Fast-track Differential Topology

Ruoceng Xu

24th. June 2023

Special thanks to my friend Sidharth Hariharan (his website with contact) for his contribution to some algebra parts (mainly in Chapter 4) and the formatting of this work. Originally a collaborative endeavour, Sid was regrettably redirected to other pressing commitments after February 2024.

Contents

1 Apéritif of topology

In this section, we will introduce 4 of the most interesting topics on topology that don't require much background knowledge. They are mainly about a concept of fundamental group, which will not be a main topic of discussion afterwards; and so,
They are NOT prerequisites for other sections, but it is recommended to have knowledge of the definitions on this page. (Definition 1.1). However it is a good chance to get an idea of what knowledge our work would base on.

Definition 1.1. Basic terms & notations in topology:

1. **(Path)** For $p, q \in X$, a path from p to q is a continuous map $\alpha : [0, 1] \to X$ such that $\alpha(0) = p$ and $\alpha(1) = q$.
2. **(Loop)** For $p \in X$, A loop based at p is a path from p to itself.
3. **(Product of paths)** For paths α and β such that $\alpha(1) = \beta(0)$, we define the product $(\alpha \cdot \beta)$ to be
$$(\alpha \cdot \beta)(s) := \begin{cases} \alpha(2s) & \text{if } 0 \le s \le \frac{1}{2} \\ \beta(2s-1) & \text{if } \frac{1}{2} \le s \le 1 \end{cases} \tag{1.1}$$
4. **(Homotopy)** Maps $f_1, f_2 \in C^0(X, Y)$ are homotopic (also written as $f_1 \approx f_2$), if there exists continuous $F : [0, 1] \times X \to Y$ such that $F|_{(0,X)} = f_1$ and $F_{(1,X)} = f_2$, and such F is called a homotopy.
5. **(Homotopy class)** For points $a, b \in X$, two paths from a to b are in the same homotopy class iff they are homotopic.
6. **(Homotopy equivalent)** $X \simeq Y$ denotes homotopy equivalent: which means there exists $f \in C^0(X, Y)$ and $g \in C^0(Y, X)$ such that
$$(f \circ g) \approx \mathrm{id}_Y \quad (g \circ f) \approx \mathrm{id}_X$$
7. **(Embedding)**: $f : X \hookrightarrow Y$ is an embedding if it's an injective homeomorphism, where its image equips the subspace topology in Y.
8. **(Quotient space)** Consider an equivalent relation $\sim$ on X, the $X/\sim$ is a topological space where $U \subset X$ is open iff $U/\sim$ is open.
9. **(Adjunction space)**
10. **(Wedge sum)** Suppose $x \in X$ and $y \in Y$, the wedge sum $X \vee Y$ by identifying x and y is the quotient space $X \sqcup Y/\sim$ where $z_1 \sim z_2$ iff $z_1, z_1 \in \{x, y\}$.

1.1 Galois correspondence in covering space

[**Convention**] In this part,

1. PC stands for path-connected
2. X is a Hausdorff and PC topological space.
3. In a point p in topological space M, N_p always represent a neighbour of p.

An interesting question that we have about our space is the definition and counting of "holes"; and the **fundamental group** is an intuitive answer.

Definition 1.2. Fix a point $x_0 \in X$, the fundamental group is defined as

$$\pi_1(X, x_0) = \{\text{loop } \gamma \text{ based at } x_0\} / \approx \tag{1.2}$$

under product of paths. We write elements in $\pi_1(X, x_0)$ as $[\gamma]$ to represent the homotopy classes of loop γ.

We can argue that in our case (X being PC), the choice of x_0 does not really matter; in other words, $\pi_1(X, x_0) \cong \pi_1(X, x_1)$: Let α be a path from x_0 to x_1, then $[\gamma] \mapsto [\alpha^{-1}\gamma\alpha] \in \pi_1(X, x_1)$ builds the isomorphism. $\square$
So there is only one fundamental group in X, and we can just write it to be $\pi_1(X)$.

Remark. Why is fundamental group intuitively tells the hole? For example, space X is simply-connected iff $\pi_1(X) \cong \{0\}$; similarly $\pi_1(\mathbb{S}^1) \cong \mathbb{Z}$, $\pi_1(\mathbb{S}^1 \vee \mathbb{S}^1) \cong \mathbb{Z} * \mathbb{Z}$ and $\pi_1(\mathbb{R}^2 \setminus \{0\}) \cong \mathbb{Z}$ corresponding to the number of "holes" they have. However, the π_1 does not capture all "holes"; for example, $\mathbb{R}^3 \setminus \{0\}$ has a "hole", but it has trivial π_1.

However *in this part, we don't really care about holes*, what we care about is the strange correspondence between π_1 and the **Deck transformation** of **covering space**.
To convince you that this is interesting, let's look at the definitions of them.

A covering space of X is a way to express X space using simpler spaces.

Definition 1.3. A topological space $\widetilde{X}$ is called a **Covering space** of X if there exists a <u>continuous</u> map $p : \widetilde{X} \to X$ such that for $\forall x \in X$, $\exists$neighbour N_x such that

$$p^{-1}(N_x) = \bigsqcup_{\alpha \in p^{-1}(x)} N_\alpha$$

where each N_α is a neighbour of α in $\widetilde{X}$.

And a deck transformation is an "isomorphism" of covering spaces,

Definition 1.4. For $x_0 \in X$ and covering spaces $p_1 : \widetilde{X}_1 \to X$ & $p_2 : \widetilde{X}_2 \to X$,

- A homeomorphism $f : \widetilde{X}_1 \to \widetilde{X}_2$ is called an **isomorphism** if $p_2 \circ f = p_1$.
- A subset of isomorphisms is said to be **base preserving** (BP) if $\exists \tilde{x}_1 \in p_1^{-1}(x_0)$ and $\tilde{x}_2 \in p_2^{-1}(x_0)$ such that $f(\tilde{x}_1) = \tilde{x}_2$ for $\forall f$ in it.
- For covering space $\widetilde{X}$, **Deck transformation** is an isomorphism to itself, we use $\mathrm{G}(\widetilde{X}, X)$ to represent the group of deck transformations.

Interestingly, as we said, each subgroup of $\pi_1(X, x_0)$ naturally corresponds to a covering space of X up to a set of BP-isomorphisms. i.e *There exists a bijection:*

$$\{\text{PC covering spaces of } X\} \Big/ \text{BP isomorphisms} \longleftrightarrow \{\text{subgroups of } \pi_1(X)\} \tag{1.3}$$

It builds an intriguing connection between the Symmetry of a space and a "hole-counting" Topological invariant. And further, we can even get

Galois correspondence of Deck transformation

Theorem 1.1. Suppose X is LPC and $\widetilde{X}$ is PC. Let $H = p(\pi_1(\widetilde{X}))$, then

$$\mathrm{G}(\widetilde{X}, X) \cong N(H)/H$$

where $N(H)$ is the normalizer of H in $\pi_1(X)$.

(where LPC means locally path-connected which we will define later.)
This is very similar to the **Galois correspondence** we have in field extensions. We will briefly talk about the underlying intuition for such an interesting analogy.
But first, let's try to prove them.

First, we introduce two background definitions:

Definition 1.1. For a topological space M, it is

- **Locally path-connected (LPC)** if for $\forall p \in M$, $\exists$neighbour N such that N is PC.
- **Semi-locally simply-connected (SLSC)** if for $\forall p \in M$, $\exists$neighbour N with natural inclusion $\iota : N \hookrightarrow X$ such that $\iota(\pi_1(N)) = \{0\}$.

The most fundamental idea in this part is that: *Each point in a covering space can be represented by a homotopy class of paths.*
This implies that, by fixing a point $x_0 \in X$, there exists bijection

$$\begin{array}{rccc} \{\text{path } \gamma \text{ in } X \text{ starting at } x_0\} / \approx & \longleftrightarrow & \widetilde{X} \\ \text{by sending} \quad [\gamma] & \mapsto & \widetilde{x} \in p^{-1}(\gamma(1)) \end{array}$$

This equivalence can be proved by something called the **Homotopy lifting property**; I will put it into the exercise, because here we only need its intuition, rather than its statement. The most direct observation of this idea is through this:

Proposition 1.1. Suppose X is PC, LPC and SLSC, there exists covering space $p : \widetilde{X} \to X$ such that $\pi_1(\widetilde{X}) = \{0\}$.

As we said, the thinking in this proof is basically to construct a topological space via homotopy classes of paths.

Proof. Fix a point $x_0 \in X$. Since X is LPC and SLSC, for each point $x \in X$, there exists PC neighbours U_x such that $\pi(U_x) = \{0\}$. Easy to verify that $\mathcal{U} = \{U_x : x \in X\}$ form a topological basis for X.
Suppose γ is path from x_0 to x. Define

$$U_x^{[\gamma]} := \{[\gamma \cdot \alpha] : \alpha(0) = \gamma(1) \text{ and } \alpha(1) \in U_x\}$$

and equip them with subspace topology of X. We can observe and show that:

1. For every $[\gamma]$, we have $U_x^{[\gamma]} \cong U_x$.
2. If $f \not\approx g$, then $U_x^{[f]} \cap U_x^{[g]} = \emptyset$.
3. $\widetilde{X} := \bigcup_x \bigcup_\gamma U_x^{[\gamma]}$ form a topological space.

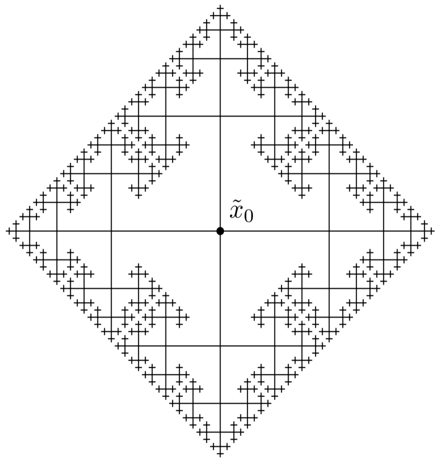
Now if we define $p : \widetilde{X} \to X$ to be $p([\gamma \cdot \alpha]) = \alpha(1)$; then by the claims (2) & (3) above, each $p|_{U_x^{[\gamma]}}$ is a homeomorphism. Hence $\widetilde{X}$ is a covering space.
Clearly $\widetilde{X}$ is PC since all its elements correspond to paths in X; now we need to show that it's simply connected.
Consider a loop γ in $\widetilde{X}$, then $p(\gamma)$ is a loop in X. Now because $\mathrm{Im}(\gamma)$ can be covered by finitely many $U_x^{[\gamma]}$ (since $[0,1]$ is compact), there exists a continuous $\phi_\bullet : [0,1] \to \widetilde{X}$ such that ϕ_t is a path from x_0 to $(\gamma(t))\,(1)$ for $\forall t \in [0,1]$.
Therefore, the map

$$\begin{aligned} [0,1] \times [0,1] &\to X \\ (s,t) &\mapsto p\left(\phi_t(s)\right) \end{aligned}$$

defines a continuous contraction from $p(\gamma)$ to x_0; hence $p(\gamma)$ is always a trivial loop. So $p\left(\pi_1(\widetilde{X})\right) \cong \{0\}$, which means $\pi_1(\widetilde{X}) \cong \{0\}$ as a subgroup of $p\left(\pi_1(\widetilde{X})\right)$. □

Very abstract, but in practice, constructing such covering space is rather down-to-earth, let's take some non-trivial examples:

Example 1.2. Consider $X = \mathbb{S}^1 \vee \mathbb{S}^1$ and point $\widetilde{x}_0 \in X$, by tracing through the path, we have that $\widetilde{X}$ can be the Cayley graph below:



where every line segment between two junctions is a loop around one of the $\mathbb{S}^1$.

This should have provided a good intuition of *How paths can generate a covering space* : Basically a simply connected $\widetilde{X}$ is like a map drawn by an amnesia patient walking in X —— each time he gets back to x_0, he won't realize that he has been there before. This is how covering space is related to loops, and thus the $\pi_1(X)$.
So we can also partially "cure" this patient by letting him know that he's just got back to the starting point in some certain loops, and then he will produce a different "map":

Proposition 1.2. Suppose X is PC, LPC and SLSC, then for $\forall H \leq \pi_1(X)$, $\exists$ covering space $p : \widetilde{X}_H \to X$ such that $p\left(\pi_1(\widetilde{X}_H)\right) = H$.

Proof. Start with a simply connected covering space $\widetilde{X}_0 = \{\text{paths in } X \text{ starting at } x_0\} / \approx$, (equipped with topology described in the proof of Proposition 1.1) Then quotient it by the equivalence $\sim_H$ where

$$[\alpha] \sim_H [\beta] \quad \iff \quad \alpha(1) = \beta(1) \text{ and } [\alpha^{-1}\beta] \in H$$

then we let $\widetilde{X}_H = \widetilde{X}_0 / \sim_H$ (verify that the covering map $p : \widetilde{X}_0 \to X$ is naturally a covering map for $\widetilde{X}_H$). Note that for any loop γ in X, the $p^{-1}([\gamma])$ are loops in $\widetilde{X}_H$ iff $[\gamma] \in H$, so we have $p(\pi_1(\widetilde{X}_H)) = H$. $\square$

An interesting similar (yet unrelated) result is that

Example 1.3. Any group is a fundamental group of some topological space.

It can be proved with the knowledge in next subsection.

Since the $\widetilde{X}$ can be constructed via taking paths in X, it makes sense that they process some kinds of *symmetry* generated by "duplicating" X. Such symmetry is demonstrated by the Deck transformation $\mathrm{G}(\widetilde{X}, X)$ which we defined earlier.

Example 1.4. For $\widetilde{X} = \mathbb{R}$ and $X = \{z \in \mathbb{C} : |z| = 1\} \cong \mathbb{S}^1$ with $p(x) = \exp(2x\pi i)$, we have that any Deck transformation must be in form of $f(x) = x^n$ for $n \in \mathbb{Z}$, so $\mathrm{G}(\mathbb{R}, \mathbb{S}^1) \cong \mathbb{Z}$.

1.2 Computation of fundamental group

Still, we assume X to be a path-connected topological space; and when we talk about $\pi_1(\cdot)$, we assume the space is path-connected.

We say $\pi_1(X)$ is a topological invariant, but under what is it invariant? Certainly it is invariant under homeomorphism, but further, it is actually invariant under homotopy equivalence; the property of being invariant under homotopy equivalence is called *homotopy invariant*.

Proposition 1.3. $\pi_1(\cdot)$ is homotopy invariant.

Proof. Conisder $X \simeq Y$ being path connected; by definition, there exists $f : X \to Y$, and $g : Y \to X$ sych that $f \circ g \approx \mathrm{id}_Y$ and $g \circ f \approx \mathrm{id}_X$.
Notice that if $\phi_1, \phi_2 \in C^0(X, Y)$ are homotopic, then obviously $\phi_1(\pi_1(X)) \cong \phi_2(\pi_1(X))$; therefore, $(f \circ g)$ and $(g \circ f)$ are homotopic to identity means that they induces automorphisms on $\pi_1(X)$ and $\pi_1(Y)$ respectively. And therefore f and g must induce isormophisms between $\pi_1(X)$ and $\pi_1(Y)$. $\square$

This simplifies things a lot. For example, by geometric deformation, we can deduce that $\pi_1(\mathbb{M}) \cong \pi_1(\mathbb{S}^1)$, and $\pi_1(\mathbb{T}^2 \setminus \{p\}) \cong \pi_1(\mathbb{S}^1 \vee \mathbb{S}^1)$ (where p is a point on $\mathbb{T}^2$).
But we will still need to be able to calculate some simple cases; for most of the use in differential topology, we only need to know these:

Proposition 1.4. We have $\pi_1(\mathbb{S}^n) \ = \{0\}$ for $\forall n \geq 2$ and

$$\pi_1(\mathbb{S}^1) \cong \mathbb{Z}$$

Proof. The $n \geq 2$ case is just easy: For any loop γ, pick $x \notin \mathrm{Im}(\gamma)$, then since $\mathbb{S}^n \setminus \{x\} \cong \mathbb{R}^n$ is simply connected, γ is a trivial loop.
For $\mathbb{S}^1$, we embed $\mathbb{S}^1 \subset \mathbb{C}$ and construct a universal cover $p : \mathbb{R} \to \mathbb{S}^1$ of it by $p(x) = e^{2\pi x i}$; then by Galois correspondence (Theorem 1.1), we have $\pi_1(\mathbb{S}^1) \cong \mathrm{G}(\mathbb{R}, \mathbb{S}^1) \cong \mathbb{Z}$. $\square$

And what's left is to deal with the spaces what are "composed" by known cases; for example, it is easy to see that: for path-connected $X_1, \cdots, X_n$, we have

$$\pi_1(X \vee Y) = \pi_1(X) * \pi_1(Y) \tag{$\star$}$$

via the map $\Phi : \pi_1(X) * \pi_1(Y) \longrightarrow \pi_1(X \vee Y)$ where $\Phi([\alpha_1] * \cdots * [\alpha_m]) = [\alpha_1 \dots \alpha_m]$ by verifying that its an isomorphism.

Then we can upgrade this to a more general setting:

Van-Kampen theorem

Theorem 1.5. Given two path-connected subspaces $U, V \subset X$ with natural inclusions $\iota : U \hookrightarrow X$ and $\tau : V \hookrightarrow X$, such that $U \cap V$ is path-connected,

$$\pi_1(X) \cong (\pi_1(U) * \pi_1(V))\,/H$$

where $H = \langle \iota(\alpha)\tau(\alpha)^{-1} : [\alpha] \in \pi_1(U \cap V)\rangle$

The proof to this theorem is classical, and almost everywhere on internet, so here we give a concise version of it, some details are omitted as exercise:

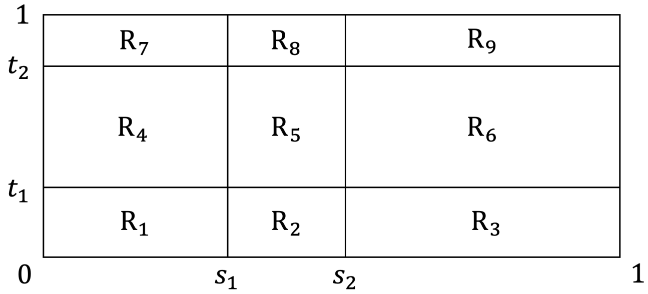
Proof. The idea is using $1^{\text{st.}}$ isomorphism theorem. Similarly, consider the map

$$\begin{aligned} \Phi : \pi_1(U) * \pi_1(V) &\to \pi_1(X) \\ [\alpha_1] * \cdots * [\alpha_m] &\mapsto [\alpha_1 \ldots \alpha_m] \end{aligned}$$

It is surjective, as every loop in X is homotopic to a product of loops $\alpha_1 \cdot \cdots \cdot \alpha_m$ where each α_j is contained entirely in U or V. (Exercise)
Obviously $H \subset \ker(\Phi)$. To show $H = \ker(\Phi)$, we define $\sim$ to be the equivalence relation in quotient group $(\pi_1(U) * \pi_1(V))\,/H$, and try showing that

$$(f_1 * \cdots * f_p) \approx (g_1 * \cdots * g_q) \Longrightarrow (f_1 * \cdots * f_p) \sim (g_1 * \cdots * g_q)$$

where each f_j, g_j is contained either in U or V.
Suppose $F : [0,1] \times [0,1] \to X$ is the homotopy from $f := (f_1 * \cdots * f_p)$ to $g := (g_1 * \cdots * g_q)$. By compactness of $[0,1] \times [0,1]$, we can always find $0 = s_0 < \cdots < s_m = 1$ and $0 = t_0 < \cdots < t_n = 1$ such that at least one of U or V contains entirely $F(R_{ij})$, where $R_{ij} := [s_{i-1}, s_i] \times [t_{j-1}, t_j]$. (Exercise)
Order these R_{ij} as $R_{11}, \ldots, R_{1n}, \ldots, R_{mn}$, and let R_k to be the k^{th} of R_{ij} under this order. Then let γ_k be a path such that the first component of γ_k is $\geq$ than those of $\mathrm{R}_1, \ldots, \mathrm{R}_k$ but $\leq$ than those of the rest. So $F(\gamma_0) = F|_{(0,\cdot)}$ and $F(\gamma_{mn}) = F|_{(1,\cdot)}$.



Easy to construct homotopy $F(\gamma_k) \approx F(\gamma_{k+1})$ for all $1 \leq k < mn$. Therefore, $F(\gamma_k) \sim F(\gamma_{k+1})$ because this is a homotopy between loops in $U \cap V$. $\square$

A classical example would be $X = \mathbb{T}^2$, the torus.
Consider a neighbor U of $p \in \mathbb{T}^2$ that is homeomorphic to $\mathbb{R}^2$, and let $V = \mathbb{T}^2 \setminus \{p\}$. We can show that $V \simeq \mathbb{S}^1 \vee \mathbb{S}^1$ (illustrated in the diagram below)

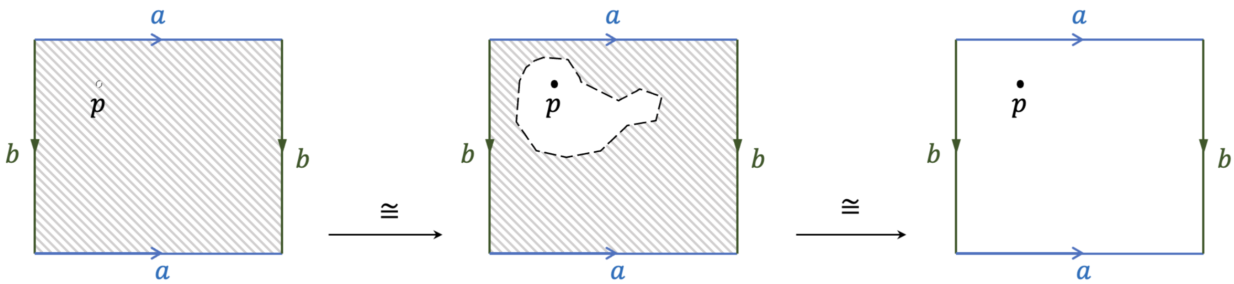


Figure 1: Torus $\mathbb{T}^2$ can be constructed as $[0,1] \times [0,1]$ by identifying two horizontal edges a and vertical edges b together. Removing a point $p \in \mathbb{T}^2$ from it would give $\mathbb{S}^1 \vee \mathbb{S}^1$;

For any $x_0 \in U \cap V$, the generator of $\pi_1(U \cap V, x_0) \cong \mathbb{Z}$ is equal to $(aba^{-1}b^{-1})$ as demonstrated in the figure (though you could try doing this formally), where $a, b \in \pi_1(V, x_0)$ are generators of $\pi_1(V, x_0) = F_{\{a,b\}}$. Hence

$$\begin{aligned} \pi_1(\mathbb{T}^2) \cong \pi_1(\mathbb{T}^2, x_0) &\cong F_{\{a,b\}}/\langle aba^{-1}b^{-1} \rangle \\ &\cong \mathbb{Z} \times \mathbb{Z} \end{aligned}$$

where F_S denotes the free group generated by a set S.

The Van-Kampen theorem enables us to compute the π_1 of almost any topological space with an explicit construction.
Mostly we care about spaces constructed by unions of lines or surfaces in differential topology; and one intuitive way to get those is via *Cell complex* (also called *CW complex*).

Definition 1.5. A **Cell complex** is the union of a sequence of topological spaces $\emptyset = X_{-1} \subset X_0 \subset X_1 \subset X_2 \subset \cdots$ such that: For $\forall n \in \mathbb{Z}_{>0}$, $\exists$ continuous maps $\phi_1, \cdots, \phi_k : \partial \mathbb{D}^n \to X_{n-1}$ such that

$$X_n = X_{n-1} \sqcup_{\phi_1} \mathbb{D}^n \sqcup_{\phi_2} \mathbb{D}^n \sqcup_{\phi_3} \cdots \sqcup_{\phi_k} \mathbb{D}^n$$

The $\mathbb{D}^n$ in it is called the n**-cell**.
The X_n in it is called the n**-skeleton**.

For example, the $\mathbb{S}^n$ is a cell complex, by letting $X_0 = \cdots = X_{n-1}$ be a point and $X_n = X_{n-1} \sqcup_\phi \mathbb{D}^n$, where $\phi(\partial \mathbb{D}^n) = X_{n-1}$. We could also construct

1. Möbius strip $\mathbb{M}$ as two 0-cells, three 1-cells, and one 2-cell.
2. Klein bottle $\mathbb{K}$ as one 0-cell, two 1-cells and one 2-cell.
3. Torus $\mathbb{T}^2$ as one 0-cell, two 1-cells, and one 2-cell.

In fact, any smooth 2-dim surface you can think of can be constructed as a cell complex. (This is not a trivial fact; later we will try to prove it after introducing *manifold* and some Morse theory.)

The good thing about describing a topological space in cell complex is that, π_1 of a cell complex is totally determined by its 2-skeleton:

Proposition 1.5. In a path-connected cell complex $X = \bigcup_k X_k$, we have

- inclusion $X_1 \hookrightarrow X$ induces surjective homomorphism $\pi_1(X_1) \to \pi_1(X)$.
- inclusion $X_2 \hookrightarrow X$ induces isomorphism $\pi_1(X_2) \to \pi_1(X)$.

Proof. For $k \geq 2$ and k-skeleton X_k, attaching any more $(k+1)$ cell will not change its fundamental group; in other words,

$$\pi_1(X_k) \cong \pi_1(X_2) \quad \text{for } \forall k \geq 2$$

this proves the second statement.

For the first statement, we thus only need to show that the inclusion $X_1 \hookrightarrow X_2$ induces surjective homomorphism $\pi_1(X_1) \to \pi_1(X_2)$. $\square$

1.3 Classification of Compact surfaces

Classification theorem of connected compact surfaces

Theorem 1.6. Suppose M is a connected compact surface,

1. If $\partial M = \emptyset$, then M is homeomorphic to

$$\mathbb{S}^2 \quad \text{or} \quad \mathbb{T}^2 \# \cdots \# \mathbb{T}^2 \quad \text{or} \quad \mathbb{RP}^2 \# \cdots \# \mathbb{RP}^2$$

2. If $\partial M \neq \emptyset$ and has r connected components, then M is homeomorphic to

$$\mathbb{S}^2 \text{ with } \quad \text{or} \quad \mathbb{T}^2 \# \cdots \# \mathbb{T}^2 \quad \text{or} \quad \mathbb{RP}^2 \# \cdots \# \mathbb{RP}^2$$

1.4 Whitehead's theorem

Whitehead's theorem

Theorem 1.7. Two CW complexes X, Y are homotopy equivalent iff

$$\pi_n(X) \cong \pi_n(Y) \quad \text{for } n = 1, 2, 3, \ldots$$

1.5 Questions

Those questions are sometimes used as a lemma in sections later on.

1. True or False:
(i) A bijective continuous map is a homeomorphism.
(ii) Path-connected topological spaces are locally path-connected.
(iii) Covering space of Hausdorff topological space is Hausdorff.
(iv) If a topological space has Hausdorff covering space, then it is Hausdorff.

2. A **local homeomorphism** from topological space X to Y is a map $\phi : X \to Y$ such that $\forall x \in X$ has a neighbour N such that $\phi(N)$ is open and $\phi|_N$ is a homeomorphism.

   1. Show that local homeomorphism maps are open maps.[1]
   2. Show that bijective local homeomorphism is a homeomorphism.
   3. Show that X is path-connected iff Y is.

3. Consider topological space X and its subspace A, a continuous map $f : X \to A$ is called a **retraction** if $f(a) = a$ for $\forall a \in A$.

   1. Show that f induces a surjective homomorphism from $\pi_1(X)$ to $\pi_1(A)$.
   2. Let $B = \overline{B_1(0)} \subset \mathbb{R}^2$, is there a retraction from B to ∂B?
   3. Given Möbius strip $\mathbb{M} := [-1, 1] \times [-1, 1]/\sim$ where $(x_1, y_1) \sim (x_2, y_2)$ iff $x_1 = -x_2 = \pm 1$ and $y_1 = -y_2$. Is there a retraction from $\mathbb{M}$ to $\partial\mathbb{M}$?

4. (Example 1.3) Show that every group can be isomorphic to the fundamental group of some topological space. *(Hint: attaching 2-Cells with edges glued along loops.)*

5. Consider a homomorphism $\phi_\gamma : \pi_1(X, a) \to \pi_1(X, b)$ defined by $\phi([\beta]) = [\gamma \cdot \beta \cdot \gamma^{-1}]$, where γ is a path from a to b.
Suppose X is path-connected, show that $\pi_1(X)$ is abelian iff ϕ_γ is unique for every $\gamma(1)$.

[1] Open maps are maps that map open sets only to open sets.

1.6 Answers

Question (1)
(i) False. Consider the identity map from $(X,\mathcal{I})$ to $(X,\mathcal{D})$, where $\mathcal{D}$ is the discrete topology on X and $\mathcal{I}$ is the indiscrete topology.
(ii) False. Consider $\bigcup_{n=0}^{n} B_{2^{-n}}(2^{-n},0) \subset \mathbb{R}^2$ equipped with subspace topology of $\mathbb{R}^2$ (iii) True. Trivial. (iv) False. Consider $X=([0,1],\tau)$ where $\tau=\{A\subseteq[0,1]:\}$

Question (2)
(i) Suppose $\phi : X \to Y$ is a local homeomorphism and $U \subset X$ is an open set. For any $y=\phi^{-1}(x)\in\phi(U)$, there exists a neighbour N of x that $\phi|_N$ is homeomorphism, which gives that $\phi(N\cap U)$ is an open neighbour of y in $\phi(U)$, so $\phi(U)$ is open.
(ii) Homeomorphism is exactly open and continuous map, we already proved open, it remains to show that $\phi : X \to Y$ in (a) is continuous (when it's bijective).
Consider any open $V \subset \mathrm{Im}(\phi) = Y$, and $x\in\phi^{-1}(V)$, there exists a neighbour N of x that $\phi|_N$ is homeomorphism, which gives that $\phi^{-1}(\phi(N)\cap V)$ is an open neighbour of x in $\phi^{-1}(V)$, thus ϕ continuous.
(iii)

Question (3)
(i) Consider natural inclusion $\iota : A \hookrightarrow X$, then $f\circ\iota : A \to A$ is an identity map, which means f must be a surjection from $\pi_1(X)$ to $\pi_1(A)$.
(ii) No. $\pi_1(B)\cong\{0\}$ but $\pi_1(\partial B)\cong\mathbb{Z}$, no surjection from $\{0\}$ to $\mathbb{Z}$, thus no retraction.
(iii) Consider natural inclusion $\iota : \partial\mathbb{M} \hookrightarrow \mathbb{M}$, then the homomorphism

2 Basics: Tensor & manifold

We will first quickly go through some basic concepts that appear everywhere in our work, which I will try to make quick and stimulating.

In short, the manifold is an intrinsic definition of a hypersurface; and a tensor is a way that makes our life easier when describing multi-variate functions (like vector fields) on a manifold, by converting multlinear maps into linear maps.

Let's start with tensor.
We are familiar with the notion of bilinear maps: they map vectors from one vector space to another linearly (in both arguments). The **Multilinear maps** are a simple generalization: we say $\phi : V_1 \times V_2 \times \cdots \times V_n \to W$ is multilinear if

$$\phi(v_1, \cdots, a_i v_i, \cdots, v_n) = a_i \phi(v_1, \cdots, v_n)$$
$$\phi(v_1, \cdots, v_n) + \phi(u_1, \cdots, u_n) = \phi(v_1 + u_1, \cdots, v_n + u_n)$$

for $i = 1, 2, \cdots, n$.

Going forward, we denote by $\mathcal{L}(V_1 \times V_2, W)$ the set of all multilinear maps from $V_1 \times V_2$ to W. Just like $\mathrm{Hom}_{\mathbb{R}}(V, W)$ for vector spaces V, W, we can give $\mathcal{L}(V_1 \times V_2, W)$ a vector space structure by letting $(f + \lambda \cdot g)(\mathbf{v}) := f(\mathbf{v}) + \lambda \cdot g(\mathbf{v})$ for $\forall f, g \in \mathcal{L}(V_1 \times V_2, W)$.

2.1 The tensor product

The tensor product is a way to write a multilinear map in form of linear map.
To get some intuition, let's consider the example of a bilinear map:

Example 2.1. Let's consider a bilinear map $Q : V_1 \times V_2 \to \mathbb{R}$ where $V_1 = V_2 = \mathbb{R}^3 = \operatorname{span}\{e_1, e_2, e_3\}$, and

$$Q(v,u) = v^T \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix}}_{A} u = \begin{matrix} 1 \cdot v_1u_1 & + & 0 \cdot v_1u_2 & + & 0 \cdot v_1u_3 & + \\ 0 \cdot v_2u_1 & + & 0 \cdot v_2u_2 & + & 1 \cdot v_2u_3 & + \\ 3 \cdot v_3u_1 & + & 2 \cdot v_3u_2 & + & 1 \cdot v_3u_3 & \end{matrix}$$

Observe that $Q(v,u)$ is a linear combination of 9 independent components determined by v and u, and so can be transformed as *a linear map from a 9-dim vector space to* $\mathbb{R}$. How do we transform it into such a linear map? First, we can define

$$f : (x_i\mathbf{e}_i, y_j\mathbf{e}_j) \mapsto x_iy_j(\mathbf{e}_i \otimes \mathbf{e}_j)$$

where we let $\{e_i \otimes e_j\}$ to be the basis of a 9-element vector space denoted as $V_1 \otimes V_2$; then we define linear $\widetilde{Q} : V_1 \otimes V_2 \to \mathbb{R}$, such that

$$\widetilde{Q} : \sum_{i,j} x_iy_j(\mathbf{e}_i \otimes \mathbf{e}_j) \mapsto \sum_{i,j} a_{ij}x_iy_j$$

where a_{ij} is the (i,j) entry of A. Then $\widetilde{Q}$ is the linear map that "represents" Q.

Given two $\mathbb{R}$-vector spaces V_1 and V_2,we wish to define a vector-space structure to transform any bilinear map $Q \in \mathcal{L}(V_1 \times V_2, \mathbb{R})$ into a linear map. Though the direct sum seems to be one way, it doesn't give the *bilinearity*: Consider $\forall u_1, u_2 \in V_1, v \in V_2$,

$$\begin{aligned}(u_1 + u_2) \oplus (v + 0) &= (u_1 \oplus v) + (u_2 \oplus 0) \\ &\neq (u_1 \oplus v) + (u_2 \oplus v)\end{aligned}$$

We therefore need *another* vector space structure on $V_1 \times V_2$, which is the **tensor product**. Informally, given bases B_1 and B_2 for V_1 and V_2 respectively, we define

$$V_1 \otimes V_2 := \operatorname{span}\{v \otimes w \ : \ v \in B_1, w \in B_2\}$$

where the tensor product $\otimes$ behaves like[2]

$$\left(\sum_{v \in B_1} a_v v\right) \otimes \left(\sum_{w \in B_2} b_w w\right) = \sum_{v \in B_1} \sum_{w \in B_2} a_v b_w \,(v \otimes w)$$

[2] It is of course possible to do things formally, but it is unnecessary for the contents afterwards. We will revisit this in the algebra section, and that is where we define $\otimes$ in an axiom-like way.

More generally, for vector spaces $V_1, \dots, V_n$, we define

$$\bigotimes_{k=1}^{n} V_k := \sum_{v_k \in V_k} v_1 \otimes v_2 \otimes \cdots \otimes v_n$$

Fixing a vector space V; we can observe that the set

$$\mathrm{T}(V) = \sum_{n=0}^{\infty} \left(\mathbb{R} \cdot \bigotimes^{n} V \right)$$

forms an algebra under $\otimes$ and $+$, and this algebra $\mathrm{T}(V)$ is called the **Tensor algebra**.

> *Remark* 1. The notation $\bigotimes^n V$ is misleading, because $\bigotimes^n V$ does NOT imply that every element can be written as $V \otimes \cdots \otimes V$. For example, let $V = \mathbb{R}^4$,
>
> $$\omega = e_1 \otimes e_2 + e_3 \otimes e_4$$
>
> is impossible to write $\omega \in \bigotimes^2 V$ in the form of $V \otimes V$.
> We call an element $\omega \in \bigotimes^k V$ *decomposable*, if ω can be expressed in form of $V \otimes \cdots \otimes V$.
> So naturally a good question is how to justify if ω is decomposable.

Now back to our original point. With this definition, we can see how $\otimes$ offers a way to transform multilinearity into linearity: Consider a map

$$\iota : V \times W \to V \otimes W$$
$$(v, w) \to v \otimes w$$

then for any $\phi \in \mathcal{L}(V \times W, U)$, there exists an unique linear map $\widetilde{\phi} : V \otimes W \to U$ such that $\phi = \widetilde{\phi} \circ \iota$. This builds a natural isomorphism

$$\mathcal{L}(V \times W, U) \cong \mathrm{Hom}_{\mathbb{R}}(V \otimes W, U) \tag{2.1}$$

demonstrating the correspondence between multilinear maps and tensor product.

Now we know the tensor product, then what is a *tensor*? A tensor is basically a tensor product of a vector space with itself and its dual space. We will talk more about it when enter into differential geometry.

Definition 2.1. The set of all type (p, q) **tensor**s of V is defined as

$$\mathrm{T}_q^p(V) = \bigotimes^p V \otimes \bigotimes^q V^*$$

Notice that for any vector spaces S, T, there is a canonical isomorphism

$$S^* \otimes T^* \cong (S \otimes T)^*$$

by sending $(s_i \mapsto a_i) \otimes (t_j \mapsto b_j) \mapsto ((s_i \otimes t_j) \mapsto a_i b_j)$ where s_i and t_j are basis vectors for S

and T respectively. So by (2.1), we have isomorphism

$$\mathrm{T}^p_q(V) \cong \mathcal{L}(\underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q}, \mathbb{R})$$

(be aware the order of V and V^* is opposite to that in definition)
In this part, we gave a non-rigorous definition for $\otimes$, and there are some more about operations on tensor; we will get to them afterwards when necessary, but I believe it is not helpful to continue on here.

2.2 Manifold and the smoothness

Let n be a positive integer.
Manifold is an intrinsic definition of a "surface", that its definition does not rely on embedding it inside a coordinate. Intuitively, standing inside a n-dim manifold M will make you feel like in $\mathbb{R}^n$, just like ancient human thought earth is plane.

Definition 2.2. Some background definitions for this subsection

1. **Locally Euclidean**: It means that any point of a topological space has a neighbour homeomorphic to an open set in $\mathbb{R}^n$ for some $n \in \mathbb{Z}_{>0}$.
2. **Manifold**: A n-dimensional manifold M is a Locally Euclidean, Second countable and Hausdorff topological space.
3. **Chart**: It is an ordered pair (U_a, φ_a), where φ_a is a homeomorphism mapping $U_a \in M$ to an open ball in $\mathbb{R}^n$.
4. **Atlas**: It is the set of charts on M. In other words, atlas defines a "location" for each $a \in M$. (However, one point can have multiple such "locations")

The chart can be thought as an "*local coordinate*" system, since assigning a neighbour of $p \in M$ into $\mathbb{R}^n$ is just like drawing a part of the world's map on paper.
Just like what manifolds are intuitively, all smooth parametric surfaces in $\mathbb{R}^m$ are manifolds; interestingly, the converse is (almost) true —— this is Whitney embedding theorem, which we will show in Section 2.6.

> *Remark* 2. Why Hausdorff and Second countable?
> **1**, The condition of *Hausdorff* for manifold is important because it enables the existence of limit and thus differentiation. (c.e. see line with two origins)
> **2**, The condition of *Second countable* is important because it allows the *Partition of unity*, which is almost everywhere in the theories concerning real manifolds. We'll get to it in subsection 2.5.

Be aware that things like $[0, 1]$ is NOT a manifold! It is a kind of "semi"-manifold called the *Manifold with boundary*; I think this is a terrible name, because it causes confusion since a manifold shouldn't have boundary.

Definition 2.3. Topological space M is a **manifold with boundary** if it is Second countable and Hausdorff, and for $\forall p \in M$, $\exists$ neighbour N_p of p such that $N_p \cong \mathbb{B}_1(0)$ or $\mathbb{H}_1(0)$.

The charts in a manifold with boundary are still well-defined, just let their pre-image be $B_1(0)$ or $H_1(0)$ with origin set to 0. We can conclude two interesting facts:

Proposition 2.1. Suppose M is a manifold with manifold, where the interior of M is a n-dim manifold;

1. ∂M is $(n-1)$-dim manifold.
2. $\partial(\partial M) = \emptyset$.

The first statement follows immediately from definition, and the second one follows immediately from the first.

How to define differentiation on manifolds?
We already know well about differentiation on $\mathbb{R}^n$, so the atlas should allow us to define that on M. However, a single point in M can be mapped to multiple "locations" in $\mathbb{R}^n$ by atlas; this suggests that differentiability depends on charts: a differentiable function in one chart may be NOT differentiable in another.
We fix this issue by introducing the concept of *smooth structure*.

Definition 2.4. Suppose an atlas $\{(U_p, \varphi_p)\}$ of manifold M satisfies that

$$(\varphi_b \circ \varphi_a^{-1}) \in C^\infty(\varphi_a(U_a), \varphi_b(U_b))$$

for $\forall a, b \in M$ with $U_a \cap U_b \neq \emptyset$, then this atlas defines a smooth structure on M.
And the function $(\varphi_b \circ \varphi_a^{-1})$ is called the **transition function**.

In other words, this property guarantees a "smooth" transition between any two charts. We call those M that process a smooth structure as **smooth manifold**.[3]

Remark. You might find it really easy to find examples of smooth manifolds: $\mathbb{S}^2$, $\mathbb{S}^3$, $\mathbb{S}^1 \times \mathbb{S}^1$, or $\mathbb{RP}^2$ ... It turns out that all manifolds of dimension ≤ 3 are smooth! (however the proof is highly nontrivial.)
So a natural question to ask is that "Is there any manifold that is NOT smooth?" There is; however, it is also hard to construct; such manifold was first constructed in paper *A Manifold which does not admit any Differentiable Structure*.

Interestingly the transition map can help us to define the *orientability* of a manifold, that is whether a manifold can be assigned with a "direction":

[3] also called **differentiable manifold** in some places

Definition 2.5. For a connected smooth manifold M, an atlas of it is **orientation-preserving** if all its transition functions have positive Jacobian.
M is **orientable** if it has an orientation-preserving alas.
Let $\mathscr{U}$ be the set of orientation-preserving atlas on M, define an equivalence relation on it by: For $\mathcal{U}_1, \mathcal{U}_2 \in \mathscr{U}$,

$$\mathcal{U}_1 \sim \mathcal{U}_2 \Longleftrightarrow \mathcal{U}_1 \cup \mathcal{U}_2 \text{ is oriental preserving}$$

Then $\mathscr{U}$ is divided into two equivalent classes: $\mathscr{U} = \mathscr{U}_+ \sqcup \mathscr{U}_1$. An **orientation** on M is an element of $\{\mathscr{U}_+, \mathscr{U}_-\}$. And M is **oriented** means that all atlas on M are assumed to be in a fixed orientation.

In this definition, it is easy to justify that manifolds like $\mathbb{R}^n$, $\mathbb{S}^n$ and any direct product of orientable manifolds are orientable (we will leave this in Exercise), we can also justify that the Möbius strip and $\mathbb{RP}^2$ are NOT orientable.
Claim. Möbius strip and $\mathbb{RP}^2$ are NOT orientable. (*The proof could be attempted after the next subsection. See Question & Answer*)

With differential structure, the smoothness of maps from M can be properly defined:

Definition 2.6. Consider manifolds X and Y and a map $F : X \to Y$, we say that F is C^k**-map** if

$$\tau_x := (\phi_y \circ \phi \circ \varphi_x^{-1}) \in C^k(\varphi_x(U_x), \phi_y(U_y))$$

for every $y = F(x) \in \mathrm{Im}(F)$ and charts φ, ϕ at x, y respectively.
In particular, a map is

- **smooth** if $k = \infty$;
- **diffeomorphism** if it is bijective C^1 map with C^1 inverse;
- **smooth scalar field** $f \in C^\infty(X)$ if f is a smooth function from X to $\mathbb{R}$.

Take an example of defining a smooth scalar field f using the definition above; Notice that in $U_a \cap U_b$ the smoothness of f suggests $\tau_a = \tau_b \circ (\varphi_b \circ \varphi_a^{-1}) \in C^\infty(\varphi(U_a))$, which means $(\varphi_b \circ \varphi_a^{-1})$ must be smooth; Hence, the definition only works when we have the differential structure.
Also, recall the definition of manifold, we defined that $\forall p \in M$, there $\exists$ neighbour U_p that the chart φ_p defines a *homeomorphism*. For a smooth manifold, it can be "upgraded" to **diffeomorphism** for the same reason.

Now we have the smooth scalar field, we can define the tangent space.

2.3 Tangent space & Pushforward

[**Convention**] In the rest of this section, we assume:
1, Manifold M is a n-dim connected smooth manifold;
2, If a neighbour of $p \in M$ is denoted as U_p, then it is the one which is homeomorphic to an open subset of $\mathbb{R}^n$.

The tangent space is the generalization of tangent lines for smooth curves and tangent planes for smooth surfaces. Clearly it's easy to define "tangent" if M is parameterized in $\mathbb{R}^m$; so the question is about defining it **intrinsically**. (i.e. without using something like "normal vector" that requires putting M into higher dimensional space.)
Consider the local coordinate defined by chart φ_p (which continuously maps every point in U_p as $(x_1, x_2, \cdots, x_n)$); in such a system,

Definition 2.7. A **tangent vector** at $p \in M$ is a map $v_p : C^\infty(U_p) \to \mathbb{R}$ defined as

$$v_p(f) = \sum_{j=1}^{n} a_j \left(\frac{\partial}{\partial x_j} (f \circ \varphi_p^{-1}) \right) \Bigg|_{p'}$$

where $a_j \in \mathbb{R}$ and $p' = \varphi_p(p)$. The **Tangent space** T_pM is the vector space of all tangent vectors at p.

which is, intuitively, all the possible directions in which we can take "partial derivative". So it is natural to form the basis of T_pM by tangent vectors with $a_j = \delta_{ij}$ for $i = 1, \cdots, n$; such basis is called a **Coordinate basis**.
So geometrically, if we consider a smooth parametric surface as a manifold, such basis can be intuitively viewed as the derivatives of parameterization:

Example 2.2. Suppose M is smoothly embedded in $\mathbb{R}^m$; For $p \in M$, its local chart $((U_p, \varphi))$ gives a parameterization of M by: $\varphi^{-1} : \varphi(U_p) \subset \mathbb{R}^n \to U_p \subset \mathbb{R}^m$. Suppose

$$\varphi^{-1}(x_1, \cdots, x_n) = \mathbf{F}(x_1, \cdots, x_n)$$

where $p = \varphi^{-1}(0)$. Take $q \in \varphi^{-1}(x_1, \cdots, x_n)$, the coordinate basis for T_qM is simply the column vectors of differential of φ^{-1} at q:

$$T_qM = \text{span}\{\partial\mathbf{F}/\partial x_1, \cdots, \partial\mathbf{F}/\partial x_n\}$$

Therefore we can also view a $v \in T_pM$ as derivative of a path starting from p; that is,

$$T_pM := \{\gamma'(0) : \text{path } \gamma \text{ starting from } p \in M\} \tag{2.2}$$

where γ' can be calculated via a local coordinate.
This way is clearly easier and intuitive to calculate, but here tangent vector is no longer a

linear functional on $C^\infty(M)$; so though being vague, we can see later that the former definition has more advantage for building theories.

The dual space of tangent space is called the **cotangent space**,

Definition 2.8. The Cotangent space T_p^*M is the dual space of T_pM.

which is also a n-dim vector space.
The coordinate basis of cotangent space is normally denoted as $\{\mathrm{d}x_1, \cdots, \mathrm{d}x_n\}$, where they are chosen so that

$$\mathrm{d}x_i(\partial_j) = \delta_{ij} \quad \text{for } \forall i, j = 1, \cdots, n$$

So given an inner product g on T_pM, the natural isomorphism between T_pM and T_p^*M is given by $\partial_i \mapsto \sum_{j=1}^n g_{ij}\,\mathrm{d}x_j$, where g_{ij} is the (i,j)-entry of g.
There seems to be no significant reason to consider a dual space of tangent space now, but after introducing the differential form in the next chapter, we can go back here and see the underlying motivation.

Let X and Y be two smooth manifolds. Given a smooth function $\phi : X \to Y$; note that $\phi(X)$ is also a smooth manifold, so ϕ naturally brings the local information of points in X to that of points in $\phi(X)$ —— like tangent spaces and cotangent spaces —— leading us to the idea of Pushforward & Pullback.
We first start from the most intuitive case, pushforward of tangent spaces

Definition 2.9. (Pull & Push) Given a C^1-map ϕ from smooth manifolds X to Y; the **Pullback** of $f \in C^\infty(Y)$ is $\phi^* f \in C^\infty(X)$ defined by

$$\phi^* f(x) = f(\phi(x))$$

for $\forall x \in X$. And based on this Pullback, the **Pushforward** map is $\phi_* : T_xX \to T_{\phi(x)}Y$ defined by: for $\forall f \in C^\infty(Y)$,

$$\left.(\phi_* v)\right|_{\phi(x)}(f) = \left.v\right|_x(\phi^* f)$$

Though the definition looks complicated, there is a straightforward geometric intuition: if X, Y are embedded in some Euclidean spaces, ϕ_* is just the differential of ϕ, so in a lot of books (and even Wikipedia), the pushforward of ϕ is also written as "$D\phi$".
Because of this, it processes the two key properties that a differential has:

Proposition 2.3. The pushforward map satisfies

1. **Linear:** $\phi_*(u+v) = \phi_*(u) + \phi_*(v)$;

2. **Chain rule:** $(\rho \circ \phi)_* = \rho_* \circ \phi_*$;

where $\phi : X \to Y$ and $\rho : Y \to Z$ are smooth maps, and X, Y, Z are smooth manifolds.

Proof. The linear property is easy to prove, so we focus on the Chain rule;
For $\phi(x) = y$, $\rho(y) = z$ and $v \in \Gamma^\infty(TX)$, we have that

$$\begin{aligned}(\rho \circ \phi)_*(v)\big|_z (f) &= v\big|_x (\rho^*(\phi^* f)) \\ &= \rho_* v\big|_y (\phi^* f) = (\rho_* \circ \phi_*) v\big|_z (f)\end{aligned}$$

for $\forall f \in C^\infty(Z)$. And similarly for $u \in \Gamma^\infty(T^*X)$,

$$\begin{aligned}(\rho \circ \phi)_*(u)\big|_z (v) &= v\big|_x (\rho_*(\phi_* v)) \\ &= \rho_* u\big|_y (\phi_* v) = (\rho_* \circ \phi_*) u\big|_z (v)\end{aligned}$$

for $\forall v \in \Gamma^\infty(TM)$. Hence it is also true for general cases of tensor fields. $\square$

Example 2.4. This language can support a very interesting interpretation of the **Cauchy-Riemann equation** in Complex analysis.
We know that a differentiable function $f(x + yi) = u(u, y) + v(x, y)i$ in $\mathbb{C}$ is essentially a differentiable $\widetilde{f} : \mathbb{R}^2 \to \mathbb{R}^2$ by looking at its components, and the pushforward $\widetilde{f}_* \in M_2(\mathbb{R})$ gives the differentiation of f:

$$\widetilde{f}_* = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix}$$

Notice that the differentiation is a linear operator, thus commute with multiplying i. So if we define $g : \mathbb{R}^2 \to \mathbb{R}^2$ to be the map given by $z \mapsto i \cdot z$, then by chain rule, we have $\widetilde{f}_* g_* = g_* \widetilde{f}_*$, which means

$$\widetilde{f}_* \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \widetilde{f}_*$$
$$\Longrightarrow \begin{cases} \partial u/\partial x = \partial v/\partial y \\ \partial u/\partial y = -\partial v/\partial x \end{cases}$$

i.e. Cauchy-Riemann Equation is equivalent to say that *pushforward of analytic functions must commute with that of multiplying i.*

In this interpretation, we can express the **Inverse Function theorem** for manifolds.

Proposition 2.2. Consider smooth manifolds X, Y with $\dim(X) = n$, and map $\phi \in C^1(X, Y)$ with $\phi(x) = y$ and $\mathrm{rank}(\phi_*\big|_x) = n$, there exists neighbour N of x such that $\phi\big|_N$ is a diffeomorphism.

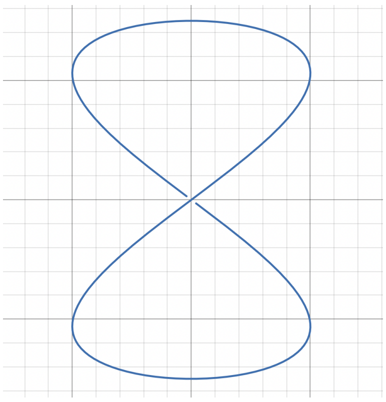
Proof. Suppose (U_1, φ_X) and (U_2, φ_Y) are charts at x and y respectively, where U_1 is a neighbour of x and U_2 is that of y. By definition, the $f = \varphi_Y \circ \phi \circ \varphi_X^{-1}$ forms a C^1-map from $\mathbb{R}^n$ to $\mathbb{R}^m$ for some $0 \leq m \leq n$.
Notice that the φ_X and φ_Y are diffeomorphism, which means that their pushforward are injective. Thus $f_* = (\varphi_X)_* \circ \phi_* \circ (\varphi_X)_*^{-1}$, is injective. By the inverse function theorem for Euclidean spaces, we have the desired statement. □

This suggests that smooth maps with full rank pushforward is a "**local embedding**":

Definition 2.10. Consider two smooth manifolds X and Y, a C^1-map $\phi : X \to Y$ is called **immersion** if its pushforward ϕ_* is an injection. i.e.

$$\operatorname{rank}(\phi_*) = \dim(X)$$

The reason that it defines a local embedding follows directly from the inverse function theorem. But then a good question is that *What does an immersion need for it to be embedding?*



Since immersion is only a local embedding, there could be self-intersecting (like trying to embed the Klein bottle in $\mathbb{R}^3$), so a close one could be "injective". However, it is NOT sufficient:
A well-known counter-example is that $X = (-\pi, \pi)$,$Y = \mathbb{R}^2$ and $\phi(x) = (\sin(2x), \sin(x))$. (in left figure)
But by observing how this counter-example works, we can cook up a sufficient one:

The problem arises in this counter-example is that, ϕ's image of neighbour of $0 \in X$ is not an open set in the subtopology of $\operatorname{Im}(\phi) \subset Y$. So the most obvious solution is that:

Lemma 2.5. If an injective immersion is an open map, then it is an embedding.

which comes immediately by checking the definition of embedding. A less obvious but more useful one is that

Proposition 2.3. A *proper* and injective immersion is embedding.

where a map is **proper** if every compact set in the image has its pre-image compact. (For example, every continuous map from a compact space to a Hausdorff space is proper.)

Proof. Consider a proper and injective immersion $\phi : X \to Y$. By lemma, the only thing that remains to show is that ϕ is an open map.

Consider an arbitrary open set $U \subset X$, and an arbitrary point $y \in \phi(U)$.
What we are going to show is that, there exists an open neighbour $N \subset Y$ of y such that $N \cap \phi(X) \subseteq \phi(U)$. Let's start with a "good" environment: Since ϕ is local embedding, suppose $\phi|_{W_1}$ is an embedding and (W_2, φ) is a local chart of y, and then let $W = W_1 \cap W_2$. Then we try to get the desired N inside W.
Pick an open ball $B_r := \varphi^{-1}(B_r(\varphi(y))) \subset W$ for $r > 0$, and let $K_r = \overline{B_r \cap \phi(X)}$; we know that K_r is compact because $\overline{B_r}$ is compact and K_r is closed.

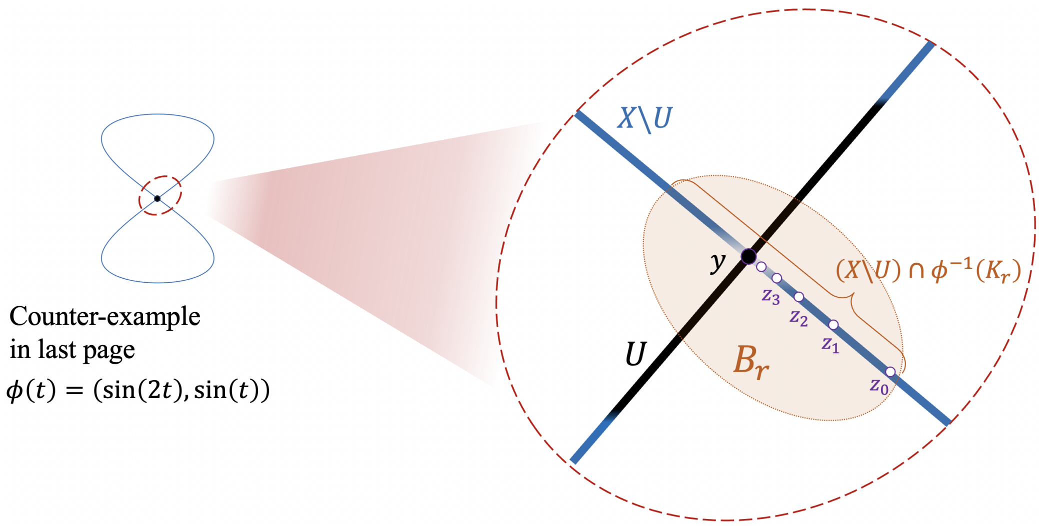


Figure 2: The intuition of this proof comes from observing the reason why our counter-example $\phi(x) = (\sin(2x), \sin(x))$ fails after adding the condition of "proper".

Now prove by contradiction. Suppose $\forall$open neighbour $N \subset Y$ of y, $\exists z \in N \cap \phi(X \setminus U)$. This means that there exists $z_0 \in B_r \cap \phi(X)$ that satisfies $\phi^{-1}(z_0) \notin U$; similarly, there will $\exists z_1, z_2, z_3, \cdots \in N \cap \phi(X \setminus U)$ that satisfy $||\varphi(z_j)|| = 2^{-j}||\varphi(z_0)||$.
Since ϕ is proper, $\phi^{-1}(K_r)$ is compact, which means $(X \setminus U) \cap \phi^{-1}(K_r)$ is compact. Notice that $\left\{\phi^{-1}(Y \setminus \overline{B_{2^{-j}r}}) : j \in \mathbb{Z}_{>0}\right\}$ forms an open cover for $(X \setminus U) \cap \phi^{-1}(K_r)$; the existence of $z_1, z_2, z_3, \cdots$ implies that there is no finite subcover, contradiction. $\square$

However, this could be hard to use if we really wish to show that a specific map is embedding —— "proper" isn't easy to check. Instead, we use a corollary.

Definition 2.1. Given a map between smooth manifolds $\phi \in C^\infty(X, Y)$, point $y \in \mathrm{Im}(\phi)$ is called a **regular value** if ϕ_* is surjection at y.

Corollary 2.1. Given $F \in C^\infty(X, Y)$, the $F^{-1}(\{y\})$ is a submanifold of X when $y \in Y$ is a regular value.

2.4 Vector field and Lie derivative

Based on the tangent space, we can generalized the idea of vector field into manifold.

Definition 2.11. Consider point $p \in M$,

1. **Tangent bundle** is the set $TM = \{(p, v_p) : p \in M \ , \ v_p \in T_pM\}$.
2. **Section** is a map $\pi : M \to TM$ that $p \mapsto (p, v_p)$
3. **Vector Field** is the field of tangent vectors produced by a particular section.
4. The set of sections of a tangent bundle is denoted as $\Gamma(TM)$.

Since the section is essentially a map between manifolds, we can define its smoothness by Definition 2.6; the vector field produced by it is called a **Smooth vector field**, the set of all smooth vector fields is denoted as $\Gamma^\infty(TM)$.
Similarly, we also have the cotangent bundle T^*M and its smooth section $\Gamma^\infty(T^*M)$.
A chart can also provide a coordinate basis for TM and T^*M locally; for example, inside a chart (U, φ), any $F \in \Gamma^\infty(TM)$ can be written in form of

$$F = \sum_{k=1}^{n} f_k \partial_k$$

where $f_k \in C^\infty(M)$ and each of $\partial_k|_p$ a coordinate basis vector of T_pM.

> *Remark* 3. An interesting question is *When does a "global" basis exist?* —— even if a global chart doesn't; for example, both $\mathbb{S}^2$ and $\mathbb{S}^1 \times \mathbb{S}^1$ don't have a global chart, but $T\mathbb{S}^2$ cannot have a global basis, yet $T(\mathbb{S}^1 \times \mathbb{S}^1)$ can!
> One may find that such a global basis exists if and only if
>
> $$TM \cong M \times \mathbb{R}^n$$
>
> This question then generalizes to the topic of *Characteristic classes.* (Section 10)

The idea of tangent bundle and vector field can then be extended to tensors, which are **tensor bundle** and **tensor field**: It is simply to replace the tangent vectors with tensors.

Definition 2.12. The set of ordered pairs

$$\mathcal{T}_r^s M = \{(p, A) \ ; \ p \in M, \ A \in \mathrm{T}_s^r(T_pM)\}$$

denotes the **tensor bundle** of M; a smooth **tensor field** is a smooth map $F : M \to \mathcal{T}_r^s M$. As always in this note, we directly interpret F as a map that $p \mapsto A$, rather than $p \mapsto (p, A)$ for simplicity.

An important example of a tensor field is the Riemannian metric.

Definition 2.13. A **Riemannian metric** g is a smooth tensor field such that $g|_p$ is a positive-definite inner product on T_pM for $\forall p \in M$.

Later, we'll often write $\langle u, v\rangle := g(u, v)$.
In the last part of this section, we will prove that M can always be embedded into $\mathbb{R}^m$ for some $m \geq n$, so every Riemannian metric in $\mathbb{R}^m$ naturally induces one on M by pullback:
Given $\phi \in C^\infty(M, \mathbb{R}^n)$ and $\langle\cdot,\cdot\rangle$ a Riemannian metric on $\mathbb{R}^m$, then

$$g(u, v) := \langle \phi_* u, \phi_* v\rangle \quad \forall u, v \in T_pM \tag{2.3}$$

gives one on M. On the other hand, when $\phi \in C^\infty(M, \mathbb{R}^n)$ satisfies (2.3), we call ϕ a **local isometry**; and if ϕ is moreover a diffeomorphism, we call it **Isometry**.

Consider $U, V \in \Gamma^\infty(TM)$.
Since tangent vectors are defined to be directional derivatives on $C^\infty(M)$, can they act on the vector field as derivatives? i.e. $V(U)$? No, because it's NOT well-defined —— not independent of the choice of coordinate.
The problem is that both U, V are changing, so we need to know how to compare values of U at nearby points. One way to achieve this is by letting the local basis "flow" with V, then pushforward U in that basis back to the origin for comparison; in other words, "pushforward then differentiate". This is the idea of Lie derivative.
Consider an ODE

$$x_* = V \quad \text{where } x(0) = p$$

for $x : \mathbb{R} \to M$. Now let $X_t(p) = x(t)$ (which is the flow), the "Rate of change of U with respect to V" is then intuitively

$$\mathcal{L}_V U := \lim_{t\to 0} \frac{1}{t}\left((X_t)_*^{-1} U_{\gamma(t)} - U_p\right) \tag{2.4}$$

where $\gamma(t) = X_t(p)$. This is known as the **Lie derivative**[4].
But there is room for simplifying. Choose a local chart (N, φ), such that $V|_{(x_1,\cdots,x_n)} = h(x_1)\partial_1$ for some $h \in C^\infty(\mathbb{R}, \mathbb{R}_{>0})$ in the induced coordinate basis. (try proving this; it's proved in Question & Answer)
In this case, if $\varphi(q) = (x_1, \cdots, x_n)$ and $X_{\tau j}$ denotes the j^{th} component of $X_\tau|_q$, then

$$\varphi\left(X_t|_q\right) = \left(x_1 + \int_0^t h\left(X_{\tau 1}\right) \mathrm{d}\tau,\ x_2, \cdots, x_n\right)$$

[4]You might worry that $(X_t)_*^{-1}$ could be undefined in (2.4). Notice that $\lim_{t\to 0}(X_t)_* = I$. So since the flow of a smooth ODE is locally smooth, there exists $\delta > 0$ such that $((X_t)_* - I)$ is sufficiently small, and $(X_t)_*^{-1}$ is thus defined with $0 \leq t < \delta$.

Observe that under the coordinate basis, we have the matrix expression

$$\left.(X_t)_*\right|_q = \operatorname{diag}\left(1 + \frac{\partial}{\partial x_1}\int_0^t h\left(X_{\tau 1}\right) \mathrm{d}\tau,\, 1, \cdots, 1\right)$$

therefore by letting $U = \sum_{j=1}^n U_j \partial_j$, we have

$$\begin{aligned}
\mathcal{L}_V U &= \lim_{t\to 0}(X_t)_*^{-1} \circ \lim_{t\to 0} \frac{U_{\gamma(t)} - U_p}{t} \;+\; \lim_{t\to 0} \frac{(X_t)_*^{-1} - I}{t} \circ U_{\gamma(t)} \\
&= \left(\sum_{j=1}^n h \frac{\partial U_j}{\partial x_1} \partial_j\right) + \lim_{t\to 0} \frac{(X_t)_*^{-1} - I}{t} \circ \lim_{t\to 0} U_{\gamma(t)} \\
&= \left(\sum_{j=1}^n h \frac{\partial U_j}{\partial x_1} \partial_j\right) - U_1 \frac{\partial h}{\partial x_1} \partial_1 \\
&= V(U) - U(V)
\end{aligned}$$

So it *might be*[5] the case that $\mathcal{L}_V U = V(U) - U(V)$:

Definition 2.14. For $X, Y \in \Gamma^\infty(TM)$, the **Lie derivative** of Y with respect to X is the operator $\mathcal{L}_X(Y) \in \Gamma^\infty(TM)$, defined by

$$\begin{aligned}
\left(\mathcal{L}_X Y\right)(f) &= [X, Y] \\
&:= X(Y(f)) - Y(X(f))
\end{aligned}$$

for $\forall f \in C^\infty(M)$. The $[X, Y]$ is called the Lie bracket.

To prove that it is consistent with the definition given before, it's sufficient to show that

Proposition 2.4. Given $U, V \in \Gamma^\infty(TM)$, the $\mathcal{L}_V U \in \Gamma^\infty(TM)$

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

$$\begin{aligned}
V(U(f)) - U(V(f)) &= \sum_{i=1}^n \sum_{j=1}^n V_i \frac{\partial}{\partial x_i}\left(U_j \partial_j(f)\right) - \sum_{i=1}^n \sum_{j=1}^n U_i \frac{\partial}{\partial x_i}\left(V_j \partial_j(f)\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \left(V_i \frac{\partial U_j}{\partial x_i} - U_i \frac{\partial V_j}{\partial x_i}\right) \partial_j(f)
\end{aligned}$$

which is a linear combination of vector fields acting on f, so $\mathcal{L}_V U \in \Gamma^\infty(TM)$. Therefore, since vector fields are independent of the choice of atlas, our previous proof to $\lim_{t\to 0} \frac{1}{t}\left((X_t)_*^{-1} U_{\gamma(t)} - U_p\right) = [V, U]$ using chart (N, φ) is true for all atlas. ☐

This is quite interesting that such an operation on vector fields will still produce a vector field. To compute it, we only need this observation:

[5] Showing that it is true in one chart doesn't imply it is true in all the others. For example, what if $\mathcal{L}_V U = V(U) - U(V) + F$, while $F = 0$ under the chart we chose?

Lemma 2.6. Given a chart (U, φ), and $\{\partial_j\}$ is the coordinate basis induced by it;

$$[\partial_j, \partial_k] = 0 \quad \text{for all } j, k = 1, \cdots, n$$

(which is trivial) and since the Lie bracket is bilinear, we can compute Lie derivatives by locally writing them in components.

The Lie derivative (or to say Lie bracket) becomes more interesting when M is a **Lie group**, which a group that is also a smooth manifold, such that group multiplication and taking inverses are both smooth maps.
The first interesting fact about M being a Lie group is that,

Proposition 2.5. When M is a Lie group, $TM \cong M \times \mathbb{R}^n$.

Proof. Denote $\mathfrak{i} \in M$ as the identity of M.
Since any $y \in M$ defines a smooth map $y : x \mapsto yx$ on M, there exists a neighbour U of $\mathfrak{i}$ such that $y|_U$ is a homeomorphism. So, by picking a local basis $\{u_1, \cdots, u_n\}$ for $T_{\mathfrak{i}}M$, we get a global basis where $T_yM = \text{span}(y_*u_1, \ldots, y_*u_n)$. □

So there exists global basis $\{\partial_1, \ldots, \partial_n\}$ for TM and thus $\Gamma^\infty(TM)$ is the $C^\infty(M)$-span of $\{\partial_1, \ldots, \partial_n\}$. However, $\{\partial_1, \ldots, \partial_n\}$ might not give a coordinate basis; in other words, it's possible that $[\partial_j, \partial_k] \neq 0$ for $j \neq k$. Clearly if there doesn't exists a set of $\{\partial_1, \ldots, \partial_n\}$ such that $[\partial_j, \partial_k]$ is constantly zero, then M cannot be a subset of $\mathbb{R}^n$; this observation give rise to a way of classifying M by studying an algebra:

Proposition 2.6. Suppose the M is also a Lie group, and $\{\partial_1, \ldots, \partial_n\}$ a global basis of TM. Let

$$\mathfrak{m} := \left\{ \sum_{k=1}^{n} a_k \partial_k : a_k \in \mathbb{R} \right\}$$

then $(\mathfrak{m}, +, [\cdot, \cdot])$ forms an algebra, the **Lie algebra** of M.

which is quite easy to verify.
A particularly natural example would be that when $M \leq \text{GL}_n(\mathbb{R})$. In this case, $\mathfrak{m} \subset \text{M}_n(\mathbb{R})$; then the Lie derivative is defined literally (i.e. $[X, Y] = XY - YX$ is given by matrix multiplication). Since Lie algebra only concerns about the $\mathbb{R}$-linear span of $\{\partial_1, \ldots, \partial_n\}$, for any $A \in \mathfrak{m} \subset \Gamma(TM)$, we only need to look at its value at a single point, for example, $\mathfrak{i}$.

Corollary 2.2. When $M \leq \text{GL}_n(\mathbb{R})$, $\mathfrak{m} = (T_I M, +, [\cdot, \cdot])$

A more interesting fact is that, in this case, we have

$$M = \exp(\mathfrak{m}) = \{\exp(A) : A \in \mathfrak{m}\}$$

where $\exp(-)$ is the matrix exponential. Proof left as an exercise.

Finally, we can generalize the pushforward & pullback to tensor fields; the construction of isometry (2.3) already provides a clue. The simplest case is that

Definition 2.2. Consider a smooth map ϕ from smooth manifold X to Y. Suppose $u \in T_x^*X$, then its **pullback** $\phi^*u \in T_{\phi(x)}^*Y$ is defined by

$$\left.(\phi^* u)\right|_{\phi(x)}(v) = \left.u\right|_x(\phi_* v)$$

for $\forall v \in T_{\phi(x)}M$.

Then for a general tensor field $T \in \mathcal{T}_s^r(TX)$, since it defines a multilinear map $(TX)^r \times (T^*X)^s \to \mathbb{R}$, we define its pushforward $\phi_*T : (TY)^r \times (T^*Y)^s \to \mathbb{R}$ to be

$$(\phi_* T)(v_1, \ldots, v_r, u_1, \ldots, u_s) := T(\phi_* v_1, \ldots, \phi_* v_r, \phi^* u_1, \ldots, \phi^* u_s)$$

for $\forall v_j \in \Gamma^\infty(TM)$ and $u_j \in \Gamma^\infty(T^*M)$.

The Lie derivative can also be generalized for other tensors, but unluckily it is NOT a tensor, as it is not a tensor at any single point of M. Therefore the method is different.
We start with the original idea of Lie derivative, that it is the differentiation after pushforward:

Definition 2.15. For $T \in \mathcal{T}_s^r(TM)$ and $V \in \Gamma(TM)$, we define

$$\mathcal{L}_V T := \lim_{t\to 0} \frac{1}{t}\left((X_{-t})_* \left.T\right|_{\gamma(t)} - \left.T\right|_p\right)$$

But to write it down it explicitly, we start with the case of $\Gamma(T^*M)$ as before. For $V \in \Gamma^\infty(TM)$ and $\omega \in \Gamma(T^*M)$, our definition dictates

$$\mathcal{L}_V \omega = \lim_{t\to 0} \frac{1}{t}\left(X_t^* \left.\omega\right|_{\gamma(t)} - \left.\omega\right|_p\right)$$

where $\gamma(t) := X_t(p)$ and $X_t(p)$ is the flow of the ODE $x_* = V$ with $x(0) = p$ as before.
So for any $F \in \Gamma(TM)$, we have

$$\begin{aligned}
(\mathcal{L}_V \omega)F &= \lim_{t\to 0} \frac{1}{t}\left(\left.\omega\right|_{\gamma(t)}((X_t)_* \left.F\right|_p) - \left.\omega\right|_p \left.F\right|_p\right) \\
&= \lim_{t\to 0} \left.\omega\right|_{\gamma(t)} \circ \lim_{t\to 0} \frac{1}{t}\left((X_t)_* \left.F\right|_p - \left.F\right|_{\gamma(t)}\right) + \lim_{t\to 0} \frac{1}{t}\left(\left.\omega\right|_{\gamma(t)} \left.F\right|_{\gamma(t)} - \left.\omega\right|_p \left.F\right|_p\right) \\
&= V(\omega(F)) - \omega(\mathcal{L}_V F)
\end{aligned}$$

so using the same logic, we have

$$\begin{aligned}(\mathcal{L}_V F)(A_1,\ldots,A_r,B_1,\ldots,B_s) &= V\left(F(A_1,\ldots,A_r,B_1,\ldots,B_s)\right) \\ &\quad -\sum_{j=1}^{r} T\left(A_1,\ldots,\mathcal{L}_V A_j,\ldots,A_r,B_1,\ldots,B_s\right) \\ &\quad -\sum_{j=1}^{s} T\left(A_1,\ldots,A_r,B_1,\ldots,\mathcal{L}_V B_j,\ldots,B_s\right)\end{aligned}$$

for $A_1,\ldots,A_r \in \Gamma(TM)$ and $B_1,\ldots,B_s \in \Gamma(T^*M)$

2.5 Partition of unity

Technical proofs in this section has no direct further applications in this book except for 2.6. In this book, the word "POU" stands for Partition of Unity.

For what we learned up to now, one application of POU is for the existence of (smooth) Riemannian metric g. More generally, it is a technique to make sure functions with local properties exist globally.

Definition 2.16. Suppose $\{U_j\}$ is an open cover of M, and the set of functions $P = \{f_j \in C^\infty(M) : 0 < f_j < 1\}$ satisfies

1. $\text{supp}(f_j) \subset U_j$ for $\forall j \in \mathbb{N}$.
2. For $\forall p \in M$, only finite number of $\text{supp}(f_j)$ contains p.
3. $\sum_{f_j \in P} f_j = 1$ at $\forall p \in M$.

then P is a **Partition of unity** subordinate to $\{U_j\}$.

where $\text{supp}(f) = \overline{\{x \in M : f(x) \neq 0\}}$, we say f is **support**ed on $\text{supp}(f)$.
Basically, POU is a bunch of locally supported functions that sum to 1 on the entire manifold, so it can be used to extend functions with good local property into a global one by multiplying its local pieces with corresponding locally supported functions in POU.
This works because interestingly POU exists on every smooth manifold:

Proposition 2.7. Any open cover of M has a POU subordinate to it.

Before proving it, we can first have a look at its most direct application:

Corollary 2.7. There exists a (smooth) Riemannian metric on M.

Proof. Consider an atlas $\{(U_j, \varphi_j)\}$, which forms an open cover on M; by Proposition 2.7, we can construct POU $\{f_j\}$ subordinate to $\{U_j\}$. By definition of POU, for each point $p \in M$ there exists a finite subset $J \subset \mathbb{N}$ such that $p \in \text{supp}(f_j)$ iff $j \in J$. Therefore,

$$g := \sum_{j \in J} f_j \cdot \left(\varphi_j^* g_{\text{eu}}\right)$$

is a (smooth) Riemannian metric on M, where g_{eu} is a Riemannian metric on $T\mathbb{R}^n$; this is because a finite sum of smooth functions is still smooth and a sum of positively defined inner product is still positively defined.
Hence the existence of POU guarantees the existence of a smooth Riemannian metric. $\square$

To prove Proposition 2.7, we need to introduce two useful concepts,

Definition 2.17. Exhaustion sets is a countable open cover $\{X_j\}_{j\in\mathbb{N}}$ of M such that each $\overline{X_j}$ is compact and $\overline{X_j} \subset X_{j+1}$.
Locally finite refinement of an open cover $\mathcal{U} = \{U_\alpha\}$ of M is a countable open cover $\{W_j\}_{j\in\mathbb{N}}$ such that each W_j is contained in some U_α, and at each $p \in M$, $\exists$neighbour N_p that intersects only finitely many W_j.

These two concepts are extremely useful in proving topological statements concerning manifolds, and they always exist:

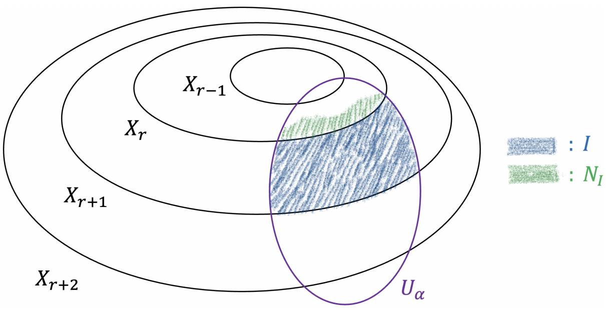
Lemma 2.8. Any M has exhaustion sets, and any open cover of M has a locally finite refinement.

Existence of Locally finite refinement can be proved by existence of an Exhaustion set.

$\exists$ **Exhaustion sets**: First notice that for each point $p \in M$ there exists a compact set $B_p \subset M$ containing it; so we can have an open cover $\{U_p\}$ of M such that each U_p is contained in B_p and each $\overline{U_p}$ is obviously compact. Since M is second countable, we can then choose a sequence $\{U_j\}_{j\in\mathbb{N}} \subset \{U_p\}$ that covers M.
Now for an open set X_1, we can take a finite number of subsets from $\{U_j\}_{j\in\mathbb{N}}$ that covers $\overline{X_1}$ (since it must be compact); then for $\forall i \in \mathbb{Z}_{>0}$, take $X_{i+1} = U_i \cup X_i$, which gives us a sequence X_i with each $\overline{X_i}$ compact. So we always have an Exhaustion set.

then based on the exhaustion sets, Locally finite refinement seems obvious:
$\exists$ **Locally finite refinement**: Consider the exhaustion sets $\{X_r\}_{r\in\mathbb{N}}$; for each $r \in \mathbb{Z}_{>0}$, take $\mathcal{I}_r := \{U_\alpha \cap X_{r+1} \setminus X_r : U_\alpha \in \mathcal{U}\}$ and $\mathcal{N}_r := \{I \cup N_I : I \in \mathcal{I}_r\}$, where N_I is the union of a set of neighbours of points on $\partial X_r \cap I$ that doesn't intersect X_{r-1}. (the N_I exists because of Hausdorff.) Now since $\overline{X_{r+1}}$ is compact, there exists a finite subset $\mathcal{W}_r \subset \mathcal{N}_r$ that covers $X_{r+1} \setminus X_r$. The $\bigcup_{r\in\mathbb{N}} \mathcal{W}_r$ is then a locally finite refinement of $\mathcal{U}$.

□



Finally we can then prove the existence of POU (Proposition 2.7).

Proof of Proposition 2.7:
First we need to show that for any open subset U of $\mathbb{R}^n$ and $p \in U$, there $\exists$ a function on U

with $p \in$ supp and supp compact. Such function can be constructed by following:

$$\text{Let } q_1(x) = \begin{cases} e^{-\frac{1}{x}} \text{ for } x > 0 \\ 0 \text{ otherwise} \end{cases} \quad ; q_2(x) = \frac{q_1(x)}{q_1(x) + q_1(1-x)}$$

where we can verify that $q_1, q_2 \in C^\infty(\mathbb{R})$, and $\mathrm{supp}(q_2) = \mathbb{R}_{\geq 0}$.
Then we can define $\beta_{a,r} \in C^\infty(\mathbb{R}^n)$ satisfying $\mathrm{supp}(\beta_{a,r}) = \overline{B_a(r)}$ by

$$\beta_{a,r}(x) = q_2\left(\frac{1 - ||x-a||^2}{1 - r^2}\right)$$

this is called a **Bump function**.
By choosing appropriate a, r and summing up corresponding $\beta_{a,r}$ together, we are able to get the desired function.

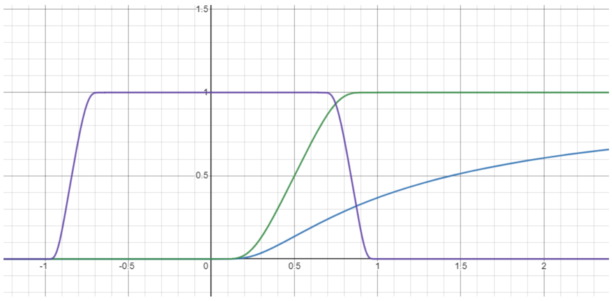


Figure 3: q_1,q_2 and β constructed

Consider an atlas $\{(U_p, \varphi_p)\}$ on M, and let $\{W_j\}_{j\in\mathbb{N}}$ be a Locally finite refinement of it. Then by our construction at the beginning of the proof, we can choose $\{\rho_j \in C^\infty(W_j)\}_{j\in\mathbb{N}}$ which satisfies that: (1), Each $\mathrm{supp}(\rho_j)$ is compact; (2), $f := \sum_{j\in\mathbb{N}} \rho_j$ is non-zero everywhere. ((2) can be satisfied because every $p \in M$ is covered by at least one W_j.)
Set $f_j = \rho_j / f$, we have that $\{f_j\}$ is a POU subordinated to $\{W_j\}$. □

The concept of POU can be easily generalized into ordinary topological spaces, just by changing the C^∞ into C^0. And by observing the axioms which we used for proving the existence of POU on M, we have the generalization below:

Corollary 2.3. Topological space X is paracompact and Hausdorff if and only if every open cover has a POU subordinate to it.

Food for thought.

2.6 Whitney embedding theorem

In this section, we derive a smooth manifold M and its properties intrinsically; but intuitively, viewing a manifold as a hyper-surface by embedding it in some $\mathbb{R}^m$ produces similar and more intuitive conclusions.
So does an embedding always exists? The answer turns out to be YES:

Whitney Embedding Theorem

Theorem 2.9. An n-dim smooth manifold can be embedded into $\mathbb{R}^{2n+1}$.

Though this is not the best result, so it is actually called *weak* Whitney embedding theorem, as this $2n + 1$ can be improved to $2n$. However, I haven't found a proof for this improved version with current knowledge.
Before starting the proof, we need to introduce a famous lemma:

Lemma 2.10. (Sard's Theorem) Suppose X, Y are smooth manifolds, $\dim(Y) = n$ and map $f : X \to Y$ is smooth; we have that

$$\lambda\left(\varphi_y \left\{y \in Y : \operatorname{rank}(f_*)|_y < n\right\}\right) = 0$$

where f_* is the pushforward of f, and λ is Lebesgue measure, and φ_y is a chart function at a neighbour of y.

In short, it says that the set of critical points of a smooth map between manifolds should be of measure zero. It can be proved by considering the case of $Y = \mathbb{R}^n$, which turns it into an analysis problem, and the rest just follows easily.

Now consider a n-dimensional smooth manifold M.
The first part of the proof uses induction, it first proves that there exists N such that M has an injective immersion into $\mathbb{R}^N$, then uses induction to lower the dimension to $2n + 1$. And finally, by using Proposition 2.3, we convert those immersions into embeddings.
Let's first look at the induction part.

Proposition 2.8. For $m > 2n + 1$, if there exists an injective immersion Φ from M to $\mathbb{R}^m$, then there must be one from M to $\mathbb{R}^{m-1}$.

Proof. The $\mathbb{R}^{m-1}$ can be embedded in $\mathbb{R}^m$ as a hyperplane passing through the origin, which is an element of $\{P_{[v]} : [v] \in \mathbb{RP}^{m-1}\}$ where

$$P_{[v]} = \{x \in \mathbb{R}^m : x \cdot v = 0\} = (\mathbb{R}v)^\perp$$

for $v \in \mathbb{R}^m$ and $[v] = \mathbb{R}v \in \mathbb{RP}^{m-1}$.

So we can construct a map $\Phi_{[v]} = \pi_{[v]} \circ \Phi$ where $\pi_{[v]} : \mathbb{R}^m \to P_{[v]}$ is the orthogonal projection of $\mathbb{R}^m$ on $P_{[v]}$; we will show that the set of $[v]$ for which $\Phi_{[v]}$ is not an injective immersion is of measure zero on $\mathbb{RP}^{m-1}$.

Suppose $\Phi_{[v]}$ is NOT injective, then there exists $p_1 \neq p_2$ where $[\Phi(p_1) - \Phi(p_2)] \in [v]$, which means all $[v]$ that make $\Phi_{[v]}$ not injective is in the image of map

$$f : M \times M \setminus \triangle_M \to \mathbb{RP}^{m-1}$$
$$\text{where } f(p_1, p_2) = [\Phi(p_1) - \Phi(p_2)]$$

where $\triangle_M = \{(p,p) : p \in M\}$. Now because $M \times M \setminus \triangle_M$ is a manifold and f is a smooth map, by Sard's theorem, $\mathrm{Im}(f)$ is of measure zero in $\mathbb{RP}^{m-1}$ because $\dim(M \times M \setminus \triangle_M) = 2n < m-1 = \dim(\mathbb{RP}^{m-1})$. So since $[v] \in \mathrm{Im}(f)$, this suggests that $\Phi_{[v]}$ is injective for almost all $[v] \in \mathbb{RP}^{m-1}$.

Similarly, Suppose $\Phi_{[v]}$ is NOT immersion, then there exists $p \in M$ such that $\mathrm{rank}((\Phi_{[v]})_*)\big|_p = \mathrm{rank}((\pi_{[v]})_* \circ \Phi_*)\big|_p = 0$, which means $\exists u_p \in T_pM$ such that

$$\left((\pi_{[v]})_* \circ \Phi_*\right)\Big|_p (u_p) = 0 = \pi_{[v]}\left(\Phi_*|_p (u_p)\right)$$

which means $\left[\Phi_*|_p (u_p)\right] = [v]$. Hence similarly, $[v]$ is in the image of map

$$g : TM \to \mathbb{RP}^{m-1}$$
$$\text{where } g(p, u_p) = \left[\Phi_*|_p (u_p)\right]$$

And $\dim(TM) = 2n < m - 1 = \dim(\mathbb{RP}^m)$, so apply Sard's theorem again, we have that the set of $[v]$ for which $\Phi_{[v]}$ is not immersion is of measure zero.

The union of two sets of measure zero must be of measure zero, so for almost all $[v] \in \mathbb{RP}^{m-1}$, the $\Phi_{[v]}$ is an injective immersion. □

Then what we need is to show the existence of the base case, that there $\exists N \in \mathbb{Z}_{>0}$ such that in which M has an injective immersion. The case of compact M would be easy: Suppose M has atlas $\{(N_p, \varphi_p)\}$, there exists a finite subcover $\{N_j\}_{1\leq j\leq r} \subset \{N_p\}$ of M. Then an injective immersion $\Phi : M \to \mathbb{R}^{rn+r}$ could be given by

$$\Phi(p) := (f_1\varphi_1, \cdots, f_r\varphi_r, f_1, \cdots, f_r)$$

where f is a POU subordinate to $\{N_j\}_{1\leq j\leq r}$.

We can easily verify that it is injective; it is also immersion, since for $\forall v \in \Gamma^\infty(TM)$, if

$$\Phi_* v|_{\Phi(p)} = (v(f_1)\varphi_1 + f_1\varphi_{1*}(v), \cdots, v(f_r)\varphi_r + f_r\varphi_{r*}(v), v(f_1), \cdots, v(f_r)) = 0$$

then $\varphi_{i*}(v) = 0$ for some $1 \leq i \leq r$, which means $v = 0$ as φ_i are all diffeomorphisms.

In other words,

Proposition 2.11. If M can be covered by a finite number of charts, then there exists an injective immersion from M to $\mathbb{R}^{2n+1}$.

But how to extend it into arbitrary M? This is the most interesting constructive step in this version of proof, and the technique it used will also be used later in the proof of **de Rham theorem**.

Proposition 2.9. There exists an injective immersion from M to $\mathbb{R}^{2n+1}$.

Proof. Consider the POU $\{(\rho_j, W_j)\}$ on M and a monotone sequence $\{a_j\}_{j\in\mathbb{N}} \subset \mathbb{R}_{>0}$ that $\lim_{j\to\infty} a_j = \infty$. Easy to see that the function

$$\rho(p) := \sum_{j=0}^{\infty} a_j \rho_j(p)$$

is smooth, and it satisfies that

$$\rho^{-1}[0, c] \subset \bigcup_{j=1}^{m} \operatorname{supp}(\rho_j)$$

for $\forall c > 0$, where $m = \min_{a_k > c}(k)$. So we can deduce that $\rho^{-1}[0, c]$ is compact, because each $\operatorname{supp}(\rho_j)$ is compact and $\rho^{-1}[0, c]$ is closed.
So consider an atlas $\{(N_{i\alpha}, \varphi_{i\alpha})\}$ of each $\rho^{-1}[i, i+1]$ for $i \in \mathbb{N}$, we are able select a finite subset $\{N_{i1}, \cdots, N_{im_i}\}$ from $\{N_{i\alpha}\}$ that covers $\rho^{-1}[i, i+1]$ (as it's compact). In this case, we can cover M by $\{M_i\}_{i\in\mathbb{N}}$, where

$$M_i := \rho^{-1}\left(i - \frac{1}{2}, i + \frac{3}{2}\right) \cap \bigcup_{j=1}^{m_i} N_{ij}$$

By Proposition 2.11, we have that each M_i has an injective immersion ϕ_i into $\mathbb{R}^{2m+1}$. Notice that $M_i \cap M_j = \emptyset$ iff $|i - j| > 1$; so we can consider the map $\Phi : M \to \mathbb{R}^{4m+4}$ by

$$\Phi(p) := \left(\sum_{i=0}^{\infty} \beta_{2i}\phi_{2i}, \sum_{i=0}^{\infty} \beta_{2i+1}\phi_{2i+1}, \sum_{i=0}^{\infty} \beta_{2i}, \sum_{i=0}^{\infty} \beta_{2i+1}\right)$$

(ϕ_i maps to the same $\mathbb{R}^{2m+1}$ for i of the same parity) where each β_i is bump function with $\operatorname{supp}(\beta_i) \subset M_i$ and $\beta_i(\rho^{-1}[i, i+1]) = \{1\}$. It is a very clever construction though using a similar principle as the compact case; we can verify that

1. Φ is injective: which is obvious;
2. Φ is immersion: suppose $p \in M_i$ and $v \in \Gamma^\infty(TM)$, then by similar logic to the compact case, $\Phi_* v|_{\Phi(p)} = 0$ suggests $\beta_i \phi_{i*} v = 0$, which gives $v = 0$, so Φ is immersion.

Thus Proposition 2.8 is true for arbitrary M, and we are done by induction. □

Remark. The function ρ in this proof is called the **Exhaustion function**; our arguments in the proof suggest that it is a proper function.

Also, as we said, the technique this proof uses will be useful later; this is basically because this proof reveals a *general* property of smooth manifolds:

Corollary 2.4. Suppose $\mathfrak{P}$ is a statement for smooth manifolds, and it is true for:

1. Euclidean spaces,
2. A union of finitely many smooth manifolds that satisfy $\mathfrak{P}$,
3. Any countable disjoint union of smooth manifolds satisfying $\mathfrak{P}$;

Then $\mathfrak{P}$ is true for all smooth manifolds.

So now we have that all M can have an injective immersion into $\mathbb{R}^{2n+1}$; what we do is to turn our injective immersion into embedding;

Proposition 2.10. There exists an embedding from M to $\mathbb{R}^{2n+1}$.

Proposition 2.3 tells us that all we need is to make our injective immersion proper, and we do so by using an exhaustion function, which happens to be a proper map itself.
So the strategy here is this:
1, Use the exhaustion function ρ to "level" the image of Φ;
2, Then use the boundedness of compact sets to show that every compact set in $\mathbb{R}^{2n+1}$ has its pre-image contained in $\rho^{-1}[0, \beta]$ for some $\beta \in \mathbb{R}$.
Let's see how it's done in practice.

Proof. Let $\widetilde{\Phi} : M \to \mathbb{R}^{2n+1}$ to be the injective immersion gained from previous proposition, and define a "shrinked" version of it:

$$\Phi(p) = h\left(\widetilde{\Phi}(p)\right) \text{ , where } h(x) = \frac{x}{|x|^2 + 1}$$

This is still an injective immersion (as h is an injective immersion), and $||\Phi(M)|| < 1$.
Now we use the exhaustion function to "level" M in $\mathbb{R}^{2n+1}$:
First, we define $\Phi_l(p) = (\Phi(p), \rho(p))$, which is an injective immersion into $\mathbb{R}^{2n+2}$; then let

$$\Psi_{[v]} = \pi_{[v]} \circ \Phi_l$$

(where $v \in \mathbb{R}^{2n+2}$) which is, explicitly,

$$\Psi_{[v]}(x) = \begin{bmatrix} \Phi(x) - ((\Phi(x) \cdot v') + v_{2n+2}\rho(x))\, v' \\ \rho(x)\left(1 - v_{2n+2}^2\right) - (\Phi(x) \cdot v')v_{2n+2} \end{bmatrix}$$

where $v' \in \mathbb{R}^{2n+1}$ is the first $(2n+1)$ components of v.
As Φ_l is an injective immersion, by the same argument used in the proof of Proposition 2.8, we have that $\underline{\Psi_{[v]} \text{ is an injective immersion into } P_{[v]} \cong \mathbb{R}^{2n+1}}$ for almost all $[v] \in \mathbb{RP}^{2n+1}$.
Our task is then to show that this $\Psi_{[v]}$ is proper.
Take any compact set $S \subset P_{[v]}$; it should be bounded, which means $\exists b > 0$ such that $S \subset$

$B_b(0)$. This suggests that all $x \in \Psi_{[v]}^{-1}(S)$ should satisfy

$$\left|\left(\Psi_{[v]}(x)\right)_{2n+2}\right| = \left|\rho(x)\left(1 - v_{2n+2}^2\right) - (\Phi(x) \cdot v')v_{2n+2}\right| \leq b$$

Notice that $||\Phi(x)||, ||v'||, ||v_{2n+1}|| < 1$, which means

$$|\rho(x)| \leq \frac{b+1}{1 - v_{2n+2}^2}$$

If we denote the RHS as β, then we have that $x \in \rho^{-1}[0, \beta]$. This suggests that $\Psi_{[v]}^{-1}(S) \subset \rho^{-1}[0, \beta]$; and since $\Psi_{[v]}^{-1}(S)$ is closed (as $\Psi_{[v]}$ is continuous), $\Psi_{[v]}^{-1}(S)$ is compact.
Hence, any compact set S has pre-image $\Psi_{[v]}^{-1}(S)$ compact, so $\Psi_{[v]}$ is proper, and thus an embedding. $\square$

2.7 Application after being embedded

Recall that the purpose of embedding is that our physical intuition becomes useful; so let's consider some applications brought by such intuition.

The first thing is that "exterior objects" like the normal vector are well-defined: suppose $M \subset \mathbb{R}^m$ is a n-dim smooth manifold with $n < m$, we can define the *normal bundle* to be $\{(p, \mathrm{N}_p M) : p \in M\}$, where

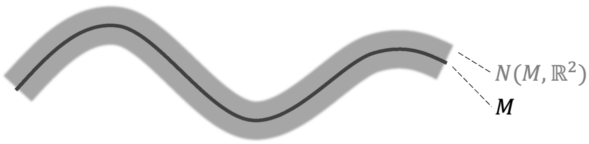
$$\mathrm{N}_p M := \{v \in \mathbb{R}^m : v \perp T_p M\}$$

We can intrinsically define it by viewing $(T_p M)^\perp \cong \mathbb{R}^m/(T_p M)$. In this way, we can give a more general definition:

Definition 2.3. For manifolds $X \subset M$, we define **Normal bundle** of X in M to be

$$\mathrm{N}(X, M) := \{(x, \mathrm{N}_x(X, M)) : x \in X\}$$

where $\mathrm{N}_x(X, M) := T_x M / T_x X$ is the normal space at x.



Like shown in the figure, the normal bundle is like "hair" growing on the manifold, where the "hair" are the vector space spanned by the normal vector (or normal space); interestingly this intuition can be formalized mathematically:

Proposition 2.11. (*Tubular neighbourhood theorem*) Suppose X is a submanifold of M, then there exists an open neighbour N of X such that

$$N \cong \mathrm{N}(X, M)$$

This comes from an intuitive principle: (Embed $M \subset \mathbb{R}^m$) For $\forall p \in M$, there will $\exists \varepsilon > 0$ such that any $x \in \mathbb{R}^m$ who satisfies $\inf_{q \in M} ||x - q|| = ||x - p|| = \varepsilon$ would have $(x - p) \in \mathrm{N}_p(M, \mathbb{R}^m)$; intuitively, this is to say that, *if your hair is short enough, then the shape of your hair is the same as the shape of your scalp.*

Formally, this is to say that:

Lemma 2.12. Suppose $M \subset \mathbb{R}^m$ with $n < m$, and consider

$$M_\varepsilon := \left\{ x \in \mathbb{R}^m : \inf_{p \in M} ||x - p|| < \varepsilon(x) \right\}$$

There $\exists \varepsilon : M \to \mathbb{R}_{>0}$ such that $M_\varepsilon \cong \mathrm{N}(M, \mathbb{R}^m)$.

The set M_ε is called the ε-**neighbourhood** of M.
The proof is a small application of POU; left as an exercise.
The ε-neighbourhood & Tubular neighbourhood are useful when dealing with problems concerning approximations, and there are two examples: *Whitney approximation theorem* and *Thom transversality theorem*.
Let's start with the first one, which is a corollary of this analysis-like statement:

Lemma 2.13. Given any $f \in C^0(M, \mathbb{R}^m)$, we have that for arbitrary $\varepsilon : M \to \mathbb{R}_{>0}$, there exists $f_\infty \in C^\infty(M, \mathbb{R}^m)$ such that

$$||f(p) - f_\infty(p)|| < \varepsilon(p) \quad \text{for } \forall p \in M$$

Proof. For $\forall x \in M$, there exists neighbour U_x such that

$$||f(U_x) - f(x)|| < \varepsilon(x)$$

the $\{U_x : x \in M\}$ gives an open cover of M, which admits a locally finite refinement $\{U_{x_k}\}$, for some $x_1, x_2, x_3, \cdots \in M$. Then by assigning a corresponding POU $\{\rho_k\}$, we can construct the map

$$f_\infty(x) := \sum_k \rho_k(x) \cdot f(x_k)$$

which is smooth, and satisfies that

$$\begin{aligned} ||f(x) - f_\infty(x)|| &= \left\| \sum_k \rho_k(x) \left(f(x_k) - f(x) \right) \right\| \\ &\leq \sum_k \rho_k(x) \left\| f(x_k) - f(x) \right\| \\ &< \varepsilon(x) \end{aligned}$$

□

We can then upgrade this in a more topological sense, that C^0-maps between manifold can always be approximated by C^∞-maps.

Proposition 2.12. *(Whitney approximation theorem)* For any $f \in C^0(X, Y)$, there exists $f_\infty \in C^\infty(X, Y)$ homotopic to f.

Proof. Embed Y into $\mathbb{R}^m$; let Y_ε be an ε-neighbourhood of Y, and $\phi : Y_\varepsilon \to \mathrm{N}(Y, \mathbb{R}^m)$ be a homeomorphism; we can construct $\pi : \mathrm{N}(Y, \mathbb{R}^m) \to Y$ by

$$\pi\left(y, v\right) = y$$

Now by previous lemma, we can find $f_\infty \in C^\infty(X, Y \subset \mathbb{R}^m)$ such that $||f_\infty - f|| < \varepsilon$, which means $\mathrm{Im}(f_\infty) \subset Y_\varepsilon$; and further, for $\forall x \in X$,

$$\begin{aligned}\inf_{y\in Y} ||f(x) + t\,(f_\infty(x) - f(x)) - y|| &\le t \cdot \inf_{y\in Y} ||f_\infty(x) - f(x)|| + \inf_{y\in Y} ||f(x) - y|| \\ &= t \cdot \inf_{y\in Y} ||f_\infty(x) - f(x)|| < t \cdot \varepsilon(x)\end{aligned}$$

for $\forall t \in [0,1]$; in other words, $f(x) + t\,(f_\infty(x) - f(x)) \in Y_\varepsilon$. Therefore, we can construct $\Phi : [0,1] \times X \to Y$ by

$$\Phi(t,x) = \pi \circ \phi\,(f(x) + t\,(f_\infty(x) - f(x)))$$

which is a homotopy from f to f_∞. □

Another application that uses the idea that small perturbation of functions using ϵ-neighbourhood is about *transversality*.

We know that given two smooth surfaces in $\mathbb{R}^3$ (or curves in $\mathbb{R}^2$), it is most likely that they are not tangent to each other, since tangent status could be broken by any slight perturbation. So Being tangent is rare.

The definition of transversality and transversality theorem formalised this idea:

Definition 2.4. Given submanifold $Y \subset M$ and map $f : X \to Y$; the $f \pitchfork Y$ denotes that f is **transversal** to Y: Either $f(X) \cap Y = \emptyset$, or

$$f_* \circ T_x X + T_y Y = T_y M$$

for any $x \in X$ and $y = f(x)$.

And we define two submanifolds $X, Y \subset M$ to be transversal if the inclusion $X \hookrightarrow M$ is transversal to Y.

Notice that this definition does NOT fully define the *transversal* that we normally think of; for example, any two curves intersecting each other in $\mathbb{R}^3$ can never be transversal in our definition, since the sum of their dimensions is always < 3.

The (parametric) transversality theorem states as following:

Proposition 2.13. (Transversality theorem) Given smooth submanifold $Z \subset Y$ and smooth map $F : (0,1) \times X \to Y$ such that $F \pitchfork Z$, we have that

$$F(c, \cdot) \pitchfork Z$$

when c is a regular value of the natural projection $\pi : F^{-1}(Z) \subset (0,1) \times X \to (0,1)$.

This suggests that $F(t,\cdot) \pitchfork Z$ for almost all t, since Sard's theorem tells us that non-regular values are of measure zero.

Proof. For simplicity let's denote $I = (0,1)$ and $f(x) = F(c,x)$.
We need to show that for any $z \in F(c,x)$ and $\xi \in T_z Y$, we have $\xi \in f_* T_x X + T_z Z$. From $F \pitchfork Z$ we know that there exists $(\delta, u) \in T_c I \times T_x X = T_{(c,x)}(I \times X)$ and $w \in T_z Z$ such that $F_*(\delta, u) + w = v$.
Since $c \in I$ is a regular value of $\pi : F^{-1}(Z) \to I$, there exists $(\delta, v) \in T_{(c,x)} F^{-1}(Z)$. We know that $F_*|_{T_x X} = f_*$ by definition, hence

$$\begin{aligned}\xi &= \underbrace{F_*(0, u - v)}_{\in f_* T_x X} - \underbrace{F_*(\delta, v)}_{\in T_z Z} + w \\ &\in f_* T_x X + T_z Z\end{aligned}$$

finishes the proof. □

Remark. Following the conditions we used in this proof, we may know that $I = (0,1)$ can be replaced by other smooth manifolds as well. The theorem would become very useful combined with the Whitney approximation theorem; as they imply that we can use small perturbation to turn a non-transversal map transversal.
This result would be very useful afterwards; but there are already 2 wonderful applications:

Corollary 2.5. Let $X \subset M$ be a smooth submanifold of dimension m.

- If $m < n - 1$, then $\pi_0(M) \cong \pi_0(M \setminus X)$.
- If $m > 2$, then $\pi_1(M) \cong \pi_1(M \setminus X)$.

proof left as an exercise.

And finally, we will propose a conclusion that would be crucial later.

Definition 2.5. Open cover $\mathcal{U}$ of topological space X is called a **good cover** if intersection of finitely many elements of $\mathcal{U}$ is either empty or contractible.

Interestingly every smooth manifold admits a good cover;

Proposition 2.14. *M always admits a good cover.*

Proof. Embed $M \hookrightarrow \mathbb{R}^m$ for some $m > n$; construct open cover $\mathcal{U} = \{U_p \subset \mathbb{R}^m : p \in M\}$ of M such that each $U_p = \mathbb{B}^m_r(p)$ for some $r > 0$ and orthogonal projection of $M \cap U_p \to T_p M$ is an embedding. Then we can verify that the $\mathcal{U}$ forms a good cover of M:
Intersection of finitely many open balls is convex, and its orthogonal projection on any linear subspace is convex, too. Since each orthogonal projection $M \cap U_p \to T_p M$ is embedding, the pre-image of that convex set must be simply connected. □

2.8 Questions

Those questions are sometimes used as a lemma in sections later on.

In questions below, we assume M an n-dim smooth manifold.

1. Given submanifolds $X, Y \subset M$, show that
   1. $X \times Y$ is a manifold.
   2. $M \setminus \overline{X}$ is a manifold.
   3. For $\dim(X) = \dim(Y)$, the $X \# Y$ is a manifold.

2. Suppose X, Y are smooth manifolds such that $X \cup Y$ is a manifold,

   1. If X and Y are orientable, is $X \times Y$ orientable?
   2. If one of X and Y is orientable, can $X \times Y$ be orientable?
   3. If none of X and Y is orientable, can $X \cup Y$ be orientable?

3. Given smooth manifolds with boundary X and Y.
   1. Show that $X \times Y$ is a manifold with boundary.
   2. If X is orientable, is ∂X orientable?
   3. Show that if X is compact, then it can be smoothly embedded into some $\mathbb{R}^m$.

4. Suppose M is connected, show that
   1. M is path-connected.
   2. For any $p, q \in M$, there exists homeomorphism $\phi : M \to M$ such that $\phi(p) = q$.
   3. If M is a Lie group, and U is a neighbour of identity $\mathfrak{i} \in M$, then
   $$M = \bigcup_{k=1}^{\infty} U^k$$

5. Definition 2.7 of tangent space is equivalent to:

   1. Define $T_pM := \{v_p \in \operatorname{Hom}(C^\infty(M), \mathbb{R}) : v_p(fg) = f \cdot v_p(g) + g \cdot v_p(f)\}$.
   2. Consider ideal $\mathcal{I} = \{f \in C^\infty(M) : f(p) = 0\} \subset C^\infty(M)$ and define $T_pM = \mathcal{I}/\mathcal{I}^2$.

6. Suppose $n > 1$, consider a nowhere zero $V \in \Gamma^\infty(TM)$, show that:

   1. For $\forall p \in M$, there exists a chart such that $V = \partial_1$ in its coordinate basis $\{\partial_1, \cdots, \partial_n\}$.
   2. Suppose γ is a smooth path starting from p such that $\dot{\gamma}(0) = V|_p$. Then there exists $W \in \Gamma(TM)$ such that
   $$V|_\gamma = W|_\gamma \text{ but } \mathcal{L}_V(\cdot) \neq \mathcal{L}_W(\cdot)$$

7. Show that Möbius strip $\mathbb{M}$ and projective space $\mathbb{RP}^2$ are NOT orientable.

8. Let $V = \mathbb{R}^n$, and define $\sim_k$ as an equivalence relation on V^k such that $A \sim_k B$ iff there exists $T \in \mathrm{GL}_k(\mathbb{R})$ such that $A = BT$. (where A, B are represented as $n \times k$ matrices)

    1. Show that $\mathbf{Gr}_k(V) := V^k / \sim_k$ is a smooth manifold.
    2. Show that $\mathbf{Gr}_k(V)$ is compact.

9.

10. Given map $F \in C^0(X, Y)$ between smooth manifolds X, Y; show that:
    1. $F \in C^\infty(X, Y)$ if and only if $F^*\left(C^\infty(Y)\right) \subseteq C^\infty(X)$
    2. X and Y are diffeomorphic if and only if $C^\infty(X) \cong C^\infty(Y)$ as rings.

This can be seen as an alternative version of *Gelfand-Kolmogorov theorem.*

2.9 Answers

(Question 1)

(i) Trivial.

(ii) $M\backslash\overline{X}$ is obviously Hausdorff and second-countable. For any $p \in M\backslash\overline{X}$, it has a neighbour $U_p \subset M$ such that $U_p \cong \mathbb{R}^n$. Since $M\backslash\overline{X}$ is open, $(M\backslash\overline{X}) \cap U_p$ is thus an Euclidean neighbour of p.

(iii)

(Question 2)

(i)

(Question 4)

(i) We can define an equivalence relation $\sim$ on M, where $a \sim b$ iff there exists path from a to b. As all points in M have a neighbour homeomorphic to an open ball in Euclidean space, we know that every equivalence class must be open.

Hence $\{[p]\}$ is an open cover of M, where $[p]$ is the equivalence class of $p \in M$; however, different equivalence classes are always disjoint, this suggests that M can only be covered by one $[p]$, so it's path connected.

This suggests that connected space that is locally path-connected is path-connected

(ii) Notice that for any $a, b \in B_1(0)$, there exists homeomorphism $f : B_1(0) \to B_1(0)$ such that $f(a) = b$ and $\lim_{||x|| \to 1} ||f(x) - x|| = 0$ by

$$f(x) = \beta g \beta^{-1} \text{ , where } \beta(x) = \frac{2\arctan ||x||}{\pi ||x||} x$$
$$\text{and } g(x) = x - \beta^{-1}(a) + \beta^{-1}(b)$$

Since M is path connected by (**i**), there exists a path $\gamma : [0,1] \to M$ from p to q. Suppose we cover this path by neighbours of points on it that are homeomorphic to $B_1(0)$; as $[0,1]$ compact, this open cover can be finite. By applying the f in each of those neighbours, we can get the desired diffeomorphism.

(**iii**)

(Question 5)

Let T_pM be the space gained by Definition 2.7, and T'_pM be the one gained by **??**. Both of them are vector spaces. Easy to verify that $T_pM \subset T'_pM$, so we only need to show that $\dim(T'_pM) \leq n$.

WLOG assume that $\varphi_p(p) = 0$. For $\forall f \in C^\infty(N_p)$ we have a closed ball $\overline{B_r(p)} \subseteq \varphi_p(N_p)$ of 0 in which $f \circ \varphi_p^{-1}$ can be approximated by a sequence of polynomials $\{f_n\} \subset \mathbb{R}[x_1, \cdots, x_n]$,

$$(f \circ \varphi^{-1})(x) = \lim_{n\to\infty} f_n$$

By the Leibniz rule, for $v_p \in T'_pM$ we have that $v_p((x_{j_1} x_{j_2} \cdots x_{j_k}) \circ \varphi_p) = 0$ for $\forall k > 1$

(where $j_1, \cdots, j_k \in \{1, \cdots, n\}$), thus T'_pM is equivalent to a subset of linear functionals on $\left\{f \in C^\infty(U) : (f \circ \varphi_p^{-1}) = \sum_{j=1}^n a_j x_j\right\}$ which has dimension n. Hence $\dim(T'_pM) \leq n$; together with the fact that T_pM is a subspace of T'_pM, we have $T_pM = T'_pM$.

(Question 7)

For convenience, given charts (U_1, φ_1) and (U_2, φ_2), for $p \in U_1 \cap U_2$, let $A_p(\varphi_1, \varphi_2)$ to be the basis transformation matrix from coordinate basis of T_pU_1 to that of T_pU_2.

The Möbius strip $\mathbb{M}$ can be defined with quotient topology $[-1,1] \times (-1,1)/\sim$ where $(x_1, y_1) \sim (x_2, y_2)$ iff $x_1 = -x_2 = 1$ and $y_1 = -y_2$, and equivalent class of (x, y) is written as $(x, y)_\sim$. Define a chart $(\{(x,y)_\sim \in \mathbb{M} : |x| \neq 1\}, \varphi_0)$, where $\varphi_0((x,y)_\sim) = (x, y)$.

Let's prove by contradiction. Suppose there exists an orientation-preserve atlas $\{(U_p, \varphi_p)\}$. Consider $S = \{(x, 0)_\sim : x \in [-1, 1]\}$. As $S \cong \mathbb{S}^1$ is compact, there exists finite subset $\{(U_j, \varphi_j)\}_{j=1,\cdots,n}$ that covers S. This suggests that there exists $-1 < x_1 < \cdots < x_n < 1$ such that each of $([x_1, x_2] \times \{0\})_\sim, \cdots, ([x_n, x_1] \times \{0\})_\sim$ is contained in one of $\{U_1, \cdots, U_n\}$; assume WLOG that

$$([x_n, x_1] \times \{0\})_\sim \subset U_1 \; ; \; ([x_1, x_2] \times \{0\})_\sim \subset U_2 \; ; \; \cdots \; ; \; ([x_{n-1}, x_n] \times \{0\})_\sim \subset U_n$$

Now let $p_j = (x_j, 0)_\sim$, and $T_j^k := A_{p_j}(\varphi_j, \varphi_k)$ for convenience. As points in the same local chart use the same coordinate basis, we have that $T_c^a T_b^c = T_b^a$, and so

$$T_n^0 T_{n-1}^n \cdots T_2^3 T_1^2 = T_1^0$$

Since $\mathbb{M}$ is orientable, $\det(T_j^{j+1}) > 0$ for $\forall j$, so $\underline{\det(T_1^0)/\det(T_n^0) > 0}$.

Now define another chart $(\{(x,y)_\sim \in \mathbb{M} : x \neq 0\}, \varphi_{-1})$ where

$$\varphi_{-1}(x,y)_\sim = \begin{cases} (x, y) & \text{for } x > (x_1 + x_n)/2 \\ (x+2, -y) & \text{for } x < (x_1 + x_n)/2 \end{cases}$$

we have that $\det(A_{p_1}(\varphi_0, \varphi_{-1})) = -1$ and $\det(A_{p_n}(\varphi_0, \varphi_{-1})) = 1$. Hence,

$$\begin{aligned} \det(T_1^0) = \det(A_{p_1}(\varphi_{-1}, \varphi_0) A_{p_1}(\varphi_1, \varphi_{-1})) &= -\det(A_{p_n}(\varphi_1, \varphi_{-1}) \\ &= -\det(A_{p_n}(\varphi_1, \varphi_0)) \\ &= -\det(T_n^0)\det(A_{p_n}(\varphi_1, \varphi_n)) = -\det(T_n^0)/\det(T_n^1)^{-1} \end{aligned}$$

which means $\det(T_1^0)/\det(T_n^0) = -\det(T_n^1)^{-1} < 0$, Contradiction.

The case for $\mathbb{RP}^2$ follows similarly, left as an exercise.

(Question 9)

(Question 10)

(**i**) Clearly we have $F^*(C^\infty(Y)) \subseteq C^\infty(X)$ when $F \in C^\infty(X, Y)$, because the composite of smooth maps is smooth.

Let's see the another direction; suppose $\dim(X) = m$, $\dim(Y) = n$, and $F \in C^0(X, Y)$ satisfies that $h \circ F \in C^\infty(X)$ for $\forall h \in C^\infty(Y)$. Notice that any chart function ϕ from a smooth atlas of Y can be written as $\phi(y) = (\phi_1(y), \cdots, \phi_n(y))$ for some $\phi_1, \cdots, \phi_n \in C^\infty(Y)$; therefore, for any chart φ from a smooth atlas of X,

$$\phi \circ F \circ \varphi^{-1}(v) = \begin{bmatrix} (\phi_1 \circ F) \circ \varphi^{-1}(v) \\ \vdots \\ (\phi_n \circ F) \circ \varphi^{-1}(v) \end{bmatrix}$$

is a smooth function from $\mathbb{R}^m$ to $\mathbb{R}^n$, which means $F \in C^\infty(X, Y)$.

(**ii**) Obviously F being a diffeomorphism implies that F^* is an isomorphism; let's see another direction; suppose F^* is an isomorphism:

- F is surjective: Suppose $y \in Y \setminus \mathrm{Im}(F)$, then by taking $f \in C^\infty(Y)$ such that $f(y) \neq 0$, we have that $g = F_*^{-1} \circ F^*(f)$ satisfies $g(0) = 0 \neq f(0)$, which means $f \neq F_*^{-1} \circ F^*$, contradiction.
- F is injective: Suppose $F(x_1) = F(x_2)$ for $x_1 \neq x_2$, then for any $f \in C^\infty(X)$ such that $f(x_1) \neq f(x_2)$ (such function exists by taking the sum of two bump functions), we conclude that $f \notin \mathrm{Im}(F^*)$, contradiction.
- F is smooth directly by (i).
- F^{-1} is smooth by some modification in the last step of (i)'s proof, left as an exercise.

Though solutions of the questions are written by me, some of the questions are NOT original.
Source for questions:

- *Introduction to Smooth Manifolds.* John M. Lee. (Question 1,4,6)

All contents in this chapter were written independently, though most of the knowledge is from the sources below.
Source for this chapter:

- *Introduction to Smooth Manifolds.* John M. Lee.
- Zuoqin Wang's lecture note on *Smooth manifolds*, see here.
- Wikipedia pages.

3 Basics: Differential Form

There are two interesting things about differential forms together with the exterior derivative;
1, It provides an elegant high-dim generalization of vector calculus.
2, The existence and quantity of certain differential forms in a space allow us to investigate the properties of space itself.

Unlike last chapter, this chapter is less technical and more insightful; however, we also have more computation details; one should be comfortable with the computations with differential forms to proceed.

[**Convention**] Throughout this section,
1, The M denotes a n-dim smooth orientable manifold.
2, Without specifying, k is an integer that $1 \leq k \leq n$.
3, When V denotes a vector space, its dimension is n and we use e_i to represent its basis.
4, We use $\mathfrak{T}_{n,k}$ to denote the set of k-element subsets of $\{1, 2, \cdots, n\}$;
5, When writing $\sigma \in \mathfrak{T}_{n,k}$, we assume that σ's elements are arranged in increasing order, and $\sigma(j)$ denotes the jth. element of σ in such order

3.1 Exterior product and Exterior algebra

[**Convention**] Suppose V is equipped with a positively-defined inner product $g = \langle \cdot, \cdot \rangle$; and under the basis $\{e_1, \cdots, e_n\}$,
1, Let $g_{ij} = \langle e_i, e_j \rangle$, the (i, j)th. entry of g.
2, Denote $\omega_g := \sqrt{|\det g|} \cdot (e_1 \wedge \cdots \wedge e_n)$.

Informally, the exterior product (also *wedge product*) $\wedge$ is a special kind of tensor product defined on V^k ($k \in \mathbb{N}$) such that

$$e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(k)} = \operatorname{sgn}(\sigma) \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_k \tag{3.1}$$

for all permutation σ of $\{1, \cdots, k\}$.
But to state it rigorously, it will be easier to introduce things in an opposite way, let's first see what an **exterior algebra** is. In the last section, we learned that vectors from a fixed vector space form a free algebra $\mathrm{T}(V)$ under $\otimes$:

$$\mathrm{T}(V) = \sum_{k=0}^{\infty} \left(\bigotimes^k V \right)$$

The exterior algebra is a quotient algebra of it:

Definition 3.1. The exterior algebra $\bigwedge(V) := \mathrm{T}(V)/I$, where ideal $I = \langle v \otimes v : v \in V \rangle$

While the **exterior product** $\wedge$ is the "multiplication" in this algebra; you can verify that multiplication $\wedge$ in it satisfies the property 3.1.
By this definition, we can express $\bigwedge(V)$ as follow:

$$\bigwedge V := \sum_{k=0}^{n} \left(\bigwedge^k V \right)$$

$$\text{where } \bigwedge^k V := \sum_{\sigma \in \mathfrak{T}_{n,k}} \mathbb{R} \left(e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \cdots \wedge e_{\sigma(k)} \right)$$

We call an element from $\bigwedge^k V$ as a k-**vector**. Aware that the vector space of k-vectors has dimension $\binom{n}{k}$, so $\bigwedge(V)$ has dimension $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

A particularly nice property of $\bigwedge V$ is that $v_1 \wedge v_2 \wedge \cdots \wedge v_k$ vanishes iff $v_1, v_2 \cdots, v_k$ are linearly dependent; so we can somehow treat an element of $\bigwedge^k V$ as a k-dimensional subspace *endowed with volume and orientation.*
Why is it intuitively a subspace? Because the exterior products of basis vectors of the same subspace are (almost) the same:

Proposition 3.1. For $u_j, v_j \in V$, if $\operatorname{span}\{v_1, \cdots, v_k\} = \operatorname{span}\{u_1, \cdots, u_k\}$, then we have $v_1 \wedge \cdots \wedge v_k = r(u_1 \wedge \cdots \wedge u_k)$ for some $r \in \mathbb{R}$.

which is quite obvious by definition.
While the reason why I said such "subspace" is equipped with a volume is quite abstract; let's consider an example:

Example 3.1. Take $v = v_1 \wedge \cdots \wedge v_k$ and $u = u_1 \wedge \cdots \wedge u_{n-k}$, the $(u \wedge v)$ must be a n-vector (also called a **volume vector**), which is in form of $r \cdot (e_1 \wedge \cdots \wedge e_n)$ for $r \in \mathbb{R}$. Interestingly, this r is equal to the (signed) "volume" spanned by the n vectors from components of v and u; that is

$$v \wedge u = \det(A)(e_1 \wedge \cdots \wedge e_n)$$

where (i, j) entry of A is $(v_i)_j$ for $i \leq k$ and $(u_{i-k})_j$ otherwise.

You might be looking for a "real life" example of $\wedge$, "Is the cross product an Exterior product? They look so similar!", but cross product is NOT an exterior product; for example, $\times$ is not associative but $\wedge$ is.

But this doesn't mean there is no connection! The cross product is actually the composition of an exterior product and a **Hodge dual**.
Recall that $\dim(\bigwedge^k V) = \dim(\bigwedge^{n-k} V)$, so there must be an isomorphism between them, and the Hodge dual is a natural construction of it. It has "dual" in its name because the Hodge dual provides a way to extend an inner product from V to $\bigwedge^k V$.

Definition 3.2. Consider $\beta = \beta_1 \wedge \cdots \wedge \beta_k \in \bigwedge^k V$, its **Hodge dual** is $(\star\beta) \in \bigwedge^{n-k} V$, which satisfies that for $\forall \alpha = \alpha_1 \wedge \cdots \wedge \alpha_k \in \bigwedge^k V$,

$$\alpha \wedge \star\beta = \det \begin{bmatrix} \langle \alpha_1, \beta_1 \rangle & \cdots & \langle \alpha_1, \beta_k \rangle \\ \vdots & \ddots & \vdots \\ \langle \alpha_k, \beta_1 \rangle & \cdots & \langle \alpha_k, \beta_k \rangle \end{bmatrix} \omega_g$$

For $\alpha, \beta \in \bigwedge^k V$, the inner product of them $\langle \alpha, \beta \rangle$ is define by $\alpha \wedge \star\beta = \langle \alpha, \beta \rangle \omega_g$.

We can write down "$\star$" explicitly: For $\omega = \omega_{r_1 \cdots r_k}(e_{r_1} \wedge \cdots \wedge e_{r_k}) \in \bigwedge^k(V)$ in ESC[6],

$$(\star\omega)_{j_{k+1}, \cdots, j_n} = \frac{\sqrt{|\det g|}}{(n-k)!} \cdot \varepsilon_{j_1 \cdots j_n} g_{r_1 j_1} \cdots g_{r_k j_k} \omega_{r_1 \cdots r_k} \left(e_{j_{k+1}} \wedge \cdots \wedge e_{j_n}\right) \tag{3.2}$$

in ESC, where we sum each $j_1, \cdots j_n$ over $1, \cdots n$.
This is easy to verify: As $\star$ is linear, we only need to prove the case of $\omega = e_{i_1} \wedge \cdots \wedge e_{i_k}$; which is to check if the value of $(e_{r_1} \wedge \cdots \wedge e_{r_k}) \wedge \star\omega$ matches the result given by Definition 3.2 for every choice of $\{r_1, \cdots r_k = 1, \cdots, n\}$.

[6] ESC means Einstein Summation Criteria

The geometrical intuition for $\wedge$ can be used to think of the Hodge dual: Since $\omega \in \bigwedge^k V$ can be viewed as a k-dim subspace of V, the $\star\omega$ can be viewed as its "*orthogonal complement*":

Proposition 3.2. For any orthonormal basis $\{v_1, \cdots, v_n\}$ of V, we have that

$$\star(v_1 \wedge \cdots \wedge v_k) = \pm(v_{k+1} \wedge \cdots \wedge v_n)$$

for $\forall k = 1, \cdots, n$.

Proof. Equivalent to find $\star(e_1 \wedge \cdots \wedge e_k)$ by setting the inner product to be $g_{ij} = \delta_{ij}$. Then the answer follows immediately by (3.2). □

This intuition brings two elegant properties of Hodge dual:

Proposition 3.2. For $\alpha, \beta \in \bigwedge^k V$, we have that

1. $\star^2(\alpha) = (-1)^{k(n-k)}\alpha$
2. $\langle \star\alpha, \star\beta \rangle = \langle \alpha, \beta \rangle$

Proof. By Proposition 3.1, for $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_k$, we can use Gram–Schmidt process to get an orthonormal set of $\{v_1, \cdots, v_k\}$ such that $\alpha = r \cdot (v_1 \wedge \cdots \wedge v_k)$ for some $r \in \mathbb{R}$; WLOG assume that $r = 1$.
By Gram-Schmidt process, we can continue to find $\{v_{k+1}, \cdots, v_n\}$ such that $\{v_1, \cdots, v_n\}$ forms an orthonormal basis for V. So we can let $\star\alpha = r_1(v_{k+1} \wedge \cdots \wedge v_n)$ and $\star(\star\alpha) = r_1 r_2(v_1 \wedge \cdots \wedge v_k)$ for $r_1, r_2 \in \{\pm 1\}$.

$$\begin{aligned}(\star\alpha) \wedge (\star^2\alpha) &= r_1^2 r_2(v_{k+1} \wedge \cdots \wedge v_n \wedge v_1 \wedge \cdots \wedge v_k) \\ &= r_2(-1)^{k(n-k)}(v_1 \wedge \cdots \wedge v_n)\end{aligned}$$

At the same time, if we choose $\{v_j\}$ as our basis for V, then $(\star\alpha) \wedge (\star^2\alpha) = (v_1 \wedge \cdots \wedge v_n)$ by definition. Hence $r_2 = (-1)^{k(n-k)}$ and $\star(\star\alpha) = (-1)^{k(n-k)}\alpha$.
Then the second statement follows as a corollary,

$$\begin{aligned}\langle \star\alpha, \star\beta \rangle &= (\star\alpha) \wedge (\star^2\beta) \\ &= (-1)^{k(n-k)}(\star\alpha) \wedge \beta \\ &= \beta \wedge (\star\alpha) = \langle \beta, \alpha \rangle = \langle \alpha, \beta \rangle\end{aligned}$$

□

Let's take some simple examples.

Example 3.3. Let $V = \mathbb{R}^3$ and g be the dot product, you can verify that

$$\star e_1 = e_2 \wedge e_3$$
$$\star e_2 = e_3 \wedge e_1$$
$$\star e_3 = e_1 \wedge e_2$$

Let $V = \mathbb{R}^4$ and $g := \mathrm{diag}(-1, 1, 1, 1)$, the Minkowski metric, we have that

$$\begin{aligned}
\star e_0 &= -e_1 \wedge e_2 \wedge e_3 \quad ; \quad \star(e_0 \wedge e_1) = -e_2 \wedge e_3 \quad \star(e_1 \wedge e_2) = e_0 \wedge e_3 \\
\star e_1 &= -e_0 \wedge e_2 \wedge e_3 \quad ; \quad \star(e_0 \wedge e_1) = -e_2 \wedge e_3 \quad \star(e_2 \wedge e_3) = e_0 \wedge e_1 \\
\star e_2 &= -e_0 \wedge e_3 \wedge e_1 \quad ; \quad \star(e_0 \wedge e_1) = -e_2 \wedge e_3 \quad \star(e_3 \wedge e_1) = e_0 \wedge e_2 \\
\star e_3 &= -e_0 \wedge e_1 \wedge e_2 \quad ;
\end{aligned}$$

an interesting application of using hodge dual and exterior derivative (which we will see soon) is that the *Maxwell equations* can be reduced to the form

$$\mathrm{d}F = 0$$
$$\star \mathrm{d} \star F = \mu_0 J$$

where F is a 2-form representing the electromagnetic tensor, and J is a 1-form representing the current 4-vector.

Notice that the $\mathbb{R}^3$ case follows the same law as the cross product: $e_i = \varepsilon_{ijk} \cdot (e_j \times e_k)$. So if we define g to be the dot product, the cross product of two vectors is actually the Hodge dual of the exterior product of them:

$$\boxed{\text{For } a, b \in \mathbb{R}^3,\ a \times b = \star(a \wedge b)}$$

This provides an interesting connection between what we are learning here and Vector calculus in $\mathbb{R}^3$, which can be extended to any 3-dim Riemannian manifold.

3.2 Differential forms

Equip M with an Riemannian metric g; let g_{ij} be its (i,j) entry.

The reason why we care about such a funny algebra in this note is to introduce a fundamental idea in calculus on smooth manifolds: *differential forms*. It is easier to first give the definition before pointing out its significance:

Definition 3.3. We define:

- A k-form at $p \in M$ is an element of $\bigwedge^k T_p^* M$.
- $\bigwedge^k T^*M = \left\{ (p, v) : p \in M, v \in \bigwedge^k T_p^* M \right\}$
- A differential k-form on M is defined as a smooth section $M \to (\bigwedge^k T^* M)$.
- The $\Omega^k(M)$ denotes the set of differential k-forms, and $\Omega(M) := \sum_{k=0}^n \Omega^k(M)$.

Like $\bigwedge^k V$, the $\Omega(M)$ also forms a graded ring under $+$ and $\wedge$.
To be more precise, if consider a local coordinate basis $\{\mathrm{d}x_1, \cdots, \mathrm{d}x_n\}$ of T^*M on $U \subset M$, then we can simply write a differential k-forms ω at $p \in U$ in form of

$$\omega|_p := \sum_{\sigma \in \mathfrak{T}_{n,k}} f_\sigma(p) \; \left(\mathrm{d}x_{\sigma(1)} \wedge \cdots \wedge \mathrm{d}x_{\sigma(k)} \right)$$

for some $f_\sigma \in C^\infty(U)$.
In our introduction of the tensor, we've seen an isomorphism $\bigotimes^k V^* \cong \mathcal{L}(V \times \cdots \times V, \mathbb{R})$, which means every element in $\mathrm{T}(V)$ gives a multilinear map from $V \times \cdots \times V$ to $\mathbb{R}$. So this isomorphism will also send $\bigwedge^k V$ to a quotient of $\mathcal{L}(V \times \cdots \times V, \mathbb{R})$:

$$\bigwedge^k V \cong \mathcal{A}\left(V^k, \mathbb{R}\right) \tag{3.3}$$

where $\mathcal{A}$ is the set of alternating multilinear maps defined by

$$\mathcal{A}(V^k, W) := \mathcal{L}(V^k, W)/J$$

where ideal $J = \left\langle \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \right\rangle$ for $f \in \mathcal{L}(V^n, W)$.
You might immediately notice that $\mathcal{A}(V^k, \mathbb{R})$ and $\bigwedge^k V^*$ appear naturally as quotient spaces of $\mathcal{L}(V^k, \mathbb{R})$ and $\bigotimes^k V^*$ respectively by the linear maps $\phi : \mathcal{L}(V^k, \mathbb{R}) \to \mathcal{L}(V^k, \mathbb{R})$ and $\varphi : \bigotimes^k V^* \to \bigotimes^k V^*$, where

$$\begin{aligned} \varphi &: u_1 \otimes \cdots \otimes u_k \longmapsto \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \cdot (u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(k)}) \\ \phi &: f \longmapsto \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \cdot f\left(v_{\sigma(1)}, \ldots, v_{\sigma(k)}\right) \end{aligned}$$

By First isomorphism theorem, we have that $\operatorname{Im}(\phi) \cong \mathcal{A}(V^k, \mathbb{R})$ and $\operatorname{Im}(\varphi) \cong \bigwedge^k V^*$. Defining in this way ensures that the construction is entirely symmetrical.

$$\begin{tikzcd}
\bigotimes^k V^* \arrow[r, "f", "\cong"'] \arrow[d, "\varphi"] & \mathcal{L}(V^k, \mathbb{R}) \arrow[d, "\phi"] \\
\bigwedge^k V^* \arrow[r] & \mathcal{A}(V^k, \mathbb{R})
\end{tikzcd}$$

So when f is defined canonically as we had:

$$f : (u_1 \otimes \cdots \otimes u_k) \longmapsto \left((v_1, \ldots, v_k) \mapsto \sum_{1 \le j_1, \ldots, j_k \le k} u_1(v_{j_1}) \cdot \cdots \cdot u_k(v_{j_k}) \right)$$

Notice that $(\phi \circ f \circ \varphi^{-1})$ is a well-defined map, so pick any $\omega = u_1 \wedge \cdots \wedge u_k \in \bigwedge^k V^*$ with pre-image $\frac{1}{k!}(u_1 \otimes \cdots \otimes u_k) \in \varphi^{-1}\omega$, we have that $(\phi \circ f \circ \varphi^{-1})\omega$ sends

$$(v_1, \ldots, v_k) \longmapsto \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) u_1(v_{\sigma(1)}) \cdot \cdots \cdot u_k(v_{\sigma(k)})$$

So though there are may other ways to give the correspondence between $\bigwedge^k V^*$ and $\mathcal{A}(V^k, \mathbb{R})$, we take this to be the standard:

Proposition 3.3. Given $\omega = u_1 \wedge \cdots \wedge u_k \in \bigwedge^k V^*$, the ω defines a multilinear map $V^k \to \mathbb{R}$, where for $(v_1, \cdots, v_k) \in V^k$,

$$\omega : \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix} \mapsto \det \begin{bmatrix} u_1(v_1) & \cdots & u_1(v_k) \\ \vdots & \ddots & \vdots \\ u_k(v_1) & \cdots & u_k(v_k) \end{bmatrix} \tag{3.4}$$

which looks like the Jacobian; indeed, this brings a geometric intuition for it: Let $V = T_p^* M$, then the action of ω on $(T_p M)^k$ is the same as computing the size of "projection" of the volume represented by ω on the space $(T_p M)^k$.

This formulation then allows us to define the pullback of differential forms using the pushforward of tangent space.

Definition 3.4. Suppose X and Y are two smooth manifolds, and $\phi \in C^\infty(X, Y)$, the pullback $\phi^* \omega$ of $\omega \in \Omega^k(Y)$ is given by

$$\phi^* \omega|_x (v_1, \cdots, v_k) := \omega|_{\phi(x)} (\phi_* v_1, \cdots, \phi_* v_k)$$

Writing down the explicit expression for arbitrary k seems difficult, as it involves solving a set of linear equations, but there is a good property to simplify the computation:
Just like pushforward, we can conclude that the pullback of differential forms also satiesfies

Linearity and Chain rule; but in addition to these, it also Commutes with $\wedge$:

Proposition 3.4. For $\forall \eta, \omega \in \Omega(X)$, $\phi^*(\eta \wedge \omega) = (\phi^*\eta) \wedge (\phi^*\omega)$

In other words, a pullback of differential forms is a ring homomorphism.

Proof. We only need to verify a general case, that $\phi^*(\omega_1 \wedge \cdots \wedge \omega_k) = (\phi^*\omega_1) \wedge \cdots \wedge (\phi^*\omega_k)$. We prove by induction, suppose it's true for $k = r - 1$, then

$$\begin{aligned}
\phi^*(\omega_1 \wedge \cdots \wedge \omega_r)\big|_x (v_1, \cdots, v_r) &= \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) \prod_{j=1}^{r} \omega_j\big|_{\phi(x)} (\phi_* v_{\sigma(j)}) \\
&= \sum_{i=1}^{r} (-1)^{i+1}\, \omega_r\big|_{\phi(x)} (\phi_* v_i) \cdot \bigwedge_{j \neq r} \phi^*\omega_j \Big|_x (v_1, \cdots, \widehat{v_i}, \cdots, v_r) \\
&= (\phi^*\omega_1 \wedge \cdots \wedge \phi^*\omega_r)\big|_x (v_1, \cdots, v_r)
\end{aligned}$$

which finishes the induction. (where the $3^{\text{rd.}}$ and $2^{\text{nd.}}$ equal sign use the inductive hypo for $k = 2$ and $(r - 1)$ respectively.) $\square$

In this way, pullback is much easier to compute based on pushforward: For $\{\mathrm{d}x_1, \cdots, \mathrm{d}x_n\}$ being the coordinate basis of T_x^*X and $\{\mathrm{d}y_1, \cdots, \mathrm{d}y_m\}$ being that of $T_{\phi(x)}^*Y$, then

$$\phi^*(\mathrm{d}y_j) = \sum_{k=1}^{m} (\phi_*)_{jk}\, \mathrm{d}x_k \tag{3.5}$$

so for 1-form, ϕ^* is numerically the transpose of ϕ_*. And for cases of $k > 1$, the rest of the work is simply expanding into components; for example, we can easily get

Example 3.4. Given smooth manifolds X, Y with $\dim(X) = \dim(Y) = n$,

$$\phi^*\omega = \det(\phi_*) \cdot (\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n)$$

for $\omega = (\mathrm{d}y_1 \wedge \cdots \wedge \mathrm{d}y_n) \in \Omega^n(Y)$.

Recall that in the Section 3.1, we defined a special n-form ω_g on M to be:

Definition 3.1. A **volume form** is a differential n-form ω_g such that

$$\omega_g := \sqrt{|\det g|}\ \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n$$

in every local coordinate basis $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$

Now we can explain it. Suppose $\phi : X \to Y$ is an immersion, and g a Riemannian metric on Y. If $\widetilde{g}$ is the pullback of g by ϕ —— which, in terms of matrices,

$$\widetilde{g} = \phi_*^T\, g\, \phi_*$$

so $\omega_{\widetilde{g}} = \phi^*\omega_g$, which means ω_g is invariant under pullback induced by orientation-preserving diffeomorphism; and therefore, ω_g exists when M is orientable (which we have assumed throughout the chapter).
Moreover, we can also deduce that ω_g does NOT exist when M is not orientable:

Proposition 3.5. A connected smooth n-dim manifold X is orientable if and only if there exists a non-vanishing $\omega \in \Omega^n(M)$.

Proof. Suppose X is orientable, which means there exists an orientation-preserving atlas. Consider a Riemannian metric g and $\omega_g \in \Omega^n(X)$ such that ω_g is given by the volume form in each of the charts in that atlas. The ω_g is well-defined since we've shown that ω_g is invariant under pullback by transition functions, which are orientation-preserving diffeomorphisms. The ω_g is a desired non-vanishing n-form by definition of the Riemannian metric.
Conversely, suppose $\omega \in \Omega^n(X)$ is a non-vanishing n-form. Let $\{U_j\}$ be an open cover of X such that each $U_j \cong \mathbb{R}^n$. For any U_j, any chart function φ_j on it always satisfies that the sign of $\omega(\partial_1, \ldots, \partial_n)$ is the same everywhere (because ω is non-vanishing and U_j is connected), so we can always choose a φ_j such that $\omega(\partial_1, \ldots, \partial_n) > 0$. Then this $\{(U_j, \varphi_j)\}$ gives an orientation-preserving atlas. $\square$

The intuition that we can get from the above reasoning is that, a non-vanishing $\omega \in \Omega^n(M)$ can indicate the orientation of M.

Proposition 3.6. Given two orientation-preserving atlas $\mathscr{U}_1$ and $\mathscr{U}_2$, by fixing a point $p \in M$ and pick coordinate basis $\mathbf{v}_1, \mathbf{v}_2 \in (T_pM)^n$ from these two atlas respectively, we have that

$$\mathscr{U}_1, \mathscr{U}_2 \text{ has same orientation} \iff \omega(\mathbf{v}_1) \cdot \omega(\mathbf{v}_2) > 0$$

3.3 Integration of differential form

Let X be an connected k-dim orientable smooth submanifold of M.
Denote $\Omega_c^k(M) := \{\omega \in \Omega^k(M) : \operatorname{supp}(\omega) \text{ compact}\}$.

We discussed that a geometrical intuition for a differential k-form acting on $(T_xX)^k$ is the size of projection of it on X. Hence we can define the integral of $\Omega^k(M)$ on a k-dim smooth submanifold as a way to measure the size of overall projection; though note that an additional requirement is that this submanifold must be orientable, we will see why.
So let's consider how to integrate k-forms on X. Our strategy is still using an atlas to put things into Euclidean space.

Definition 3.5. Consider an orientation-preserving atlas $\{(N_j, \varphi_j)\}$ that is locally finite and a POU $\{\rho_j\}$ subordinate to it.
For any $\omega = f\,(\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k) \in \Omega_c^k(X)$ we define

$$\int_X \omega := \sum_j \int_{\varphi_j(N_j)} (f \cdot \rho_j) \circ \varphi_j^{-1} \tag{3.6}$$

where the integrals on RHS are integrals on $\mathbb{R}^k$.

Though it seems messy, it is actually a very natural formulation, as all we do is pull the form into $\mathbb{R}^n$. So intuitively it should be *independent of the choice of atlas*:

Proposition 3.7. Under the same orientation, the value of $\int_X \omega$ in Definition 3.6 is invariant under different choices of atlas and POU.

Proof. First, it is easy to show that the statement is true for two atlas of the same open cover and POU: Consider $\{(U_i, \varphi_i)\}$ and $\{(U_i, \vartheta_i)\}$ with POU $\{\rho_i\}$, we have that by change of variable formula in $\mathbb{R}^k$ and Example 3.4,

$$\begin{aligned}\int_{\varphi_i(U_i)} (f \cdot \rho_i) \circ \varphi_i^{-1} &= \int_{\vartheta_i(U_i)} \frac{1}{\det\left(\vartheta_i \circ \varphi_i^{-1}\right)^*} \cdot \left(f \cdot \rho_i \cdot \left|\det(\vartheta_i \circ \varphi_i^{-1})_*\right|\right) \circ \vartheta_i^{-1} \\ &= \int_{\vartheta_i(U_i)} (f \cdot \rho_i) \circ \vartheta_i^{-1}\end{aligned}$$

so these two atlas produce the same $\int_X \omega$. Then we consider the general case;
Suppose $\{(U_i, \varphi_i)\}$ and $\{(W_j, \vartheta_j)\}$ are two atlas with POU $\{\rho_i\}$ and $\{\varrho_j\}$ respectively. Notice that $\{(U_i \cap W_j, \varphi_i)\}$ is also an atlas with POU $\{\rho_i \varrho_j\}$, which means

$$\begin{aligned}\sum_i \int_{\varphi_i(U_j)} (f \cdot \rho_i) \circ \varphi_i^{-1} &= \sum_i \int_{\varphi_i(U_i)} \left(f \cdot \rho_i \sum_j \varrho_j\right) \circ \varphi_i^{-1} \\ &= \sum_i \sum_j \int_{\varphi_i(U_i \cap W_j)} (f \cdot \rho_i \varrho_j) \circ \varphi_i^{-1}\end{aligned}$$

Because we have just shown that the integral is independent of choice of atlas, we then have, by symmetry

$$\begin{aligned}\sum_i \int_{\varphi_i(U_i)} (f \cdot \rho_i) \circ \varphi_i^{-1} &= \sum_i \sum_j \int_{\varphi_i(U_i \cap W_j)} (f \cdot \rho_i \varrho_j) \circ \varphi_i^{-1} \\ &= \sum_j \sum_i \int_{\vartheta_j(U_i \cap W_j)} (f \cdot \rho_i \varrho_j) \circ \vartheta_j^{-1} = \sum_j \int_{\vartheta_j(W_j)} (f \cdot \varrho_j) \circ \vartheta_j^{-1}\end{aligned}$$

□

Remark. This is a clever proof, as a direct proof of such statement can be complicated, while an use of symmetry simplifies the problem.

Using this fact, we can freely write down the change of variable formula:

Proposition 3.5. Suppose $\phi : X \to Y$ is diffeomorphism, then

$$\int_Y \omega = \pm \int_X \phi^* \omega \tag{3.7}$$

where $\pm = 1$ when ϕ is orientation-preserving, and -1 when reversing.

Proof. First notice that given an atlas $\{(U_j, \varphi_j)\}$ with POU $\{\rho_j\}$ on X, the $\{(\phi(U_j)\ ,\ \varphi_j \circ \phi^{-1})\}$ and $\{\rho_j \circ \phi^{-1}\}$ will give an atlas and POU on Y respectively.
By the change of variable formula in $\mathbb{R}^k$, the integral of ω on each $\phi(U_j)$ is the same (up to sign) as that of $\phi^*\omega$ on each U_j respectively. So the statement is true for this pair of atlas and POU, which means it is generally true by Proposition 3.7. □

This formula immediately allow us to integrate a r-form on a r-dim submanifold ($r \leq \dim(Y)$); and one may find that the integration in vector calculus fits well into this theory:

Example 3.8. Take a vector-calculus-like example. Suppose $X = \mathbb{R}^2$ and $Y = \mathbb{R}^3$, let $\omega = F_1(\mathrm{d}y_2 \wedge \mathrm{d}y_3) + F_2(\mathrm{d}y_3 \wedge \mathrm{d}y_1) + F_3(\mathrm{d}y_1 \wedge \mathrm{d}y_2) \in \Omega_c^2(Y)$, and $\phi : X \hookrightarrow Y$ smooth embedding, we have that

$$\begin{aligned}\int_{\phi(X)} \omega &= \int_X \det \begin{pmatrix} F_1 & F_2 & F_3 \\ (\phi_*)_{11} & (\phi_*)_{21} & (\phi_*)_{31} \\ (\phi_*)_{12} & (\phi_*)_{22} & (\phi_*)_{32} \end{pmatrix} (\mathrm{d}x_1 \wedge \mathrm{d}x_2) \\ &= \int_{\phi(X)} \vec{F} \cdot \vec{n}\ dS\end{aligned}$$

where $\vec{F} = (F_1, F_2, F_3)$ and $\vec{n} = (\phi_* \partial_1) \times (\phi_* \partial_2)$, and the integral after the $2^{\text{nd.}}$ "=" is written in Vector calculus' convention.

This is another evidence showing that theory of differential form can generalize vector calculus, and we will see the ultimate one in the next part.

3.4 Exterior derivative

For convenience, since the important part of a differential k-form $p \mapsto (p, \omega)$ is the ω part, we will only write the ω part of a differential form.

The exterior derivative is defined in a strange way,

Definition 3.6. Fix a local coordinate basis $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_n\}$ for T^*M; an exterior derivative is a map $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$ defined as

$$\mathrm{d} : \sum_{\sigma \in \mathfrak{T}_{n,k}} f_\sigma \cdot \mathrm{d}x_\sigma \mapsto \sum_{\sigma \in \mathfrak{T}_{n,k}} \sum_{j=1}^{n} \frac{\partial f_\sigma}{\partial x_j} \cdot \mathrm{d}x_j \wedge \mathrm{d}x_\sigma$$

where $\mathrm{d}x_\sigma = \mathrm{d}x_{\sigma(1)} \wedge \mathrm{d}x_{\sigma(2)} \wedge \cdots \wedge \mathrm{d}x_{\sigma(k)}$ and $f_\sigma \in C^\infty(M)$.

It can also be defined using axioms without coordinates: for each $0 \le k \le n$, an exterior derivative is a linear map $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$ (for any $0 \le k \le n$) that satisfies

- $(\mathrm{d}f)(v) = v(f)$ for any 0-form f.
- $\mathrm{d}^2(f) = 0$ for any 0-form f.
- $\mathrm{d}(\alpha \wedge \beta) = (\mathrm{d}\alpha) \wedge \beta + (-1)^q \alpha \wedge (\mathrm{d}\beta)$ for $\alpha \in \Omega^q(M)$

for $\forall v \in T_pM$ and $\beta \in \Omega(M)$. These 2 definitions are **equivalent**:

Proof. Easy to show that the Definition 3.6 satisfies the 3 conditions in the axiom-definition; we only need to prove the converse.
Prove by induction. Suppose it is true for all $k < r$, now consider $k = r$.
Notice that we only need to show the case of $\mathrm{d}(\omega)$ for $\omega = f \cdot \mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \cdots \wedge \mathrm{d}x_r$ as other cases follow directly by symmetry and linearity of d. By inductive hypo,

$$\begin{aligned} \mathrm{d}(\omega) &= \mathrm{d}\left(f \cdot \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1}\right) \wedge \mathrm{d}x_r + (-1)^{r-1} \omega \wedge \underbrace{\mathrm{d}(\mathrm{d}x_r)}_{=0} \\ &= \sum_{j=r}^{n} \frac{\partial f}{\partial x_j} \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{r-1} \wedge \mathrm{d}x_j \end{aligned}$$

thus it is true for $k = r$.
Since the case for $k = 1$ is true by definition (Condition 1), it is true for all k by induction. □

This means that d is independent of the choice of coordinates.
As an intrinsic property, an interesting property of it is that it also commutes with the pullback map produced by a diffeomorphism:

Proposition 3.6. If $\phi \in C^\infty(X, Y)$ for smooth manifolds X and Y, then

$$\mathrm{d} \circ \phi^* = \phi^* \circ \mathrm{d}$$

Proof. Since $(\phi^*(\mathrm{d}f))(v) = (\mathrm{d}f)(\phi^* v) = v(\phi^* f) = (\mathrm{d}(\phi^* f))(v)$, the statement is true for $k = 0$; we can then use induction:
Suppose it's true for $k \leq r$. Since both d and ϕ^* are linear, we only need to verify the case where ω is in form of $f\,(\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \cdots \wedge \mathrm{d}x_r)$ where $f \in C^\infty(Y)$. So denote $\omega_r = f\,(\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_r)$, we have

$$\begin{aligned}
\mathrm{d} \circ \phi^*(\omega) &= \mathrm{d}(\phi^*(\omega_r) \wedge \phi^*(\mathrm{d}x_{r+1})) \\
&= \mathrm{d} \circ \phi^*(\omega_r) \wedge \phi^*(\mathrm{d}x_{r+1}) + (-1)^{r-1} \phi^*(\omega_r) \wedge (\mathrm{d}(\phi^* \mathrm{d}x_{r+1})) \\
&= \phi^*(\mathrm{d}(\omega_r) \wedge \mathrm{d}x_{r+1}) + (-1)^{r-1} \phi^*(\omega_r) \wedge (\phi^*(\mathrm{d}(\mathrm{d}x_{r+1}))) \\
&= \phi^* \left(\mathrm{d}(\omega_r) \wedge \mathrm{d}x_{r+1} + (-1)^{r-1} \omega_r \wedge \mathrm{d}(\mathrm{d}x_{r+1}) \right) = \phi^* \circ \mathrm{d}(\omega)
\end{aligned}$$

by applying the $3^{\text{rd.}}$ property in the axiom-definition of d. which finishes the induction. □

The exterior derivative is a way of considering the differentiation of $\Omega^k(M)$, but since $\Omega^k(M)$ are essentially tensor fields, the Lie derivative tells something as well.
For $\omega \in \Omega^k(M)$, $F \in \Gamma^\infty(TM)$ and $V_1, \ldots, V_k \in \Gamma^\infty(TM)$,

$$(\mathcal{L}_F \omega) = F\left(\omega(V_1, \ldots, V_k)\right) - \sum_{j=1}^{k} \omega\left(V_1, \ldots, \mathcal{L}_F V_j, \ldots, V_k\right)$$

by definition. This is difficult to compute; luckily a magical formula gives a nice formulation on it by introducing an "anti-derivative":

Definition 3.2. Given $F \in \Gamma^\infty(M)$, for $k = 1, 2, 3, \cdots$, the **interior derivative** is the map $F \lrcorner : \Omega^k(M) \to \Omega^{k-1}(M)$ defined by

$$\left(F \lrcorner\, \omega\right)\left(W_1, \ldots, W_{k-1}\right) = \omega\left(F, W_1, \ldots, W_{k-1}\right)$$

for $\forall W_1, \ldots, W_{k-1} \in \Gamma^\infty(M)$. And when $k = 0$, we set $F \lrcorner\, \omega = 0$

Proposition 3.7. (Cartan magic formula) For any $V \in \Gamma^\infty(M)$ and $\omega \in \Omega(M)$,

$$\mathcal{L}_V \omega = \left((V \lrcorner) \circ \mathrm{d} + \mathrm{d} \circ (V \lrcorner)\right) \omega$$

I don't yet know a good intuition that can lead to this formula; though a proof given by Shiing-Shen Chern explain some algebraic reasoning of it:

Proof. □

As the exterior derivative is given, it is possible to express concepts like grad, curl and div in any 3-dim differentiable manifold N.

Definition 3.3. Suppose M is a 3-dimensional manifold with inner product $\langle , \rangle$ defined on tangent space on each point of M, we define

$$(\nabla) : C^\infty(M) \to \Gamma(TM) \text{ such that } \nabla f = (\mathrm{d}f)^\sharp$$
$$(\nabla\times) : \Gamma(TM) \to \Gamma(TM) \text{ such that } \nabla \times F = (\star\mathrm{d}(F^\flat))^\sharp$$
$$(\nabla\cdot) : \Gamma(TM) \to C^\infty(M) \text{ such that } \nabla \cdot F = \star\mathrm{d}(\star F^\flat)$$

They are just simple generalizations of the $M = \mathbb{R}^3$ case, as we can check that they work when $M = \mathbb{R}^3$: The first one is simple; here are the cases of $(\nabla\times)$ and $(\nabla\cdot)$:

Proof. Suppose $(\partial_i)^\flat = \mathrm{d}x_i$, then for $F = f_1\partial_1 + f_2\partial_2 + f_3\partial_3 \in \Gamma^\infty(TM)$, we have

$$\begin{aligned}\left(\star\mathrm{d}(F^\flat)\right)^\sharp &= \left(\star\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\partial f_i}{\partial x_j}(\mathrm{d}x_i \wedge \mathrm{d}x_j)\right)^\sharp \\ &= \left(\star\sum_{\{i,j\}\in\mathfrak{T}_{3,2}} \mathrm{sgn}(i\ j\ k)\left(\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i}\right)(\mathrm{d}x_i \wedge \mathrm{d}x_j)\right)^\sharp \equiv \nabla \times F\end{aligned}$$

and for divergence,

$$\begin{aligned}\star\mathrm{d}(\star F^\flat) &= \star\mathrm{d}\left(\sum_{\mathrm{sgn}(i\ j\ k)=1} f_i(\mathrm{d}x_j \wedge \mathrm{d}x_k)\right) \\ &= \star\sum_{\mathrm{sgn}(i\ j\ k)=1}\frac{\partial f_i}{\partial x_i}(\mathrm{d}x_i \wedge \mathrm{d}x_j \wedge \mathrm{d}x_k) \equiv \nabla \cdot f\end{aligned}$$

□

If we neglect the Hodge star and music isomorphisms, they are simply representations of the exterior derivative in a 3-dimensional manifold, so exterior derivative is a more generalized yet abstract way to think about "derivative".

3.5 Stoke's Theorem

Denote $\Omega_c^k(M) : \left\{\omega \in \Omega^k(M) : \mathrm{supp}(\omega) \text{ compact}\right\}$ as before.
Let $X \subset M$ be an orientable k-dimensional submanifold with boundary.

Based on our knowledge about exterior derivative d, the ∇, $\nabla\times$ and $\nabla\cdot$ are essentially vector-algebra versions of d in $\mathbb{R}^3$; Similarly, the *Green's formula, Stoke's formula & divergence theorem* can be viewed as that of *Generalized Stoke's Theorem* in $\mathbb{R}^3$.

Generalized Stoke's Theorem

Theorem 3.9. Let $\iota : X \hookrightarrow M$ be the natural inclusion; for any $\omega \in \iota^* \circ \Omega_c^{k-1}(M)$,

$$\int_X \mathrm{d}\omega = \int_{\partial X} \omega$$

where the orientation of ∂X is induced by ι.

Or we can write the theorem explicitly by just letting $\omega \in \Omega_c^k(X)$.

Proof. Any point $p \in M$ is inside either $X \setminus \partial X$ or ∂X or $M \setminus X$. In these 3 cases, there exists a neighbour U_p and a diffeomorphism φ from $(U_p \cap X)$ to $\emptyset$ or $\mathbb{B}_1^k(0)$ or $\mathbb{H}_1^k(0)$ respectively. Cover $\mathrm{supp}(\omega)$ by those U_p. Since $\mathrm{supp}(\omega)$ is compact, it has a finite subcover $\mathcal{U}$. Now denote

$$\mathcal{B} := \mathcal{U} \cap \{U_p : p \in X \setminus \partial X\};$$
$$\mathcal{H} := \mathcal{U} \cap \{U_p : p \in \partial X\}$$

Now for each $U_p \in \mathcal{B}$ or $\mathcal{H}$, choose chart φ where $\varphi(U_p) = \mathbb{B}_1(0)$ or $\mathbb{H}_1(0)$ respectively, so that when in the case of $U_p \in \mathcal{H}$,

$$\varphi(\partial X) = \left\{x \in \mathbb{H}_1^k(0) : x_k = 0\right\}$$

and $\{(U_p, \varphi)\}$ gives an orientation-preserving atlas.
Since $|\mathcal{B} \cup \mathcal{H}|$ is finite, we can assign a POU $\{\rho_p\}$ subordinate to them. So we can write

$$\int_X \mathrm{d}\omega = \sum_{U_p \in \mathcal{B}} \int_{U_p} \mathrm{d}(\rho_p\omega) + \sum_{U_p \in \mathcal{H}} \int_{U_p \cap X} \mathrm{d}(\rho_p\omega) \tag{3.8}$$

$$\text{and} \quad \int_{\partial X} \omega = \sum_{U_p \in \mathcal{H}} \int_{\partial X \cup U_p} \rho_p\omega \tag{3.9}$$

For any $U_p \in \mathcal{B}$, consider $\rho_p\omega := \sum_{r=1}^n (-1)^{r+1}(f_r \circ \varphi_p)\left(\mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_r} \wedge \cdots \wedge \mathrm{d}x_n\right)$ where $f_r \in C^\infty(\mathbb{B}_1(0))$. Because $\mathrm{supp}(\rho_p)$ is compact, we have $\lim_{||x|| \to 1} f_r(x) = 0$, which means

$$\int_{U_p} \mathrm{d}\left(\rho_p\omega\right) = \int_{\mathbb{B}_1^k(0)} \left(\sum_{r=1}^n \frac{\partial f_r}{\partial x_r}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n = \sum_{r=1}^n \int_{\mathbb{B}_1^k(0)} \frac{\partial f_r}{\partial x_r} \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0$$

And similarly for any $U_p \in \mathcal{H}$, let $\rho_p\omega := \sum_{r=1}^n (-1)^{r+1}(f_r \circ \varphi_p)\left(\mathrm{d}x_1 \wedge \cdots \wedge \widehat{\mathrm{d}x_r} \wedge \cdots \wedge \mathrm{d}x_n\right)$

where $f_r \in C^\infty(\mathbb{H}_1(0))$; we have $\lim_{||x||\to 1} f_r(x) = 0$, and

$$\begin{aligned}\int_{U_p\cap X} \mathrm{d}\left(\rho_j\omega\right) &= \sum_{r=1}^{n}\int_{\mathbb{H}_1^k(0)}\frac{\partial f_r}{\partial x_r}\mathrm{d}x_1\cdots\mathrm{d}x_n \\ &= \sum_{r=1}^{n}\int_{\mathbb{B}_1^{k-1}(0)} -f_r(x_1,\cdots,x_{k-1},\mathbf{0})\ \mathrm{d}x_1\cdots\mathrm{d}x_{n-1} = \int_{\partial X\cup U_p}\rho_p\omega\end{aligned}$$

Plugging them into (3.8) gives the desired result. □

Stoke's theorem can explain the definition of d, as we can say that Exterior derivative is the unique linear operator that makes Stoke's theorem work. There is a simple explanation.

Proposition 3.10. Suppose $\mathrm{d} : \Omega^*(M) \to \Omega^{*+1}(M)$ is a linear operator such that

$$\int_X \mathrm{d}\omega = \int_{\partial X}\omega$$

for every compact submanifold $X \subset M$ and $\omega \in \iota^*\Omega(M)$, where $\iota : X \hookrightarrow M$ is the natural inclusion. Then d is the exterior derivative.

Proof. Consider arbitrary $p \in M$ and a local chart $\varphi : \mathbb{R}^n \to M$ with $\varphi(0) = p$, and suppose $\omega = (f\circ\varphi^{-1})(\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_k) \in \Omega^k(M)$ in the coordinate basis.
For every $\sigma \in \mathfrak{T}_{n,k+1}$, construct $X_\sigma = \varphi(B_\sigma)$ and natural inclusion $\iota : X_\sigma \hookrightarrow M$, where

$$B_\sigma = \mathrm{span}_{\mathbb{R}}\left\{e_{\sigma(1)},\ldots,e_{\sigma(k+1)}\right\}\cap\overline{\mathbb{B}_\varepsilon^n(0)}$$

for some $\varepsilon > 0$. Since $B_\sigma \cong \overline{\mathbb{B}_1^{k+1}(0)}$, consider diffeomorphism $\phi : B_\sigma \to \overline{\mathbb{B}_1^{k+1}(0)}$ by sending $\varepsilon\cdot e_{\sigma(j)} \mapsto e_j$, and let $\pi : \mathbb{B}_1^{k+1}(0) \to \mathbb{D}^k$ be the orthogonal projection on equatorial plane. Denote $\{\mathrm{d}y_1,\ldots,\mathrm{d}y_k\}$ be the coordinate basis of $T\mathbb{D}^k$.
Notice that $\iota^*(\mathrm{d}x_{\tau(1)}\wedge\cdots\wedge\mathrm{d}x_{\tau(k+1)}) \neq 0$ only when $\tau = \sigma$, and $\iota^*\omega = 0$ when $\{1,\ldots,k\} \not\subset \sigma$; so there exists some $h_r \in C^\infty(M)$ such that

$$\mathrm{d}\omega = \sum_{\substack{r=1,\ldots,n\\ r\notin\{1,\ldots,k\}}} h_r(\mathrm{d}x_r\wedge\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_k)$$

To find out those h_r, we consider $\sigma = \{1,\ldots,k,r\}$. In this case, Let $\theta = \phi\circ\varphi|_{X_\sigma}$; we have that $\theta^*\omega|_y = 0$ for all y on the equator, so

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^{k+1}}\int_{\partial X_\sigma}\omega &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{\mathbb{D}^k}\rho(x)\cdot\left(f(x_+)-f(x_-)\right)(\mathrm{d}y_1\wedge\cdots\wedge\mathrm{d}y_k) \\ &= a_r\cdot\frac{\partial f}{\partial x_r} \qquad \text{(by exchanging limit and integral)}\end{aligned}$$

for some $\rho \in C^\infty(\mathbb{D}^k,\mathbb{R}_{\geq 0})$ and $a_r \in \mathbb{R}_{>0}$, where $\{x_+,x_-\} = \pi^{-1}(\{x\})$. Then similarly

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^{k+1}}\int_{X_\sigma}\mathrm{d}\omega = b_r\cdot h_r(p)$$

So by Stoke's theorem,

$$1 = \lim_{\varepsilon \to 0} \frac{\int_{\partial X_\sigma} \omega}{\int_{X_\sigma} \mathrm{d}\omega} = \frac{a_r}{b_r} \cdot \frac{1}{h_r(p)} \cdot \frac{\partial f}{\partial x_r}$$

Hence we now have,

$$\mathrm{d} : f \cdot (\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k) \longmapsto \sum_{\substack{r=1,\ldots,n \\ r \notin \{1,\ldots,k\}}} \frac{a_r}{b_r} \cdot \frac{\partial f}{\partial x_r} \cdot (\mathrm{d}x_r \wedge \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k)$$

Stoke's theorem tells us that $\mathrm{d}^2 = 0$, which means a_r/b_r are all equal; then by computing an explicit example, we can conclude that $a_r/b_r = 1$ and we're done by linearity of d. □

This explains how natural the definition of d is.
Since we have shown that d is a generalization of ∇, $\nabla\times$ and $\nabla\cdot$ in $\mathbb{R}^3$, we can also Stoke's theorem a generalization of those familiar theorems in vector calculus:

Example 3.11. By the principle we have in Definition 3.3, we can find the vector calculus version of Stoke's theorem in $\mathbb{R}^2$ and $\mathbb{R}^3$:

- **Green's theorem**: By letting $\omega = L\,\mathrm{d}x_1 + M\,\mathrm{d}x_2$,

$$\begin{aligned} \oint_{\partial X} L\,\mathrm{d}x_1 + M\,\mathrm{d}x_2 &:= \int_{\partial X} \omega = \int_X \mathrm{d}\omega \\ &=: \int_X \left(\frac{\partial M}{\partial x_1} - \frac{\partial L}{\partial x_2} \right) \mathrm{d}x_1 \wedge \mathrm{d}x_2 = \iint_X \left(\frac{\partial M}{\partial x_1} - \frac{\partial L}{\partial x_2} \right) \mathrm{d}x_1\,\mathrm{d}x_2 \end{aligned}$$

- **Stoke's theorem**: By letting $\omega = F_1\,\mathrm{d}x_1 + F_2\,\mathrm{d}x_2 + F_3\,\mathrm{d}x_3$,

$$\iint_X (\nabla \times F) \cdot \widehat{n}\ \mathrm{d}A := \int_X \mathrm{d}\omega = \int_{\partial X} \omega = \oint_{\partial X} F \cdot \mathrm{d}\mathbf{s}$$

- **Divergence theorem**: By letting $\omega = F_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + F_2\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 + F_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2$,

$$\iint_{\partial X} F \cdot \widehat{n}\ \mathrm{d}A := \int_{\partial X} \omega = \int_X \mathrm{d}\omega = \iiint_X \nabla \cdot F\ \mathrm{d}V$$

3.6 "$\mathrm{d}^2 = 0$"

[**Convention**] In this subsection, we write

- $\mathrm{d}_k = \mathrm{d}|_{\Omega^k(M)}$ for clarity.
- id_M as the identity map in $C^\infty(M, M)$.

In last subsection, we found out that ∇, $\nabla\times$ and $\nabla\cdot$ are simply representations of exterior derivative in a 3-dim manifold, and we also know in vector calculus that

$$\nabla \times (\nabla(f)) = 0 \quad \text{and} \quad \nabla \cdot (\nabla \times (F)) = 0$$

so Is it true that *two exterior derivatives always result in 0?* Surprisingly it is! This can be proved easily by induction using the axiom version of the definition of d.

Proposition 3.8. For $\forall \omega \in \bigwedge(T_p^* M)$, we always have $\mathrm{d}^2(\omega) = 0$.

How interesting that everything vanishes after taking 2 derivatives! *Recall a similar fact in homology: For $\forall c$ in a Chain complex, we always have $\partial^2(c) = 0$; is there any connection?* If we think further; since a form that is the exterior derivatives of the another has 0 derivative, a good question is

$$\text{"Is there any } \omega \text{ with } \mathrm{d}\omega = 0 \text{ but is NOT a derivative of others?"} \tag{$\dagger$}$$

Clearly the answer is NO for $M = \mathbb{R}^3$ (except for $\omega \in \Omega^0(M)$), but what about in other manifolds? To answer this question, we first define

Definition 3.7. The **closed k-form** is set $Z^k(M) := \ker(\mathrm{d}_k)$, and the **exact k-form** is set $B^k(M) := \mathrm{Im}(\mathrm{d}_{k-1})$.

$Z^k(M)$ are those with 0 derivative, and $B^k(M)$ are those which are the derivatives of other forms. Clearly we have that

- $B^k(M) \subseteq Z^k(M)$
- $B^k(M)$ and $Z^k(M)$ are vector spaces (thus groups under $+$)

These suggest that the non-trivial elements in the quotient group $Z^k(M)/B^k(M)$ will answer the question of "*What form has zero derivative but is NOT a derivative of another form?*". And we define this quotient group as **de-Rham Cohomology group**:

$$\boxed{H^k_{\mathrm{dR}}(M) := Z^k(M)/B^k(M)}$$

For example, by conclusions from Multivariable Calculus, we have $H^1_{\mathrm{dR}}(\mathbb{R}^3), H^2_{\mathrm{dR}}(\mathbb{R}^3) = \{0\}$. What's interesting about this group is that it is able to reveal some *topological traits* of a manifold: It is an *invariant under diffeomorphism*:

Proposition 3.9. If there is a diffeomorphism ϕ from smooth manifold X to Y, then we have $H^k_{\mathrm{dR}}(X) \cong H^k_{\mathrm{dR}}(Y)$ induced by pullback ϕ^*.

which is quite obvious, because diffeomorphism induces isomorphisms $\Omega^k(X) \cong \Omega^k(Y)$, and d commutes with ϕ^*.
Let's compute a few examples of simple M to find more properties;
Example 3.12. For $M = \mathbb{R}^2$,

$$H^0_{\mathrm{dR}}(\mathbb{R}^2) \cong \mathbb{R} \quad ; \quad H^2_{\mathrm{dR}}(\mathbb{R}^2) \cong \{0\}$$

$$\begin{aligned} H^1_{\mathrm{dR}}(\mathbb{R}^2) &= \left\{ f_1\,\mathrm{d}x_1 + f_2\,\mathrm{d}x_2 : \frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1} \right\} / \left\{ \frac{\partial g}{\partial x_1}\mathrm{d}x_1 + \frac{\partial g}{\partial x_2}\mathrm{d}x_2 : g \in C^\infty(\mathbb{R}^2) \right\} \\ &= \left\{ \frac{\partial g}{\partial x_1}\mathrm{d}x_1 + \frac{\partial g}{\partial x_2}\mathrm{d}x_2 : \frac{\partial^2 g}{\partial x_1 \partial x_2} = \frac{\partial^2 g}{\partial x_2 \partial x_1} \right\} / \left\{ \frac{\partial g}{\partial x_1}\mathrm{d}x_1 + \frac{\partial g}{\partial x_2}\mathrm{d}x_2 : g \in C^\infty(\mathbb{R}^2) \right\} \cong \{0\} \end{aligned}$$

Example 3.13. For $M = \mathbb{R}^3$, we can know from Multivariable calculus that "zero divergence means it is the curl of something" and "zero curl means it is the gradient of something"; so easily deduce that

$$H^0_{\mathrm{dR}}(\mathbb{R}^3) \cong \mathbb{R} \quad ; \quad H^1_{\mathrm{dR}}(\mathbb{R}^3), H^2_{\mathrm{dR}}(\mathbb{R}^3), H^3_{\mathrm{dR}}(\mathbb{R}^3) \cong \{0\}$$

Example 3.14. For $M = \mathbb{S}^1$, with coordinate expression $M = [-1,1]/\{\pm 1\}$, notice that $B^1(\mathbb{S}^1) = \{g'(x)\ \mathrm{d}x : g \in C^\infty(\mathbb{S}^1)\} = \left\{ f(x)\mathrm{d}x :\ \int_{-1}^x f(t)\mathrm{d}t \in C^\infty(\mathbb{S}^1) \right\}$, which are exactly those $f\,\mathrm{d}x \in \Omega^1(\mathbb{S}^1)$ that satisfy $\int_{-1}^1 f(t)\mathrm{d}t = 0$. And now since

$$\begin{aligned} Z^1(\mathbb{S}^1) &= \left\{ f(x)\,\mathrm{d}x :\ c \in \mathbb{R}\ ,\ \int_{-1}^1 f(t)\mathrm{d}t = c \right\} \\ &= \left\{ \left(f(x) + \frac{c}{2} \right)\ \mathrm{d}x :\ c \in \mathbb{R}\ ,\ \int_{-1}^1 f(t)\mathrm{d}t = 0 \right\} \end{aligned}$$

we have that $H^0_{\mathrm{dR}}(\mathbb{S}^1), H^1_{\mathrm{dR}}(\mathbb{S}^1) \cong \mathbb{R}$.
Example 3.15. For $\mathbb{R}^2 \setminus \{0\}$, we can verify that it is diffeomorphic to a ring shape, with coordinate representation $D_1 = ([-1,1] \times \mathbb{R}) / \sim$ where $(x_1, y_1) \sim (x_2, y_2)$ iff $x_1 = -x_2 = 1$ and $y_1 = y_2$.
Obviously $H^0_{\mathrm{dR}}(D_1) \cong \mathbb{R}$, $H^2_{\mathrm{dR}}(D_1) \cong \{0\}$.
For $H^1_{\mathrm{dR}}(D_1)$, we know that if we let $R := [-1,1] \times \mathbb{R}$, then

$$B^1(D_1) = \left\{ g\ \mathrm{d}x_1 + \left(\int_{-1}^{x_1} \frac{\partial g}{\partial x_2} dx_1 \right) \mathrm{d}x_2 :\ g \in C^\infty(R)\ ,\ \int_{-1}^1 g\ dx_1 = 0 \right\}$$

thus we assert that

$$\begin{aligned} H^1_{\mathrm{dR}}(D_1) &= \left\{ (c + f_1)\,\mathrm{d}x_1 + f_2\,\mathrm{d}x_2 :\ c \in \mathbb{R}\ ;\ \frac{\partial f_1}{\partial x_2} = \frac{\partial f_2}{\partial x_1} \right\} / B^1(D_1) \\ &= \left\{ (c + g)\,\mathrm{d}x_1 + \left(\int_{-1}^{x_1} \frac{\partial g}{\partial x_2} dx_1 \right) \mathrm{d}x_2 : c \in \mathbb{R}\ ;\ g \in C^\infty(R) \right\} / B^1(D_1) \cong \mathbb{R} \end{aligned}$$

3.7 Homotopy invariance & Poincaré's Lemma

By observing these examples, we can conjecture that the de rham cohomology group is a **homotopic invariance**. This is clearly much stronger than homeomorphic invariant which we concluded earlier.

Proposition 3.10. (Homotopic Invariant) Suppose smooth manifolds X to Y are homotopy equivalent, then $H^k_{\mathrm{dR}}(X) \cong H^k_{\mathrm{dR}}(Y)$.

(The homotopic equivalence is defined at Definition 1.1)
This conclusion follows from the fact that homotopic smooth maps give the same pullback:

Lemma 3.16. If $\phi_1, \phi_2 \in C^\infty(X, Y)$ are homotopic, then $\phi_1^*, \phi_2^* \in \mathrm{Hom}\left(H^k_{\mathrm{dR}}(Y), H^k_{\mathrm{dR}}(X)\right)$ are precisely equal.

Proof. Write d_X to be exterior derivative on $\Omega(X)$ and d_Y to be that on $\Omega(Y)$.
By Whitney Approximation theorem, there exists a *smooth* homotopy $\Phi : \mathbb{R} \times X \to Y$ from ϕ_1 to ϕ_2, where $\Phi(0, \cdot) = \phi_1$ and $\Phi(1, \cdot) = \phi_2$.
Let $\iota : X \hookrightarrow \mathbb{R} \times X$ be the natural inclusion; and let $\{\mathrm{d}x_1, \mathrm{d}x_2, \cdots\}$ be a basis on T^*X, $\{\mathrm{d}t, \mathrm{d}x_1, \mathrm{d}x_2, \cdots\}$ be a basis on $T^*(\mathbb{R} \times X)$, so that $\iota^*(\mathrm{d}t) = 0$ and $\iota^*(\mathrm{d}x_j) = \mathrm{d}x_j$.
Notice that for any $\omega \in Z^k(Y)$, we can write

$$(\Phi^*\omega)(t) = \rho(t) + \mathrm{d}t \wedge \theta(t)$$

where $\rho : \mathbb{R} \to \Omega^k(X)$ and $\theta : \mathbb{R} \to \Omega^{k-1}(X)$. As $\mathrm{d}_Y\omega = 0$, we have that

$$\begin{aligned} 0 = \Phi^*(\mathrm{d}_Y\omega) &= \mathrm{d}(\Phi^*\omega) \\ &= \mathrm{d}t \wedge \left(\frac{\mathrm{d}\rho}{\mathrm{d}t} + \mathrm{d}_X(\theta)\right) + \mathrm{d}_X(\rho) \end{aligned}$$

which suggests $\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\mathrm{d}_X(\theta)$. So

$$\begin{aligned} \phi_2^*\omega - \phi_1^*\omega = \iota^*\Phi^*\omega\big|_0^1 &= \int_0^1 \frac{\mathrm{d}\rho}{\mathrm{d}t}\ \mathrm{d}t \\ &= \mathrm{d}_X\left(-\int_0^1 \theta\ \mathrm{d}t\right) \in B^k(X) \end{aligned}$$

which gives $\phi_1^*([\omega]) = \phi_2^*([\omega])$. □

Then homotopy invariance (Proposition 3.10) follows immediately: If $f \in C^\infty(X, Y)$ and $g \in C^\infty(Y, X)$ is the pair of maps which satsifies that $(f \circ g)$ and $(g \circ f)$ are homotopic to identity, then by our lemma both $(f^* \circ g^*)$ and $(g^* \circ f^*)$ are group automorphism, which suggests that f^* and g^* are bijective, and thus isomorphisms.

This result immediately implies that all Euclidean spaces have the same de-rham cohomology group, as they are all homotopic.

Corollary 3.1. (Poincaré's Lemma) For $n \in \mathbb{Z}_{>0}$,

$$H^k_{\mathrm{dR}}(\mathbb{R}^n) \cong \begin{cases} \mathbb{R} & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

If we take a closer look at the last equation in the proof of Lemma 3.16, what we can find is an operator called the **Homotopy operator**. Consider the situation as in the lemma; the homotopy operator $h : \Omega^k(Y) \to \Omega^{k-1}(X)$ (for $k \in \mathbb{Z}_{>0}$) is a linear map such that

$$(\phi_1^* - \phi_2^*)\omega = (\mathrm{d} \circ h + h \circ \mathrm{d})\omega \tag{3.10}$$

for $\forall \omega \in \Omega^k(Y)$.

$$\begin{tikzcd}
\cdots \arrow[r, "\mathrm{d}"] & \Omega^{n-1}(X) \arrow[r, "\mathrm{d}"] & \Omega^n(X) \arrow[r, "\mathrm{d}"] & \Omega^{n+1}(X) \arrow[r, "\mathrm{d}"] & \cdots \\
\cdots \arrow[r, "\mathrm{d}"] & \Omega^{n-1}(Y) \arrow[r, "\mathrm{d}"] \arrow[u, shift left, "\phi_1^*"] \arrow[u, shift right, "\phi_2^*"'] & \Omega^n(Y) \arrow[r, "\mathrm{d}"] \arrow[u, shift left, "\phi_1^*"] \arrow[u, shift right, "\phi_2^*"'] \arrow[ul, "h"'] & \Omega^{n+1}(Y) \arrow[r, "\mathrm{d}"] \arrow[u, shift left, "\phi_1^*"] \arrow[u, shift right, "\phi_2^*"'] \arrow[ul, "h"'] & \cdots
\end{tikzcd}$$

Explicitly, for $\Phi^*\omega = (\rho + \mathrm{d}t \wedge \theta) \in \Omega^k(\mathbb{R} \times X)$, we have its expression

$$\begin{aligned} h : \Omega^k(Y) &\to \Omega^{k-1}(X) \\ \omega &\mapsto \int_0^1 \theta(\tau, \cdot)\ \mathrm{d}\tau \end{aligned}$$

we can easily verify that this is a well-defined homomorphism that satisfies the (3.10).
But how did we come up with such an operator? We can build our intuition by putting the simplest example into the proof of Lemma 3.16:

Example 3.17. Suppose $X = \mathbb{R}$ and $Y = \mathbb{R}^2$; maps $\phi_1(x) = (0, x)$ and $\phi_2(x) = (1, x)$ are homotopic by $\Phi(t, x) = (t, x)$.
Then for $\omega = f_1\ \mathrm{d}y_1 + f_2\ \mathrm{d}y_2 \in \Omega^1(Y)$, we have that

$$\Phi^*\omega|_{(t,x)} = f_1(t, x)\mathrm{d}x + f_2(t, x)\mathrm{d}t$$

which gives that

$$\begin{aligned} (h \circ \mathrm{d})\omega &= f_1(0, \cdot)\mathrm{d}x - f_1(1, \cdot)\mathrm{d}x - \left(\int_0^1 \frac{\partial f_2}{\partial x}\ \mathrm{d}t\right) \mathrm{d}x \\ (\mathrm{d} \circ h)\omega &= \left(\int_0^1 \frac{\partial f_2}{\partial x}\ \mathrm{d}t\right) \mathrm{d}x \end{aligned}$$

and so $(\phi_1^* - \phi_2^*)\omega = f_1(0, \cdot)\mathrm{d}x - f_1(1, \cdot)\mathrm{d}x = (h \circ \mathrm{d} + \mathrm{d} \circ h)\,\omega$.

so intuitively, $(\mathrm{d} \circ h)$ is the "correction" $(h \circ \mathrm{d})$:

The h can be thought as the line integral along the path $\gamma(t) = \Phi(t, x)$ for each fixed $x \in X$, and d is like the "differential". Based on this interpretation, $(h \circ \mathrm{d})\omega$ is like the line integral of the "differential" of ω, so it calculates the $(\phi_1^* - \phi_2^*)\omega$; but since the "differential" contains some irrelevant terms, we add $(\mathrm{d} \circ h)\omega$ to eliminate these impurities.

And this is where our introduction to background knowledge ends. We now know something about Calculus on the neighbour of a point on a manifold, which provides us with derivatives; but the way to generalize higher derivatives and integrals remains a question; Luckily they are all based on what we have learned, and we will answer this question respectively.

3.8 (Special) More about Hodge Dual

In this part, we will first discuss the reasoning behind the formation of Hodge dual, and then a short visit on *Hodge theory*. The technical part in this subsection contains no necessary knowledge for further development, but it is strongly recommended to understand the concept.

Previously we visualize the k-form over $V = \mathbb{R}^n$ as a k-dim subspace of $\mathbb{R}^n$, but does this intuition have a theoretical basis? There is, and it is called the **Plücker embedding**.
First, we come up with a topological way to define "the set of k-dim subspace of a vector space", the *Grassmannian*;

Definition 3.8. Define $\sim_k$ as an equivalence relation on V^k such that $A \sim_k B$ iff there exists $T \in \mathrm{GL}_k(\mathbb{R})$ such that $A = BT$. (where A, B are represented as $n \times k$ matrices)
The **Grassmannian** is the quotient space

$$\mathbf{Gr}_k(V) := V^k / \sim_k$$

where the V^k is endowed with the topology induced by metric $d(\mathbf{u}, \mathbf{v}) = \left(\sum_{j=1}^{k} ||u_j - v_j||^2\right)^{1/2}$.
Easy to see that there is a bijection between $\mathbf{Gr}_k(V)$ and the set of k-dim subspaces of V; the important part is that $\mathbf{Gr}_k(V)$ equips topology; for example, $\mathbf{Gr}_1(V) \cong \mathbb{RP}^{n-1}$, $\mathbf{Gr}_n(V) \cong \{0, 1\}$ with topology $\{\emptyset, \{1\}, \{0, 1\}\}$, and

$$\mathbf{Gr}_k(V) \cong \mathbf{Gr}_{n-k}(V)$$

What makes this interesting is that $\mathbf{Gr}_k(V)$ actually forms a smooth manifold:

Proposition 3.11. For $0 < k < n$, the $\mathbf{Gr}_k(V)$ is a $k(n-k)$-dim smooth manifold.

Proved in Q8 of Q&A in Section 2.

3.9 Questions

Those questions are sometimes used as a lemma in sections later on.

In questions below, we assume M an n-dim smooth manifold.

1. We know that wedge product is closely related to determinant;

(i) Suppose $A \in M_n(\mathbb{Z})$, show that for any $\lambda \in \mathbb{Z}$ such that $\lambda\mathbb{Z}^n \subset \operatorname{Im}(A)$, we have

$$\det(A)|\lambda^n$$

(ii) Show that $A \in M_n(\mathbb{C})$ is of rank k iff all its $k \times k$ minors all vanish.

2. Show that M is orientable if and only if there exists a no-where-zero volume form.

3.10 Answers

Question (1)

4 Basics: Algebra

In this section, we will study some bad languages. Abstract yet insightful when we observe similar structures in totally separated areas in this note.
This section is particularly boring if I don't introduce any application, but introducing them will blur my purpose; so this is a section for "tools", which means everything in this section can be treated like exercises and will be used somewhere else.

Example 4.1. There are some common phrases that we will use in this section,

1. Given a diagram (like the one below) with sets and arrows,

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"] & B \arrow[d, "h"] \\ C \arrow[r, "l"] & D \end{tikzcd}$$

   we say "**This diagram commutes**", if going from one set to another via different paths of arrows will give the same result. For example, in this case, it means

$$h \circ f = l \circ g$$

2. Consider a sequence of groups or rings or modules $\{A_1, A_2, A_3, \cdots\}$ equipped with homomorphism $f_k : A_k \to A_{k+1}$:

$$\cdots \longrightarrow A_{k-1} \xrightarrow{f_{k-1}} A_k \xrightarrow{f_k} A_{k+1} \xrightarrow{f_{k+1}} \cdots$$

   we say that this sequence is **Exact** if

$$\mathrm{Im}(f_k) = \ker(f_{k+1}) \quad \text{for } \forall k$$

3. Consider a sequence of groups or rings or modules $\{A_1, A_2, A_3, \cdots\}$ equipped with homomorphism $f_k : A_k \to A_{k+1}$:

$$\cdots \longrightarrow A_{k-1} \xrightarrow{f_{k-}} A_k \xrightarrow{f_k} A_{k+1} \xrightarrow{f_{k+1}} \cdots$$

   we call that this sequence a **chain complex** if

$$\mathrm{Im}(f_k) \subseteq \ker(f_{k+1}) \quad \text{for } \forall k$$

We will also encounter the reversed version of chain complex:

Definition 4.1. A **cochain complex** $A^\bullet$ is a sequence of abelian groups

$$\cdots \xleftarrow{f_{n+1}} A^{n+1} \xleftarrow{f_n} A^n \xleftarrow{f_{n-1}} A^{n-1} \xleftarrow{f_{n-2}} A^{n-2} \xleftarrow{f_{n-3}} \cdots$$

where f_k are homomorphism such that $\mathrm{im}(f_{k-1}) \subseteq \ker(f_k)$ for all k.

In context related to topology, the f_k is usually called *boundary maps* for chain complex or *differentials* for cochain complex.

Exact sequence is a very nice way to present a commutative algebraic structure via maps from other known structure; for example, we can express the idea of the kernal and cokernal of a linear map $\varphi : V \to W$ like this:

$$0 \longrightarrow \ker(\varphi) \xhookrightarrow{\subset} V \xrightarrow{\varphi} W \longrightarrow \operatorname{coker}(\varphi) \longrightarrow 0$$

Exact sequence basically helps to simplify our language.
Although *chain complex* gets motivation from topology, one can see that it has one thing in common with exact sequence: They always have

$$f^2 = 0$$

this somehow can correspond to the principal in topology, that $\partial^2 = \emptyset$ —— the boundary has no boundary; anyway, we will see about this later.

Like we had for $H^\bullet_{\mathrm{dR}}$, the *(co)homology* groups can measure the "exactness" of the boundary/differential maps of a (co)chain complex.

Definition 4.1 (Homology Groups)**.** Given a chain complex of R-modules

$$\cdots \xrightarrow{\partial_{k+2}} A_{k+1} \xrightarrow{\partial_{k+1}} A_k \xrightarrow{\partial_k} A_{k-1} \xrightarrow{\partial_{k-1}} A_{k-2} \xrightarrow{\partial_{k-2}} \cdots$$

we define the kth homology group to be

$$H_k(A) := (\ker(\partial_k) \,/\, \operatorname{im}(\partial_{k+1})\,, +)$$

The **cohomology** is the "dual" of homology; the main observation is that if we let $A^k := \operatorname{Hom}(A_k, R)$ and $\delta_k : A^k \to A^{k+1}$ be defined as $\delta_k(\omega) := \omega \circ \partial_k$, then the sequence

$$\cdots \xrightarrow{\delta_{k-2}} A^{k-1} \xrightarrow{\delta_{k-1}} A^k \xrightarrow{\delta_k} A^{k+1} \xrightarrow{\delta_{k+1}} A^{k+2} \xrightarrow{\delta_{k+2}} \cdots$$

gives a cochain complex. And similarly, the kth **cohomology group** is defined as

$$H^k(A) := (\ker(\delta_k) \,/\, \operatorname{im}(\delta_{k-1})\,, +)$$

Viewing separately, cochain and cohomology group have the same construction as chain and homology, so it seems to be meaningless. But we will actually see the difference, both algebraically and topologically.

4.1 Fibre bundles

This is to formalize our ideas about bundles and vector fields. The tangent bundle is a special case of some thing called Fibre bundle, which is a more topological construction;

Definition 4.2. A **Fibre Bundle** (E, B, π, F) is defined by topological spaces E, B, F and a continuous surjection $\pi : E \to B$, so that $\forall x \in E$, $\exists$ neighbour $U \subset B$ such that $\exists$ homeomorphism $\Phi : \pi^{-1}(U) \to U \times F$ which satisfies

$$(\mathrm{proj}_U \circ \Phi)(p) = \pi(p)$$

for $\forall p \in \pi^{-1}(U)$. (where $\mathrm{proj}_U : (u, x) \mapsto x$)

For convenience, we usually use E to denote the fiber bundle (E, B, π, F).
In this definition, E is called **total space**, B is called **base space**, F is called **fibre**, and Φ is called **local trivialization**. The local trivialization tells that, locally E looks like $B \times F$, though the fibers $\pi^{-1}(x)$ may be continuously "twisted".

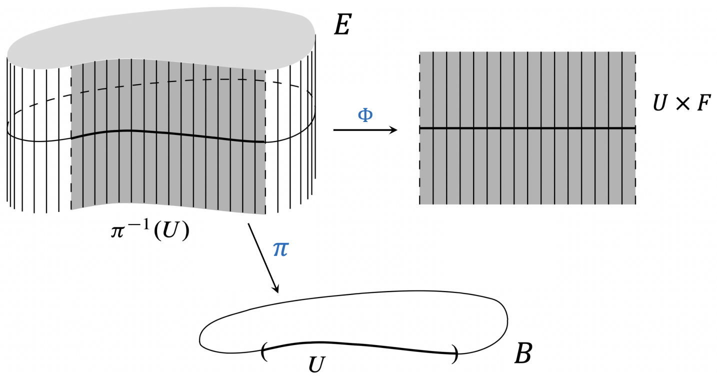


Figure 4: Digram from *Introduction to Smooth Manifold*, John Lee

This definition might seem to be a bunch of abstract nonsense, but it could be quite intuitive if we imagine about its special cases:

Example 4.2. The Tangent bundle is a fibre bundle for $E = TM$, $B = M$, $F = \mathbb{R}^n$.

Previously we define $TM = \{(p, v) : p \in M, v \in T_pM\}$, which can also be used in arbitrary E: For arbitrary fibre bundle (E, B, π, F), we can write

$$E := \left\{(x, v) : x \in B, v \in \pi^{-1}(x)\right\}$$

Another example seen before is the normal bundle $\mathrm{N}(M, X)$, where its fibre $F \cong \mathbb{R}^m$ for $m = \dim(X) - \dim(M)$. More generally, the TM and $\mathrm{N}(M, X)$ are **Vector bundles**, which are fibre bundles with fibre being a vector space.

The maps between fibre bundle over the same space are called *morphisms* (or *bundle maps*), which are basically maps between the total space with based space preserved.

Definition 4.2. (Bundle morphism) Given fibre bundles (ξ, X, π_1, F_1) and (ζ, Y, π_2, F_2), the pair of continuous maps $\phi : \xi \to \zeta$ and $\varphi : X \to Y$ is a morphism if the following diagram commutes

$$\begin{tikzcd} \xi \arrow[r, "\phi"] \arrow[d, "\pi_1"] & \zeta \arrow[d, "\pi_2"] \\ X \arrow[r, "\varphi"] & Y \end{tikzcd}$$

Further if $Y = X$, an **isomorphism** between ξ and ζ is a homeomorphism $\phi : \xi \to \zeta$ such that (ϕ, id_X) is a morphism.

Example 4.3. Suppose E_1, E_2 are fibre bundles over X, with a homeomorphism $\phi : E_1 \to E_2$ such that $\phi|_X = \mathrm{id}$; is it possible that E_1, E_2 are NOT isomorphic?
The answer is yes; but can you find one example?

When thinking about the question above, you might face the problem of constructing a given topological space into a fibre bundle; another question might come up is, given a total space and fibre, or given a total space and base space, is it always possible to construct a fibre bundle using them?
Surely not, a rather simple example which we will see later is $E = \mathbb{R}^2 \setminus \{p, q\}$ and $F = \mathbb{R}$. An important property of fibre bundle that helps in this problem is the *homotopy lifting property*:

Definition 4.3. Given topological spaces $X, \widetilde{Y}, Y$, map $\pi \in C^0(\widetilde{Y}, Y)$ has **homotopy lifting property** with respect to X if: For any homotopy $\Phi : [0,1] \times X \to Y$ and $\phi \in C^0(X, \widetilde{Y})$ such that $\pi \circ \phi = \Phi(0, \cdot)$, there exists a homotopy $\widetilde{\Phi} : [0,1] \times X \to \widetilde{Y}$ such that

- $\widetilde{\Phi}(0, x) = \phi$
- $\Phi(t, x) = \pi \circ \widetilde{\Phi}(t, x)$

This is lengthy, but would be intuitive: The set of possible images of Φ covers points in Y, and by picking different ϕ, we get local covering spaces of Y formed by $\widetilde{\Phi}$; then the union of all these local covering space makes $\widetilde{Y}$ looks like a fibre bundle locally.
Actually we define a "semi fibre bundle" using this, and it is called the *fibration*:

Definition 4.3. Given two topological spaces E, B, a **fibration** is a map $\pi : E \to B$ that has homotopy lifting property with respect to any X.

A fibre bundle over M (or any paracompact space) gives a fibration:

Proposition 4.4. A fibre bundle $F \to E \xrightarrow{\pi} B$ for a paracompact B is a fibration.

Proof. □

But the converse is not necessarily true: Consider $E = \{(x, y) \in \mathbb{R}^2 : |x| \geq |y|\}$, $B = \mathbb{R}\times\{0\} \subset \mathbb{R}^2$ and $\pi : (x, y) \mapsto (x, 0)$. (Left as an exercise to verify that it's a fibration.) The key difference is that the fibre bundle requires the fibre at each point be homeomorphic, yet fibration only requires homotopic equivalent.

The definition of fibration provides some insights in algebraic topology of fibre bundle:

Proposition 4.1. Given a fibre bundle $F \xrightarrow{\iota} E \xrightarrow{\pi} B$ and point $v_0 \in F = \pi^{-1}(x_0)$, we have a long exact sequence of homotopy groups

$$\begin{aligned} &\cdots \longrightarrow \pi_k(F, v_0) \xrightarrow{\iota} \pi_k(E, x_0) \xrightarrow{\pi} \pi_k(B, x_0) \longrightarrow \pi_{k-1}(F, v_0) \xrightarrow{\iota} \dots \\ &\cdots \longrightarrow \pi_0(F, v_0) \xrightarrow{\iota} \pi_0(E, x_0) \end{aligned}$$

The proof of this is not hard (see Page 376 in Hatcher's [2])

4.2 Vector bundle and Principal bundle

In this part, let M be a smooth manifold.

Vector bundles are fibre bundles with fibre being a vector space.

Definition 4.4. Giving a vector bundle $\mathbb{R}^r \to E \xrightarrow{\pi} X$, a local trivialization $\Phi : \pi^{-1}(U) \to X \times \mathbb{R}^r$, and a set of basis $\mathbb{R}^r = \operatorname{span}_{\mathbb{R}} \{\partial_1, \dots, \partial_r\}$, we call the $\{\Phi_*^{-1}\partial_1, \dots, \Phi_*^{-1}\partial_n\} \subset \Gamma(E)$ as a **local frame** of E.

Since fibres are vector spaces, we can extend the operations between vector spaces on vector bundles, like $\oplus$ and $\otimes$: For vector bundles E_1 and E_2 over the same base space X, we write

$$\begin{aligned} E_1 \oplus E_2 &:= \left\{(x, v_1 \oplus v_2) : x \in X \text{ and } (v_1, v_2) \in \pi_1^{-1}(x) \times \pi_2^{-1}(x)\right\} \\ E_1 \otimes E_2 &:= \left\{(x, v_1 \otimes v_2) : x \in X \text{ and } (v_1, v_2) \in \pi_1^{-1}(x) \times \pi_2^{-1}(x)\right\} \\ E_1^* &:= \left\{(x, v) : x \in X \text{ and } v \in \pi_1^{-1}(x)^*\right\} \end{aligned}$$

So we can decompose a vector bundle; together with the fact that sections of a vector bundle $(E, X, \pi, \mathbb{R}^r)$ can form a $C(X)$-module, this leads to some interesting algebra questions which we will talk about later.
Vector bundle over paracompact spaces can also be equipped with an inner product (an elment of $\Gamma(E^* \otimes E^*)$ such that it's an inner product when restricted to every fibre):

Proposition 4.2. Every vector bundle E over paracompact & Hausdorff space X admits an inner product.

It's the same as proving the existence of Riemannian metric.

Proof. It's clear that inner product exists on trivial E.
Recall that POU exists in paracompact Hausdorff spaces, so consider open cover $\{U_x\}$ with local trivializations $\Phi_x : \pi^{-1}(U_x) \to U_x \times \mathbb{R}^m$, and let $\{\rho_x\}$ be a POU. There exists an inner product g_x on each of $U_x \times \mathbb{R}^m$; and therefore the summing $g = \sum_x \rho_x g_x$ we have that g is a inner product on the entire X. □

Notice that vector bundle of a manifold is a manifold, but vector bundle of a *smooth* manifold is not necessarily smooth, so we define:

Definition 4.4. (*Smooth* Vector Bundle) A vector bundle $(E, M, \pi, \mathbb{R}^m)$ with M being a smooth manifold is *smooth* if for any $U \cap V \neq \emptyset$ with local trivializations Φ_U & Φ_V, there exists smooth $\varphi : U \cap V \to \mathrm{GL}_n(\mathbb{R})$ such that

$$\Phi_U \circ \Phi_V^{-1}(x, v) = (x, (\varphi(x))(v))$$

Verifying that the smooth vector bundle itself is a smooth manifold. So we can also construct some manifolds as the vector bundles on a manifold of smaller dimension: For example, the Mobius strip M can be a vector bundle on $\mathbb{S}^1$.
But it's clearly impossible for compact manifolds to be vector bundles (with rank > 0), so a good question is: Is every smooth *non-compact* manifold a vector bundle E of some M with $\operatorname{rank}(E) > 0$?
The answer is NO.

Example 4.5. The $\mathbb{R}^2 \setminus \{p, q\}$ cannot be the smooth vector bundle of any M with non-zero rank. $(p \neq q)$

Proof. This can be proved by two conclusions: For any connected smooth n-dim manifold M,

1. $\dim(H^n_{\mathrm{dR}}(M)) \in \{0, 1\}$, which will be proved in Section 5.3;
2. $H^k_{\mathrm{dR}}(M) \cong H^k_{\mathrm{dR}}(E)$ for $k = 1, \cdots, n$; left as an exercise

By these two conclusions, the fact $H^1_{\mathrm{dR}}(\mathbb{R}^2 \setminus \{p, q\}) \cong \mathbb{R}^2$ suggests that $\mathbb{R}^2 \setminus \{p, q\}$ cannot be any rank-1 vector bundle. □

This is to say, the vector bundle of manifolds is unable to equip too "complicated" structure; another way to view this is that there is a "restriction" on the rank of vector bundles:

Proposition 4.3. Given a rank-m smooth vector bundle E over M; if $m > n$, then there $\exists$ rank-n smooth vector bundle ξ such that

$$E = \xi \oplus \varepsilon$$

for some rank-$(m - n)$ trivial bundle ε.

This is to say, though $\operatorname{rank}(E)$ can be arbitrarily large, the "non-trivial part" is limited!
The proof is very geometrical and intuitive.

Proof. The key observation is that, the E has a rank-1 trivial subbundle, iff there exists a no-where-zero smooth section of E.
Equip E with an inner product g. Fix some $V \in \Gamma^\infty(E)$, notice that its image $W := \{(x, V(x)) : x \in X\}$ forms an n-dim submanifold of E. Now define $f \in C^\infty(E)$ by

$$f(x, v) := g|_x (V(x), V(x))$$

The restriction $f|_W$ is also a smooth function on W.
At point $(x, V(x)) \in W$, one can decompose $T_{(x,V(x))}E = T_{(x,V(x))}W \oplus T_{V(x)}(\pi^{-1}\{x\})$. Notice that since $n = \dim(T_{(x,V(x))}W) < \dim(T_{V(x)}(\pi^{-1}\{x\})) = m$, the set

$$Q_x := \left\{v \in T_{V(x)}(\pi^{-1}\{x\}) : (T_{(x,V(x))}W) \circ v(f) > 0\right\}$$

is non-empty for $\forall x \in M$.
For each $x \in X$, we have a canonical isomorphism $\phi : T_{V(x)}(\pi^{-1}\{x\}) \to \pi^{-1}\{x\}$.
Now we claim that there exists $\theta \in \Gamma^\infty(E)$ such that $\theta(x) \in \phi(Q_x)$ for $\forall x \in M$. (Because we can verify that is true in any local frame, it's sufficient to glue them together using a POU.)
By definition of Q_x, there exists $\varepsilon \in C^\infty(M)$ such that $f(x, V(x) + \varepsilon \cdot \theta(x)) > f(x, V(x))$, which means $V(x) + \varepsilon \cdot \theta(x)$ is a no-where-zero section.
Therefore, E has a rank-1 trivial subbundle ε, and so we can decomposite $E = \varepsilon^\perp \oplus \varepsilon$. As long as $\operatorname{rank}(\varepsilon^\perp) > n$, we can keep finding such rank-1 subbundle and decomposition, so we're done by induction. □

We will see (though very later) that this principle has a straight-forward algebraic explanation using the *Stiefel-Whitney class*.

Lastly, as a lot of topological spaces have symmetry in their structure, a group structure can be added to a fibre bundle to form **Principal bundle**, as it turns out that a lot of topological spaces can be constructed as the Principal bundle of some base space.

Definition 4.5. A **Principal G-bundle** is a fibre bundle (E, X, π, G) with topological group G acting on E by right multiplication that satisfies:

1. For $\forall g \in G$ and $x \in X$, we have $(\pi^{-1}\{x\})\, g \subseteq \pi^{-1}\{x\}$
2. The group action is free & transitive.

Intuitively, a principal G-bundle E is basically a fibre bundle with all fibre homeomorphic to G, and the action of $g \in G$ on E is a homeomorphism on every fibre. You may ask "why does the group action by G use the *right* multiplication?"; the first of the examples below might provide a good reason:

- Giving Lie group L and Lie subgroup G, the $(L, L/G, \pi, G)$ gives a principal G-bundle, where $\pi(x) = xG$. In this way, the group action of $g \in G$ must be right multiplication in order to satisfies the $1^{\text{st.}}$ condition.
- Given a vector bundle $(E, X, \pi, \mathbb{R}^n)$, the frame bundle of it is a principal $\mathrm{GL}_n(\mathbb{R})$-bundle $(F(E), X, \widetilde{\pi}, \mathrm{GL}_n(\mathbb{R}))$, constructed as
$$F(E) := \bigsqcup_{x \in X} F(\pi^{-1}(x))$$
where $\widetilde{\pi}(F(\pi^{-1}(x))) = \{x\}$ and
$$F(\pi^{-1}(x)) := \left\{(e_1, \ldots, e_n) \in \pi^{-1}(x)^n : \operatorname{span}(e_1, \ldots, e_n) \cong \mathbb{R}^n\right\}$$
An element $g \in \mathrm{GL}_n(\mathbb{R})$ acts on $F(E)$ by $g(e_1, \ldots, e_n) = (ge_1, \ldots, ge_n)$.
- We can construct $\mathbb{R}^{n+1}$ as a principal $\mathbb{R}$-bundle over $\mathbb{RP}^n$, and similarly, $\mathbb{S}^n$ as a principal $\mathbb{Z}/2\mathbb{Z}$-bundle over $\mathbb{RP}^n$.

- Another very important example is the **Hopf bundle** $(\mathbb{S}^3, \mathbb{S}^2, \pi, \mathbb{S}^1)$, which is a principal $\mathbb{S}^1$-bundle, where the group structure $\mathbb{S}^1 \cong \{z \in \mathbb{C} : ||z|| = 1\}$ under multiplication. It could be defined by letting

$$\mathbb{S}^3 = \left\{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\right\}$$
$$\pi(z_1, z_2) = \left(2 \cdot \Re(z_1\overline{z_2})\ ,\ 2 \cdot \Im(z_1\overline{z_2})\ ,\ |z_1|^2 - |z_2|^2)\right)$$

  and the group action of $s \in \mathbb{S}^1$ be $s \cdot (z_1, z_2) = (sz_1, sz_2)$.

The principal bundles are really not our concerns; it's just a good language to use.

4.3 Module and tensor product

Let R to be a non-trivial commutative ring (with unit).

A module is a "vector space" over a ring (rather than a field); this brings a lot of non-trivial properties that we take for granted in vector spaces. The definition of a module can be found in Wikipedia.

The module that is closest to a vector space is a *free module*;

Definition 4.6. A **Free R-module** M is a module isomorphic to the direct sum of countably many R; if it's finite, say n, then we denote $M = R^n$.

Remark. One that makes free modules very different from vector spaces is that, the "dimension" n in the above definition is NOT invariant! i.e. there $\exists R$ such that $R^n \cong R^m$ for $n \neq m$. An example of such R can be seen in Questions & Answers.

For non-free modules, we can still generate some of them as a span of some basis: Fix some elements $\alpha_1, \cdots, \alpha_n$ of a R-module M and define a surjective homomorphism $\phi : \mathbb{R}^n \to M$ such that,

$$\phi(r_1, \cdots, r_n) = \sum_{k=1}^{n} r_k \alpha_k$$

so all elements in M can be expressed in n coefficients. If such $\{\alpha_1, \cdots, \alpha_n\}$ exists, then we call M **Finitely Generated**, a module-version of *finite-dimensional*.

In this case, by first isomorphism theorem[7], we have

$$M \cong \mathbb{R}^n / \ker(\phi)$$

In linear algebra (when R is a field), this means $R^n = M \oplus \ker(\phi)$, but clearly this is not always true for modules (e.g When $n = 1$, $R = \mathbb{Z}/4\mathbb{Z}$ and $\ker(\phi) = \langle 2 \rangle$); but if it is satisfied, then we call M a *projective module*.

Definition 4.7. For a ring R, a R-module M is a **Projective module** if there exists a R-module N such that $M \oplus N$ is free.

The reason why there is a *"projective"* in its name is that, in Linear algebra (which is when R is field), the definition for ϕ being a "projection" is satisfying $\phi^2 = \phi$, which is then equivalent to $R^n = \operatorname{Im}(\phi) \oplus \ker(\phi)$. (try verifying this)

There are some other ways in which we define projective module, and some of them also bring some intuition into this concept:

[7] The 4 isomorphism theorems also have their corresponding version for modules

Example 4.6. Saying that M is a projective module is equivalent to say:

1. Given any two R-modules $\widetilde{X}$ and X with $p \in \mathrm{Hom}_R(\widetilde{X}, X)$ and $f \in \mathrm{Hom}_R(M, X)$, if p is subjective, then there $\exists \widetilde{f} \in \mathrm{Hom}_R(M, \widetilde{X})$ such that $p \circ \widetilde{f} = f$.
2. Any short exact sequence of R-modules $0 \to A \to B \to M \to 0$ splits.
3. There exists $\{(m_j, f_j) : j \in J\} \subset M \times \mathrm{Hom}(M, R)$ such that for $\forall z \in M$,

$$z = \sum_{j \in J} f_j(z) \cdot m_j$$

The proof can be seen here. (from Proofwiki)
We can also see that The functor $\mathrm{Hom}_R(M, \cdot)$ is exact.

One of the most important examples of projective module is the *smooth section* of a vector bundle. The simplest case is that the

Proposition 4.4. On any smooth vector bundle ξ over a compact smooth manifold X, $\Gamma^\infty(\xi)$ is a projective $C^\infty(X)$-module.

Proof. Clearly any smooth section of vector bundle is a $C^\infty(X)$-module, so we only need to show that $\Gamma^\infty(\xi)$ is projective.
Any finitely generated free $C^\infty(X)$-module is isomorphic to $\Gamma^\infty(\zeta)$ of some trivial bundle $\zeta = X \times \mathbb{R}^m$. Also notice that

$$\zeta \cong \xi \oplus \eta \implies \Gamma^\infty(\zeta) \cong \Gamma^\infty(\xi) \oplus \Gamma^\infty(\eta)$$

which will finish the proof.
Therefore, we wish to find such decomposition of ζ. Since we proved that vector bundle over smooth manifolds can have an inner product, we can use orthogonal decomposition.
Since there exists trivial bundle ζ such that ξ is a subbundle of it (exercise, see Question 3), we have that $\zeta = \xi \oplus \xi^\perp$. Verifying that $\xi^\perp$ is a smooth vector bundle (Exercise), the $\Gamma(\xi)$ is thus a projective module. $\square$

By our discussion on POU, X can be replaced with any paracompact Hausdorff space, and Γ^∞ with just Γ.

> *Remark* 4. In fact, the converse is also true, and can be generalized: Every finitely generated projective $C^1(X)$-module is isomorphic to $\Gamma^1(\xi)$ of some vector bundle ξ over a compact Hausdorff space X; this is known as the *Serre-Swan theorem.*

There are some even more "weaker" versions of free modules (see below), and interestingly all of them have some geometric meaning; however, up to this stage, it is not necessary to proceed beyond projective modules.

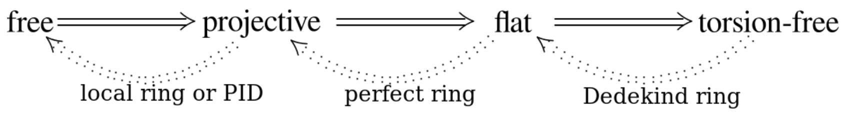


Figure 5: Diagram from Wikipedia.

Let's return to our original point, that using a free module F_0 (not necessarily finitely generated), a homomorphism $\phi_0 : F_0 \to M$ can represent a module M via $M \cong R^n/\ker(\phi)$. This gives the exact sequence

$$0 \longrightarrow \ker(\phi_0) \xrightarrow{\subseteq} F_0 \xrightarrow{\phi} M \longrightarrow 0$$

Note that $\ker(\phi)$ can also be represented the same way, so there exists another free module F_1 and $\phi_1 : F_1 \to \ker(\phi_0) \subseteq F_0$ that extends our exact sequence to

$$0 \longrightarrow \ker(\phi_1) \xrightarrow{\subseteq} F_1 \xrightarrow{\phi_1} F_0 \xrightarrow{\phi_0} M \longrightarrow 0$$

keep doing this, we can get a sequence of free modules, and this is a way to generate something called the *Free resolution* of M;

Definition 4.8. Given a R-module M, a (left) **resolution** of it is an exact sequence of R-modules

$$\cdots \xrightarrow{\phi_3} N_2 \xrightarrow{\phi_2} N_1 \xrightarrow{\phi_1} N_0 \xrightarrow{\phi_0} M \longrightarrow 0$$

where the ϕ_0 is called the augmentation map.
A free resolution requires N_j to be free modules.

Surely a resolution of M can tell something about M in some ways; one of such way is through (co)homology. Obviously one module can have more than one free resolutions, but we can show that they are all "homotopy equivalent":

Lemma 4.7. Consider the two resolutions of M

$$\begin{tikzcd} \cdots \arrow[r, "\phi_3"] & E_2 \arrow[r, "\phi_2"] \arrow[d, "f_2"] & E_1 \arrow[r, "\phi_1"] \arrow[d, "f_1"] & E_0 \arrow[r, "\phi_0"] \arrow[d, "f_0"] & M \arrow[r] \arrow[d, "\mathrm{id.}"] & 0 \\ \cdots \arrow[r, "\varphi_3"] & F_2 \arrow[r, "\varphi_2"] & F_1 \arrow[r, "\varphi_1"] & F_0 \arrow[r, "\varphi_0"] & M \arrow[r] & 0 \end{tikzcd} \tag{4.1}$$

and maps $\{f_k\} \subset \mathrm{Hom}_R(E_k, F_k)$ such that the diagram commutes.
If another $\{g_k\} \subset \mathrm{Hom}_R(E_k, F_k)$ also makes the diagram commutes, then $f_\bullet$ and $g_\bullet$ are chain homotopic.

Proof. To treat things more generally, let $M = E_{-1} = F_{-1}$ and $f_{-1} = g_{-1} = \mathrm{id}_M$; as a starter, we can let $h_{-1} : M \to F_0$ be a zero map.
Then for $\forall k \in \mathbb{N}$, we wish to construct $h_k : E_k \to F_{k+1}$ such that

$$(g_k - f_k) = (h_{k-1} \circ \phi_k - \varphi_{k+1} \circ h_k)$$

inductively (since h_{-1} is known). This would require that

$$h_k(x) \in \varphi_{k+1}^{-1}\left(h_{k-1} \circ \phi_k(x) + (f_k - g_k)(x)\right) \quad \text{for } \forall x \in E_k \tag{4.2}$$

which, by exactness, is equivalent to requiring $\varphi_k\left(h_{k-1} \circ \phi_k + (f_k - g_k)\right) = 0$. Notice that because the diagram commutes, we have that

$$\begin{aligned}\varphi_k\left(h_{k-1} \circ \phi_k + (f_k - g_k)\right) &= \varphi_k \circ h_{k-1} \circ \phi_k + (f_{k-1} - g_{k-1}) \circ \phi_k \\ &= \varphi_k \circ h_{k-1} \circ \phi_k - (h_{k-2} \circ \phi_{k-1} - \varphi_k \circ h_{k-1}) \circ \phi_k \\ &= -h_{k-2} \circ \phi_{k-1} \circ \phi_k = 0\end{aligned}$$

So (4.2) is always satisfied. Now by letting $\{x_1, x_2, x_3, \ldots\}$ be a basis of E_k, and pick y_j from the RHS of (4.2) to give $y_j = h_k(x_j)$, we construct the h_k inductively. □

We have encountered that chain homotopy implies the isomorphisms on (co)homology; but in our case in the lemma, the two rows are exact already, so homologies are trivial. However, it turns out that if we take the "dual" on the sequence, then we get isomorphisms between cohomologies:

Corollary 4.1. Let A be a commutative ring, then apply $\mathrm{Hom}(\cdot, A)$ on the commutative diagram (4.1); we have isomorphism

$$H^k(E_\bullet; A) \cong H^k(F_\bullet; A)$$

induced by f_k for all $k = 0, 1, 2, \ldots$.

This yields an invariant of those free resolutions of a module; and we call it the Ext functor[8].

Definition 4.9. Consider R-modules M and A, and a free resolution: $\cdots \to E_1 \xrightarrow{f_1} E_0 \xrightarrow{f_0} M \to 0$. Applying $\mathrm{Hom}_R(\cdot, A)$ produces

$$\cdots \longleftarrow \mathrm{Hom}_R(E_1, A) \xleftarrow{f_1^*} \mathrm{Hom}_R(E_0, A) \xleftarrow{f_0^*} \mathrm{Hom}_R(M, A) \longleftarrow 0$$

then **Ext functor** $\mathrm{Ext}_R^k(M, A) := \ker(f_{k+1}^*)/\mathrm{Im}(f_k^*)$, the cohomology at $\mathrm{Hom}_R(E_k, A)$.

This concept provides a way to characterize a free resolution. The Wikipedia page of it summarizes some key examples of it; for example, we always have $\mathrm{Ext}_R^0(M, A) = \mathrm{Hom}_R(M, A)$, and when M is a projective module, $\mathrm{Ext}_R^k(M, A) = 0$ for $\forall k > 0$.
The Wikipedia page also explains the reason for the "Ext" in its name, it is worthy to know, though it's unrelated to our development.

[8]you can see that there exists a slight difference between our definition and that in Wikipedia; but it's easy to verify that they are actually the same.

Finally, we now re-visit the tensor product, which we initially defined over vector spaces. Previously the definition of tensor product and tensor algebra is purely intuitive, by setting $V \otimes W = \operatorname{span}_{\mathbb{R}} \{v_j \otimes w_k\}$ (where v_j and w_k are basis of V and W respectively) that satisfies the bi-linearity:

$$(x + \lambda y) \otimes z = x \otimes z + \lambda(y \otimes z)$$
$$x \otimes (z + \lambda w) = x \otimes z + \lambda(x \otimes w)$$

for $\forall x, y \in V, z, w \in W$ and $\lambda \in \mathbb{R}$.
We can formalize and generalize this thought as:

Definition 4.10. Consider R-modules M_1 and M_2, let

$$H := \operatorname{span}_R \left\{ \begin{matrix} (u_1 + v_1, v_2) - (u_1, v_2) - (v_1, v_2)\,, \\ (u_1, u_2 + v_2) - (u_1, u_2) - (u_1, v_2)\,, \\ \lambda\,(v_1, v_2) - (\lambda v_1, v_2)\,, \\ \lambda\,(v_1, v_2) - (v_1, \lambda v_2) \end{matrix} : u_i, v_i \in M_i, \lambda \in R \right\}$$

Then, the tensor product of M_1 and M_2 is defined as

$$M_1 \otimes_R M_2 := {R[M_1 \times M_2]} \big/ {H}$$

where $R[M_1 \times M_2]$ is the free-module generated by $M_1 \times M_2$.

and as before, the tensor algebra $T(M)$ of M is the set

$$T(M) := \bigoplus_{n=0}^{\infty} \sum_{h_1, \ldots, h_n \in M} h_1 \otimes_R \cdots \otimes_R h_n$$

under $+$ and $\otimes_R$.

There are a lot of interpretations of the tensor product, as it becomes hard to really specify the real intuition when this operation is generalized to arbitrary ring. But there is one that could be useful here: For an R-module A, we have

$$R^n \otimes_R A \cong A^n$$
$$\text{via} \quad e_j \otimes a \mapsto a e_j$$

where $\{e_1, \ldots, e_n\}$ is basis of R^n. In other words, $\otimes_R A$ has the effect of "change of coefficient".

4.4 Some category and Functor

A category $\mathcal{C}$ is basically a set of algebraic structures, called "objects" $\mathrm{ob}(\mathcal{C})$, with maps between those objects, called "morphisms" $\mathrm{mor}(\mathcal{C})$, where some conditions applies, see definition in Wikipedia.
There are some common categories that will be mentioned in this part:

Set	all sets	maps
Top	all topological spaces	continuous map
Ring	all commutative rings with unit	ring homomorphism
Ab	all abelian group	group homomorphism
$\mathrm{Open}(X)$	open sets in X	inclusion
$\mathbf{Mod}_R$	all R-module	homomorphism
$\mathbf{Man}^\infty$	all smooth manifolds	smooth maps
$\mathbf{Vect}_\mathbb{F}$	all $\mathbb{F}$-vector fields	linear maps
$\mathbf{Vect}^n_\mathbb{R}(M)$	all rank n $\mathbb{R}$-vector bundles over M	bundle morphisms

Table 1: The category, its objects, and its morphisms

Categories can also be constructed by existing categories like:

Definition 4.5. Given categories $\mathcal{C}, \mathcal{D}$, and some $X \in \mathrm{ob}(C)$

1. $\mathcal{C}^{\mathrm{op}}$ denotes the Opposite category of $\mathcal{C}$, where $\mathrm{ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{ob}(\mathcal{C})$, and there exists a bijection $\mathrm{mod}(\mathcal{C}^{\mathrm{op}}) \longleftrightarrow \mathrm{mod}(\mathcal{C})$ such that each $f^{\mathrm{op}} : X \to Y$ corresponds to a $f : Y \to X$, and $f^{\mathrm{op}} \circ g^{\mathrm{op}} = (g \circ f)^{\mathrm{op}}$.
2. $\mathcal{C} \times \mathcal{D}$ denotes the Product category of $\mathcal{C}$ and $\mathcal{D}$, where $\mathrm{ob}(\mathcal{C} \times \mathcal{D}) = \mathrm{ob}(\mathcal{C}) \times \mathrm{ob}(\mathcal{D})$ and $\mathrm{mor}(\mathcal{C} \times \mathcal{D}) = \mathrm{mor}(\mathcal{C}) \times \mathrm{mor}(\mathcal{D})$.
3. $\mathcal{C}/X$ denotes the Slice Category of $\mathcal{C}$ over X, where $\mathrm{ob}(\mathcal{C}/X) = \{f \in \mathrm{mor}(\mathcal{C}) : \mathrm{cod}(f) = X\}$ and a morphism from elements $f : Y \to X$ to $g : Z \to X$ is a morphism $h : Y \to Z$ such that $f = g \circ h$.

The best example of opposite category is in basic algebraic geometry: The category of affine schemes is equivalent to $\mathbf{Ring}^{\mathrm{op}}$: Given any $A, B \in$ **Ring** and $f \in \mathrm{Hom}(A, B)$, the $f^{-1} \in \mathrm{Hom}(\mathrm{Spec}(B), \mathrm{Spec}(A))$ gives a map from schemes B to schemes in A, based on the fact that pre-image of a prime ideal is still a prime ideal.

The interesting point about Category is the mapping between Categories, like what we saw in the Galois correspondence in covering space and fundamental groups.

Definition 4.11. Given two categories $\mathcal{A}$ and $\mathcal{B}$, a (covariant) **Functor** F is a map from $\text{ob}(\mathcal{A})$ to $\text{ob}(\mathcal{B})$, so that for $\forall f, g \in \text{mor}(\mathcal{A})$, the $F(f), F(g) \in \text{mor}(\mathcal{B})$ satisfy

1. $F(\text{id}_{\mathcal{A}}) = \text{id}_{\mathcal{B}}$
2. $F(f \circ g) = F(f) \circ F(g)$

On the other hands, if F is a **Contravariant Functor**, then

1. $F(\text{id}_{\mathcal{A}}) = \text{id}_{\mathcal{B}}$
2. $F(f \circ g) = F(g) \circ F(f)$

basically, functors are "homomorphisms" in the level of category; they provide an analogy between different kinds of mathematical objects a formal language.
We can also view contravariant functor $\mathcal{C} \to \mathcal{D}$ as a (covariant) functor from $\mathcal{C}^{\text{op}} \to \mathcal{D}$. But still, it's just a simpler way to describe it.

Functors are actually everywhere in our previous discussion. There are also many good examples of functors listed in the Wikipedia page, here are some simple and crucial ones:

Example 4.8. Given R a commutative ring,

- Tensor product is a functor from $\mathbf{Mod}_R \times \mathbf{Mod}_R$ to $\mathbf{Mod}_R$, where it sends (V, W) to $V \otimes_R W$ and morphisms (f_V, f_W) to $f_V \otimes f_W$; where
$$(f_V \otimes f_W)(v \otimes w) := f_V(v) \otimes f_W(w)$$
- Also, fix an R-module A, the $\cdot \otimes A : \mathbf{Mod}_R \to \mathbf{Mod}_R$ is also a functor.
- Let $\mathbf{Man}^\infty$ be the category of smooth manifolds, where morphisms are smooth maps. The de-rham cohomology H_{dR} is a functor from $\mathbf{Man}^\infty$ to $\mathbf{Ring}$.
- Given smooth manifold M, the global section $\Gamma^\infty(\cdot)$ is a functor from the category of smooth vector bundles over M to $\mathbf{Mod}_{C^\infty(M)}$.
Actually in algebraic geometry, Given a topological space, $\Gamma(\cdot)$ is a contravariant functor called Sheaf. The sheaf is a functor from $\text{Open}(X)^{\text{op}}$ to a category like groups, or modules or just set depending on the context.
- Given any category C, the Hom functor $\text{Hom}(\cdot, C)$ is a contravariant functor from C to $\mathbf{Set}$. We will talk more about it.

Definition 4.6. For abelian groups A and B, we denote by $\text{Hom}(A, B)$ the set of all homomorphisms from A to B. Further, for any abelian group C and morphism $f : A \to C$, there is a natural map

$$\text{Hom}(f, B) : \text{Hom}(C, B) \to \text{Hom}(A, B) : \varphi \mapsto \varphi \circ f \tag{4.3}$$

which turns out to be a morphism of abelian groups.

For a fixed abelian group B, $\text{Hom}(\bullet, B)$ turns out to be a contravariant endofunctor in the category of abelian groups. With a picture,

$$
\begin{tikzcd}
 & A_2 \arrow[dr, "a_2"] & \\
A_1 \arrow[ur, "a_1"] \arrow[rr, "a_3"'] & & A_3
\end{tikzcd}
\quad \rightsquigarrow \quad
\begin{tikzcd}
 & \mathrm{Hom}(A_2, B) \arrow[dl, "\mathrm{Hom}(a_1\, B)"'] & \\
\mathrm{Hom}(A_1, B) & & \mathrm{Hom}(A_3, B) \arrow[ul, "\mathrm{Hom}(a_2\, B)"'] \arrow[ll, "\mathrm{Hom}(a_3\, B)"]
\end{tikzcd}
$$

An important property which we care about functors in abelian category is the **exactness**.

> **Definition 4.12.** Given $X, Y, Z \in \mathbf{ab}$, a short exact sequence $0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$, a functor F is called
>
> - **Exact**, if $0 \longrightarrow F(X) \xrightarrow{F(f)} F(Y) \xrightarrow{F(g)} F(Z) \longrightarrow 0$ is exact.
> - **Left exact**, if $0 \longrightarrow F(X) \xrightarrow{F(f)} F(Y) \xrightarrow{F(g)} F(Z)$ is exact.
> - **Right exact**, if $F(X) \xrightarrow{F(f)} F(Y) \xrightarrow{F(g)} F(Z) \longrightarrow 0$ is exact.

But the $\mathrm{Hom}(\cdot, R)$ functor is exact when R is an *injective module*, which won't be introduced here. Free modules are injective modules, and it is easy to verify this special case:

> **Lemma 4.9.** Endofunctor $\mathrm{Hom}(\cdot, R)$ of the category of free R-modules is exact.

An easy example to see that this is NOT true for non-free module: Take $R = \mathbb{Z}$ and $0 \longrightarrow \mathbb{Z} \xrightarrow{x \mapsto 2x} \mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow 0$, one can see that after applying $\mathrm{Hom}(\cdot, \mathbb{Z})$, the sequence $0 \longrightarrow 0 \longrightarrow \mathbb{Z} \xrightarrow{x \mapsto 2x} \mathbb{Z} \longrightarrow 0$ is NOT exact.

And finally, we'll briefly introduce the *limit* in algebra.
The intuition for algebraic limit is almost the same as that in analysis —— in fact, our definition of the algebraic limit could be generalized to any category, which would give a generalization to the limit in analysis.
The definition works for any abelian category $\mathcal{C}$;

> **Definition 4.13.** Given a directed set $(S, \leq)$.
> Suppose there exist $\{X_s : s \in S\} \subset \mathrm{ob}(\mathcal{C})$ and morphisms $f_{st} : X_s \to X_t$ for any $s \leq t$, such that
>
> - For any $s \in S$, the f_{ss} is identity map;
> - For any $a \leq b \leq c$, we have $f_{ac} = f_{bc} \circ f_{ab}$;
>
> Then $X_\bullet$ is called a **direct system**.
> Let $X := \bigsqcup_{s \in S} X_s$ and define an equivalence relation $\sim$ on X such that for $x_a \in X_a$ and $x_b \in X_b$, we have $x_a \sim x_b$ iff there $\exists c \in S$ such that $f_{ac}(x_a) = f_{bc}(x_b)$.
> Then $\varinjlim X_\bullet := X/\sim$ is called the **direct limit** of $X_\bullet$.

4.5 Some homological algebra (1)

Now we will turn to some theories; in the following discussion:

1. As how they're constructed, (co)homology groups can be seen as R-modules; therefore, from now on we assume them to be so.
2. In topology, (co)homology is usually produced by a (co)chain of **free** modules; therefore, from now on we assume them to be so.

The cohomology is costructed from cochains, which is the dual of chains, while chains give homology; so naturally we ask, *Can we get cohomology solely using homology?* By observing examples, there might indeed be an ambiguous relation.

Example 4.10. Let $R = \mathbb{Z}$, and consider a chain complex $C_\bullet(X)$

$$0 \longrightarrow \mathbb{Z}^2 \xrightarrow{\partial_2} \mathbb{Z}^3 \xrightarrow{\partial_1} \mathbb{Z}^3 \xrightarrow{\partial_0} \mathbb{Z}^3 \longrightarrow 0$$

with the boundary operators represented as matrices

$$\partial_2 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 2 & 0 \end{bmatrix} \quad \partial_1 = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 0 & 5 & 0 \end{bmatrix} \quad \partial_0 = \begin{bmatrix} 2 & 1 & 0 \\ 6 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and the cochain complex $C^\bullet(X)$:

$$0 \longrightarrow \mathbb{Z}^3 \xrightarrow{\delta_0} \mathbb{Z}^3 \xrightarrow{\delta_1} \mathbb{Z}^3 \xrightarrow{\delta_2} \mathbb{Z}^2 \longrightarrow 0$$

where $\mathbb{Z}^k \cong \mathrm{Hom}(\mathbb{Z}^k, \mathbb{Z})$ via $(a_1, \cdots, a_k) \mapsto \left((x_1, \cdots, x_k) \mapsto \sum_{j=1}^k a_j x_j\right)$ with the coboundary operators represented as matrices $\delta_j = \partial_j^T$ for $j = 1, 2, 3$. We can get that these give the (co)homology

$$\begin{aligned} &H_0(X) \cong \mathbb{Z}^2 \oplus \mathbb{Z}/3\mathbb{Z}, \quad H_1(X) \cong \mathbb{Z}/5\mathbb{Z}, \quad H_2(X) \cong \mathbb{Z}/2\mathbb{Z}, \quad H_3(X) \cong \mathbb{Z} \\ &H^0(X) \cong \mathbb{Z}^2, \quad H^1(X) \cong \{0\}, \quad H^2(X) \cong \mathbb{Z}/5\mathbb{Z}, \quad H^3(X) \cong \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} \end{aligned}$$

Though it seems that $H^k(X)$ depends on $H_{k-1}(X)$ and $H_k(X)$, the precise relation remains unclear. So let's investigate using some reasoning instead.

Let's consider a chain $\mathrm{C}_\bullet(X)$ and its cochain $\mathrm{C}_\bullet(X)$.
Since $\mathrm{C}^k(X) = \mathrm{C}_k(X)^\star$, elements in $H^k(X)$ naturally corresponds to those in $H_k(X)^\star$ via a map

$$\begin{aligned} \widetilde{h} : H^k(X) &\to H_k(X)^\star \\ [\omega] &\mapsto ([c] \mapsto \omega(c)) \end{aligned}$$

Easily checking that it's well-defined, and is a surjective homomorphism.
But this is not clearly enough without knowing $\ker(\widetilde{h})$. To find it out, we remove some

"restriction" of $\widetilde{h}$: The $\widetilde{h}$ is actually the composition $\widetilde{h} = h \circ \tau$, where

$$\begin{aligned} h : &H^k(X) \to Z_k(X)^\star \\ &[\omega] \mapsto (c \mapsto \omega(c)) \end{aligned}$$

and τ the quotient $c \mapsto [c]$. This h could be *no longer surjective*, as any $f \in Z_k(X)^\star$ is contained in its image iff $f|_{B_k(X)} = 0$. Hence,

$$\operatorname{Im}(h_k) = \ker(\mathscr{C}_k) \tag{4.4}$$

where $\mathscr{C}_k : Z_k(X)^\star \to B_k(X)^\star$ sends $f \mapsto f|_{B_k(X)}$.
Now consider $\ker(h_k)$; obviously $[\omega] \in \ker(h_k)$ when there $\exists \eta \in \mathrm{C}^{k-1}(X)$ such that $\omega = \eta \circ \partial_k$; which means, if we consider $\phi_{k-1} : B_{k-1}(X)^\star \to H^k(X)$ by sending $\eta \mapsto [\eta \circ \partial_k]$, then $\operatorname{Im}(\phi_{k-1}) \subseteq \ker(h_k)$.
Further, the "$\subseteq$" here can be improved to "$=$": Note that for $\forall \partial_k c \in B_{k-1}(X)$, we have that $\partial_k^{-1} \circ \partial_k c = c + Z_k(X)$, which means that $\forall [\omega] \in \ker(h_k)$, the $\omega \circ \partial_k^{-1} \in B_{k-1}(X)^\star$; therefore

$$\ker(h_k) = \operatorname{Im}(\phi_{k-1}) \tag{4.5}$$

Continue on, we can also find that $\operatorname{Im}(\mathscr{C}_k) = \ker(\phi_k)$. Together with (4.4) and (4.5), this builds a long exact sequence:

$$\begin{tikzcd} \cdots \arrow[r, "\mathscr{C}_{k-1}"] & B_{k-1}(X)^\star \arrow[r, "\phi_{k-1}"] & H^k(X) \arrow[r, "h_k"] & Z_k(X)^\star \arrow[dlll, "\mathscr{C}_k"] \\ B_k(X)^\star \arrow[r, "\phi_k"] & H^{k+1}(X) \arrow[r, "h_{k+1}"] & Z_{k+1}(X)^\star \arrow[r, "\mathscr{C}_{k+1}"] & \cdots \end{tikzcd}$$

the important part can then be extracted if we only care about $H^k(X)$:

$$0 \longrightarrow \operatorname{coker}(\mathscr{C}_{k-1}) \xrightarrow{\phi_{k-1}} H^k(X) \xrightarrow{h_k} \ker(\mathscr{C}_k) \longrightarrow 0$$

Doing this is useful only when $\operatorname{coker}(\mathscr{C}_{k-1})$ and $\ker(\mathscr{C}_k)$ can be determined by $H_\bullet(X)$. Luckily it's easy to see that $\ker(\mathscr{C}_k) \cong H_k(X)^\star$ via map $\omega \mapsto \omega \circ \tau$; in other words, we have exact sequence.

Lemma 4.11. For $k > 0$, there exists short exact sequence

$$0 \longrightarrow \operatorname{coker}(\mathscr{C}_{k-1}) \xrightarrow{\phi_{k-1}} H^k(X) \xrightarrow{\widetilde{h}_k} \operatorname{Hom}\left(H_k(X), R\right) \longrightarrow 0$$

Hence we'll need to worry about $\operatorname{coker}(\mathscr{C}_{k-1})$.
The key observation is that our $\mathscr{C}_{k-1}$ is "dual" to the natural inclusion $B_{k-1}(X) \hookrightarrow Z_{k-1}(X)$, which produces short exact sequence

$$0 \longrightarrow B_{k-1}(X) \xrightarrow{\subset} Z_{k-1}(X) \xrightarrow{\tau} H_{k-1}(X) \longrightarrow 0$$

applying $\mathrm{Hom}(\cdot, R)$ on it will produce $\mathscr{C}_{k-1}$:

$$0 \longrightarrow H_{k-1}(X)^\star \xrightarrow{\tau_*} Z_{k-1}(X)^\star \xrightarrow{\mathscr{C}_{k-1}} B_{k-1}(X)^\star \longrightarrow 0$$

Be aware that sequence above is *the dual of a free resolution* of $H_{k-1}(X)$, and our $\mathrm{coker}(\mathscr{C}_{k-1})$ is the $1^{\text{st.}}$ cohomology group of it. i.e.

$$\mathrm{coker}(\mathscr{C}_{k-1}) = \mathrm{Ext}^1_R(H_{k-1}(X), R)$$

By Corollary 4.1, it depends solely on $H_{k-1}(X)$ —— so we are done:

Proposition 4.5. For $k > 0$, there exists a short exact sequence

$$0 \longrightarrow \mathrm{Ext}^1_R(H_{k-1}(X), R) \xrightarrow{\phi_{k-1}} H^k(X) \xrightarrow{\widetilde{h}_k} \mathrm{Hom}\left(H_k(X), R\right) \longrightarrow 0$$

achieving our goal of expressing cohomology solely usinig homology.
I believe that the derivation of this conclusion is very satisfactory, and it gives an idea of how to use the tool of exact sequence in analyzing structures.
If you observe further, two more features can be improved in our conclusion:

1. This exact sequence actually splits;
2. $\mathrm{Hom}(\cdot, R)$ in our derivation can be replaced with $\mathrm{Hom}(\cdot, A)$ for arbitrary R-module A.

Then we get to a fundamental theorem in homological algebra:

Universal Coefficient theorem (cohomology)

Theorem 4.12. Consider an R-module A, then for $k = 1, 2, 3, \dots$, there exists a split short exact sequence

$$0 \longrightarrow \mathrm{Ext}^1_R(H_{k-1}(X), A) \xrightarrow{\phi_{k-1}} H^k(X; A) \xrightarrow{\widetilde{h}_k} \mathrm{Hom}\left(H_k(X), A\right) \longrightarrow 0$$

(where the definitions of ϕ_k and $\widetilde{h}_k$ stay the same, just replace $\mathrm{Hom}(\cdot, R)$ with $\mathrm{Hom}(\cdot, A)$.)
The proof is left as an exercise.
Now we can explain the pattern found at the beginning, the case of $A = R = \mathbb{Z}$:
Any finitely generated $\mathbb{Z}$-module is of the form $T \oplus \mathbb{Z}^r$ for some finite abelian group T, which we call the *torsion part*, and $r \in \mathbb{N}$ is called the *rank*. So suppose $H_k(X)_{\mathrm{tor}}$ and r_k is the torsion part and the rank of $H_k(X)$ respectively, then

$$H^k(X) \cong H_{k-1}(X)_{\mathrm{tor}} \oplus \mathbb{Z}^{r_k} \tag{4.6}$$

which is a very satisfactory result.

4.6 Some homological algebra (2)

In this part, we write $\otimes_R$ as $\otimes$ for simplicity.

In topology, homology over $\mathbb{Z}$ usually tells more topological traits than over other rings (there is actually a good reason for this), and of course finding it is also harder. The Universal coefficient theorem (UCT) also makes it algebraically more useful —— it states that if we know $H_\bullet(X;\mathbb{Z})$, then we know $H_\bullet(X;A)$ for any $\mathbb{Z}$-modules A.

Universal Coefficient theorem (homology)

Theorem 4.13. Consider an R-module A, then for $k = 1, 2, 3, \ldots$, there exists a split short exact sequence

$$0 \longrightarrow H_k(X;R) \otimes A \xrightarrow{\widetilde{h}_k} H_k(X;A) \xrightarrow{\phi_k} \operatorname{Tor}_1^R(H_{k-1}(X), A) \longrightarrow 0$$

The statement is very much alike the cohomology version in last part; and you may find that algebraically they share a similar reasoning.

Proof. The story starts with the property that $R^n \otimes A \cong A^n$, and since $\otimes$ is a right-exact functor, which means tensor product $\mathrm{C}_\bullet(X) \otimes A$ is still a chain complex, where we have

$$\partial : c \otimes a \mapsto (\partial c) \otimes a$$

Therefore, since $\mathrm{C}_k(X) \otimes A \cong \mathrm{C}_k(X;A)$, the $H_k(X;A)$ is simply the homology of $\mathrm{C}_\bullet(X) \otimes A$, which means there exists a natural map

$$\begin{gathered}\widetilde{h} : H_k(X;R) \otimes A \to H_k(X;A) \\ [c] \otimes a \mapsto [c \otimes a]\end{gathered}$$

for $c \in Z_k(X)$ and $a \in A$.
Now as before, we consider its "reduced" version: $h : Z_k(X;R) \otimes A \to H_k(X;A)$, and extend it to a long exact sequence by studying its kernal and image

$$\begin{tikzcd} \cdots \arrow[r, "\mathscr{C}_{k+1}"] & Z_{k+1}(X) \otimes A \arrow[r, "h_{k+1}"] & H_{k+1}(X;A) \arrow[r, "\phi_{k+1}"] & B_k(X) \otimes A \arrow[dll, "\mathscr{C}_k"] \\ & Z_k(X) \otimes A \arrow[r, "h_k"] & H_k(X;A) \arrow[r, "\phi_k"] & B_{k-1}(X) \otimes A \arrow[r, "\mathscr{C}_{k-1}"] & \cdots \end{tikzcd}$$

where the maps are $h : c \otimes a \mapsto [c \otimes a]$, $\phi : [c \otimes a] \mapsto \partial c \otimes a$, and $\mathscr{C} : c \otimes a \mapsto \iota(c) \otimes a$, where $\iota : B_k(X) \hookrightarrow Z_k(X)$ is the natural inclusion. (The derivation left as an exercise. You might be confused that *why aren't $\phi = 0$ and $\mathscr{C}$ always injective?* It would help by studying a specific example.)
Then at each level, we can extract a short exact sequence

$$0 \longrightarrow \operatorname{coker}(\mathscr{C}_k) \xrightarrow{h_k} H_k(X;A) \xrightarrow{\phi_k} \ker(\mathscr{C}_{k-1}) \longrightarrow 0 \tag{4.7}$$

Easy to see that $\mathrm{coker}(\mathscr{C}_k) = H_k(X) \otimes A$. Then notice that by applying $\cdot \otimes A$ on the free resolution $0 \longrightarrow B_{k-1}(X) \xrightarrow{\subset} Z_{k-1}(X) \longrightarrow H_{k-1}(X) \longrightarrow 0$, we have

$$\ker(\mathscr{C}_{k-1}) = \mathrm{Tor}_1^R(H_{k-1}(X), A)$$

by definition. And we're done. $\square$

Example 4.14. Consider $R = \mathbb{Z}$ and $A = \mathbb{Z}/p\mathbb{Z}$ for a prime number p; suppose each $H_k(X)$ is finitely generated and is decomposed as

$$H_k(X) = \bigoplus_{j=1}^{m} (\mathbb{Z}/n_{kj}\mathbb{Z}) \oplus \mathbb{Z}^{r_k}$$

for some $n_{kj} \in \mathbb{Z}_{>1}$. Now we let $c_k := \left|\{n_{kj} : p|n_{kj}\}\right|$.
Then using the fact that $\mathrm{Tor}_\bullet^R(P \oplus Q, A) = \mathrm{Tor}_\bullet^R(P, A) \oplus \mathrm{Tor}_\bullet^R(Q, A)$, we have

$$\begin{aligned} \mathrm{Tor}_1^R(H_{k-1}(X), A) &= \bigoplus_{j=1}^{m} \mathrm{Tor}_1^R(\mathbb{Z}/n_{(k-1)j}\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}) \\ &= (\mathbb{Z}/p\mathbb{Z})^{c_{k-1}} \end{aligned}$$

hence, $H_k(X; A) \cong (\mathbb{Z}/p\mathbb{Z})^{c_k + c_{k-1} + r_k}$

In both UCT, the key in the proofs is a long exact sequence which comes from the study of $\mathrm{coker}(h)$ and $\ker(h)$; there must be some common construction.
You could observe the similarities. In both two long exact sequences, the three terms in each row are the homology of the chain complexes in the three columns below respectively:

$$\begin{tikzcd}
 & \vdots \arrow[d, "F\circ\partial"] & \vdots \arrow[d, "F\circ\partial"] & \vdots \arrow[d, "F\circ\partial"] & \\
0 \arrow[r] & F \circ Z_{k+1}(X) \arrow[r, "\subset"] \arrow[d, "F\circ\partial"] & F \circ \mathrm{C}_{k+1}(X) \arrow[r, "\tau"] \arrow[d, "F\circ\partial"] & F \circ B_k(X) \arrow[r] \arrow[d, "F\circ\partial"] & 0 \\
0 \arrow[r] & F \circ Z_k(X) \arrow[r, "\subset"] \arrow[d, "F\circ\partial"] & F \circ \mathrm{C}_k(X) \arrow[r, "\tau"] \arrow[d, "F\circ\partial"] & F \circ B_{k-1}(X) \arrow[r] \arrow[d, "F\circ\partial"] & 0 \\
 & \vdots & \vdots & \vdots &
\end{tikzcd} \tag{4.8}$$

where F is the functor $\mathrm{Hom}(\cdot, R)$ and $\otimes A$ in the two UCT respectively. For example, in case of $F = \otimes A$, the $F \circ \partial$ on the $1^{\text{st.}}$ and $3^{\text{rd.}}$ column are zero maps, so the homologies are just $B_\bullet(X) \otimes A$ and $Z_\bullet(X) \otimes A$ respectively; and the homology on the middle column would be $H_\bullet(X; A)$, the homology of $\mathrm{C}_\bullet(X) \otimes A$.
Then, to form a long exact sequence, a map from $F \circ B_k(X) \to F \circ Z_k(X)$ is required. From the two UCT, we may conjecture that the map is defined by

$$\mathscr{C} = \iota^{-1} \circ (F \circ \partial) \circ \tau^{-1}$$

Let's verify it in a general setting.

Snake Lemma (homology)

Theorem 4.15. Suppose we have a short exact sequence of complexes $A^\bullet, B^\bullet, C^\bullet$

$$0 \longrightarrow A^\bullet \xrightarrow{f} B^\bullet \xrightarrow{g} C^\bullet \longrightarrow 0$$

such that homomorphisms f, g commute with boundary maps, then there exists maps $\mathscr{C}_k : H_k(C) \to H_{k-1}(A)$ producing a long exact sequence

$$\begin{tikzcd} \cdots \arrow[r, "\mathscr{C}"] & H_k(A) \arrow[r, "\widetilde{f}"] & H_k(B) \arrow[r, "\widetilde{g}"] & H_k(C) \arrow[dll, "\mathscr{C}"] & \\ & H_{k-1}(A) \arrow[r, "\widetilde{f}"] & H_{k-1}(B) \arrow[r, "\widetilde{g}"] & H_{k-1}(C) \arrow[r, "\mathscr{C}"] & \cdots \end{tikzcd}$$

where maps $\widetilde{f}$ and $\widetilde{g}$ induced by f and g.

Proof. Set $\widetilde{f}([\omega]) = [f(\omega)]$ and $\widetilde{g}([\omega]) = [g(\omega)]$; easy to verify that they are well-defined. Then define $\mathscr{C}_k([c]) := g_k \circ \partial_k \circ f_k([c])$. We need to verify 2 things:

1. $\mathscr{C}_k : H_k(C) \to H_{k-1}(A)$ is well-defined homomorphism.
2. The sequence is exact.

□

We begin with an important result on combinations of exact sequences.

Proposition 4.6 (Four Lemmas)**.** Consider the following commutative diagram with exact rows. Let α be surjective and δ be injective.

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] \arrow[d, two heads, "\alpha"] & A_2 \arrow[r, "f_2"] \arrow[d, "\beta"] & A_3 \arrow[r, "f_3"] \arrow[d, "\gamma"] & A_4 \arrow[d, hook, "\delta"] \\ B_1 \arrow[r, "g_1"] & B_2 \arrow[r, "g_2"] & B_3 \arrow[r, "g_3"] & B_4 \end{tikzcd} \tag{4.9}$$

1. If γ is surjective, then so is β.
2. If β is injective, then so is γ.

Proof. The proof of this kind of statement is simply diagram tracing, very simple yet tedious. So we won't present the proof explicitly except for this time.

1. For the first case, try showing that arbitrary $b \in B_2$ satisfies $b \in \mathrm{im}(\beta)$.
   By surjectivity of γ, $\exists a_3 \in A_3$ such that $\gamma(a_3) = g_2(b) =: b_3$. Since $b_3 \in \mathrm{im}(g_2) = \ker(g_3)$, $g_3(b_3) = 0$. Since (4.9) commutes, $f_3(a_3) \in \ker(\delta) = \{0\}$ because δ is injective.

Therefore, $a_3 \in \ker(f_3) = \operatorname{im}(f_2)$. Then, $\exists a_2 \in A$ such that $f_2(a_2) = a_3$.
Indeed, by commutativity of (4.9), we have that $g_2(\beta(a_2)) = b_3 = g_2(b) = \gamma(f_2(a_2))$. Then, we must have that $b - \beta(a_2) \in \ker(g_2) = \operatorname{im}(g_1)$. Therefore, there exists $b_1 \in B_1$ such that $g_1(b_1) = b - \beta(a_2)$. By surjectivity of α, there additionally exists $a_1 \in A_1$ such that $\beta(f_1(a_1)) = \beta(a_2) - b$. Therefore, $b = \beta(a_2 - f_1(a_1)) \in \operatorname{im}(\beta)$, as required.

2. For the second case, fix $a \in \ker(\gamma)$ and try showing that $a = 0$.
   That $g_3(\gamma(a)) = 0$ is evident. Since (4.9) commutes, $\delta(f_3(a)) = 0$ as well. Since δ is injective, $f_3(a) = 0$ and therefore $a \in \ker(f_3) = \operatorname{im}(f_2)$, meaning that $\exists a_2 \in A_2$ such that $f_2(a_2) = a$.
   Now since (4.9) commutes, $g_2(\beta(a_2)) = \gamma(f_2(a_2)) = 0$. Therefore, $b_2 := \beta(a_2) \in \ker(g_2) = \operatorname{im}(g_1)$. Hence, $\exists b_1 \in B_1$ such that $g_1(b_1) = b_2$. Then, by surjectivity of α, $\exists a_1 \in A_1$ such that $\alpha(a_1) = b_1$. Therefore, by commutativity of (4.9), we get that $\beta(a_2) = g_1(\alpha(a_1)) = \beta(f_1(a_1))$, meaning that $f_1(a_1) - a_2 \in \ker(\beta)$. Since β is injective, $\ker(\beta) = \{0\}$, meaning that $f_1(a_1) = a_2$. Hence, $a_2 \in \operatorname{im}(f_1) = \ker(f_2)$, meaning that $a = f_2(a_2) = 0$ as required.

$\square$

This yields the following useful Corollary:

Corollary 4.2 (Five Lemma)**.** Consider the following commutative diagram with exact rows.

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] \arrow[d, two heads, "\alpha"] & A_2 \arrow[r, "f_2"] \arrow[d, "\rho", "\wr"'] & A_3 \arrow[r, "f_3"] \arrow[d, "\gamma"] & A_4 \arrow[r, "f_4"] \arrow[d, "\eta", "\wr"'] & A_5 \arrow[d, hook, "\beta"] \\ B_1 \arrow[r, "g_1"] & B_2 \arrow[r, "g_2"] & B_3 \arrow[r, "g_3"] & B_4 \arrow[r, "g_4"] & B_5 \end{tikzcd} \tag{4.10}$$

Suppose α is surjective, β is injective, and ρ and η are isomorphisms, then γ is an isomorphism as well.

We leave the proof as an exercise to the reader, offering only the hint that the Five Lemma can be seen as a combination of the Four Lemmas.

4.7 Singular homology

It's recommended to have knowledge about simplicial homology.
[*Convention*]: In this part, we denote:
- X as a non-empty topological space;
- $\triangle_k$ as the k-simplex, and $\mathrm{C}_k^{\triangle}(\cdot)$ as a k-simplicial complex;
- $\partial_k^{\triangle}$ as the boundary map in $\mathrm{C}_k^{\triangle}(\cdot)$.

The homology groups were first invented for distinguishing the surfaces as a topological invariant by Henri Poincaré, so it's time now to have some topological perspectives on what we learned in last subsection.
Singular homology is nicely explained in Wikipedia, though we will still state its definition here for convenience:

Definition 4.14. The **singular complex** $\mathrm{C}_k(X)$ is a free $\mathbb{Z}$-module generated by $\{\sigma \in C^0(\triangle^k, X)\}$ equipped with boundary map $\partial_k : \sigma \mapsto \sigma \circ \partial_k^{\triangle}$.
The **Singular homology** is defined as the quotient group

$$H_k(X) := \ker(\partial_k)/\mathrm{Im}(\partial_{k+1}) \quad \text{for } k \in \mathbb{N}$$

There is little physical intuition in singular homology, and *almost no elementary way to compute* —— unlike in simplicial case, we compute by triangulation. However, also due to the "freeness" in such definition, singular homology works for any topological space, and therefore it's more suitable for developing theory.
Singular homology can be seen as an abtract generalization of simplicial homology, as

Lemma 4.16. For any topological space X being a union of simplicial simplex,

$$H_k^{\triangle}(X) \cong H_k(X) \quad \text{for } k \in \mathbb{N}$$

(Proof in Question & Answer)
So in short: Simplicial homology computes simple cases, singular homology helps us to develop theory on more complicated cases, based on those simple cases.

As a topological invariant, the $H_\bullet(X)$ is homotopy invariant.

Proposition 4.7. For homotopy equivalent X, Y, we have $H_\bullet(X) \cong H_\bullet(Y)$.

We have seen this before in de-rham cohomology, and the idea of proof is similar: prove that homotopic maps $f, g : X \to Y$ have the same pushforward, which is done by **homotopy**

operator (also *prism operator*).

Proof. Let $\Phi : X \times [0,1] \to Y$ be a homotopy from f to g and $\iota : X \hookrightarrow X \times [0,1]$ be the natural inclusion. Consider any singular simplex $\sigma : \triangle_n \to X$ in X; the $\triangle_n \times [0,1]$ is clearly not a simplex, but it can be divided into $(n+1)$ of $\triangle_{n+1}$ by:

$$[e_0, \cdots, e_n] \times [0,1] = \bigcup_{k=1}^{n} [e_0, \cdots, e_k, \widetilde{e}_k, \cdots, \widetilde{e}_n]$$

where $[e_0, \cdots, e_n] = \triangle_n \times \{0\}$ and $[\widetilde{e}_0, \cdots, \widetilde{e}_n] = \triangle_n \times \{1\}$. (e.g. When $n = 2$, this operation is like dividing a prism into 3 tetrahedrons)
Then we define the homotopy operator $P : \mathrm{C}_n(X) \to \mathrm{C}_{n-1}(Y)$ as

$$P : \sigma \mapsto \sum_{k=0}^{n} (-1)^k \cdot \Phi \circ (\sigma \times \mathrm{id}_{[0,1]})[e_0, \cdots, e_k, \widetilde{e}_k, \cdots, \widetilde{e}_n] \tag{4.11}$$

where $(\sigma \times \mathrm{id}_{[0,1]}) : \triangle_n \times [0,1] \to X$ sends $(v,t) \mapsto (\sigma(v), t)$.

$$\begin{tikzcd}
\cdots \arrow[r, "\partial"] & \mathrm{C}_{n-1}(X) \arrow[r, "\partial"] \arrow[d, shift right, "f_*"'] \arrow[d, shift left, "g_*"] & \mathrm{C}_n(X) \arrow[r, "\partial"] \arrow[dl, "P"'] \arrow[d, shift right, "f_*"'] \arrow[d, shift left, "g_*"] & \mathrm{C}_{n+1}(Y) \arrow[r, "\partial"] \arrow[dl, "P"'] \arrow[d, shift right, "f_*"'] \arrow[d, shift left, "g_*"] & \cdots \\
\cdots \arrow[r, "\partial"] & \mathrm{C}_{n-1}(Y) \arrow[r, "\partial"] & \mathrm{C}_n(Y) \arrow[r, "\partial"] & \mathrm{C}_{n+1}(Y) \arrow[r, "\partial"] & \cdots
\end{tikzcd}$$

As we've seen before, it satisfies $\partial \circ P + P \circ \partial = g_* - f_*$, verifying this could be very lengthy to write down here, so I left as an exercise.
So $\forall [\sigma] \in H_n(X)$, we then have

$$(g_* - f_*)[\sigma] = (\partial \circ P)[\sigma] + (P \circ \underbrace{\partial)[\sigma]}_{=0} \in B_n(X)$$

therefore $f_* = g_*$ for all $f \approx g$, implies that $H_\bullet(X) \cong H_\bullet(Y)$. □

As we said, perhaps there is no direct intuition for $H_\bullet$, or an elementary way to count. It is for proving more complicated theories just like the theorem above. Singular homology brings two advanced techniques: *relative homology* and *Mayer-Vietoris sequence.*

It is generally difficult to calculate the homology on a quotient space X/A by counting; the *relative homology* provides a strong tool. It basically suggests that the homology of $\mathrm{C}_\bullet(X/A)$ is *almost* the same as the homology of $\mathrm{C}_\bullet(X)/\mathrm{C}_\bullet(A)$:

Definition 4.15. Given a subspace A of X, the **relative homology group** is defined as $H_n(X,A) := \ker(\partial_n)/\mathrm{Im}(\partial_{n+1})$ for

$$\cdots \xrightarrow{\partial_3} \mathrm{C}_2(X)/\mathrm{C}_2(A) \xrightarrow{\partial_2} \mathrm{C}_1(X)/\mathrm{C}_1(A) \xrightarrow{\partial_1} \mathrm{C}_0(X)/\mathrm{C}_0(A) \xrightarrow{\partial_0} 0$$

Check that the ∂ here is well-defined: for $c+\mathrm{C}_k(A) \in \mathrm{C}_k(X)/\mathrm{C}_k(A)$, we have $\partial\left(c+\mathrm{C}_k(A)\right) \subseteq \partial(c)+\mathrm{C}_{k-1}(A) \in \mathrm{C}_{k-1}(X)/\mathrm{C}_{k-1}(A)$.

The relative homology is easy to compute: Notice that we have a natural short exact sequence

$$0 \longrightarrow \mathrm{C}_k(A) \xrightarrow{\subset} \mathrm{C}_k(X) \longrightarrow \mathrm{C}_k(X)/C_k(A) \longrightarrow 0$$

which, by Snake lemma, gives a long exact sequence

$$\begin{tikzcd}
\cdots \arrow[r, "\partial"] & H_k(A) \arrow[r, "\iota_*"] & H_k(X) \arrow[r, "\tau_*"] & H_k(X,A) \arrow[dll, "\partial"] & \\
& H_{k-1}(A) \arrow[r, "\iota_*"] & H_{k-1}(X) \arrow[r, "\tau_*"] & H_{k-1}(X,A) \arrow[r, "\partial"] & \cdots
\end{tikzcd} \tag{$\star$}$$

Pick any $x_0 \in X$; we have that $H_n(\{x_0\}) = 0$ for $n > 0$ and $H_0(\{x_0\}) \cong \mathbb{Z}$. By ($\star$),

$$\begin{aligned} H_0(X) &\cong H_0(X,\{x_0\}) \oplus \mathbb{Z} \\ H_n(X) &\cong H_n(X,\{x_0\}) \qquad \text{for } n > 0 \end{aligned}$$

We call this $H_n(X,\{x_0\})$ to be the **reduced homology**, denoted to be $\widetilde{H}_n(X)$.

Now let's see how it can be applied to compute $H_\bullet(X/A)$.

The main conclusion that connects relative homology with topology is the Excision theorem, stating that removing some "small" part of A from A and X simultaneously will not change their relative homology:

Lemma 4.17. (Excision theorem) Given subspaces $Z \subset A \subset X$ such that $\overline{Z} \subset A \setminus \partial A$, we have isomorphism

$$H_\bullet(X \setminus Z, A \setminus Z) \cong H_\bullet(X,A)$$

induced by the natural inclusion $\iota : (X \setminus Z, A \setminus Z) \hookrightarrow (X,A)$

A detailed proof can be found in *Algebraic Topology* by Hatcher, Page 119 - 124; such proof involves a lot of lengthy geometrical reasoning, and doesn't interfere with our development, so it won't be presented here.

As mentioned, an important use of relative homology is computing the homology of some quotient spaces, in particular the quotient by *good* subspace.

Definition 4.7. A subspace $X \subseteq Y$ is a **strong deformation retract** of Y if there exists continuous $F : [0,1] \times Y \to Y$ such that

- $F|_{\{0\}\times Y} = \text{id.}$ and $F|_{\{1\}\times Y} \approx \text{id.}$
- $F|_{[0,1]\times X} = \text{id.}$

And we call a closed subspace $A \subset X$ a **good** subspace if A has an open neighbour strong deformation retract to A.

Any good subspace $A \subset X$ satisfies $\overline{A} \subset X \setminus \partial X$, but the converse is not necessarily true. (e.g. consider discrete topology) However, by ϵ-neighborhood (Lemma 2.12), the converse is also true when A and X are smooth manifolds (with or without boundary).
So good subspace is a slightly stronger condition for the Excision theorem to work, and it has a good property in terms of homology: For a neighbour U that strong deformation retracts to A, we have

$$H_n(X,A) \cong H_n(X,U) \cong H_n(X \setminus A, U \setminus A)$$

and if we consider the quotient map $\vartheta : X \to X/A$ acting on X, U and A, then let it act on these three homology groups, we have commutative diagram

$$\begin{tikzcd}
H_n(X,A) \arrow[r, "\cong"] \arrow[d, "\vartheta_*"] & H_n(X,U) \arrow[r, "\cong"] \arrow[d, "\vartheta_*"] & H_n(X \setminus A, U \setminus A) \arrow[d, "\vartheta_*", "="'] \\
H_n(X/A, A/A) \arrow[r, "\cong"] & H_n(X/A, U/A) \arrow[r, "\cong"] & H_n(X/A \setminus A/A, U/A \setminus A/A)
\end{tikzcd} \tag{4.12}$$

As the 3rd. vertical map is just the identity map, we have that the other two vertical maps are isomorphism. Therefore,

$$H_n(X,A) \cong H_n(X/A, A/A)$$

which eventually means

Proposition 4.8. Given any good subspace $A \subset X$, we have

$$H_n(X,A) \cong \widetilde{H}_n(X/A)$$

In other words, by $(\star)$, given a good subspace $A \subset X$ with natural inclusion $f : A \hookrightarrow X$ and quotient map $h : X \to X/A$, we have long exact sequence

$$\begin{tikzcd}
\cdots \arrow[r, "\partial"] & \widetilde{H}_k(A) \arrow[r, "f_*"] & \widetilde{H}_k(X) \arrow[r, "h_*"] & \widetilde{H}_k(X/A) \arrow[dll, "\partial"] & \\
& \widetilde{H}_{k-1}(A) \arrow[r, "f_*"] & \widetilde{H}_{k-1}(X) \arrow[r, "h_*"] & \widetilde{H}_{k-1}(X/A) \arrow[r, "\partial"] & \cdots
\end{tikzcd}$$

As subspaces of manifolds in which we're interested are mostly good subspaces, this makes finding homology of quotient spaces an easy task.

Example 4.18. For example, consider topological spaces $X_1, X_2, X_3, \ldots$ with points $x_j \in X_j$ such that $\{x_j\}$ are good subspaces,

$$\widetilde{H}_n\left(\bigvee_k X_k\right) \cong \bigoplus_k \widetilde{H}_n(X_k)$$

where $\bigvee_k X_k = \bigsqcup_k X_k / \{x_1, x_2, x_3, \ldots\}$.

Finally, we review the most useful conclusion in computing $H_\bullet$, the *Mayer-Vietoris sequence* (**M-V sequence** for short); it is to compute the $H_\bullet(U \cup V)$ using $H_\bullet(U)$ and $H_\bullet(V)$ —— just like the Van-Kampen theorem for the fundamental group; actually, you'll see that the M-V sequence can be derived from the Van-Kampen theorem; see in Question & Answer for more detail.

Given topological spaces U, V and $X = U \cup V$, we can construct a short exact sequence

$$0 \longrightarrow \mathrm{C}_k(U \cap V) \xrightarrow{p} \mathrm{C}_k(U) \oplus \mathrm{C}_k(V) \xrightarrow{q} \mathrm{C}_k(X) \longrightarrow 0$$

where $p(\sigma) = (\sigma, \sigma)$ and $q(\sigma, \tau) = \sigma - \tau$. Then by Snake lemma, this extends into a long exact sequence, the M-V sequence

Proposition 4.9. Given topological spaces U, V and $X = U \cup V$, there is a long exact sequence,

$$\cdots \xrightarrow{\partial_*} H_{k+1}(U \cap V) \xrightarrow{D_0} H_{k+1}(U) \oplus H_{k+1}(V) \xrightarrow{D_1} H_{k+1}(X) \xrightarrow{\partial_*} H_k(U \cap V) \xrightarrow{D_0} H_k(U) \oplus H_k(V) \xrightarrow{D_1} H_k(X) \xrightarrow{\partial_*} \cdots$$

where the maps are defined by

$$\begin{aligned} D_0 &: [\sigma] \mapsto ([\sigma], [\sigma]) \\ D_1 &: [\sigma, \tau] \mapsto [\sigma - \tau] \end{aligned}$$

However, the map ∂_* cannot be written down explicitly like D_0 and D_1. (Though there is an explanation on the construction of it in Wikipedia done by baycentric division.)

By M-V sequence, we can compute almost all the homology group of smooth manifolds with an explicit construction from atlas or known components; we will have a table listing the homology groups of some seen manifolds.

4.8 Singular cohomology

As the construction of cohomology is almost the same as the homology, some conclusions from homology work in cohomology by some small modifications; we will go through the M-V sequence and some

Though the

Finally, as an important lemma for the de-rham theorem in later sections, we have

> **Proposition 4.10.** The inclusion map $\iota : \mathrm{C}_k^\infty(M) \hookrightarrow \mathrm{C}_k(M)$ induces isomorphism $H_k(M) \cong H_k^\infty(M)$.

4.9 Questions

Those questions are sometimes used as a lemma in sections later on.

In questions below, we assume that R is a non-trivial ring.

1. We say that R has the Invariant Basis Number property (IBN) if $R^n \cong R^m$ if and only if $m = n$. Interestingly not all rings have IBN.

(i) Consider ring K and free K-module M generated by a countably infinite basis $\{x_j : j \in \mathbb{N}\}$; show that $R := \mathrm{End}_K(M)$ doesn't satisfy IBN.

(ii) Consider a two-sided ideal $I \subset R$, Show that if R/I satisfy IBN, then so do R.

(iii) Show that the converse is NOT true.

(iv) Suppose R is commutative and has a maximal ideal, then it has IBN.

2. Given an R-module E, its symmetric algebra is $S_+(E) := T(E)/\langle x \otimes y - x \otimes y\rangle$; and its anti-symmetric algebra is $S_-(E) := T(E)/\langle x \otimes y + x \otimes y\rangle$.

(i) Suppose $R = \mathbb{R}$, $E = R^n$, show that $S_+(E) \cong \mathbb{R}[X_1, \cdots, X_n]$.

(ii) Suppose $R = \mathbb{R}$, $E = R^n$, show that $S_-(E) \cong \bigwedge E$.

(iii) Suppose $R = \mathbb{C}[x_1, x_2]$, consider presentation $R^2 \to R^2 \to E \to 0$ where

$$\varphi = \begin{pmatrix} x_2 & x_1 - x_2 \\ 0 & -x_1 \end{pmatrix}$$

find $S_+(E)$ in terms of the quotient ring of a polynomial ring.

(iv) Suppose $R = \mathbb{Z}$, $E = R^2/\langle (3,0), (2,2)\rangle$, find $S_-(E)$.

3. Let X be an n-dim smooth compact manifold, ξ is a vector bundle over X. Prove that:

(i) There exists embedding of ξ into some trivial bundle ζ.

(ii) Show that $H^k_{\mathrm{dR}}(\xi) \cong H^k_{\mathrm{dR}}(X)$ for $k = 0, \cdots, n$.

4. Given a path connected topological space X, let $\pi_1(X)$ be the fundamental group (See Section 1.1) and denote $[a, b] := aba^{-1}b^{-1}$.

(i) Consider a map $h : \pi_1(X) \to H_1(X)$ be defined as $h([\gamma]) = \gamma_*(1)$ where $\gamma_* : H_1(\mathrm{Im}(\gamma)) \cong \mathbb{Z} \to H_1(X)$. Show that h is well-defined.

(ii) Show that $H_1(X) \cong \pi_1(X)/[\pi_1(X), \pi_1(X)]$.

5.

4.10 Answers

Question 1

(i) We can show that $R^2 \cong R$. Consider $\phi : R \to R^2$ by

$$([w_1, w_2, w_3, \cdots] \mapsto [z_1, z_2, z_3, \cdots]) \mapsto \left(\begin{bmatrix} w_1, w_3, w_5, \cdots \\ w_2, w_4, w_6, \cdots \end{bmatrix} \mapsto \begin{bmatrix} z_1, z_3, z_5, \cdots \\ z_2, z_4, z_6, \cdots \end{bmatrix} \right)$$

First, ϕ is clearly an injective homomorphism, which then implies that it is a surjection: For any $(f, g) \in R^2$, check that $\phi^{-1}(f, 0)$ and $\phi^{-1}(0, g)$ exists and unique, so $\phi^{-1}(f, g)$ exists. This gives us an isomorphism $R^2 \cong R$.

(ii) Suppose $R^m \cong R^n$ via $\phi : R^m \to R^n$. Now if we define $\phi_* : R^m/I^m \to R^n/I^n$ as

$$\phi_* (a_1 + I, \cdots, a_m + I) = \phi(a_1, \cdots, a_n) + I^n$$

then we can verify that ϕ_* is also an isomorphism.

Notice that $R^k/I^k = (R/I)^k$ for $\forall k \in \mathbb{Z}_{>0}$. Therefore $R^m \cong R^n$ implies

$$(R/I)^m = R^m/I^m \cong R^n/I^n = (R/I)^n$$

which means $m = n$ by IBN of R/I.

(iii) Key observation is that if R satisfy IBN, then so does $(R \times K)$ for any non-zero ring K. Notice that $(R \times K)^n$ is isomorphic to $R^n \times K^n$ via

$$((r_1, k_1), \cdots, (r_n, k_n)) \mapsto (r_1, \cdots, r_n) \times (k_1, \cdots, k_n)$$

Therefore $(R \times K)$ will satisfy IBN if R does.

So for non-IBN K, the $(R \times K)/(R \times \{0_K\}) \cong K$ cannot be IBN.

(iv) Let $\mathfrak{m} \subset R$ be a maximal ideal, then $F := R/\mathfrak{m}$ is a field; notice that there exists isomorphism $R^n/\mathfrak{m}^n \to F^n$, so if $R^m \cong R^n$, then $F^n \cong F_m$, which means $m = n$.

Question 2

(i) Consider map $\phi : S_+(\mathbb{R}^n) \to \mathbb{R}[X_1, \ldots, X_n]$ by sending each base vector $e_j \mapsto X_j$ and

$$\lambda(e_{j_1} \otimes \cdots \otimes e_{j_k}) \mapsto \lambda \cdot X_{j_1} \cdots X_{j_k}$$

Question 3

(i) Suppose ξ is of rank m.

For $\forall x \in X$, there exists a neighbour U_x of it with local trivialization, so we have an open cover $\{U_x\}$ of M; let $\{U_1, \ldots, U_r\}$ be a finite subcover of it, then assign a POU $\{\rho_1, \ldots, \rho_r\}$ subordinate to it.

Let $\phi_j : \pi^{-1}(U_j) \to U_j \times \mathbb{R}^m$ be the local trivialization on U_j; we extend it into a smooth map $\varphi_j : \pi^{-1}(U_j) \to (\mathbb{R}^m)^r$ by embedding the R^m component of ϕ_j into the jth component of $(\mathbb{R}^m)^r$. This will give $\varphi_j \cap \varphi_k = \emptyset$ for $j \neq k$.

So finally we are able to construct

$$\Phi : \xi \to X \times \mathbb{R}^m$$
$$(x,v) \mapsto \left(x, \sum_{j=1}^{r} \rho_j \cdot \varphi_j(x,v) \right)$$

verify that this is indeed an embedding.

Question 4

(i) Fix a point $x_0 \in X$ and denote γ_x as a path from x_0 to $x \in X$.
Notice that any singular 1-simplex σ is a path, so we can define $\phi(\sigma) = \gamma_{\sigma(0)} \cdot \sigma \cdot \gamma_{\sigma(1)}^{-1}$ and it induces a map $\phi_* : H_1(X) \to \pi_1(X)$ by sending $[\sigma]$ to $[\gamma_{\sigma(0)} \cdot \sigma \cdot \gamma_{\sigma(1)}^{-1}]$, we need to check that such map is well-defined: Suppose $\sigma_1, \sigma_2 \in [\sigma]$, which means there $\exists \tau \in \mathrm{C}_2(X)$ such that $[\gamma_{\sigma(0)} \cdot \sigma_1 \cdot \gamma_{\sigma(1)}^{-1}]$,

5 de-Rham Cohomology

Throughout this section, denote M a n-dim oriented smooth manifold. And we write

1. POU as Partition of Unity, definition see subsection 2.5.
2. $H^k(M)$ as **smooth** singular cohomology group.
3. Without specifying, cohomology groups are assumed to have coefficients in $\mathbb{R}$.

Recall that, intuitively the Stoke's theorem builds a "dual" relation between ∂ and d; to be more precise, consider $\psi : \Omega^k(M) \to \mathrm{C}^k(M;\mathbb{R})$ defined by

$$\psi(\omega)(c) := \int_c \omega$$

we have that by Stoke's theorem, $\psi(\mathrm{d}\omega)(c) = \psi(\omega)(\partial c)$, a perfect symmetry.
Further, we can check that ψ is a homomorphism from $\Omega^\bullet(M)$ to $\mathrm{C}^\bullet(M;\mathbb{R})$: Clearly ψ is additive, and "multiplicative" if we define $\psi([\omega]) \smile \psi([\omega]) := \psi([\omega] \wedge [\rho])$.
So ψ induces a group homomorphism on cohomology, the **de Rham homomorphism** $\Psi : H^k_{\mathrm{dR}}(M) \to H^k(M;\mathbb{R})$, which is given by

$$\Psi([\omega])(c) := \psi(\omega)(c)$$

We can easily check that Ψ is well-defined; and interestingly, it is an isomorphism.

de Rham Theorem

Theorem 5.1. The Ψ defines a ring isomorphism

$$(H^\bullet_{\mathrm{dR}}(M), +, \wedge) \cong (H^\bullet(M;\mathbb{R}), +, \smile)$$

This can be intuitive. As previously described, $\Omega^k(M)$ are somehow just "volumes" endowed with scalar and "orientation", just like the complexes; and integration on such "volumes" is the same as linearly mapping those complexes into $\mathbb{R}$.
Formally speaking, the $\{\Omega^k(M)\}$ equipped with the coboundary operator d forms a cochain complex, which will be called the **de Rham complexes** in the rest of this subsection. And the coboundary operator δ_k defined as $\delta_k(\Psi(\omega))(c) = \Psi(\omega)(\partial c)$ satisfies that $\Psi \circ \mathrm{d} = \delta \circ \Psi$:

$$\begin{tikzcd}
H^k_{\mathrm{dR}}(M) \arrow[r, "\Psi"] \arrow[d, "\mathrm{d}"] & H^k(M) \arrow[d, "\delta"] \\
H^{k+1}_{\mathrm{dR}}(M) \arrow[r, "\Psi"] & H^{k+1}(M)
\end{tikzcd}$$

Viewing $\Omega^k(M)$ this way, the theorem seems very likely and natural.

5.1 Calculation of $H^k_{\mathrm{dR}}(M)$

Before proving the **de Rham theorem** 5.1, let's first treat it as "*something that might be true*", and observe if those major theorems and computational tools (last section) in singular cohomology can be applied in H_{dR}.
As before, we start with the Mayer-Vietoris Sequence, which helps to decompose computations of $H^k_{\mathrm{dR}}(M)$ into those of its open cover.

Proposition 5.1. (Mayer-Vietoris Sequence) For open sets $U, V \subset M$ such that $M = U \cup V$, we have exact sequence

$$\begin{tikzcd} \cdots \arrow[r, "\delta"] & H^k_{\mathrm{dR}}(M) \arrow[r, "\alpha_*"] & H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V) \arrow[r, "\beta_*"] & H^k_{\mathrm{dR}}(U \cap V) \arrow[dll, "\vartheta"] \\ & H^{k+1}_{\mathrm{dR}}(M) \arrow[r, "\alpha_*"] & H^{k+1}_{\mathrm{dR}}(U) \oplus H^{k+1}_{\mathrm{dR}}(V) \arrow[r, "\beta_*"] & H^{k+1}_{\mathrm{dR}}(U \cap V) \arrow[r, "\delta"] & \cdots \end{tikzcd}$$

For simplicity, we will call it the **M-V sequence** later on.

The proof is similar to what we did in singular cohomology: Using Snake lemma by replacing the k-complexes with differential k-form.

Proof. Consider the inclusion relations

$$\begin{tikzcd} U \cap V \arrow[r, hook, "\iota_2"] \arrow[d, hook, "\iota_1"] & V \arrow[d, hook, "\tau_2"] \\ U \arrow[r, hook, "\tau_1"'] & M \end{tikzcd}$$

and the 4 de Rham complexes $\Omega^k(M), \Omega^k(U) \oplus \Omega^k(V)$ and $\Omega^k(U \cap V)$; to apply the Snake lemma, we need to construct the maps α_k and β_k for the exact sequence

$$0 \longrightarrow \Omega^k(M) \xrightarrow{\alpha_k} \Omega^k(U) \oplus \Omega^k(V) \xrightarrow{\beta_k} \Omega^k(U \cap V) \longrightarrow 0$$

where α_k, β_k are constructed by pullback of inclusion maps. Recall that to make a short sequence exact, all we need from this construction are: $\underline{\alpha_k \text{ injective}}$, $\underline{\beta_k \circ \alpha_k = 0}$ and $\underline{\beta_k \text{ surjective}}$. And luckily there exists a very natural one:

$$\alpha_k(\omega) := \left(\tau_1^*\omega|_U \, , \, \tau_2^*\omega|_V\right) \quad ; \quad \beta_k(\mu, \rho) := \iota_1^*\mu|_{U\cap V} - \iota_2^*\rho|_{U\cap V}$$

it's easy to see that this construction satisfies all three criteria. So the sequence is exact, which proves the theorem by Snake lemma. $\square$

Remark: Recall that the link map of M-V sequence for singular (co)homology doesn't admit an explicit expression, yet it does in the case for de-rham cohomology.

The α_* and β_* are simply the induced homomorphism as we saw in the proof:

$$\alpha_* : [\omega] \mapsto ([\tau_1^*\omega|_U], [\tau_2^*\omega|_V]) \quad ; \quad \beta_* : ([\omega], [\rho]) \mapsto [\iota_1^*\omega|_{U\cap V} - \iota_2^*\rho|_{U\cap V}]$$

While the link map ϑ is a bit complicated. Let's do the same thing as in (co)homology, tracing through the diagram;

$$\begin{tikzcd}
0 \arrow[r] & \Omega^k(M) \arrow[r, "\alpha_*"] \arrow[d, "\mathrm{d}"] & \Omega^k(U) \oplus \Omega^k(V) \arrow[r, "\beta_*"] \arrow[d, "\mathrm{d}"] & \Omega^k(U\cap V) \arrow[r] \arrow[d, "\mathrm{d}"] & 0 \\
0 \arrow[r] & \Omega^{k+1}(M) \arrow[r, "\alpha_*"] & \Omega^{k+1}(U) \oplus \Omega^{k+1}(V) \arrow[r, "\beta_*"] & \Omega^{k+1}(U\cap V) \arrow[r] & 0
\end{tikzcd} \tag{5.1}$$

Follow the proof of snake lemma: First start with an $\omega \in Z^k(U\cap V)$ and consider $\beta_*^{-1}\omega$. Be careful that ω might not be smooth on the entire U or V, so use a POU $\{\rho_U, \rho_V\}$ so that $\rho_U\omega$ is smooth on V and $\rho_V\omega$ is smooth on U. Hence

$$\beta_*^{-1}\omega = (\rho_V\omega, -\rho_U\omega)$$

Then consider $(\alpha_*^{-1} \circ \mathrm{d} \circ \beta_*^{-1})\omega$, which gives

$$\vartheta[\omega] := \left[(\alpha_*^{-1} \circ \mathrm{d} \circ \beta_*^{-1})\omega\right] \tag{5.2}$$

where

$$(\alpha_*^{-1} \circ \mathrm{d} \circ \beta_*^{-1})\omega = \begin{cases} \mathrm{d}\,(\rho_V\omega) & \text{when in } U \\ -\mathrm{d}\,(\rho_U\omega) & \text{when in } V \end{cases}$$

Though it still remains to verify that such ϑ is well-defined, independent of choice of POU, and exact in the sequence. (see Question & Answers)

Remark. All we need is just a homomorphism from $H^k_{\mathrm{dR}}(U\cap V)$ to $H^{k+1}(M)$, so why can't it simply be $\vartheta = \mathrm{d}$? Because $\omega \in \Omega^k(U\cap V)$ might not be continuous on entire M, whereas multiplying it with ρ_V on U and ρ_U on V ensures the smoothness. Besides, even if it is smooth, letting $\vartheta = \mathrm{d}$ will make it a zero map.

As we said, M-V sequence can be immediately used to compute $H^\bullet_{\mathrm{dR}}(M)$ for some common M. For example:

Example 5.2. For $\forall m \geq 1$, we have $H^k_{\mathrm{dR}}(\mathbb{S}^m) \cong \begin{cases} \mathbb{R} & \text{for } k = 0 \text{ or } m \\ \{0\} & \text{otherwise} \end{cases}$

Proof. Consider $\mathbb{S}^m := \{x : ||x|| = 1\}$ with $U = \mathbb{S}^m \setminus \{0, \cdots, 1\}$ and $V = \mathbb{S}^m \setminus \{0, \cdots, -1\}$. Notice that $U, V \cong \mathbb{R}^m$ and $U \cap V \approx \mathbb{S}^{m-1}$; so we can use induction:

Claim. For $\forall k > 1$, $H^k_{\mathrm{dR}}(\mathbb{S}^m) \cong H^{k-1}_{\mathrm{dR}}(\mathbb{S}^{m-1})$

By applying the result for $H^k_{\mathrm{dR}}(\mathbb{S}^2)$, our statement follows immediately after this claim.
Proof of the Claim: Since $U, V \cong \mathbb{R}^m$, $H^{k-1}_{\mathrm{dR}}(U), H^{k-1}_{\mathrm{dR}}(V) \cong \{0\}$ for $k > 1$; so the M-V sequence breaks into several short exact sequences like

$$0 \xrightarrow{\beta_{k-1}} H^{k-1}_{\mathrm{dR}}(U \cap V) \xrightarrow{\vartheta} H^k_{\mathrm{dR}}(M) \xrightarrow{\alpha_k} 0$$

By exactness, the ϑ is an isomorphism, proves the claim. □

Example 5.3. Consider $P_n = \mathbb{R}^m \setminus \{p_1, \cdots, p_n\}$, the Euclidean space with n distinct points removed, we have, for $m > 1$,

$$H^k_{\mathrm{dR}}(P_n) \cong \begin{cases} \mathbb{R} & \text{for } k = 0 \\ \mathbb{R}^n & \text{for } k = m-1 \\ 0 & \text{otherwise} \end{cases}$$

Proof. WLOG assume that $p_j = (j, 0, \cdots, 0)$ for $j = 1, 2, \cdots, n$.
Let $U = \left\{x \in P_n : x_1 < n - \frac{1}{3}\right\}$ and $V = \left\{x \in P_n : x_1 > n + \frac{1}{3}\right\}$, we have that $U \cap V \cong \mathbb{R}^m$, $U \cong P_{n-1}$ and $V \cong P_1$, which yield the first 6 terms in M-V sequence:

$$\mathbb{R} \xrightarrow{\alpha_0} \mathbb{R}^2 \xrightarrow{\beta_0} \mathbb{R} \xrightarrow{\vartheta} H^1_{\mathrm{dR}}(P_n) \xrightarrow{\alpha_1} H^1_{\mathrm{dR}}(P_{n-1}) \oplus H^1_{\mathrm{dR}}(P_1) \xrightarrow{\beta_1} 0$$

Since $\dim(\mathrm{Im}(\alpha_0)) \le 1$, we have that $\dim(\ker(\beta_0)) \le 1$, which means β_0 is surjective and thus $\vartheta = 0$. So $H^1_{\mathrm{dR}}(P_n) \cong H^1_{\mathrm{dR}}(P_{n-1}) \oplus H^1_{\mathrm{dR}}(P_1)$.
And for $k > 1$, we have short exact sequence

$$0 \xrightarrow{\vartheta} H^1_{\mathrm{dR}}(P_n) \xrightarrow{\alpha_{k+1}} H^1_{\mathrm{dR}}(P_{n-1}) \oplus H^1_{\mathrm{dR}}(P_1) \xrightarrow{\beta_{k+1}} 0$$

so $H^k_{\mathrm{dR}}(P_n) \cong H^k_{\mathrm{dR}}(P_{n-1}) \oplus H^k_{\mathrm{dR}}(P_1)$ is then true for all k. Applying the result of $H^k_{\mathrm{dR}}(P_1)$, our statement follows immediately by induction. □

Remark. We can take a step further, which is find the relation between $H^k_{\mathrm{dR}}(M \setminus \{p\})$ and $H^k_{\mathrm{dR}}(M)$ for arbitrary M and $p \in M$.
By considering $U = M \setminus \{p\}$ and $V = N_p \cong B_1(0)$, easy to get that for $n > 1$,

$$H^k_{\mathrm{dR}}(M) \cong H^k_{\mathrm{dR}}(M \setminus \{p\}) \text{ for } k \le n-2$$

however there are (precisely) two scenarios for $k = n-1$ and $k = n$:

$$\textbf{Case 1:} \begin{cases} H^{n-1}_{\mathrm{dR}}(M \setminus \{p\}) & \cong H^{n-1}_{\mathrm{dR}}(M) \oplus \mathbb{R} \\ H^n_{\mathrm{dR}}(M \setminus \{p\}) & \cong H^n_{\mathrm{dR}}(M) \end{cases}$$

$$\textbf{Case 2:} \begin{cases} H^{n-1}_{\mathrm{dR}}(M \setminus \{p\}) & \cong H^{n-1}_{\mathrm{dR}}(M) \\ H^n_{\mathrm{dR}}(M \setminus \{p\}) \oplus \mathbb{R} & \cong H^n_{\mathrm{dR}}(M) \end{cases}$$

But how to tell which case is it in practice?
It is difficult to tell without much information about $H^n_{\mathrm{dR}}(M)$; but luckily later (in Section 5.3) we will show that $\dim(H^n_{\mathrm{dR}}(M)) \in \{0,1\}$ and $\dim(H^n_{\mathrm{dR}}(M)) = 1$ iff M is compact.
Based on this, we have that

Example 5.4. (For M being connected & oriented) When $n > 1$,

$$H^k_{\mathrm{dR}}(M \setminus \{p\}) \cong \begin{cases} H^k_{\mathrm{dR}}(M) & \text{for } k < n-1 \\ H^k_{\mathrm{dR}}(M) & \text{for } k = n-1 \text{ and } M \text{ compact} \\ H^k_{\mathrm{dR}}(M) \oplus \mathbb{R} & \text{for } k = n-1 \text{ and } M \text{ non-compact} \\ 0 & \text{for } k = n \end{cases}$$

This immediately enables the computation of cohomology groups of the connected sum of manifolds (though it is complicated to write down an explicit formula). For example, the genus-m surface has $H^1_{\mathrm{dR}}(\mathbb{T}^2 \# \mathbb{T}^2 \# \cdots \# \mathbb{T}^2) \cong \mathbb{R}^{2m}$.

We also have an interesting way to deal with the product of manifolds.

Proposition 5.2. (Künneth formula) Given smooth manifolds X and Y, we have

$$H^n_{\mathrm{dR}}(X \times Y) \cong \bigoplus_{k=0}^{n} H^k_{\mathrm{dR}}(X) \otimes H^{n-k}_{\mathrm{dR}}(Y)$$

In other words, $H^\bullet_{\mathrm{dR}}(X \times Y) \cong H^\bullet_{\mathrm{dR}}(X) \otimes H^\bullet_{\mathrm{dR}}(Y)$ as graded rings. The proof of this uses almost the same thinking as that of the de-Rham theorem, we will leave it as an exercise after seeing through the proof in the next part.

5.2 de-Rham Theorem

[**Convention**] For simplicity, if a manifold satisfies the de-rham theorem, we say it is DR.

The key observation in the proof is that: If two submanifolds and their intersection are DR, then their union is DR.

Proposition 5.3. If open subsets $U, V \subset M$ and $U \cap V$ are DR, then $U \cup V$ is DR.

We can quickly extend this using induction:

Corollary 5.5. Suppose $\{U_1, \cdots U_r\}$ are all DR open subsets of M, and if the intersection of finitely many elements from it is always DR, then $\bigcup_{j=1}^{r} U_j$ is DR.

Before proving the proposition, notice that this can already prove the case of compact M; but what if M is not? This puts us into a very similar situation as in the proof of Whitney Embedding theorem (Corollary 2.4); so we can apply the same strategy:

Proposition 5.4. *(Upgraded Proposition 5.3)*
Suppose $\{U_1, U_2, U_3, \cdots\}$ are all DR open sets in M, and intersection of finitely many elements from it is always DR, then $\bigcup_{j=1}^{\infty} U_j$ is DR.

Proof. Set $W = \bigcup_{j=1}^{\infty} U_j$, it is a submanifold of M.
Similar to Proposition 2.9 in proof of Whitney Embedding theorem; Consider atlas with POU $\{(U_j, \varphi_j), \rho_j\}_{j\in\mathbb{N}}$ of W. Then pick a divergent monotone increasing sequence $\{a_j\}_{j\in\mathbb{N}}$ to construct an exhaustion function

$$F(p) := \sum_{j\in\mathbb{N}} a_j \rho_j(p)$$

It is smooth and satisfies that $F^{-1}[c, d] \subset W$ is compact for $\forall c < d$. So we can give each $F^{-1}[l, l+1]$ a finite open cover $\{N_{lj} : 1 \le j \le m_l\}$ and let

$$W_l = F^{-1}\left(l - \frac{1}{2}, l + \frac{3}{2}\right) \cap \bigcup_{j=1}^{m_l} N_{lj}$$

we have that $W_a \cap W_b = \emptyset$ if $|a - b| > 1$.
Now notice that the countable union of disjoint DR manifolds is DR (why?), which means $\bigcup_{j=1}^{\infty} W_{2l}$, $\bigcup_{j=1}^{\infty} W_{2j-1}$ and their intersection are all DR, so $W = \left(\bigcup_{j=1}^{\infty} W_{2l}\right) \cup \left(\bigcup_{j=1}^{\infty} W_{2l-1}\right)$ is DR. $\square$

Then the de rham theorem follows immediately —— Because $\mathbb{R}^n$ is DR and M always admits a countable good cover $\{U_j\}$, the M is DR by the proposition above.

So the only thing remains is to prove the Proposition 5.3.

Proof. of Proposition 5.3: List the M-V sequences for $H^\bullet_{\mathrm{dR}}(M)$ and $H^\bullet(M)$ in two columns respectively, we can have a diagram of homomorphism:

$$\begin{tikzcd}
H^{k-1}_{\mathrm{dR}}(U) \oplus H^{k-1}_{\mathrm{dR}}(V) \arrow[r, "\Psi \oplus \Psi"] \arrow[d] & H^{k-1}(U) \oplus H^{k-1}(V) \arrow[d] \\
H^{k-1}_{\mathrm{dR}}(U \cap V) \arrow[r, "\Psi"] \arrow[d] & H^{k-1}(U \cap V) \arrow[d] \\
H^{k}_{\mathrm{dR}}(U \cup V) \arrow[r, "\Psi"] \arrow[d] & H^{k}(U \cup V) \arrow[d] \\
H^{k}_{\mathrm{dR}}(U) \oplus H^{k}_{\mathrm{dR}}(V) \arrow[r, "\Psi \oplus \Psi"] \arrow[d] & H^{k}(U) \oplus H^{k}(V) \arrow[d] \\
H^{k}_{\mathrm{dR}}(U \cap V) \arrow[r, "\Psi"] & H^{k}(U \cap V)
\end{tikzcd}$$

If this diagram commutes, then we are done by applying the Five lemma (Corollary 4.2). To show that this diagram commutes, we only need to consider two kinds of scenarios:

Lemma 5.6. 1. If $f : X \to Y$ is a smooth map between two manifolds X, Y, then this diagram commutes:

$$\begin{tikzcd}
H^{k}_{\mathrm{dR}}(X) \arrow[r, "\Psi"] \arrow[d, "f^*"] & H^{k}(X) \arrow[d, "f^*"] \\
H^{k}_{\mathrm{dR}}(Y) \arrow[r, "\Psi"] & H^{k}(Y)
\end{tikzcd}$$

2. If U, V are open subsets of M, then this diagram commutes:

$$\begin{tikzcd}
H^{k}_{\mathrm{dR}}(U \cap V) \arrow[r, "\Psi"] \arrow[d, "\vartheta"] & H^{k}(U \cap V) \arrow[d, "\theta_k"] \\
H^{k+1}_{\mathrm{dR}}(U \cup V) \arrow[r, "\Psi"] & H^{k+1}(U \cup V)
\end{tikzcd}$$

where ϑ is the link map of the M-V sequence for $H^\bullet_{\mathrm{dR}}$, and θ_k is that of $H^\bullet$.

- For the first case: Consider $[\omega] \in H^k_{\mathrm{dR}}(X)$, for any $[\sigma] \in H_k(Y)$,

$$(\Psi \circ f^*[\omega])([\sigma]) = \int_\sigma f^*\omega = \int_{f(\sigma)} \omega = \Psi([\omega])([f(\sigma)]) = (f^* \circ \Psi[\omega])([\sigma])$$

- For the second case: Recall that the linking map θ in M-V sequence of singular cohomology can be derived from that of homology: $\varsigma : H_k(M) \to H_{k-1}(U \cap V)$ defined by

$$\varsigma([\sigma]) = [\partial \sigma_U] = [-\partial \sigma_V]$$

for $\sigma = \sigma_U + \sigma_V$, where $\mathrm{Im}(\sigma_U) \subset U$ and $\mathrm{Im}(\sigma_U) \subset U$; then let $\theta([u]) := [u] \circ \varsigma$. Hence

$$(\Psi \circ \vartheta[\omega])([\sigma]) = \int_{\sigma_U} \mathrm{d}(\rho_V \omega) - \int_{\sigma_V} \mathrm{d}(\rho_U \omega)$$

$$
\begin{aligned}
&= \int_{\partial\sigma_U} \rho_V \omega + \int_{-\partial\sigma_V} \rho_U \omega \\
&= (\Psi[\omega]) \circ \varsigma[\sigma] \\
&= (\theta \circ \Psi[\omega])([\sigma])
\end{aligned}
$$

□

By this lemma, our diagram commutes; so the Proposition 5.3 immediately follows by Five lemma, and then the de-rham theorem. QED

5.3 $H^n_{\mathrm{dR}}(M)$ tells the compactness

We know that $H^0_{\mathrm{dR}}(M)$ tells the number of connected components, but what about $H^n_{\mathrm{dR}}(M)$? The answer turns out to be **compactness**.

Proposition 5.5. For connected M, we have

$$H^n_{\mathrm{dR}}(M) \cong \begin{cases} \mathbb{R} & \text{if } M \text{ compact} \\ \{0\} & \text{otherwise} \end{cases}$$

but it turns out that such a conclusion is hard to prove directly. To solve this problem, we develop the **compactly supported** de rham cohomology group $H^\bullet_c(M)$.

Definition 5.1. We define the compactly supported k-form on M as $\Omega^k_c(M) := \{\omega \in \Omega^k(M) : \mathrm{supp}(\omega) \text{ compact.}\}$, and let

$$\begin{aligned} Z^k_c(M) &:= \Omega^k_c(M) \cap \mathrm{d}_k^{-1}\{0\} \\ B^k_c(M) &:= \Omega^k_c(M) \cap \mathrm{d}_{k-1}(\Omega^{k-1}_c(M)) \end{aligned}$$

The compactly supported de-rham cohomology group is $H^k_c(M) := Z^k_c(M)/B^k_c(M)$.

Verify that $(H^\bullet_c(M), +, \wedge)$ forms a ring, and $\underline{H^k_c(M) = H^k_{\mathrm{dR}}(M) \text{ for compact } M}$.
Most conclusions for $H^\bullet_{\mathrm{dR}}$ hold similarly for $H^\bullet_c$, BUT there are some key differences.

(1) Homotopy invariant: Given two smooth manifolds X, Y, smooth map $f : X \to Y$ doesn't always induce homomorphism $f_* : H^\bullet_c(X) \to H^\bullet_c(Y)$, because for $\eta \in \Omega^k_c(Y)$, it's possible that $f^*\eta \notin \Omega^k_c(X)$. So $H^\bullet_c$ is NOT always homotopy invariant!
But this can be immediately fixed by letting f be *proper*, and thus $H^\bullet_c$ is invariant under *proper Homotopy equivalence*[9]. (i.e $\exists$ proper $f \in C^\infty(X, Y), g \in C^\infty(Y, X)$ such that $f \circ g \approx \mathrm{id}_Y$ and $g \circ f \approx \mathrm{id}_X$.)

(2) $H^0_c(M)$: Notice that $H^0_c(M)$ no longer counts #(connected components):

$$\dim(H^0_c(M)) = \#\,(\text{compact connected components})$$

(3) Poincaré Lemma: Interestingly, the H^k_c is "opposite" to H^k_{dR} for $\mathbb{R}^m$:

$$H^k_c(\mathbb{R}^m) \cong \begin{cases} \mathbb{R} & \text{for } k = m \\ 0 & \text{otherwise} \end{cases} \tag{5.3}$$

Also, it's proved differently. (We will see in a minute)

[9]The proof works the same as for H_{dR}, but will need to prove the Whitney approximation theorem for proper smooth maps, which is a good exercise.

(4) M-V sequence: The M-V sequence for H_{dR} won't work for H_c, as the pullback of $\Omega_c^k(M)$ by inclusion $U \hookrightarrow M$ can be non-compactly supported!
However, notice that any $\eta \in \Omega_c^k(U)$ is naturally in $\Omega_c^k(M)$ by setting $\eta|_{M\setminus U} = 0$! (but this is NOT true for Ω_k) Hence, there is a "reversed" M-V sequence:

$$\begin{tikzcd}
\cdots \arrow[r, "\vartheta"] & H_c^k(U\cap V) \arrow[r, "\alpha"] & H_c^k(U)\oplus H_c^k(V) \arrow[r, "\beta"] & H_c^k(M) \arrow[dll, "\vartheta"] \\
& H_c^{k+1}(U\cap V) \arrow[r, "\alpha"] & H_c^{k+1}(U)\oplus H_c^{k+1}(V) \arrow[r, "\beta"] & H_c^{k+1}(M) \arrow[r, "\vartheta"] & \cdots
\end{tikzcd}$$

where $\alpha([\omega]) = ([\omega], -[\omega])$ and $\beta([\xi],[\eta]) = [\xi+\eta]$; while for the link homomorphism ϑ, consider a POU $\{\rho_U, \rho_V\}$, and let $\vartheta([\omega]) = [\mathrm{d}(\rho_U\omega)] = [-\mathrm{d}(\rho_V\omega)]$.

This idea which we used in (4) can be used to prove (3):

Proof. Pick $p \in \mathbb{S}^m$; since $\mathbb{S}^m$ is compact (which means $H_c^k(\mathbb{S}^m) \cong H_{\mathrm{dR}}^k(\mathbb{S}^m)$) and $P := \mathbb{S}^m \setminus \{p\} \cong \mathbb{R}^m$, so we try finding $H_c^k(P)$.
First, by (2) we prove the case $k = 0$; now let's try $0 < k < m$.
Pick $\omega \in Z_c^k(P)$, which is naturally in $Z_c^k(\mathbb{S}^m)$. We know that $Z_c^k(\mathbb{S}^m) = Z^k(\mathbb{S}^m) = B^k(\mathbb{S}^m) = B_c^k(\mathbb{S}^m)$, so there $\exists \eta \in \Omega_c^k(\mathbb{S}^m)$ such that $\mathrm{d}\eta = \omega$. We wish to modify η to let it satisfy $p \notin \operatorname{supp}(\eta)$. To achieve this, construct a bump function g such that $\{x \in \mathbb{S}^m : g(x) = 1\}$ contains a neighbour around p, and $\nu \in \Omega^{k-1}(\mathbb{S}^m)$ such that $\mathrm{d}\nu = \eta|_p$. Let

$$\eta' = \eta - \mathrm{d}(g\cdot\nu)$$

Verify that $\eta' \in \Omega_c^k(P)$ and $\mathrm{d}(\eta') = \omega$. Hence $\omega \in B_c^k(P)$.
Finally, for $H_c^m(P)$; recall that the isomorphisms $H_{\mathrm{dR}}^1(\mathbb{S}^1), H_{\mathrm{dR}}^2(\mathbb{S}^2) \cong \mathbb{R}$ were actually given by integrals on $\mathbb{S}^1$ and $\mathbb{S}^2$, so we consider

$$\int_P : H_c^m(P) \to \mathbb{R}$$
$$[\omega] \mapsto \int_P [\omega]$$

verify that it's a well-defined homomorphism: 1, $\operatorname{supp}(\omega)$ is always compact; 2, Since $\partial P = \emptyset$, we have that $\int_P B_c^m(P) \subset \int_{\partial P} \Omega_c^{m-1}(P) = \{0\}$ by Stokes' theorem.
The $\int_P$ is clearly surjective, and only need to show that it's injective. Suppose $\eta \in Z_c^n(P)$ satisfies that $\int_P \eta = 0$. Notice that by $H_c^m(\mathbb{S}^m) \cong \mathbb{R}$, we have that $\int_{\mathbb{S}^m}$ is an injection. Similarly, because η is naturally in $Z_c^m(\mathbb{S}^m)$, we have that $\int_{\mathbb{S}^m} \eta = 0 \implies \eta \in B_c^m(\mathbb{S}^m)$, which means $\eta \in B_c^m(P)$ as $\operatorname{supp}(\eta) \subset P$. Hence $\int_P$ is injective, $H_c^m(P) \cong \mathbb{R}$. □

Note that isomorphism $H_c^n(\mathbb{R}^n) \cong \mathbb{R}$ is induced by $\int_{\mathbb{R}^m}$, so is it true generally?
Indeed, it is the main theory in this topic:

Proposition 5.6. For connected M, map $\int_M : H^n_c(M) \to \mathbb{R}$ is isomorphism.

Proof. Still, $\int_M$ is clearly surjective, thus we only need to show that

$$\int_M \omega = 0 \implies \omega \in B^n_c(M)$$

So let's consider $\omega \in \Omega^n_c(M)$ that satisfies $\int_M \omega = 0$. Since $\operatorname{supp}(\omega)$ is compact and M is connected, the $\operatorname{supp}(\omega)$ must has a connected finite good cover $\{U_1, \cdots, U_r\}$, where $U_j \cong \mathbb{R}^n$. Since our statement is correct for $r = 1$ by (3), let's do induction:
Suppose the statement is true for $r < m - 1$, now let's prove for $r = m$.
We argue that for any given $\{U_1, \cdots, U_m\}$, there exists U_k such that $\{U_1, \cdots, \widehat{U_k}, \cdots, U_m\}$ is connected (Why?). Now let $U = U_k$ and $V = U_1 \cup \cdots \cup \widehat{U_k} \cup \cdots \cup U_m$, and consider a POU $\{\rho_U, \rho_V\}$ subordinate to $\{U, V\}$.
As our statement holds inside U, V and $U \cap V$; we can try "moving" the support of $[\omega] = [\rho_U \omega] + [\rho_V \omega]$ into one of them. Construct an $\eta \in \Omega^n_c(U \cap V)$ such that

$$\int_U \eta = \int_U \rho_U \omega$$

As $U \cong \mathbb{R}^n$, this gives that $[\eta] = [\rho_U \omega]$, and thus $[\eta + \rho_V \omega] = [\omega]$. Therefore

$$\int_V (\eta + \rho_V \omega) = \int_M \omega = 0$$

we have that $(\eta + \rho_V \omega) \in B^n_c(V)$ by inductive hypothesis, which gives $[\omega] = [\eta + \rho_V \omega] = [0]$; hence, the statement also holds for $r = m$. □

This immediately proves the first part of our conclusion at the beginning, as when M is compact, $H^n_{\mathrm{dR}}(M) \cong H^n_c(M) \cong \mathbb{R}$. So let's prove the non-compact case:

Proof of Proposition 5.5: As the compact case is solved, let's consider a non-compact M. First, consider a countable open cover $\{U_j\}$ with POU $\{f_j\}$, then construct an exhaustion function (used in proofs of Proposition 2.9 and de-Rham theorem),

$$F := \sum_{j \in \mathbb{N}} a_j f_j$$

(where $\{a_j\} \subset \mathbb{R}$ is a divergent monotone increasing sequence) Then construct an open cover $\{V_j\}_{j \in \mathbb{N}}$ with POU $\{\rho_j\}$, where $V_j = F^{-1}(j-1, j+1)$; the $\overline{V_j}$ is always compact because F is smooth & proper.
Now the trick is like playing *pass the parcel*, throwing non-exact parts of ω infinitely far away: Partition $\omega = \sum_{j \in \mathbb{N}} \rho_j \omega$, then construct $\eta_j \in \Omega^n_c(F^{-1}(j, j+1))$ such that $\int_M \eta_j = 1$.

Notice that if we let $c_j = c_{j-1} + \int_M \rho_j \omega$ with $c_{-1} = 0$, then for $\forall j > 0$,

$$\int_M (\rho_j \omega - c_j \eta_j + c_{j-1}\eta_{j-1}) = 0$$

which means by Prop 5.6, there $\exists \xi_j \in \Omega_c^{n-1}(V_j)$ such that $\mathrm{d}\xi_j = (\rho_j\omega - c_j\eta_j + c_{j-1}\eta_{j-1})$. Hence

$$\begin{aligned}\mathrm{d}\left(\sum_{j=0}^{\infty} \xi_j\right) &= (\rho_0\omega - c_0\eta_0) + \sum_{j=1}^{\infty}(\rho_j\omega - c_j\eta_j + c_{j-1}\eta_{j-1}) \\ &= \sum_{j=0}^{\infty} \rho_j \omega = \omega\end{aligned}$$

Check that every point in M is covered by no more than two of $\{V_j\}$, which means $\sum_{j=0}^{\infty} \xi_j \in \Omega^{n-1}(M)$, so $\omega \in B^n(M)$. $\square$

Now we have being consider the case in which M is a smooth *oriented* manifold, but what if M is not oriented? The conclusion is intuitively that $H^n_{\mathrm{dR}}(M) = \{0\}$.

Corollary 5.7. For smooth connected n-dim manifold X,

$$H^n_{\mathrm{dR}}(X) \cong \begin{cases} \mathbb{R} & \text{for } X \text{ being both compact } \textit{and} \text{ orientable} \\ \{0\} & \text{otherwise} \end{cases}$$

But to prove this, we would need some treatment on orientablity, which would be discussed in next part.

5.4 Something interesting about $(M, M \setminus K)$

It's related to two topics: $H_c^\bullet(M)$ & Fundamental Class.
Besides the 4 remarks we made about the differences between H_c and H_{dR}, there is one last issue: What is $H_c^\bullet(M; R)$ for other R? Certainly it is easy to do analogy: Define the complex $\mathrm{C}_c^k(M, R)$ to be the set of cochains with compact support, and then

$$\begin{aligned} Z_c^k(M; R) &:= \mathrm{C}_c^k(M; R) \cap Z^k(M; R) \\ B_c^k(M; R) &:= \mathrm{C}_c^k(M; R) \cap \delta_{k-1}\left(\mathrm{C}_c^{k-1}(M; R)\right) \\ H_c^k(M; R) &:= Z_c^k(M; R)/B_c^k(M; R) \end{aligned}$$

Be aware that it's NOT checked that this really coincides with the definition of $H_c(M)$ when $R = \mathbb{R}$, but they do coincide, and we'll show this. (i.e. $H_c(M; \mathbb{R}) \cong H_c(M)$)
Let's write down the formal definition of $\mathrm{C}_c^k(M, R)$:

$$\mathrm{C}_c^k(M, R) = \left\{\eta \in \mathrm{C}^k(M; R) : \exists \text{ compact } K \subset M \text{ such that } \omega \circ \mathrm{C}_k(M \setminus K; R) = \{0\}\right\}$$

So we have

$$\mathrm{C}_c^k(M, R) = \bigcup_{K \subseteq M \text{ compact}} \mathrm{C}^k(M, M \setminus K, R)$$

Therefore, consider a sequence $K_1 \subseteq K_2 \subseteq K_3 \subseteq \cdots$ of compact subsets of M such that $\bigcup_j K_j = M$, and let $h_{ij} : K_j \to K_i$ be the natural inclusions for each $i > j$. Each h_{ij} induces quotient map $\mathrm{C}^k(M, M \setminus K_i, R) \to \mathrm{C}^k(M, M \setminus K_j, R)$, which gives a homomorphism $f_{ij} : H^k(M, M \setminus K_i, R) \to H^k(M, M \setminus K_j, R)$. Hence,

Proposition 5.7. $H_c^k(M; R) = \varinjlim H^k(M, M \setminus K_\bullet; R)$

Proof. We claim that $\varinjlim H^k(M, M \setminus K_\bullet; R) \subseteq H_c^k(M; R)$: For any $\omega \in \mathrm{C}^k(M, M \setminus K_j; R)$, we have that $\operatorname{supp}\omega \subseteq K_j$ compact, so every element in $H^k(M, M \setminus K_j; R)$ would be a class of compactly supported forms.
We also have $H_c^k(M; R) \subseteq \varinjlim H^k(M, M \setminus K_\bullet; R)$, since for any $\omega \in \mathrm{C}_c^k(M)$, there exists $r \in \mathbb{N}$ such that $\operatorname{supp}(\omega) \subset K_r$. □

Take a computational example of $M = \mathbb{R}^n$, we can use $K_j = \overline{\mathbb{B}_j^n(0)} \subset M$. For any j, we know that $H^k(M, M \setminus K_j; R) \cong 0$ for $k < n$, and R for $k = n$; therefore, since $f_{ij} : H^k(M, M \setminus K_i; R) \to H^k(M, M \setminus K_j; R)$ is an embedding for any $i \geq j$, we have

$$H_c^k(\mathbb{R}^n; R) \cong \begin{cases} R & \text{for } k = n \\ 0 & \text{otherwise} \end{cases}$$

5.5 Poincaré Duality

Consider $X \subset M$ a smooth connected oriented submanifold.

By comparing $H_c^k(\mathbb{R}^n)$ and M-V sequence for $H_c^\bullet$ with that of $H_{\mathrm{dR}}^\bullet$, it's natural to conjecture that $H_c^k(M) \cong H_{\mathrm{dR}}^{n-k}(M)$; though it turns out that this is only true when $\dim(H_{\mathrm{dR}}^\bullet(M)) < \infty$, it is quite close:

Proposition 5.8. (Poincaré duality) We have isomorphism

$$H_{\mathrm{dR}}^k(M) \cong \left(H_c^{n-k}(M)\right)^*$$

via map $\mathcal{P}$ defined as

$$\mathcal{P} : [\omega] \mapsto \left([\eta] \mapsto \int_M \eta \wedge \omega\right)$$

As a corollary, $H_{\mathrm{dR}}^k(M) \cong H_c^{n-k}(M)$ when their dimension $< \infty$, but otherwise this never holds, because for any $\mathbb{R}$-vector space V with $\dim(V) = \infty$, cardinality of $\dim(V^*)$ is always greater than that of $\dim(V)$. This also means

$$\left(H_{\mathrm{dR}}^k(M)\right)^* \cong H_c^{n-k}(M) \iff H_{\mathrm{dR}}^k(M) < \infty$$

The proof of Poincaré duality resembles that of the de-rham theorem and Künneth formula. A detailed proof is left to Question & Answer.

By de-Rham theorem and universal coefficient theorem, our Poincaré duality is equivalent to:

$$(H_c^k(M;\mathbb{R}))^* \cong (H_{n-k}(M;\mathbb{R}))^*$$

Based on their similarity in M-V sequences and basic examples like $\mathbb{R}^m$ and $\mathbb{S}^m$ we found previously, one would expect that a much stronger conclusion:

Example 5.8. (general Poincaré Duality) For any commutative ring R, we have that, for $k = 0, 1, 2, \ldots,$

$$H_c^k(M;R) \cong H_{n-k}(M;R)$$

where the isomorphism is given by sending $\omega \mapsto [M] \frown \omega$.

I'll skip the proof, since this version of Poincaré dual isn't our focus here after all. If curious, there are a lot of sources for this classical theorem (e.g There is a proof in Hatcher's *Algebraic topology* P245-249)

This general version provides a valuable insight: It suggests that one could represent a compact submanifold in differential forms. A k-dim compact submanifold $X \subset M$ can always represents a homology class in $H_k(X;R)$, while integration on X always defines a linear functional $\Omega^k(M)$, with $B^k(M) \mapsto \{0\}$. In other words, X gives an element of $(H^k_{\mathrm{dR}}(M))^*$, which, by our Poincaré duality, corresponds to a unique element of $H_c^{n-k}(X)$:

Corollary 5.1. For any smooth k-dim oriented compact submanifold $X \subset M$, there exists unique $[\eta] \in H_c^{n-k}(M)$ such that

$$\int_M \eta \wedge \omega = \int_X \iota^* \omega \quad \text{for } \forall [\omega] \in H^k_{\mathrm{dR}}(M)$$

where $\iota : X \hookrightarrow M$ is the natural inclusion.

For convenience, we call $[\eta]$ in this corollary the **dual forms** of X.
Let's demonstrate with an example: Consider unit circle $X = \mathbb{S}^1 \subset M = \mathbb{R}^2 \setminus \{0\}$. Let $\rho \in C^\infty(M)$ be a bump function such that $\int_{\mathbb{R}^2} \rho = 1$, and

$$\eta|_{(x,y)} = \rho \cdot (x \ \mathrm{d}x + y \ \mathrm{d}y) \in \Omega^1_c(M)$$

We know that $H^1_{\mathrm{dR}}(M) \cong \mathbb{R}$, which is the $\mathbb{R}$-span of $[\omega]$, where

$$\omega|_{(x,y)} = \frac{-y \ \mathrm{d}x + x \ \mathrm{d}y}{x^2 + y^2}$$

then we have that

$$\int_M \eta \wedge \omega = \int_M \rho \ \mathrm{d}x \wedge \mathrm{d}y = 1 = \int_X \iota^* \omega$$

as desired.
Though we could do more examples, this one is sufficient to give a general intuition on how η is like: The $[\eta]$ contains the $(n-k)$-forms "spanning" the normal subspace of each point on X in M. So we wonder if this could be done literally —— construct an η that lies inside a tubular neighbourhood of $X \subset M$.

Lemma 5.9. For any tubular neighbourhood $N \subset M$ of X, there exists a dual form $\eta \in Z_c^{n-k}(M)$ such that $\mathrm{supp}(\eta) \subset N$.

Proof. Fix any $[\omega] \in H^k_{\mathrm{dR}}(M)$ and a dual form $\theta \in Z_c^{n-k}(M)$ of X. Since M is oriented, take a volume form ω_g of it and suppose $\theta \wedge \omega = \widetilde{f} \cdot \omega_g$ for some $\rho \in C_c^\infty(M)$. Then there exists $f \in C_c^\infty(M)$ such that $\mathrm{supp}(\varrho) \subset N$ and

$$\int_M f \cdot \omega_g = \int_M \widetilde{f} \cdot \omega_g$$

and so it satisfies $(\widetilde{f} - f)\omega_g \in B_c^k(M)$.
Notice that there exists $\overline{\eta} \in \Omega_c^{n-k}(M)$ such that $(\overline{\eta} \wedge \omega)|_N = \omega_g|_N$ (because it exists locally, and we can use POU to extend it around N). Then by letting $\eta = f \cdot \overline{\eta}$, we get the desired η. (Verify that $\eta \in Z_c^{n-k}(M)$ and η is indeed a dual form of X.) $\square$

This restriction can be further improved: Since the tubular neighbourhood is diffeomorphic to the normal bundle, we can consider η inside of it:

Lemma 5.10. Let N be any tubular neighbourhood of X and $\phi : \mathrm{N}(X, M) \to N$ an orientation-preserving diffeomorphism. Fix a local chart $U \subset X$ and a local trivialization, then we can pick a dual form η of X such that

$$\phi^*\eta\big|_x = f(x)\ \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_{n-k} \qquad \text{for } \forall x \in U$$

for some $f \in C^\infty(\mathrm{N}(X, M))$, where $\{\mathrm{d}u_1, \ldots, \mathrm{d}u_{n-k}\}$ is a local frame for the chosen local trivialization.

Proof. Giving a local chart and a local trivialization, denote local coordinate basis of TX as $\{\mathrm{d}x_1, \ldots, \mathrm{d}x_k\}$ and local frame as $\{\mathrm{d}u_1, \ldots, \mathrm{d}u_{n-k}\}$; together they form a local basis for $T^*\mathrm{N}(X, M)$.

By previous lemma, we can construct the dual form $\widetilde{\eta}$ that $\mathrm{supp}(\widetilde{\eta}) \subset N$; now suppose we decompose $\widetilde{\eta}$ such that in local frame,

$$\phi^*\widetilde{\eta} = \underbrace{f \cdot (\mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_{n-k})}_{\eta} + \alpha$$

where α are forms with $\mathrm{d}x_j$ components. Similarly, we can also decompose any $\omega \in \Omega^k(M)$ such that $\omega|_{\pi^{-1}(U)} = \phi^*\left(h \cdot (\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k) + \beta\right)$ where β are forms with $\mathrm{d}u_j$ components. Now since ϕ is orientation-preserving, we have

$$\begin{aligned}\int_M \widetilde{\eta} \wedge \omega &= \int_{\mathrm{N}(X,M)} (f \cdot h)\ \mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}x_k + \int_{\mathrm{N}(X,M)} \alpha \wedge \beta \\ &= \int_X h\ \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_k\end{aligned}$$

By fixing $\widetilde{\eta}$ and h, we have that the terms $\int_M \alpha \wedge \beta$ must be constantly 0. Therefore by Corollary 5.1, we have η is in the same cohomology class as $\widetilde{\eta}$. □

In our proof, we could actually observe that, if we pick the dual form like this, then integral of η along the fibre at each point is equal to 1:

Corollary 5.2. The dual form η chosen in Lemma 5.10 satisfies

$$\int_{\pi^{-1}(x)} \phi^*\eta = 1 \qquad \text{for all } x \in X \tag{5.4}$$

And this provides almost complete description on what η is like.

The intuition that "a dual form span the normal subspace" is suggesting an interesting application in *intersection theory*.

Definition 5.1. Given smooth oriented submanifolds $X, Y \subset M$ with $\dim(X)+\dim(Y) = n$, let ω_X and ω_Y be the orientation forms on X and Y respectively. Suppose $|X \cap Y| < \infty$, then for each $p \in X \cap Y$, define

$$\varepsilon_p(X, Y) := \begin{cases} 1 & \text{for } \left.(\omega_X \wedge \omega_T)\right|_p(\mathbf{v}) > 0 \\ -1 & \text{for } \left.(\omega_X \wedge \omega_T)\right|_p(\mathbf{v}) < 0 \\ 0 & \text{otherwise} \end{cases}$$

for some $\mathbf{v} \in (T_pM \setminus \{0\})^n$. Then define the **intersection number** of X and Y to be

$$\mathcal{I}(X, Y) := \sum_{p \in X \cap Y} \varepsilon_p(X, Y)$$

This $\mathcal{I}(X, Y)$ does not count the number of intersection but the "*minimal number*" of intersections: Intuitively, suppose $\dim(M) = 2$ and X, Y are smooth curves with finitely many intersection, then by moving X and Y inside M you could somehow reduce the number of intersections, but you can never reduce that number below $|\mathcal{I}(X, Y)|$.
Based on our definition and the interpretation of Poincaré duality, we could draw down an intuitive conclusion on $\mathcal{I}(X, Y)$:

Proposition 5.9. Consider compact oriented connected submanifolds $X, Y \subset M$ such that $\dim(X) + \dim(Y) = n$. Suppose $X \pitchfork Y$ and $|X \cap Y| < \infty$, let η_X be a dual form of X, and η_Y be a dual form of Y; then we have

$$\mathcal{I}(X, Y) = \int_M \eta_X \wedge \eta_Y$$

The trick is shrinking the support of dual forms as much as desirable.

Proof. Let $X \cap Y = \{p_1, \ldots, p_m\}$, it is left as a topology exercise to show that there exists:

- A tubular neighbourhood N of Y in M such that

$$N \cap X = \bigsqcup_{j=1}^{m} \{p_j\} \times \mathbb{R}^{n-k}$$

- A diffeomorphism $\phi : \mathrm{N}(X, M) \to N$ such that each of $\phi \circ \pi^{-1}(\{p_j\})$ is a connected component of $N \cap X$.

Now take N and ϕ as described above.
Notice that we can assign the orientation of N to be a choice of orientation of M such that the natural embedding $Y \hookrightarrow M$ is orientation-preserving.

Pick the dual form η_Y as in Lemma 5.10, giving that

$$\int_M \eta_X \wedge \eta_Y = \sum_{j=1}^{m} \int_{\pi^{-1}(\{p_j\})} \tau_j^* \circ \eta_Y \tag{5.5}$$

where $\tau_j : \pi^{-1}(\{p_j\}) \subset N \cap X \to X$ are embeddings given by $\tau_j = \phi|_{\pi^{-1}(\{p_j\})}$. By the orientation we assigned to N, the τ_j is orientation-preserving for $\varepsilon_{p_j}(X, Y) = 1$ and reversing for $\varepsilon_{p_j}(X, Y) = -1$. Hence, by Corollary 5.2, we have

$$\int_{\pi^{-1}(\{p_j\})} \tau_j^* \circ \eta_Y = \varepsilon_{p_j}(X, Y)$$

Plugging this back into (5.5) finishes the proof. □

5.6 Application: deg & χ

Now we haven't yet got a significant application for the knowledge of this section; but there are two interesting topological invariants related: deg and χ.

In this section, let X, Y, Z be connected smooth manifolds with dimension n.
Recall that because $H_c^n(X), H_c^n(Y) \cong \mathbb{R}$, the homomorphism between them induced by any proper smooth map $f : X \to Y$ should be an endomorphism of $\mathbb{R}$. So there exists a constant $\deg(f)$ such that

$$\int_X f^*\omega = \deg(f) \int_Y \omega$$

for $\forall \omega \in \Omega_c^n(Y)$. This $\deg(f)$ is called the **Mapping degree** of f.

Example 5.11. Let's consider some simple cases:

- If f is a diffeomorphism, then

$$\deg(f) = \begin{cases} 1 & f \text{ orientation-preserving} \\ -1 & \text{otherwise} \end{cases}$$

- If f is NOT subjective, then $\deg(f) = 0$.
- If $X, Y \subseteq \mathbb{R}^n$, then $\deg(f) = 1$.
- If there $\exists$ a proper homotopy from f to g, then $\deg(f) = \deg(g)$.
- If $f : X \to Y$ is a covering space, then $\deg(f) = \#\{\text{sheets}\}$.

The last two cases basically say that, a smooth & proper map f can be viewed as a "global" covering space with some local "overlapping", as it behaves well at limit points on $\mathrm{Im}(f)$; and therefore $\deg(f)$ can be thought as $\#$(sheets) of this covering space.
We conjecture that $\deg(f)$ should possess 2 key properties of $\#$(sheets).

Proposition 5.10. Given proper smooth maps $f : X \to Y$ and $g : Y \to Z$,

1. $\deg(f) \in \mathbb{Z}$
2. $\deg(g \circ f) = \deg(g) \deg(f)$

Proof. The **2** follows immediately by the chain law of pullback, while the **1** can be seen as a corollary of Sard's theorem (Proposition 2.10).
Define $\{[a]_f : a \in X\}$ to be the equivalent classes on X where $[a]_f \cong [b]_f$ iff $f(a) = f(b)$. Notice that since f is proper, any $[a]_f$ can contain only finitely many points.
Now consider a local chart (U, φ), and let $C \subset U$ be the set of f's critical points in U; by

Sard's theorem, we have that

$$\lambda\left(\varphi\left(\bigcup_{C_j\in[C]_f} C_j\cap U\right)\right)=0$$

therefore, there exists $y\in Y$ such that $f_*|_x$ is full rank for $\forall x\in f^{-1}(y)$. Since $f^{-1}(y)$ is finite, by inverse function theorem, there exists neighbour $N\cong\mathbb{R}^n$ of y such that

$$f^{-1}(N)=\bigsqcup_j N_j$$

for some $N_j\subset X$ with charts $\varphi_j:N_j\to\mathbb{R}^n$. So by taking $\omega\in\Omega_c^n(N)$, the $\deg(f)$ equals to $\#\,(\text{orientation-preserving }\varphi_j)-\#\,(\text{orientation-reversing }\varphi_j)$, which is an integer. □

This simple invariant actually tells a lot of things, involving two most well-known theorems in topology & analysis: *Hairy ball theorem* and *Brouwer fixed-point theorem*;

Corollary 5.3. (**Hairy ball theorem**) Any $F\in\Gamma(T\mathbb{S}^{2n})$ has a zero.

To prove it, let's assume $\mathbb{S}^{2n}:=\{x\in\mathbb{R}^{2n+1}:||x||=1\}$ and thus $T_x\mathbb{S}^{2n}:=x^\perp$. Suppose there does exist a nowhere-zero $F\in\Gamma(\mathbb{S}^{2n})$, then it's possible to construct a homotopy

$$\begin{aligned}\Phi:[0,\pi]\times\mathbb{S}^{2n}&\to\mathbb{S}^{2n}\\(t,x)&\mapsto\cos(t)x+\sin(t)\frac{F}{||F||}\end{aligned}$$

from $\text{id.}=\Phi(0,x)$ to $(-\text{id.})=\Phi(\pi,x)$. Notice that $\deg(\text{id.})=1$ while $\deg(-\text{id.})=-1$ by checking that $(-\text{id.})$ is orientation-reversing on $\mathbb{S}^{2n}$; however, since degree is invariant under smooth homotopy, contradiction. □

A great generalization to this result is due to J.F.Adam's *Vector fields on Spheres*, where he stated and proved that: Suppose $n=(2a+1)2^b$ and $b=c+4d$ for $c\in\{0,1,2,3\}$, then $T\mathbb{S}^{n-1}$ at most (2^c+8d-1) no-where-zero sections that are linearly independent.
Another famous application is:

Corollary 5.4. (**Brouwer fixed-point theorem**) Any continuous map f from $\mathbb{D}^n$ to itself has fixed point.

There are too many proofs for it, including one just by the Hairy ball theorem, so I won't present a proof for it. Also because of the number of proofs it has, this theorem comes with a lot of generalizations.
One generalization is the *Poincaré-Hopf theorem*, and further the *Lefschetz-Hopf theorem*.

An innovative thought for the fixed point theorem is to translate the fixed point problem in terms of intersection theory.

How? Notice that for any $\phi \in C^\infty(M)$, it naturally gives a map $\Phi \in C^\infty(M, M \times M)$ by $\Phi(x) := (x, \phi(x))$, so that we'll have

$$\{\text{fixed points of } \phi\} \longleftrightarrow \{\text{intersection between } \Phi(M) \text{ and } \triangle_M\}$$

where $\triangle_M = \{(x,x) : x \in M\} \subset M \times M$. So we propose the *Lefschetz number*:

Definition 5.2. Given a smooth manifold X and $f \in C^\infty(X, X)$, the **Lefschetz number** Λ_f of map f is defined as

$$\Lambda_f := \mathcal{I}(\triangle_X, F(X))$$

where $F : X \to X \times X$ is given by $F(x) = (x, f(x))$.

The intuition for this number could be obtained from our interpretation of intersection number, but it is no doubt more complicated to be accurate. It is helpful in theory than in physical intuition, though we can at least know that $|\Lambda_f| \neq 0$ means f has fixed point(s).
Hence, combined with the intersection theory we got in the last part, we can establish a link between $H^\bullet_{\mathrm{dR}}(M)$ and the number of fixed points.

Proposition 5.11. (**Lefschetz fixed-point theorem**) For compact M and smooth map $f : M \to M$, we have

$$\Lambda_f = \sum_{k=0}^{\infty} (-1)^k \cdot \mathrm{tr}(f^*|_{H^k_{\mathrm{dR}}(M)})$$

Proof. □

5.7 (Special) Čech cohomology and Spectral sequence

This part contains no necessary knowledge for further development; it's solely something interesting that deserves a good introduction.

Clearly, the M-V sequence can be generalized to $M = U_1 \cup \cdots \cup U_m$ for m beyond 2; but one may ask if m could be infinite —— If so, then we'll have a more natural way to compute (co)homology. Why? Because that means the (co)homology is computed directly from the open cover itself, a construction based on topology. Hence, this idea is useful in algebraic topology, somewhere the topology of spaces is coarser than smooth manifolds.

Let's suppose M has an open cover $\{U_\alpha : \alpha \in A\}$ (for some countable set A); as an analogy to the case where there are only two open sets, we denote

$$\begin{aligned} U_{\alpha_1 \ldots \alpha_r} &:= U_{\alpha_1} \cap \cdots \cap U_{\alpha_r} \\ \omega_{\alpha_1 \ldots \alpha_r} &:= \text{restriction of } \omega \text{ on } U_{\alpha_1 \ldots \alpha_r} \end{aligned}$$

then define $\theta : \Omega^\bullet(M) \to \bigoplus_\alpha \Omega^\bullet(U_\alpha)$, so that the α component of $\theta(\omega)$ is the restriction of ω on U_α; and

$$\begin{aligned} \theta_{\alpha_1 \ldots \alpha_r} : \bigoplus_{\alpha_1, \ldots, \alpha_r} \Omega^\bullet(U_{\alpha_1 \ldots \alpha_r}) &\longrightarrow \bigoplus_{\alpha_1, \ldots, \alpha_{r+1}} \Omega^\bullet(U_{\alpha_1 \ldots \alpha_{r+1}}) \\ \omega_{\alpha_1 \ldots \alpha_r} &\longmapsto \end{aligned}$$

Proposition 5.12.

5.8 Questions

In these questions, we assume M to be a n-dim smooth manifold.

1. Show that the linking homomorphism ϑ (5.2) is well-defined, independent of choice of POU, and exact in the M-V sequence.

2. Here is a proof of the Poincaré duality that we emitted; for recollection, we defined the homomorphism $\mathcal{P} : H^k_{\mathrm{dR}}(M) \to \left(H^{n-k}_c(M)\right)^*$ by

$$\mathcal{P} : [\omega] \longmapsto \left([\eta] \mapsto \int_M \omega \wedge \eta\right)$$

where $[\omega] \in H^k_{\mathrm{dR}}(M)$ and $[\eta] \in H^{n-k}_c(M)$. Let's say *smooth manifold X is PD* if $\mathcal{P}$ is an isomorphism in case of $M = X$.

1. Consider two open sets $U, V \subset M$ such that U, V and $U \cap V$ are all PD, then we have that $U \cup V$ is PD.
2. Verify that $\mathbb{R}^n$ is PD.
3. By Corollary 2.4, prove that $\mathcal{P}$ is always isomorphism.

5.9 Answers

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & H^k_{\mathrm{dR}}(M) & \xrightarrow{\alpha_*} & H^k_{\mathrm{dR}}(U) \oplus H^k_{\mathrm{dR}}(V) & \xrightarrow{\beta_*} & \Omega^k(U \cap V) & \longrightarrow & \cdots \\
 & & \big\downarrow \mathrm{d} & & \big\downarrow \mathrm{d} & & \big\downarrow \mathrm{d} & & \\
\cdots & \longrightarrow & H^{k+1}_c(M) & \xrightarrow{\alpha_*} & H^{k+1}_c(U) \oplus H^{k+1}_c(V) & \xrightarrow{\beta_*} & H^{k+1}_c(U \cap V) & \longrightarrow & \cdots
\end{array}$$

where the rows are the M-V sequences of $H^{\cdot}_{\mathrm{dR}}$ and $H^{\cdot}_c$ respectively.

6 "Transport" of Derivatives

Welcome to the physics part.
All we have in Chapters 1 and 2 deal with operations around a single point on M; surely a neighbour is enough for defining a derivative, but clearly not enough for a second and higher derivative. The dynamics on M is then very limited.

The way we solve this is to send a "messenger" travelling from one point to another, carrying a "frame" without any reshaping or twisting it when travelling. This provides a fair way to compare things between 2 points on M. We achieve such an idea by something called the **Levi-Cita connection**.

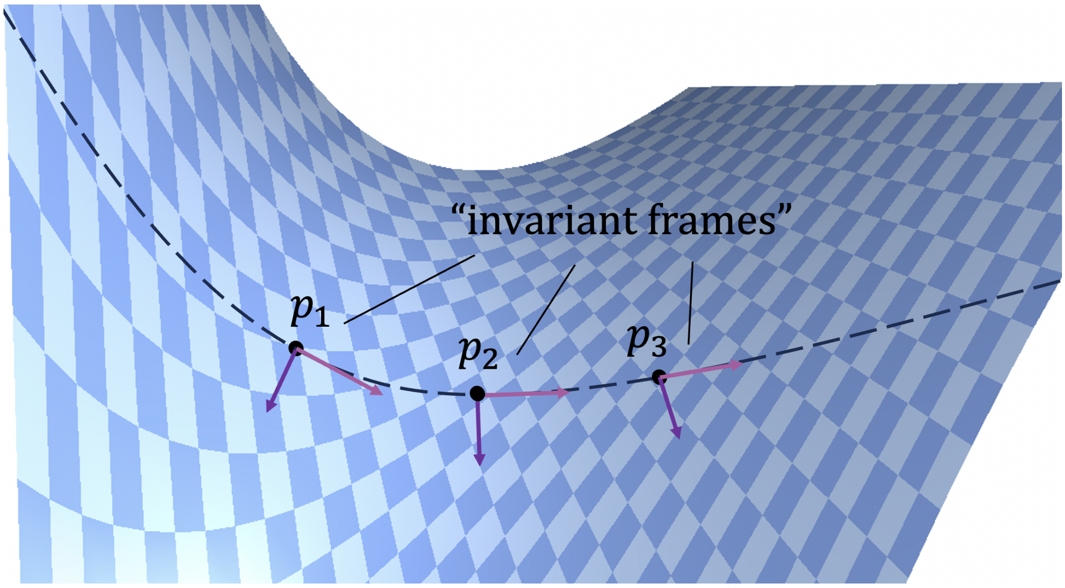


Figure 6: The messenger carries a "frame", which we can regarded as 2 "rods" representing a coordinate. If we have a vector in this frame in tangent space p_1, then we can transport it into p_2 and p_3 by moving this messenger together with this frame.

The main problem is how to properly define a *"frame" travels without reshaping or twisting*, though it is pretty easy to do this when the space is flat or curved in a simple way (e.g sphere); so this will be our main objective.

[Convention]: In this section,

- M is a n-dimensional connected Riemannian manifold.
- We will say "in ESC" when we use the Einstein Summation convention;
- The basis for vector space V or $\mathbb{R}^n$ is denoted as e_i, while for V^* it will be e^i.
- Let $\{\partial_i\}$ be represent a coordinate basis for T_pM, and $\{\mathrm{d}x_j\}$ for T_p^*M.
- When talking about vector fields, we assume that all of them are smooth.

6.1 More about tensor

Before moving on, we need to know more about tensors.
Let's start with its expression under a given basis. Suppose V is a n-dim $\mathbb{R}$-vector space, we know that a type (p,q) tensor over V is in form of $\mathrm{T}^p_q(V) := \bigotimes^p V \otimes \bigotimes^q V^*$, which means any $T \in \mathrm{T}^p_q(V)$ can be written as an array with n^{p+q} entries,

$$T = T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q}(e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}) \tag{6.1}$$

(in ESC) where each $T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q} \in \mathbb{R}$. A good thing about using this expression is that we can directly write down a multilinear map it corresponds to: For example, the tensor in (6.1) defines a multilinear map

$$T(e^{i_1}, \cdots, e^{i_p}, e_{j_1}, \cdots, e_{j_q}) = T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q}$$

Since $V \cong V^*$, we can change one input of V to an input of V^* by $e_i \mapsto e^i$, and wish that the final result doesn't change; this process is called **raising indices**; and alternatively, we also have **lowering indices**. Take an example,

Example 6.1. For a type $(1,1)$-tensor with components T^j_i, the tensor obtained by **lowering** and **raising** its indices is defined as

$$T_{ab} = g_{aj}T^b_j \quad \text{and} \quad T^{ab} = g^{aj}T^b_j$$

(in ESC) respectively.

However, be aware that as geometrical objects, tensors are **independent of choices of basis**. In other words, when it defines a multilinear map, this map is equipped with a change of basis operation from the Definition below:

Definition 6.1. If we change the basis of V by $e_i \mapsto Ae_i$,(where $A \in \mathrm{GL}_n(\mathbb{R})$) then if T with new basis is represented by the array $T^{i'_1,\cdots,i'_p}_{j'_1,\cdots,j'_q}$, we have, in ESC,

$$T^{i'_1,\cdots,i'_p}_{j'_1,\cdots,j'_q} = (A^{-1})^{i_1}_{i'_1} \cdot \cdots \cdot (A^{-1})^{i_p}_{i'_p} \cdot T^{i_1,\cdots,i_p}_{j_1,\cdots,j_q} \cdot A^{j_1}_{j'_1} \cdot \cdots \cdot A^{j_p}_{j'_p}.$$

where A^r_c represents the (r,c) entry of A.

The reasoning behind this formulation is that, when the basis of tensor and input are changed, the output stays unchanged; in other word, independent of the choice of basis.
Hence this definition is based on the fact that the new basis of V obeys $(e')_i = A^i_j e_i$ in ESC, which suggests that new basis of V^* obeys $(e')^i = (A^{-1})^i_j e^i$ since $(e')_i(e')^i = 1$; the remained derivation is straightforward but lengthy, so I left as an exercise.

Next, we introduce a rather important idea that "reduces" a tensor by pairing the vector component and the dual component, we call this **tensor contraction** $\mathcal{C} : \mathrm{T}^p_q(V) \to \mathrm{T}^{p-1}_{q-1}(V)$:

Definition 6.2. Suppose $\mathcal{C}(T)$ is the tensor contraction of $T \in \mathrm{T}^p_q(V)$, where $\mathcal{C}(T) = \tilde{T}^{i_1,\cdots,i_{p-1}}_{j_1,\cdots,j_{q-1}}(e_{i_1} \otimes \cdots \otimes e_{i_{p-1}} \otimes e^{j_1} \otimes \cdots \otimes e^{j_{q-1}}) \in \mathrm{T}^{p-1}_{q-1}(V)$, then

$$\tilde{T}^{i_1,\cdots,i_{p-1}}_{j_1,\cdots,j_{q-1}} = \sum_{r=1}^{\dim V} T(e^{i_1}, \cdots, e^{i_{p-1}}, e^r, e_{j_1}, \cdots, e_{j_{q-1}}, e_r)$$

The simplest example is when T is a linear transform (which is when $T \in \mathrm{T}^1_1$); in this case $\mathcal{C}(T) = \mathrm{tr}(T)$. You can also try other examples to get more familiar with this strange operation.

Example 6.2. Consider a tensor $T \in \mathrm{T}^1_2(\mathbb{R}^3)$ written in basis $\{e_i, e^i\}$ defined as

$$T = (i - j - k)(e_i \otimes e^j \otimes e^k) \text{ in ESC}$$

then $[\mathcal{C}(T)](e_j) = \sum_{r=1}^3 (r + j - r) = 3j$, which gives $\mathcal{C}(T) = (3, 6, 9) \in V^*$.

We can view this operation as a generalization of **how dual vectors acting on vectors**. For example, since $(V^* \otimes V^*)$ is a dual of $(V \otimes V)^*$, the tensor contraction provides a natural duality:

$$\begin{aligned} &(e^i \otimes e^j) \text{ "acting" on } (e_k \otimes e_l) \\ :=&\mathcal{C}^2(e_k \otimes e_l \otimes e^i \otimes e^j) \\ =&\delta_{ik}\, \delta_{jl} \end{aligned}$$

But in practice, we do NOT always write tensor contraction in above way; because we can just write things in ESC; for example, for a metric tensor $g \in \mathrm{T}^0_2$ acting on $(u, v) \in V$, which is $C^2((u \otimes v) \otimes g)$, it is simply equal to $g_{ij}u_iv_j$ in ESC.

6.2 Main idea: Parallel transport

Let’s formalize our intuition: defining the idea of “*moving on M without reshaping or twisting*” that we described at the beginning. We will give a math definition afterwards, but here let’s call it **Parallel transport**.

We know that M must be path-connected since it is connected, which means for any two points $a, b \in M$, there must be a continuous path $\gamma : [0,1] \to M$ with $\gamma(0) = a$, $\gamma(1) = b$ that can be covered by **finite number** of local coordinates.

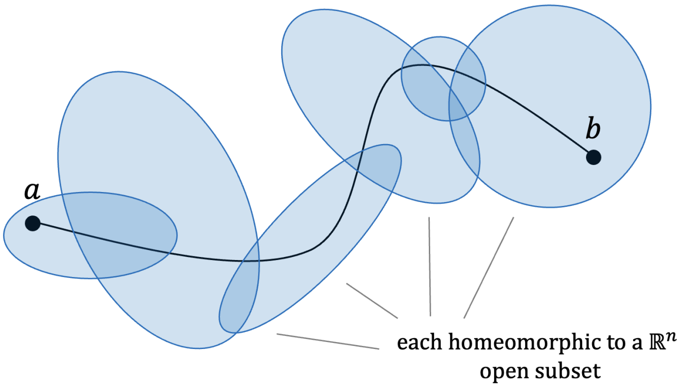


Figure 7: By compactness of $[0,1]$, any path from a to b can be covered by a finite number of neighbours that homeomorphic to open subsets of $\mathbb{R}^n$.

This suggests that it is sufficient to *define the parallel transport inside a single coordinate.* Consider a smooth vector field U, and a path $\gamma(s)$ from a to b. And suppose there is a vector field $U_\parallel$ such that $U_\parallel\big|_{\gamma(s)}$ is “parallel” to $U\big|_{\gamma(0)}$ at $\gamma(s)$.
We know that in $M = \mathbb{R}^n$, this process “$U_\parallel$” of transporting a point with velocity $U(s)$ along $\gamma(s)$ is always “parallel” if and only if

$$\left(\nabla_{\gamma'(s)} U_\parallel\right)\big|_{\gamma(s)} = 0 \text{ for } \forall s \in [0,1]$$

So accordingly we can make a generalization, assuming we have already defined “∇”:

Definition 6.3. Suppose neighbour $N_p \in M$ is homeomorphic to an open subset of $\mathbb{R}^n$, and $q \in N_p$ has a continuous path $\gamma : [0,1] \to N_p$ connecting p.
The **Parallel transport** of $U(p)$ from p to q is the section $U_\parallel : \gamma(s) \mapsto u_{\gamma(s)} \in T_{\gamma(s)}M$ that $U_\parallel(\gamma(0)) = U(p)$, and satisfies

$$\left(\nabla_{\gamma'(s)} U_\parallel\right)\big|_{\gamma(s)} = 0 \text{ for } \forall s \in [0,1]$$

Thus our task now is to generalize “partial derivative” ∇ in arbitrary M, or more precisely, in arbitrary neighbour in M that is homeomorphic to $\mathbb{R}^n$. And we name this generalized version of ∇ as **Covariant Derivative**.

Consider $U, V \in \Gamma(TM)$ for arbitrary M, we use the same symbol "$\nabla_V U$" as partial derivative for covariant derivative. Since we define parallel transport as a transform along which the covariant derivative is zero; the $\nabla_V U$ calculates the rate of change of the U with respect to its parallel transport.

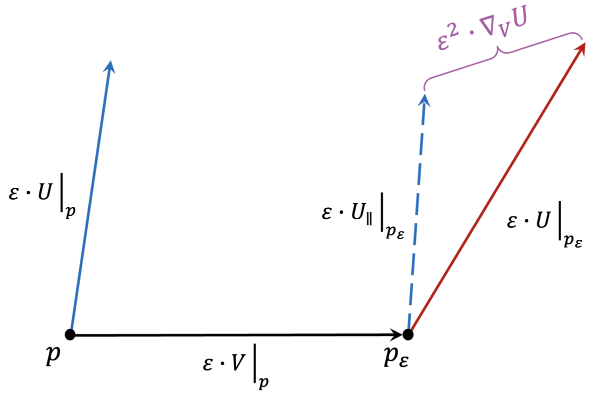


Figure 8: A not rigorous interpretation of $\nabla_V U$. The p_ε represents the location after travelling along $V|_p$ by ε from p, where $\varepsilon \to 0$.

To get a formal definition, we can do similar things as we did in other idea generalizations: Let $\nabla_V U$ preserve all the important properties of partial derivatives: <u>U-linear</u>, <u>V-linear</u>, and <u>Leibniz rule</u>.

But notice that under these rules the $\nabla_V U$ may not be unique; so the rest of the work is to find the unique one under the condition of *"no reshaping and twisting"*.

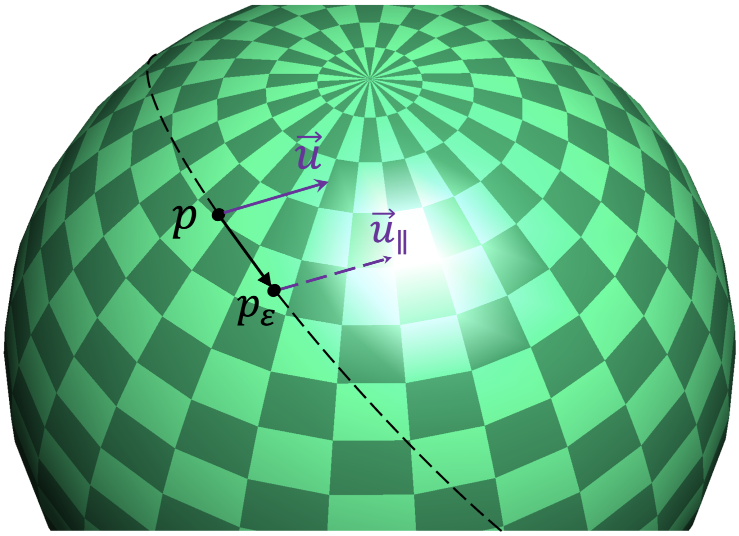


Figure 9: How to parallel transport a rod pointing u on a sphere? This is easy for a human to tell. Because if we go straight, our path must be an equator; the $u_{\parallel}$ is simply the vector with the same length and has an invariant angle to the equator.

6.3 Covariant Derivative

As we said, we can define covariant derivatives as a function that preserves all the important properties of partial derivatives; (which yields the first 3 items); but here we add another property (item 4) which makes it work on tensor fields.

Definition 6.4. Consider arbitrary $F \in \mathcal{T}_s^r M$, the Covariant derivative $\nabla_V(F)$ is a map $\nabla_V : \mathcal{T}_s^r M \to \mathcal{T}_s^r M$ that satisfies

1. **F-Linear:** For $\forall r \in \mathbb{R}$, we have $\nabla_V(F + rG) = \nabla_V(F) + r\nabla_V(G)$
2. **V-Linear:** For $\forall f \in C^\infty(M)$, we have $\nabla_{W+rV}(F) = \nabla_W(F) + f \cdot \nabla_V(F)$
3. **Leibniz rule:** $\nabla_V(F \otimes G) = \nabla_V(F) \otimes G + F \otimes \nabla_V(G)$
4. **Commute with contraction** $\nabla_V(\mathcal{C}(F)) = \mathcal{C}(\nabla_V(F))$

where $V \in \Gamma(TM)$, and F, G are arbitrary tensor fields.

Let's start with the simplest case where $F \in \Gamma(TM)$.
For simplicity, let's denote $\nabla_k := \nabla_{\partial_k}$. By the 2 linearities, we can conclude that a covariant derivative is uniquely determined by the values of $\nabla_k(\partial_j) \in \Gamma(TM)$ for $j, k = 1, \cdots, n$. So we can define some $\Gamma^i_{jk} \in C^\infty(M)$ such that

$$\boxed{\nabla_k(\partial_j) = \Gamma^i_{jk}\partial_i}$$

(in ESC) these Γ^i_{jk} are called the **Christoffel symbol**.
This symbol encodes all the information of a covariant derivative —— as $\nabla_V F$ isn't defined explicitly and uniquely, it is expressed and determined by values of Γ^i_{jk}.

Example 6.3. Let's compute $\nabla_V U$, where $U, V \in TM$:
Set $U = U_j\partial_j$ and $V = V_k\partial_k$ in ESC, then by the numbering (that represents which condition in definition 6.4) above the equal sign, we have

$$\begin{aligned}
\nabla_V(U) &\overset{2}{=} V_k\nabla_k(U_j\partial_j) \\
&\overset{3}{=} V_k\left(\partial_j\nabla_k(U_j) + U_j\Gamma^i_{jk}\partial_i\right) \\
&\overset{3}{=} V_k\left(\partial_j \cdot \partial_k(U_j) + U_j\Gamma^i_{jk}\partial_i\right) = \left(V_k \cdot \partial_k(U_i) + V_kU_j\Gamma^i_{jk}\right)\partial_i
\end{aligned}$$

in ESC, where the last equal sign comes from switching order of summation.

Before moving on, let's think about the meaning of our result.

$$(\nabla_V U)_i = \underbrace{V_k \cdot \partial_k(U_i)}_{\substack{\text{Change produced} \\ \text{by } U \text{ itself}}} + \underbrace{V_kU_j\Gamma^i_{jk}}_{\substack{\text{Change produced} \\ \text{by } M}} \tag{6.2}$$

In Vector calculus, we know that such expression only contains the terms of $V_k \cdot \partial_k(U_i)$ when M has a "flat" global coordinate; hence we can deduce that Γ^i_{jk} reveals the difference between *definition of "straight"* in view of coordinate and of our natural physical intuition.

We can also calculate the result for cotangent vectors:

Example 6.4. Let's start from a simpler case, $\nabla_l \, \mathrm{d}x_m$ for a specific pair of (l, m). We know that the result should be in form of $\sum_j \mathbb{R} \cdot \mathrm{d}x_j$;
Notice that by the 4^{th} condition,

$$\begin{aligned}\nabla_l \left(\mathcal{C}(\partial_j \otimes \mathrm{d}x_m)\right) = 0 &= \mathcal{C}\left(\nabla_l(\partial_j \otimes \mathrm{d}x_m)\right) \\ &= \mathcal{C}\left(\nabla_l(\partial_j) \otimes \mathrm{d}x_m\right) + \mathcal{C}\left(\partial_j \otimes \nabla_l(\mathrm{d}x_m)\right)\end{aligned}$$

Thus by taking j from 1 to n, we have: (in ESC)

$$\nabla_l(\mathrm{d}x_m) = -\left(\Gamma^m_{jl}\right) \mathrm{d}x_j$$

Then use linearity, we have that for $W = W_j \, \mathrm{d}x_j$ and $V = V_k \partial_k$ in ESC,

$$\begin{aligned}\nabla_V(W) &= V_k \left(\nabla_k(W_j) \, \mathrm{d}x_j + W_j \nabla_k(\mathrm{d}x_j)\right) \\ &= V_k \left(\partial_k(W_j) \, \mathrm{d}x_j - W_j \Gamma^j_{ik} \, \mathrm{d}x_i\right) \\ &= \left(V_k \cdot \partial_k(W_i) - V_k W_j \Gamma^j_{ik}\right) \mathrm{d}x_i\end{aligned}$$

(in ESC) by similar approach.

Put them together, that is

$$(\nabla_k(F))_i = \begin{cases} \partial_k(F_i) + F_j \Gamma^i_{jk} & \text{for } F \in \Gamma(TM) \\ \partial_k(F_i) - F_j \Gamma^j_{ik} & \text{for } F \in \Gamma(T^*M) \end{cases} \tag{6.3}$$

Using this idea, we are able to get a formula for all $F \in \mathcal{T}^r_s M$:

Corollary 6.5. For $F = F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s}(\partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x_{j_1} \otimes \cdots \otimes \mathrm{d}x_{j_s}) \in \mathcal{T}^r_s M$, we have

$$\begin{aligned}(\nabla_k F)^{i_1, \cdots, i_r}_{j_1, \cdots, j_s} = \partial_k(F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s}) &+ F^{l, \cdots, i_r}_{j_1, \cdots, j_s} \Gamma^{i_1}_{lk} + \cdots + F^{i_1, \cdots, l}_{j_1, \cdots, j_s} \Gamma^{i_r}_{lk} \\ &- F^{i_1, \cdots, i_r}_{l, \cdots, j_s} \Gamma^l_{j_1 k} - \cdots - F^{i_1, \cdots, i_r}_{j_1, \cdots, l} \Gamma^l_{j_s k}\end{aligned}$$

in ESC. (summing over $l = 1, 2, \cdots, n$)

Proof. Notice that by Leibniz rule, if T_i are tensor fields, then

$$\nabla_k \left(\bigotimes_{0 \le m \le r+s} T_m \right) = \sum_{0 \le m \le r+s} (\nabla_k T_m) \otimes \left(\bigotimes_{\substack{1 \le d \le r+s \\ d \ne m}} T_d \right)$$

By plugging in the $T_0 = F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s}$, $F_a = \partial_{i_a}$, and $F_b = \partial_{j_b}$, we have

$$\begin{aligned}&\nabla_k F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s}(\partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x_{j_1} \otimes \cdots \otimes \mathrm{d}x_{j_s}) \\ &\qquad = \partial_k(F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s}) + F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s} \Gamma^l_{i_1 k} \partial_l \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x_{j_1} \otimes \cdots \mathrm{d}x_{j_s} \\ &\qquad\quad + \cdots + F^{i_1, \cdots, i_r}_{j_1, \cdots, j_s} \Gamma^l_{i_r k} \partial_{i_1} \otimes \cdots \otimes \partial_l \otimes \mathrm{d}x_{j_1} \otimes \cdots \mathrm{d}x_{j_s}\end{aligned}$$

$$- F^{i_1,\cdots,i_r}_{j_1,\cdots,j_s}\Gamma^{j_1}_{lk}\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes \mathrm{d}x_l\otimes\cdots \mathrm{d}x_{j_s}$$
$$-\cdots - F^{i_1,\cdots,i_r}_{j_1,\cdots,j_s}\Gamma^{j_s}_{lk}\partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes \mathrm{d}x_{j_1}\otimes\cdots \mathrm{d}x_l$$

(in ESC, summing over $l = 1, 2, \cdots, n$) Then by switching order of summation, we get the desired result. □

Example 6.6. Though it is a long formula, but it has simple pattern; for example, we can easily write down the case for Riemannian metric $g = g_{ij}(\mathrm{d}x_i \otimes \mathrm{d}x_j)$:

$$(\nabla_k g)_{ij} = \partial_k g_{ij} - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{il}$$

in ESC, summing over $l = 1, \ldots, n$.

Now we can see that the ∇ depends solely on the value of Christoffel symbols, our goal narrows down to restricting these $n \times n \times n$ functions so that the parallel transport produced fits the physics law.

And this brings us to the next step: Defining *reshaping* & *twisting*.

6.4 Levi-Civita Connection

As discussed, a covariant derivative ∇ that fits into our naïve physical intuition should demonstrate the idea of:

1. **no compression**: so that the frame itself will not reshape.
2. **no torque**: so that the direction of it will not change (i.e. any twisting)

The $1^{\text{st.}}$ idea is quite simple, it can be achieved by preserving the metric, so that the relative location of vectors in the tangent space will not change when moving; this is to say, $\nabla_V g = 0$ for $\forall V \in \Gamma(TM)$.
While the second idea needs some elaborate thinking; we call it "**Torsion-free property**".

Imagine you are driving a car on a 2-dim surface. How can you tell that you are going straight? One way is to observe the speed difference between two tires —— if your left tire is slower, then you are turning left;

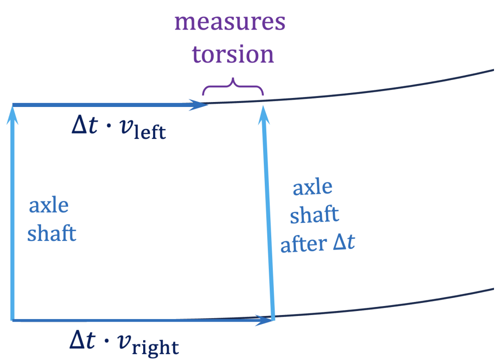


Figure 10: v_{right} stands for the velocity of the right tire, and v_{left} is that of the left tire.

Now we can express the infinitesimal right tire displacement as εv, and the driving shaft as εu, so that the parallel transport $\varepsilon v_{\parallel}$ gives infinitesimal left tire displacement.

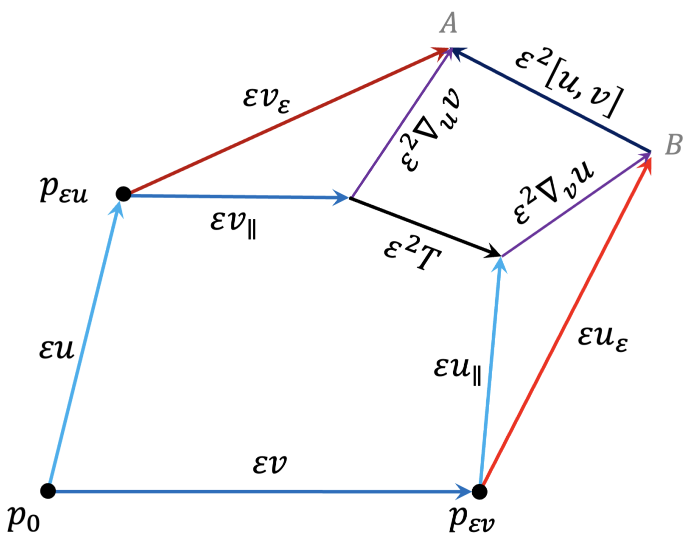


Figure 11: Illustration for the idea of torsion tensor. In this figure, $v_\varepsilon = V(p_{\varepsilon u})$ and $v_\parallel$ is the parallel transport of v at $p_{\varepsilon u}$; similar for u_ε and $u_\parallel$. The $[u,v](f) = u(v(f)) - v(u(f))$ (for $\forall f \in C^\infty$) is the Lie bracket.

Then we need to express T in objects we've already defined. First, when $\varepsilon \to 0$, the covariant derivatives can be approximated as $\nabla_u v \approx (v_\varepsilon - v_\|)/\varepsilon$ and $\nabla_v u \approx (u_\varepsilon - u_\|)/\varepsilon$. Next, consider the vector representing $\vec{AB}$;
Notice that u, v are tangent vectors, which naturally act as partial derivatives in their directions respectively; hence

$$\vec{AB} = (\varepsilon v_\varepsilon + \varepsilon u) - (\varepsilon u_\varepsilon + \varepsilon v) = \varepsilon(v_\varepsilon - v) - \varepsilon(u_\varepsilon - u) \approx \varepsilon^2[u, v]$$

which gives $T = \nabla_u v - \nabla_v u - [u, v]$. Write this idea formally, it would be

Definition 6.5. Consider $U, V \in \Gamma(TM)$, the **torsion tensor** $T(U, V)$ is a tensor field (more precisely a vector field) defined as

$$T(U, V) = (\nabla_U V - \nabla_V U) - [U, V]$$

The covariant derivative on M is said to be **torsion-free** iff $T(U, V) = 0$ for $\forall U, V$.

If we are using a coordinate basis, the lie bracket vanishes, so we have $T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj}$, which means torsion-free is **equivalent** to say $\Gamma^i_{jk} = \Gamma^i_{kj}$.
These $n^2(n-1)/2$ equations restrict the number of free Γ^i_{jk} values to $n^2(n+1)/2$. Together with the $n^2(n+1)/2$ equations from $\nabla_i g = 0$, it is now possible to define a unique covariant derivative:
it is the solution to equations

$$\begin{cases} \nabla_k g = 0 \\ \Gamma^i_{jk} = \Gamma^i_{kj} \end{cases} \tag{6.4}$$

for $\forall i, j, k = 1, 2, \cdots, n$. (not in ESC) Now we can solve this set of equations, and the solution we get is the Christoffel symbol for **Levi-Civita connection.**

Definition 6.6. The Levi-Civita connection is the covariant derivative computed by the Christoffel Symbols computed by:

$$\Gamma^i_{jk} = \frac{g^{ir}}{2}\left(\partial_k g_{rj} - \partial_r g_{jk} + \partial_j g_{rk}\right)$$

(in ESC) where g^{ir} is the (i, r) entry of the inverse of matrix representation of g.

Proof. By Example 6.6, we have that,

$$(\nabla_k g)_{ij} = 0 = \partial_k(g_{ij}) - g_{il}\Gamma^l_{jk} - g_{lj}\Gamma^l_{ik}$$

(in ESC) Now suppose we let $\Gamma_{ijk} = g_{il}\Gamma^l_{jk}$ in ESC, which means $\Gamma^i_{jk} = g^{ir}\Gamma_{rjk}$; above equation is then $\partial_k(g_{ij}) = \Gamma_{ijk} + \Gamma_{jik}$.

Notice that $\Gamma_{ijk} = \Gamma_{ikj}$, which suggests

$$\begin{aligned}(\Gamma_{ijk} + \Gamma_{jik}) - (\Gamma_{jki} + \Gamma_{kji}) + (\Gamma_{kij} + \Gamma_{ikj}) &= 2\Gamma_{ijk} \\ &= \partial_k g_{ij} - \partial_i g_{kj} + \partial_j g_{ik}\end{aligned}$$

Hence we have $\Gamma^i_{jk} = \frac{1}{2} g^{ir} (\partial_k g_{rj} - \partial_r g_{jk} + \partial_j g_{rk})$. □

Computing the Christoffel symbols is generally a time taking task even for simple cases:

Example 6.7. Take two examples, sphere surface $\mathbb{S}^2$ and torus $\mathbb{T}^2$.
$\mathbb{S}^2$: Consider an open subset $U \subset \mathbb{S}^2$ with chart given by the spherical coordinate $p \mapsto (\theta(p), \phi(p))$. Define the metric tensor g as

$$g = \mathrm{d}\theta \otimes \mathrm{d}\theta + \sin(\theta)^2 \mathrm{d}\phi \otimes \mathrm{d}\phi$$

(so that it describes the infinitesimal length of the path between two points on a unit sphere, as $\mathrm{d}s = \sqrt{\mathrm{d}\theta^2 + \sin(\theta)\mathrm{d}\phi^2}$.)
In this case,

$$\Gamma^i_{jk} = \begin{cases} -\sin(\theta)\cos(\theta) & \text{for } (i,j,k) = (1,2,2) \\ \sin(\theta)^{-1}\cos(\theta) & \text{for } (i,j,k) = (2,2,1) \text{ or } (2,1,2) \\ 0 & \text{otherwise} \end{cases}$$

You can see that the $\Gamma^{\cdot}_{jk} = 0$ only when $(j,k) = (1,1)$; this is actually quite intuitive as $\Gamma^{\cdot}_{11}$ represents the coordinate changing when moving longitudinally.
$\mathbb{T}^2 \cong \mathbb{S}^1 \times \mathbb{S}^1$ Consider $\mathbb{S}^1 \times \mathbb{S}^1$ with $\mathbb{S}^1 := [0, 2\pi]/\sim$ (where $x \sim y$ iff $x, y \in \{0, 2\pi\}$) and $V = (0, 2\pi)/\sim\ \subset \mathbb{S}^1$. Let $U = V \times V$, equipped with chart $([x]_\sim, [y]_\sim) \mapsto (x, y)$. Define the metric tensor g as

$$g = (b + a\cos(y))\,\mathrm{d}x \otimes \mathrm{d}x + a^2 \mathrm{d}y \otimes \mathrm{d}y$$

(so that it describes the infinitesimal length of the path on $\mathbb{S}^1 \times \mathbb{S}^1$ embedded in $\mathbb{R}^3$ as a standard torus with aspect ratio b/a.)
In this case,

$$\Gamma^i_{jk} = \begin{cases} \sin(y)\left(\frac{b}{a} + \cos(y)\right) & \text{for } (i,j,k) = (2,1,1) \\ -\sin(y)\left(\frac{b}{a} + \cos(y)\right)^{-1} & \text{for } (i,j,k) = (1,1,2) \text{ or } (1,2,1) \\ 0 & \text{otherwise} \end{cases}$$

You can see that the $\Gamma^{\cdot}_{jk} = 0$ only when $(j,k) = (2,2)$; this is actually quite intuitive as $\Gamma^{\cdot}_{22}$ represents the coordinate changing when moving in the outer circle.

which are a bit useless in developing general theory. So to avoid Chrisoffel symbols in covariant derivative computation, we make two modifications to our approach

1. Instead of using a coordinate, use the axioms of the Levi-Civita connection (6.4) directly.
2. Since their Riemannian metrics are induced by their embedding in $\mathbb{R}^3$, so should their Levi-Civita connections.

We demonstrate these with an useful example.
Equip $\mathbb{R}^{n+1}$ with some $\overline{g} = \langle \cdot, \cdot \rangle$ as its Riemannian metric, and let $\overline{\nabla}$ be the corresponding Levi-Civita connection. Then suppose $M \subset \mathbb{R}^m$, and g is the Riemannian metric induced by embedding.
There should be a relation between ∇ and $\overline{\nabla}$; using some physical intuition, we could guess that ∇ is the orthogonal projection of $\overline{\nabla}$ on M:

Proposition 6.1. Given $V, F, N \in \Gamma(T\mathbb{R}^{n+1})$ such that $V|_M, F|_M \in \Gamma(TM)$ and $N|_p$ is the unit normal vector at any $p \in M$. We have

$$\overline{\nabla}_V F = \nabla_V F + \langle \overline{\nabla}_V F, N \rangle \; N$$

Proof. First, the ∇ here obviously satisfies the 2 linearities and Leibniz rule by corresponding property of $\overline{g}$ and $\overline{\nabla}$, making it a valid covariant derivative[10]. The remaining is to verify it's Levi-Civita.
- The ∇ is torsion-free:

$$\begin{aligned}\nabla_V F - \nabla_F V - [V, F] &= \underbrace{\overline{\nabla}_V F - \overline{\nabla}_F V - [V, F]}_{=0} + \langle \overline{\nabla}_V F - \overline{\nabla}_F V, N \rangle N \\ &= \langle \underbrace{\overline{\nabla}_V F - \overline{\nabla}_F V - [V, F]}_{=0}, N \rangle N + \langle \underbrace{[V, F]}_{\in N^\perp}, N \rangle N = 0\end{aligned}$$

- The g is invariant under ∇: Notice that by Leibniz rule, for $\forall V \in \Gamma(TM)$ and $X, Y \in \Gamma(T\mathbb{R}^{n+1})$ such that $X|_M, Y|_M \in \Gamma(TM)$,

$$\begin{aligned}(\nabla_V g)(X, Y) &= \nabla_V g(X, Y) - g(\nabla_V X, Y) - g(\nabla_V Y, X) \\ &= \overline{\nabla}_V \langle X, Y \rangle - \langle \overline{\nabla}_V X, Y \rangle - \langle \overline{\nabla}_V Y, X \rangle\end{aligned}$$

which is 0 since $\langle \cdot, \cdot \rangle$ is invariant under $\overline{\nabla}$.
As Levi-Civita connection is unique under a fixed g, thus the ∇ is the Levi-Civita connection on M. □

This proof avoids the use of Christoffel symbols, though I believe it is a good exercise trying to prove it using them.

Remark. This conclusion can immediately be upgraded by replacing $\mathbb{R}^{n+1}$ with arbitrary

[10] Here we only deal with covariant derivative of vector fields, so the condition of commuting with tensor contraction is omitted

smooth manifold X with $\dim X > n$. However, the equation may not exist since $\mathrm{N}(M, X)$ may not have a global basis like $\mathrm{N}(M, \mathbb{R}^{n+1})$ —— but of course, this approach still works locally.

This formulation is easier for computing covariant derivatives on surfaces being embedded in Euclidean space. For example, the covariant derivative $\mathbb{S}^n$:

Example 6.8. Embed $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ as a unit sphere surface, consider $V, F \in \Gamma(T\mathbb{R}^{n+1})$ such that $V|_{\mathbb{S}^n}, F|_{\mathbb{S}^n} \in \Gamma(T\mathbb{S}^n)$,

$$\nabla_V F = \overline{\nabla}_V F + \langle V, F \rangle \vec{p}$$

for $\forall p \in \mathbb{S}^n$.

Proof. By using the fact that $N = \vec{p}$, we have $\nabla_V F = \overline{\nabla}_V F - \langle \overline{\nabla}_V F, \vec{p} \rangle \vec{p}$. Then by Leibniz rule, we know that $\langle \overline{\nabla}_V F, \vec{p} \rangle = \overline{\nabla}_V \langle F, \vec{p} \rangle - \langle F, \overline{\nabla}_V \vec{p} \rangle = -\langle F, V \rangle$, and we're done. $\square$

6.5 Parallel transport and Geodesics

We can now answer our question at the beginning, about how to parallel transport a tangent vector on M without "external force".

Recall that, consider a path $\gamma : [0,1] \to M$ in a neighbour N_p that is homeomorphic to an open subset of $\mathbb{R}^n$, the parallel transport $U_\parallel$ is the solution to $\nabla_{\gamma'(s)} U_\parallel = 0$.
Now set $\gamma_k(s)$ as the $k^{\text{th.}}$ components in coordinate expression of γ, and $U_\parallel = U_j \partial_j$ in ESC; then by the Levi-Civita connection,

$$\begin{aligned}\nabla_{\gamma'(s)} U_\parallel = 0 &= \nabla_{\frac{\mathrm{d}\gamma_k}{\mathrm{d}s}\partial_k} (U_j \partial_j) \\ &= \frac{\mathrm{d}\gamma_k}{\mathrm{d}s} \left(\partial_k (U_i) + U_j \Gamma^i_{jk} \right) \partial_i\end{aligned}$$

(in ESC) which gives the differential equation for **parallel transport** in N_p: For $i = 1, \cdots, n$,

$$\boxed{\frac{\mathrm{d}U_i}{\mathrm{d}s} + U_j \Gamma^i_{jk} \frac{\mathrm{d}\gamma_k}{\mathrm{d}s} = 0} \tag{6.5}$$

(in ESC, but fixing i, summing over j and k) Under Levi-Civita connection, this equation answers our question at the beginning. It is a linear autonomous ODE, so since Γ^i_{jk} are smooth, the solution exists for any smooth path.

We can also define what it means by *"going straight"* now: When a path is straight, its direction at the next infinitesimal moment is simply the parallel transport of its direction now. This idea is called the **Geodesic**; express it formally,

Definition 6.7. Consider a neighbor N_p in M that is homeomorphic to an open subset of $\mathbb{R}^n$, the path $\gamma : [0,1] \to N_p$ is a Geodesic if

$$\nabla_{\gamma'(s)} \gamma'(s) = 0 \text{ for } \forall s \in [0,1]$$

We can also expand it as equations for components like (6.5): For $i = 1, \cdots, n$,

$$\boxed{\frac{\mathrm{d}^2\gamma_i}{\mathrm{d}s^2} + \Gamma^i_{jk} \frac{\mathrm{d}\gamma_j}{\mathrm{d}s} \frac{\mathrm{d}\gamma_k}{\mathrm{d}s} = 0} \tag{6.6}$$

(in ESC, but fixing i, summing over j and k)

We can verify that the geodesic calculated in this way is exactly the same as by the Euler-Lagrange equation with Lagrangian $\mathcal{L}(s, \gamma, \gamma') = \sqrt{[g](\gamma', \gamma')}$:

Proposition 6.2. Among paths connecting two points in the neighbour N_p, the path satisfying (6.6) has the minimal length.

We just need to show that $\mathcal{L}(s,\gamma,\gamma') = \sqrt{[g](\gamma',\gamma')}$ satsifies the E-L equation as mentioned. However, doing this directly will be very long and painful (though possible), here a simpler way to prove it is by realising that:

Lemma 6.9. $\mathcal{L} = \sqrt{[g](\gamma',\gamma')}$ satisfies the E-L equation when $\mathcal{L}^2$ does.

Proof. If $\mathcal{L}^2$ satisfies the E-L equation, then

$$\begin{aligned}\frac{\mathrm{d}\mathcal{L}^2}{\mathrm{d}s} &= \frac{\partial \mathcal{L}^2}{\partial \gamma_i'}\gamma_i'' + \frac{\partial \mathcal{L}^2}{\partial \gamma_i}\gamma_i' \\ &= \frac{\partial \mathcal{L}^2}{\partial \gamma_i'}\gamma_i'' + \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma_i'}\right)\gamma_i' = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma_i'}\cdot\gamma_i'\right)\end{aligned}$$

(in ESC), this is known as the *Beltrami Identity*.
Notice that when $\mathcal{L} = \sqrt{[g](\gamma',\gamma')} = \sqrt{g_{jk}\gamma_j'\gamma_k'}$ (in ESC), we have

$$\frac{\mathrm{d}\mathcal{L}^2}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma_i'}\cdot\gamma_i'\right) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\gamma_i'\sum_k g_{ik}\gamma_k' + \gamma_i'\sum_j g_{ji}\gamma_j'\right) = 2\cdot\frac{\mathrm{d}\mathcal{L}^2}{\mathrm{d}s}$$

suggesting that $\frac{\mathrm{d}\mathcal{L}^2}{\mathrm{d}s} = 0$, which gives $\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}s} = 0$. This can be intuitively interpreted as "the infinitesimal length by each ds is constant". Therefore, we have

$$\begin{aligned}\frac{\partial \mathcal{L}^2}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma'}\right) &= 2\mathcal{L}\left(\frac{\partial \mathcal{L}}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}}{\partial \gamma'}\right)\right) - 2\underbrace{\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}s}}_{=0}\frac{\partial \mathcal{L}}{\partial \gamma'} \\ \implies \frac{\partial \mathcal{L}}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}}{\partial \gamma'}\right) &= 0\end{aligned}$$

□

With this conclusion, we can then prove Proposition 6.2:

$$\begin{aligned}\frac{\partial \mathcal{L}^2}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma'}\right) &= \partial_k(g_{ij})\gamma_i'\gamma_j' - 2\cdot\frac{\mathrm{d}}{\mathrm{d}s}\left(g_{ij}\gamma_j'\right) \\ &= \partial_k(g_{ij})\gamma_i'\gamma_j' - 2\cdot\partial_k(g_{ij})\gamma_k'\gamma_j' - 2g_{ij}\gamma_j'' \\ &= \partial_j(g_{ki})\gamma_k'\gamma_i' - \partial_k(g_{ij})\gamma_k'\gamma_i' - \partial_j(g_{ik})\gamma_k'\gamma_i' - 2g_{lj}\gamma_l''\end{aligned}$$

(in ESC, summing i,j,k,l over $1,2,\cdots,n$) where the 3$^{\text{rd.}}$ equal sign comes from switching order of summation; hence

$$\frac{\partial \mathcal{L}^2}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial \mathcal{L}^2}{\partial \gamma'}\right) = -2g_{jl}\left(\Gamma^l_{ik}\gamma_i'\gamma_k' + \gamma_l''\right) = 0$$

(in ESC) So by lemme 6.9, we conclude that $\mathcal{L}$ satisfies the E-L equation, which finishes the proof of Proposition 6.2.

A good thing to notice is that,

Proposition 6.3. Suppose γ is a geodesic, and $U_\parallel$ is a parallel transport produced along γ. The value of $g\left(\gamma'(s), U_\parallel(s)\right)$ is invariant.

This fact is easy to prove: By Leibniz rule, we have $\nabla_{\gamma'(s)}(\gamma' \otimes U_\parallel \otimes g) = 0$, which means $\nabla_{\gamma'(s)} g(\gamma', U_\parallel) = \mathcal{C}^2\left(\nabla_{\gamma'(s)}(\gamma' \otimes U_\parallel \otimes g)\right) = 0$, so it's an invariant along $\gamma(s)$.
This conclusion is kind of a generalization of the *Corresponding angle rule* of the parallel lines in $\mathbb{R}^2$ we learned in primary school. It suggests that when $n = 2$, the parallel transport along the geodesic is easy to compute: at each point it is just a vector with known angle to the geodesics' tangent.

Per our motivation, the parallel transport defines "parallel" in $\mathbb{R}^n$, which is an equivalence relation; but does it define the "parallel" as an equivalence relation on all manifolds? The answer is NO:
Claim. Parallel transporting a vector via different paths may NOT give the same result.

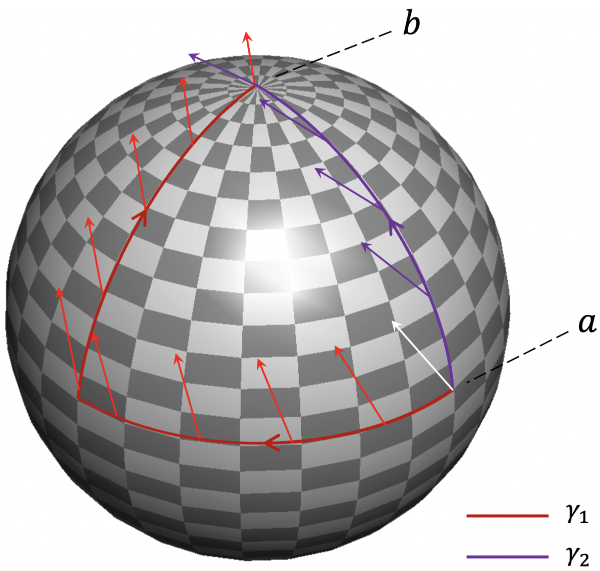


Figure 12: Consider unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$, let a be a point on equator and b at the north pole in spherical coordinate. Set two paths from a to b, where γ_1 goes along the longitude, and γ_2 goes along the equator by $\pi/2$ then goes along longitude.
Under the metric induced by embedding in $\mathbb{R}^3$, the γ_1 and γ_2 are (almost) geodesics.
So by proposition 6.3, parallel transporting $v = (0, -1, 1) \in T_a\mathbb{S}^2 \subset T_a\mathbb{R}^3$ along γ_1 produces $(-1, -1, 0)$ at b, BUT γ_2 produces $(-1, 1, 0)$, which is different.

Intuitively, the flatter the surface is, the smaller the difference would be produced; so if we observe this locally around a single point, then such difference can measure the extent by which the surface curves; this idea leads to the *Riemannian curvature*.

6.6 Exponential & Completeness

[Convention]: In this subsection, we assume or use

1. M being connected.
2. Levi-civita connection for covariant derivative.

Geodesics brings us to the study of "shortest path". A good way to describe the "distance" between two points without embedding M into Euclidean space is by defining to be the length of the shortest path between them:

Definition 6.1. For $p, q \in M$, their **Riemannian distance** d_g is defined by

$$d_g(p,q) := \inf_{\substack{\text{path } \gamma \\ \text{from } p \text{ to } q}} \int_0^1 ||\gamma'(t)|| \, \mathrm{d}t$$

And we can show that this is indeed a metric:

Proposition 6.4. (M, d_g) forms a metric space, and $(M, d_g) \cong M$.

The first part is trivial, and the second part is put in Question & Answer.
For a metric space, we naturally care about its completeness; unfortunately not all (M, d_g) are complete —— e.g. $(\mathbb{R}^2 \setminus \{0\}, ||.||)$. Since geodesics are shortest paths, we can conjecture that: (M, d_g) is complete $\iff$ For any two points, there always exists a geodesics connecting them.
This idea leads to another completeness, *Geodesically complete.*

A good way to discuss the existence of geodesics is by talking about when the *Exponential map* is defined. We will briefly talk about the interesting reason why there is an *exponential* in its name later.

Definition 6.8. Given $(p, v) \in M \times T_pM$, and geodesic γ with $\gamma(0) = p, \gamma'(0) = v$, the **exponential map** is defined by $\exp_p(v) := \gamma(1)$.

Remark. We didn't say how the geodesic is parameterized —— this is because $||\dot{\gamma}||$ is constant along the entire geodesics: $\nabla_{\dot{\gamma}}||\dot{\gamma}||^2 = \mathcal{C}^2(\nabla_{\dot{\gamma}}(\dot{\gamma} \otimes \dot{\gamma} \otimes g)) = 0$; therefore $\dot{\gamma}(0)$ already determines the "speed" of the entire γ.

We can deduce that $\exp_p$ is locally a diffeomorphism at 0. This is easy to proof: Given any ODE $\dot{x} = f(x)$ of $x : \mathbb{R} \to \mathbb{R}^n$ where $f \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$, the solution is smoothly dependent on $x(0)$ and f's derivatives. So inside a local chart of p, by letting $x = \dot{\gamma}$ and plugging Γ^i_{jk} into the geodesic equation (6.6) to produce f, the $\gamma(1)$ should be smoothly dependent on $\gamma'(0)$. And since $(\exp_p)_* = \text{id.}$, it is locally a diffeomorphism.

But there are two issues:

1. It is NOT always defined; for example, in $(\mathbb{R}^2 \setminus \{0\}, ||.||)$, the $\exp_{(1,0)}(-1,0)$ is undefined.
2. It could be NOT continuous; for example, consider the south pole p on the unit sphere surface $\mathbb{S}^2$, $\exp_p$ is singular at $\{v : ||v|| = \pi\}$.

Let's consider the 1$^{\text{st.}}$ one first.

Definition 6.2. We say M to be **geodesically complete** if there exists Riemannian metric g such that $\exp_p(v)$ is defined for $\forall v \in T_pM$ at any $p \in M$.

This definition is actually independent of the choice of g, you'll soon know why.
Also as we mentioned, there is a more intuitive way to define geodesically completeness:

Proposition 6.5. The M is geodesically complete if and only if for $\forall p, q \in M$, there exists a geodesics connecting them.

The proof is quite technical; I put it in question at the end of the section.

Remark. In our naïve physical intuition, this conclusion suggests that "M is geodesically complete" means if you are standing inside M, then for any target inside M, you can fire a bullet at a certain direction to hit it.
In other words, M is geodesically complete iff $\exp_p$ is surjective.

Interestingly, the Hopf-Rinow theorem points out that these two completeness are indeed the same, and gives another equivalent statement to make justification easier.

Hopf-Rinow Theorem

Theorem 6.10. There 3 statements are equivalent:

1. (M, d_g) is complete
2. M is geodesically complete
3. Closed & bounded sets in (M, d_g) are compact

This theorem is nice, because it is so intuitive and meaningful yet easy to prove.

Proof. (1) $\implies$ (2): Fix $p \in M$, and suppose $\exp_p$ is defined on $E \subset T_pM$.
Since M is locally Euclidean, E at least contains a neighbour of $0 \in T_pM$. So let's consider the set $E_v := \mathbb{R}v \cap E$ for some $v \in E \setminus \{0\}$.
Easy to see that E_v should be connected and open, so $E_v = (a, b)v$ for some $a, b \in \mathbb{R}$. But we also claim that E_v is closed: Consider any sequence $\{b_k\} \subset (a, b)$ such that $\lim_{k\to\infty} b_k = b$;

then by completeness of (M, d_g), we have that

$$\exp_p(bv) = \lim_{k \to \infty} \exp_p(b_k v) \in M$$

So a, b cannot be finite, which means $E_v = \mathbb{R}v$, and $E = T_p M$.

(2) $\implies$ (3): Consider a closed and bounded $B \subset M$ and $p \in B$. Let $r := \sup_{x \in B} d_g(x, p)$. Since geodesical completeness implies that $\exp_p$ is surjective, we have that

$$B \subseteq \exp_p \{v : ||v|| \leq r\}$$

Note that the RHS is compact, so as a closed subset, B is compact.

(3) $\implies$ (1): Fix an arbitrary cauchy sequence $\{p_k\}$ in (M, d_g), which must be bounded; so we pick a bounded subset $B \subset M$ that contains all $\{p_k\}$. Because $\overline{B}$ is compact, and thus sequentially compact, there exists a subsequence of $\{p_k\}$ that converges inside $\overline{B} \subset M$; therefore $\{p_k\}$ converges inside M. $\square$

This demonstrates a good example of the similarity between (M, d_g) and $\mathbb{R}^n$ as a metric space. It also shows how "natural" the Riemann distance is. And,

Corollary 6.1. A closed submanifold of a complete M is complete.

A direct corollary is that, given $f \in C^\infty(\mathbb{R}^n)$ such that 0 is a regular value, then $f^{-1}(0)$ is a complete Riemannian manifold.

As the issue of when $\exp_p$ is well-defined solved, we would like to know what benefit it brings. We will start with a problem of $\exp_p$, that it could be *not smooth*.
Assume that M is geodesically complete.

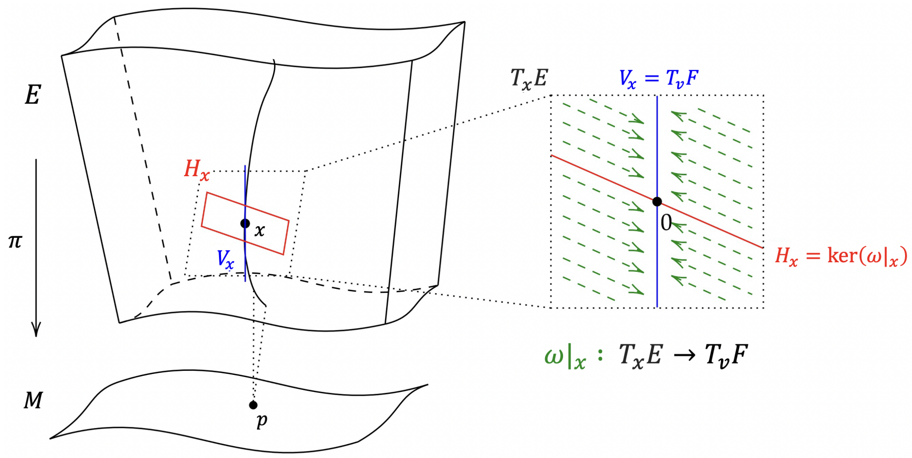
The $\exp_p$, one would expect that the singular point of it happens when it is non-injective

6.7 Connections in vector bundles

Let (E, M, π, F) be a smooth vector bundle over M.
The reason that defining “how a vector field changes with respect to another” was so difficult is that it's not a clear question in the case of $E = TM$, while it is actually quite clear as long as we consider a general case.
Consider a section $V \in \Gamma(E)$ and point $x = (p, v) \in E$. We can define that, if we are to move x on a given direction $u \in T_xE$, then a connection $\boldsymbol{\omega}$ is the projection of change of $V(x)$ on to $\pi^{-1}(p) \cong F$. In other words, it's a smooth map $\boldsymbol{\omega} : \Gamma(E) \to \Gamma(T^*E) \otimes \Gamma(TE)$,

$$\boldsymbol{\omega} : x \longmapsto (\text{projection map } T_xE \twoheadrightarrow T_vF \subset T_xE)$$

A linear projection map is fully determined by its kernel, which is $\ker(\boldsymbol{\omega}|_x) \subset T_xE$ in our case, and it satisfies $T_xE = \ker(\boldsymbol{\omega}|_x) \oplus T_vF$. Hence, choosing a connection is equivalent to choosing such kernel.



This interpretation explains the underlying confusion we stated at beginning, as the covariant derivative actually uses twice of F: $\nabla_V F = \boldsymbol{\omega}|_F(F, V)$ for $(F, V) \in T_xE$, where T_xE is given with the decomposition $T_xE = T_pM \oplus T_pM$.
This also tells that, from a practical point of view, our definition for $\boldsymbol{\omega}$ is redundant in defining a covariant derivative[11], so for $E = TM$, we can simplify $\boldsymbol{\omega}$ to a linear map $\Gamma(E) \to \Gamma(T^*M) \otimes \Gamma(E)$. So we define:

Definition 6.3. A **connection** is a linear map $\mathrm{d}_E : \Gamma(E) \to \Gamma(T^*M) \otimes \Gamma(E)$ that satisfies

$$\mathrm{d}_E(f \cdot \xi) = \mathrm{d}f \otimes \xi + f \cdot \mathrm{d}_E(\xi)$$

for all $f \in C^\infty(M)$ and $\xi \in \Gamma(E)$.

So by setting

[11] For example, in $\boldsymbol{\omega} : \Gamma(E) \to \Gamma(T^*E) \otimes \Gamma(TE)$, we're only interested in the part $\Gamma(T^*M)$ of $\Gamma(T^*E)$, and the fibre part of $\Gamma(TE)$, which could simply be given by $\Gamma(E)$ since $TF \cong F \cong \mathbb{R}^m$

$$\mathrm{d}_E(\xi) = \sum_{j=1}^{m} \mathrm{d}x_j \otimes \nabla_j \xi$$

we get a valid covariant derivative ∇, making our theory consistent as before.
Note that by writing ξ under a basis, the information of d_E is encoded inside an m by m matrix, which we call the *connection form*.

Definition 6.9. Let $\{\xi_j, \cdots, \xi_m\}$ be local basis of fibre in E, then there exists differential 1-forms $\{\omega_{ij}\}_{i,j=1,\cdots,m} \subset \Omega^1(M)$ such that

$$\mathrm{d}_E(\xi_j) = \sum_{i=1}^{m} \omega_{ij} \otimes \xi_i$$

this matrix of 1-forms $\boldsymbol{\omega} = (\omega_{ij})$ is called the **Connection form**.

For our original case of $E = TM$, we have $\omega_{jk} = \sum_{k=1}^{n} \Gamma_{jk}^{i} \mathrm{d}x_i$.
The name is misleading, as $\boldsymbol{\omega}$ isn't a form but a matrix of forms. But then why do we call it a "form"? This is because: Given a principal G-bundle on M, and $\mathfrak{g} = T_{\mathfrak{i}} M$ being the Lie algebra of G, connection form $\boldsymbol{\omega}$ is a $\mathfrak{g}$-valued 1-form.
This brings us to studying our theory of the special case of principal bundles.

Consider a smooth principal bundle (E, M, π, G), where Lie group G has identity $\mathfrak{i}$ and Lie algebra $\mathfrak{g}$. Since we haven't gone through enough knowledge in Lie groups, the discussion below will be mostly about intuition rather than derivation.

We can still start with $\boldsymbol{\omega} : \Gamma(E) \to \Gamma(T^*E) \otimes \Gamma(TE)$.
Fix a point $x = (p, g) \in E$ and $v \in T_x E$. As we mentioned, the $\boldsymbol{\omega}|_x$ is a linear projection $T_x E \twoheadrightarrow T_g G$; so for any smooth path $\widetilde{\gamma}$ such that $\widetilde{\gamma}(0) = x$ and $\widetilde{\gamma}'(0) = v$, the $\boldsymbol{\omega}$ can locally project it on a path γ on $\pi^{-1}\{p\}$, so that

$$\boldsymbol{\omega}|_x (v) = \boldsymbol{\omega}|_x (\gamma'(0)) = \gamma'(0)$$

Notice that $T_g G \cong \mathfrak{g} = T_{\mathfrak{i}} G$ for $\forall g \in G$, so the $\boldsymbol{\omega}$ could be simplified as a smooth map $\boldsymbol{\omega} : \Gamma(E) \to \Gamma(T^*E \otimes \mathfrak{g})$. To do this, notice that we can write $\gamma(t) = g \cdot \alpha(t)$ for some smooth path α starting from $\mathfrak{i}$; hence $\gamma'(0) = g \cdot \alpha'(0)$ and

$$\boldsymbol{\omega}|_x (v) = \alpha'(0) \in \mathfrak{g}$$

however, the $\mathfrak{g}$ embeds differently for different $g \in G$. So we will need to make sure that $\boldsymbol{\omega}$ is compatible with the pushforward of $\mathfrak{g}$ to different points. For arbitrary $h \in G$,

$$\begin{aligned}\boldsymbol{\omega}|_{xh} (v \cdot h) = \boldsymbol{\omega}|_{xh} (\gamma'(0) \cdot h) &= \boldsymbol{\omega}|_{xh} (g \cdot \alpha'(0) \cdot h) \\ &= \boldsymbol{\omega}|_{xh} \left(g \cdot h \cdot h^{-1} \cdot \alpha'(0) \cdot h\right)\end{aligned}$$

$$= h^{-1} \cdot \alpha'(0) \cdot h$$

So there are two requirements $\boldsymbol{\omega}$: The first one is that $\boldsymbol{\omega}|_x$ is a linear projection as mentioned; the second one is the equation above. Write them formally, we have:

Definition 6.10. Consider a smooth principal G-bundle E on M. A **connection form** is a $\omega \in C^\infty(E, T^*E \otimes \mathfrak{g})$ that satisfies

1. For $\forall x = (p, g) \in E$ and $v \in T_gG \subset T_xE$, we have $\boldsymbol{\omega}|_x(v) = v$.
2. For $\forall h \in G$, we have

$$R_h^* \boldsymbol{\omega} = \mathrm{Ad}(h^{-1}) \circ \boldsymbol{\omega}$$

where R_h^* is the pullback by $R_h : x \mapsto xh$, and $\mathrm{Ad}(h^{-1}) : \mathfrak{g} \to \mathfrak{g}$ is the pushforward of map $\mathrm{ad}_{h^{-1}} : G \to G$ defined by $x \mapsto h^{-1}xh$.

The simplest example would be the Maurer-Cartan connection, the canonical connection form for trivial principal bundle $E = M \times G$. The **Maurer-Cartan connection** ω_{mc} is a connection form ω_{mc} given by

$$\omega_{\mathrm{mc}}|_{(p,g)} := (L_{g^{-1}})^* \circ \pi$$

where $L_{g^{-1}} : E \to E$ sends $x \mapsto g^{-1}x$ and $\pi : E \to M$ is the canonical projection $(p, g) \mapsto g$. And when $G \leq \mathfrak{gl}_m$, we have that

$$\omega_{\mathrm{mc}}|_{(p,g)} = g^{-1} \cdot (R_g)_*$$

Actually, a connection form on a principal bundle is always locally Maurer-Cartan form by an appropriate choice of local trivializations.

So if we take an explicit construction of a smooth principal bundle, we can always get a connection form by taking the Maurer-Cartan form locally. Take a well-known example for non-trivial bundle:

Example 6.11. Re-call the Hopf bundle $U(1) \to \mathbb{S}^3 \xrightarrow{\pi} \mathbb{S}^2$, where $\mathbb{S}^3$ is embedded as $\{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ with

$$\pi(z_1, z_2) = \big(2 \cdot \Re(z_1\overline{z_2})\ ,\ 2 \cdot \Im(z_1\overline{z_2})\ ,\ |z_1|^2 - |z_2|^2)\big)$$

and $s \in U(1)$ acting on $(z_1, z_2) \in \mathbb{S}^3$ by $()s$

6.8 Questions

In these questions, we assume M to be a n-dim smooth manifold.

1. Given Riemannian metrics g and h on M,
(i) Show that (M, d_g) has the same topology as M.
(ii) Show that (M, d_g) always has the same topology as (M, d_h).
(iii) Show that geodesically completeness is independent of the choice of g.

2. Prove the Jacobian identity

$$[\nabla_a, [\nabla_b, \nabla_c]] + [\nabla_b, [\nabla_c, \nabla_a]] + [\nabla_c, [\nabla_a, \nabla_b]] = 0$$

3. Show that M is geodesically complete if and only if For $\forall p, q \in M$, there exists a geodesics connecting them.

6.9 Answers

Question 3 The "if" part is trivial, let's see "only if": Fix two points $p, q \in M$ with $L := d_g(p,q)$, we wish to construct a geodesic from p to q.
By smoothness, there exists $r > 0$ such that $\exp_p$ is homoemorphism on $\{v : ||v|| \leq r\} \subset T_pM$, which means that for each point on $S_r(p) := \{b \in M : d_g(b,p) = r\}$, there exists a geodesic connecting with p.
Since d_g is continuous and B_r is compact, there exists $b \in S_r(p)$ such that

$$d_g(b,q) = \min_{x \in S_r(p)} d_g(x,q)$$

and $\gamma : [0, L] \to M$ the geodesics starting from p such that $\gamma(r) = b$, we wish that $\gamma(L) = q$.
By triangle inequality, $d_g(b,q) + r \geq L$; but notice that any path between p and q need to pass through B_r, so

$$d_g(b,q) + r = L$$

as b minimizes distance to q. Similarly, we can deduce that $d_g(\gamma(t), q) + t = L$ for $\forall t \in [0, r]$.
So let's denote

$$T := \{t \in [0, L] : d_g(\gamma(t), q) + t = L\}$$

By smoothness of $d_g(\gamma(t), q) + t$, the T is closed. However, T is also open:
Given any $s \in T$, there $\exists \delta > 0$ and $S_\delta(\gamma(s))$ where every point has a geodesic connecting it with $\gamma(s)$. Similarly, we can find $b' \in S_\delta(\gamma(s))$ that minimizes its distance with q; so

$$d_g(b', q) + s + \delta = L$$

which means $d_g(p, b') \geq s + \delta$; but since there is already a path of length $(s + \delta)$, the γ should pass through b', therefore $s + \delta \in T$
Hence $T = [0, L]$ as it's both open and closed in $[0, L]$; the γ is thus a geodesic from p to q.

7 Some curvatures

An interesting fact we had in the last section is that parallel transports on different paths are different; interestingly this gives a possible way to compute the concept of curvature.
But before all these, we will introduce some efforts that Gauss made on curvature before the **Riemann Curvature Tensor** is introduced.

[**Convention**]**:** In this section,

1. Covaiant derivative ∇ is assumed to be Levi-Cvita.
2. "$\cdot$" is assumed to be the dot product.

Different from what we are doing in the previous section, mathematicians computed geometries about (hyper-)surfaces by embedding them in a Euclidean space before Riemann geometry. This is no doubt a natural way, but it turns out that some quantities are independent of how we describe our surfaces, and the Theorema egregium is a perfect example.
The **Theorema egregium** of Gauss stated that *Gaussian curvature of a surface can be expressed solely in terms of the first fundamental form and its derivatives.* So let's first look at what are the two fundamental forms.

7.1 Curvatures on Surface

Consider a 2-dimensional compact smooth surface $\mathcal{M}$ with a finite open cover so that each open set in it can be expressed as a parameterized surface; let's pick one of them, say

$$M := \left\{\mathbf{r}(u,v) : (u,v) \in U \subset \mathbb{R}^2\right\}$$

We denote the tangent plane at a point $p = (u,v)$ as the T_pM in this subsection; T_pM in this case is a plane spanned by $\{\mathbf{r}_u, \mathbf{r}_v\}$, where $\mathbf{r}_u = \partial\mathbf{r}/\partial u$.

Definition 7.1. Let $E = \mathbf{r}_u \cdot \mathbf{r}_u$, $F = \mathbf{r}_u \cdot \mathbf{r}_v$ and $G = \mathbf{r}_v \cdot \mathbf{r}_v$, consider point $p = (u,v)$ on M; The First fundamental form "**I**" is the quadratic form

$$\mathbf{I}(x,y) = x^T \underbrace{\begin{bmatrix} E & F \\ F & G \end{bmatrix}}_{g} y$$

Now consider $\widehat{n} = (\mathbf{r}_u \times \mathbf{r}_v)/|\mathbf{r}_u \times \mathbf{r}_v|$, let $J = \widehat{n} \cdot \mathbf{r}_{uu}$, $K = \widehat{n} \cdot \mathbf{r}_{uv}$ and $L = \widehat{n} \cdot \mathbf{r}_{vv}$ where $\mathbf{r}_u u = \partial^2\mathbf{r}/\partial u^2$. The Second Fundamental form "**II**" is the quadratic form

$$\mathbf{II}(x,y) = x^T \underbrace{\begin{bmatrix} J & K \\ K & L \end{bmatrix}}_{S} y$$

both are expressed under basis $\{\mathbf{r}_u, \mathbf{r}_v\}$.

The reason we use g to represent the matrix of $\mathbf{I}$ is that, the $\mathbf{I}$ is simply the induced Riemannian metric on TM by its embedding in $\mathbb{R}^3$. But a Riemannian metric can exist without embedding; in other words, $\mathbf{I}$ is an *Intrinsic* quantity,.
For the $\mathbf{II}$, its matrix S can be thought of as the x^2 term of Taylor expansion of M at p; thus it describes how the surface "curves" *in the way we embed it*, as its computation uses the normal vector. This suggests that $\mathbf{II}$ is an *Extrinsic* quantity.

The $\mathbf{II}$ describes the curving of the surface intuitively: the *Shape operator*, a simple formulation that tells *How the normal vector changes* when moving on the surface.

Definition 7.2. Consider $\widehat{n} : U \to \mathbb{R}^3$, where $\widehat{n}(u,v)$ gives unit normal vector of M at $\mathbf{r}(u,v)$. The **shape operator** $\mathrm{Sh}_p : T_pM \to T_pM$ at p is defined as

$$\mathrm{Sh}_p(v) = -d\widehat{n} \circ d\mathbf{r}^{-1}(v)$$

where $d\widehat{n}$ and $d\mathbf{r}$ are the differential of $\widehat{n}$ and $\mathbf{r}$ respectively.

We can try to write it in form of a matrix.

Notice that $\widehat{n} \cdot \mathbf{r}_u = \widehat{n} \cdot \mathbf{r}_v = 0$, which by Leibniz rule, gives that

$$J = (\widehat{n} \cdot \mathbf{r}_{uu}) = \mathrm{Sh}_p(\mathbf{r}_u) \cdot \mathbf{r}_u \quad ; \quad K = (\widehat{n} \cdot \mathbf{r}_{uv}) = \mathrm{Sh}_p(\mathbf{r}_v) \cdot \mathbf{r}_u$$
$$K = (\widehat{n} \cdot \mathbf{r}_{vu}) = \mathrm{Sh}_p(\mathbf{r}_u) \cdot \mathbf{r}_v \quad ; \quad L = (\widehat{n} \cdot \mathbf{r}_{vv}) = \mathrm{Sh}_p(\mathbf{r}_v) \cdot \mathbf{r}_v$$

thus if the basis $\{\mathbf{r}_u, \mathbf{r}_v\}$ is orthonormal, then $\mathrm{Sh}_p = S$; but in general, we have

$$\boxed{\mathrm{Sh}_p(v) = Sg^{-1}v}$$

For convenience, we denote the (i,j) component of Sh_p as s_{ij}. A good thing to be noticed about Sh_p is that it is **independent of parameterization**; this means that

Proposition 7.1. Given two smooth parameterizations $\mathbf{x} : U \to N_p$ and $\mathbf{y} : U' \to N_p$, with $\phi : C^\infty(U, U')$ such that $\mathbf{x} \circ \phi = \mathbf{y}$.
If we express Sh_p as matrices $\tilde{H}$ and H under basis $\{\mathbf{x}_u, \mathbf{x}_v\}$ and $\{\mathbf{y}_u, \mathbf{y}_v\}$ respectively, then $H = \tilde{H}$ numerically.

This can be proved easily by noticing that by the Chain rule, we have $\tilde{H} = d\widehat{n}_{\mathbf{y}} \circ d\mathbf{y}^{-1} = (d\widehat{n}_{\mathbf{x}} \circ d\phi) \circ (d\mathbf{x} \circ d\phi)^{-1} = d\widehat{n}_{\mathbf{x}} \circ d\phi \circ d\phi^{-1} \circ d\mathbf{x}^{-1} = H$; there is a change of basis at the second equal sign, and $\widehat{n}_{\mathbf{x}}$ & $\widehat{n}_{\mathbf{y}}$ are $\widehat{n}$ under parameterizations $\mathbf{x}$ & $\mathbf{y}$ respectively.

The shape operator helps to define the curvatures on a surface.
We know that the (signed) curvature of a parametric curve $\gamma : [0,1] \to \mathbb{R}^3$ is equal to $\widehat{n} \cdot \widehat{\dot{\gamma}}$.
So we can measure the curvature of a surface by taking paths on it.
But on a surface, different paths generate different curvatures, what we do is to describe the curvature of a surface as two components.

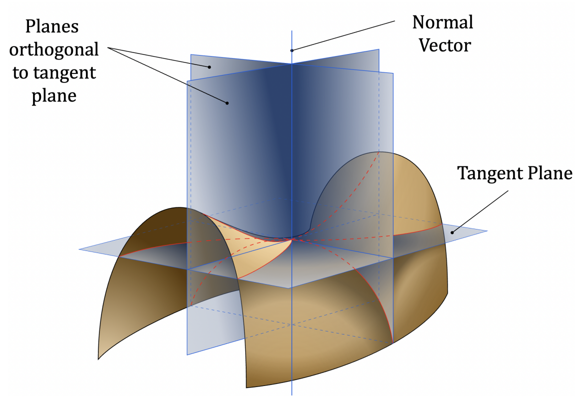


Figure 13: The principle curvatures are the max and min curvatures of the sections on those planes perpendicular to tangent plane.

Based on what we have now, we can construct such idea by **Principle curvatures** $\{\kappa_1, \kappa_2\}$.
Consider the planes intercept and perpendicular to tangent plane at p, and M has a section (which is a curve) on each of such planes, then *the minimum and maximum curvatures of those sections at* p are the Principle curvatures.
The shape operator gives a systematic way to compute κ_1 and κ_2:

Proposition 7.2. The principle curvatures at point $p \in M$ are the two eigenvalues of the shape operator Sh_p.

Proof. By the formula for curvatures of curves on a plane, the principle curvatures are the extreme value of the quadratic form $v^T \mathrm{Sh}_p v$ for $||v|| = 1$. Now since $\{v \in T_pM : ||v|| = 1\}$ is compact, we have that extreme value of $v^T \mathrm{Sh}_p v$ is taken when $w^T \mathrm{Sh}_p r = 0$ for $\forall ||w|| = 1$ with $w \perp v$, thus v must be eigenvectors of Sh_p. $\square$

The **Gauss curvature** κ is another way to measure the extent of curving; it can be thought of as a simple generalization of curvature of 1-dim curve by this interpretation:

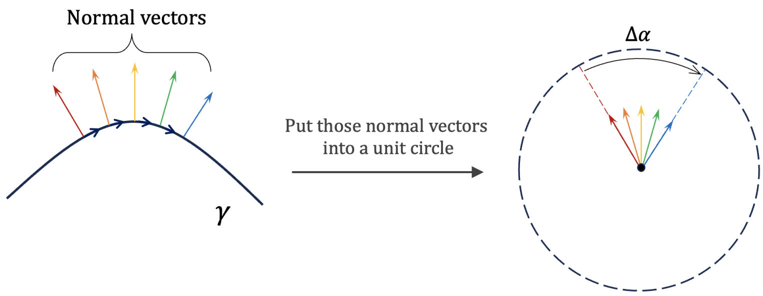


Figure 14: Imagine you move on the curved path by an infinitesimal distance $\triangle s$; if we record the unit normal vector along your journey, it swipes an arc on a unit circle. The (signed) length of this arc $\triangle\alpha$ over the length of your journey gives the curvature $\kappa \approx \triangle\alpha / \triangle s$.

Think of the Gauss curvature as an analogy; replace our infinitesimal 1-dim "journey" by a 2-dimensional region $\triangle S$, and replace the "arc" swiped by an "area" $\triangle A$ swiped; and the Gauss Curvature is the ratio $\triangle A / \triangle S$ when $\triangle S \to 0$.

We can easily find the formulation of this limit: Suppose $p = \mathbf{r}(u_0, v_0)$, and let $N = [u_0, u_0 + \varepsilon] \times [v_0, v_0 + \varepsilon]$; we have

$$\kappa := \lim_{\varepsilon \to 0} \frac{\text{signed area of } \widehat{n}(N)}{\text{area of } \mathbf{r}(N)} = \frac{\widehat{n} \cdot (\mathrm{Sh}_p(\mathbf{r}_u) \times \mathrm{Sh}_p(\mathbf{r}_v))}{\sqrt{\det(g)}}$$

By some vector algebra, we can compute in a easier way that

$$\begin{aligned}
\widehat{n} \cdot (\mathrm{Sh}_p(\mathbf{r}_u) \times \mathrm{Sh}_p(\mathbf{r}_v)) &= \frac{(\mathbf{r}_u \times \mathbf{r}_v) \cdot (\mathrm{Sh}_p(\mathbf{r}_u) \times \mathrm{Sh}_p(\mathbf{r}_v))}{||\mathbf{r}_u \times \mathbf{r}_v||} \\
&= \frac{(\mathbf{r}_u \cdot \mathrm{Sh}_p(\mathbf{r}_u))(\mathbf{r}_v \cdot \mathrm{Sh}_p(\mathbf{r}_v)) - (\mathbf{r}_u \cdot \mathrm{Sh}_p(\mathbf{r}_v))(\mathbf{r}_v \cdot \mathrm{Sh}_p(\mathbf{r}_u))}{||\mathbf{r}_u \times \mathbf{r}_v||} \\
&= \frac{JL - K^2}{\sqrt{\det(g)}} = \frac{\det(S)}{\sqrt{\det(g)}}
\end{aligned}$$

Hence we finally obtain $\kappa = \det(S)/\det(g)$, and can define:

Definition 7.3. The Gauss Curvature κ is defined as

$$\kappa = \det(\mathrm{Sh}_p) = \frac{\det(S)}{\det(g)} = \kappa_1 \kappa_2$$

Based on this definition, it seems that computation of κ depends on *how we embed* M (though normal vector can be expressed as $\mathbf{r}_u \times \mathbf{r}_v$, the cross product is undefined without embedding); however, it turns out that computation of κ does NOT depend on embedding. And this leads us to the Gauss' Theorema Egredium.

Remark. Based on this intuition for the definition of Gaussian curvature, we can intuitively say that, if M is a compact and convex subset with smooth boundary in $\mathbb{R}^3$, we have that

$$\int_{\partial M} \kappa \, \mathrm{d}A = 4\pi$$

and this turns out to be a special case of the *Gauss-Bonnet theorem*.

7.2 Gauss' Theorema egregium

(continue from previous subsection) And this property of "Independent of embedding" is described as the Theorema egregium (in Latin, means *an Excellent theorem*):

Theorema egregium

Theorem 7.1. The Gauss Curvature κ is invariant under local isometry.

Remark. It's interesting why such a conclusion could be "egregium" for someone like Gauss, so I put my answer at the **end** of this subsection.

We will prove this theorem in the way of calculus, and then translate things into words of manifolds, which gives the insight of generalization.

By Proposition 7.1, for every $x \in U$ we can always choose a neighbour N and a parameterization $\mathbf{r}$ such that $g = \operatorname{diag}(\lambda_1, \lambda_2)$ in N.
Under this parameterization, the set $\left\{\mathbf{r}_u/\sqrt{\lambda_1},\ \mathbf{r}_v/\sqrt{\lambda_2},\ \widehat{n}\right\}$ forms an orthonormal basis of $\mathbb{R}^3$; for convenience, let's denote them $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$, if we have a look at the rate of change of them, we have that

Lemma 7.2.

$$\begin{bmatrix}(\mathbf{f}_1)_u\\(\mathbf{f}_2)_u\\(\mathbf{f}_3)_u\end{bmatrix} = \begin{bmatrix} 0 & -(\lambda_1^{1/2})_v\lambda_2^{-1/2} & \lambda_1^{1/2}h_{11}\\ (\lambda_1^{1/2})_v\lambda_2^{-1/2} & 0 & \lambda_2^{1/2}h_{21}\\ -\lambda_1^{1/2}h_{11} & -\lambda_2^{1/2}h_{21} & 0\end{bmatrix}\begin{bmatrix}\mathbf{f}_1\\\mathbf{f}_2\\\mathbf{f}_3\end{bmatrix}$$
$$\begin{bmatrix}(\mathbf{f}_1)_v\\(\mathbf{f}_2)_v\\(\mathbf{f}_3)_v\end{bmatrix} = \begin{bmatrix} 0 & (\lambda_2^{1/2})_u\lambda_1^{-1/2} & \lambda_1^{1/2}h_{12}\\ -(\lambda_2^{1/2})_u\lambda_1^{-1/2} & 0 & \lambda_2^{1/2}h_{22}\\ -\lambda_1^{1/2}h_{12} & -\lambda_2^{1/2}h_{22} & 0\end{bmatrix}\begin{bmatrix}\mathbf{f}_1\\\mathbf{f}_2\\\mathbf{f}_3\end{bmatrix}$$

where h_{ij} is the (i,j) entry of Sh_p under basis $\{\mathbf{r}_u, \mathbf{r}_v\}$.

Proof. Here we only need to show the first equation, as the second one follows symmetrically.
First, the diagonal of the matrix must be all 0 since $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ are orthonormal;
Second, it is easy to calculate that the third row and column of it are $[-\lambda_1^{1/2}h_{11}, -\lambda_2^{1/2}h_{21}, 0]$ and $[\lambda_1^{1/2}h_{11}, \lambda_2^{1/2}h_{21}, 0]^T$ respectively.
So we only need to know the $(1,2)$ and $(2,1)$ entries, which are $(\mathbf{f}_1)_u \cdot \mathbf{f}_2$ and $(\mathbf{f}_2)_u \cdot \mathbf{f}_1$.
Notice that by $\mathbf{r}_u \cdot \mathbf{r}_v = \mathbf{r}_v \cdot \mathbf{r}_u = 0$,

$$(\mathbf{r}_u \cdot \mathbf{r}_v)_u = \mathbf{r}_{uu} \cdot \mathbf{r}_v + \frac{1}{2}(\lambda_1)_v = 0$$

this suggests that $(\mathbf{f}_1)_u \cdot \mathbf{f}_2 = -(\lambda_1^{1/2})_v\lambda_2^{-1/2}$ and $(\mathbf{f}_2)_u \cdot \mathbf{f}_1 = (\lambda_1^{1/2})_v\lambda_2^{-1/2}$.
Plug them in, we get the desired result for $[(\mathbf{f}_1)_u, (\mathbf{f}_2)_u, (\mathbf{f}_3)_u]^T$. $\square$

Now notice that

$$\begin{aligned}\lambda_1^{1/2}\lambda_2^{1/2}\cdot\det(\mathrm{Sh}_p) &= (\mathbf{f}_1)_u\cdot(\mathbf{f}_2)_v-(\mathbf{f}_1)_v\cdot(\mathbf{f}_2)_u\\ &= (\mathbf{f}_1\cdot(\mathbf{f}_2)_v)_u-(\mathbf{f}_1\cdot(\mathbf{f}_2)_u)_v\\ &= -\frac{\partial}{\partial u}\left((\lambda_2^{1/2})_u\lambda_1^{-1/2}\right)\end{aligned}$$

which gives $\det(\mathrm{Sh}_p) = -\lambda_1^{-1/2}\lambda_2^{-1/2}\frac{\partial}{\partial u}\left((\lambda_2^{1/2})_u\lambda_1^{-1/2}\right)$, suggesting that $\kappa = \det(\mathrm{Sh}_p)$ can be solely determined by λ_1 and λ_2. Combined with the fact that the shape operator is independent of parameterization, the Gaussian curvature is thus invariant under local isometry, $\mathbb{QED}$.

Remark. Now we know that Gauss curvature is solely determined by E, F, G (First Fundamental form), so can we really write down such an expression for Gauss Curvature? The answer is Yes, it can be calculated by the **Brioschi formula**:

$$\kappa = \frac{1}{\det(g)^2}\left(\det\begin{vmatrix}F_{uv}-\frac{1}{2}E_{vv}-\frac{1}{2}G_{uu} & \frac{1}{2}E_u & F_u-\frac{1}{2}E_v\\ F_v-\frac{1}{2}G_u & E & F\\ \frac{1}{2}G_v & F & G\end{vmatrix}-\det\begin{vmatrix}0 & \frac{1}{2}E_v & \frac{1}{2}G_u\\ \frac{1}{2}E_v & E & F\\ \frac{1}{2}G_u & F & G\end{vmatrix}\right)$$

So why was Gauss so happy to reach this conclusion?
Because Theorema egregium suggests that, though curvature seems like a higher dimensional property, a 2-dim creature can still calculate the curvature of its own space! This leads to the possibility that **human can know how the universe is curved with the data in our own 3-dim space**, as long as we can generalize the Theorema egreguim.

7.3 Formulation into manifolds

We developed last two parts using elementary calculus, which is simple but brings little insight about how to extent similar ideas into higher dimensions. So the purpose of this section is "translating" the last two parts; or more precisely, giving another proof of Theorema Egregium in words of **Levi-Civita connection**.
Still let's consider an open subset M of a 2-dim compact smooth surface $\mathcal{M}$, and $M := \{\mathbf{r}(u,v) : (u,v) \in U \subset \mathbb{R}^2\}$ such that $\operatorname{rank}(d\mathbf{r}) = 2$ everywhere.

Since the g is essentially a Riemannian metric, by the formula of Christoffel symbols for the Levi-Civita connection (Definition 6.6),

$$\begin{bmatrix} \Gamma^1_{jk} \\ \Gamma^2_{jk} \end{bmatrix} = \frac{1}{2} g^{-1} \begin{bmatrix} \partial_k g_{1j} - \partial_1 g_{jk} + \partial_j g_{1k} \\ \partial_k g_{2j} - \partial_2 g_{jk} + \partial_j g_{2k} \end{bmatrix}$$

which gives the following identities:

Proposition 7.3. (For symbolic consistency, we use notations $\Gamma^1_{11} := \Gamma^u_{uu}$, $\Gamma^1_{12} := \Gamma^u_{uv}$, $\Gamma^1_{21} := \Gamma^u_{vu}$ and so on.)

$$\begin{bmatrix} \Gamma^1_{11} \\ \Gamma^2_{11} \end{bmatrix} = g^{-1} \begin{bmatrix} \frac{1}{2}E_u \\ F_u - \frac{1}{2}E_v \end{bmatrix} = \frac{1}{2\det g} \begin{bmatrix} GE_u - 2FF_u + FE_v \\ -FE_u + 2EF_u - EE_v \end{bmatrix}$$
$$\begin{bmatrix} \Gamma^1_{12} \\ \Gamma^2_{12} \end{bmatrix} = g^{-1} \begin{bmatrix} \frac{1}{2}E_v \\ \frac{1}{2}G_u \end{bmatrix} = \frac{1}{2\det g} \begin{bmatrix} GE_v - FG_u \\ EG_u - FE_v \end{bmatrix}$$
$$\begin{bmatrix} \Gamma^1_{22} \\ \Gamma^2_{22} \end{bmatrix} = g^{-1} \begin{bmatrix} F_v - \frac{1}{2}G_u \\ \frac{1}{2}G_v \end{bmatrix} = \frac{1}{2\det g} \begin{bmatrix} -FG_v + 2GF_v - GG_u \\ EG_v - 2FF_v + FG_u \end{bmatrix}$$

which means we can prove the Theorema Egregium if it's possible to find an expression of κ in terms of the Christoffel symbols.

Now in order to get the shape operator, we focus on the quantities related to second derivatives. Recall that $(\mathbf{r}_{uu} \cdot \mathbf{r}_u) = \frac{1}{2}E_u$, $(\mathbf{r}_{uv} \cdot \mathbf{r}_v) = \frac{1}{2}G_u$, $(\mathbf{r}_{vv} \cdot \mathbf{r}_u) = F_v - \frac{1}{2}G_u$ and so on; by above proposition, these suggests that

$$\begin{aligned} \mathbf{r}_{uu} &= \Gamma^1_{11}\mathbf{r}_u + \Gamma^2_{11}\mathbf{r}_v + J\widehat{n} \\ \mathbf{r}_{uv} &= \Gamma^1_{12}\mathbf{r}_u + \Gamma^2_{12}\mathbf{r}_v + K\widehat{n} \\ \mathbf{r}_{vv} &= \Gamma^1_{22}\mathbf{r}_u + \Gamma^2_{22}\mathbf{r}_v + L\widehat{n} \end{aligned}$$

Notice that $(\mathbf{r}_{uu})_v = (\mathbf{r}_{uv})_u$, so we can differentiate two sides of the first two of above equations by v and u respectively, which gives us:

$$
\begin{aligned}
(\mathbf{r}_{uu})_v &= \begin{bmatrix} \partial_2\Gamma^1_{11} + \Gamma^1_{11}\Gamma^1_{12} + \Gamma^2_{11}\Gamma^1_{22} \\ \Gamma^1_{11}\Gamma^2_{12} + \partial_2\Gamma^2_{11} + \Gamma^2_{11}\Gamma^2_{22} \\ \Gamma^1_{11}K + \Gamma^2_{11}L + J_v \end{bmatrix} + J\cdot \mathrm{Sh}_p(\mathbf{r}_v) \\
= (\mathbf{r}_{uv})_u &= \begin{bmatrix} \partial_1\Gamma^1_{12} + \Gamma^1_{12}\Gamma^1_{11} + \Gamma^2_{12}\Gamma^1_{12} \\ \Gamma^1_{12}\Gamma^2_{11} + \partial_1\Gamma^2_{12} + \Gamma^2_{12}\Gamma^2_{12} \\ \Gamma^1_{12}J + \Gamma^2_{12}K + K_u \end{bmatrix} + K\cdot \mathrm{Sh}_p(\mathbf{r}_u)
\end{aligned}
$$

written in basis $\mathbb{R}^3 = \operatorname{span}\{\mathbf{r}_u, \mathbf{r}_v, \widehat{n}\}$.
Now if we compare the coefficients of $\mathbf{r}_u$ and $\mathbf{r}_v$ term of them respectively, we have that

$$
\begin{aligned}
Js_{12} - Ks_{11} &= \partial_1\Gamma^1_{12} - \partial_2\Gamma^1_{11} + \Gamma^1_{12}\Gamma^1_{11} - \Gamma^1_{11}\Gamma^1_{12} + \Gamma^2_{12}\Gamma^1_{12} - \Gamma^2_{11}\Gamma^1_{22} \\
Js_{22} - Ks_{21} &= \partial_1\Gamma^2_{12} - \partial_2\Gamma^2_{11} + \Gamma^1_{12}\Gamma^2_{11} - \Gamma^1_{11}\Gamma^2_{12} + \Gamma^2_{12}\Gamma^2_{12} - \Gamma^2_{11}\Gamma^2_{22}
\end{aligned}
$$

(where s_{ij} is the (i,j) entry of $\mathrm{Sh}_p = Sg^{-1}$ as before). Notice that

$$
\begin{aligned}
Js_{12} - Ks_{11} &= \frac{1}{\det(g)}\left(FLJ - FK^2\right) = F\kappa \\
Js_{22} - Ks_{21} &= \frac{1}{\det(g)}\left(EK^2 - JLE\right) = -E\kappa
\end{aligned}
$$

which gives two expressions of κ:

Proposition 7.4. The Gaussian Curvature κ can be written as

$$
\begin{aligned}
\kappa &= \frac{1}{F}\left(\partial_1\Gamma^1_{12} - \partial_2\Gamma^1_{11} + \Gamma^1_{12}\Gamma^1_{11} - \Gamma^1_{11}\Gamma^1_{12} + \Gamma^2_{12}\Gamma^1_{12} - \Gamma^2_{11}\Gamma^1_{22}\right) \\
&= \frac{-1}{E}\left(\partial_1\Gamma^2_{12} - \partial_2\Gamma^2_{11} + \Gamma^1_{12}\Gamma^2_{11} - \Gamma^1_{11}\Gamma^2_{12} + \Gamma^2_{12}\Gamma^2_{12} - \Gamma^2_{11}\Gamma^2_{22}\right)
\end{aligned}
$$

By symmetry, there would be one more, which is just our second expression after switching 1 & 2 and u & v, then change E to G.
Because the Christoffel symbols can be written in form of the $\mathbf{I}$ (by Proposition 7.3), the expressions in Proposition 7.4 give another proof of Theorema Egregium, $\mathbb{QED}$.

What's interesting about our expression of κ is that it can be written in another form (up to a scalar multiple). Notice that,

$$
\begin{aligned}
(\nabla_i\nabla_j - \nabla_j\nabla_i)\partial_k &= \nabla_i(\Gamma^p_{kj}\partial_p) - \nabla_j(\Gamma^q_{ki}\partial_q) \\
&= \partial_i(\Gamma^p_{kj})\partial_p - \partial_j(\Gamma^q_{ki})\partial_q + \Gamma^p_{kj}\Gamma^a_{pi}\partial_a - \Gamma^q_{ki}\Gamma^b_{qj}\partial_b \\
\implies ((\nabla_i\nabla_j - \nabla_j\nabla_i)\partial_k)_r &= \partial_i\Gamma^r_{kj} - \partial_j\Gamma^r_{ki} + \Gamma^p_{kj}\Gamma^r_{pi} - \Gamma^q_{ki}\Gamma^r_{qj}
\end{aligned}
$$

(in ESC by fixing i, j, k and summing over p, q, a, b.) And this is the same as the formula of κ ! (with $1/F$ or $-1/E$ removed)

What this implies is another geometric intuition of κ:

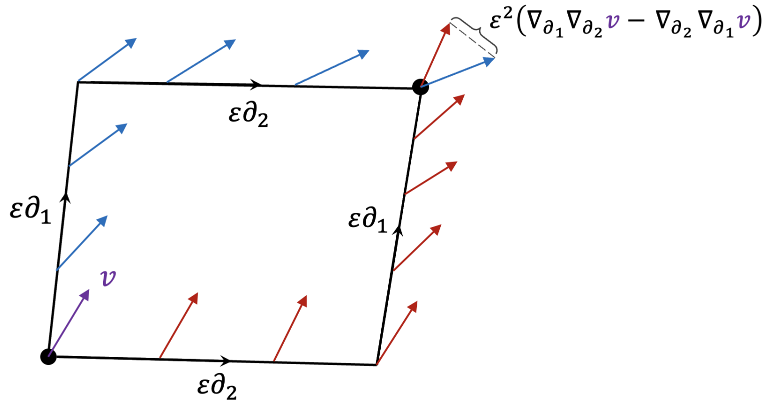


Figure 15: There is a difference between the results of parallel transport by *Change the* $1^{st.}$ *then* $2^{nd.}$ *component* and *Change the* $2^{nd.}$ *then* $1^{st.}$ *component*; by viewing this difference infinitesimally, it measures the Gaussian curvature.

Since $(\nabla_i\nabla_j - \nabla_j\nabla_i)\partial_k$ can be extended into higher dimension, this could be a generalization of the Gaussian curvature, the *Riemannian curvature*.

7.4 Riemann Curvature Tensor

Let M be a smooth n-dim manifold with Riemannian metric $g = \langle \cdot, \cdot \rangle$.
Just as how it is introduced in the last part, the Riemann curvature can be viewed as a high-dim generalization of the Gaussian curvature, with an intuitive geometric meaning:

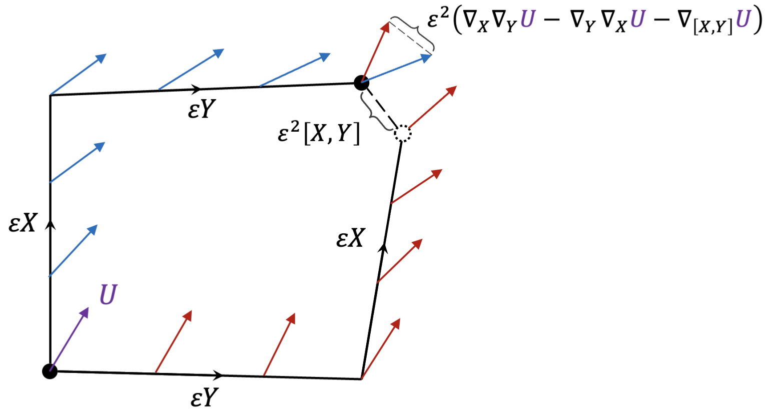


Figure 16: There is a difference between the results of parallel transport by *Change the* $i^{th.}$ *then* $j^{th.}$ *component* and *Change the* $j^{th.}$ *then* $i^{th.}$ *component*; by viewing this difference infinitesimally, it measures the Riemann curvature.

And thus formally,

Definition 7.4. Consider the Levi-civita connection ∇ on M, the Riemann Curvature $R : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ is

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

where $X, Y, Z \in \Gamma(TM)$.

We write it in form of components as R^a_{bcd}, where $R^a_{bcd}\partial_a = \mathrm{R}(\partial_c, \partial_d)\partial_b$ in ESC by summing over $a = 1, \ldots, n$; represent it in Christoffel symbols:

$$R^a_{bcd} = \partial_c \Gamma^a_{bd} - \partial_d \Gamma^a_{bc} + \Gamma^a_{jc}\Gamma^j_{bd} - \Gamma^a_{dj}\Gamma^j_{bc}$$

in ESC, summing j over $1, \cdots, n$.
It is sometimes more convenient to consider this tensor by lowering one index: $R_{abcd} = g_{jd}R^j_{abc}$ in ESC (summing over $j = 1, \cdots, n$); in other words, $R_{abcd} = \langle \mathrm{R}(\partial_b, \partial_c)\partial_a, \partial_d \rangle$.
For convenience, let's denote $R_{abcd} = \mathrm{Rm}(\partial_c, \partial_b, \partial_a, \partial_d)$, where

$$\mathrm{Rm}(X, Y, Z, W) := -\langle \mathrm{R}(X, Y)Z, W \rangle$$

4 important properties can be derived from the symmetry of this expression:

Proposition 7.5. For $\forall a, b, c, i, j = 1, \cdots, n$, we have

1. $R^i_{abc} = -R^i_{acb}$
2. $R_{abcd} = -R_{dbca}$
3. $R^i_{abc} + R^i_{bca} + R^i_{cab} = 0$
4. $R_{abcd} = R_{cdab}$
5. (Bianchi Identity) $\partial_a(R^i_{jbc}) + \partial_b(R^i_{jca}) + \partial_c(R^i_{jab}) = 0$

Proof. The $1^{\text{st.}}$ and $3^{\text{rd.}}$ statements are easy, just by observing the permutation of indices; for the $2^{\text{nd.}}$, note that $\mathrm{R}(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$ is skew-symmetric when viewed as a linear transform, therefore

$$R_{abcd} = \langle \mathrm{R}(\partial_c, \partial_d)\partial_b, \partial_a \rangle = -\langle \partial_b, \mathrm{R}(\partial_c, \partial_d)\partial_a \rangle = -R_{bacd}$$

The $4^{\text{th.}}$ identity follows directly from the first three, though it's bit lengthy to write down, left as an exercise.
And finally, the $5^{\text{th.}}$ identity: Notice that for fixed a, b, c,

$$\begin{aligned}
[\nabla_a, [\nabla_b, \nabla_c]]\partial_i &= \nabla_a\left([\nabla_b, \nabla_c]\partial_i\right) - [\nabla_b, \nabla_c]\left(\nabla_a \partial_i\right) \\
&= \nabla_a(R^i_{jbc}\partial_j) - [\nabla_b, \nabla_c](\Gamma^j_{ia}\partial_j) \\
&= (\nabla_a R^i_{jbc})\partial_j + \underbrace{R^i_{jbc}\nabla_a(\partial_j) - \Gamma^j_{ia} \cdot [\nabla_b, \nabla_c]\partial_j}_{=0} \qquad \text{(by Leibniz Law)}
\end{aligned}$$

(in ESC, summing over j) which, by applying the Jacobian identity, yields that

$$\begin{aligned}
0 &= [\nabla_a, [\nabla_b, \nabla_c]]\partial_i + [\nabla_b, [\nabla_c, \nabla_a]]\partial_i + [\nabla_c, [\nabla_a, \nabla_b]]\partial_i \\
&= (\nabla_a R^i_{jbc})\partial_j + (\nabla_b R^i_{jca})\partial_j + (\nabla_c R^i_{jab})\partial_j
\end{aligned}$$

so $\forall j$, we have $\nabla_a R^i_{jbc} + \nabla_b R^i_{jca} + \nabla_c R^i_{jab} = 0$. $\square$

How did such an identity (Bianchi identity) come out? Actually, the Bianchi identity has a straightforward form when we view the things in **Curvature form**, which we will discuss in the next few parts.

Also by those symmetries, we can see that

Corollary 7.1. There are $n^2(n^2-1)/12$ independent components in Riemann curvature tensor R_{abcd}.

Proof. It would be easier to consider the R_{abck} version of the tensor. (The result would be the same since $\det(g) \neq 0$.)
Notice that there are $\frac{1}{2}\binom{n}{2}\left(\binom{n}{2}+1\right)$ components independent from $1^{\text{st.}}$, $2^{\text{nd.}}$ and $4^{\text{th.}}$ symme-

tries, by choosing two pairs of indices on the ab and cd respectively. Among these, the 3[rd.] symmetry then eliminates $\binom{n}{4}$ of them, and so

$$\#\,(\text{independent components}) \leq \frac{1}{2}\binom{n}{2}\left(\binom{n}{2}+1\right) - \binom{n}{4} = \frac{n^2(n^2-1)}{12}$$

the $\leq$ can be turned to $=$ by considering specific cases; left as an exercise. $\square$

As a result, for $\dim(M)$, the R_{abcd} only have one component independent, while others are linearly dependent on that component. So each R_{abcd} should be "proportional" to the Gaussian curvature by our discussion in the last part.
Indeed, from Proposition 7.4, we know that

$$\begin{aligned}\kappa &= \frac{1}{g_{12}}\left(\partial_1\Gamma^1_{12} - \partial_2\Gamma^1_{11} + \Gamma^1_{12}\Gamma^1_{11} - \Gamma^1_{11}\Gamma^1_{12} + \Gamma^2_{12}\Gamma^1_{12} - \Gamma^2_{11}\Gamma^1_{22}\right)\\ &= \frac{1}{g_{12}}R^1_{112}\end{aligned}$$

also $\kappa = -R^2_{112}/g_{11}$ and $\kappa = -R^1_{221}/g_{22}$. Based on these, we have that

Example 7.3. Given $\dim(M) = 2$ and Gaussian curvature κ,

$$R_{abcd} = \left(g_{ac}g_{bd} - g_{ad}g_{bc}\right)\kappa$$

For example, a quite symmetrical one would be $\det(g)\kappa = R_{1212}$.

A more physical intuition of Riemann curvature is from the *geodesic deviation*. For convenience, let's assume M to be geodesically complete.
A **geodesic variation** is a map $\gamma \in C^\infty(\mathbb{R}\times\mathbb{R}, M)$ such that $\gamma(s,\cdot)$ is a geodesics for all $s \in \mathbb{R}$. The geodesics deviation says, with the presence of curvature, the "angle" between two adjacent geodesics (say $\gamma(s,\cdot)$ and $\gamma(s+\delta s,\cdot)$) will change over time; in other words, there exists a relation between $\partial(\partial\gamma/\partial s)/\partial t^2$ and the curvature tensor. Interestingly, this relation is elegant:

Proposition 7.6. Suppose $J(t) = \partial\gamma/\partial s|_{(0,t)}$ and $V(t) = \partial\gamma/\partial t|_{(0,t)}$,

$$\nabla_V\nabla_V J = \mathrm{R}(V,J)V$$

Proof. Inside a local coordinate basis of M, let J_k, V_k and γ_k be the $k^{\text{th.}}$ components of J, V and γ respectively. By definition,

$$\partial\gamma_k/\partial s|_{(0,t)} = J_k \quad\text{and}\quad \partial\gamma_k/\partial t|_{(0,t)} = V_k$$

Notice that $\partial^2\gamma/\partial s\partial t = \partial^2\gamma/\partial t\partial s$ in the coordinate, meaning that $[V, J] = 0$, so

$$\begin{aligned}\nabla_V J &= V(J) + \Gamma^i_{jk} V_k J_j \partial_i \\ &= J(V) + \Gamma^i_{jk} V_k J_j \partial_i = \nabla_J V\end{aligned}$$

(in ESC) Then we have

$$\begin{aligned}\nabla_V \nabla_V J = \nabla_V \nabla_J V &= [\nabla_V, \nabla_J]V + \nabla_J \nabla_V V \\ &= \mathrm{R}(V, J)V + \nabla_{[V,J]} V \\ &= \mathrm{R}(V, J)V\end{aligned}$$

where the first equality is by definition of geodesics, and third equality is by $[V, J] = 0$. $\square$

This provides a physical intuition for the Riemann curvature tensor: if two objects begin moving along initially parallel trajectories, the presence of a tidal gravitational force causes their paths to bend toward or away from each other. This bending results in a relative acceleration between the objects, precisely as described by geodesic deviation.

7.5 Taylor expansion and curvatures

For convenience, in this part we assume:

- M to be geodesically complete WLOG, as our conclusions could be easily modified to non-geodesically-complete cases.
- For any function $f \in C^\infty(\mathbb{R} \times \mathbb{R}, M)$, we denote $f' := f_* e_1$ and $\dot{f} := f_* e_2$.
- For path $\gamma \in C^\infty(M)$, the $\dot{\gamma}(t)$ denotes the pushforward of $\partial_1 \in T_t\mathbb{R}$ by γ.

The geodesic derivation (Proposition 7.6) provides an intuitive yet formal way of interpreting (a part of) Riemann curvature; now we take a closer look at that J in our equation. Let's call it the *Jacobi field*:

Definition 7.5. Given a geodesic $\gamma \in C^\infty(\mathbb{R}, M)$, smooth map $J : \mathbb{R} \to TM$ is a **Jacobi field** along γ if it satisfies

$$\nabla_{\dot{\gamma}(t)} \nabla_{\dot{\gamma}(t)} J = \mathrm{R}(\dot{\gamma}(t), J)\dot{\gamma}(t) \quad \text{for } \forall t \in \mathbb{R}$$

We can prove that Jacobi fields are given exactly by geodesic variation by reversing the proof of Proposition 7.6. In other words, for any Jacobi field J along γ, there exists a geodesic variation $f \in C^\infty(\mathbb{R} \times \mathbb{R}, M)$ such that $f(0, \cdot) = \gamma(\cdot)$ and $f'(0, t) = J(t)$ —— in this case we call J the Jacobi field given by f.
In the case of $J(0) = 0$, the geodesic variation who gives J is just a bunch of geodesics starting from $p := \gamma(0)$, and it takes a rather simple form: Consider Jacobi field J along γ given by geodesic variation $f \in C^\infty(\mathbb{R} \times \mathbb{R}, M)$, then

$$f(s,t) = \exp_{\gamma(0)} \left(t \left(\dot{\gamma}(0) + s \cdot J'(0)\right)\right) \tag{7.1}$$

This immediately gives that, by taking derivative,

$$J(t) = \left.\frac{\partial f}{\partial s}\right|_{(0,t)} = \left.(\exp_{\gamma(0)})_*\right|_{t\cdot\dot{\gamma}(0)} (t \cdot J'(0)) \tag{7.2}$$

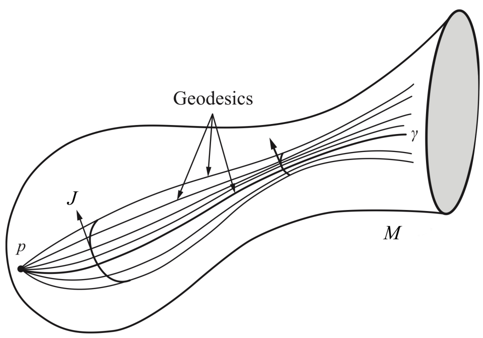


Figure 17: Jacobi fields with $J(0) = 0$; diagram from [1]

This expression (7.2) has a straight-forward meaning: The $J(t)$ is the partial derivative of $\exp_{\gamma(0)}(\mathbf{x})$ in the direction of $t \cdot J'(0)$ at $\mathbf{x} = t \cdot \dot{\gamma}(0)$.

Therefore, if we fix $p \in M$, the $\exp_p$ would give a local chart centered at p; then for $v \in \mathbb{R}^n$, the Jacobi field $J(t)$ along $\gamma(t) := \exp_p(t\mathbf{x})$ such that $J(0) = 0$ and $J'(0) = v$ gives the pushfoward of tv by $\exp_p$ at $\gamma(t)$.
This leads to the main thought of the Jacobi field: The $\exp_p$ gives a local coordinate, while the definition of the Jacobi field enable us to measure the curving of this coordinate in terms of R_{abcd}. This suggests that we can give a full description of the second derivative of g in terms of the Riemannian curvature by studying J:

Taylor expansion of g

Theorem 7.4. Under a local coordinate $(U, \exp_p^{-1})$, for $\mathbf{x} = (x_1, \ldots, x_n)$,

$$g_{ij}(\mathbf{x}) = \delta_{ij} - \frac{1}{3} R_{ikjr}(\mathbf{x}) x_k x_r + \mathcal{O}(||\mathbf{x}||^3)$$

(in ESC, summing over $k, r = 1, \ldots, n$.)

Proof. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of $\mathbb{R}^n$. Fix $\mathbf{x} \in \mathbb{R}^n$ and denote $J_k(t)$ as the Jacobi field along $\gamma(t) := \exp_p(t\mathbf{x})$ such that $J_k(0) = 0$ and $J_k'(0) = e_k$. By (7.2),

$$g_{ij}(t\mathbf{x}) = \frac{1}{t^2} \langle J_i(t), J_j(t) \rangle \tag{7.3}$$

Let's denote $f(t) := \langle J_i(t), J_j(t) \rangle$.
By (7.2), we have $J_k'(0) = e_k$; and by definition of Jacobi field, we also have $J_k''(0) = \mathrm{R}(V, J_k(0))V = 0$ since $J_k(0) = 0$. So by Leibniz rule, we have

$$\begin{aligned}
f(0) = f'(0) &= f^{(3)}(0) = 0 \\
f''(0) &= 2\langle J_i'(0), J_j'(0) \rangle = 2\delta_{ij} \\
f^{(4)}(0) &= 4\langle e_i, J_j^{(3)}(0) \rangle - 4\langle J_i^{(3)}(0), e_j \rangle
\end{aligned}$$

To deal with $J_k^{(3)}(t) = \nabla_{\dot\gamma(t)} \mathrm{R}(\dot\gamma(t), J_k(t))\dot\gamma(t)$, notice that when $t = 0$,

$$\begin{aligned}
\nabla_{\dot\gamma} \left(\dot\gamma_b \dot\gamma_c J_{kd} R^a_{bcd} \partial_a \right) &= R^a_{bcd} \gamma_b \gamma_c \gamma_d \partial_a + \underbrace{\text{multiples of } J_{kd}}_{=0} \\
&= \mathrm{R}(\dot\gamma, J_k)\dot\gamma
\end{aligned}$$

in ESC (summing over a, b, c, d), where J_{kd} is the $d^{\text{th.}}$ component of J_k.
Hence $f^{(4)}(0) = 8 \cdot \mathrm{Rm}(\dot\gamma, \partial_j, \dot\gamma, \partial_i)$. Then by applying the Maclaurin Series on (7.3), we get the desired result. □

So geometrically, the $\exp_p$ gives a more meaningful local coordinate, since geodesics define being "straight". We call a chart $(U, \exp_p^{-1})$ the **normal coordinate**, and U the **normal neighbourhood**.
And naturally, the Rm takes a simple form in this coordinate:

Corollary 7.2. Under a normal coordinate $(U, \exp_p^{-1})$,

$$R_{ikjr}\big|_p = \frac{1}{2}\left(\partial_i\partial_r g_{kj} + \partial_k\partial_j g_{ir} - \partial_i\partial_j g_{kr} - \partial_k\partial_r g_{ij}\right)(0)$$

However, these are NOT implying that the Rm can determine g, even locally: Consider normal coordinates $(U, \exp_x^{-1})$ and $(V, \exp_y^{-1})$ such that $U \cong V \cong \mathbb{R}^n$, then even if R_{abcd} take the same value inside both coordinates, U and V could be not isometric! For example, R_{abcd} has only one independent component in case of $n = 2$, yet the second derivatives of g_{ij} have 3 terms, so it would be absurd that R_{abcd} can dominate the change of g_{ij}.
But we can still make the idea work, but by using the fact that geodesics defines a unique parallel coordinate, rather than using the normal coordinate:
Consider n-dim smooth manifolds X, Y, points $(x, y) \in X \times Y$ and a linear isometry $\phi_x : T_xX \to T_yY$. Consider normal neighbourhoods $B_r(x) \subset X$ and $B_r(y) \subset Y$ and a bijection $\Phi : B_r(x) \to B_r(y)$ defined by

$$\Phi = \exp_y \circ \phi_x \circ \exp_x^{-1}$$

For $p \in B_r(x)$, construct $\phi_p : T_pX \to T_{\Phi(p)}Y$ by

$$\phi_p = \mathrm{prl}_\beta \circ \phi_x \circ \mathrm{prl}_\alpha^{-1}$$

where prl_γ denotes the parallel transport from $\gamma(0)$ to $\gamma(1)$, path α is geodesic from x to p, and β is geodesic from y to $\Phi(p)$.

Proposition 7.7. (Cartan's theorem) If for any $U, V, W \in \Gamma(TM)$,

$$\phi_p \circ \mathrm{R}_X(U, V)W = \mathrm{R}_Y(\phi_p(U), \phi_p(V))\phi_p(W)$$

then $\Phi_*\big|_p = \phi_p$ and Φ gives an isometry.

The proof is basically repeating our thinking.

Proof. Fix a basis $\{\partial_1, \ldots, \partial_n\}$ of T_xX, a point $p \in B_r(x)$, and non-zero $v \in T_pX$. Let α be geodesics from x to p, and β be geodesic from y to $\Phi(p)$.
Now for every $\tau \in [0, 1]$, we construct basis

$$\begin{aligned} T_{\alpha(\tau)}X &= \mathrm{span}\left\{\mathrm{prl}_{\alpha(t\cdot\tau)}\partial_1, \ldots, \mathrm{prl}_{\alpha(t\cdot\tau)}\partial_n\right\} \\ T_{\beta(\tau)}Y &= \mathrm{span}\left\{\mathrm{prl}_{\beta(t\cdot\tau)}\phi_x(\partial_1), \ldots, \mathrm{prl}_{\beta(t\cdot\tau)}\phi_x(\partial_n)\right\} \end{aligned}$$

Notice that $\Phi(\alpha(\tau)) = \beta(\tau)$, and the $\phi_{\alpha(\tau)}$ sends

$$\phi_{\alpha(\tau)} : \mathrm{prl}_{\alpha(\tau)}\partial_j \mapsto \mathrm{prl}_{\beta(\tau)}\phi(\partial_j)$$

for $j = 1, \ldots, n$; it is an isometry by definition of parallel transport.

Consider two Jacobi fields J_x and J_y along α and β respectively, such that

$$J_x(0) = x \ , \ J_x'(0) = \dot{\alpha}(0) \ ; \quad J_y(0) = y \ , \ J_y'(0) = \dot{\beta}(0)$$

Now note that under the basis we constructed for $T_{\alpha(\tau)}X$ and $T_{\beta(\tau)}Y$, the $J_x(0)$ & $J_y(0)$ and $J_x'(0)$ & $J_y'(0)$ all have the same coordinate. Also, by

$$\begin{aligned}\phi_{\alpha(\tau)} \circ J_x''(t) &= \phi_{\alpha(\tau)} \circ \mathrm{R}_X(U, V)W \\ &= \mathrm{R}_Y(\phi_{\alpha(\tau)}(U), \phi_{\alpha(\tau)}(V))\phi_{\alpha(\tau)}(W) = J_y''(t)\end{aligned}$$

so $J_x(t)$ and $J_y(t)$ are the same in respective basis of $T_{\alpha(t)}X$ and $T_{\beta(t)}Y$; and therefore $\Phi_*|_p = \phi_p$. Hence we have that Φ is an isometry, as ϕ_p is always a linear isometry. $\square$

From the Taylor expansion of g, there are some simpler ideas of curvatures that could be realized from simple expressions using Riemannian curvatures.

Definition 7.6. (Curvatures) Given $X, Y, Z \in \Gamma(TM)$,

- The Ricci curvature $\mathrm{Ric} : \Gamma(TM) \times \Gamma(TM) \to \mathbb{R}$ is defined as

$$\mathrm{Ric}(Y, Z) := \mathrm{tr}\left(X \mapsto \mathrm{R}(X, Y)Z\right)$$

  Its components are written as R_{ij}, and have the values $\mathrm{R}_{ij} = \sum_{k=1}^{n} R_{ikj}^{k}$.
- The Scalar curvature $\mathrm{Scal} : M \to \mathbb{R}$ is defined as

$$\mathrm{Scal}(p) := \mathrm{tr}_g \mathrm{Ric} := \sum_{j,k} g^{jk} \mathrm{R}_{jk}$$

- The Sectional Curvature $\mathrm{K} : \Gamma(TM) \times \Gamma(TM) \to \mathbb{R}$ is defined as

$$\mathrm{K}(X, Y) := \frac{\langle \mathrm{R}(X, Y)Y, X\rangle}{\langle X, X\rangle\langle Y, Y\rangle - \langle X, Y\rangle^2}$$

Their expressions are simple yet tells almost nothing, but those expressions can be derived from some straight-forward geometrical quantities; we will see one by one.

The Ricci curvature describes the local behavior of $\det(g)$.

Proposition 7.8. In the Riemann normal coordinate $\mathbf{x} = (x_1, \ldots, x_n)$, we have Taylor expansion

$$\sqrt{\det g} = 1 - \frac{1}{6}\mathrm{R}_{jk}(\mathbf{x})x^j x^k + \mathcal{O}(||\mathbf{x}||^3)$$

Proof. □

The Scalar curvature describes the local behavior of the volume of geodesics ball.

Proposition 7.9.

The Sectional curvature describes the local behavior of the perimeter of geodesics ball.

Proposition 7.10.

7.6 Curvature form

Let E be a rank-r smooth vector bundle over M.

Recall that we used the derivative $\mathrm{d}_E : \Gamma(E) \to \Gamma(T^*M) \otimes \Gamma(E)$ to define the connection on a vector bundle. In this chapter, the Riemannian curvature was demonstrated as the "*change of a connection*"; so we can try to define and generalize the curvature in terms of "$\mathrm{d}_E \circ \mathrm{d}_E$". But first we need to extend d_E; like that the exterior derivative d is an extension of its restriction on $C^\infty(M) = \Omega^0(M)$, we can obtain d_E from the generalizing exterior derivative to vector bundles.

Definition 7.1. Consider a smooth vector bundle (E, M, π), we define the set of differential k-form on E to be the smooth sections $\Omega^k(E) := \Omega^k(M) \otimes \Gamma(E)$, equipped with wedge product

$$\begin{aligned} \wedge : \Omega^j(M) \times \Omega^k(E) &\to \Omega^{j+k}(E) \\ (\omega, \eta \otimes v) &\mapsto (\omega \wedge \eta) \otimes v \end{aligned}$$

Define **exterior derivative** as an $\mathbb{R}$-linear map $\mathrm{d}_E : \Omega^k(E) \to \Omega^{k+1}(E)$ that satisfies

$$\mathrm{d}_E(\omega \otimes v) := \mathrm{d}\omega \otimes v + (-1)^k \omega \wedge \mathrm{d}_E(v)$$

and for $k = 1$, the d_E is the connection chosen.

Basically, the idea is to link the value of the exterior derivative with some information of the vector bundle constructed on M.

We already know that $\mathrm{d}_E : \Omega^0(E) \to \Omega^1(E)$ produces the connection, so as a differential operator, we would expect that applying it again would produce a curvature-related quantity.

Definition 7.7. Given a connection, the **curvature form** $F_\nabla : \Omega^0(E) \to \Omega^2(E)$ on E is given by $F_\nabla(\xi) := \nabla \circ \nabla(\xi)$.

We can write this explicitly:

$$\begin{aligned} F_\nabla(\xi) &= (\mathrm{d} + \boldsymbol{\omega}) \circ (\mathrm{d} + \boldsymbol{\omega})(\xi) \\ &= \boldsymbol{\omega} \circ \mathrm{d}(\xi) + \mathrm{d} \circ \boldsymbol{\omega}(\xi) + \boldsymbol{\omega} \circ \boldsymbol{\omega}(\xi) \\ &= (\mathrm{d} \circ \boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega})(\xi) \end{aligned}$$

7.7 Chern-Gauss-Bonnet Theorem

As we conjectured at the end of the introduction to Gaussian curvature, the integral of κ over a convex & compact 2-dim manifold is just the surface area of a 2-dim sphere (4π); but what about non-convex cases?
It turns out that we have a general theorem:

Proposition 7.11. (Gauss-Bonnet theorem) Suppose M is a compact 2-dim orientable Riemannian manifold with boundary. If k_g is the geodesic curvature on ∂M, we have that

$$\int_M \kappa \ \mathrm{d}A + \int_{\partial M} k_g \mathrm{d}s = 2\pi\chi(M)$$

There is a very pleasing theorem, though the proof is painful.

7.8 Questions

1. Given $\mathbb{S}^n$ embedded in $\mathbb{R}^{n+1}$ as a unit sphere, show that for $X, Y \in \Gamma(T\mathbb{R}^{n+1})$ such that $X|_{\mathbb{S}^n}, Y|_{\mathbb{S}^n}$, we have

$$\mathrm{Rm}(X, Y, Z, W) = \langle X, Z\rangle\langle Y, W\rangle - \langle Y, Z\rangle\langle X, W\rangle$$

at every $p \in \mathbb{S}^n \subset \mathbb{R}^{n+1}$.

7.9 Answers

Question 1 Using our conclusion from Example 6.8 which says $\nabla_V F = \overline{\nabla}_V F + \langle V, F\rangle \vec{p}$,

$$\begin{aligned}
\nabla_X \nabla_Y Z &= \overline{\nabla}_X \nabla_Y Z + \langle \nabla_Y Z, X\rangle \vec{p} \\
&= \overline{\nabla}_X \overline{\nabla}_Y Z + X\left(\langle Y, Z\rangle\right)\vec{p} + \langle Y, Z\rangle \overline{\nabla}_X \vec{p} + \langle \overline{\nabla}_Y Z, X\rangle \vec{p} \\
&= \overline{\nabla}_X \overline{\nabla}_Y Z + X\left(\langle Y, Z\rangle\right)\vec{p} + \langle Y, Z\rangle X + Y\left(\langle Z, X\rangle\right)\vec{p} - \langle \overline{\nabla}_Y X, Z\rangle \vec{p}.
\end{aligned}$$

then using the fact that $\overline{R}(X,Y)Z = [\overline{\nabla}_X, \overline{\nabla}_Y]Z - \overline{\nabla}_{[X,Y]}Z = 0$ (since Christoffel symbols in $\mathbb{R}^{n+1}$ are all zero), we have

$$\begin{aligned}
\mathrm{R}(X,Y)Z &= \overline{\nabla}_{[X,Y]}Z + \langle \overline{\nabla}_X Y, Z\rangle \vec{p} - \langle \overline{\nabla}_Y X, Z\rangle \vec{p} - \nabla_{[X,Y]}Z + \langle Y, Z\rangle X - \langle X, Z\rangle Y \\
&= \langle Y, Z\rangle X - \langle X, Z\rangle Y
\end{aligned}$$

Hence, $\mathrm{Rm}(X,Y,Z,W) = \langle X, Z\rangle\langle Y, W\rangle - \langle Y, Z\rangle\langle X, W\rangle$.

8 Physics trou normand

Let's have a rest to see some related physics.

This is a section about physics and reality, so instead of definitions, propositions, theorems, proofs and so on, we will add an element called **fact**s, which are usually contained in blue text boxes.

[Convention]: In this chapter, we assume:

1. M denotes a smooth n-dim connected manifold.

8.1 Some general relativity

S-T coordinate stands for Space-Time coordinate; it is a tuple of 4 real numbers (t, x, y, z) representing an event, where t is this event's time and x, y, z are its spacial coordinate.

As we know, the Galilei transformation —— which is how we describe movements in space & time turns out to be just a low-speed approximation of the **Lorentz transformation**, which is the fact below

Fact 8.1. Suppose in the S-T coordinate of an object B, an object $A = (t, x, y, z)$ started moving in constant velocity $\mathbf{v}$ in the X direction from the origin. If $B = (t', x', y', z')$ in the S-T coordinate of A, then we have

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \underbrace{\begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\Lambda} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

where $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta = v/c$.

This description is equivalent to the combination of 3 natural facts:

1, (Linear) Any object that moves in constant velocity relative to A must moves in constant velocity relative to B.
2, (Eigenvector) Any objects that travels at constant speed c has the same speed in every observer's S-T coordinate.
3, (Symmetry) For A, B moving in constant velocity, the trajectory of B in S-T coordinate of A is the same as the trajectory of A in S-T coordinate of B, but in opposite direction.

The derivation is left as an exercise.
In classical physics, we normally use time as infinitesimal for defining velocity, acceleration and so on, but now time is unreliable since the Lorentz transform changes the time scale. So instead, we use a quantity called **proper time** τ:

Definition 8.1. In a S-T coordinate, the **Minkowski metric** η is an inner product given by matrix $\operatorname{diag}(1, -1, -1, -1)$.
Consider an object with trajectory $q(t) = (ct, x(t), y(t), z(t))$.

1. The proper time of it is $\tau := \int \frac{1}{c}\sqrt{\eta(\dot{q}, \dot{q})}\ \mathrm{d}t$
2. The Four-velocity of it is $\mathbf{v} := \frac{\mathrm{d}q}{\mathrm{d}\tau}$.
3. The Four-acceleration of it is $\mathbf{a} := \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\tau}$.

Take two computational examples,

Example 8.2. ∘ For an object moving in constant velocity $\vec{u}$,

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma; \quad \mathbf{v} = (\gamma c, \gamma \vec{u})$$

∘ Suppose an object (initially at rest) now has a uniform Four-acceleration of $(ct, a, 0, 0)$ starting from 0, then its trajectory in S-T coordinate is

$$\begin{aligned} q(\tau) &= \left(\tau\sqrt{c^2 + a^2\tau^2/4}, \frac{a\tau^2}{2}, 0, 0\right) \\ &= \left(ct, \frac{c}{a}(\sqrt{c^2 + a^2t^2} - c), 0, 0\right) \end{aligned}$$

which is approximately the parabola $x = at^2/2$ at small t (which is what we would have in classical physics), and then asymptotically approaches $x = ct$.

The reason that we use Four-velocity & acceleration is that in General relativity, we regard *everything as an object moving in a 4-dim manifold*, (i.e even "static" objects are moving in Four-velocity $(c, 0, 0, 0)$) —— the time has NO difference from other coordinates.

So the most important property of τ (or to say Minkowski metric) is that it is invariant under the Lorentz transform, which means it can be a Riemannian metric;

Proposition 8.1. For two events with S-T coordinates r and q,

$$\eta(\Lambda(r), \Lambda(q)) = \eta(r, q)$$

Proof: Just by matrix multiplication. □

In summary, we can interpret:

$$\begin{aligned} \text{our universe with time} &\equiv \text{4-dim smooth manifold} \\ \text{S-T coordinate} &\equiv \text{tangent space} \\ \text{Minkowski metric} &\equiv \text{a Riemannian metric} \\ \text{Four velocity} &\equiv \text{tangent vector with length measured by } \eta \end{aligned}$$

Now we add an ingredient into our theory: Stress-momentum tensor T^{ij}. There are different interpretations of this tensor depending on the context, in our scenario,

Definition 8.1. Fix a given time in a S-T coordinate, let ρ be the energy density, $\vec{F}$ be the flux of energy, and $\vec{p}$ be the momentum density; we define

$$T^{00} = \rho \quad T^{0j} = \frac{1}{c}F_j \quad T^{j0} = cp_j$$

(and $T^{ij} = 0$ otherwise) to be the 4×4 components of **Stress-energy tensor**.

Be clear that the word of "energy" and "momentum" can be misleading —— since there is no general definition of them (though it is possible to put it in particular cases such as electromagnetic theory), we need to make some assumptions on them:

Fact 8.3. We assume that $\partial_j T^{ij} = 0$ (in ESC) for $\forall i$.

In other words, we are assuming the diffusion equation $\frac{\mathrm{d}}{\mathrm{d}t}\rho + \nabla\vec{F} = 0$ and conservation of momentum holds.

The Einstein field equation is not from reasoning but from an educated guess; and this guess is based on two fundamental facts:

1. The conservation of the Einstein tensor coincides with the conservation of T^{ij};
2. The geodesic derivation caused by Riemann curvature coincides with the effect of gravity.

The first one yields the guess that $G^{ij} = rT^{ij}$ for some $r \in \mathbb{R}$, while the second one gives the value of r —— which involves some calculations.

Einstein Field Equation

Theorem 8.4. For the Einstein tensor $G^{ij} = \mathrm{R}^{ij} - \frac{1}{2}\mathcal{R}g^{ij}$ and Stress-energy tensor T^{ij}, we (probably) have

$$G^{ij} = \frac{8\pi G}{c^4} T^{ij}$$

Proof. (Prove that it's reasonable to guess that the coefficient is $\frac{8\pi G}{c^4}$.)
Assume that the coefficient is $\mu \in \mathbb{R}$, that is,

$$\mathrm{R}^{ij} - \frac{1}{2}\mathcal{R}g^{ij} = \mu T^{ij} \tag{8.1}$$

In a local chart U with low and static gravitational field strength, consider a low-speed particle with trajectory $\gamma : (-1, 1) \to U$, we have that $\frac{\mathrm{d}\tau}{\mathrm{d}t} \approx 1$ and $\frac{\mathrm{d}\gamma_\bullet}{\mathrm{d}\tau} \approx (1, 0, 0, 0)$, which by the geodesic equation, yields that

$$\frac{\mathrm{d}^2\gamma_i}{\mathrm{d}t^2} \approx \frac{\mathrm{d}^2\gamma_i}{\mathrm{d}\tau^2} = -\Gamma^i_{00}$$

Now turn to Newtonian mechanics; suppose the only force applies on this particle is gravity imposed by static matters around, and $\rho(\vec{x})$ is the density of matters at $\vec{x}$, then

$$\phi(\vec{x}) := -G \int_{\vec{r}\in U} \frac{\rho(\vec{r})}{||\vec{x} - \vec{r}||}$$

where ϕ represents the gravitational potential. The acceleration of this particle then equals $-\nabla\phi$ which is also $\approx -(\Gamma^1_{00}, \Gamma^2_{00}, \Gamma^3_{00})$.

The key observation is that by Gauss' theorem,

$$4\pi G\rho = -\nabla\cdot\nabla\phi \approx \sum_{j=1}^{3}\partial_j\Gamma^j_{00}$$

The terms on the RHS can be seen in the explicit expression of R_{00}, so let's find it by equation (8.1) —— which is difficult, but here is a trick: Lowering the indices then contracting both sides; this gives that $\mathcal{R}-\frac{1}{2}\mathcal{R}\cdot 4 = -\mathcal{R} = \mu\mathcal{T}$, where $\mathcal{T} := C(T^i_j)$; hence

$$\mathrm{R}^{ij} = \mu\left(T^{ij} - \frac{1}{2}\mathcal{T}g^{ij}\right)$$

Note that since Christoffel symbols are infinitesimal quantities in weak gravitational field, we have that

$$\begin{aligned}\mathrm{R}_{00} = \sum_{k=0}^{3} R^k_{0k0} &= \sum_{k=0}^{3}\partial_k\Gamma^k_{00} - \partial_0\Gamma^k_{k0} + \underbrace{\sum_{j=0}^{3}\Gamma^k_{jk}\Gamma^j_{00} - \Gamma^k_{0j}\Gamma^j_{0k}}_{\approx 0}\\ &= \sum_{k=1}^{3}\partial_k\Gamma^k_{00} = 4\pi G\rho\end{aligned}$$

(the $3^{\text{rd.}}$ equality is because the time derivative of Christoffel symbols should be 0 since the field is static) Hence we have that

$$\mu = \frac{4\pi G\rho}{T^{00} - \mathcal{T}g^{00}/2} = \frac{8\pi G}{c^4}$$

□

We don't know if this equation is true or not, but it explains & predicts a lot of things that were used to be a mystery using Newtonian mechanics; e.g The anomalous perihelion shift of Mercury & the prediction of the deflection of light by the Sun. So we think it is true.

8.2 Morse theory

This section will provides a background knowledge for later sections.
The main thinking and definition of Morse homology is filled with physical intuition

Definition 8.2. Given a scalar function $f \in C^\infty(M)$,

- (Critical value) For $x \in M$, the $f(x)$ is a critical value if $\mathrm{d}f|_x = 0$.
- (Hessian) Under a local coordinate $(x_1, \dots, x_n)$, the matrix $\mathrm{Hess}_x(f) := (\partial^2 f / \partial x_i \partial x_j)$ is called the Hessian of f at x.
- (Morse function) The f is a **Morse function** if the Hessian every critical point is non-degenerate.
- (Gradient) Gradient f is a vector field $\mathrm{grad}(f) \in \Gamma^\infty(TM)$ such that

$$\mathrm{grad}(f)|_x f = \max\{v(f) : v \in T_x M, ||v|| = 1\}$$

  for all $x \in M$.
- (Gradient flow) Given Morse function f, the gradient flow of f is the flow $\phi : \mathbb{R} \times M \to M$ of the ODE $\dot{\gamma}(f) = \mathrm{grad}(f)$ for $\gamma : \mathbb{R} \to M$.
- (Stable submanifold) Given Morse function f, the stable submanifold at x is

$$W_+(x) = \left\{p \in M : \lim_{t \to \infty} \phi(t, p) = x\right\}$$

- (Unstable submanifold) Given Morse function f, the unstable submanifold at x is

$$W_-(x) = \left\{p \in M : \lim_{t \to -\infty} \phi(t, p) = x\right\}$$

In these definitions, it is left as an exercise to show that: **1**, the Morse function is well-defined (Independent of the choice of local coordinates which are used to define Hessian); **2**, The gradient flow always exists (That is, the solution to that ODE exists on entire $\mathbb{R}$) **3**, W_- and W_+ are indeed submanifolds.

The Morse theory basically tells how critical points of f captures the topological information.

Definition 8.2. (Moduli space of gradient flows) For two critical points $x, y \in M$, denote

$$\widetilde{\mathcal{M}}(x, y) := \left\{\gamma \text{ gradient flow of } f : \lim_{t \to -\infty} \gamma(t) = x \text{ and } \lim_{t \to \infty} \gamma(t) = y\right\}$$

Define $\gamma_1 \sim \gamma_2$ if $\exists r \in \mathbb{R}$ such that $\gamma_1(\cdot + r) = \gamma_2(\cdot)$, and called $\mathcal{M}(x, y) := \widetilde{\mathcal{M}}(x, y)/\sim$ the Moduli space of gradient flows.

9 (Special) What if Complex?

This section is not a prerequisite for contents afterwards, so feel free to skip it. But I believe that this section explains a number of underlying reasons and motivations of de-rham cohomology and knowledge afterwards.

What will happen if the $\mathbb{R}$ for locally Euclidean of manifolds is changed to another locally ringed space; say for example, $\mathbb{C}$. How will our theory change?
It turns out there are HUGE differences! Let's review some of our previous theories.

Definition 9.1. Topological space M is a **Complex manifold** if M is Hausdorff, second countable, locally homeomorphic to an *open ball in* $\mathbb{C}^n$ for some $n \in \mathbb{Z}_{>0}$, and its transition functions are holomorphic.

Why do I highlight the *open ball in* $\mathbb{C}^n$?
Consider the "diffeomorphism" for complex manifolds: **biholomorphism**, which is the same definition as diffeomorphism but replaces smooth maps with holomorphic maps; the most important difference produced is that the "open ball in $\mathbb{C}^n$" in our definition of the complex manifold cannot be replaced by $\mathbb{C}^n$ itself!

Example 9.1. Open ball $\{z \in \mathbb{C} : ||z|| < 1\}$ is NOT biholomorphic to $\mathbb{C}$.

a direct corollary of Liouville theorem. Actually, these 3 topologically identical spaces are all NOT biholomorphic:

1. $\mathbb{C}^n := \mathbb{C} \times \cdots \times \mathbb{C}$
2. Open ball $\{(z_1, \cdots, z_n) \in \mathbb{C}^n : \sum_{k=1}^n ||z_k||^2 < 1\}$
3. Polydisk $\{(z_1, \cdots, z_n) \in \mathbb{C}^n : ||z_k|| < 1\}$

In other words, the topology is not the only thing we concern about complex manifolds when comparing them.

Luckily, things like (co)tangent space, pullback/pushforward, differential forms and de-rham cohomology are still well-defined by changing $\mathbb{R}$ to $\mathbb{C}$; however, POU is NOT guaranteed in complex manifolds! The main ingredient missing is that, we used the bump function to construct POU, but bump function doesn't exist in $\mathbb{C}^n$:

Lemma 9.2. If $f \in C^\infty(\mathbb{C})$ has compact support, then $f = 0$.

which is a direct corollary of the maximum modulus principle. And a serious consequence is that Whitney's Embedding theorem won't be true for complex manifolds. [12]

Example 9.3. To sum up,

1. POU doesn't exist.
2. As a result, theoerms like Whitney embedding theorem and de-rham theorem no longer hold.
3. Open ball is not biholomorphic to $\mathbb{C}^n$
4. Since topologically $\mathbb{C} \cong \mathbb{R}^2$ and holomorphic maps are orientation-preserving when viewing them as maps on $\mathbb{R}^2$, all complex manifolds are said to be orientable.

Most of the differences that changes our theory are based one single intuition: Functions in complex manifolds have rigidity. In other words, how functions behave locally decides how they behave globally (in conneced M).

This leads to that the study in complex geometry mainly relies on the study of functions on it, and can be thought as the initial insights of using functions on a space to study the space itself —— e.g de-rham cohomology. And moreover, the de-rham cohomology, which doesn't give much information on (real) manifolds, could give (alomst) all topological information on complex manifolds.

And this is generally why this section would be more related to algebraic geometry, and be placed as a special section.

[12] More interestingly, being able to be embedded into some $\mathbb{C}^m$ is very rare for complex manifolds —— they even have a special name called Stein manifold.

9.1 Basics on Riemann surface

Consider 1-dim complex manifolds, which are also called the **Riemann Surface**s; we know that 1-dim connected smooth manifolds are quite trivial —— just $\mathbb{R}$ or $\mathbb{S}^1$; but Riemann Surfaces are far more interesting to imagine.

Example 9.4. We've already known 2 distinct Riemann surfaces, $\mathbb{C}$ and open ball; another example is the **Riemann sphere** $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$, where ∞ is defined to be a point in $\mathbb{C}_\infty$ such that $\mathbb{C}_\infty \setminus \{0\}$ is biholomorphic to $\mathbb{C}$ via $z \mapsto 1/z$.

Let's start with the remark we had in the introduction, that *an open ball* $\mathbb{B} \subset \mathbb{C}$ *is NOT biholomorphic to* $\mathbb{C}$; but clearly $\mathbb{D}$ is biholomorphic to some other simply connected open set, while the Riemann mapping theorem ensures that $\mathbb{C}$ is the only exception.

Riemann Mapping theorem

Theorem 9.5. Simply connected open $Q \subsetneq \mathbb{C}$ is biholomorphic to an open ball.

Luckily, the *conventional* proof is already presented in the Wikipedia page and many other sources (like this); so here we will only present its main idea.

Let's fix a point $q \in Q$ and an open ball $\mathbb{B} := \{z \in \mathbb{C} : ||z|| < 1\}$. This proof can be divided into two construction & one analysis steps:

1. ((Construction) For any Q, there exists a non-empty set
$$\mathcal{F} := \{ \text{ analytic and injective } f : Q \to \mathbb{D} \text{ that } f(q) = 0\}$$
2. (Analysis) Since $\mathcal{F}$ is non-empty, let $M := \sup_{f\in\mathcal{F}} |f'(q)|$ (it might be ∞), and there exists a sequence $\{f_k\} \subset \mathcal{F}$ such that $\lim_{k\to\infty} |f_k'(q)| = M$;
Then prove that it has a subsequence $\{f_{\alpha(j)}\}$ that uniformly converges to some $f \in \mathcal{F}$ on any compact subset of Q;
3. (Construction) Clearly the existence of such f means that $M = |f'(q)| < \infty$; then we show that the f is surjective by contradiction.

And since a bijective holomorphic map has holomorphic inverse (try verify this), these finish the proof to our theorem.

We won't dive into the analysis step (though it's the hardest part in this proof); we will instead consider the $1^{\text{st.}}$ and $3^{\text{rd.}}$ steps which contains most of the intuition.

Proof. ($1^{\text{st.}}$ step) The trick here is to use Möbius transformation to move (possibly) unbounded Q into a bounded region: Since $Q \neq \mathbb{C}$, there $\exists\ y \in \mathbb{C} \setminus Q$, so we can construct

$$\phi : z \mapsto (z - y)^{-1}$$

This ϕ would be bounded, analytic and injective; so

$$\psi(z) := \frac{\phi(z) - \phi(q)}{\sup_{s \in Q} |\phi(s) - \phi(q)|}$$

gives a valid element in $\mathcal{F}$. (i.e. $\mathcal{F} \neq \emptyset$) $\square$

Proof. (3$^{\text{rd.}}$ step) This step uses an idea from an interesting conclusion in geometrical complex analysis, which says

Lemma 9.6. For $c \in \mathbb{B}$, define $\varphi_c : \mathbb{B} \to \mathbb{B}$ to be

$$\varphi_c : z \mapsto \frac{z - c}{1 - \bar{c}z}$$

Then the set of bijective analytic functions from $\mathbb{B}$ to $\mathbb{B}$ is exactly $\{\varphi_c : c \in \mathbb{B}\}$

this can be proved by Schwarz lemma, left as an exercise.
Now, let's prove the 3$^{\text{rd.}}$ step by contradiction: suppose there $\exists\, b \in \mathbb{B} \setminus \mathrm{Im}(f)$; and we hope to construct some $f_+ \in \mathcal{F}$ such that $|f_+'(q)| > |f'(q)|$.
If our statement is true, then for elements in $\mathcal{F}$, the bigger the size of their derivative at q, the bigger the size of their image, so the trick is using a function $\mathbb{B} \to \mathbb{B}$ to "enlarge" $\mathrm{Im}(f)$; the $\sqrt{\ }$ would be a good candidate.
Be careful that $\sqrt{\ }$ is not holomorphic at 0, so we first use φ_b (in the lemma) to move b to 0, apply the $\sqrt{\ }$, and then move the original 0 back:

$$f_+(z) := \varphi_{\widetilde{b}}\left(\sqrt{\varphi_b \circ f(z)}\right)$$

where $\widetilde{b} = \sqrt{\varphi_b \circ f(q)} = \sqrt{-b}$. Checking that

$$|f_+'(q)| = |f'(q)| \cdot \frac{1}{2}\left(\left|\sqrt{-b}\right| + \left|\sqrt{-b}\right|^{-1}\right) > |f'(q)|$$

contradict to the assumption $|f'(q)| = M$. $\square$

By going through these 3 steps, we get a proof of Riemann Mapping theorem.
Now we can further ask is that, is there any other simply connnected Riemann surfaces, except for the open ball, $\mathbb{C}$, and $\mathbb{C}_\infty$?

Example 9.7. (Uniformization theorem) Every simply connected Riemann surface is biholomorphic to an open ball, $\mathbb{C}$ or $\mathbb{C}_\infty$.

This is hard, but is definitely one of our main interest; the proof will be given when we're ready. Besides this, we will have an overview on other things that may spark your interest in Riemann surface.

When studying complex analysis, you might have notice that there are lots of analytic functions that couldn't be defined on the entire $\mathbb{C}$, like $\sqrt{}$ and ln; a key observation of Riemann was that, we can usually define those functions as an entire function on a covering space of $\mathbb{C}$.

10 Characteristic Classes

Theory with smooth vector bundles is interesting as it is where some differential geometry met with algebraic topology, and those theories on vector bundles eventually reveals some essence of the cohomology on manifolds.

Cohomology classes $\longleftrightarrow$ Obstructions on vector bundles

[**Convention**]: In this section, we assume/denote

1. M being a n-dim connected smooth manifold;
2. ξ being a rank r smooth $\mathbb{R}$-vector bundle over M;
3. $H_k(X, A)$ and $H^k(X, A)$ denote the relative (co)homology of (X, A) over integer ring $\mathbb{Z}$.

The best starting point of this chapter is to ask this question: When does a k-rank trivial subbundle fail to exist in ξ? In other words, how many (globally) linearly independent non-zero section can there exists, and how to determine it via information of M?

Let's first have a small attempt; let's say M satisfies $\mathcal{P}_k$ if there exists at least k linearly independent sections.

First, we can reformulate this problem. Given a (ordered) list of k linearly independent vectors in $\mathbb{R}^r$, the choice (or to say, order) of these vectors, by definition gives an element in Stiefel manifold $V_k(\mathbb{R}^r)$. So if we define

Definition 10.1. Stiefel bundle associated with ξ is

$$V_k(\xi) := \bigsqcup_{x \in M} V_k(\pi^{-1}(x))$$

with $\widetilde{\pi} : (x, M) \mapsto x$ and local trivialization inherited from ξ's.

then a list of k linearly independent sections in ξ is one section in $V_k(\xi)$, which means

$$\mathcal{P}_k \iff \Gamma(V_k(\xi)) \neq \emptyset$$

Let's suppose M is equipped with a CW complex structure, the idea of justifying $\mathcal{P}_k$ is to see if $\mathcal{P}_k$ is hold for j-skeletons $M_0 \subset \cdots \subset M_j$ one by one.

Clearly $\mathcal{P}_k$ always holds for M_0, and it's also easy to extend it to M_1: For any 1-cell l connecting two 0-cells a_1 and a_2, there exists a path from arbitrary $f_1 \in V_k(a_1)$ to $f_2 \in V_k(a_2)$ by the path connectness of $V_k(\mathbb{R}^r)$; and this path gives an element in $\Gamma(V_k(l))$.

But the problem becomes non-trivial at M_2.

Given a 2-cell $\sigma \subset M$ and $f \in \Gamma\left(V_k(\pi^{-1}(\partial\sigma))\right)$, suppose there does exists $\widetilde{f} \in \Gamma\left(V_k(\pi^{-1}(\sigma))\right)$ such that $f = \widetilde{f}|_{\partial\sigma}$, then $f(\partial\sigma) = \widetilde{f}(\partial\sigma)$ must gives a contractible loop in $V_k(\mathbb{R}^r)$, but this is not guarenteed: For example, $\pi_1(V_r(\xi)) \cong \mathbb{Z}$ is non-trivial.

And therefore the difficulty of extending f to a valid $\widetilde{f}$ is determined by the kernal of map

$$\begin{gathered}\phi_f : \mathrm{C}_2(M; \mathbb{Z}) \to \pi_1(V_r(\mathbb{R}^r)) \cong \mathbb{Z} \\ \sigma \mapsto [f(\partial\sigma)]\end{gathered}$$

Easy to verify that this is in a cohomology class $[\phi_f] \in H^2(M; \mathbb{Z})$. In addition, ϕ_f are in the same cohomology class for all choices of f: For two non-vanishing $p, q \in \Gamma(V_r(\pi^{-1}(M_1)))$, there exists

$$(\phi_p - \phi_q)\sigma = [p(\partial\sigma)] - [q(\partial\sigma)]$$

10.1 Stiefel-Whitney class

From our discussion in the introduction, we argued that the difficulty of constructing k linearly independent sections of ξ on an m-skeleton M_m is given by the trivial-ness of the map

$$\begin{aligned} \phi_f : \mathrm{C}_m(M;\mathbb{Z}) &\to \pi_{m-1}(V_k(\mathbb{R}^r)) \\ \sigma &\mapsto [f(\partial\sigma)] \end{aligned}$$

for some non-vanishing section $f \in \Gamma(V_k(\pi^{-1}(M_m)))$. That is, if the map is zero for all m, then there must exists k linearly independent sections on ξ. And, we also know that this ϕ_f gives a cohomology class $[\phi_f] \in H^m(M;\pi_{m-1}(V_k(\mathbb{R}^r)))$.
It will be nice if we know the homotopy group $\pi_k(V_{r-k+1}(\mathbb{R}^n))$; a conclusion that we had is

Lemma 10.1. For $k = 1, 2, 3, \ldots$ we have that $\pi_m(V_k(\mathbb{R}^r)) = 0$ for $\forall m < r-k$.

$$\pi_{r-k}(V_k(\mathbb{R}^r)) \cong \begin{cases} \mathbb{Z} & \text{for } 2|(r-k) \text{ or } k=1 \\ \mathbb{Z}/2\mathbb{Z} & \text{otherwise} \end{cases}$$

which suggests that there shouldn't be any problem to construct k linearly independent sections on any m-skeleton for $m \leq r-k$.
So this ϕ_f would be a very useful indicator, and to make things easy and uniform, we reduce the cases where the coefficient ring is $\mathbb{Z}$ to $\mathbb{Z}/2\mathbb{Z}$ coefficient cohomology by mod 2 reduction. And this (roughly) gives the Stiefel-Whitney class, by considering upon the necessary conditions that $[\phi_f]$ satisfies:

Definition 10.2. The kth **Stiefel-Whitney class** of ξ is an element $\mathrm{w}_k(\xi) \in H^k(M;\mathbb{Z}/2\mathbb{Z})$, that satisfies the following 4 axioms; the $\mathrm{w}(\xi)$ is the $\mathbb{Z}/2\mathbb{Z}$-algebra generated by $\{\mathrm{w}_0(\xi), \mathrm{w}_1(\xi), \mathrm{w}_2(\xi), \ldots\}$.

- $\mathrm{w}_0(\xi) = 1$.
- For bundle morphism $F : \xi \to \zeta$ with $f : M \to N$, $\mathrm{w}_k(\zeta) = f^* \circ \mathrm{w}(\xi)$
- For any vector bundle ζ over M, $\mathrm{w}(\xi \oplus \zeta) = \mathrm{w}(\xi) \smile \mathrm{w}(\zeta)$.
- $\mathrm{w}_1(E_1(\mathbb{R})) \neq 0$

The first 3 conditions are quite reasonable, while the last one is confusing —— why can't I replace $E_1(\mathbb{R})$ with other non-trivial vector bundles? We will answer that somehow later, but here a simple explanation is from the "universal" nature of $E_1(\mathbb{R})$

10.2 Orientation and Thom isomorphism

The most ob Like manifolds, vector bundles also have the problem of orientation when considering some global properties or operations. The definition of orientation for smooth vector bundles is quite similar to that of manifolds

Definition 10.3. Consider open cover $\{U_j\}$ of M where each U_j is equipped with local trivialization $\Phi_j : \pi^{-1}(U_j) \to U_j \times \mathbb{R}^r$, smooth vector bundle ξ is **oriented** if: For $\forall U_a \cap U_b \neq \emptyset$ and $(x, v) \in \pi^{-1}(U_a \cap U_b)$ such that

$$\Phi_a \circ \Phi_b^{-1}(x, v) = (x, Av)$$

for some $A \in \mathrm{GL}_r(\mathbb{R})$, the $\det(A) > 0$.

Try verifying that such property is independent of the choice of local trivialization.
So we know that, since M is connected, the choice of basis at one point in M determines the orientation of the entire ξ. So formally,

Definition 10.1. When ξ is oriented, fix $p \in M$, an **orientation** is an equivalence class of surjective linear maps $\phi : \mathbb{R}^r \to \pi^{-1}(p)$, where $\phi_1 \sim \phi_2$ iff $\det(\phi_1^{-1} \circ \phi_2) > 0$.

Clearly there can only be two orientations for ξ.
Previously we went through the idea of representing the orientation of M through the fundamental class, so let's try working out a similar way for vector bundles.
The idea begins with a trivial observation:

Example 10.2. Choosing the orientation of $\mathbb{R}^r$ is equivalent to choosing a generator of $H^r(\mathbb{R}^r, \mathbb{R}^r \setminus \{0\}\,; \mathbb{Z}) \cong \mathbb{Z}$. This is because any basis $\{v_1, \ldots, v_r\}$ of $\mathbb{R}^r$ at 0 forms a simplex σ based at 0, so an element $\omega \in H^r(\mathbb{R}^r, \mathbb{R}^r \setminus \{0\}\,; \mathbb{Z})$ tells if that basis agrees with the chosen orientation by the sign of $\omega|_0\,(\sigma)$.
So for a trivial bundle ξ of M, the orientation of ξ is given by the choice of two generators of $H^r(\xi, \xi \setminus M; \mathbb{Z})$.

10.3 Euler Class

10.4 Chern-Weil theory

10.5 Characteristic classes from $\mathbf{Gr}_\infty$

There are tons of ways to interpret the characteristic classes. At the beginning, we introduce them via obstructions; here, with enough examples, we can consider a more universal way of thinking through those classes.
We can

References

[1] Do Carmo. *Differential geometry of curves and surfaces*. Dover Publications, 1976.

[2] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.