G \subseteq F(V \times W)$ be defined as follows:

$$\begin{aligned} G := \{ & \alpha(v, w) - (\alpha v, w), \\ & \alpha(v, w) - (v, \alpha w), \\ & (v_1 + v_2, w) - (v_1, w) - (v_2, w), \\ & (v, w_1 + w_2) - (v, w_1) - (v, w_2) \\ & \quad : \ v, v_1, v_2 \in V,\ w, w_1, w_2 \in W \}. \end{aligned}$$

Now define the tensor product is as

$$V \otimes W := F(V \times W)/G.$$

Note that $\mu : V \times W \to V \otimes W$ is bilinear by the construction of G. Now consider any bilinear map $f : V \times W \to Z$ for some vector space Z. Note that since $\widetilde{\mu}$ is injective we have that there exists a unique linear map $\widetilde{F} : F(V \times W) \to Z$ such that $\widetilde{F} \circ \widetilde{\mu} = f$. It is easy to check that

$$\ker f \subseteq G,$$

and so we have that there exists a unique map $F : V \otimes W \to Z$ such that

$$F = \widetilde{F} \circ \pi_{F(V \times W) \to V \otimes W},$$

and we have that F makes the diagram in the definition of the tensor product commute. So we see that $V \times W$ is indeed a tensor product of V and W, as desired. □

Remark 2.3. It is almost never very useful to use the actual construction of the tensor product and we can just use the universal mapping property of the tensor product to obtain very powerful results very easily! ⋄

Proposition 2.4. *The tensor product is unique (up to linear isomorphism).*

Proof. Let V and W be vector spaces and suppose that A and B are both tensor products of V and W (and let $\mu_A : V \times W \to A$ and $\mu_B : V \times W \to B$ denote their respective multiplications). Then we obtain unique linear maps $\alpha : A \to B$ and $\beta : B \to A$ such that the following diagram commutes:

$$\begin{tikzcd}
& A \arrow[d, dashed, "\alpha"] \\
V \times W \arrow[ur, "\mu_A"] \arrow[r, "\mu_B"] \arrow[dr, "\mu_A"'] & B \arrow[d, dashed, "\beta"] \\
& A
\end{tikzcd}$$

Now we obtain that $\alpha \circ \beta : A \to A$ is the *unique morphism* such that $\mu_A \circ (\alpha \circ \beta) = \mu_A$. Note that since $\mathrm{id}_A : A \to A$ satisfies this we obtain that $\alpha \circ \beta = \mathrm{id}_A$. We do the similar for B and find that the maps $\alpha : A \to B$ and $\beta : B \to A$ induce a linear isomorphism of tensor products. □

Remark 2.5. Basic tensors in $V \otimes W$ will be denoted as $v \otimes w$, which is the equivalence class which contains (v, w) (identified in $F(V \times W)$). Furthermore, note that not all tensors are of the above form, i.e. there are tensors of the form

$$\sum_{i=1}^{k} \alpha_i (v_i \otimes w_i), \qquad \alpha_i \in \mathbf{R},\ v_i \in V,\ w_i \in W.$$

◇

The tensor product has many very natural properties. One is given in the following proposition, and it also demonstrates the power of the universal mapping property.

Proposition 2.6. *Let V_1, V_2, W_1, W_2 be vector spaces. Let $f : V_1 \to W_1$ and $g : V_2 \to W_2$ be linear maps. Then there exists a unique linear map $f \otimes g : V_1 \otimes V_2 \to W_1 \otimes W_2$ satisfying*

$$(f \otimes g)(v_1 \otimes v_2) = f(v_1) \otimes g(v_2).$$

Proof. Consider the following diagram

$$\begin{tikzcd}
& V_1 \otimes V_2 \arrow[dr, dashed, "f \otimes g"] & \\
V_1 \times V_2 \arrow[ur, "\mu_{V_1 \times V_2}"] \arrow[r, "f \times g"'] & W_1 \times W_2 \arrow[r, "\mu_{W_1 \times W_2}"'] & W_1 \otimes W_2
\end{tikzcd}$$

The map $f \otimes g$ comes from the universal mapping property of $V_1 \otimes V_2$ and since $(\mu_{W_1 \times W_2}) \circ (f \times g)$ is a bilinear map into a vector space $W_1 \otimes W_2$. □

Example 2.7. Let $f : [0,1] \to \mathbf{R}$ and $g : [0,1] \to \mathbf{R}$ be smooth functions. Consider the covectors $F, G \in (L^1([0,1]))^*$ given by

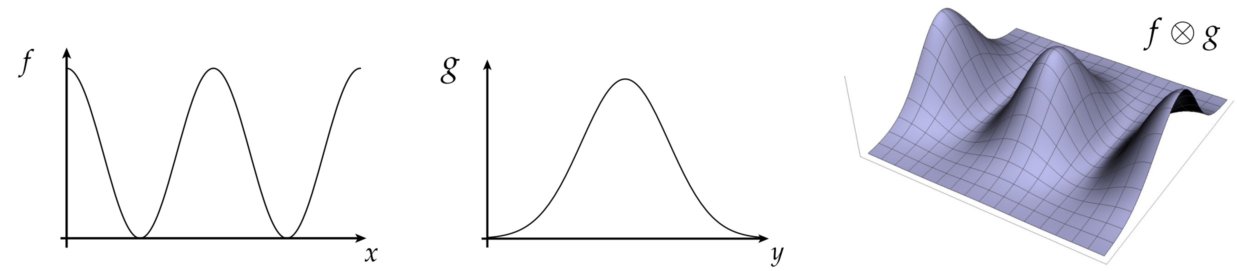
$$F(u) := \int_0^1 f(x) u(x)\, \mathrm{d}x,$$

$$G(v) := \int_0^1 g(y) v(y)\, \mathrm{d}y.$$

Note that F and G are clearly linear maps. Now we can explicitly compute $F \otimes G : (L^1([0,1]) \otimes L^1([0,1])) \to \mathbf{R} \otimes \mathbf{R} \cong \mathbf{R}$ to be the bilinear form

$$(F \otimes G)(u \otimes v) = \int_{[0,1]^2} u(x) f(x) g(y) v(y)\, \mathrm{d}x\, \mathrm{d}y.$$

Visually, we have the following:



Note that here $f \otimes g$ is does not actually mean the tensor product of f and g as in the previous proposition, since f and g are not linear functions from $L^1([0,1]) \to \mathbf{R}$. This notation is simply meant as a type of intuition used in visualizing the tensor product. ⋄

Proposition 2.8. *If $V \cong P \oplus Q$ is a vector space (with P, Q vector spaces) and W is any other vector space then*

$$V \otimes W \cong (P \otimes W) \oplus (Q \otimes W).$$

Proof. Consider the map $\iota : (P \oplus Q) \times W \to (P \otimes W) \oplus (Q \otimes W)$ given by

$$\iota(p+q, w) = p \otimes w + q \otimes w.$$

This map is clearly bilinear, and so by the universal property of the tensor product we obtain a unique linear map $\iota^{\otimes} : (P \oplus Q) \otimes W \to (P \otimes W) \oplus (Q \otimes W)$ satisfying

$$\iota^{\otimes}((p+q) \otimes w) = p \otimes w + q \otimes w.$$

We will now use the universal property of the tensor product to explicitly construct the inverse to $\iota^{\otimes}$. Consider the maps $\alpha : P \times W \to (P \oplus Q) \otimes W$ and $\beta : Q \times W \to (P \oplus Q) \otimes W$ given by

$$\alpha(p, w) = p \otimes w = (p+0) \otimes w,$$
$$\beta(q, w) = q \otimes w = (0+q) \otimes w.$$

Note that α and β are both bilinear maps, and so they induce linear maps $\alpha^{\otimes} : P \otimes W \to (P \oplus Q) \otimes W$ and $\beta^{\otimes} : Q \otimes W \to (P \oplus Q) \otimes W$ such that

$$\alpha^{\otimes}(p \otimes w) = p \otimes w = (p+0) \otimes w,$$
$$\beta^{\otimes}(q \otimes w) = q \otimes w = (0+q) \otimes w.$$

Now consider the morphism $\gamma : (P \otimes W) \oplus (Q \otimes W) \to (P \oplus Q) \otimes W$ given by

$$\gamma(p \otimes w_1 + q \otimes w_2) = \alpha^{\otimes}(p \otimes w_1) + \beta^{\otimes}(q \otimes w_2) = (p+0) \otimes w_1 + (0+q) \otimes w_2.$$

We see that γ is indeed a well defined linear morphism since $\alpha^{\otimes}$ and $\beta^{\otimes}$ are. It is easy to see that γ and $\iota^{\otimes}$ are indeed inverses, and so we are done. □

Remark 2.9. Of course by induction we have that the above result holds for any finite number of direct sums. More generally, it can be shown that the result holds for arbitrary direct sums. ⋄

Proposition 2.10. $\dim(V \otimes W) = (\dim V)(\dim W)$.

Proof. Let $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^m$ be bases for V and W, respectively. In light of Proposition 2.8 we see that the elements $\{e_i \otimes f_j\}_{i,j}$ form a basis of $V \otimes W$. Hence

$$\dim(V \otimes W) = (\dim V)(\dim W).$$

□

Proposition 2.11. *Suppose that either $\dim V < +\infty$ or $\dim W < +\infty$, and let $\mathscr{L}_2(V \times W)$ be the space of all bilinar functions from $V \times W \to \mathbf{R}$. Then*

$$(V \otimes W)^{\star} \cong \mathscr{L}_2(V \times W) \cong V^{\star} \otimes W^{\star}.$$

Definition 2.12. Let R be a ring. We say that R is **graded** if there is a distinguished decomposition

$$R = \bigoplus_{n=0} R_n,$$

where each R_n is a subgroup of R with the property that $R_i \cdot R_j \subseteq R_{i+j}$. We call R_n the **homogeneous part of degree** n and all $r \in R_n$ a **homogeneous element of degree** n. ⋄

Definition 2.13. A **graded R-algebra** A is a graded ring such that the homogeneous parts are R-submodules. In the case where R is a field, that is to say that each homogeneous part is a vector space. ⋄

Definition 2.14. The *contravariant tensor algebra* is the graded algebra

$$\bigotimes\nolimits_* V := \bigoplus_{n=0}^{\infty} V^{\otimes n},$$

where $V^{\otimes 0} = \mathbf{R}$, $V^{\otimes 1} = V$, $V^{\otimes 2} = V \otimes V$, and so on. The multiplication on $\bigotimes_* V$ is given such that the restriction to $V^{\otimes m} \times V^{\otimes n}$ is simply the following bilinar composition:

$$V^{\otimes m} \times V^{\otimes n} \longrightarrow V^{\otimes m} \otimes V^{\otimes n} \longrightarrow V^{\otimes (m+n)}$$

⋄

Proposition 2.15 (UMP of the tensor algebra)**.** *For any graded associative algebra A with a unit element, and a linear map $f : V \to A_1$, there exists a* unique *extension to a unit preserving algebra homomorphism*

$$F : \bigotimes\nolimits_* V \to A,$$

that preserves the grading.

Proof. Let $A = \bigoplus_{n=0}^{\infty} A_n$ be a graded associative algebra with a unit. Let $f : V \to A_1$ be a linear map. Now define $F : \bigotimes_* V \to A$ as the linear extension satisfying the following: for each $n \in \mathbf{N}$ let

$$F(v_1 \otimes \cdots \otimes v_n) = \prod_{i=1}^{n} f(v_i),$$

where $v_i \in V$, and the multiplication on the right is understood to be in A. Similarly, let

$$F(c) = c \cdot \mathbf{1}_A,$$

for $c \in \mathbf{R} = V^{\otimes 0}$ and $\mathbf{1}_A$ is the unit element in A. Clearly F is a well-defined algebra homomorphism from $\bigotimes_* V$ into A, satisfying $f = F \circ \iota$, where $\iota : V \to \bigotimes_* V$ is the inclusion.

Now it suffices to show that F is unique. Suppose that we have two such algebra homomorphism $F : \bigotimes_* V \to A$ and $G : \bigotimes_* V \to A$ satisfying $F \circ \iota = f = G \circ \iota$. Note that since F and G are algebra homomorphisms we have that $F(1) = 1_A = G(1)$. Then we see that if $\iota_n : V^{\otimes n} \to \bigotimes_* V$ is the inclusion map we have

$$F \circ \iota_n = G \circ \iota_n,$$

for all $n \geq 1$. So we see that

$$F = \sum_{i=1}^{\infty} F \circ \iota_n,$$

and

$$G = \sum_{i=1}^{\infty} G \circ \iota_n.$$

Now we can deduce that $F = G$ since $F \circ \iota_n = G \circ \iota_n$ for all $n \geq 0$. □

Remark 2.16. Note that when we say algebra homomorphism we always mean a *unital* algebra homomorphism.

The universal property of the contravariant tensor algebra does not hold in general if we consider non-unital algebra homomorphisms, in particular the uniqueness fails! An explicit counter example is given by $V = \mathbf{R}$ and consider the map $V \to \mathbf{R} \oplus \mathbf{R}$ given by $r \mapsto (r, 0) = r + 0$. Now we can extend this to an non-unital algebra homomorphism $V \oplus V^{\otimes 2} \oplus \cdots \to A$ in the obvious way. Now we see that there are two distinct ways to extend to the space $V^{\otimes 0} \cong \mathbf{R}$, either by putting it into the first coordinate or pushing it diagonally through. ⋄

Remark 2.17. Note that the contravariant tensor algebra is *exactly* the free algebra on a vector space. If we perform the same construction on the dual V^* we get an algebra of functions. If $\dim V = n$ then we obtain an isomorphism
$$T(V^*) \cong \mathbf{R}\{x_1, \dots, x_n\},$$
which is the algebra of noncommutative polynomials in n-variables. ⋄

Proposition 2.18. *Every linear map $f : V \to W$ can be uniquely extended to a unit preserving algebra homomorphism*
$$\bigotimes_* f : \bigotimes_* V \to \bigotimes_* W,$$
such that
$$\bigotimes_* f = \bigoplus_{n=1}^{\infty} \bigotimes_n f,$$
where $\bigotimes_n f : V^{\otimes n} \to W^{\otimes m}$.

Proof. This follows immediately from the universal property of the tensor algebra. □

The universal property of the tensor algebra tells us that we have a bijection
$$\hom_{\mathbf{R}}(V, A) \cong \hom_{\mathbf{R}\text{-Algebra}}(\bigotimes_* V, A).$$
Those of you with a bit of a background in category theory might note that this looks a bit like an adjunction. Indeed, there is an adjunction here, and the proof follows immediately from the universal property of the tensor algebra.

Proposition 2.19. *The functor $\bigotimes_*$: $\mathbf{R}$-Vect $\to$ $\mathbf{R}$-Algebra is left adjoint to the forgetful functor U : $\mathbf{R}$-Algebra $\to$ $\mathbf{R}$-Vect.*

Definition 2.20. Consider the following tensor product
$$V_r^s := V^{\otimes r} \otimes (V^*)^{\otimes s} = \underbrace{V \otimes \cdots \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s \text{ times}}.$$
Elements of V_r^s are called **(r,s)-type tensors**. In particular, elements of V_r^0 are called **contravariant** tensors of degree r, and elements of V_0^s are called **covariant** tensors of degree s. ⋄

Definition 2.21. The **tensor algebra** is given by
$$T(V) := \bigoplus_{r,s=0}^{\infty} V_r^s.$$
$T(V)$ is endowed with two natural operations:

- **contraction** is a mapping $V_r^s \to V_{r-1}^{s-1}$ given by
$$v_1 \otimes \cdots \otimes v_r \otimes \alpha_1 \otimes \cdots \otimes \alpha_s \mapsto \alpha_1(v_r)(v_1 \otimes \cdots \otimes v_{r-1} \otimes \alpha_2 \otimes \cdots \otimes \alpha_s).$$
- **tensor product** is a mapping $V_{r_1}^{s_1} \times V_{r_2}^{s_2} \to V_{r_1+r_2}^{s_1+s_2}$ given by
$$\begin{aligned}(v_1 \otimes \cdots \otimes v_{r_1} \otimes \alpha_1 \otimes \cdots \otimes \alpha_{s_1}, w_1 \otimes \cdots \otimes w_{r_2} \otimes \beta_1 \otimes \cdots \otimes \beta_{s_2})&\\ \mapsto v_1 \otimes \cdots \otimes v_{r_1} \otimes w_1 \otimes \cdots \otimes w_{r_2} \otimes \alpha_1 \otimes \cdots \otimes \alpha_{s_1} \otimes \beta_1 \otimes \cdots \otimes \beta_{s_2}&.\end{aligned}$$

⋄

Example 2.22. A special case of a contraction is when $\boldsymbol{t} \in V_1^1$ is a (1,1)-type tensor. Let $\{v_i\}_{i=1}^n$ be a basis of V and let $\{\alpha_i\}_{i=1}^n$ be the associated dual basis for V^* such that $\alpha_i(v_j) = \delta_{i,j}$. Write $\boldsymbol{t}$ in this basis as
$$\boldsymbol{t} = \sum_{i=1}^n \sum_{j=1}^n t_{i,j} v_i \otimes \alpha_j.$$
Then we see that
$$\text{contract}(\boldsymbol{t}) = \sum_i \sum_j t_{i,j} \alpha_j(v_i) = \sum_i t_{i,i},$$

and so we see that the contraction is simply a generalization (or coordinate-free representation) of the trace operator. ◇

Definition 2.23. $\mathrm{Mult}_r^s(V,W)$ is the vector space of all multilinear functions

$$\underbrace{V\times\cdots\times V}_{r\text{ times}}\times\underbrace{V^*\times\cdots\times V^*}_{s\text{ times}}.$$

◇

We now recall a basic result from linear algebra.

Definition 2.24. Let V and W be two vector spaces. A **pairing** is a bilinear map $\langle\cdot,\cdot\rangle : V\times W\to\mathbf{R}$. A pairing is **non-degenerate** if

$$\langle v_0,w\rangle=0\ \forall w\in W\implies v_0=0,$$

and

$$\langle v,w_0\rangle=0\ \forall v\in V\implies w_0=0.$$

◇

Lemma 2.25. *If $B:V\times W\to\mathbf{R}$ is a non-degenerate pairing, then $V\cong W^*$ and $W\cong V^*$.*

Proof. Consider the mapping $b_1:V\to W^*$ given by

$$b_1(v)=(w\mapsto B(v,w)).$$

Note that b_1 is linear since B is bilinear, and

$$\ker b_1=\{v_0\in V\ :\ b_1(v_0)=0\}=\{v_0\in V\ :\ B(v_0,w)=0\ \forall w\in W\}=\{0\}.$$

So b_1 is injective and so $\dim V\le\dim W^*=\dim W$. By an identical argument we find that $\dim W\le\dim V^*=\dim V$. Hence $\dim V=\dim W$, and so b_1 is an isomorphism. By the same argument applied to $b_2:W\to V^*$ given by $b_2(w)=(v\mapsto B(v,w))$ we obtain that b_2 is an isomorphism. Hence $V\cong W^*$ and $W\cong V^*$ as desired. □

Proposition 2.26. $(V^*)_r^s\cong(V_r^s)^*\cong\mathrm{Mult}_r^s(V,\mathbf{R})$.

Proof. To show the first isomorphism we construct a non-degenerate pairing $(V^*)_r^s\times V_r^s\to\mathbf{R}$ given on basic tensors

$$v^*=\alpha_1\otimes\cdots\otimes\alpha_r\otimes v_1\otimes\cdots\otimes v_s\in(V^*)_r^s,$$

and

$$u=u_1\otimes\cdots\otimes u_r\otimes\beta_1\otimes\cdots\otimes\beta_s\in V_r^s$$

by

$$(v^*,u)=\alpha_1(u_1)\cdots\alpha_r(u_r)\beta_1(v_1)\cdots\beta_s(v_s),$$

and extended multilinearly. This pairing is clearly non-singular. Hence, by Proposition 2.25 we have an isomorphism

$$(V^*)_r^s\cong(V_r^s)^*.$$

On the other hand, by the universal property of the tensor product we have that

$$(V_r^s)^*\cong\mathrm{Mult}_r^s(V,\mathbf{R}).$$

Note that under this isomorphism if $h\in(V_r^s)^*$ then the corresponding multilinear function $H\in\mathrm{Mult}_r^s(V,\mathbf{R})$ satisfies

$$H(v_1,\dots,v_r,\alpha_1,\dots,\alpha_s)=h(v_1\otimes\cdots\otimes v_r\otimes\alpha_1\otimes\cdots\otimes\alpha_s).$$

Putting this all together we obtain the desired isomorphism $(V^*)_r^s\cong\mathrm{Mult}_r^s(V,\mathbf{R})$. □

Notation 2.27. We will adopt the *Einstein summation convention*: whenever a pair of variables is indexed by the same letter i in both "lower" and "upper" indices, we interpret this as a sum over all possible values of i. For example:

$$\alpha_i v^i=\sum_i\alpha_i v^i.$$

Proposition 2.28 (Change of variables for (r,s)-tensors)**.** *Let V be a vector space and let $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^n$ be two bases of V. Let $B = (b_i^j)$ be a transition matrix between the two bases, i.e.,*

$$f_i = b_i^j e_j.$$

Then for any (r,s)-type tensor $\boldsymbol{t} \in V_r^s$, the change of coordinates between the corresponding two bases of V_r^s is given by

$$\tilde{t}^{i_1,\dots,i_r}_{j_1,\dots,j_s} = t^{k_1,\dots,k_r}_{l_1,\dots,l_s} \beta^{i_1}_{k_1} \dots \beta^{i_r}_{k_r} b^{l_1}_{j_1} \dots b^{l_s}_{j_s},$$

where (β_j^i) is the inverse of B.

Proof. By the multilinearity of tensors we have that

$$\begin{aligned}
\boldsymbol{t} &= t^{k_1,\dots,k_r}_{l_1,\dots,l_s} e_{k_1} \otimes \cdots \otimes e_{k_r} \otimes e^{*l_1} \otimes \cdots e^{*l_s} \\
&= t^{k_1,\dots,k_r}_{l_1,\dots,l_s} \left(\beta^{i_1}_{k_1} f_{i_1}\right) \otimes \cdots \otimes \left(\beta^{i_r}_{k_r} f_{i_r}\right) \otimes \left(b^{l_1}_{j_1} f^{*j_1}\right) \otimes \cdots \left(b^{l_s}_{j_s} f^{*j_s}\right) \\
&= t^{k_1,\dots,k_r}_{l_1,\dots,l_s} \beta^{i_1}_{k_1} \cdots \beta^{i_r}_{k_r} b^{l_1}_{j_1} \cdots a^{l_s}_{j_s} f_{i_1} \otimes \cdots \otimes f_{i_r} \otimes f^{*j_1} \otimes \cdots \otimes f^{*j_s}.
\end{aligned}$$

So we have the desired change of coordinates formula. □

Remark 2.29. From the change of coordinates formula we see why the vector components of V_r^s are called contravariant and the dual components of V_r^s are called covariant – the change of coordinates for the vector components is given by the inverse of the transition matrix (contravariantly) while the change of coordinates for the dual components is given by the original transition matrix (covariantly). ⋄

2.2. **Exterior algebras.** We first review a bit of algebra.

Definition 2.30. Let R be a ring. An (two-sided) **ideal** $\mathfrak{a}$ is an additive subgroup of R such that for all $x \in R$ we have $x\mathfrak{a} \subseteq \mathfrak{a}$. ⋄

Note that every element $r \in R$ can uniquely be written as

$$r = \sum_{n=0}^{\infty} r_n,$$

where $r_n \in R_n$.

Definition 2.31. A **homogeneous ideal** of a ring R is an ideal $\mathfrak{a}$ such that

$$\mathfrak{a} = \bigoplus_{n=0} \mathfrak{a}_n,$$

where $\mathfrak{a}_n := \mathfrak{a} \cap R_n$. ⋄

Note that if $\mathfrak{a}$ is a homogeneous ideal of the ring R then $R/\mathfrak{a}$ is naturally a graded ring with homogeneous part of degree n simply being $R_n/\mathfrak{a}$. If A is a graded R-algebra then this ideal $\mathfrak{a}$ will also be a submodule, and so the quotient will also be a graded R-module with the already specified grading.

Proposition 2.32. *Let R be a graded ring. A two-sided ideal $\mathfrak{a}$ is homogeneous if and only if it is generated by homogeneous elements.*

Proof. Note that if $\mathfrak{a}$ is homogeneous then it is clearly generated by the homogeneous components of all its elements. Conversely, suppose that $\mathfrak{a}$ is generated by a set of homogeneous elements $\{a_i\}_{i\in I}$ for some index set I and $a_i \in R_{n_i}$. Now let $x \in \mathfrak{a}$, and write

$$x = \sum_{i=1}^{n} y_i a_i z_i,$$

where $y_i, z_i \in R$. Then for any $m \in \mathbf{N}$ we have $x_m = \sum_i (y_i a_i z_i)_m$ and so it suffices to consider yaz for some $a \in \{a_i\}_{i\in I} \cap R_n$, $y, z \in R$. Note that

$$(yaz)_m = \sum_{i+j=m} y_i (az)_j = \sum_{i+j=m} y_i a z_{j-n},$$

where $z_k = 0$ for $k < 0$. So we see that $(yaz)_m \in \mathfrak{a}$. Now we see that $(y_i a_i z_i)_m \in \mathfrak{a}_m$ for all $m \in \mathbf{N}$ and so $\mathfrak{a}$ is a two-sided homogeneous ideal. $\square$

The reason we are concerned with homogeneous ideals is that we know that $\bigotimes_* V$ is a graded $\mathbf{R}$-algebra, and so we can start modding out by homogeneous ideals to get new graded $\mathbf{R}$-algebras – in particular, the exterior and symmetric algebras.

Construction 2.33 (Exterior Algebra)**.** Consider the two-sided homogeneous ideal

$$\mathfrak{A}V := \langle v \otimes v \ : \ v \in V \rangle.$$

This is indeed a homogeneous ideal by Proposition 2.32. The **exterior algebra** is defined to be the following graded $\mathbf{R}$-algebra:

$$\bigwedge^* V := \bigotimes\nolimits_* V / \mathfrak{A}V.$$

The homogeneous part of degree k is denoted as $\bigwedge^k V$, and elements of $\bigwedge^k V$ are called k-vectors of V. Note that since $\mathfrak{A}V$ only contains elements of degree 2 or more we have that $\bigwedge^0 V = \mathbf{R}$ and $\bigwedge^1 V = V$. The multiplication in $\bigwedge^* V$ is called **exterior multiplication** and is denoted by $\wedge$.

One can think about the exterior algebra as the tensor algebra where we have forced antisymmetry in the sense that we have something like $v_1 \otimes \cdots \otimes v_n = \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_n$, where $\sigma : [n] \to [n]$ is a permutation. To see that this holds note that

$$(x+y) \otimes (x+y) = x \otimes x + y \otimes y + x \otimes y + y \otimes x,$$

and so

$$x \otimes y + y \otimes x = (x+y) \otimes (x+y) - x \otimes x - y \otimes y \in \mathfrak{A}V.$$

Hence the coset

$$(x \otimes y + y \otimes x) + \mathfrak{A}V = 0 + \mathfrak{A}V \in \bigwedge^* V.$$

Now let

$$v_1 \wedge \cdots \wedge v_n := v_1 \otimes \cdots \otimes v_n + \mathfrak{A}V.$$

So we see that

$$x \wedge y + y \wedge x = 0, \quad \text{and so} \quad x \wedge y = -y \wedge x.$$

To see that both notions of the $\wedge$ coincide (both as an equivalence class/coset and as an exterior multiplication of vectors in V) note that the canonical homomorphism $\bigotimes_* V \to \bigwedge^* V$ maps $v_1 \otimes \cdots \otimes v_n$ onto $v_1 \wedge \cdots \wedge v_n$.

Note that since we have the anticommutativity of the wedge product of two elements $x, y \in V$ we have that for any permutation $\sigma : [n] \to [n]$ that

$$v_1 \wedge \cdots \wedge v_n = \operatorname{sgn}(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}.$$

Proposition 2.34. *For $v_1, \ldots, v_k \in V$ we have that*

$$v_1 \wedge \cdots \wedge v_k = \frac{1}{k!} \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(k)},$$

where Σ_k is the set of permutations on the set $[k] := \{1, \ldots, k\}$.

Proof. Let $\sigma \in \Sigma_k$ be any permutation. Then we see that

$$v_1 \wedge \cdots \wedge v_k = \operatorname{sgn}(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(k)}.$$

Since $|\Sigma_k| = k!$ the result follows immediately

$$\frac{1}{k!} \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(k)} = \frac{1}{k!} v_1 \wedge \cdots \wedge v_k = v_1 \wedge \cdots \wedge v_k.$$

$\square$

Proposition 2.35. *Let $\omega \in \bigwedge^k V$ and $\eta \in \bigwedge^\ell V$. Then*

$$\omega \wedge \eta = (-1)^{k\ell} (\eta \wedge \omega).$$

Proof. By the bilinearity of $\wedge$ it suffices to prove this on simple k-vectors $\omega = v_1 \wedge \cdots \wedge v_k$ and simple ℓ-vectors $\eta = w_1 \wedge \cdots \wedge w_\ell$. The result immediately follows since it requires $k\ell$ transpositions to switch $\omega \wedge \eta$ to $\eta \wedge \omega$. □

Now we see that if k is odd and $\omega \in \bigwedge^k V$ that $\omega \wedge \omega = 0$.

The exterior algebra inherits the following universal mapping property from the tensor algebra:

Theorem 2.36 (UMP of the exterior algebra)**.** *For every anticommutative associative graded algebra A with a unit, every linear map $f : V \to A_1$ can be uniquely extended to an algebra homomorphism of $F : \bigwedge^* V \to A$, which preserves the grading (i.e. $\bigwedge^k V \to A^k$).*

Proof. Let $f : V \to A_1$ be a linear map. By the universal property of the tensor algebra we obtain a unique extension $\widetilde{F} : \bigotimes_* V \to A$, which is an algebra homomorphism that preserves the grading. Since A is anticommutative and $\mathbf{R}$ has a characteristic that is different from 2, we have that $a^2 = 0$ whenever $a \in \mathfrak{A}V$. Hence $\mathfrak{A}V \subseteq \ker \widetilde{F}$, and so $\widetilde{F}$ is divisible by the canonincal homomorphism $\bigotimes_* V \to \bigwedge^* V$. □

We also have a universal mapping property for the n-fold exterior product.

Definition 2.37. A multilinear function $f : \underbrace{V \times \cdots \times V}_{n \text{ times}} \to W$ is called **alternating** if and only if

$$f(v_1, \ldots, v_k) = \operatorname{sgn}(\sigma) h(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$$

for every permutation $\sigma : [n] \to [n]$. ⋄

Definition 2.38. Let $\operatorname{Alt}^n(V, W)$ be the vector space of all alternating multilinear functions $f : V^n \to W$. ⋄

Note that if $f \in \operatorname{Alt}^n(V, W)$ and $\widetilde{F} : V^{\otimes n} \to W$ is the corresponding linear function induced by the universal property of the tensor product then we see $\mathfrak{A}V \cap V^{\otimes n} \subseteq \ker \widetilde{F}$ and so we have a unique linear map $F : \bigwedge^n V \to W$ satisfying

$$F(v_1 \wedge \cdots \wedge v_n) = f(v_1, \ldots, v_n)$$

for $v_1, \ldots, v_n \in V$.

Now by associating f with F we obtain the following natural isomorphism:

$$\operatorname{Alt}^n(V, W) \cong \hom\left(\bigwedge^n V, W\right).$$

This is exactly the universal property of the exterior algebra, which we state as a theorem below:

Theorem 2.39. *Let V be a vector space over $\mathbf{R}$. The **exterior product** is a vector space $\bigwedge^k V$ and an alternating bilinear map, called the exterior multiplication, $\nu : V^k \to \bigwedge^k V$ such that for any vector space W and any alternating multilinear map $f : V^k \to W$ we obtain a* unique *linear map $F : \bigwedge^k V \to W$ such that the following diagram commutes:*

$$\begin{tikzcd} \bigwedge^k V \arrow[dr, dashed, "F"] & \\ V^k \arrow[u, "\nu"] \arrow[r, "f"'] & W \end{tikzcd}$$

In light of this universal mapping property we se that an elements $v_1 \wedge \cdots \wedge v_k$ is nonzero in $\bigwedge^k V$ if and only if there exists some vector space W and an alternating multilinear map $V^k \to W$ such that $(v_1, \ldots, v_k)$ is not sent to zero. In particular, this gives us an interesting view of what it means for $\bigwedge^k V$ to be the zero module: it means that the only alternating maps out of V^k to any vector space is the zero map.

Proposition 2.40. $\bigwedge^k(V^*) \cong \operatorname{Alt}^k(V, \mathbf{R}) \cong (\bigwedge^k V)^*$

Proof. To show the first isomorphism consider the pairing $B : \bigwedge^k(V^*) \times \bigwedge^k(V) \to \mathbf{R}$ given by

$$B(\alpha_1 \wedge \cdots \wedge \alpha_k, v_1 \wedge \cdots \wedge v_k) = \det(\alpha^i(v_j))_{i,j=1}^n,$$

extended bilinearly. To see that B is non-degenerate it suffices to evaluate the pairing on the respective bases. Hence by Lemma 2.25 we deduce that $\bigwedge^k(V^*) \cong (\bigwedge^k V)^*$.

$$\mathrm{Alt}^k(V, \mathbf{R}) \cong \hom\left(\bigwedge^k V, \mathbf{R}\right) = \left(\bigwedge^k V\right)^*.$$

□

Elements of $\mathrm{Alt}^k(V, \mathbf{R}) \cong \bigwedge^k(V^*)$ are called k**-covectors**.

Definition 2.41. Let $f : V \to W$ be a linear map. The **pullback map** is the map $f^* : \bigwedge^k(W^*) \to \bigwedge^k(V^*)$ given by

$$(f^*\varphi)(v_1, \ldots, v_k) = \varphi(f(v_1), \ldots, f(v_k)),$$

where $\varphi \in \bigwedge^k(W^*)$ and $v_i \in V$. Hence, we use Proposition 2.40 to consider φ and $f^*\varphi$ as alternating k-linear maps. ⋄

Proposition 2.42. *$f^* : \bigwedge^*(W^*) \to \bigwedge^*(V^*)$ is an algebra homomorphism.*

Proof. Note that $W^* = \bigwedge^1(W^*)$ and so $f^* : W^* \to \bigwedge^1(V_*)$ is a linear map. So by the universal property of the tensor algebra we find that f^* extends *uniquely* to an algebra homomorphism $\widetilde{f^*} : \bigwedge^*(W^*) \to \bigwedge^*(V^*)$. The uniqueness ensures that $\widetilde{f^*} = f^*$. □

Now we would like to determine the dimension of $\bigwedge^k V$. Note that if $\{v_1, \ldots, v_n\}$ is a basis for V then $\{v_{i_1} \wedge \cdots \wedge v_{i_k}\}_{i_1,\ldots,i_k \in [n]}$ is a spanning set for $\bigwedge^k V$ since these elements with $\otimes$ replacing $\wedge$ are a spanning set for $V^{\otimes k}$. Of course, most of these elements are zero. For example, we can immediately replace this set with

$$\{v_{i_1} \wedge \cdots \wedge v_{i_k} : i_1, \ldots, i_k \in [n],\ i_\ell \neq i_j \text{ for all } \ell \neq j.\}.$$

This set is still redundant since indexed the elements by injections $[k] \to [n]$, but since the wedge product is anticommutative we can index by combinations of k elements of n. In particular, we have that

$$\{v_{i_1} \wedge \cdots \wedge v_{i_k} : i_1, \ldots, i_k \in [n],\ i_1 < \cdots < i_k\}$$

spans $\bigwedge^k V$. So we have the following result:

Proposition 2.43. *Let V be a vector space of dimension n. For $k \le n$ we have that*

$$\dim\left(\bigwedge^k V\right) = \binom{n}{k}.$$

For $k > n$ we have that

$$\bigwedge^k V = 0.$$

The fact that for $k > n$ we have $\bigwedge^k V = 0$ makes sense since there are no nonzero alternating multilinear maps out of V^k.

Example 2.44. Let $V = \mathbf{R}^3$. Then $\bigwedge^2 V \cong \mathbf{R}^3$ and in particular $e_1 \wedge e_2 \mapsto e_3$, $e_1 \wedge e_3 \mapsto -e_2$, and $e_2 \wedge e_3 \mapsto e_1$ defines an isomorphism. Under such identification the wedge is the cross product

$$v_1 \wedge v_2 = v_1 \times v_2.$$

Furthermore,

$$v_1 \wedge v_2 \wedge v_3 = \det([v_1\ v_2\ v_3]).$$

This isomorphism is exactly the Hodge star ($\star$) operation, which we will define soon. ⋄

In light of the fact that $\bigwedge^k V = 0$ for all $k > \dim n$ we see

$$\bigwedge^* V = \bigoplus_{k=0}^{\infty} \bigwedge^k V = \bigoplus_{k=0}^{n} \bigwedge^k V.$$

In particular, we see that $\bigwedge^* V$ is an $\mathbf{R}$-algebra with

$$\dim\left(\bigwedge^* V\right) = \sum_{k=0}^{\infty} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Unsurprisingly, the construction of $\bigwedge$ is functorial, in the sense that the $\bigwedge$ operation can not only take exterior powers of vector spaces, but also exterior powers of linear maps. If $f : V \to W$ is a linear map, then we should be able to construct some linear map $\bigwedge^k f : \bigwedge^k V \to \bigwedge^k W$ which respects compositions and identities.

Theorem 2.45. *For all $k \in \mathbf{N}$ the operator $\bigwedge^k$ is a functor from the category of $\mathbf{R}$-vector spaces to $\mathbf{R}$-vector spaces.*

Proof. Let $f : V \to W$ be a linear map between vector spaces. Then by the universal property of the tensor product we get a map $\bigotimes_k f : \bigotimes_k V \to \bigotimes_k W$ such that

$$\bigotimes_k f(v_1 \otimes \cdots \otimes v_k) = f(v_1) \otimes \cdots \otimes f(v_k).$$

Now by considering the composition

$$\wedge \circ \bigotimes_k f \circ \otimes : V^k \to \bigotimes_k V \to \bigwedge^k W,$$

where $\wedge : \bigotimes_k W \to \bigwedge^k W$ is the quotient map and where $\otimes : V^k \to \bigotimes_k V$ is the multiplication map, we get an alternating k-linear map $V^k \to \bigwedge^k W$ which satisfies

$$(v_1, \ldots, v_k) \mapsto f(v_1) \wedge \cdots \wedge f(v_k).$$

Hence, by the universal property of the exterior product we obtain a unique linear map

$$\begin{gathered} \bigwedge^k f : \bigwedge^k V \to \bigwedge^k W \\ v_1 \wedge \cdots \wedge v_k \mapsto f(v_1) \wedge \cdots \wedge f(v_k). \end{gathered}$$

It is easy to see that $\bigwedge^k (g \circ f) = \bigwedge^k g \circ \bigwedge^k f$ and that $\bigwedge^k (\mathrm{id}_V) = \mathrm{id}_{\bigwedge^k V}$. So we see that $\bigwedge^k$ is indeed a functor. $\square$

Proposition 2.46. *Let V and W be $\mathbf{R}$-vector spaces.*

(1) Let $f : V \to W$ be a surjective linear map. Then, $\bigwedge^k f : \bigwedge^k V \to \bigwedge^k W$ is surjective.
(2) Let $f : V \to W$ be an injective linear map. Then $\bigwedge^k f : \bigwedge^k V \to \bigwedge^k W$ is injective.

Proof.

(1) Let $W_0 \subseteq W$ be a generating set for W. Since f is surjective, we know that there is some subset $V_0 \subseteq V$ such that $f(V_0) = W_0$. Now it is trivial to see that

$$S_W := \{w_{i_1} \wedge \cdots \wedge w_{i_k} \, : \, w_{i_1}, \ldots, w_{i_k} \in W_0\}$$

is a generating set for $\bigwedge^k V$, and moreover that if

$$S_V := \{v_{i_1} \wedge \cdots \wedge v_{i_k} \, : \, v_{i_1}, \ldots, v_{i_k} \in V_0\}$$

then

$$\bigwedge^k (f)(S_V) = S_W.$$

So $\bigwedge^k f$ hits a generating set of $\bigwedge^k W$, and thus by linearity we deduce that $\bigwedge^k f$ is surjective.

(2) Note that since $\operatorname{im}(f)$ is a vector subspace of W have have a split short exact sequence

$$0 \longrightarrow M \xrightarrow{f} \underbrace{\operatorname{im}(f) \oplus Z}_{W} \xrightarrow{\pi} Z \longrightarrow 0,$$

where Z is the complement of $\operatorname{im}(f)$ in W. Now the result immediately follows since functors take retractions to retractions. More explicitly, since our sequence splits we get a retraction map $s : W \to V$ such that $s \circ f = \mathrm{id}_V$. Then $\bigwedge^k(s) \circ \bigwedge^k(f) = \mathrm{id}_{\bigwedge^k V}$, and therefore $\bigwedge^k(f)$ is injective.

□

Determinants. The exterior powers of maps relate to determinants in a very concrete way. This definition of a determinant will be used when we define the Hodge dual.

Let V be a vector space with $\dim V = n$, let $\mathscr{B} = \{v_1, \dots, v_n\}$ be a basis of V, and consider a linear map $f : V \to V$. Recall that associated to f is a unique matrix $[f]_{\mathscr{B}} \in \mathbf{R}^{n\times n}$ defined by making that $(i,j)^{\text{th}}$ entry of $[f]_{\mathscr{B}}$ the unique $a_{i,j} \in \mathbf{R}$ making

$$f(v_j) = \sum_{i=1}^{n} a_{i,j} v_i.$$

Moreover, the association $f \mapsto [f]_{\mathscr{B}}$ is actually an $\mathbf{R}$-algebra isomorphism $\operatorname{End}_{\mathbf{R}}(V) \to \mathbf{R}^{n\times n}$. Moreover, we know that if we define ϕ to be the unique isomorphism $V \to \mathbf{R}^n$ with

$$\phi(v_i) = e_i$$

then $f = \phi^{-1} \circ [f]_{\mathscr{B}} \circ \phi$ where $[f]_{\mathscr{B}}$ acts on $\mathbf{R}^n$ in the usual way (i.e. matrix-vector multiplication).

So we see that we can think about endomorphisms $V \to V$ as matrices, as long as we keep track of bases. Moreover, we can define the determinant of a matrix $A = (a_{i,j})$ in $\mathbf{R}^{n\times n}$ as follows:

$$\det(A) = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Moreover, we define the adjugate matrix $\operatorname{adj}(A)$ to be the matrix whose $(i,j)^{\text{th}}$ coordinate is computed by taking the determinant of A after removing the j^{th} row and i^{th} column, and multiplying by $(-1)^{i+j}$. Now from basic linear algebra we see that

$$\operatorname{adj}(A)A = A\operatorname{adj}(A) = \det(A) I_{n\times n}.$$

So we see that $A \in \mathbf{GL}_n(\mathbf{R})$ if and only if $\det(A) \neq 0$.

Now what does this have to do with the exterior algebra, well note that $\dim \bigwedge^n V = 1$ and so we see that for every linear map $f : V \to V$ that

$$\bigwedge^n f : \bigwedge^n V \to \bigwedge^n V \cong \mathbf{R},$$

and so there is a unique constant $C_f \in \mathbf{R}$ such that

$$\bigwedge^n f(v_1 \wedge \cdots \wedge v_n) = C_f(v_1 \wedge \cdots \wedge v_n).$$

Theorem 2.47.

$$\bigwedge^n f(v_1 \wedge \cdots \wedge v_n) = \det([f]_{\mathscr{B}})(v_1 \wedge \cdots \wedge v_n).$$

Proof. Write $[f]_{\mathscr{B}} = (a_{i,j})$. We now have

$$\begin{aligned}\bigwedge^n f(v_1 \wedge \cdots \wedge v_n) &= f(v_1) \wedge \cdots \wedge f(v_n) \\ &= \left(\sum_{i_1=1}^{n} a_{i_1,1} v_{i_1}\right) \wedge \cdots \wedge \left(\sum_{i_n=1}^{n} a_{i_n,n} v_{i_n}\right) \\ &= \sum_{i_1,\dots,i_n=1}^{n} a_{i_1,1} \cdots a_{i_n,n} (v_{i_1} \wedge \cdots \wedge v_{i_n}).\end{aligned}$$

Now note that any $(i_1, \dots, i_n)$ with repeated numbers results in $v_{i_1} \wedge \cdots \wedge v_{i_n} = 0$, and so we may consider the last sum only over indices which are in bijection with $[n]$, i.e. only permuations of $[n]$, and so we see that

$$\bigwedge^n(f)(v_1 \wedge \cdots \wedge v_n) = \sum_{\sigma \in \Sigma_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}(v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}),$$

but since $v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)} = \operatorname{sgn}(\sigma) v_1 \wedge \cdots \wedge v_n$ we have that

$$\begin{aligned}
\bigwedge^n(f)(v_1 \wedge \cdots \wedge v_n) &= \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}(v_1 \wedge \cdots \wedge v_n) \\
&= \left(\sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \right)(v_1 \wedge \cdots \wedge v_n) \\
&= \det\left([f]_{\mathscr{B}}^{\top}\right)(v_1 \wedge \cdots \wedge v_n) \\
&= \det([f]_{\mathscr{B}})(v_1 \wedge \cdots \wedge v_n).
\end{aligned}$$

□

Note that if we have selected a different basis $\mathscr{B}'$ for V then we would know from the above theorem that $\bigwedge^n(f)$ is simply multiplication by $\det([f]_{\mathscr{B}'})$. In particular, we see that

$$\det([f]_{\mathscr{B}})(v_1 \wedge \cdots \wedge v_n) = \bigwedge^n(f)(v_1 \wedge \cdots \wedge v_n) = \det([f]_{\mathscr{B}'})(v_1 \wedge \cdots \wedge v_n),$$

and since $v_1 \wedge \cdots \wedge v_n \neq 0$ we deduce that

$$\det([f]_{\mathscr{B}'}) = \det([f]_{\mathscr{B}}).$$

In other words, the determinant is a well-defined invariant of a linear transformation, i.e. the determinant does not depend on the choice of basis.

The functoriality of $\bigwedge^n$ gives us a very quick proof of the multiplicativeness of the determinant.

Proposition 2.48. *Let $f, g : V \to V$ be linear maps. Then* $\det(g \circ f) = \det(g)\det(f)$*.*

Proof. Let $\boldsymbol{v} \in \bigwedge^n V$. Then we know that

$$\begin{aligned}
\det(g \circ f)\boldsymbol{v} &= \bigwedge^n(g \circ f)(\boldsymbol{v}) \\
&= \left(\bigwedge^n(g) \circ \bigwedge^n(f)\right)(\boldsymbol{v}) \\
&= \bigwedge^n(g)(\det(f)\boldsymbol{v}) \\
&= \det(g)\det(f)\boldsymbol{v}.
\end{aligned}$$

□

To tie this back to geometry, we prove a basic fact from linear algebra and measure theory.

Theorem 2.49. *Let $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^n$ be linear. Then for every Lebesgue measurable set $E \subseteq \mathbf{R}^n$, $\mathbf{L}(E)$ is Lebesgue measurable and*

$$\mathscr{L}^n(\mathbf{L}(E)) = |\det(\mathbf{L})| \mathscr{L}^n(E).$$

Proof. If $\det(\mathbf{L}) = 0$ then $\mathbf{L}(\mathbf{R}^n)$ is a subspace of $\mathbf{R}^n$ of dimension less than n, and in this case $\mathscr{L}^n(\mathscr{L}(\mathbf{R}^n)) = 0$. Since $\mathscr{L}^n$ is a complete measure, we have for every measurable set E that

$$0 = \mathscr{L}^n(\mathbf{L}(E)) = |\det(\mathbf{L})| \mathscr{L}^n(E),$$

as desired.

Now suppose that $\det(\mathbf{L}) \neq 0$. So we have that $\mathbf{L}$ is invertible with continuous inverse. In particular, $\mathbf{L}(E)$ is the inverse image of the set E through the continuous function $\mathbf{L}^{-1}$; so if E is Borel, then so is $\mathbf{L}(E)$. So we can define the measure $\mu(E) := \mathscr{L}^n(\mathbf{L}(E))$ for $E \in \mathscr{B}(\mathbf{R}^n)$. By the linearity of $\mathbf{L}$ we deduce that μ is translation invariant.

Since $\mathscr{L}^n$ is the unique translation invariant Borel measure we have that there exists some constant $c \geq 0$ such that

$$\mu(E) = c\mathscr{L}^n(E), \qquad E \in \mathscr{B}(\mathbf{R}^n). \tag{1}$$

If $E \subseteq \mathbf{R}^n$ is just Lebesgue measurable then we can find two Borel sets F and G such that $F \subseteq E \subseteq G$ such that $\mathscr{L}^n(G \setminus F) = 0$. By (1) we see that $\mathscr{L}^n(\mathbf{L}(G \setminus F)) = 0$, and so since $E \setminus F \subseteq G \setminus F$ we have that $\mathbf{L}(E \setminus F) \subseteq \mathbf{L}(G \setminus F)$, which, again by completeness, implies that $\mathbf{L}(E \setminus F)$ is measurable. In turn, $\mathbf{L}(E) = \mathbf{L}(F) \cup \mathbf{L}(E \setminus F)$ is Lebesgue measurable, since $\mathbf{L}(F)$ is Borel.

Note that if $\mathbf{L}$ is a rotation, then $\mathbf{L}(B(\mathbf{0},1)) = B(\mathbf{0},1)$, and so

$$\mathscr{L}^n(B(\mathbf{0},1)) = \mathscr{L}^n(\mathbf{L}(B(\mathbf{0},1))) = c\mathscr{L}^n(B(\mathbf{0},1)),$$

which implies that $c = 1$. Since $|\det \mathbf{L}| = 1$, we have that the desired formula holds in this case.

Now recall that any invertible linear transformation can be written as a composition of linear invertible transformations of three basic types:

$$\begin{aligned}
\mathbf{T}_s(\boldsymbol{x}) &:= (sx_1, \ldots, x_2, \ldots, x_n),\\
\mathbf{A}(\boldsymbol{x}) &:= (x_1, x_1 + x_2, \ldots, x_n),\\
\mathbf{S}_{ij}(\boldsymbol{x}) &:= \mathbf{S}_{ij}(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n).
\end{aligned}$$

Now since the determinant of the composition of two linear transformations is their product, it suffices to verify the result for these three basic types.

- Note that $|\det(\mathbf{T}_s)| = |s|$, while $\mathbf{T}_s([0,1]^n) = [0,s] \times [0,1]^{n-1}$ if $s > 0$, and $\mathbf{T}_s([0,1]^n) = [s,0] \times [0,1]^{n-1}$ if $s < 0$. In both cases
$$\mathscr{L}^n(\mathbf{T}_s([0,1]^n)) = |s| = c,$$
and so the result holds.
- Note that $|\det \mathbf{A}| = 1$, and $\mathbf{A}([0,1)^n) = \{\boldsymbol{y} \in \mathbf{R}^n \,:\, y_1 \leq y_2 < y_1 + 1, y_i \in [0,1), \text{ for all } i \neq 2\}$. Let
$$\begin{aligned}
F_1 &:= \{\boldsymbol{x} \in \mathbf{A}([0,1)^n) \,:\, x_2 < 1\} = \{\boldsymbol{y} \in \mathbf{R}^n \,:\, y_1 \leq y_2 < y_1 + 1,\ y_i \in [0,1) \text{ for all } i \neq 2,\ y_2 < 1\}.,\\
F_2 &:= \mathbf{A}([0,1)^n) \setminus F_1 = \{\boldsymbol{y} \in \mathbf{R}^n \,:\, y_1 \leq y_2 < y_1 + 1,\ y_i \in [0,1) \text{ for all } i \neq 2,\ y_2 \geq 1\}.
\end{aligned}$$
Then
$$-e_2 + F_2 = \{\boldsymbol{z} \in \mathbf{R}^n \,:\, z_1 - 1 \leq z_2 < z_1,\ z_i \in [0,1) \text{ for all } i \neq 2,\ z_2 \geq 0\},$$
and so $F_1 \cup (-e_2 + F_2) = [0,1)^n$, and $F_1 \cap (-e_2 + F_2) = \emptyset$. Hence,
$$\begin{aligned}
c\mathscr{L}^n([0,1)^n) &= \mathscr{L}^n(\mathbf{A}([0,1)^n)) = \mathscr{L}^n(F_1 \cup F_2) = \mathscr{L}^n(F_1) + \mathscr{L}^n(F_2)\\
&= \mathscr{L}^n(F_1) + \mathscr{L}^n(-e_2 + F_2) = \mathscr{L}^n(F_1 \cup (-e_2 + F_2)) = \mathscr{L}^n([0,1)^n) = 1.
\end{aligned}$$
Again, the result holds in this case.
- Finally, $|\det \mathbf{S}_{ij}| = 1$, and $S_{ij}([0,1]^n) = [0,1]^n$, and so again $c = 1$, and the result holds in this case.

□

This theorem shows us that the absolute value of the determinant measures the volume spanned by the vectors $L(e_i)$ where e_i is the standard basis of $\mathbf{R}^n$. More generally, the determinant provides us with a notion of signed volume which depends on the orientation of the vectors. In particular, if we are working in the tangent space of a manifold then the geometric interpretation of the determinant is the same. This is one way to intuit the definition of volume form, i.e. a measure on the manifold.

2.3. **Tensor fields and differential forms.** To tie this back to geometry and manifold theory, we would like to have tensor fields that vary smoothly across our manifolds. To introduce this we will consider tensor bundles, which are in some sense generalizations of the tangent and cotangent bundles.

Definition 2.50. The (r,s)**-tensor bundle** $T^s_r M$ is the vector bundle over M for which a fiber over a point $p \in M$ is $(T_pM)^s_r = (T_pM)^{\otimes r} \otimes (T_pM^*)^{\otimes s}$. ⋄

We have a natural coordinate chart associated to the (r,s)-tensor bundle. In particular, if $(U,\varphi : U \to M)$ is a coordinate chart for M then we have $\Phi : U \times \mathbf{R}^{n^{r+s}} \to T_r^s M$ given by

$$(\boldsymbol{x},\boldsymbol{a}) \mapsto \left(\phi(\boldsymbol{x}), a^{i_1,\dots,i_r}_{j_1,\dots,j_s} \frac{\partial}{\partial x_{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x_{i_r}} \otimes dx_{j_1} \otimes \cdots \otimes dx_{j_s}\right),$$

where we used the Einstein summation convention, and where $\dim M = n$.

Definition 2.51. A (smooth) **tensor field** (with abuse of terminology we also call these just tensors) is a (smooth) section of a tensor bundle. We simply write $T \in T_r^s M$ to denote that T is a smooth (r,s)-tensor field. ⋄

Remark 2.52. Note that sections of $T_1^0 M$ are simply vector fields on M. ⋄

Let (U,φ) and (V,ψ) be two coordinate charts such that there exists some $p \in \varphi(U) \cap \psi(V)$. Let $T \in T_r^s M$ be a smooth tensor field. Then in a neighborhood of p we have smooth functions $a^{i_1,\dots,i_r}_{j_1,\dots,j_s}$ and $b^{k_1,\dots,k_r}_{l_1,\dots,l_s}$ such that

$$\begin{aligned} T &= a^{i_1,\dots,i_r}_{j_1,\dots,j_s} \frac{\partial}{\partial x_{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x_{i_r}} \otimes dx_{j_1} \otimes \cdots \otimes dx_{j_s} \\ &= b^{i_1,\dots,i_r}_{j_1,\dots,j_s} \frac{\partial}{\partial y_{k_1}} \otimes \cdots \otimes \frac{\partial}{\partial y_{k_r}} \otimes dy_{l_1} \otimes \cdots \otimes dy_{l_s}, \end{aligned}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the coordinates induced by φ and ψ, respectively. Now we want to find the relationship between the coefficients. A direct application of Proposition 2.28 tells us that

$$b^{m_1,\dots,m_r}_{q_1,\dots,q_s} = a^{i_1,\dots,i_r}_{j_1,\dots,j_s} \frac{\partial y_{m_1}}{\partial x_{i_1}} \cdots \frac{\partial y_{m_r}}{\partial x_{i_r}} \frac{\partial x_{j_1}}{\partial y_{q_1}} \cdots \frac{\partial x_{j_s}}{\partial y_{q_s}}.$$

Now we can construct the exterior bundle.

Definition 2.53. The **exterior k-bundle** $\bigwedge^k M^*$ is the fiber bundle with fibers $\bigwedge^k(T_p M^*)$. The **exterior bundle** $\bigwedge M$ is the fiber bundle with fibers $\bigwedge(T_p M^*)$. ⋄

In a local chart (U,φ) of M we see that the natural basis of $\bigwedge^k M^*$ at a point $p \in \varphi(U)$ is given by

$$\{dx^{i_1} \wedge \cdots \wedge dx^{i_k} \;:\; 1 \le i_1 < \cdots < i_k \le n\}.$$

It isn't hard to check that $\bigwedge^k M^*$ is a smooth manifold of dimension $n + \binom{n}{k}$. Similarly, the dimension of $\bigwedge M$ is $n + 2^n$.

Definition 2.54. A section of the exterior k-bundle $\bigwedge^k M^*$ is called **smooth k-form** on M. A **differential form** on M is a section of the exterior bundle $\bigwedge(M)$. The space of k-forms is denoted by $\Omega^k(M)$ and the space of all differential forms is denoted by $\Omega(M)$. ⋄

Note that by definition $\Omega(M)$ is simply $\mathscr{C}^\infty(M)$, the space of smooth functions on M. Also note that every differential form ω can be uniquely written as

$$\omega = \omega_0 + \omega_1 + \cdots + \omega_\ell,$$

where each ω_k is a k-form in $\Omega^k(M)$. There is a canonical exterior product structure on the $\mathbf{R}$-vector space $\Omega(M)$ induced by the exterior product over each fiber pointwise. It turns out that $\Omega(M)$ is an infinite rank graded algebra over $\mathbf{R}$.

Remark 2.55. From here on out, we will always consider the differential df of a smooth function $f \in \mathscr{C}^\infty(M)$ as a 1-form on M. Note that if $f : \mathbf{R}^n \to \mathbf{R}$ is a smooth function, then in the standard coordinates we have

$$df = \frac{\partial f}{\partial x_1} dx^1 + \cdots + \frac{\partial f}{\partial x_n} dx^n.$$

⋄

Consider a differential k-form $\omega \in \Omega^k(M)$. Then $\omega_p \in \bigwedge^k(T_pM^*)$ for all $p \in M$, and by the duality pairing 2.40 we can consider ω_p as an alternating multilinear function on T_pM. So if $X_1, \ldots, X_k \in \mathfrak{X}(M)$ are vector fields on M, then $\omega(X_1, \ldots, X_k)$ is a smooth function on M and is given by

$$\omega(X_1, \ldots, X_k)(p) = \omega_p(X_1|_p, \ldots, X_k|_p),$$

alternatively, we also have

$$\omega(X_1, \ldots, X_k) = \langle \omega, X_1 \wedge \cdots \wedge X_k \rangle.$$

In particular, since $\mathfrak{X}(M)$ is a smooth $\mathscr{C}^\infty(M)$ module on M, we see that

$$\omega : \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k \text{ times}} \to \mathscr{C}^\infty(M)$$

is an alternating multilinear map over $\mathfrak{X}(M)$ into $\mathscr{C}^\infty(M)$. It is important to note that ω is multilinear over the *$\mathscr{C}^\infty$ module* $\mathfrak{X}(M)$, i.e.

$$\omega(X_1, \ldots, fX + gY, \ldots, X_k) = f\,\omega(X_1, \ldots, X, \ldots, X_k) + g\,\omega(X_1, \ldots, Y, \ldots, X_k),$$

for $f, g \in \mathscr{C}^\infty(M)$ and $X_1, \ldots, X_k, X, Y \in \mathfrak{X}(M)$.

Proposition 2.56. *Every alternating $\mathscr{C}^\infty(M)$ multilinear map over the module $\mathfrak{X}(M)$ into $\mathscr{C}^\infty(M)$ defines a differential form.*

Proof. Let $\omega : \mathfrak{X}(M)^k \to \mathscr{C}^\infty(M)$ be a k-linear map over the $\mathscr{C}^\infty(M)$ module $\mathfrak{X}(M)$. We claim that $\omega(X_1, \ldots, X_k)(p)$ only depends on the values of $X_i|_p$. If this holds, then ω gives rise to an alternating multilinear function ω_p on T_pM, and hence gives rise to an element of $\bigwedge^k(T_pM^*)$ in the following way: let $v_1, \ldots, v_k \in T_pM$, and consider smooth vector fields $V_1, \ldots, V_k \in \mathfrak{X}(M)$ satisfying $V_i(p) = v_i$ for $i = 1, \ldots, k$. Now let

$$\omega_p(v_1, \ldots, v_k) := \omega(V_1, \ldots, V_k)(p).$$

By our claim, we see that $\omega_p(v_1, \ldots, v_k)$ is well defined. So we see that ω gives rise to a smooth section $p \mapsto \omega_p$ of $M \to \bigwedge^k M^*$. Hence ω is a smooth k-form.

Now we prove the claim. Fix $p \in M$, and let $\{E_i\}$ be a frame field defined on a neighborhood U of p, that is a vector field $E_i \in \mathfrak{X}(M)$ such that for all $p \in M$ we have that $\{E_i(p)\}_{i=1}^n$ defines a basis of T_pM. Writing each vector field $X_1, \ldots, X_k$ in terms of the basis gives us $X_j = E_i(X_j)^i$ for some smooth functions $(X_j)^i$. By the multilinearity of ω over $\mathfrak{X}(M)$ we find that

$$\omega(X_1, \ldots, X_k) = (X_1)^{i_1} \cdots (X_k)^{i_k} \omega(E_{i_1}, \ldots, E_{i_k}).$$

Hence,

$$\omega(X_1, \ldots, X_k)(p) = (X_1)^{i_1}(p) \cdots (X_k)^{i_k}(p) \omega(E_{i_1}, \ldots, E_{i_k})(p).$$

Now let $Y_1, \ldots, Y_k \in \mathfrak{X}(M)$ be another set of vector fields on U satisfying $Y_j(p) = X_j(p)$ for $j = 1, \ldots, k$. In particular, we have that $(Y_j)^i(p) = (X_j)^i(p)$ for $j = 1, \ldots, k$ and $i = 1, \ldots, n$. Then we see that

$$\begin{aligned} \omega(Y_1, \ldots, Y_k)(p) &= (Y_1)^{i_1} \cdots (Y_k)^{i_k} \omega(E_{i_1}, \ldots, E_{i_k}) \\ &= (X_1)^{i_1} \cdots (X_k)^{i_k} \omega(E_{i_1}, \ldots, E_{i_k}). \end{aligned}$$

So the claim holds, and the proof is complete. □

Summarizing the above results, we have the following theorem.

Theorem 2.57. *$\Omega^k(M)$ is canonically isomorphic as a $\mathscr{C}^\infty(M)$-module to the $\mathscr{C}^\infty(M)$-module of alternating $\mathscr{C}^\infty(M)$-multilinear maps*

$$\underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{k \text{ times}} \to \mathscr{C}^\infty(M).$$

We have a similar result for tensor fields $T \in \Gamma(T^s_rM)$ by using the duality pairing 2.26.

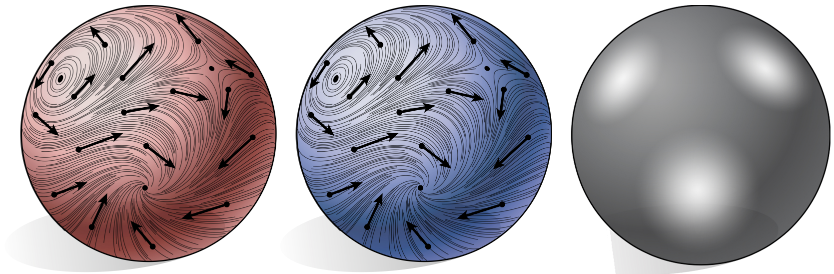


FIGURE 5. A visualization of the duality between differential forms and vector fields

Theorem 2.58. *The set of (r,s)-type tensor fields, $\Gamma(T^s_r M)$, is canonically isomorphic as a $\mathscr{C}^\infty(M)$-module to the $\mathscr{C}^\infty(M)$-module of $\mathscr{C}^\infty(M)$-multilinear maps*

$$\underbrace{\Omega^1(M)\times\cdots\times\Omega^1(M)}_{r\text{ times}}\times\underbrace{\mathfrak{X}(M)\times\cdots\times\mathfrak{X}(M)}_{s\text{ times}}\to\mathscr{C}^\infty(M).$$

Remark 2.59. There is another interpretation of (r,s)-tensor fields that I don't like as much since it depends on the specific representation of vectors and covectors on M.

Recall that a tangent vector $v\in T_pM$ is a derivation, i.e. a linear map $v:\mathscr{C}^\infty(M)\to\mathbf{R}$. Similarly, covectors are simply linear functionals $\alpha: T_pM\to\mathbf{R}$. So if we have $v_1,\dots,v_r\in T_pM$ and $\alpha_1,\dots,\alpha_s\in(T_pM)^*$ we can use the functoriality of the tensor product to constuct a linear map

$$v_1\otimes\cdots\otimes v_r\otimes\alpha_1\otimes\cdots\otimes\alpha_s:(T^s_rM)_p\to\mathbf{R},$$

satisfying

$$(v_1\otimes\cdots\otimes v_r\otimes\alpha_1\otimes\cdots\otimes\alpha_s)(f_1\otimes\cdots\otimes f_r\otimes u_1\otimes\cdots\otimes u_s)=v_1(f_1)\cdots v_r(f_r)\alpha_1(u_1)\cdots\alpha_s(u_s),$$

and extended linearly using the universal property of the tensor product. Since linear maps from the tensor product are in natural bijection with multilinear maps from the Cartesian product, we obtain the desired multilinear map.

I will not use this interpretation, but since I have seen this being used several times I decided to comment on it. ⋄

Definition 2.60. Given a smooth map $f:M\to N$ we can pullback covariant tensors and differential forms on N back onto M. Let S be a $(0,s)$-type tensor on N. The **pullback** of S by f is the $(0,s)$-tensor f^*S on M defined as follows:

$$f^*S|_p(v_1,\dots,v_s)=S_{f(p)}(\mathrm{d}f(v_1),\dots,\mathrm{d}f(v_s)),$$

for $v_1,\dots,v_s\in T_pM$. Note that here we use the interpretation of a $(0,s)$-tensor as a s-multilinear map over the $\mathscr{C}^\infty(M)$ module $\mathfrak{X}(M)$. ⋄

Proposition 2.61. *Let $f:M\to N$ and $g:N\to P$ be smooth maps between manifolds. Let $S_1\in\Gamma(T^{s_1}_0N)$ and $S_2\in\Gamma(T^{s_2}_0N)$, $\omega,\omega_1,\omega_2\in\Omega(N)$. The follow properties hold:*

(i) $f^*(S_1\otimes S_2)=f^*(S_1)\otimes f^*(S_2)$,
(ii) $f^*(\omega_1\wedge\omega_2)=f^*(\omega_1)\wedge f^*(\omega_2)$,
(iii) $(g\circ f)^*=f^*\circ g^*$.

Proof.

(i) Consider $\boldsymbol{v}=(v_1,\dots,v_{s_1})$ and $\boldsymbol{w}=(w_1,\dots,w_{s_2})$. By definition we have

$$\begin{aligned}f^*(S_1\otimes S_2)(\boldsymbol{v},\boldsymbol{w})&=(S_1\otimes S_2)(\mathrm{d}f(\boldsymbol{v}),\mathrm{d}f(\boldsymbol{w}))\\&=S_1(\mathrm{d}f(\boldsymbol{v}))S_2(\mathrm{d}f(\boldsymbol{w}))\\&=(f^*S_1)(\boldsymbol{v})(f^*S_2)(\boldsymbol{w})\\&=((f^*S_1)\otimes(f^*S_2))(\boldsymbol{v},\boldsymbol{w}).\end{aligned}$$

Since $\boldsymbol{v}$ and $\boldsymbol{w}$ were arbitrary we conclude that

$$f^*(S_1 \otimes S_2) = f^*(S_1) \otimes f^*(S_2).$$

(ii) Follows immediately from (i).

(iii) Let S be a $(0,s)$-type tensor field on P, and $\boldsymbol{v} = (v_1, \ldots, v_s) \in (T_pM)^s$.

$$(g \circ f)^* S(\boldsymbol{v}) = S(\mathrm{d}(g \circ f)(\boldsymbol{v})) = S(\mathrm{d}g \circ \mathrm{d}f(\boldsymbol{v})) = g^*(S)(\mathrm{d}f(\boldsymbol{v})) = (f^* \circ g^*)(S)(\boldsymbol{v}).$$

□

Example 2.62. Consider the polar coordinates map $f : (0, \infty) \times (0, 2\pi) \to \mathbf{R}^2$ given by

$$f(r, \theta) = (x(r, \theta), y(r, \theta)) = (r \cos\theta, r \sin\theta).$$

f is a diffeomorphism onto the open set

$$U := \mathbf{R}^2 \setminus \{(x, 0) : x \geq 0\}.$$

Now we see that

$$f^*(dx) = d(r\cos\theta) = \cos\theta\, dr - r \sin\theta\, d\theta,$$
$$f^*(dy) = d(r\sin\theta) = \sin\theta\, dr + r \cos\theta\, d\theta.$$

By using Proposition 2.61 we can easily compute

$$\begin{aligned} f^*(dx \wedge dy) &= f^*(dx) \wedge f^*(dy) \\ &= (\cos\theta\, dr - r\sin\theta\, d\theta) \wedge (\sin\theta\, dr + r\cos\theta\, d\theta) \\ &= r(\cos^2\theta + \sin^2\theta) dr \wedge d\theta = r\, dr \wedge d\theta. \end{aligned}$$

We can also compute

$$f^*(dx \otimes dx + dy \otimes dy) = dr \otimes dr + r^2 d\theta \otimes d\theta.$$

⋄

2.4. **Exterior differentiation.** We want to turn the exterior algebra into a natural framework for calculus on manifolds. The exterior derivative allows us to differentiate arbitrary k-forms – this is one of the central tools when doing analysis on manifolds.

Theorem 2.63. *There is a unique linear map $d : \Omega(M) \to \Omega(M)$ such that*

(i) $d(\Omega^k(M)) \subseteq \Omega^{k+1}(M)$,
(ii) For any smooth function $f \in \Omega^0(M) = \mathscr{C}^\infty(M)$, df is the differential of f,
(iii) $d^2 = d \circ d = 0$,
(iv) $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{|\alpha|} \alpha \wedge d\beta$, where $|\alpha|$ is the rank of α.

Proof.

Step 1. Uniqueness.

Let U be a local chart on M with local coordinates $(x_1, \ldots, x_n)$. On U, any k-form $\omega \in \Omega^k(M)$ can be written as

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \omega_{i_1, \ldots, i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

By the product rule and since $d \circ d = 0$ we see that

$$\begin{aligned} d\omega &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} (d\omega_{i_1, \ldots, i_k}) \wedge (dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \\ &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} \left(\frac{\partial \omega_{i_1, \ldots, i_k}}{\partial x_i} dx^i \right) \wedge (dx^{i_1} \wedge \cdots \wedge dx^{i_k}). \end{aligned}$$

So $d\omega$ is uniquely determined on any coordinate chart of M, and the uniqueness of d immediately follows.

Step 2. Existence.

Let $\omega \in \Omega^k(M)$ be a k-form. For each coordinate chart U in M we define $d\omega|_U$ via the above formula:

$$d\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \left(\frac{\partial \omega_{i_1,\ldots,i_k}}{\partial x_i} dx^i \right) \wedge (dx^{i_1} \wedge \cdots \wedge dx^{i_k}).$$

Now we need to show that $d\omega$ is well defined in the sense that $d\omega|_U = d\omega|_V$ on $U \cap V$ for coordinate charts U, V in M such that $U \cap V \neq \emptyset$.

Let $(x_1, \ldots, x_n)$ be the local coordinates in U and $(y_1, \ldots, y_n)$ be the local coordinates in V. Then we see that in U we have

$$\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \omega_{i_1,\ldots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

and so

$$d\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \left(\frac{\partial \omega_{i_1,\ldots,i_k}}{\partial x_i} dx^i \right) \wedge (dx^{i_1} \wedge \cdots \wedge dx^{i_k}).$$

In V we have

$$\omega = \sum_{1 \le j_1 < \cdots < j_k \le n} \widetilde{\omega}_{j_1,\ldots,j_k} dy^{j_1} \wedge \cdots \wedge dy^{j_k},$$

and so

$$d\omega = \sum_{1 \le j_1 < \cdots < j_k \le n} \left(\frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial y_j} dy^j \right) \wedge (dy^{j_1} \wedge \cdots \wedge dy^{j_k}).$$

Since the exterior product is anticommutative we see that the component functions $\omega_{i_1,\ldots,i_k}$ and $\widetilde{\omega}_{j_1,\ldots,j_k}$ are skew-symmetric, i.e. for all $\sigma \in \Sigma_k$ we have

$$\omega_{i_{\sigma(1)},\ldots,i_{\sigma(k)}} = \operatorname{sgn}(\sigma) \omega_{i_1,\ldots,i_k},$$

and similarly for $\widetilde{\omega}_{j_1,\ldots,j_k}$. Since $\omega|_U = \widetilde{\omega}|_V$ on $U \cap V$ we have

$$\omega_{i_1,\ldots,i_k} = \frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \widetilde{\omega}_{j_1,\ldots,j_k}.$$

So we deduce that

$$\frac{\partial \omega_{i_1,\ldots,i_k}}{\partial x_i} = \sum_{\ell=1}^{k} \left(\frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial^2 y_{j_\ell}}{\partial x_i \partial x_{i_\ell}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \widetilde{\omega}_{j_1,\ldots,j_k} + \frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial x_i} \right).$$

We deduce that

$$\begin{aligned}\sum_i \sum_{i_1,\ldots,i_k} \frac{\partial \omega_{i_1,\ldots,i_k}}{\partial x_i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} &= \sum_i \sum_{\ell=1}^{k} \frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial^2 y_{j_\ell}}{\partial x_i \partial x_{i_\ell}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \widetilde{\omega}_{j_1,\ldots,j_k} dx^i \wedge \cdots dx^{i_1} \wedge \cdots \wedge dx^{i_k} \\ &\quad + \sum_i \sum_{\ell=1}^{k} \frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial x_i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.\end{aligned}$$

Observe that mixed partials commute: $\partial_{x_i, x_{i_k}} = \partial_{x_{i_k}, x_i}$, but that exterior products anticommute: $dx^i \wedge dx^{i_k} = -dx^{i_k} \wedge dx^i$. So we see that the first term in the right hand side of the above expression is identically zero. So on $U \cap V$ we have that

$$\begin{aligned}\frac{\partial \omega_{i_1,\ldots,i_k}}{\partial x_i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} &= \frac{\partial y_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial y_{j_k}}{\partial x_{i_k}} \frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial x_i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \\ &= \left(\frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial x_i} dx^i \right) \wedge \left(\frac{\partial y_{j_1}}{\partial x_{i_1}} dx^{i_1} \right) \wedge \cdots \wedge \left(\frac{\partial y_{j_k}}{\partial x_{i_k}} dx^{i_k} \right) \\ &= (d\widetilde{\omega}_{j_1,\ldots,j_k}) \wedge dy^{j_1} \wedge \cdots \wedge dy^{j_k} \\ &= \frac{\partial \widetilde{\omega}_{j_1,\ldots,j_k}}{\partial y_j} dy^j \wedge dy^{j_1} \wedge \cdots \wedge dy^{j_k}.\end{aligned}$$

So we see that $d\omega|_U = d\omega|_V$ on $U \cap V$. Hence d is a well defined map. Note that d is clearly linear and is a map $\Omega^k(M) \to \Omega^{k+1}(M)$ for all $k \in \mathbf{N}$. Furthermore, clearly df is the differential of f for all smooth functions $f \in \mathscr{C}^\infty(M)$.

Step 3. Verifying the product rule.
Since d is linear and a local operator it suffices to prove the product rule locally on simple differential forms. Let $\alpha = f\,dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ and $\beta = g\,dx^{j_1} \wedge \cdots \wedge dx^{j_\ell}$. Then we have

$$\begin{aligned}
d(\alpha \wedge \beta) &= d(f g dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell}) \\
&= d(fg) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell} \\
&= (df \cdot g + f \cdot dg) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell} \\
&= df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell} + \\
&\qquad (-1)^k (f \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \wedge (dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_\ell}) \\
&= d\alpha \wedge \beta + (-1)^{|\alpha|} \alpha \wedge d\beta.
\end{aligned}$$

Step 4. $d^2 = 0$.
As before, it suffices to check this on simple k-forms. Let $\alpha \in \Omega^k(M)$ be given by

$$\alpha = f\,dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Then

$$\begin{aligned}
d^2(\alpha) &= d^2(f\,dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \\
&= d(df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \\
&= (d^2 f) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.
\end{aligned}$$

So it suffices to show that $d^2 f = 0$ for all smooth functions $f \in \mathscr{C}^\infty(M)$. We have htat

$$d^2 f = \frac{\partial^2 f}{\partial x_i \partial x_j} dx^i \wedge dx^j = \frac{\partial^2 f}{\partial x_j \partial x_i} dx^i \wedge dx^j = -\frac{\partial^2 f}{\partial x_j \partial x_i} dx^j \wedge dx^i.$$

After reindexing, the result follows immediately since $d^2 f = -d^2 f$, and so $d^2 f = 0$.

□

Proposition 2.64. *Let $f : M \to N$ be a smooth map. Then $df^* = f^* d$.*

Proof. Consider any k-form ω on N. Write

$$\omega = \sum_I \omega_I dx^I,$$

where the sum ranges over all multindices $1 \le i_1 < \cdots < i_k \le \dim N$. Then we see that

$$d_M(f^*\omega) = \sum_I \left(d_M(f^*\omega_I) \wedge f^*(dx^I) + f^*\omega_I \wedge d_M(f^*(dx^I)) \right).$$

Now the usual chain rule gives us that $d_M(f^*\omega_I) = f^*(d_N \omega_I)$. In local coordinates $(x_1, \ldots, x_n)$ we see that f looks like a collection of n functions $f_i \in \mathscr{C}^\infty(N)$, and so we get

$$f^*(dx^I) = df^I = d_M f^{i_1} \wedge \cdots \wedge d_M f^{i_k}.$$

In particular, we see that $d_M(df^I) = 0$. Putting all of the above results together we deduce that

$$d_M(f^*\omega) = f^*(d_N \omega_I) \wedge df^I = f^*(d_M \omega_I) \wedge f^*(dx^I) = f^*(d_M \omega).$$

□

Example 2.65. Consider $M = \mathbf{R}^3$, and let (x, y, z) denote the standard bases of $\mathbf{R}^3$. Then for a smooth function $f : \mathbf{R}^3 \to \mathbf{R}$ we have that

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz,$$

which is exactly the differential in local coordinates. Furthermore, the coefficients of df are exactly the coefficients of the gradient.

Consider a one form $\omega = a dx + b dy + c dz$. We then can compute

$$\begin{aligned} d\omega &= \left(\frac{\partial a}{\partial x}dx + \frac{\partial a}{\partial y}dy + \frac{\partial a}{\partial z}dz\right) \wedge dx + \left(\frac{\partial b}{\partial x}dx + \frac{\partial b}{\partial y}dy + \frac{\partial b}{\partial z}dz\right) \wedge dy + \left(\frac{\partial c}{\partial x}dx + \frac{\partial c}{\partial y}dy + \frac{\partial c}{\partial z}dz\right) \wedge dz \\ &= \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dy \wedge dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right) dz \wedge dx + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right) dx \wedge dy, \end{aligned}$$

which are exactly the coefficients of the curl of a vector field (a, b, c).

Finally, for a 2-form

$$\omega = a dy \wedge dz + b dz \wedge dx + c dx \wedge dy,$$

we have that

$$d\omega = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right) dx \wedge dy \wedge dz.$$

We see that the coefficient of $d\omega$ is the gradient of the vector field (a, b, c).

Note that in this case the property of $d^2 = 0$ allows us to immediately recover the following facts from vector calculus:

$$\operatorname{curl}(\nabla f) = 0 \quad \text{and} \quad \operatorname{div}(\operatorname{curl} X) = 0.$$

$\diamond$

Definition 2.66. Let $\alpha \in \Omega^p(M)$. If $\alpha = d\beta$ for some $\beta \in \Omega^{p-1}(M)$ then we say that α is **exact**. We say that α is **closed** if $d\alpha = 0$. $\diamond$

Note that every exact form is closed, but not every closed form is exact. This is the basis of de Rham cohomology.

The exterior derivative d is a derivation on the space of differential forms. Another derivation is the *interior derivation*.

Definition 2.67. Let X be a vector field on M. The **interior product** is a mapping from $(0, s)$-type tensor fields to $(0, s-1)$-type tensor fields. It is given by

$$(i_X S)|_p(v_1, \dots, v_{s-1}) = S|_p(X(p), v_1, \dots, v_{s-1}).$$

$\diamond$

The interior product of differential forms is just that which is induced from the interior product of covariant tensor fields.

Proposition 2.68. *Let $X, Y \in \mathfrak{X}(M)$, $f \in \mathscr{C}^\infty(M)$, $\alpha \in \Omega^k(M)$, and $\beta \in \Omega^\ell(M)$.*

(i) $i_X(df) = df(X) = X[f]$,
(ii) $i_X \circ i_Y = -i_Y \circ i_X$ *on* $\Omega(M)$, *and so* $i_X \circ i_X = 0$,
(iii) $i_X(\alpha \wedge \beta) = (i_X \alpha) \wedge \beta + (-1)^k \alpha \wedge (i_X \beta)$.

Proof.

(i) Note that $df \in \Omega^1(M)$ and so $i_X(df) \in \Omega^0(M) = \mathscr{C}^\infty(M)$. In particular, we have that

$$i_X(df)(p) = df_p(X(p)),$$

and so $i_X(df) = df(X)$.

(ii) Let $\omega \in \Omega^k(M)$. Then we have

$$i_Y(\omega)|_p(v_1,\dots,v_{k-1}) = \omega|_p(Y(p),v_1,\dots,v_{k-1}),$$

and consequently

$$\begin{aligned}(i_X \circ i_Y)(\omega)|_p(v_1,\dots,v_{k-2}) &= \omega|_p(Y(p),X(p),v_1,\dots,v_{k-2})\\ &= -\omega|_p(X(p),Y(p),v_1,\dots,v_{k-2})\\ &= -(i_Y \circ i_X)(\omega)|_p(v_1,\dots,v_{k-2}).\end{aligned}$$

(iii) By the bilinearity of the interior product and the wedge product, it suffices to prove the result for decomposable α and β. In particular, if $\omega_i \in \Omega^1(M)$ for $i = 1,\dots,k$ we see that

$$i_X(\omega_1 \wedge \cdots \wedge \omega_k) = \sum_{i=1}^{k} (-1)^{k-1}(i_X\omega_i)\omega_1 \wedge \cdots \wedge \widehat{\omega_i} \wedge \cdots \wedge \omega_k,$$

where the $\widehat{\omega_i}$ indicates that i^{th} term is removed. Note that for vector fields $X_2,\dots,X_k \in \mathfrak{X}(M)$, and letting $X_1 = X$, we have that

$$(\omega_1 \wedge \cdots \wedge \omega_k)(X_1,X_2,\dots,X_k) = \sum_{i=1}^{k} (-1)^{i-1}(i_X\omega_i)(\omega_1 \wedge \cdots \wedge \widehat{\omega_i} \wedge \cdots \wedge \omega_k)(X_2,\dots,X_k).$$

Note that the left hand side of the above is simply the determinant of the matrix $W := (\omega_i(X_j))_{i,j=1}^k$, and the right hand side is the Laplace expansion of the determinant of W with respect to the first column of W. The result now immediately follows.

□

2.5. **The Lie derivative.** In Definition 1.93 we introduced the Lie derivative of vector fields. Intuitively, the Lie derivative of Y with respect to X, $\mathscr{L}_X Y$, measures how much Y changes with respect to an observer who is flowing along X. We would like to generalize this idea to take Lie derivatives of arbitrary tensor fields with respect to a vector field X. Recall that the pullback of a $(0,s)$-tensor field $S \in T^s_0 N$ via a map $f : M \to N$ is given by $f^*S = S \circ (df)^{\otimes s}$. The pushforward of a $(r,0)$-tensor field is a bit mor subtle.

Definition 2.69. Let $f : M \to N$ be a smooth map, and T be a $(r,0)$-type tensor field on M. The **pushforward** of T via f is the section

$$f_*T := (df)^{\otimes r} \circ T \in \Gamma(f^*TN^{\otimes r}).$$

⋄

Note that if f is a diffeomorphism, then we can identify $\Gamma(f^*TN^{\otimes r})$ with $\Gamma(TN^{\otimes r})$; moreover, in this case we can even pushforward $(0,s)$-type tensor field by pulling back via the inverse. So if f is a diffeomorphism, we can pushforward and pullback arbitrary (r,s)-type tensor fields.

Definition 2.70. Let T be a tensor field on M, and X a vector field. The **Lie derivative** of T with respect to X is defined by

$$\mathscr{L}_X T := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_{-t})_* T = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_t)^* T,$$

where Φ is the local flow of X. ⋄

As an example, we consider the Lie derivative of a $(0,1)$-type tensor field α, i.e. a 1-form. For any smooth vector field Y, we have that

$$\begin{aligned}(\mathscr{L}_X\alpha)(Y)(p) &= \lim_{t\to 0} \frac{\Phi_t^*\alpha_{\Phi_t(p)} - \alpha_p}{t}(Y_p)\\ &= \lim_{t\to 0} \frac{\alpha_{\Phi_t(p)}((d\Phi_t)_p(Y_p)) - \alpha_p(Y_p)}{t}\\ &= \lim_{t\to 0} \frac{\alpha_{\Phi_t(p)}((d\Phi_t)_p(Y_p)) - \alpha_{\Phi_t(p)}(Y_{\Phi_t(p)})}{t} + \lim_{t\to 0}\frac{\alpha_{\Phi_t(p)}(Y_p) - \alpha_p(Y_p)}{t}.\end{aligned}$$

Since

$$\begin{aligned}\lim_{t\to 0}\alpha_{\Phi_t(p)}\left(\frac{(d\Phi_t)_p(Y_p)-Y_{\Phi_t(p)}}{t}\right) &= \lim_{t\to 0}\alpha_{\Phi_t(p)}\left((d\Phi_t)_p\left(\frac{(Y_p)-(d\Phi_{-t})_{\Phi_t(p)}Y_{\Phi_t(p)}}{t}\right)\right)\\ &= \lim_{t\to 0}(\Phi_t^*\alpha_{\Phi_t(p)}\left(\frac{(Y_p)-(d\Phi_{-t})_{\Phi_t(p)}Y_{\Phi_t(p)}}{t}\right)\\ &= -\alpha(\mathscr{L}_X Y)(p),\end{aligned}$$

and

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\alpha_{\Phi_t(p)}(Y_{\Phi_t(p)}) = X_p(\alpha(Y)),$$

we deduce that

$$(\mathscr{L}_X\alpha)(Y) = X(\alpha(Y)) - \alpha(\mathscr{L}_X Y).$$

Now we see that $\mathscr{L}_X\alpha$ is well-defined and smooth in p.

More generally, if T is a (r,s)-type tensor field on M, locally it is a linear combination of monomials of the form $X_1\otimes\cdots\otimes X_r\otimes\alpha^1\otimes\cdots\otimes\alpha^s$. By the definition of the Lie derivative, it is straightforward to see that

$$\begin{aligned}\mathscr{L}_X(X_1\otimes\cdots\otimes X_r\otimes\alpha^1\otimes\cdots\otimes\alpha^s) &= \sum_{i=1}^r X_1\otimes\cdots\otimes\mathscr{L}_X X_i\otimes\cdots\otimes X_r\otimes\alpha^1\otimes\cdots\otimes\alpha^s\\ &\quad+\sum_{j=1}^s X_1\otimes\cdots\otimes X_r\otimes\alpha^1\otimes\cdots\mathscr{L}_X\alpha^j\otimes\cdots\alpha^s.\end{aligned}$$

Again, by the previous discussion for the case of 1-forms, we deduce that the Lie derivatives of arbitrary tensor fields is well defined and smooth in p. Summarizing the above results, we have the following theorem.

Theorem 2.71. *Let S, T be two tensor fields on M. Then*

- $\mathscr{L}_X(S\otimes T) = \mathscr{L}_X S\otimes T + S\otimes\mathscr{L}_X T$.
- *If T is a (r,s)-type tensor field, then for any 1-forms $\alpha^1,\dots\alpha^r$ and smooth vector fields $Y_1,\dots,Y_s$ we have*

$$\begin{aligned}\mathscr{L}_X T(\alpha^1,\dots,\alpha^r,Y_1,\dots,Y_s) &= X(T(\alpha^1,\dots,\alpha^r,Y_1,\dots,Y_s)) - \sum_{i=1}^r T(\alpha^1,\dots,\mathscr{L}_X\alpha^i,\dots,\alpha^r,Y_1,\dots,Y_s)\\ &\quad+\sum_{j=1}^s T(\alpha^1,\dots,\alpha^r,Y_1,\dots,\mathscr{L}_X Y_j,\dots,Y_s).\end{aligned}$$

Some of the most important applications of the Lie derivative are when it acts on differential forms. Since every p-form ω can be regarded as an antisymmetric $(0,p)$-type tensor field, we see that $\mathscr{L}_X\omega$ makes sense, and so by Theorem 2.71 we deduce that

$$\mathscr{L}_X(\omega_1\wedge\omega_1) = \mathscr{L}_X\omega_1\wedge\omega_2 + \omega_1\wedge\mathscr{L}_X\omega_2.$$

Recall that by Proposition 2.61 we have that $\Phi_t^*(d\omega) = d\Phi_t^*\omega$. Hence,

$$\mathscr{L}_X(d\omega) = \lim_{t\to 0}\frac{\Phi_t^*(d\omega)-d\omega}{t} = \lim_{t\to 0}\frac{d(\Phi_t^*\omega)-d\omega}{t} = d\mathscr{L}_X\omega.$$

The following theorem is now just a straightforward exercise. It follows since these rules are enough to fully compute the Lie derivative locally for any differential form.

Theorem 2.72. *The Lie derivative on differential forms is the unique operator that satisfies*

- $\mathscr{L}_X f = X[f]$,
- $\mathscr{L}_X(\alpha\wedge\beta) = \mathscr{L}_X\alpha\wedge\beta + \alpha\wedge\mathscr{L}_X\beta$,
- $\mathscr{L}_X\circ d = d\circ\mathscr{L}_X$.

Using this characterization we immediately deduce the following theorem.

Theorem 2.73 (Cartan's magic formula)**.** $\mathscr{L}_X = d\circ i_X + i_X\circ d$.

2.6. **Integration of differential forms.** Let M be a smooth n-dimensional manifold and let $\alpha \in \Omega^n(M)$ be a compactly supported differential form. In local coordinates, we can write

$$\alpha = f\ dx^1 \wedge \cdots \wedge dx^n,$$

and it seems natural to want to integrate α. By changing local coordinates via a diffeomorphism ϕ we see that the representation of α becomes

$$\alpha = (f \circ \phi)(\det(d\phi))\ dx^1 \wedge \cdots \wedge dx^n.$$

This is almost the same as the change of variables formula for integration, but we would need to ensure that $\det(d\phi) > 0$. This is the motivation for requiring integration to be performed *only* on orientable manifolds.

Definition 2.74. A **volume form** is a nowhere vanishing n-form. ⋄

Theorem 2.75. *M is orientable if and only if M admits a volume form.*

Proof. First assume that M is orientable. Then we have coordinate charts such that $\det(\partial y_i/\partial x_j) > 0$ on the intersections. Consider a smooth partition of unity $\{\rho_\alpha\}$ subordinated to this orientation, and let

$$\nu = \sum \rho_\alpha\ dy_1^\alpha \wedge dy_2^\alpha \wedge \cdots \wedge dy_n^\alpha,$$

where $\{y_j^\alpha\}$ denotes the local coordinates on the domain of ρ_α. Now if we consider a coordinate chart U_β with coordinates $x_1, \ldots, x_n$, we have

$$\nu|_{U_\beta} = \sum \rho_\alpha \det\left(\frac{\partial y_i^\alpha}{\partial x_j}\right) dx_1 \wedge \cdots \wedge dx_n.$$

Since $\rho_\alpha \geq 0$ and $\det(dy_i^\alpha/dx_j) > 0$ we see that ν vanishes nowhere. Hence ν is a volume form.

Conversely, suppose that M admits a volume form ν. In coordinates, we can write

$$\nu = f(x_1, \ldots, x_n)\ dx_1 \wedge \cdots \wedge dx_n.$$

Without loss of generality, we can assume that $f > 0$, since we can just make a coordinate change $x_1 \mapsto c - x_1$. Now let $\{y_j\}$ be an overlapping set of local coordinates. Then we have

$$\begin{aligned} \nu &= g(y_1, \ldots, y_n)\ dy_1 \wedge \cdots \wedge dy_n \\ &= g(y_1(x), \ldots, y_n(x)) \det\left(\frac{\partial y_i}{\partial x_j}\right) dx_1 \wedge \cdots \wedge dx_n \\ &= f(x_1, \ldots, x_n)\ dx_1 \wedge \cdots \wedge dx_n. \end{aligned}$$

Since $f > 0$ and $g > 0$ we deduce that $\det(\partial y_i/\partial x_j) > 0$, and hence M is orientable. □

Definition 2.76. Let $\{\rho_\alpha\}$ be a smooth partition of unity suboordinated to a locally finite open cover $\{V_\alpha\}$ by images of an orientation, i.e. $V_\alpha = \varphi_\alpha(U_\alpha)$. Let $\omega \in \Omega^n(M)$ have compact support. Locally, write $f_\alpha\ dx_1 \wedge \cdots \wedge dx_n = \varphi_\alpha^*(\rho_\alpha \omega)$. Then we define

$$\int_M \omega = \sum_\alpha \int_{U_\alpha} f_\alpha(x_1, \ldots, x_n)\ dx_1 \cdots dx_n.$$

Note that $\int_M : \Omega^n(M) \to \mathbf{R}$ is a linear map. ⋄

Intuitively, the above definition says we integrate top level forms simply by partitioning the manifold into charts and integrating in these charts using the usual Lebesgue integral. We still need to show that the above definition is well defined, in the sense that it does not depend on the choice of locally finite cover and on the choice of partition of unity. This follows immediately from the change of variables for the Lebesgue integral and the fact that the determinant of the transition functions is positive.

If M is not compact, we can still integrate n-forms which are not compactly supported in certain cases. Consider the decomposition $\omega = \omega^+ - \omega^-$, and then if at least one of $\int_M \omega^+$ or $\int_M \omega^-$ is finite, we define

$$\int_M \omega := \int_M \omega^+ - \int_M \omega^-.$$

If both of them are infinite, we say that ω is not integrable. Note that ω^+ and ω^- are not necessarily smooth, but we can still integrate differential forms with less regularity.

Note that we can integrate lower level forms on lower dimensional submanifolds. For example, consider a smooth curve in M. Let $\gamma : [a,b] \to M$ be a smooth parametric curve and $\alpha \in \Omega^1(M)$. Then we define

$$\int_\gamma \alpha = \int_a^b \gamma^* \alpha = \int_a^b \alpha(\dot{\gamma}(t))\, dt.$$

If $\alpha = df$ is an exact 1-form, then

$$\int_\gamma df = \int_a^b \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\, dt = f(\gamma(b)) - f(\gamma(a)).$$

So we see that the integral of an exact 1-form only depends on it's endpoints. In particular, if γ is a loop then $\gamma(a) = \gamma(b)$ and so $\int_\gamma df = 0$.

Definition 2.77. M is **contractible** if a smooth map $H : M \times \mathbf{R} \to M$ exists such that $H(p,1) = p$ and $H(p,0) = p_0$ for a fixed p_0 and all $p \in M$. ⋄

Clearly, $\mathbf{R}^n$ is contractible with map $H(p,t) = tp$.

Lemma 2.78 (Poincaré's Lemma.). *If M is contractible and α is a smooth closed k-form, $k \geq 1$, then there exists a smooth $(k-1)$-form η such that $\alpha = d\eta$.*

Proof. Let $N = M \times \mathbf{R}$ and consider the vector field $Z = \frac{\partial}{\partial t}$, where t is the variable associated to $\mathbf{R}$ component of N. The flow associated to Z is $\Phi_t(p,s) = (p, s+t)$ on N. Now define the map $j_t : M \to N$ by $j_t(p) = (p,t)$. Let $\mathfrak{D}$ be the map from smooth k-forms on N to smooth $(k-1)$-forms on M given by

$$(\mathfrak{D}\xi)_p = \int_0^1 (j_t^*(i_Z \xi))_p\, dt.$$

Now let ξ be a smooth k-form on N, and define $\zeta = i_Z \xi$ and $\xi_1 = \xi - dt \wedge \zeta$. Then

- $i_Z \zeta = 0$ since $i_Z \circ i_Z = 0$.
- $i_Z \xi_1 = 0$ since $i_Z(dt \wedge \zeta) = dt(Z)\xi = \xi = i_Z \xi$.
- $j_1^* \xi - j_0^* \xi = d\mathfrak{D}\xi + \mathfrak{D}d\xi$. This follows since $\Phi_s \circ j_t = j_{t+s}$, and so

$$\begin{aligned}
d\mathfrak{D}\xi + \mathfrak{D}d\xi &= d\int_0^1 j_t^*(i_Z\xi)\, dt + \int_0^1 j_t^*(i_Z d\xi)\, dt \\
&= \int_0^1 j_t^*(d(i_Z\xi) + i_Z d\xi)\, dt \\
&= \int_0^1 j_t^* \mathscr{L}_Z \xi\, dt \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} (\Phi_s \circ j_t)^* \xi\, dt \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} j_t^* \xi\, dt = j_1^*\xi - j_0^*\xi.
\end{aligned}$$

Now let $H : M \times \mathbf{R} \to M$ be a contraction map and define $\xi = H^*\alpha$. Note that $d\xi = dH^*\alpha = H^* d\alpha = 0$ and $j_1^*\xi = (H \circ j_1)^*\alpha = \mathrm{id}^*\alpha = \alpha$. Similarly, we have $j_0^*\xi = 0$. Then $\alpha = j_1^*\xi - j_0^*\xi = d\mathfrak{D}\xi$. Hence, the result holds with $\eta = \mathfrak{D}\xi$. □

The treatment of integration we have taken above is nice from a theoretical standpoint, but is unwieldy to use for actual applications – no one wants to explicitly compute partitions of unity! Now we refine the theory a bit more to make integration on manifolds essentially the same as integration over Euclidean space.

The first point of business is to reconsider the space of partitions of unity that we admit in our definitions of integration. The following theorem be *very* important when we consider integration on manifolds with boundary. Note that in the definition of the integral, we require the partitions of unity to have support contained in coordinate charts *and* to have compact support. We want to relax these two assumptions to consider any family $\{\rho_\alpha\}$ such that each ρ_α

(i) is non-negative and smooth,
(ii) has locally finite support,
(iii) $\sum_\alpha \rho_\alpha = 1$.

This is the most general class of partitions of unity that we should ever expect to use.

Theorem 2.79. *Let $\{\rho_\alpha\}$ be a collection of non-negative smooth functions on M whose supports form a locally finite collection of closed sets in M satisfying $\sum_\alpha \rho_\alpha = 1$. For any n-form $\omega \in \Omega^n(M)$, ω is absolutely integrable if and only if all of the $\rho_\alpha\omega$ are absolutely integrable and $\sum_\alpha \int_M |\rho_\alpha\omega|$ is finite – in this case, the sum is equal to $\int_M |\omega|$. Furthermore, if ω is absolutely integrable then $\sum_\alpha \int_M \rho_\alpha\omega$ is absolutely convergent and equal to $\int_M \omega$.*

Proof. Let $\{\eta_\beta\}$ be a smooth partition of unity with locally finite and compact supports contained in coordinate charts. We first assume that ω is absolutely integrable, and we want to show that all of the $\rho_\alpha\omega$ are absolutely integrable and the sum of their integrals is equal to $\int_M |\omega|$. Since ω is absolutely integrable, by definition we have

$$\int_M |\omega| = \sum_\beta \int_M |\eta_\beta\omega|.$$

Since each η_β is compactly supported in a coordinate chart we see that $\eta_\beta\rho_\alpha\omega = 0$ for all but finitely many α depending on β, and that

$$\int_M |\eta_\beta\omega| = \sum_\alpha \int_M |\eta_\beta\rho_\alpha\omega|$$

since the integral is linear, and hence finitely additive. Putting this together we see that the following sum is convergent and

$$\int_M |\omega| = \sum_\beta\sum_\alpha \int_M |\eta_\beta\rho_\alpha\omega| = \sum_\alpha\sum_\beta \int_M |\eta_\beta\rho_\alpha\omega|,$$

where the rearrangement of the series is justified since all of the terms are non-negative. In particular, for all α we have that $\sum_\beta \int_M |\eta_\beta\rho_\alpha\omega|$ is convergent, and by definition of the integral we deduce that

$$\int_M |\rho_\alpha\omega| = \sum_\beta \int_M |\eta_\beta\rho_\alpha\omega|.$$

So we conclude that each $\rho_\alpha\omega$ is absolutely integrable and that $\sum_\alpha \int_M |\rho_\alpha\omega| = \int_M |\omega|$.

Conversely, suppose that each $\rho_\alpha\omega$ is absolutely integrable with $\sum_\alpha \int_M |\rho_\alpha\omega|$ finite. By definition, for any α,

$$\int_M |\rho_\alpha\omega| = \sum_\beta \int_M |\eta_\beta\rho_\alpha\omega|,$$

hence the sum $\sum_\alpha\sum_\beta \int_M |\eta_\beta\rho_\alpha\omega|$ is also convergent. Again, since all of the terms are non-negative we can rearrange the series to find that $\sum_\alpha\sum_\beta \int_M |\eta_\beta\rho_\alpha\omega| = \sum_\beta\sum_\alpha \int_M |\eta_\beta\rho_\alpha\omega|$, and in particular the right hand side is convergent. Since each $\eta_\beta\omega$ is compactly supported in a coordinate chart, we can write

$$\eta_\beta\omega = f\ dx_1 \wedge \cdots \wedge dx_n,$$

and so

$$\int_M |\eta_\beta\omega| = \int_{U_\beta} |f|\ dx_1 dx_2 \cdots dx_n.$$

Since $|f|$ is compactly supported we see that $\rho_\alpha|f| = 0$ for all but finitely many α, and so $|f| = \sum_\alpha |\rho_\alpha f|$ as a *finite* sum. By the finite additivity of the Lebesgue integral we deduce that

$$\int_M |\eta_\beta \omega| = \int_{U_\beta} |f|\, dx_1 dx_2 \cdots dx_n = \sum_\alpha \int_{U_\beta} |\rho_\alpha f|\, dx_1 dx_2 \cdots dx_n = \sum_\alpha \int_M |\eta_\beta \rho_\alpha \omega|.$$

This yields the result since

$$\int_M |\omega| = \sum_\beta \int_M |\eta_\beta \omega| = \sum_\beta \sum_\alpha \int_M |\eta_\beta \rho_\alpha \omega| = \sum_\alpha \sum_\beta \int_M |\eta_\beta \rho_\alpha \omega| = \sum_\alpha \int_M |\rho_\alpha \omega|.$$

Finally, since each $\rho_\alpha \omega$ is absolutely integrable we deduce, by the above, that the series $\sum_\alpha \int_M \rho_\alpha \omega$ is absolutely convergent. By definition, $\int_M \rho_\alpha \omega = \sum_\beta \int_M \eta_\beta \rho_\alpha \omega$ for all α, and this sum is also absolutely convergent. Hence,

$$\sum_\alpha \int_M \rho_\alpha \omega = \sum_\alpha \sum_\beta \int_M \eta_\beta \rho_\alpha \omega,$$

and this double sum is absolutely convergent since $\sum_\alpha \sum_\beta \int_M |\eta_\beta \rho_\alpha \omega|$ is finite (by the assumption that ω is absolutely integrable). So we can rearrange this series to obtain $\sum_\alpha \sum_\beta \int_M \eta_\beta \rho_\alpha \omega = \sum_\beta \sum_\alpha \int_M \eta_\beta \rho_\alpha \omega$. Note that $\int_M \omega = \sum_\beta \eta_\beta \omega$, and so we just need to show that $\sum_\alpha \int_M \eta_\beta \rho_\alpha \omega = \int_M \eta_\beta \omega$ for all β. But this again reduces to the Euclidean case since η_β is compactly supported in a coordinate chart. $\square$

The second point of order is to come up with a way for us to actually compute integrals. Can we compute an integral over the sphere as the sum of integrals over complementary hemispheres?

Lemma 2.80. *Let $f : M \to N$ be a diffeomorphism between two smooth oriented n-manifolds. Let $\omega \in \Omega^n(M)$, and write $\widetilde{\omega} = f^*\omega$. Then ω is absolutely integrable over M if and only if $\widetilde{\omega}$ is absolutely integrable over N, in which case $\int_M |\omega| = \int_N |\widetilde{\omega}|$. Moreover, if f is orientation preserving then $\int_M \omega = \int_N \widetilde{\omega}$.*

Proof. Let $\{(U_\alpha, \varphi_\alpha)\}$ be a smooth atlas for M consisting of a locally finite collection of open subsets U_α, and write $V_\alpha = f^{-1}(U_\alpha)$ and $\psi_\alpha = \varphi_\alpha \circ f|_{V_\alpha}$. Let $\{\rho_\beta\}$ be a smooth partition of unity on M with compact supports subordinated to $\{U_\alpha\}$. Since the definitions of integration on M and N are defined in terms of integrals on subsets of Euclidean space, the result is reduced to the case of compactly supported smooth functions on open subsets of $\mathbf{R}^n$. The fact that f is orientation preserving implies that the determinant of the Jacobian is everywhere positive, hence is equal to its absolute value. So the usual change of variables theorem gives the desired result. $\square$

This lemma allows us to shift the integration problem from one manifold to another diffeomorphic one.

Theorem 2.81. *Let $M_1, \dots, M_r$ be finitely many smooth manifolds of the same dimension, n. Let $f_i : M_i \to M$ be smooth injective immersions that are homeomorphisms onto their images. Assume that the $f(M_i)$ are pairwise disjoint, and that closure of the complement of the union of the $f(M_i)$ has measure zero in M. For any $\omega \in \Omega^n(M)$, ω is absolutely integrable if and only if each $\omega_i = f_i^*\omega$ is absolutely integrable on M_i, in which case $\int_M |\omega| = \sum_i \int_{M_i} |\omega_i|$. Furthermore, assume that M is oriented and endow each M_i with the pullback orientation given by the bundle isomorphism $TM_i \cong f_i^*(TM)$, then $\int_M \omega = \sum_i \int_{M_i} f_i^*\omega$.*

Proof. By Lemma 2.1 we see that $\omega|_{f_i(M_i)}$ is absolutely integrable over the open submanifold $f_i(M_i)$ in M if and only if $f_i^*\omega$ is absolutely integrable over M_i, in which case $\int_{M_i} |f_i^*\omega| = \int_{f_i(M_i)} |\omega|$, and moreover in the case that M is oriented we have that $\int_{M_i} f_i^*\omega = \int_{f_i(M_i)} \omega$.

In light of this, we can replace M_i with $f_i(M_i)$ and so the theorem reduces to showing the following: if M is a smooth manifold and the M_i are pairwise disjoint open submanifolds then ω is absolutely integrable if and only if it is absolutely integrable over each open subset M_i, in which case $\int_M |\omega| = \sum_i \int_{M_i} |\omega|$. To see this let $\{\rho_\alpha\}$ be a smooth partition of unity with compact supports subordinated to an orientation on M, so that $\{\rho_\alpha|_{M_i}\}$ satisfy the same property on M_i except that the supports on M_i are no longer necessary compact. In light of Theorem 2.79 we can use these non-compactly supported partitions of unity to integrate over M_i. Hence, we only need to consider the above problem for each $\rho_\alpha \omega$. In particular, we can assume that ω has compact support contained in an open coordinate chart of M, which we denote as U. Write $U_i := M_i \cap U$. This implies that the differential geometric integral is simply the usual Lebesgue integral. In particular, write $\omega = f\ dx_1 \wedge \cdots \wedge dx_n$, and we see that since $U \setminus \bigcup U_i$ is closed with Lebesgue measure zero in U, we see that

$$\int_M \omega = \int_U f\ dx_1 dx_2 \cdots dx_n = \sum_i \int_{U_i} f\ dx_1 dx_2 \cdots dx_n = \sum_i \int_{M_i} \omega.$$

□

Example 2.82. Now we can compute the integral over a sphere as a sum of integrals over a pair of complementary hemispheres, or over the open locus on which spherical coordinates are defined. Similarly, for any manifold endowed with a coordinate chart complementary to a closed set of measure zero, all integration problems on the manifold can be shifted to the coordinate chart. We use this all the time when computing integrals in the plane using polar coordinates, or integrals on the torus using angle parameterization.

Also, on Grassmannian manifolds any of the standard open subsets $U_{\mathbf{I}}$ has complement that is closed with measure zero, and so any integration on a Grassmanian manifold can be computed by working in a single $U_{\mathbf{I}}$. ⋄

Finally, we want to understand how the classical notation of vector calculus can be rigorously understood. Let S be an embedded oriented smooth surface with boundary in $\mathbf{R}^3$, and γ be an embedded oriented curve with boundary in $\mathbf{R}^n$. We often see expressions like

$$\int_S P\ dxdy + Q\ dxdz + R\ dzdy, \qquad P, Q, R \in \mathscr{C}^\infty(U),\ S \subseteq U \subseteq \mathbf{R}^3$$

$$\int_C \sum_j f_j\ dx_j, \qquad f_1, \ldots, f_n \in \mathscr{C}^\infty(\widetilde{U}),\ C \subseteq \widetilde{U} \subseteq \mathbf{R}^n.$$

in classical vector calculus. The question we want to answer is what does this notation actually mean?

From the modern point of view, if we let $i : S \hookrightarrow U \subseteq \mathbf{R}^3$ and $\sigma : C \hookrightarrow \widetilde{U} \subseteq \mathbf{R}^n$ denote the embedding maps then the integrals should be understood to mean

$$\int_S i^*(P\ dx \wedge dy + Q\ dx \wedge dz + R\ dz \wedge dy) \text{ and } \int_C \sigma^*\Big(\sum f_j\ dx_j\Big).$$

2.7. **Stokes' Theorem.** Before proving the celebrated Stokes' theorem we need to define manifolds with boundaries. Define the half-space

$$\mathbf{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbf{R}^n\ :\ x_1 \geq 0\}.$$

The half-space is the model manifold with boundary, just as $\mathbf{R}^n$ is the model manifold. Manifolds with boundary are defined in terms of local coordinates whose domains are now open subsets of the half-space instead of $\mathbf{R}^n$. This introduces two distinct types of points on M: the interior points, which ine in the image of the parameterization whose domain is contained in $\mathbf{R}^n_+$ minus the boundary, i.e. $x_1 > 0$; and boundary points, which are the image under the coordinate maps of points on the boundary of $\mathbf{R}^n_+$. This dichotomy between interior points and boundary points of a manifold with boundary rely on the *invariance of domain theorem*.

Theorem 2.83 (Invariance of domain)**.** *If $U \subseteq \mathbf{R}^n$ is open and $f : U \to \mathbf{R}^n$ is an injective and continuous map, then $V := f(U)$ is open and f is a homeomorphism between U and V.*

We postpone the proof for now since we will need tools from algebraic topology.

A function f defined on a subset of $\mathbf{R}^n_+$ is said to be differentiable at a point p on the boundary of $\mathbf{R}^n_+$ if f can be extended to a differentiable function past the boundary. Similarly, we can talk about smooth functions in regions of $\mathbf{R}^n_+$. With this all in mind, we can define a smooth manifold with boudnary in the same way we did for regular manifolds, but where the open sets of M are homeomorphic to open sets in $\mathbf{R}^n_+$.

It is not hard to show that the boundary ∂M of M is itself a smooth manifold of dimension $n-1$. If M is orientable with positive orientation determined by some nowhere vanishing n-form ω, then the boundary ∂M is also orientable. We define the **positive boundary orientation** to be the one given by the $(n-1)$-form $\sigma = \iota_X \nu$, where X is any non-vanishing continuous vector field on ∂M such that for every $p \in \partial M$ $X(p) \in T_pM \setminus T_p(\partial M)$ pointing out of M. The notion of pointing out of M means that in any coordinate chart the image of $X(p)$ in $\mathbf{R}^n$ points into the negative x_1 half-space. This means that if $dx_1 \wedge \cdots \wedge dx_n$ defines a positive orientation on $\mathbf{R}^n_+$ then $-dx_2 \wedge \cdots \wedge dx_n$ defines the induced positive orientation on the boundary of $\mathbf{R}^n_+$. We finally are able to prove Stokes' Theorem.

Theorem 2.84 (Stokes' Theorem)**.** *Let M be an n-dimensional smooth, oriented, manifold with boundary ∂M. Let μ be a smooth $(n-1)$-form on M with compact support. By an abuse of notation we also write μ for $i^*\mu$ where $i : \partial M \to M$ is the inclusion map. Then*

$$\int_{\partial M} \mu = \int_M d\mu.$$

Proof. As usual, let $\{\rho_\alpha\}$ be a smooth partition of unity of M subordinated to the locally finite open cover $\{V_\alpha\}$. On each coordinate chart with coordinates $(x_1, \dots, x_n)$ we write

$$\rho_\alpha \mu = \sum_{j=1}^n f_j dx_1 \wedge \cdots \widehat{dx_j} \wedge \cdots \wedge dx_n,$$

for smooth functions $f_j \in \mathscr{C}^\infty_c(\mathbf{R}^n_+)$. The hat over dx_j indicates that this factor is omitted from the wedge product. by the linearity of the integral we may assume that $\mu = \rho_\alpha \mu$ and ignore the summation over α.

There are two types of coordinate neighborhoods: those parameterized over open sets in $\mathbf{R}^n_+$ that do not intersect the boundary, and those parameterized by open sets that contain boundary points of $\mathbf{R}^n_+$. The argument below using the fundamental theorem of calculus shows that the integral $\rho_\alpha \mu$ over V_α of the first kind is zero. So we only consider coordinate neighborhoods of the second kind.

Identifying for simplicity of notation $d(\rho_\alpha \mu)$ with its pullback to $\mathbf{R}^n$ under the coordinate map we may write

$$d(\rho_\alpha \mu) = \sum_{j=1}^n \frac{\partial f_j}{\partial x_j} dx_j \wedge dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n = \sum_{j=1}^n (-1)^{j-1} \frac{\partial f_j}{\partial x_j} dx_1 \wedge \cdots \wedge dx_n.$$

Note that if $j \neq 1$ then

$$\int_{\mathbf{R}^n_+} \frac{\partial f_j}{\partial x_j} dx_1 \wedge \cdots \wedge dx_n = 0.$$

This follows since f_j is compactly supported, so

$$\int_{-\infty}^{\infty} \frac{\partial f_j}{\partial x_j} dx_j = 0$$

for any fixed set of values of $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$. For $j = 1$ on the other hand, we have

$$\begin{aligned} \int_{\mathbf{R}^n_+} \frac{\partial f_1}{\partial x_1} dx_1 \wedge \cdots \wedge dx_n &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\int_0^{\infty} \frac{\partial f_1}{\partial x_1}\, dx_1 \right) dx_2 \cdots dx_n \\ &= -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_1(0, x_2, \dots, x_n) dx_2 \cdots dx_n. \end{aligned}$$

Keeping in mind the orientation conventions for M and its boundary, we have that the last integral in the above sequence of equalities is simply the integral $\int_{\partial M} \rho_\alpha \mu$. ☐

2.8. **Riemannian metrics and volume forms.** For each point $p \in M$ we would like to assign a scalar product

$$g_p : T_pM \times T_pM \to \mathbf{R}$$

in a smooth way. That is, the map $p \mapsto g_p$ should be smooth. For this the language of the tensor bundle and tensor fields is extremely natural.

Definition 2.85. A **Riemannian metric** on M is a section of the tensor bundle $T^2_0 M = TM^* \otimes TM^*$

$$\begin{aligned} g : M &\to TM^* \otimes TM^*, \\ p &\mapsto g_p \in (T_p M)^* \otimes (T_p M)^*, \end{aligned}$$

such that the bilinear map $g_p : T_p M \times T_p M \to \mathbf{R}$ is symmetric and positive definite. ⋄

Theorem 2.86. *If M has a countable basis of open neighborhoods, then M admits a Riemannian metric.*

Proof. By assumption about countable basis, there exists a partition of unity. That is, given an open covering $M = \bigcup_i U_i$, there is a **subordinated locally finite** covering $M = \bigcup_k V_k$ with k belonging to a countable set, so that

- each V_k is contained in one of U_i's,
- each point $p \in M$ has a neighborhood $W \ni p$ such that $W \cap V_k \neq \emptyset$ only for a finite number of k.

There exists a countable family of smooth functions (η_k) with $\eta_k \geq 0$ such that $\operatorname{supp} \eta_k \subset V_k$, and

$$\sum_k \eta_k(p) = 1 \quad \text{for all } p \in M.$$

We may assume that V_k's are the charts of M together with coordinate functions $\phi_k : V_k \to \Omega^k \subset \mathbf{R}^n$. Given a point $p \in V_k$, we define a map $g_k : V_k \to TM^* \otimes TM^*$ by

$$g_k(u, v) = \langle d\phi_k(u), d\phi_k(v) \rangle \quad \text{for } (u, v) \in T_p M \times T_p M,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product on $\mathbf{R}^n$.

Now we put

$$g = \sum_k \eta_k \, g_k.$$

Here $\eta_k \, g_k$ is zero away from V_k. Given a point $p \in M$, there exists k such that $V_k \ni p$, and $\eta_k(p) > 0$. For any $u \in T_p M \setminus \{0\}$ one has

$$g_p(u, u) = \sum_k \eta_k(p) \, g_k(u, u) \geq \eta_k(p) \, g_k(u, u) > 0.$$

Hence g_p is positive definite. □

Remark 2.87. An alternative, very short proof uses the Whitney embedding theorem. Since every manifold can be embedded in Euclidean space for large enough n, we can just pullback the usual Euclidean metric under this embedding, and check that it is symmetric and positive definite. ⋄

For any tangent vectors $u, v \in T_p M$ with local coordinates on a chart (U, ϕ) given by

$$u = \sum_{1 \leq i \leq n} u^i \, \frac{\partial}{\partial x^i}(p), \quad v = \sum_{1 \leq j \leq n} v^j \, \frac{\partial}{\partial x^j}(p)$$

one has

$$g_p(u, v) = \sum u^i \, v^j \, \underbrace{g_p\left(\frac{\partial}{\partial x^i}(p), \frac{\partial}{\partial x^j}(p) \right)}_{= g_{ij}(p)}.$$

Here $g_{ij} : U \to \mathbf{R}$ is a smooth function. Since g is symmetric we have $g_{ij} = g_{ji}$. We write

$$g = \sum g_{ij} \, \underbrace{dx^i \otimes dx^j}_{\in TM^* \otimes TM^*}.$$

Sometimes the sign "⊗" is omitted and one writes "$g = \sum g_{ij} \, dx^i \, dx^j$".

In the above, we let $dx^1, \ldots dx^d$ be the basis which is dual to $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^d}$, that is

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_{ij}.$$

Example 2.88. One can check that the Euclidean metric on $\mathbf{R}^2$ in polar coordinates is given by

$$g = dr^2 + r^2\,d\theta^2,$$

where $dr^2 = dr \otimes dr$ and $d\theta^2 = d\theta \otimes d\theta$. The polar coordinates are related to the usual coordinates by $x = r\,\cos\theta$ and $y = r\,\sin\theta$. ⋄

Definition 2.89. The **Riemannian volume form** is the unique volume form $\mathrm{vol}_g \in \Omega^n(M)$ satisfying

$$\mathrm{vol}_g(e_1, \ldots, e_n) = 1,$$

whenever $(e_1, \ldots, e_n)$ is a positive orthonormal basis of T_pM. ⋄

By an abuse of notation, we write for $f \in \mathscr{C}^\infty(M)$

$$\int_M f := \int_M f \ \mathrm{vol}_g\,.$$

Proposition 2.90. *There is a unique Riemannian volume form.*

Proof. First we show that this form is unuiquely determined. Let ν be a Riemannian volume form. Let $\{e_1, \ldots, e_n\}$ be a positive orthonormal frame of TM in a neighborhood of p. Let $\{e_1^*, \ldots, e_n^*\}$ be the dual coframe. Locally, we can write

$$\nu = F\ e_1^* \wedge \cdots \wedge e_n^*.$$

Since ν is a Riemannian volume form, we deduce that $F \equiv 1$, and so the form is uniquely determined.

Now we prove the existence. Let $\{e_1, \ldots, e_n\}$ be a positive orthonormal frame of TM in a neighborhood of p, and define

$$\nu_p = e_1^* \wedge \cdots \wedge e_n^*.$$

This is well defined since if we have another positive orthonormal frame $\{f_1, \ldots, f_n\}$ then

$$\nu_p = e_1^* \wedge \cdots \wedge e_n^* = \det(\tau)\ f_1^* \wedge \cdots \wedge f_n^*,$$

where τ is the transition matrix between the bases. Since both bases are orthonormal we have that $\tau \in \mathbf{O}(n)$ and hence $\det(\tau) = \pm 1$. Since both bases are positive we determine that $\det(\tau) > 0$, and so in particular $\det(\tau) = 1$. Therefore ν_p is well defined. This is smooth in p, and can be seen by considering a smooth orthonormal basis. So in particular $\nu \in \Omega^n(M)$, and by construction ν is a Riemannian volume form. □

Proposition 2.91. *Let $(x_1, \ldots, x_n)$ be a positive local coordinate system, and suppose the metric is given by $g_{ij} := g(\partial_{x_i}, \partial_{x_j})$. Then the Riemannian volume form is given by*

$$\mathrm{vol}_g := \sqrt{\det(g_{ij})}\ dx_1 \wedge \cdots \wedge dx_n.$$

Proof. Let $\{x_1, \ldots, x_n\}$ be positively oriented local coordinates in a neighborhood of $p \in M$. Let $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right\}$ and $\{dx_1, \ldots, dx_n\}$ be the associated basis of T_pM and T_pM^*, respectively. Write $g_{ij}(x) = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right)(x)$, where g is the Riemannian metric of M.

In these coordinates, we can write

$$\mathrm{vol}_g = F\ dx_1 \wedge \cdots \wedge dx_n,$$

for some smooth positive function F. Now let $\{e_1, \ldots, e_n\}$ be any positive orthonormal frame of TM in a neighborhood of p, and let $\{e_1^*, \ldots, e_n^*\}$ be the dual coframe. In particular, we can write

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^{n} c_{i,j} e_j,$$

for some smooth functions $c_{i,j} \in \mathscr{C}^\infty(M)$. Now we can compute

$$F = \mathrm{vol}_g\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right) = e_1^* \wedge \cdots \wedge e_n^*\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right) = \det\left(e_j^*\left(\frac{\partial}{\partial x_i}\right)\right) = \det(C),$$

where $C = (c_{ij})$ We also have

$$g_{ij}(x) = g\left(c_{i,k}e_k, c_{j,\ell}e_\ell\right) = \sum_k \sum_\ell c_{i,k}c_{j,\ell}\, g(e_k, e_\ell) = \sum_k c_{i,k}c_{j,k}$$

Note that $(C^\top C)_{i,j} = \sum_k c_{i,k}c_{j,k}$, and so we deduce

$$\det(g_{ij}) = \det(C^\top C) = \det(C^\top)\det(C) = (\det C)^2.$$

Hence, $\det(C) = \pm\sqrt{\det(g_{ij})}$, but since $\{e_1, \ldots, e_n\}$ and $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right\}$ are both positively oriented we deduce that $\det(C) = \sqrt{\det(g_{ij})}$, and so $F = \sqrt{\det(g_{ij})}$. □

2.9. **The musical isomorphisms.** Since $\dim T_pM = \dim T_pM^*$ we should expect a natural isomorphism between these two spaces. In fact, we get such a natural isomorphism via the Riemannian metric on (M, g).

Note that $g_p : T_pM \times T_pM \to \mathbf{R}$ is a nondegenerate symmetric bilinear form, and so it induces an isomorphism

$$\begin{aligned} \flat &: T_pM \to T_pM^* \\ \flat(v) &= \langle v, \cdot \rangle, \end{aligned}$$

with inverse $\sharp : T_pM^* \to T_pM$. In local coordinates, if

$$v = a^i \frac{\partial}{\partial x^i}, \text{ then } \flat(v) = v^\flat = g_{ij}a^j dx^i.$$

This is known as *lowering the indices*.

Similarly, given a cotangent vector $\alpha \in T_pM^*$ we can *raise the indices* to obtain the corresponding tangent vector via the $\sharp$ map. In local coordinates, if

$$\alpha = a_i dx^i, \text{ then } \sharp(\alpha) = \alpha^\sharp = g^{ij}a_j \frac{\partial}{\partial x^i}.$$

These isomorphisms are known as the *musical isomorphisms*.

Remark 2.92. I haven't been too careful about my placement of indices of coefficients, but I should have always used lower indices when talking about covectors, and upper indices when talking about vectors. Similarly, I should have always written dx^i and $\frac{\partial}{\partial x^i}$ instead of dx_i and $\frac{\partial}{\partial x_i}$ – the vector bases $\frac{\partial}{\partial x^i}$ have indices lower, and the 1-form bases dx^i have the indices up. Being careful about this notation allows us to take advantage of the Einstein summation notation much more easily. ⋄

The placement of indices provides a nice mnemonic for the musical isomorphisms. In musical notation $\sharp$ indicates a half-step increase in pitch, corresponding to an upward movement on the staff. Therefore, to go from a 1-form to a vector we *raise* the indices. Similarly, $\flat$ indicates a decrease in pitch and a downward motion on the staff, and so $\flat$ lowers the indices of a vector to give use a 1-form.

Example 2.93. In a flat space we don't have to worry about the metric since $g_{ij} = \delta_{ij}$, and so for a 1-form

$$\alpha = \alpha_1 dx^1 + \cdots + \alpha_n dx^n,$$

we have

$$\alpha^\sharp = \alpha_1 \frac{\partial}{\partial x^1} + \cdots + \alpha_n \frac{\partial}{\partial x^n}.$$

⋄

These isomorphisms allow us to transport the metric $\langle \cdot, \cdot \rangle : T_pM \times T_pM \to \mathbf{R}$ to a metric $\langle \cdot, \cdot \rangle : T_pM^* \times T_pM^* \to \mathbf{R}$. Let $(x^1, \ldots, x^n)$ be a smooth coordinate system on an open subset $U \subseteq M$, and in these coordinates write

$$\langle \cdot, \cdot \rangle = g_{ij} dx^i \otimes dx^j.$$

Then on the cotangent space, we have that

$$\langle dx^i, dx^j \rangle = g^{ij}.$$

2.10. **The Hodge star operator.** Let V be an n-dimensional vector space with inner product $\langle\cdot,\cdot\rangle$. We introduce a natural inner product on the exterior powers $\bigwedge^p V$ given by

$$\langle v_1 \wedge \cdots \wedge v_p, w_1 \wedge \cdots \wedge w_p\rangle = \det\left(\langle v_i, w_j\rangle\right),$$

where the inner product is extended bilinearly. In particular, we see that if $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $(V, \langle\cdot,\cdot\rangle)$ that $\{e_{i_1} \wedge \cdots \wedge e_{i_p} : 1 \le i_1 < i_2 < \cdots < i_p \le n\}$ forms an orthonormal basis of $\bigwedge^p V$.

Remark 2.94. We extend the inner product to the entire exterior algebra $\bigwedge^* V$ be considering elements of different degrees as orthogonal. $\diamond$

Note that $\dim \bigwedge^k V = \binom{n}{k} = \binom{n}{n-k} = \dim \bigwedge^{n-k} V$, and so we should expect an isomorphism between these two exterior products. Throughout this section we consider a fixed orientation on V.

Definition 2.95. The **Hodge star operator** is the linear map $\star : \bigwedge^p V \to \bigwedge^{n-p} V$ given by

$$\star(e_{i_1} \wedge \cdots \wedge e_{i_p}) = e_{j_1} \wedge \cdots \wedge e_{j_{n-p}},$$

where $j_1, \ldots, j_{n-p}$ is chosen such that $e_{i_1}, \ldots, e_{i_p}, e_{j_1}, \ldots, e_{j_{n-p}}$ is a positive orthonormal basis of V. $\diamond$

Note that the $\star 1 = e_1 \wedge \cdots \wedge e_n$ and $\star(e_1 \wedge \cdots \wedge e_n) = 1$ if $e_1, \ldots, e_n$ is a positive orthonormal basis. We still need to show that $\star$ is well defined, in the sense that it does not depend on the choice of positive orthonormal basis. Let A be a $n \times n$-matrix, and consider $v_1, \ldots, v_n \in V$. Then it follows from basic linear algebra that

$$\star(Av_1 \wedge \cdots \wedge Av_n) = (\det A) \star (v_1 \wedge \cdots \wedge v_n).$$

From this we see that the Hodge star does not depend on the choice of positive orthonormal basis, since any two such bases are related by a transition matrix in $\mathbf{SO}(n)$.

Proposition 2.96. *The Hodge star operator satisfies the following properties:*

(i) $\star\star = (-1)^{k(n-k)} : \bigwedge^k V \to \bigwedge^k V$,

(ii) for $v, w \in \bigwedge^k V$,

$$\langle v, w\rangle = \star(w \wedge \star v) = \star(v \wedge \star w),$$

(iii) for an arbitrary positive basis $\{v_1, \ldots, v_n\}$ *we have*

$$\star 1 = \frac{1}{\sqrt{\det\left(\langle v_i, v_j\rangle\right)}} v_1 \wedge \cdots \wedge v_n.$$

Proof.

(i) Clearly $\star\star : \bigwedge^k V \to \bigwedge^k V$. Suppose

$$\star(e_{i_1} \wedge \cdots \wedge e_{i_k}) = e_{j_1} \wedge \cdots \wedge e_{j_{n-k}}.$$

By definition of the Hodge star operator we see that

$$\star\star(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \pm e_{i_1} \wedge \cdots \wedge e_{i_k},$$

where the ambiguity in the sign depends on whether $e_{j_1}, \ldots, e_{j_{n-k}}, e_{i_1}, \ldots, e_{i_k}$ is a positive or negative basis of V. Using the antisymmetry of the wedge product we have that

$$e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge e_{j_1} \wedge \cdots \wedge e_{j_{n-k}} = (-1)^{k(n-k)} e_{j_1} \wedge \cdots \wedge e_{j_{n-k}} \wedge e_{i_1} \wedge \cdots \wedge e_{i_p}.$$

So by considering the basis transformation matrix τ from the positive basis $e_{i_1}, \ldots, e_{i_p}, e_{j_1}, \ldots, e_{j_{n-p}}$ to the unknown signed basis $e_{j_1}, \ldots, e_{j_{n-k}}, e_{i_1}, \ldots, e_{i_k}$, we see that $\det(\tau) = (-1)^{k(n-k)}$; the result now follows.

(ii) It suffices to prove this for basis elements of $\bigwedge^k V$. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of V, and consider the basis $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$ of $\bigwedge^k V$. Note that for any two distinct elements v, w in this basis we have $w \wedge \star v = 0$. So we are left to compute $\star(v \wedge \star v)$. For any choice of $1 \le i_1 < \cdots < i_k \le n$,

$$\star(e_{i_1} \wedge \cdots \wedge e_{i_k} \wedge \star(e_{i_1} \wedge \cdots \wedge e_{i_k})) = \star(e_1 \wedge \cdots \wedge e_n) = 1.$$

The result follows by the bilinearity of the inner product, the bilinearity of the wedge product, and the linearity of the $\star$ operator.

(iii) Since

$$v_1 \wedge \cdots \wedge v_n = \left(\det\left(\langle v_i, v_j\rangle\right)\right)^{\frac{1}{2}} e_1 \wedge \cdots \wedge e_n$$

and $\star 1 = e_1 \wedge \cdots \wedge e_d$ the result follows.

□

Now we extend the Hodge star to a $\mathscr{C}^\infty(M)$-linear map over $\bigwedge^k M^*$. Let M be an *oriented* Riemannian manifold of dimension n. Since M is oriented, we can find a consistent orientation of all of the tangent and cotangent spaces, T_pM and T_pM^* respectively. Since M is a Riemannian manifold, we have an inner product $\langle \cdot, \cdot \rangle$ on each T_pM^* given by $(g^{ij}) = (g_{ij})^{-1}$, and so we obtain a Hodge star operator

$$\star : \bigwedge^k (T_pM^*) \to \bigwedge^{n-k} (T_pM^*).$$

In particular, we can view this operator as a basepoint preserving operator $\star : \Omega^k(M) \to \Omega^{n-k}(M)$. By the previous lemma, in local coordinates we see that

$$\star 1 = \sqrt{\det(g_{ij})}\, dx_1 \wedge \cdots \wedge dx_n.$$

In particular, this is the *Riemannian volume form*!

Corollary 2.97. *For a k-form $\alpha \in \Omega^k(M)$, $\star\alpha \in \Omega^{n-k}(M)$ is the unique $(n-k)$-form such that*

$$\beta \wedge \star\alpha = \langle \beta, \alpha \rangle_g \ \mathrm{vol}_g,$$

for all $\beta \in \Omega^k(M)$.

Now we use the Hodge star to induce an L^2-inner product on $\Omega^k(M)$.

Definition 2.98. Let $\alpha, \beta \in \Omega^k(M)$ have compact support. The **L^2-inner product** of α, β is given by

$$\langle\langle \alpha, \beta \rangle\rangle := \int_M \langle \alpha, \beta \rangle \star 1 = \int_M \alpha \wedge \star\beta.$$

⋄

This inner product is clearly positive and positive definite. We define the L^2-norm as

$$\|\alpha\| = \langle\langle \alpha, \alpha \rangle\rangle^{\frac{1}{2}}.$$

Note that the space of differential forms is not complete with respect to this norm, and hence it is not a Hilbert space. Later we will consider the completion of this space and consider L^p and Sobolev spaces of sections of vector bundles.

We now give a taste of integration by parts on a Riemannian manifold.

Definition 2.99. The **codifferential** $\delta : \Omega^p(M) \to \Omega^{p-1}(M)$ is defined by

$$\delta\omega = (-1)^{n(p+1)+1} \star d \star \omega.$$

⋄

We claim that δ is the formal adjoint of d.

Theorem 2.100. *Let (M, g) be a compact manifold without boundary. Then*

$$\int_M \langle \delta\alpha, \beta \rangle \ \mathrm{vol}_g = \int_M \langle \alpha, d\beta \rangle \ \mathrm{vol}_g,$$

for $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^{p-1}(M)$.

Proof. A direct computation shows

$$\begin{aligned}\int_M \langle \alpha, d\beta\rangle \ \mathrm{vol}_g &= \int_M d\beta \wedge \star\alpha \\ &= \int_M \left(d(\beta\wedge\star\alpha) + (-1)^p \beta \wedge d\star\alpha\right) \\ &= \int_M (-1)^{p+(n-p+1)(p-1)} \beta \wedge \star\star d\star\alpha \\ &= \int_M \langle \beta, (-1)^{n(p+1)+1} \star d \star \alpha\rangle \ \mathrm{vol}_g \\ &= \int_M \langle \beta, \delta\alpha\rangle \ \mathrm{vol}_g \,.\end{aligned}$$

□

For a Riemannian manifold we can define the curl and the divergence as follows:

Definition 2.101. Let M be an 3-dimensional manifold. Let $X \in \Gamma(TM)$ be a smooth vector field. Then define the **divergence** of X as

$$\nabla\cdot X = \star d \star \flat X = \delta(X^{\beta}).$$

⋄

This definition will suffice for now, and we will discuss the divergence of tensor fields more once we introduce the notion of a linear connection.

Definition 2.102. Let M be an 3-dimensional manifold. Let $X \in \Gamma(TM)$ be a smooth vector field. The **curl** of X is the vector field

$$\nabla \times X = \left(\star dX^{\flat}\right)^{\sharp}.$$

⋄

It is a straightforward exercise to show that these definitions coincide with the usual definitions when $M = \mathbf{R}^n$ (for the divergence) or $M = \mathbf{R}^3$ for the curl.

We now conclude this section on differential forms by providing an example of how differential forms can be used to describe natural physical forces.

2.11. **Maxwell's equations.** As an instructive example, we want to consider a reformulation of Maxwell's equations in terms of the Hodge-star and the d-operator on differential forms in flat Minkowski space-time. We will see that there are two equations in differential forms that encode the four classical equation, and conservation of charge ($\partial_t\rho + \nabla\cdot \boldsymbol{j} = 0$) is easily deduced from these much like in the classical setup with vector calculus.

The reformulation in the language of differential forms has several advantages: it clarifies the close link between topology and the theory of electromagnetic potential fields, it gives a straightforward explanation of the relationship between Maxwell's equations and special relativity, and by replacing the pseudo-Riemannian tangent bundle with a more general structure (vector bundle with connection over the spacetime manifold) the Maxwell equations in terms of differential forms become a special case of the Yang-Mills equations (a fact that is difficult to perceive when the equations are expressed in coordinatized form instead of in terms of the common mathematical language of vector bundles over manifolds). I won't discuss the Yang-Mills theory any further here (for now).

Of course, to really understand the physical relevance of the equations and where they come from one must approach these matters through action principles and other ideas from physics (Lagrangians and Hamiltonians, etc.). The purpose here is not to derive physical laws from first principles, but rather to explain how to put some known physical laws in a convenient mathematical form that enables us to better understand some of there mathematical properties.

2.11.1. *Classical Setup and Orientations.* We fix a 3-dimensional *flat* Riemannian manifold with corners S and let $X = S \times \mathbf{R}$ but endowed with the Lorentzian *product metric* that comes fromt he metric tensor on S and the negative-definite flat metric on $\mathbf{R}$ induced by the quadratic form $-t^2$. Thus, for all $x = (s,t) \in X$ the vector space decomposition $T_xX = T_s(S) \oplus T_t\mathbf{R}$ is orthogonal. The classical case is $S = \mathbf{R}^3$ with its standard inner product (and associated flat metric). One puzzling feature of the classical case is that there should be no preferred point in space, and so in particular no meaningful "linear structure" on space. It is therefore a bit peculiar to say that classically S is a vector space. In classical physics what happens is that at the beginning of every physical problem one chooses an origin and somehow this choice never affects the answer. It would be better to have a framework in which there is no need to be choosing random origins, but I won't discuss the matter any further here; our space S is a smooth manifold and so the issue of an origin and linear structure on S is eliminated (but we retain this information of the tangent spaces at points, which is really what matters).

It will be convenient to assume S is connected and oriented (as in the classical case), but in the end we will get equations that do not require an orientation on S or even that S is orientable. Note that since $X = S \times \mathbf{R}$ with $\mathbf{R}$ oriented in the canonical way, orientability of S is equivalent to that of X. An orientation on either of S or X uniquely determines an orientation on the other so that X has a product orientation. I'll speak in the language of orientations on S. The Lorentzian manifold with corners X is called *spacetime*. The flatness of the metric tensor on S will be essential for everything we do. The role of flatness is to permit us to carry out local calculations in *flat* coordinate systems, in terms of which the metric tensor acquires the same simple form as in the classical case. The case of non-flat metrics is the framework of General Relativity (as opposed to special relativity, which is essentially the context in which we are working).

Our goal is to write down Maxwell's equations in the language of differential forms on X in the case when there is no magnetic or polarized material present and the units are chosen to trivialize natural constants: $c = \varepsilon_0 = \mu_0 = 1$. We first make an important definition that encodes the fact that the force fields of classical physics exhibit a time dependence in their evolution but they do not point in a "time direction."

Definition 2.103. A smooth vector field $\vec{v}$ over an open set U in $X = S \times \mathbf{R}$ is *spacelike* if for each $u = (s,t) \in U$ the tangent vector $\vec{v}(u) \in T_uX \cong T_sS \oplus T_r\mathbf{R}$ lies in the hyperplane T_sS. ⋄

In other words, if we choose local coordinates $\{x,y,z\}$ on S and write a vector field locally as a smooth linear combination of $\partial_x, \partial_y, \partial_z$, and ∂_t then the spacelike condition on $\vec{v}$ says that the ∂_t-component of $\vec{v}$ vanishes. More globally, if $p : X \to S$ is the natural projection then $p^*(TS)$ is naturally a subbundle (even direct summand) of TX and the sections of this subbundle are the spacelike vector fields.

Definition 2.104. For an open set $U \subseteq X$ and a vector field $\vec{v} \in \Gamma(TU \subseteq TX)$, let $\omega_{\vec{v}} \in \Omega^1_X(U)$ be the 1-form that is dual to $\vec{v}$ under the Lorentz metric. That is, for all $u \in U$, the linear functional $\omega_{\vec{v}}(u)$ on T_uX given by $\langle \cdot, \vec{v}(u)\rangle_u$. ⋄

Note in particular that since we give $X = S \times \mathbf{R}$ a product metric, if the vector field $\vec{v}$ is spacelike then for any $u_0 = (s_0, t_0) \in U$ the functional $\omega_{\vec{v}}(u_0)$ kills the line $T_{t_0}\mathbf{R}$ in $T_{u_0}X$ and hence

$$\omega_{\vec{v}}(u_0) \in T_{s_0}S^\vee \subseteq T_{s_0}S^\vee \oplus T_{t_0}\mathbf{R}^\vee = T_{u_0}X^\vee.$$

Explicitly, if we choose a local flat coordinate system $\{x,y,z\}$ on $U_0 \subseteq S$ then for a smooth vector field $\vec{v} = f_1\partial_x + f_2\partial_y + f_3\partial_z + f_4\partial_t$ over an open set $U \subseteq U_0 \times \mathbf{R}$ (with $f_i \in \mathscr{C}^\infty(U)$) we compute pointwise

$$\omega_{\vec{v}} = f_1dx + f_2dy + f_3dz - f_4dt \in \Omega^1_X(U).$$

The correctness of this calculation rests crucially on the fact that $\{\partial_x, \partial_y, \partial_z\}$ is an orthonormal frame for $TS|_{U_0}$ and that $\langle \partial_t, \partial_t\rangle = 1$.

The classical operators of divergence, gradient, and curl for vector fields over open subsets of $\mathbf{R}^3$ can be generalized to 4-dimensional Lorentzian spaces by working in 3-dimensional time slices:

Definition 2.105. Let S be an oriented 3-dimensional Riemannian manifold with boundary or corners. Let $X = S \times \mathbf{R}$ and let $p : X \to S$ be the standard projection. For an open set $U \subseteq X$ and a smooth U-section $\vec{v}$ of the subbundle $p^*(TS) \subseteq TX$ of spacelike vector fields, the **spacelike curl** $\nabla_S \times \vec{v} \in \Gamma(TU)$ is the spacelike vector field given on each time slice $U_t = U \cap (S \times \{t\})$ by the ordinary curl applied to the smooth vector field $\vec{v}|_{U_t} \in \Gamma_S(U_t)$. The spacelike divergence and spacelike gradient are defined analogously. ⋄

Explicitly, if $\{x, y, z\}$ are local oriented flat coordinates on S then in the local coordinate system $\{x, y, z, t\}$ the above three spacelike operators are given by the habitual formulas in each time slice (using the differential operators $\partial_x, \partial_y, \partial_z$, and not ∂_t). Hence, we see that smoothness is preserved by these operations. Note that the spacelike divergence and spacelike gradient are independent of the orientation (as this is true on each time slice), and so they make sense without any orientability hypotheses on S (by globalizing from the local orientable case). In contrast, the spacelike curl only makes sense in the orientable case and negating the orientation cause it to change by a sign.

Beware that the spacelike divergence on spacelike vector fields $\vec{v}$ is rather different from the "usual" generalized divergence $\star_4 d \star_1 \omega_{\vec{v}} \in \mathscr{C}^\infty(U)$ that one would get through the global pseudo-Riemannian structure on X in the sense that the generalized divergence involves a t-derivative of the (usually t-dependent) local coefficient functions in local oriented flat coordinates $\{x, y, z, t\}$.

Remark 2.106. If $f \in \mathscr{C}^\infty(U)$ is a smooth function, then a simple calculation using local flat coordinates on S yields the identity

$$df = \omega_{\nabla_S f} + (\partial_t f)dt$$

as the unique decomposition of $df \in \Omega^1_X(U)$. Indeed, if $\{x, y, z\}$ is a local *flat* coordinate system then this identity is just the expansion

$$df = (\partial_x f)dx + (\partial_y f)dy + (\partial_z)dz + (\partial_t f)dt$$

that one has for an arbitrary local smooth coordinate system on X. This global decomposition identity for df will come up later in our considerations of potential functions for the electric field and vector potentials for the magnetic field. ⋄

The *classical Maxwell equations* on open sets U in X are as follows: for spacelike vector fields $\mathbf{E}$ and $\mathbf{B}$ on U expressing the electric and magnetic fields as functions of position and time (so $\mathbf{B}$ is sign-dependent on the choice of orientation on S),

$$\begin{aligned}
\nabla_S \cdot \mathbf{B} &= 0, \\
\nabla_S \times \mathbf{E} + \partial_t \mathbf{B} &= 0, \\
\nabla_S \times \mathbf{B} &= \boldsymbol{j} + \partial_t \mathbf{E}, \\
\nabla_S \cdot \mathbf{E} &= \rho,
\end{aligned}$$

where $\rho : U \to \mathbf{R}$ is called the electric charge density and $\boldsymbol{j}$ is a spacelike vector field that is called the current density over U. These equations are respectively called *non-existence of magnetic monopoles*, *Faraday's law of induction*, *Ampere's Law*, and *Gauss' law for electricity*. In the classical case, these are the tradition equations of Maxwell's theory.

The supplementary law of conservation of charge $\partial_t \rho + \nabla_S \cdot \boldsymbol{j} = 0$ is an immediate consequence of taking the t-partial derivative of Gauss' law for electricity and the divergence of Ampere's law: the two sides of the identity for conservation of charge are simply two different ways to compute $\partial_t(\nabla_S \cdot \mathbf{E}) = \nabla_S \cdot (\partial_t E))$.

Remark 2.107. There is much more to classical electrostatics than Maxwell's equations, such as Coloumb's law and the action principles that construct potential fields *a priori*. ⋄

Observe that just as the definition of $\mathbf{B}$ is sign-dependent on a choice of orientation for S, the spacelike curl also has such sign dependence. This is good, because one sees by inspection that all four classical Maxwell equations are thereby independent of the choice of orientation: the left side of the second equation is sign-dependent but the property that it equal 0 is thereby unaffected, and the third equation has no orientation intervention on the right side but has it intervening twice (in the formation of $\nabla_S \times \mathbf{B}$) and thereby cancelling out on the left side.

We conclude that the classical Maxwell equations only require the flat Riemannian structure and orientability of S; they do not depend on the *choice* of orientation. Since equations in physics should not be coordinate-dependent, the above coordinate-free equations clarify the underlying geometrical aspects of the classical Maxwell theory. However, there are some defects. First of all, the equations involve the input $\mathbf{B}$ whose definition necessitates a choice of global orientation. Since there does not seem to be a natural orientation in the real world, it is preferable if we can formulate the equations without such a choice (even if the real world is orientable). Also, we would like to understand how the theory of electromagnetic potential is controlled by geometry in spacetime, and the classical formulation is not well-suited for such questions. We shall recast the equations in terms of differential forms on the abstract flat Lorentzian manifold $X = S \times \mathbf{R}$ for any abstract flat Riemannian 3-dimensional manifold with corners S, and in so doing we will be able to eliminate orientation conditions and the topological input will become clearer.

2.11.2. *Some useful identities.* Let us temporarily assume S is orientable and connected. Choose an orientation (there are two), whence we get a product orientation on $X = S \times \mathbf{R}$. It will be seen that the equations we get in the end will not depend on this choice, so the equations will globalize to the case of possibly non-orientable S. The choice of orientation gives rise to the Hodge-star bundle isomorphisms $\star_r : \Omega^r_X \cong \Omega^{4-r}_X$ that satisfy $\star_1 4 - r \circ \star_r = (-1)^{r+1}$. Letting vol denote the volume form on X arising from the Lorentz structure and the orientation, in local oriented flat coordinates $\{x, y, z\}$ over an open $U_0 \subseteq S$ we have $\mathrm{vol} = dx \wedge dy \wedge dz \wedge dt$ over $U_0 \times \mathbf{R}$. Of course, as always we have $\mathrm{vol} = \star(1)$. and so we will usually write "$\star(1)$" rather than "vol". In particular, if we change the orientation of S then vol is negated since the only other orientation on S is the opposite one (as S is connected). Recall that $\star_r$ is characterized by the local identities $\omega \wedge \star\eta = \langle \omega, \eta \rangle\, \mathrm{vol}$. Hence, if we negate the orientation on S then each $\star_r$ is also negated. So we know that the operator $\delta = \star_{5-r} \circ d \circ \star_r : \Omega^r_X \to \Omega^{r-1}_X$ for $1 \le r \le 4$ is independent of the orientation and so globalizes to the case when there are no orientability (or connectivity) hypotheses on S.

The following lemma is the result of a simple direct computation.

Lemma 2.108. *Fixing an orientation on S and using the induced product orientation on $X = S \times \mathbf{R}$, if $\{x, y, z\}$ is a local oriented flat coordinate system on an open $U_0 \subseteq S$, then*

$$\star_1(dx) = dy \wedge dz \wedge dt, \qquad \star_1(dy) = -dx \wedge dz \wedge dt, \qquad \star_1(dz) = dx \wedge dy \wedge dt, \qquad \star_1(dt) = dx \wedge dy \wedge dz$$

in $\Omega^3_X(U_0 \times \mathbf{R})$ and

$$\star_2(dx \wedge dy) = dz \wedge dy, \quad \star_2(dx \wedge dz) = -dy \wedge dt, \quad \star(dx \wedge dt) = -dy \wedge dz,$$
$$\star_2(dy \wedge dz) = dx \wedge dt, \quad \star_2(dy \wedge dt) = dx \wedge dz, \quad \star(dz \wedge dt) = -dx \wedge dy$$

in $\Omega^2_X(U_0 \times \mathbf{R})$. Moreover, $\star_2^2 = -1$, $\star_1 \circ \star_3 = 1$, and $\star_3 \circ \star_1 = 1$.

We next need some global identities that relate the spacelike divergence and curl (for spacelike vector fields) with the d and Hodge-star operators. We first have to define the time-derivative of a spacelike vector field. This goes as follows:

If $\vec{v} \in \Gamma_X(TU)$ is an arbitrary smooth spacelike vector field, then there is a unique identity of the form

$$d\omega_{\vec{v}} = -\theta \wedge dt + \eta$$

where the 1-form θ is a section of $p^*(\Omega^2_S)$ and the 2-form η is a section of $p^*(\Omega^1_S)$. By the duality between 1-forms and vector fields provided by the Lorentz metric, we can therefore uniquely write $\theta = \omega_{\partial_t \vec{v}}$ for a unique smooth *spacelike* vector field $\partial_t \vec{v}$ over U. Explicitly, for local coordinates $\{x, y, z\}$ on S (unrelated to the orientation and Riemannian structure) we can uniquely write $\vec{v} = f_1\partial_x + f_2\partial_y + f_3\partial_z$ with $f_i \in \mathscr{C}^\infty(U)$ and it is easy to check directly that

$$\partial_t(\vec{v}) = (\partial_t f_1)\partial_x + (\partial_t f_2)\partial_y + (\partial_t f_3)\partial_z \in \Gamma_X(U).$$

Lemma 2.109. *Assume S is orientable and fix an orientation. Let $U \subseteq X = S \times \mathbf{R}$ be an open subset and $\vec{v} \in \Gamma_X(U)$ a spacelike smooth vector field. The following hold:*

(1) $\delta(\omega_{\vec{v}}) = -\nabla_S \cdot \vec{v}$,
(2) $(d\omega_{\vec{v}}) \wedge dt = \star_1(\omega_{\nabla_S \times \vec{v}})$,
(3) $\delta(\omega_{\vec{v}} \wedge dt) = -(\nabla_S \cdot \vec{v})dt - \omega_{\partial_t \vec{v}}$,
(4) $d\omega_{\vec{v}} = \star_2(\omega_{\nabla_S \times \vec{v}} \wedge dt) - \omega_{\partial_t \vec{v}} \wedge dt$.

Remark 2.110. Note the important consistency check that both sides of (1) and (3) are independent of the orientation on S. Also the left sides of (2) and (4) are independent of the orientation on S, and so are the right sides becuase they involve a single Hodge star but also a single spacelike curl operator. ⋄

Proof. It is possible to deduce these identities by pure thought using (i) the definitions of the spacelike divergence and curl in terms of the 3-dimensional counterparts on time slices, and (ii) the relations between the classical divergence and curl with d and Hodge star in the classical 3-dimensional case on $\mathbf{R}^3$. However, it is a notational pain to give such an "intrinsic" proof. Hence, we instead leave it to the reader to carry out the pleasant exercise of verifying the identities by coordinate calculation upon picking local oriented flat coordinates on S. This is essentially a mechanical exercise once one has available the identities in the previous lemma (and one knows how the 3-dimensional curl and divergence work out in such local coordinates). ☐

We have one final lemma.

Lemma 2.111. *Fix an orientation on S. For any open set $U \subseteq X = S \times \mathbf{R}$ and smooth differential forms $F \in \Omega^2_X(U)$ and $J \in \Omega^1_X(U)$ there exists a unique smooth function $\rho \in \mathscr{C}^\infty(U)$ and unique spacelike smooth vector fields $\mathbf{E}, \mathbf{B}, \boldsymbol{j} \in \Gamma_X(U)$ such that*

$$F = \star_2(\omega_{\mathbf{B}} \wedge dt) - \omega_{\mathbf{E}} \wedge dt, \qquad J = \rho\, dt - \omega_{\boldsymbol{j}}.$$

The vector field $\mathbf{E}$ and $\boldsymbol{j}$ are orientation-independent, as is the function ρ, but $\mathbf{B}$ changes by a sign if we negate the orientation on S.

Proof. As with the proof of the previous lemma, one can give a proof without mentioning any local coordinates but we take the quick way out. Pick local oriented flat coordinates on S and write out the "general form" of the right side in terms of ρ and coefficient functions of the vector fields. One sees by inspection that these are just an encoding of the coefficient functions of the differential forms F and J. This gives the asserted existence/uniqueness results locally, and due to the local uniqueness it follows that the local solutions agree on overlaps and hence globalize. As for the sign-dependence on the orientation of S, this is immediate from the fact that the only ingredient in the "shape" of the formulas that depends on the orientation is the Hodge star $\star_2$. This changes by a sign if we change the orientation of S, and so by uniqueness it follows that $\mathbf{B}$ must also exhibit the same sign dependence in order for the effect to cancel out and give the initial choice of F. ☐

2.11.3. *The modern formulation of Maxwell's equations.* We now drop all orientability and connectivity hypothesis on S. Choose an open set $U \subseteq X$ and pick smooth differential forms $F \in \Omega^2_X(U)$ and $J \in \Omega^1_X(U)$. We call these the *electromagnetic form* and the *current density form*, respectively. The *abstract Maxwell equations* are then

$$dF = 0, \tag{2}$$

$$\delta F = J, \tag{3}$$

where $\delta = \star d \star$ over orientable open subsets (for any choice of orientation). The first equation encodes the non-existence of magnetic monopoles and Faraday's law of induction, and the second encodes Gauss' law for electricity and Ampere's law, as we'll show below.

It is the second equation ($\delta F = J$) that encodes the serious physical information, in the sense that the first equation ($dF = 0$) is a physical triviality: action principles that logically precede the electromagnetic theory provide an electromagnetic potential and it will be seen below that in terms of differential forms this leads to the condition $F = dA$ as *a-priori* input in the theory from a physical point of view. This makes the first equation physically uninteresting since it follows from the general identity $d \circ d = 0$. On the other hand, the second equation takes the form $\delta dA = J$ and this turns out to be of enormous physical significance.

The local calculation

$$d \star_1 (J) = d(\star_1 \circ \star_4 \circ d \star_2 (F)) = dd(\star_2 F) = 0 \tag{4}$$

resting on the second equation 3 will turn out to be a repackaging of the identity for conservation of charge. The calculation above is a disguised version of the classical deduction of this conservation law from Maxwell's equation. Finally, the reason for the names of F and J is that they will turn out to encode precisely the information of the electromagnetic fields and the charge/current densities.

How do we make the translation from the classical equations in the oriented case? By Lemma 2.111 if we assume S is orientable and we pick an orientation of S then we get a unique $\rho \in \mathscr{C}^\infty(U)$ and unique spaceklike vector fields $\mathbf{E}, \mathbf{B}, \boldsymbol{j} \in \Gamma_X(TU)$ such that they recover F and J via the formulas in Lemma 2.111. In particular, $\rho, \mathbf{E}$, and $\boldsymbol{j}$ are orientation-independent, but $\mathbf{B}$ changes by a sign on a connected component of U if we change the orientation on the component. Upon fixing an orientation to get a definite $\mathbf{B}$ and to be able to write $\delta = \star d \star$ we now use Lemma 2.109 to compute

$$\begin{aligned} dF &= d\star(\omega_{\mathbf{B}} \wedge dt) - d(\omega_{\mathbf{E}} \wedge dt) \\ &= \star(\star d\star(\omega_{\mathbf{B}} \wedge dt)) - d\,(\omega_{\mathbf{E}} \wedge dt) \\ &= \star\left(-(\nabla_S \cdot \mathbf{B})dt - \omega_{\partial_t \mathbf{B}}\right) - \star\omega_{\nabla_S \times \mathbf{E}} \\ &= (-\nabla_S \cdot \mathbf{B})(\star dt) - \star\left(\omega_{\partial_t \mathbf{B} + \nabla_S \times \mathbf{E}}\right). \end{aligned}$$

Thus, the condition $dF = 0$ says exactly $\nabla_S \cdot \mathbf{B} = 0$ and $\partial_t \mathbf{B} + \nabla_S \times \mathbf{E} = 0$. These are exactly the non-existence of magnetic monopoles and Faraday's law, as promised.

Since $\star_2 \circ \star_2 = -1$, we similarly compute via Lemma 2.109 that

$$\begin{aligned} \delta F &= (\star d \star)(F) \\ &= -\star d(\omega_{\mathbf{B}} \wedge dt) - (\star d \star)(\omega_{\mathbf{E}} \wedge dt) \\ &= -\star(\star\omega_{\nabla_S \times \mathbf{B}}) + ((\nabla_S \cdot \mathbf{E})dt + \omega_{\partial_t \mathbf{E}}) \\ &= \omega_{-\nabla_S \times \mathbf{B} + \partial_t \mathbf{E}} + (\nabla_S \cdot \mathbf{E})dt, \end{aligned}$$

so the condition $\delta F = J := \rho\ dt - \omega_{j} = \rho\ dt + \omega_{-j}$ says $\nabla_S \cdot \mathbf{E} = \rho$ and $\nabla_S \times \mathbf{B} = \partial_t \mathbf{E} + \boldsymbol{j}$. These two identities are respectively Gauss' law for electricity and Ampere's law.

Finally, since $\langle dt, dt \rangle = -1$ we have $dt \wedge \star dt = -\star(1)$, and hence

$$d(f \cdot (\star dt)) = df \wedge \star(dt) = \partial_t f \cdot \star 1$$

for any $f \in \mathscr{C}^\infty(U)$. Taking $f = \rho$, one compute (using $\star_0 \circ \star_4 = -1$) that

$$\begin{aligned} d\star J &= d(\rho\star(dt)) - \star\omega_{j} \\ &= d\rho \wedge \star(dt) + \star(\star d\star\omega_{-j}) \\ &= -(\partial_t \rho + \nabla_S \cdot \boldsymbol{j})\star(1), \end{aligned}$$

where the last step again follows from Lemma 2.109. Thus, the *abstract* calculation in (4) that $d(\star J) = 0$ says precisely that $\partial_t \rho + \nabla_S \cdot \boldsymbol{j} = 0$, and this latter identity is simply conservation of charge. The vanishing of $d \circ d$ as used in (4) is just a repackaging of the vanishing of $\nabla_S \cdot (\nabla_S \times (\cdot))$ on spacelike vector fields, and this latter vanishing (applied on time slices) is the content of the classical deduction of the conservation of charge from Maxwell's equations.

Remark 2.112. Suppose S is oriented. On open subsets $U \subseteq X$ on which the current density J vanishes, the 2-form $\star F \in \Omega^2_X(U)$ is closed and so defines a cohomology class in $H^2_{dR}(U)$ that is supposed to have some physical significance. ⋄

2.11.4. *Topological consequences: potential fields.* So far, the above has just been clever linguistic repackaging. The question now is: does the rephrasing in terms of differential forms tell us anything interesting?

Theorem 2.113. *Let $U \subseteq X = S \times \mathbf{R}$ be an open subset that is smoothly contractible to a point. If $F \in \Omega^2_X(U)$ is a closed 2-form then it can be expressed as $F = dA$ for a smooth 1-form $A \in \Omega^1_X(U)$ that is unique up to an additive term df for some $f \in \mathscr{C}^\infty(U)$.*

Proof. This is an immediate consequence of the Poincaré lemma, i.e. that F is exact. Upon writing $F = dA$, the extent to which $A \in \Omega^1_X(U)$ is non-unique is adding a closed 1-form to A. Again, since U is contractible, we deduce that this one form must be of the form df for some $f \in \mathscr{C}^\infty(U)$. □

What is the meaning of A in the classical theory? I claim that it encodes the theory of electromagnetic potential. Since $X = S \times \mathbf{R}$ we can uniquely write $A = \eta + \phi\ dt$ with $\phi \in \mathscr{C}^\infty(U)$ and $\eta \in \Omega^1_X(U)$ a 1-form such that $\eta(u) \in T_u X^\vee \cong S^\vee \oplus \mathbf{R}^\vee$ has vanishing $(dt)(u)$-component for all $u \in U$. We may write $\eta = \omega_{\mathbf{A}}$ for a unique smooth spacelike vector field $\mathbf{A} \in \Gamma_X(U)$. Since df likewise has the spacetime decomposition

$$df = \omega_{\nabla_S f} + (\partial_t f)dt,$$

replacing A with $A + df$ corresponds to the change

$$(\mathbf{A}, \phi) \mapsto (\mathbf{A} + \nabla_S f, \phi + \partial_t f).$$

The spacelike vector field $\mathbf{A}$ and the smooth function $\phi : U \to \mathbf{R}$ are orientation-independent and are together unique up to a linked change in terms of $f \in \mathscr{C}^\infty(U)$.

The physical meaning is seen as follows: taking S to be oriented now, we have

$$\star(\omega_{\mathbf{B}} \wedge dt) - \omega_{\mathbf{E}} \wedge dt = F = dA = d(\omega_{\mathbf{A}} + \phi\ dt) = d\omega_{\mathbf{A}} + d\phi \wedge dt.$$

Since $d\phi = \omega_{\nabla_S \phi} + (\partial_t \phi)\ dt$, clearly $d\phi \wedge dt = \omega_{\nabla_S \phi} \wedge dt$. Also, the last identity in Lemma 2.109 gives

$$\mathbf{B} = \nabla_S \times \mathbf{A}, \qquad \mathbf{E} = -\nabla_S \phi + \partial_t \mathbf{A}$$

with $\mathbf{A}$ a spacelike vector field on U and $\phi \in \mathscr{C}^\infty(U)$ a function such that the pair $(\mathbf{A}, \phi)$ is uniquely determined up to adding $\nabla_S f$ to $\mathbf{A}$ and $\partial_t f$ to ϕ for some $f \in \mathscr{C}^\infty(U)$.

This is *exactly* the classical theory of electromagetric potential: the vector field $\mathbf{A}$ is called the *vector potential* for the magnetic field and ϕ is called the *electrostatic potential function*. In the absence of magnetic fields ($\mathbf{B} = 0$) we may take the vector potential $\mathbf{A}$ to vanish and so the electrostatic potential function ϕ is uniquely determined up to adding a function $\partial_t f$ such that the spacelike gradient $\nabla_S f$ vanishes (to retain the condition of vanishing vector potential). But the condition of vanishing for the spacelike gradient says precisely that f is locally "independent of the space variables", and so in the special case that $U = U_0 \times I$ for an open set $U_0 \subseteq S$ it follows that f is a smooth function of time and hence the electrostatic potential ϕ is unique up to adding an arbitrary function of time. In this case we may fix the value of ϕ to be a specific constant at one point $u_0 \in U_0$ for all time $t \in I$, and this eliminates all of the ambiguity. This is precisely the classical device of uniquely determining the electrostatic potential (in the absence of magnetic forces) by requiring it to be zero at some point of U_0.

Of course, since there are other physical laws such as Coloumb's law and action principles, it is always possible to infer the existence of an electrostatic potential function even when the de Rham cohomology is nonzero. That is, the physics tells us a lot more than what is mathematically deducible from Maxwell's theory alone. In particular, the electromagnetic potential is much more than just a device for extracting the electromagnetric field, and so the mathematics is not the whole story.

Example 2.114. In the case of a time-dependent magnetic field complementary to a line in space, there is a vector potential (since the relevant de Rham cohomology is an H^2 that vanishes) but its non-uniqueness is controlled by an H^1 that is nonzero and so it is possible to change the choice of the vector potential by more than just vector fields that are spacelike gradients. That is, in such cases there are spacelike vector fields $\vec{v}$ on the domain that are not gradients and yet have vanishing curl, so we can add such a $\vec{v}$ to the vector potential without affecting its property of having curl equal to the magnetic field. (By taking $\vec{v}$ to be constant in time, so $\partial_t \vec{v} = 0$, this modification of the vector potential does not force any changes in the choice of electrostatic potential function.) But is this physically relevant? After all, presumably the potential is chosen according to principles that go beyond just Maxwell's theory, and so the additive modification by a curl-free non-gradient field $\vec{v}$ as suggested above may well be unreasonable on physical grounds. I don't have the technical background to say anything else intelligent about this. ⋄

The main point is this: the existence of electromagnetic potential can be understood in many situations purely based on topological properties of the domain under consideration, and when the region is topologically complicated (i.e., has nonvanishing higher de Rham cohomology) then there can be rather intricate ways in which the non-uniqueness of the solution to the equation $F = dA$ manifests itself. (That is, non-uniqueness can occur by more operations that naive ones that are available on contractible domains.) However, it appears that A is more fundamental than F, and so "solving for A given F" may be physically unsound. One last point worth noting is that if we take the 1-form J and the 1-form A as the primary objects of study (as seems to be the case in physics) then the only interesting Maxwell equation is $\delta dA = J$ yet this turns out to be exactly the Euler-Lagrange equation arising from the action principle. Hence, in a sense the Maxwell equations are consequences of more fundamental physical principles applied to the current form and potential form, coupled with mathematical trivialities such as $d \circ d = 0$. Lemma 2.111 provides the link between these abstractions and the classical formulation of the theory (with oriented S).

3. Metric Connections and Geodesics.

3.1. **Linear connections on vector bundles.** This section is done in full generality. The idea of connections is to differentiate a vector field in the direction given by another vector field, in a way compatible with the Riemannian metric.

Definition 3.1. Let $E \to M$ be a vector bundle. A **connection** (also called a **covariant derivative**) on the space $\Gamma(E)$ of global sections is a bilinear map

$$D\colon \Gamma(TM) \times \Gamma(E) \to \Gamma(E),$$
$$(X, s) \mapsto D_X s,$$

such that for any smooth function f

- $D_{fX}s = f D_X s$,
- $D_X(fs) = X(f)s + f D_X s$.

$\diamond$

Proposition 3.2. *For a point $p \in M$ the value $(D_X s)(p)$ depends only on $X(p)$, not on X.*

Proof. $D_x s$ is $\mathscr{C}^\infty(M)$-linear with respect to X. It is enough to show that if $X(p) = 0$, then $(D_X s)(p) = 0$.

Let $U \ni p$ be an open neighborhood of p giving a local chart. Let f be a function with support contained in U such that $f(p) = 1$. Then
$$(D_{fX}s)(p) = f(p)(D_X s)(p) = (D_X s)(p).$$
So we may assume that $X = 0$ outside of U.

Pick V such that $\overline{V} \subset U$ and $f \equiv 1$ on V and $f \equiv 0$ outside from U.

$$Xf = \sum f X^i \frac{\partial}{\partial x^i}.$$

We can replace X with Xf.

$$(D_X s)(p) = \sum f(p) X^i(p) \left(D_{\frac{\partial}{\partial x^i}} s\right)(p) = 0.$$

$\square$

Example 3.3. On $\mathbf{R}^n$ the tangent space is $T\mathbf{R}^n \cong \mathbf{R}^n \times \mathbf{R}^n$, and a vector field is just a function $X\colon \mathbf{R}^n \to \mathbf{R}^n$ (we should write $X(p) = (p, X_p) \in \mathbf{R}^n \times \mathbf{R}^n$, but we forget about the base point). Now $D_X Y = dY(X)$ satisfies the definition of being a connection on $\mathbf{R}^n$. $\diamond$

Example 3.4. Let $S \subset \mathbf{R}^{d+n}$ be a submanifold of dimension d. For $X \in \Gamma(TS)$ and $p \in S$ we say that $\widetilde{X} \in \Gamma(T\mathbf{R}^{d+n})$ is a **local extension** of X around p if there exists a neighborhood $U \ni p$ in $\mathbf{R}^{d+n}$ such that $\widetilde{X}\big|_{U\cap S} = X|_{U \cap S}$. Any $X \in \Gamma(TS)$ has a local extension at any point $p \in S$.

For $X, Y \in \Gamma(TS)$ and $p \in S$ we define

$$(D_X Y)(p) = (\underbrace{D_{\widetilde{X}} \widetilde{Y}}_{=d\widetilde{Y}(\widetilde{X})})^\top(p),$$

where $\widetilde{X}$ and $\widetilde{Y}$ are local extensions of X and Y around p, and $\top$ denotes the orthogonal projection $\mathbf{R}^{d+n} \to T_p S$.

We get

$$D\colon \Gamma(TS) \times \Gamma(TS) \to \Gamma(TS).$$

We have to check that $(D_{\widetilde{X}} \widetilde{Y})^\top$ does not depend on local extensions $\widetilde{X}$ and $\widetilde{Y}$, only on X and Y.

Choose $\mathbf{R}^{d+s} \supset U \xrightarrow{\phi} V \subset \mathbf{R}^{d+n}$ such that $\phi(S \cap U) = \mathbf{R}^d \times \{0\} \cap V$. $\phi = (x^1, \ldots, x^{d+n})$. We note that for a chart on S we may assume $\phi|_{U\cap S} = (x^1, \ldots, x^d, 0, \ldots, 0)$.

$$(D_{\widetilde{X}} \widetilde{Y})(p) = \left(D_{\sum_{1\le i\le d+n} \widetilde{X}^i \frac{\partial}{\partial x^i}} \sum_j \widetilde{Y}^j \frac{\partial}{\partial x^j}\right)(p) = \sum_{1\le i,j\le d+n} \widetilde{X}_i(p) \left(D_{\frac{\partial}{\partial x^i}} \widetilde{Y}^j \frac{\partial}{\partial x^j}\right)(p).$$

Now $\widetilde{X}(p) = X(p) = \sum_{1 \le i \le d} X^i(p) \frac{\partial}{\partial x^i}$.

$$(D_{\widetilde{X}}\widetilde{Y})(p) = \sum_{\substack{1 \le i \le d \\ 1 \le j \le d+n}} X^i(p) \left(D_{\frac{\partial}{\partial x^i}} \widetilde{Y}^j \frac{\partial}{\partial x^j} + \widetilde{Y}^j D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right).$$

For some curve c, such that $[c] = \frac{\partial}{\partial x^i}$

$$\frac{\partial}{\partial x^i} \widetilde{Y}^j = d\widetilde{Y}^j \left(\frac{\partial}{\partial x^i} \right) = \frac{d}{dt} \, (\widetilde{Y}^j \cdot c(t)) \Big|_{t=0}.$$

Now $\frac{\partial}{\partial x^i}$ is tangent to S (lies in T_pS). We choose $c \colon T \to S$, $c(0) = p$, $c'(0) = \frac{\partial}{\partial x^i}$.

$$\widetilde{Y}^j \cdot c(t) = \begin{cases} Y^j \cdot c(t), & j \le d, \\ 0, & j > d. \end{cases}$$

$$\frac{\partial}{\partial x^i} \widetilde{Y}^j = \begin{cases} \frac{\partial}{\partial x^i} Y^j, & j \le d, \\ 0, & j > d. \end{cases}$$

(On $S \cap U$ we have $\widetilde{Y} = Y$.)

$$(D_{\widetilde{X}}\widetilde{Y})(p) = \sum_{\substack{1 \le i \le d \\ 1 \le j \le d}} X^i(p) \left(\frac{\partial}{\partial x^i} Y^j \frac{\partial}{\partial x^j} + Y^j D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right).$$

So this indeed does not depend on $\widetilde{X}$ and $\widetilde{Y}$, only on X and Y. ⋄

Example 3.5. Let $S \subset \mathbf{R}^{d+n}$ be a submanifold of dimension d. Consider the normal bundle

$$N(S) := \coprod_{p \in S} N_p(S),$$

where $N_p(S) := (T_pS)^\perp \subset \mathbf{R}^{d+n}$. In fact $N(S)$ is a vector bundle of rank n.

We define

$$\begin{aligned} D \colon \Gamma(TS) \times \Gamma(NS) &\to \Gamma(NS), \\ (X, s) &\mapsto D_X s, \end{aligned}$$

where $D_X s := (\underbrace{D_{\widetilde{X}}\widetilde{s}}_{=d\widetilde{s}(\widetilde{x})})^\perp$, where $\widetilde{X}, \widetilde{s}$ is a local extension of X, s, and $\perp$ is the orthogonal projection to NS.

This definition is similar to the definition of connection on the tangent bundle. ⋄

3.2. **Levi-Civita connection.** We are interested in a specific connection that is compatible with a given Riemannian structure.

Theorem 3.6 (Fundamental Theorem of Riemannian Geometry). *Let (M, g) be a Riemannian manifold. Then there exists a unique connection*

$$\begin{aligned} \Gamma(TM) \times \Gamma(TM) &\to \Gamma(TM), \\ (X, Y) &\mapsto D_X Y. \end{aligned}$$

such that

(1) the connection is "torsion free" (or "symmetric"), meaning that

$$D_X Y - D_Y X = [X, Y] \quad \text{for all } X, Y \in \Gamma(TM),$$

(2) it preserves the metric,

$$X \cdot g(Y, Z) = g(D_X Y, Z) + g(Y, D_X Z) \quad \text{for all } X, Y, Z \in \Gamma(TM).$$

Here $X \cdot g(Y, Z)$ denotes the derivative of $g(Y, Z)$ in the direction X.

In particular, (1) applied to $X = \frac{\partial}{\partial x^i}$ and $Y = \frac{\partial}{\partial x^j}$ gives a formula

$$D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = D_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i},$$

which is useful in computations.

Definition 3.7. This connection is called the **Levi Civita connection** (or the **canonical connection**) on (M, g). ⋄

Proof. Assuming D exists, it must satisfy the relations (2):

$$X \cdot g(Y,Z) = g(D_X Y, Z) + g(Y, D_X Z), \tag{5}$$

$$Y \cdot g(Z,X) = g(D_Y Z, X) + g(Z, D_Y X), \tag{6}$$

$$Z \cdot g(X,Y) = g(D_Z X, Y) + g(X, D_Z Y). \tag{7}$$

Now we have (using the fact that $g(\cdot,\cdot)$ is bilinear and symmetric) for (5) + (7) - (7):

$$X \cdot g(Y,Z) + Y \cdot g(Z,X) - Z \cdot g(X,Y) = g(D_X Y + D_Y X, Z) + g(\underbrace{D_Y Z - D_Z Y}_{[Y,Z]}, X) + g(\underbrace{D_X Z - D_Z X}_{[X,Z]}, Y).$$

Using relation (1), we have

$$\begin{aligned} D_X Y + D_Y X &= 2\, D_X Y - [X,Y], \\ D_Y Z - D_Z Y &= [Y,Z], \\ D_X Z - D_Z X &= [X,Z], \end{aligned}$$

hence

$$X \cdot g(Y,Z) + Y \cdot g(Z,X) - Z \cdot g(X,Y) = 2\, g(D_X Y, Z) - g([X,Y],Z) + g([Y,Z],X) + g([X,Z],Y).$$

This gives an expression

$$g(D_X Y, Z) = \frac{1}{2}\left(X \cdot g(Y,Z) + Y \cdot g(Z,X) - Z \cdot g(X,Y) + g([X,Y],Z) - g([Y,Z],X) - g([X,Z],Y)\right).$$

Since Z is arbitrary, the latter defines $D_X Y$ uniquely. Conversely, $D_X Y$ satisfying the last expression is a Levi Civita connection, i.e. satisfies (1) and (2). □

Example 3.8. For $\mathbf{R}^n$ (with respect to the usual scalar product $\langle \cdot, \cdot \rangle$ as a Riemannian metric) the Levi Civita connection is given by $D_X Y = dY(X)$.

We check the property (1):

$$\begin{aligned} D_X Y - D_Y X &= dY(X) - dX(Y) \\ &= d(Y^1, \ldots, Y^n)(X) - d(X^1, \ldots, X^n)(Y) \\ &= (dY^1(X), \ldots, dY^n(X)) - (dX^1(Y), \ldots, dX^n(Y)) \\ &= \sum_i \left(dY^i(X) - dX^i(Y)\right) \frac{\partial}{\partial x^i} \\ &= \sum_i \left(X \cdot Y^i - Y \cdot X^i\right) \frac{\partial}{\partial x^i} = [X,Y]. \end{aligned}$$

And we check the property (2):

$$\begin{aligned} X \cdot \langle Y, Z \rangle &= X \cdot \sum Y^i Z^i \\ &= \sum_i \underbrace{(X \cdot Y^i) Z^i}_{dY^i(X)} + \underbrace{(X \cdot Z^i) \cdot Y^i}_{dZ^i(X)} \\ &= \langle dy(x), z \rangle + \langle dz(X), y \rangle \\ &= \langle D_X Y, Z \rangle + \langle Y, D_X Z \rangle. \end{aligned}$$

⋄

Example 3.9. Let $S \subset \mathbf{R}^{d+n}$ be a submanifold of dimension d. Let g be the metric on S induced by the usual scalar product on $\mathbf{R}^n$:

$$g_p(u,v) := \langle u, v\rangle\,.$$

Let $D_X Y := (D_{\widetilde{X}}\widetilde{Y})^\top$ where $\top$ is the orthogonal projection to TS (see Example 3.4 above).

This is the canonical connection on (S,g). To verify this choose (U,f) with $U \ni p$ an open neighborhood in $\mathbf{R}^{d+n}$ and $f = (x^1,\dots,x^{d+n})\colon U \to V \subset \mathbf{R}^{d+n}$ is a diffeomorphism such that $f(U\cap S) = V \cap \mathbf{R}^d \times \{0\}$.

Then $\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^d}$ restricted to $S\cap U$ are in TS, and $\frac{\partial}{\partial x^{d+1}},\dots,\frac{\partial}{\partial x^{d+n}}$ are transverse to TS (but not necessarily orthogonal).

Let $\widetilde{X} = \sum_{1\le i\le d+n}\widetilde{X}^i\frac{\partial}{\partial x^i}$ and $\widetilde{Y} = \sum_{1\le i\le d+n}\widetilde{Y}^i\frac{\partial}{\partial x^i}$.

$$D_X Y - D_Y Z = (D_{\widetilde{X}}\widetilde{Y} - D_{\widetilde{Y}}\widetilde{X})^\top = [\widetilde{X},\widetilde{Y}]^\top.$$

We have $[\widetilde{X},\widetilde{Y}] = [X,Y]$ on $S\cap U$ (cf. Example 3.4), so the first property of the Levi Civita connection is satisfied.

For the second property, let $X,Y,Z\in\Gamma(TS)$ and let $\widetilde{X},\widetilde{Y},\widetilde{Z}$ be local extensions. On S we have

$$X\cdot g(Y,Z) = \widetilde{X}\cdot\left\langle\widetilde{Y},\widetilde{Z}\right\rangle = \left\langle D_{\widetilde{X}}\widetilde{Y},\widetilde{Z}\right\rangle + \left\langle\widetilde{Y},D_{\widetilde{X}},\widetilde{Z}\right\rangle = \left\langle(D_{\widetilde{X}}\widetilde{Y})^\top,\widetilde{Z}\right\rangle + \left\langle\widetilde{X},(D_{\widetilde{X}}\widetilde{Z})^\top\right\rangle = \langle D_X Y, Z\rangle + \langle X, D_X Z\rangle\,.$$

⋄

3.3. **Christoffel symbols.** We can express the Levi Civita connection D in terms of local coordinates. Let $p\in M$ and let (U,ϕ) be a chart on M around p. Here $\phi = (x^1,\dots,x^d) : U \to V \subset \mathbf{R}^d$. Let $g(\cdot,\cdot)$ denote the Riemannian metric. We define functions $\Gamma_{ij}^k : U \to \mathbf{R}$ for $(1\le i,j,k\le d)$, called **Christoffel symbols**, by

$$D_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \sum_{1\le k\le d}\Gamma_{ij}^k\frac{\partial}{\partial x^k}.$$

If in local coordinates $X = \sum X^i\frac{\partial}{\partial x^i}$ and $Y = \sum Y^j\frac{\partial}{\partial x^j}$, then we calculate using the linearity

$$\begin{aligned}
D_X Y &= D_{\sum_i X^i\frac{\partial}{\partial x^i}}\sum_j Y^j\frac{\partial}{\partial x^j}\\
&= \sum_i\sum_j X^i D_{\frac{\partial}{\partial x^i}}\left(Y^j\frac{\partial}{\partial x^j}\right)\\
&= \sum_{i,j} X^i\left(D_{\frac{\partial}{\partial x^i}}Y^j\frac{\partial}{\partial x^j} + Y^j D_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j}\right)\\
&= \sum_{i,k} X^i D_{\frac{\partial}{\partial x^i}}Y^k\frac{\partial}{\partial x^k} + \sum_{i,j,k}X^i Y^j\Gamma_{ij}^k\frac{\partial}{\partial x^k}\\
&= \sum_k\left(\sum_i\left(X^i D_{\frac{\partial}{\partial x^i}}Y^k + \sum_j X^i Y^j\Gamma_{ij}^k\right)\right)\frac{\partial}{\partial x^k}.
\end{aligned}$$

(CHECK THIS???) The symbols Γ_{ij}^k determine the connection.

One can compute Γ_{ij}^k in terms of the $d\times d$ matrix $(g_{ij})_{i,j} = (g(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}))_{i,j}$, derivatives of g_{ij} and the inverse matrix $(g_{ij})^{-1}$

Example 3.10. In polar coordinates (r,θ) one has $g = dr^2 + f^2(r)\,d\theta$. We get

$$g\left(D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right) = \frac{1}{2}\frac{\partial}{\partial r}\underbrace{g\left(\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right)}_{=1\text{ constant}} = 0.$$

And so $\Gamma^1_{11} = 0$.

$$g\left(D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r},\frac{\partial}{\partial\theta}\right) = \frac{\partial}{\partial r}\underbrace{g\left(\frac{\partial}{\partial r},\frac{\partial}{\partial\theta}\right)}_{=0} - g\left(\frac{\partial}{\partial r},\underbrace{D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta}}_{=0}\right) = -\frac{1}{2}\frac{\partial}{\partial\theta}g\left(\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right) = 0.$$

We have $\Gamma^2_{11} = 0$.

$D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r} = 0$.

$$g\left(D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right) = \frac{1}{2}\frac{\partial}{\partial r}\underbrace{g\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right)}_{=f^2(r)} = f'f.$$

$$D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta} = \Gamma^1_{12}\frac{\partial}{\partial r} + \Gamma^2_{12}\frac{\partial}{\partial\theta}.$$

$$g\left(D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right) = \Gamma^2_{12}\underbrace{g\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right)}_{=f^2}$$

Hence we have $\Gamma^2_{12} = f'/f$.

TODO: finish and check that $D_{\frac{\partial}{\partial r}}\frac{\partial}{\partial\theta} = \frac{f'}{f}\frac{\partial}{\partial\theta}$ and $D_{\frac{\partial}{\partial\theta}}\frac{\partial}{\partial\theta} = -f'f\frac{\partial}{\partial r}$. ◇

3.4. **Covariant derivatives of arbitrary tensor fields.** Let $E \to M$ and $F \to M$ be vector bundles over M, with covariant derivative operators ∇, and $\widetilde{\nabla}$, respectively. The covariant derivative operators on $E \otimes F$ and $\hom(E,F)$ are

$$\nabla_X(s\otimes\widetilde{s}) = (\nabla_X s)\otimes\widetilde{s} + s\otimes(\widetilde{\nabla}_X\widetilde{s}) \tag{8}$$

$$(\nabla_X L)(s) = \widetilde{\nabla}_X(L(s)) - L(\nabla_X s), \tag{9}$$

for $s \in \Gamma(E)$, $\widetilde{s} \in \Gamma(F)$, and $L \in \Gamma(\hom(E,F))$. Note that the covariant derivative operator in $\bigwedge(E)$ is given by

$$\nabla_X(s_1\wedge\cdots\wedge s_r) = \sum_{i=1}^r s_1\wedge\cdots\wedge\nabla_X s_i\wedge\cdots\wedge s_r,$$

for $s_i \in \Gamma(E)$.

Now if we want to enforce the product rule of the covariant derivative for (r,s)-tensor we obtain the following:

Definition 3.11. Let $T \in \Gamma(T^s_r M)$, then the **covariant derivative** ∇T is an $(r,s+1)$ tensor given by

$$\nabla T(X,Y_1,\dots,Y_s) = \nabla_X(T(Y_1,\dots,Y_s)) - \sum_{i=1}^s T(Y_1,\dots,\nabla_X Y_i,\dots,Y_s).$$

◇

We now consider the above definitions in components for (r,s)-tensors. If $X \in \Gamma(TM)$ is a vector field, then ∇X is a $(1,1)$-type tensor field. By the definition of a connection, we have

$$\nabla_m X := \nabla_{\partial_m} X = \nabla_{\partial_m}(X^j\partial_j) = (\partial_m X^j)\partial_j + X^j\Gamma^l_{mj}\partial_l = (\nabla_m X^i + X^l\Gamma^i_{ml})\partial_i.$$

In other words,

$$\nabla X = \nabla_m X^i(dx^m\otimes\partial_i),$$

where

$$\nabla_m X^i = \partial_m X^i + X^l\Gamma^i_{ml}.$$

However, for a 1-form ω, the tensorial properties of the covariant derivative imply

$$\nabla\omega = (\nabla_m \omega_i) dx^m \otimes dx^i,$$

with

$$\nabla_m \omega_i = \partial_m \omega_i - \omega_l \Gamma^l_{im}.$$

The Leibniz product rule then implies for a general (r,s)-type tensor field S,

$$\nabla_m S^{i_1,\dots,i_r}_{j_1,\dots,j_s} = \partial_m S^{i_1,\dots,i_r}_{j_1,\dots,j_s} + S^{l,i_2,\dots,i_r}_{j_1,\dots,j_s}\Gamma^{i_1}_{ml} + \cdots + S^{i_1,\dots,i_{r-1},l}_{j_1,\dots,j_s}\Gamma^{i_r}_{ml} - S^{i_1,\dots,i_r}_{l,j_2,\dots,j_s}\Gamma^l_{mj_1} - \cdots - S^{i_1,\dots,i_r}_{j_1,\dots,j_{s-1},l}\Gamma^l_{mj_s}.$$

Notice the following computations:

$$(\nabla g)(X,Y,Z) = Xg(Y,Z) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0,$$

so the metric is parallel. Note that in coordinates this says

$$0 = \nabla_m g_{ij} = \partial_m g_{ij} - \Gamma^p_{mi} g_{pj} - \Gamma^p_{mj} g_{ip},$$

which yield the formula

$$\partial_k g_{ij} = \Gamma^p_{ki} g_{pj} + \Gamma^p_{kj} g_{ip}.$$

This is sometimes written as

$$\partial_k g_{ij} = [ki,j] + [kj,i],$$

where

$$[ij,k] = \frac{1}{2}\left(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\right).$$

These are often known as the Christoffel symbols of the first kind.

Definition 3.12. Let $T \in \Gamma(T^s_r M)$ be an (r,s)-type tensor field. The **second covariant derivative** of T is the $(r,s+2)$-type tensor field

$$\nabla^2 T = \nabla\nabla T.$$

$\diamond$

Now it is easiest to compute the second covariant derivative of a vector field, i.e. when T is a $(1,0)$-type tensor field. We simply have

$$(\nabla\nabla T)(X,Y) = \nabla(\nabla T)(X,Y) = \nabla_X(\nabla T(Y)) - \nabla T(\nabla_X Y) = \nabla_X \nabla_Y T - \nabla_{\nabla_X Y} T.$$

Proposition 3.13. *Let T be an (r,s)-type tensor field, then the double covariant derivative satisfies*

$$\nabla\nabla T(X,Y) = \nabla_X \nabla_Y T - \nabla_{\nabla_X Y} T.$$

Proof. Let T be an (r,s)-type tensor field, then the covariant derivative ∇T is the $(r,s+1)$-type tensor field is given by

$$\nabla T(X,\vec{Z}) = (\nabla_X T)(\vec{Z}) = \nabla_X(T(\vec{Z})) - \sum_{i=1}^{s} T(Z_1,\dots,\nabla_X Z_i,\dots,Z_s),$$

where $\vec{Z} = (Z_1,\dots,Z_s)$ is a s-tuple of vector fields on M.

We try and do the same computation for arbitrary (r,s)-type tensor fields that we did for vector fields.

$$(\nabla\nabla T)(X,Y,\vec{Z}) = \nabla_X(\nabla T(Y,\vec{Z})) - \nabla T(\nabla_X Y,\vec{Z}) - \sum_{i=1}^{s}\nabla T(Y,Z_1,\dots,\nabla_X Z_i,\dots,Z_s),$$

where $\vec{Z} = (Z_1, \dots, Z_s)$ is an s-tuple of vectors. We also compute

$$\begin{aligned}\nabla_X(\nabla_Y T)(\vec{Z}) - (\nabla_{\nabla_X Y} T)(\vec{Z}) &= \nabla_X(\nabla_Y T(\vec{Z})) - \sum_{i=1}^{s} (\nabla_Y T)(Z_1, \dots, \nabla_X Z_i, \dots, Z_s) \\ &\quad - \nabla_{\nabla_X Y}(T(\vec{Z})) + \sum_{i=1}^{s} T(Z_1, \dots, \nabla_{\nabla_X Y} Z_i, \dots, Z_s) \\ &= \nabla_X(\nabla T(Y, \vec{Z})) - \sum_{i=1}^{s} \nabla T(Y, Z_1, \dots, \nabla_X Z_i, \dots, Z_s) \\ &\quad - \nabla T(\nabla_X Y, \vec{Z}).\end{aligned}$$

So we have shown

$$(\nabla\nabla T)(X, Y, \vec{Z}) = \nabla_X(\nabla_Y T)(\vec{Z}) - (\nabla_{\nabla_X Y} T)(\vec{Z}).$$

□

Definition 3.14. The **Hessian** of T with respect to $X, Y \in \Gamma(TM)$ is the (r,s)-type tensor field $\nabla\nabla_{(X,Y)} T$ characterized by

$$(\nabla\nabla_{(X,Y)} T)(\vec{Z}) = \nabla\nabla T(X, Y, \vec{Z}).$$

Sometimes we will write

$$\nabla\nabla_{(X,Y)} T = \nabla^2 T(X, Y) = \nabla\nabla T(X, Y).$$

⋄

3.5. **Gradients, divergence, and integration by parts.**

Definition 3.15. Let (M, g) be a Riemannian manifold, and $f \in \mathscr{C}^\infty(M)$. The **gradient** of f is the vector field ∇f defined by

$$g(\nabla f, \cdot) = df.$$

More concisely, we can write

$$\nabla f = \sharp(df).$$

⋄

Note that $\nabla f = df$, where this ∇ is the covariant derivative. So we need to be very careful about this notation.

Definition 3.16. The **Hessian** of a smooth function $f \in \mathscr{C}^\infty(M)$ is the $(0,2)$-tensor defined by the double covariant derivative of f, i.e.

$$\nabla^2 f(X, Y) = \nabla\nabla f(X, Y) = \nabla_X \nabla_Y f - \nabla_{\nabla_X Y} f = XYf - (\nabla_X Y)f.$$

⋄

In components, we have

$$\nabla^2 f(\partial_i, \partial_j) = \nabla_i \nabla_j f = \partial_i \partial_j f - \Gamma_{ij}^k (\partial_k f).$$

Definition 3.17. The **Laplacian** of f is

$$\Delta f = \operatorname{tr}\big(X \mapsto \sharp\big(\nabla^2 f(X, \cdot)\big)\big).$$

⋄

In coordinates, we have

$$\Delta f = g^{ij} \nabla_i \nabla_j f.$$

Now we define the divergence of an arbitrary tensor field.

Definition 3.18. If T is an (r,s)-type tensor, then the **divergence** of T is the $(r, s-1)$-type tensor field

$$(\operatorname{div} T)(Y_1, \dots, Y_{s-1}) = \operatorname{tr}(X \mapsto \sharp(\nabla T)(X, \cdot, Y_1, \dots, Y_{s-1})),$$

that is, we trace the covariant derivative on the first two covariant indices. ⋄

Remark 3.19. Note that the divergence should really be called the divergence with respect to some index, and that there is no canonical choice of which indices we should be tracing over. It might seem like it is natural to trace over the first two indices, but this is really just an inconvenient result of the notation we use. ⋄

In coordinates, we have

$$(\operatorname{div} T)^{i_1,\dots,i_r}_{j_1,\dots,j_{s-1}} = g^{ij}\nabla_i T^{i_1,\dots,i_r}_{j,j_1,\dots,j_{s-1}}.$$

Using a local orthonormal frame $\{e_i\}_{i=1}^n$, the divergence can also be written as

$$(\operatorname{div} T)(Y_1,\dots,Y_{s-1}) = \sum_{i=1}^n (\nabla_{e_i} T)(e_i, Y_1, \dots, Y_{s-1}).$$

If X is a vector field, we define

$$(\operatorname{div} X) = \operatorname{tr}(\nabla X),$$

which in coordinate is

$$\operatorname{div} X = \delta^i_j \nabla_i X^j = \nabla_j X^j.$$

Proposition 3.20. *For a vector field X,*

$$\operatorname{div} X = \operatorname{div}(\flat X).$$

Proof.

$$\operatorname{div} X = \delta^i_j \nabla_i X^j = \delta^i_j \nabla_i g^{jl} X_l = \delta^i_j g^{jl} \nabla_i X_l = g^{il}\nabla_i X_l = \operatorname{div}(\flat X).$$

□

In a local orthonormal frame $\{e_i\}$ the divergence of the 1-form is given by

$$\operatorname{div}\omega = \sum_{i=1}^n (\nabla_{e_i}\omega)(e_i) = \sum_{i=1}^n e_i(\omega(e_i)) - \omega\left(\sum_{i=1}^n \nabla_{e_i} e_i\right),$$

whereas the divergence of a vector field is given by

$$\operatorname{div} X = \sum_{i=1}^n \langle \nabla_{e_i} X, e_i\rangle.$$

Now we build up to prove the divergence theorem on Riemannian manifolds.

Proposition 3.21. *Let $X \in \Gamma(TM)$. Then*

$$\star(\operatorname{div} X) = (\operatorname{div} X)\ \mathrm{vol}_g = d(i_X \mathrm{vol}_g) = \mathscr{L}_X \mathrm{vol}_g.$$

In coordinates, we have

$$\operatorname{div} X = \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial x^i}\left(X^i\sqrt{\det(g)}\right).$$

Proof. Fix a point $p \in M$ and let $\{e_i\}$ be an orthonormal basis of T_pM. In a small neighborhood of p, consider the geodesic frame e_i induces by parallel transportation of this frame along radial geodesics. In such a frame we clearly have

$$\nabla_{e_i} e_j = 0.$$

Let $\{\omega^i\}$ be the associated dual frame field. We then have

$$\begin{aligned}
\mathscr{L}_X \operatorname{vol}_g &= (d i_X + i_X d)\operatorname{vol}_g \\
&= d(i_X \operatorname{vol}_g) \\
&= \sum_i \omega^i \wedge \nabla_{e_i}(i_X(\omega^1 \wedge \cdots \wedge \omega^n)) \\
&= \sum_i \omega^i \wedge \nabla_{e_i}\left((-1)^{j-1}\sum_{j=1}^n \omega^j(X)\omega^1 \wedge \cdots \wedge \widehat{\omega}^j \wedge \cdots \wedge \omega^n\right) \\
&= \sum_{ij}(-1)^{j-1} e_i\left(\omega^j(X)\right)\omega^i \wedge \omega^1 \wedge \cdots \wedge \widehat{\omega}^j \wedge \cdots \wedge \omega^n \\
&= \sum_i \omega^i\left(\nabla_{e_i}X\right)\operatorname{vol}_g \\
&= (\operatorname{div} X)\operatorname{vol}_g \\
&= \star(\operatorname{div} X).
\end{aligned}$$

Applying the Hodge star once more to this formula gives us the expression in local coordinates:

$$\begin{aligned}
\operatorname{div} X &= \star d(i_X \operatorname{vol}_g) \\
&= \star d\left(i_X\left(\sqrt{\det(g)}dx^1 \wedge \cdots \wedge dx^n\right)\right) \\
&= \star d\left(\sum_{j=1}^n (-1)^{j-1}X^j\sqrt{\det(g)}dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n\right) \\
&= \star\left(\partial_i\left(X^i\sqrt{\det(g)}\right)dx^1 \wedge \cdots \wedge dx^n\right) \\
&= \star\left(\partial_i\left(X^i\sqrt{\det(g)}\right)\frac{1}{\sqrt{\det(g)}}\operatorname{vol}_g\right) \\
&= \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial x^i}\left(X^i\sqrt{\det(g)}\right).
\end{aligned}$$

□

Theorem 3.22 (Divergence theorem.). *Let (M,g) be a compact, orientable, manifold with boundary. If $X \in \Gamma(TM)$, and $f \in \mathscr{C}^1(M)$ then*

$$\int_M (\operatorname{div} X)f \ \operatorname{vol}_g = -\int_M df(X)\ \operatorname{vol}_g + \int_{\partial M}\langle X, \boldsymbol{n}\rangle f \ d\sigma,$$

where $\boldsymbol{n}$ is the outer unit normal, and $d\sigma$ is the induced volume form on ∂M.

If $\omega\Omega^1(M)$, then

$$\int_M (\operatorname{div}\omega)f \ \operatorname{vol}_g = -\int_M \langle \omega, df\rangle \ \operatorname{vol}_g + \int_{\partial M}\omega(\boldsymbol{n})f \ d\sigma.$$

If $u, v \in \mathscr{C}^\infty(M)$, then

$$\int_M v\Delta u \ \operatorname{vol}_g = -\int_M \langle \nabla u, \nabla v\rangle_g \ \operatorname{vol}_g + \int_{\partial M}\langle \nabla u, \boldsymbol{n}\rangle v \ d\sigma,$$

and

$$\int_M (v\Delta u - u\Delta v)\operatorname{vol}_g = \int_{\partial M}(\langle \nabla u, \boldsymbol{n}\rangle v - \langle \nabla v, \boldsymbol{n}\rangle u)\, d\sigma.$$

Proof. We compute

$$d(f i_X \operatorname{vol}_g) = df \wedge (i_X \operatorname{vol}_g) + f\, d(i_X \operatorname{vol}_g).$$

Now using Stokes theorem and Proposition 3.21, we find

$$\int_M (\operatorname{div} X) f \ \operatorname{vol}_g + \int_M df \wedge (i_X \operatorname{vol}_g) = \int_{\partial M} f \ i_X \operatorname{vol}_g .$$

A similar computation shows

$$df \wedge (i_X \operatorname{vol}_g) = df(X) \operatorname{vol}_g .$$

Now on ∂M we can decompose $X = X^\top + X^\perp$, where $X^\top$ is tangential to ∂M and $X^\perp$ is normal to ∂M. Then

$$i_X \operatorname{vol}_g = \operatorname{vol}_g \left(X^\top + X^\perp, \cdots\right) = \operatorname{vol}_g \left(\langle X, \boldsymbol{n}\rangle \boldsymbol{n}, \cdots\right) = \langle X, \boldsymbol{n}\rangle \ d\sigma,$$

since the volume form on ∂M is given by $d\sigma = i_{\boldsymbol{n}} \operatorname{vol}_g$. This gives us the first equality.

The divergence theorem for 1-forms is dual, but identical to the previous argument. The first Green formula follows by taking $\Delta u = \operatorname{div}(\nabla u)$, and the second follows immediately from the first. □

Remark 3.23. Note that the above integration by parts formula gives us another way to derive the coordinate expression of the divergence: Fix a local coordinate system on M, and assume that X and f have compact support in these coordiantes. Then

$$\begin{aligned}
\int_M (\operatorname{div} X) f \ \operatorname{vol}_g &= -\int_M df(X) \ \operatorname{vol}_g \\
&= -\int_M \partial_i f \, dx^i (X^j \partial_j) \sqrt{\det(g)} \ d\boldsymbol{x} \\
&= -\int_{\mathbf{R}^n} \partial_i f X^i \sqrt{\det(g)} \ d\boldsymbol{x} \\
&= \int_{\mathbf{R}^n} f \partial_i \left(X^i \sqrt{\det(g)}\right) d\boldsymbol{x} \\
&= \int_M f \frac{1}{\sqrt{\det(g)}} \partial_i \left(X^i \sqrt{\det(g)}\right) \operatorname{vol}_g .
\end{aligned}$$

Since this holds for any f, we deduce that

$$\operatorname{div} X = \frac{1}{\sqrt{\det(g)}} \partial_i \left(X^i \sqrt{\det(g)}\right).$$

◇

Corollary 3.24. *In local coordinates, the Laplace-Beltrami operator of a function is given by*

$$\Delta f = \frac{1}{\sqrt{\det(g)}} \partial_i \left(g^{ij} \partial_j f \sqrt{\det(g)}\right).$$

Proof. This follows since $\Delta f = \operatorname{div}(\nabla f)$, and so we can take $X^i = g^{ij} \partial_j f$ in the above coordinate expression of the divergence of X. □

Now we prove integration by parts for (r,s)-type tensor fields.

Theorem 3.25. *Let (M, g) be a compact manifold without boundary. Let $T \in \Gamma(T^s_r M)$ be an (r,s)-type tensor field, and $S \in \Gamma(T^{s+1}_r M)$. Then*

$$\int_M \langle \nabla T, S\rangle \ \operatorname{vol}_g = -\int_M \langle T, \operatorname{div} S\rangle \ \operatorname{vol}_g .$$

Proof. We consider the inner product $\langle T, S\rangle$ as a 1-form ω. In local coordinates, we have

$$\omega = \langle T, S\rangle = T^{j_1,\dots,j_s}_{i_1,\dots,i_r} S^{i_1,\dots,i_r}_{j,j_1,\dots,j_s} dx^j .$$

Note that the indices on T are reversed since we are taking an inner product. Taking the divergence, and noting that g is parallel, we compute

$$\begin{aligned}\operatorname{div}(\langle T,S\rangle) &= \nabla^j\left(T^{j_1,\dots,j_s}_{i_1,\dots,i_r} S^{i_1,\dots,i_r}_{j,j_1,\dots,j_s}\right)\\ &= \nabla^j\left(T^{j_1,\dots,j_s}_{i_1,\dots,i_r}\right) S^{i_1,\dots,i_r}_{j,j_1,\dots,j_s} + T^{j_1,\dots,j_s}_{i_1,\dots,i_r} \nabla^j S^{i_1,\dots,i_r}_{j,j_1,\dots,j_s}\\ &= \langle \nabla T, S\rangle + \langle T, \operatorname{div} S\rangle.\end{aligned}$$

The theorem now follows since $\operatorname{div}(X) = \operatorname{div}(\flat X)$ for a vector field X and the usual integration by parts theorem for vector fields. □

Remark 3.26. Often texts consider operator $\nabla^* = -\operatorname{div}$. Then,

$$\int_M \langle \nabla T, S\rangle \ \mathrm{vol}_g = \int_M \langle T, \nabla^* S\rangle \ \mathrm{vol}_g,$$

so we see that ∇^* is the (formal) L^2-adjoint of ∇. ⋄