Basic Geometry and Topology

Stephan Stolz

December 9, 2019

Contents

1 Pointset Topology

1.1 Open subsets of $\mathbb{R}^n$

Definition 1.1. For $x \in \mathbb{R}^n$ and $r > 0$, let

$$B_r(x) := \{y \in \mathbb{R}^n \mid \mathrm{dist}(x, y) < r\}$$

be the open ball $B_r(x)$ of radius r around x. Here

$$\mathrm{dist}(x, y) := ||x - y|| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

is the distance between the points x and y.

A subset $U \subset \mathbb{R}^n$ is *open* if for each point there is some $r > 0$ such that $B_r(x)$ is contained in U. Equivalently, U is open if and only if U is a union of open balls.

The point of this definition is that it makes it possible to give a very compact definition of continuity of maps $f : \mathbb{R}^m \to \mathbb{R}^n$ which is equivalent to the usual ϵ-δ definition.

Definition 1.2. A map $f : \mathbb{R}^m \to \mathbb{R}^n$ is *continuous* if for every open subset $U \subset \mathbb{R}^n$ the preimage $f^{-1}(U)$ is open in $\mathbb{R}^m$. More generally, if $V \subset \mathbb{R}^m$, $W \subset \mathbb{R}^n$ are open subsets a map $f : V \to W$ is *continuous* if for every open subset $U \subset \mathbb{R}^n$ the preimage $f^{-1}(U)$ is open in $\mathbb{R}^m$.

Examples of continuous maps.

1. From calculus we know that the following maps $f\colon \mathbb{R} \supset V \to \mathbb{R}$ are continuous: polynomials, exponential functions, rational functions, trigonometric functions. Here $V \subset \mathbb{R}$ is the natural domain of these functions.

2. The maps $\mathbb{R}^2 \to \mathbb{R}$ given by $(x_1, x_2) \mapsto x_1 + x_2$ or $(x_1, x_2) \mapsto x_1 x_2$.

3. The projection maps $p_k\colon \mathbb{R}^m \to \mathbb{R}$ given by $(x_1, \dots, x_m) \mapsto x_k$.

Warning. The open set characterization of continuity is great for more abstract statements, like showing that the composition of continuous maps is continuous. However, checking that a given map f is continuous by verifying that $f^{-1}(U)$ is open for an open subset U of the codomain of f is usually cumbersome. A much better strategy is to recognize a given map as "built from simpler maps" that we already know to be continuous. The following three lemmas illustrate what we mean by "built from".

Lemma 1.3. *The composition of continuous maps is continuous.*

We leave the simple proof to the reader.

Lemma 1.4. *A map $f\colon V \to \mathbb{R}^n$, $f(x) = (f_1(x), \dots, f_n(x))$ is continuous if and only if all its component maps $f_k\colon V \to \mathbb{R}$ are continuous.*

The proof of this statement will follow from the much more general *continuity criterion for maps to a product*, which we will prove after introducing the product topology (see Lemma 1.19).

Lemma 1.5. *Let $f_1, f_2\colon \mathbb{R} \supset V \to \mathbb{R}$ be continuous maps. Then also $f_1 + f_2$ and $f_1 \cdot f_2$ are continuous.*

Proof. Let $f\colon V \to \mathbb{R}^2$ be the map with components maps f_1, f_2; i.e., $f(x) = (f_1(x), f_2(x))$. The map f is continuous since it component maps are continuous. The map $f_1 + f_2\colon V \to \mathbb{R}$ can be factored as

$$V \xrightarrow{f} \mathbb{R}^2 \xrightarrow{+} \mathbb{R}$$

and hence is continuous as the composition of continuous maps. Replacing the map $\mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto x_1 + x_2$ by the map $(x_1, x_2) \mapsto x_1 x_2$ similarly shows that $f_1 \cdot f_2$ is continuous. $\square$

Example 1.6. (More Examples of continuous maps.)

1. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a polynomial map, i.e.,

$$f(x) = \sum_{i_1,\dots,i_n} a_{i_1,\dots,i_n} x_1^{i_1} \cdots x_n^{i_n} \qquad \text{for } x = (x_1,\dots,x_n) \in \mathbb{R}^n \text{ and coefficients } a_{i_1,\dots,i_n} \in \mathbb{R}$$

   We observe that f is a sum of functions, and each summand is a product of projection maps $x \mapsto x_k$ and the constant map $x \mapsto a_{i_1,\dots,i_n}$. Hence the continuity of the project maps and constant maps imply by Lemma 1.5 the continuity of each summand, which in turn implies the continuity of f.

2. Let $M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n^2}$ be the vector space of $n \times n$ matrices. Then the map

$$M_{n\times n}(\mathbb{R}) \times M_{n\times n}(\mathbb{R}) \longrightarrow M_{n\times n}(\mathbb{R}) \qquad (A,B) \mapsto AB$$

   given by matrix multiplication is continuous. To see this, it suffices by Lemma 1.4 to check that each component map is continuous. This is the case, since each matrix entry of AB is a polynomial and hence a continuous function of the matrix entries of A and B.

1.2 Topological spaces

The characterization 1.2 of continuous maps $f\colon \mathbb{R}^m \to \mathbb{R}^n$ in terms of open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ suggests that we can define what we mean by a *continuous map* $f\colon X \to Y$ between sets X, Y, once we pick collections $\mathcal{T}_X$, $\mathcal{T}_Y$ of subsets of X resp. Y that we consider the "open subsets" of these sets. The next result summarizes the basic properties of open subsets of a metric space X, which then motivates the restrictions that we wish to put on such collections $\mathcal{T}$.

Lemma 1.7. *Open subsets of a metric space X have the following properties.*

(i) X and $\emptyset$ are open.

(ii) Any union of open sets is open.

(iii) The intersection of any finite number of open sets is open.

Definition 1.8. A *topological space* is a set X together with a collection $\mathcal{T}$ of subsets of X, called *open sets* which are required to satisfy conditions (i), (ii) and (iii) of the lemma above. The collection $\mathcal{T}$ is called a *topology* on X. The sets in $\mathcal{T}$ are called the *open sets*, and their complements in X are called *closed sets*. A subset of X may be neither closed nor open, either closed or open, or both.

A map $f\colon X \to Y$ between topological spaces X, Y is *continuous* if the inverse image $f^{-1}(V)$ of every open subset $V \subset Y$ is an open subset of X.

It is easy to see that the composition of continuous maps is again continuous.

Example 1.9. (Examples of topological spaces.)

1. Let $\mathcal{T}$ be the collection of open subsets of $\mathbb{R}^n$ in the sense of Definition 1.1. Then $\mathcal{T}$ is a topology on $\mathbb{R}^n$, the *standard topology on* $\mathbb{R}^n$ or *metric topology on* $\mathbb{R}^n$ (since this topology is determined by the metric $\text{dist}(x, y) = ||x - y||$ on $\mathbb{R}^n$).

2. Let X be a set. Then $\mathcal{T} = \{\text{all subsets of } X\}$ is a topology, the *discrete topology.* We note that *any* map $f : X \to Y$ to a topological space Y is continuous. We will see later that the only continuous maps $\mathbb{R}^n \to X$ are the constant maps.

3. Let X be a set. Then $\mathcal{T} = \{\emptyset, X\}$ is a topology, the *indiscrete topology.*

Sometimes it is convenient to define a topology $\mathcal{U}$ on a set X by first describing a smaller collection $\mathcal{B}$ of subsets of X, and then defining $\mathcal{U}$ to be those subsets of X that can be written as *unions* of subsets belonging to $\mathcal{B}$. We've done this already when the topology on $\mathbb{R}^n$: Let $\mathcal{B}$ be the collection of all open balls $B_r(x) \subset \mathbb{R}^n$; we recall that $B_r(x) = \{y \in X \mid \text{dist}(x, y) < r\}$. The standard topology on $\mathbb{R}^n$ consists of those subsets U which are unions of subsets belonging to $\mathcal{B}$.

Lemma 1.10. *Let $\mathcal{B}$ be a collection of subsets of a set X satisfying the following conditions*

1. *Every point $x \in X$ belongs to some subset $B \in \mathcal{B}$.*

2. *If $B_1, B_2 \in \mathcal{B}$, then for every $x \in B_1 \cap B_2$ there is some $B \in \mathcal{B}$ with $x \in B$ and $B \subset B_1 \cap B_2$.*

Then $\mathcal{T} := \{\text{unions of subsets belonging to } \mathcal{B}\}$ is a topology on X.

Definition 1.11. If the above conditions are satisfied, we call the collection $\mathcal{B}$ is called a *basis for the topology* $\mathcal{T}$ or we say that $\mathcal{B}$ *generates the topology* $\mathcal{T}$.

1.2.1 Subspace topology

Definition 1.12. Let X be a topological space, and $A \subset X$ a subset. Then

$$\mathcal{T} = \{A \cap U \mid U \underset{open}{\subset} X\}$$

is a topology on A called the *subspace topology.*

Example 1.13. (Examples of subspaces of $\mathbb{R}^n$) Here are examples of subspaces of $\mathbb{R}^n$ (i.e., subsets of $\mathbb{R}^n$ equipped with the subspace topology) we will be talking about during the semester:

1. The *n-disk* $D^n := \{x \in \mathbb{R}^n \mid |x| \le 1\} \subset \mathbb{R}^n$, and $D^n_r := \{x \in \mathbb{R}^n \mid |x| \le r\}$, the *n-disk of radius* $r > 0$.

2. The *n-sphere* $S^n := \{x \in \mathbb{R}^{n+1} \mid |x| = 1\} \subset \mathbb{R}^{n+1}$.

3. The *torus* $T = \{v \in \mathbb{R}^3 \mid \operatorname{dist}(v, C) = r\}$ for $0 < r < 1$. Here
$$C = \{(x, y, 0) \mid x^2 + y^2 = 1\} \subset \mathbb{R}^3$$
is the unit circle in the xy-plane, and $\operatorname{dist}(v, C) = \inf_{w \in C} \operatorname{dist}(v, w)$ is the distance between v and C.

4. The *general linear group*
$$\begin{aligned} GL_n(\mathbb{R}) &= \{\text{vector space isomorphisms } f\colon \mathbb{R}^n \to \mathbb{R}^n\} \\ &\longleftrightarrow \{(v_1, \dots, v_n) \mid v_i \in \mathbb{R}^n,\ \det(v_1, \dots, v_n) \neq 0\} \\ &= \{\text{invertible } n \times n\text{-matrices}\} \subset M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n^2} \end{aligned}$$
Here we think of $(v_1, \dots, v_n)$ as an $n \times n$-matrix with column vectors v_i, and the bijection is the usual one in linear algebra that sends a linear map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ to the matrix $(f(e_1), \dots, f(e_n))$ whose column vectors are the images of the standard basis elements $e_i \in \mathbb{R}^n$.

5. The *special linear group*
$$SL_n(\mathbb{R}) = \{(v_1, \dots, v_n) \mid v_i \in \mathbb{R}^n,\ \det(v_1, \dots, v_n) = 1\} \subset M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n^2}$$

6. The *orthogonal group*
$$\begin{aligned} O(n) &= \{\text{linear isometries } f\colon \mathbb{R}^n \to \mathbb{R}^n\} \\ &= \{(v_1, \dots, v_n) \mid v_i \in \mathbb{R}^n,\ v_i\text{'s are orthonormal}\} \subset M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n^2} \end{aligned}$$
We recall that a collection of vectors $v_i \in \mathbb{R}^n$ is *orthonormal* if $|v_i| = 1$ for all i, and v_i is perpendicular to v_j for $i \neq j$.

7. The *special orthogonal group*
$$SO(n) = \{(v_1, \dots, v_n) \in O(n) \mid \det(v_1, \dots, v_n) = 1\} \subset M_{n\times n}(\mathbb{R}) = \mathbb{R}^{n^2}$$

8. The *Stiefel manifold*
$$V_k(\mathbb{R}^n) = \{(v_1, \dots, v_k) \mid v_i \in \mathbb{R}^n,\ v_i\text{'s are orthonormal}\} \subset M_{n\times k}(\mathbb{R}) = \mathbb{R}^{nk}$$

Lemma 1.14. (Continuity criterion for maps to a subspace.) *Let X, Y be topological spaces and let A be a subset of Y equipped with the subspace topology.*

- *The inclusion map $i\colon A \to Y$ is continuous.*
- *A map $f\colon X \to A$ is continuous if and only if the composition $X \xrightarrow{f} A \xrightarrow{i} Y$ is continuous.*

Proof. Homework □

Example 1.15. (Examples of continuous maps involving subspaces.)

1. The map $GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$, $A \mapsto A^{-1}$ is continuous. Homework problem. Hint: by the above lemma, it suffices to prove continuity of the composition $GL_n(\mathbb{R}) \to GL_n(\mathbb{R}) \hookrightarrow M_{n\times n}(\mathbb{R})$, which in turn by Lemma 1.4 amounts to checking continuity of each matrix componentof A^{-1} as a function of the matrix components of A.

2. Let G be one of the groups $SL_n(\mathbb{R})$, $O(n)$, $SO(n)$, equipped with the subspace topology as subsets of $M_{n\times n}(\mathbb{R})$. Then the map $G \to G$, $A \mapsto A^{-1}$ is continuous. To see that this map is continuous, we note it is the restriction of the continuous map $A \mapsto A^{-1}$ on $GL_n(\mathbb{R})$ to the subspace $G \subset GL_n(\mathbb{R})$ and use the following handy fact.

Lemma 1.16. *Let $f\colon X \to Y$ be a continuous map. If $A \subset X$, $B \subset Y$ are subspaces with $f(A) \subset B$, then the* restriction *$f_{|A}\colon A \to B$ is continuous (with respect to the subspace topology on X and Y.*

Proof. Consider the commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{f_{|A}} & B \\ {\scriptstyle i}\downarrow & & \downarrow{\scriptstyle j} \\ X & \xrightarrow{f} & Y \end{array}$$

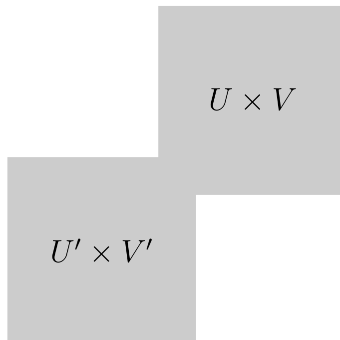
where i, j are the obvious inclusion maps. These inclusion maps are continuous w.r.t. the subspace topology on A, B by Lemma 1.14. The continuity of f and i implies the continuity of $f \circ i = j \circ f_{|A}$ which again by Lemma 1.14 implies the continuity of $f_{|A}$. □

1.2.2 Product topology

Definition 1.17. The *product topology* on the Cartesian product $X \times Y = \{(x, y) \mid x \in X,\, y \in Y\}$ of topological spaces X, Y is the topology generated by the subsets

$$\mathcal{B} = \{U \times V \mid U \underset{open}{\subset} X, V \underset{open}{\subset} Y\}$$

The collection $\mathcal{B}$ obviously satisfies property (1) of a basis (see Definition 1.11); property (2) holds since $(U \times V) \cap (U' \times V') = (U \cap U') \times (V \cap V')$. We note that the collection $\mathcal{B}$ is *not* a topology since the union of $U \times V$ and $U' \times V'$ is typically not a Cartesian product. For example, if $X = Y = \mathbb{R}$ and U, U', V, V' are open intervals the products $U \times V$ and $U' \times V'$ are (open) rectangles whose union might look like the shaded region in the figure below.



There is obviously a plethora of examples of product spaces, e.g., the product of any two of the eight spaces of Example 1.13. Sometimes, the product topology on a product agrees with a topology described in a different way, for example:

Lemma 1.18. *The product topology on $\mathbb{R}^m \times \mathbb{R}^n$ (with each factor equipped with the metric topology) agrees with the metric topology on $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$.*

Proof: homework.

Other product spaces might be homeomorphic to topological spaces constructed completely differently. For example, we will see that the product $S^1 \times S^1$ is homeomorphic to the torus T of Example 1.13(3). To work with product spaces, it is very useful to have the following recognition principal for continuity of map to a product.

Lemma 1.19. (Continuity criterion for maps to a product.) *Let X, Y_1, Y_2 be topological spaces.*

- *The projection maps $p_i \colon Y_1 \times Y_2 \to Y_i$ are continuous.*
- *A map $f \colon X \to Y_1 \times Y_2$ is continuous if and only if the compositions*

$$X \xrightarrow{f} Y_1 \times Y_2 \xrightarrow{p_i} Y_i$$

are continuous for $i = 1, 2$.

We note that the composition $p_i \circ f$ is the i-th component map of f. So according to the above lemma a map to a product is continuous if and only if all its component maps are continuous. This is a far reaching generalization of Lemma 1.4 which was about maps with target space $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$.

For the proof of Lemma 1.19, as well as in many other situations, it will be helpful to use the following simple result, the reader is charged with proving.

Lemma 1.20. *Let $f\colon X \to Y$ be a map be topological spaces. Suppose the topology on the codomain Y is generated by a basis $\mathcal{B}$. Then f is continuous if and only if $f^{-1}(U)$ is open in X for every $U \in \mathcal{B}$.*

Proof of Lemma 1.19. To show that the projection map $p_1\colon Y_1 \times Y_2 \to Y_1$ is is continuous, suppose that $U \subset Y_1$ is an open subset. Then $p_1(U) = U \times Y_2$, which is an open subset of $U \times Y_2$ by construction of the product topology (in fact this is a product of open subsets of Y_1 resp. Y_2, i.e., it belongs to the collection of subsets $\mathcal{B}$ that generates the product topology). The argument that p_2 is continuous is completely analogous.

If $f\colon X \to Y_1 \times Y_2$ is continuous, then the component maps $f_i := p_i \circ f$ are continuous, since they are compositions of the continuous maps p_i and f. Conversely, assume that the component maps f_1, f_2 are continuous. To show that f is continuous it suffices by the previous lemma to show that $f^{-1}(U)$ is open where U belongs to the basis $\mathcal{B}$ that generated the product topology. In other words, U is a product $U = U_1 \times U_2$ of open subsets $U_1 \subset Y_1$, $U_2 \subset Y_2$. Then

$$f^{-1}(U) = f^{-1}(U_1 \times U_2) = f_1^{-1}(U_1) \cap f_2^{-1}(U_2) \subset X,$$

is an open subset of X, since $f_i^{-1}(U_i)$ is open in X by the assumed continuity of f_i. $\square$

The following result is consequence of the Continuity criterion for maps to a product; its proof is a good illustration of how the criterion is used.

Lemma 1.21. *Let G be one of the groups $GL_n(\mathbb{R})$, $SL_n(\mathbb{R})$, $O(n)$, $SO(n)$, equipped with the subspace topology as subsets of $M_{n\times n}(\mathbb{R})$. Then G is a* topological group*, i.e., G is a topological space and a group, and the topology and the group structure are compatible in the sense that*

- *The multiplication map $G \times G \xrightarrow{\mu} G$ is continuous, and*
- *the map $G \to G$, $g \mapsto g^{-1}$ is continuous.*

Proof. We discussed continuity of the inverse map in Example 1.15. To prove continuity of the multiplication map μ, we consider the commutative diagram

$$\begin{tikzcd} G \times G \arrow[r, "\mu"] \arrow[d, "i\times i"'] & G \arrow[d, "i"] \\ M_{n\times n}(\mathbb{R}) \times M_{n\times n}(\mathbb{R}) \arrow[r, "m"] & M_{n\times n}(\mathbb{R}) \end{tikzcd}$$

where i is the inclusion map, and m is matrix multiplication which is continuous by Example 1.6. It might be tempting to argue that μ is the restriction of the continuous map m, and hence it is continuous by Lemma 1.16. However, that assumes that $G \times G$ is equipped with its *subspace topology* as a subset of $M_{n\times n}(\mathbb{R}) \times M_{n\times n}(\mathbb{R})$, rather than as equipped with the *product topology*. Proving that these topologies in fact agree is one way to finish the proof.

Alternatively, using Lemma 1.19, we argue that the map $i \times i \colon G \times G \to M_{n\times n}(\mathbb{R}) \times M_{n\times n}(\mathbb{R})$ is continuous since its component maps are: the first component map is the composition of the continuous maps

$$G \times G \xrightarrow{p_1} G \xrightarrow{i} M_{n\times n}(\mathbb{R})$$

and hence continuous; similarly for the second component map. Hence $m \circ (i \times i)$ is continuous, which equals $i \circ \mu$ by the commutativity of the diagram. It follows that μ is continuous by the criterion for continuity of a map to a subspace 1.14 $\square$

1.2.3 Quotient topology.

Definition 1.22. Let X be a topological space and let $\sim$ be an equivalence relation on X. We denote by $X/\sim$ be the set of equivalence classes and by

$$p \colon X \to X/\sim \qquad x \mapsto [x]$$

be the projection map that sends a point $x \in X$ to its equivalence class $[x]$. The *quotient topology* on $X/\sim$ is given by the collection of subsets

$$\mathcal{T} = \{U \subset X/\sim \mid p^{-1}(U) \text{ is an open subset of } X\}.$$

The set $X/\sim$ equipped with the quotient topology is called the *quotient space*.

The quotient topology is often used to construct a topology on a set Y which is not a subset of some Euclidean space $\mathbb{R}^n$, or for which it is not clear how to construct a metric. If there is a surjective map

$$p \colon X \longrightarrow Y$$

from a topological space X, then Y can be identified with the quotient space $X/\sim$, where the equivalence relation is given by $x \sim x'$ if and only if $p(x) = p(x')$. In particular, $Y = X/\sim$ can be equipped with the quotient topology. Here are important examples.

Example 1.23. (Examples of quotient spaces).

1. Let A be a subset of a topological space X. Define a equivalence relation $\sim$ on X by $x \sim y$ if $x = y$ or $x, y \in A$. We use the notation X/A for the quotient space $X/\sim$. A concrete example is provided by D^n/S^{n-1}, which is homeomorphic to the sphere S^n, as we will see later.

2. The *real projective space of dimension n* is the set

$$\mathbb{RP}^n := \{1\text{-dimensional subspaces of } \mathbb{R}^{n+1}\}.$$

The map

$$S^n \longrightarrow \mathbb{RP}^n \qquad \mathbb{R}^{n+1} \ni v \mapsto \text{subspace generated by } v$$

is surjective, leading to the identification

$$\mathbb{RP}^n = S^n/(v \sim \pm v),$$

and the quotient topology on $\mathbb{RP}^n$.

3. Similarly, working with complex vector spaces, we obtain a quotient topology on the the *complex projective space*

$$\mathbb{CP}^n := \{1\text{-dimensional subspaces of } \mathbb{C}^{n+1}\} = S^{2n+1}/(v \sim zv), \qquad z \in S^1$$

4. Generalizing, we can consider the *Grassmann manifold*

$$G_k(\mathbb{R}^{n+k}) := \{k\text{-dimensional subspaces of } \mathbb{R}^{n+k}\}.$$

There is a surjective map

$$V_k(\mathbb{R}^{n+k}) = \{(v_1, \dots, v_k) \mid v_i \in \mathbb{R}^{n+k},\ v_i\text{'s are orthonormal}\} \quad \twoheadrightarrow \quad G_k(\mathbb{R}^{n+k})$$

given by sending $(v_1, \dots, v_k) \in V_k(\mathbb{R}^{n+k})$ to the k-dimensional subspace of $\mathbb{R}^{n+k}$ spanned by the v_i's. Hence the subspace topology on the Stiefel manifold $V_k(\mathbb{R}^{n+k}) \subset \mathbb{R}^{(n+k)k}$ gives a quotient topology on the Grassmann manifold $G_k(\mathbb{R}^{n+k}) = V_k(\mathbb{R}^{n+k})/\sim$. The same construction works for the complex Grassmann manifold $G_k(\mathbb{C}^{n+k})$.

As the example 1.25(1) shows, a quotient space $X/\sim$ might be homeomorphic to a topological space Z constructed in a different way. To establish the homeomorphism between $X/\sim$ and Z, we need to construct continuous maps

$$f\colon X/\sim \longrightarrow Z \qquad g\colon Z \to X/\sim$$

that are inverse to each other. The next lemma shows that it is easy to check continuity of the map f, the map *out of the quotient space*.

Lemma 1.24. (Continuity criterion for a map out of a quotient space).

- *The projection map $p\colon X \to X/\sim$ is continuous.*

- *A map $f\colon X/\sim \to Z$ to a topological space Z is continuous if and only if the composition $X \xrightarrow{p} X/\sim \xrightarrow{f} Z$ is continuous.*

Proof: homework

As we will see in the next section, there are many situations where the continuity of the inverse map for a continuous bijection f is automatic. So in the examples below, and for the exercises in this section, we will defer checking the continuity of f^{-1} to that section.

Example 1.25. (1) We claim that the quotient space $[-1,+1]/\{\pm 1\}$ is homeomorphic to S^1 via the map $f\colon [-1,+1]/\{\pm 1\} \to S^1$ given by $[t] \mapsto e^{\pi i t}$. Geometrically speaking, the map f wraps the interval $[-1,+1]$ once around the circle. Here is a picture.

It is easy to check that the map f is a bijection. To see that f is continuous, consider the composition

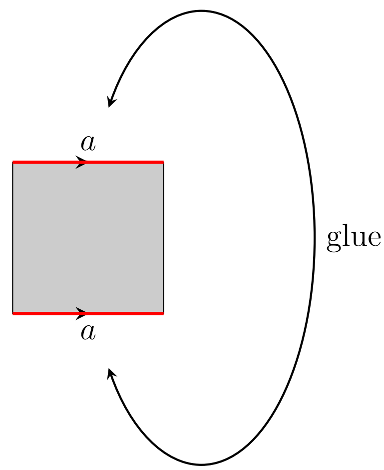
$$[-1,+1] \xrightarrow{p} [-1,+1]/\{\pm 1\} \xrightarrow{f} S^1 \xrightarrow{i} \mathbb{C} = \mathbb{R}^2,$$

where p is the projection map and i the inclusion map. This composition sends $t \in [-1,+1]$ to $e^{\pi i t} = (\cos \pi t, \sin \pi t) \in \mathbb{R}^2$. By Lemma 1.19 it is a continuous function, since its component functions $\sin \pi t$ and $\cos \pi t$ are continuous functions. By Lemma 1.24 the continuity of $i \circ f \circ p$ implies the continuity of $i \circ f$, which by Lemma 1.14 implies the continuity of f. As mentioned above, we'll postpone the proof of the continuity of the inverse map f^{-1} to the next section.

(2) More generally, D^n/S^{n-1} is homeomorphic to S^n. (proof: homework)

(3) Consider the quotient space of the square $[-1,+1]\times[-1,+1]$ given by identifying $(s,-1)$ with $(s,1)$ for all $s \in [-1,1]$. It can be visualized as a square whose top edge is to be glued with its bottom edge. In the picture below we indicate that identification by

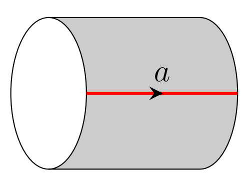
labeling those two edges by the same letter.



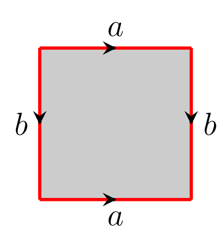
The quotient $([-1,+1]\times[-1,+1])\,/(s,-1)\sim(s,+1)$ is homeomorphic to the cylinder

$$C=\{(x,y,z)\in\mathbb{R}^3 \mid x\in[-1,+1], y^2+z^2=1\}.$$

The proof is essentially the same as in (1). A homeomorphism from the quotient space to C is given by $f([s,t])=(s,\sin\pi t,\cos\pi t)$. The picture below shows the cylinder C with the image of the edge a indicated.



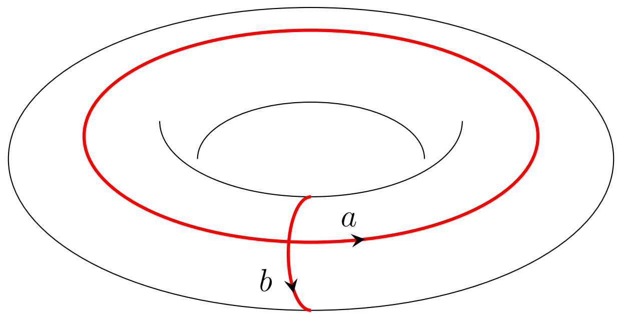
(4) Consider again the square, but this time using an equivalence relations that identifies more points than the one in the previous example. As before we identify $(s,-1)$ and $(s,1)$ for $s\in[-1,1]$, and in addition we identify $(-1,t)$ with $(1,t)$ for $t\in[-1,1]$. Here is the picture, where again corresponding points of edges labeled by the same letter are to be identified.



We claim that the quotient space is homeomorphic to the torus

$$T:=\{x\in\mathbb{R}^3 \mid d(x,K)=d\},$$

where $K = \{(x_1, x_2, 0) \mid x_1^2 + x_2^2 = 1\}$ is the unit circle in the xy-plane and $0 < d < 1$ is a real number (see) via a homeomorphism that maps the edges of the square to the loops in T indicated in the following picture below.



Exercise: prove this by writing down an explicit map from the quotient space to T, and arguing that this map is a continuous bijection (as always in this section, we defer the proof of the continuity of the inverse to the next section).

(5) We claim that the quotient space $D^n/\sim$ with equivalence relation generated by $v \sim -v$ for $v \in S^{n-1} \subset D^n$ is homeomorphic to the real projective space $\mathbb{RP}^n$. More precisely, let $f\colon D^n \to S^n$ be the embedding of the n-disk as the upper hemisphere of S^n. Explicitly, $f(x)$ for $x = (x_1, \dots, x_n)$ is given by the formula

$$f(x_1, \dots, x_n) := (x_1, \dots, x_n, \sqrt{1 - (x_1^2 + \cdots + x_n^2)})$$

Lemma 1.26. *The map* $\bar{f}\colon D^n/\sim \;\to\; \mathbb{RP}^n = S^n/\sim$ *given by* $[x] \mapsto [f(x)]$ *is a continuous bijection.*

With more tools at our disposal in the next section we will argue that this map is in fact a homeomorphism.

Proof. To check that $\bar{f}$ is well-defined, we note that get identified in D^n are $x \sim -x$ for $x \in \partial D^n = S^{n-1}$. For such x, $f(x) = (x_1, \dots, x_n, 0)$ and $f(-x) = -(x_1, \dots, x_n, 0)$, showing that $\bar{f}$ is well-defined.

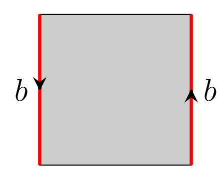
Next we argue that $\bar{f}$ is continuous. The map f is continuous since its components are continuous functions. By construction of $\bar{f}$ we have the commutative diagram

$$\begin{tikzcd}
D^n \arrow[r, "f"] \arrow[d, "p_1"'] & S^n \arrow[d, "p_2"'] \\
D^n \sim \arrow[r, "\bar{f}"] & S^n/\sim
\end{tikzcd}$$

where the vertical maps are the projection maps. Since f is continuous, so is the composition $p_2 \circ f = p_1 \circ \bar{f}$, and hence $\bar{f}$ (a map out of a quotient space is continuous if and only if its pre-composition with the projection map is).

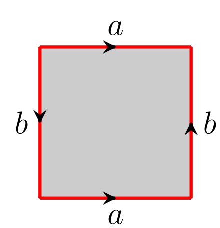
The map f provides a bijection between D^n and the upper hemisphere of S^n (including the equator); the inverse map is given by sending a point $(x_1, \dots, x_{n+1})$ in the upper hemisphere to $(x_1, \dots, x_n)$. Since every equivalence class in S^n can be represented by a point in the upper hemisphere, this implies that $\bar{f}$ is surjective. Since the only points in the upper hemisphere that are identified by the equivalence relation on S^n are antipodal points on the equator, this implies that $\bar{f}$ is injective. $\square$

(6) The quotient space $[-1, 1] \times [-1, 1]/ \sim$ with the equivalence relation generated by $(-1, t) \sim (1, -t)$ is represented graphically by the following picture.



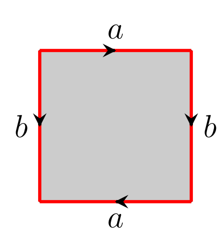
This topological space is called the *Möbius band*. It is homeomorphic to a subspace of $\mathbb{R}^3$ shown by the following picture

(7) The quotient space of the square by edge identifications given by the picture



is the *Klein bottle*. It is harder to visualize, since it is not homeomorphic to a subspace of $\mathbb{R}^3$ (which can be proved by the methods of algebraic topology).

(8) The quotient space of the square given by the picture



is homeomorphic to the real projective plane $\mathbb{RP}^2$. Exercise: prove this (hint: use the statement of example (5)). Like the Klein bottle, it is challenging to visualize the real projective plane, since it is not homeomorphic to a subspace of $\mathbb{R}^3$.

1.3 Properties of topological spaces

In the previous subsection we described a number of examples of topological spaces X, Y that we claimed to be homeomorphic. We typically constructed a bijection $f\colon X \to Y$ and argued that f is continuous. However, we did not finish the proof that f is a homeomorphism, since we defered the argument that the inverse map $f^{-1}\colon Y \to X$ is continuous. We note that not every continuous bijection is a homeomorphism.

For example, the map

$$f\colon [0,1) \longrightarrow S^1 \subset \mathbb{R}^2 = \mathbb{C} \qquad \text{given by} \qquad t \mapsto e^{2\pi i t} \tag{1.27}$$

is a bijection. It is the restriction of the map $\widetilde{f}\colon \mathbb{R} \to \mathbb{R}^2$ given by the same formula; $\widetilde{f}$ is continuous since its component functions $\cos 2\pi i t$ and $\sin 2\pi i t$ are continuos, and hence f is continuous (with the respect to the subspace topology on $[0,1) \subset \mathbb{R}$ and $S^1 \subset \mathbb{R}^2$). The inverse map $g\colon S^1 \to [0,1)$ is *not continuous*, since $[0,1/2) \subset [0,1)$ is open, but $g^{-1}([0,1/2)) = f([0,1/2))$ consists of the lower semicircle (the intersection of the lower open halfplane $\{(x,y) \in \mathbb{R}^2 \mid y < 0\}$ with S^1) and the point $(1,0)$) which we claim is not an open subset of S^1. To prove this, assume that $f([0,1/2))$ is in fact open in the subspace topology, i.e., $f([0,1/2)) = S^1 \cap U$ for some open subset $U \subset \mathbb{R}^2$. Since $(1,0) \in U$ and U is open, there is radius $r > 0$ such that the ball $B_r((1,0))$ is contained in U, and hence $S^1 \cap B_r((1,0)) \subset S^1 \cap U = f([0,1/2))$. This is the desired contradiction, since no point with positive y coordinate belongs to $f([0,1/2))$.

Fortunately, there are situations where the continuity of the inverse map is automatic as the following proposition shows.

Proposition 1.28. (Criterion for continuity of inverse). *Let $f\colon X \to Y$ be a continuous bijection. Then f is a homeomorphism provided X is compact and Y is Hausdorff.*

This result does not apply to the function (1.27) since the domain of the map is noncompact.

The goal of this section is to define these notions, prove the proposition above, and to give a tools to recognize that a topological space is compact and/or Hausdorff.

1.3.1 Hausdorff spaces

Definition 1.29. Let X be a topological space, $x_i \in X$, $i = 1, 2, \dots$ a sequence in X and $x \in X$. Then *x is a limit of the x_i's* if for any open subset $U \subset X$ containing x there is some N such that $x_i \in U$ for all $i \geq N$.

Caveat: If X is a topological space with the indiscrete topology 1.9, *every point* is the limit of every sequence. There is at most one limit of the x_i if the topological space has the following property:

Definition 1.30. A topological space X is *Hausdorff* if for every $x, y \in X$, $x \neq y$, there are disjoint open subsets $U, V \subset X$ with $x \in U$, $y \in V$.

Lemma 1.31. *The Euclidean space $\mathbb{R}^n$ is Hausdorff. More generally, any subspace $U \subset \mathbb{R}^n$ is Hausdorff.*

Proof. Let $x, y \in U$ with $x \neq y$. Then the balls $B_r(x)$, $B_r(y)$ are open subsets in $\mathbb{R}^n$ which are disjoint if we choose the radius r small enough; for example the choice $r := \operatorname{dist}(x, y)/2$ works. Then $B_r(x) \cap U$ and $B_r(y) \cap U$ are disjoint open neighborhoods of x resp. y in U, showing that U is Hausdorff. $\square$

Lemma 1.32. *Let X be a topological space and A a closed subspace of X. If $x_n \in A$ is a sequence with limit x, then $x \in A$.*

Proof. Assume $x \notin A$. Then x is a point in the open subset $X \setminus A$ and hence by the definition of limit, all but finitely many elements x_n must belong to $X \setminus A$, contradicting our assumptions. $\square$

1.3.2 Compact spaces

Definition 1.33. An *open cover* of a topological space X is a collection of open subsets of X whose union is X. If for every open cover of X there is a finite subcollection which also covers X, then X is called *compact*.

Some books (like Munkres' *Topology*) refer to open covers as *open coverings*, while newer books (and wikipedia) seem to prefer the above terminology, probably for the same reasons as me: to avoid confusion with *covering spaces*, a notion we'll introduce soon.

Example 1.34. (Example of a non-compact space.) The real line $\mathbb{R}$ with the metric topology is non-compact, since the collection of open intervals $(n-1, n+1) \subset \mathbb{R}$ for $n \in \mathbb{Z}$ form an open cover of $\mathbb{R}$, but it does not admit a finite subcover. Indeed, removing just any one interval $(k-1, k+1)$ from the cover, this is no longer a cover of $\mathbb{R}$, since the point $k \in \mathbb{R}$ is not contained in any interval $(n-1, n+1)$ for $n \neq k$.

While it is easy to show that a topological space X is non-compact (by finding an open cover without a finite subcover), showing that X is compact from the definition of compactness is hard: you need to ensure that *every open cover* has a finite subcover. That sounds like a lot of work... Fortunately, there is a very simple classical characterization of compact subspaces of Euclidean spaces, see Theorem 1.37.

Next we will prove some useful properties of compact spaces and maps between them, which will be the essential ingredients of the proof of Proposition 1.28 as well Proposition **??** which guarantees the existence of minima and maxima of a continuous function $f\colon X \to \mathbb{R}$ on a compact space X.

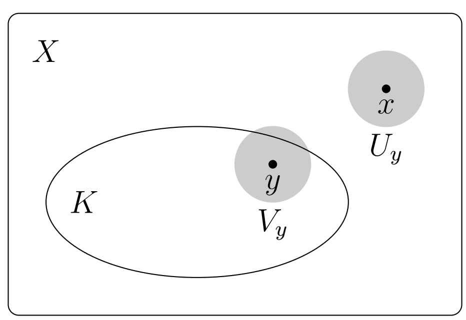
Lemma 1.35. *If $f\colon X \to Y$ is a continuous map and X is compact, then the image $f(X)$ is compact. In particular, if X is compact, then any quotient space $X/\sim$ is compact, since the projection map $X \to X/\sim$ is continuous with image $X/\sim$.*

Proof. To show that $f(X)$ is compact assume that $\{U_a\}$, $a \in A$ is an open cover of the subspace $f(X)$. Then each U_a is of the form $U_a = V_a \cap f(X)$ for some open subset $V_a \in Y$. Then $\{f^{-1}(V_a)\}$, $a \in A$ is an open cover of X. Since X is compact, there is a finite subset A' of A such that $\{f^{-1}(V_a)\}$, $a \in A'$ is a cover of X. This implies that $\{U_a\}$, $a \in A'$ is a finite cover of $f(X)$, and hence $f(X)$ is compact. □

Lemma 1.36. *1. If K is a closed subspace of a compact space X, then K is compact.*

2. If K is compact subspace of a Hausdorff space X, then K is closed.

Proof. The proof of part (1) is a homework problem. To prove (2), we need to show that $X \setminus K$ is open. So let $x \in X \setminus K$, and we aim to find an open neighborhood U of x which is contained in $U \setminus K$. Since X is Hausdorff, and $x \notin K$, for each $y \in K$ there are disjoint open neighborhoods V_y of y and U_y of x. This situation is illustrated in the following figure.



Then $V_y \cap K$ is an open subset of K, and the collection of subsets $\{V_y \cap K\}_{y \in K}$ is an open cover of K. The compactness of K guarantees that this contains a finite subcover, i.e., there are points $y_1, \dots, y_n \in K$ such that $\bigcup_{i=1,\dots,n} V_{y_i} \cap K = K$. In particular, $K \subset \bigcup_{i=1,\dots,n} V_{y_i}$. Then $U := \bigcap_{i=1,\dots,n} U_{y_i}$ is an open subset containing x; by construction,

$$U \cap \bigcup_{i=1,\dots,n} V_{y_i} = \emptyset \qquad \text{and hence} \qquad U \cap K = \emptyset,$$

which proves that U is an open subset in $U \setminus K$. □

Proof of Proposition 1.28. We need to show that the map $g\colon Y \to X$ inverse to f is continuous, i.e., that $g^{-1}(U) = f(U)$ is an open subset of Y for any open subset U of X. Equivalently (by passing to complements), it suffices to show that $g^{-1}(C) = f(C)$ is a closed subset of Y for any closed subset C of C.

Now the assumption that X is compact implies that the closed subset $C \subset X$ is compact by part (1) of Lemma 1.36 and hence $f(C) \subset Y$ is compact by Lemma 1.35. The assumption that Y is Hausdorff then implies by part (2) of Lemma 1.36 that $f(C)$ is closed. □

Now we want to apply Proposition 1.28 to show that the continuous bijections that we constructed in Example 1.25 and Lemma 1.26 are in fact homeomorphism. This requires that we are able to show that the domain of the map is compact, which is often done using the the following compactness criterion for subspaces of Euclidean space $\mathbb{R}^n$.

Theorem 1.37. (Heine-Borel Theorem) *A subspace $K \subset \mathbb{R}^n$ is compact if and only if K is a closed subset of $\mathbb{R}^n$ and bounded, i.e., there is some $R > 0$ such that K is contained in the ball $B_R(0)$ of radius R around the origin.*

With this tool in hand, we now revisit Example 1.25(1) and (5):

Example 1.25(1) We have constructed a continuous bijection $f\colon [-1,+1]/\{\pm 1\} \longrightarrow S^1$. The domain of f is compact since $[-1,+1]$ is a closed and bounded subset of $\mathbb{R}$ and hence compact by the Heine-Borel Theorem. It follows that the quotient space $[-1,+1]/\{\pm 1\}$ is compact by Lemma 1.35. The codomain of f is the circle S^1 which is Hausdorff as a subspace of $\mathbb{R}^2$ by Lemma 1.31. Hence f is a homeomorphism by Proposition 1.28.

Example 1.25(5) We have constructed a continuous bijection $f\colon D^n/\sim \longrightarrow \mathbb{RP}^n$. The domain is compact, since it is a quotient of the closed bounded subspace $R^n \subset \mathbb{R}^n$. So it remains to show that the codomain $\mathbb{RP}^n$ is Hausdorff. It might be tempting to argue that $\mathbb{RP}^n$ is Hausdorff, since it is a quotient of the Hausdorff space $S^n \subset \mathbb{R}^n$. Alas, Hausdorff is not a property inherited by quotient spaces as the example below shows. So a more detailed argument is needed.

Lemma 1.38. *The projective space $\mathbb{RP}^n$ is Hausdorff.*

Proof. Let $p\colon S^n \to \mathbb{RP}^n$ be the projection map. For $x \in S^n$ let $[x] = p(x) \in \mathbb{RP}^n$ be the equivalence class of x, consisting of the pair of antipodal points $\{x, -x\} \subset S^n$. If $[x] \neq [y] \in \mathbb{RP}^n$, then x, $-x$, y, $-y$ are four distinct points in S^n. Hence for sufficiently small r the four balls of radius r around these points are pairwise disjoint. In particular,

$$U := (B_r(x) \cup B_r(-x)) \cap S^n \qquad \text{and} \qquad V := (B_r(y) \cup B_r(-y)) \cap S^n$$

are disjoint open subsets of S^n. Then $p(U)$, $p(V)$ are disjoint open subsets of $\mathbb{RP}^n$ since $p^{-1}(p(U)) = U$ and $p^{-1}(p(V)) = V$. □

Example 1.39. (Example of a Hausdorff space a quotient of which is not Hausdorff). The interval $(-1,1)$ is a subspace of $\mathbb{R}$ and so we can form the quotient space $X := \mathbb{R}/(-1,1)$ where all points belonging to $(-1,1)$ are identified. We claim that X is not Hausdorff; more precisely, we claim that the points $[-1],[1] \in X$ do not have disjoint open neighborhoods $U \ni [-1]$, $V \ni [1]$. To prove this, assume that there are disjoint open neighborhoods. Then their preimages $p^{-1}(U)$, $p^{-1}(V)$ under the projection map $p\colon \mathbb{R} \to X$ are disjoint open subsets of $\mathbb{R}$ with $-1 \in p^{-1}(U)$ and $1 \in p^{-1}(V)$. Due to these being open subsets of $\mathbb{R}$, it follows that $p^{-1}(U)$ must contain some point $x \in (-1,1)$ and that $p^{-1}(V)$ must contain some point $y \in (-1,1)$. It follows that $U \ni p(x) = p(y) \in V$ contradicting the assumption that U and V are disjoint.

The proof of the Heine-Borel Theorem is based on the following two results.

Lemma 1.40. *A closed interval $[a,b]$ is compact.*

This lemma has a short proof that can be found in any pointset topology book, e.g., [Mu].

Theorem 1.41. *If $X_1,\dots,X_n$ are compact topological spaces, then their product $X_1 \times \cdots \times X_n$ is compact.*

For a proof see e.g. [Mu, Ch. 3, Thm. 5.7]. The statement is true more generally for a product of *infinitely many* compact space (as discussed in [Mu, p. 113], the correct definition of the product topology for infinite products requires some care), and this result is called *Tychonoff's Theorem*, see [Mu, Ch. 5, Thm. 1.1].

Proof of the Heine-Borel Theorem. Let K be a compact subspace of $\mathbb{R}^n$. Then K is closed by Lemma 1.36(2). The collection $B_r(0) \cap K$, $r \in (0,\infty)$, is an open cover of K. By compactness, K is covered by a *finite* number of these balls; if R is the maximum of the radii of these finitely many balls, this implies $K \subset B_R(0)$, i.e., K is bounded.

Conversely, let $K \subset \mathbb{R}^n$ be closed and bounded, say $K \subset B_r(0)$. We note that $B_r(0)$ is contained in the n-fold product

$$P := [-r,r] \times \cdots \times [-r,r] \subset \mathbb{R}^n$$

which is compact by Theorem 1.41. So K is a closed subset of P and hence compact by Lemma 1.36(1). $\square$

Here is another interesting consequence of (the easier part of) the Heine-Borel Theorem.

Proposition 1.42. *If $f\colon X \to \mathbb{R}$ is a continuous function on a compact space X, then f has a maximum and a minimum.*

Proof. $K = f(X)$ is a compact subset of $\mathbb{R}$. Hence K is bounded, and thus K has an infimum $a := \inf K \in \mathbb{R}$ and a supremum $b := \sup K \in \mathbb{R}$. The infimum (resp. supremum) of K is the limit of a sequence of elements in K; since K is closed (by Lemma 1.36 (2)), the limit points a and b belong to K by Lemma 1.32. In other words, there are elements $x_{min}, x_{max} \in X$ with $f(x_{min}) = a \leq f(x)$ for all $x \in X$ and $f(x_{max}) = b \geq f(x)$ for all $x \in X$. □

1.3.3 Connected spaces

Definition 1.43. A topological space X is *connected* if it can't be written as decomposed in the form $X = U \cup V$, where U, V are two non-empty disjoint open subsets of X.

For example, if a, b, c, d are real numbers with $a < b < c < d$, consider the subspace $X = (a, b) \amalg (c, d) \subset \mathbb{R}$. The topological space X is not connected, since $U = (a, b)$, $V = (c, d)$ are open disjoint subsets of X whose union is X. This remains true if we replace the open intervals by closed intervals. The space $X' = [a, b] \amalg [c, d]$ is not connected, since it is the disjoint union of the subsets $U' = [a, b]$, $V' = [c, d]$. We want to emphasize that while U' and V' are not open as subsets of $\mathbb{R}$, they are *open subsets of* X', since they can be written as

$$U' = (-\infty, c) \cap X' \qquad V' = (b, \infty) \cap X',$$

showing that they are open subsets for the subspace topology of $X' \subset \mathbb{R}$.

Lemma 1.44. *Any interval I in $\mathbb{R}$ (open, closed, half-open, bounded or not) is connected.*

Proof. Using proof by contradiction, let us assume that I has a decomposition $I = U \cup V$ as the union of two non-empty disjoint open subsets. Pick points $u \in U$ and $v \in V$, and let us assume $u < v$ without loss of generality. Then

$$[u, v] = U' \cup V' \qquad \text{with} \qquad U' := U \cap [u, v] \quad V' := U \cap [u, v]$$

is a decomposition of $[u, v]$ as the disjoint union of non-empty disjoint open subsets U', V' of $[u, v]$. We claim that the supremum $c := \sup U'$ belongs to both, U' and V', thus leading to the desired contradiction. Here is the argument.

- Assuming that c doesn't belong to U', for any $\epsilon > 0$, there must be some element of U' belonging to the interval $(c - \epsilon, c)$, allowing us to construct a sequence of elements $u_i \in U'$ converging to c. This implies $c \in U'$ by Lemma 1.32, since U' is a *closed* subspace of $[u, v]$ (its complement V' is open).
- By construction, every $x \in [u, v]$ with $x > c = \sup U'$ belongs to V'. So we can construct a sequence $v_i \in V'$ converging to c. Since V' is a closed subset of $[u, v]$, we conclude $c \in V'$.

□

Theorem 1.45. (Intermediate Value Theorem) *Let X be a connected topological space, and $f\colon X \to \mathbb{R}$ a continuous map. If elements $a, b \in \mathbb{R}$ belong to the image of f, then also any real number c between a and b belongs to the image of f.*

Proof. Assume that c is not in the image of f. Then $X = f^{-1}(-\infty, c) \cup f^{-1}(c, \infty)$ is a decomposion of X as a union of non-empty disjoint open subsets. □

There is another notion, closely related to the notion of connected topological space, which might be easier to think of geometrically.

Definition 1.46. A topological space X is *path connected* if for any points $x, y \in X$ there is a path connecting them. In other words, there is a continuous map $\gamma\colon [a, b] \to X$ from some interval to X with $\gamma(a) = x$, $\gamma(b) = y$.

Lemma 1.47. *Any path connected topological space is connected.*

Proof. Using proof by contradiction, let us assume that the topological space X is path connected, but not connected. So there is a decomposition $X = U \cup V$ of X as the union of non-empty open subsets $U, V \subset X$. The assumption that X is path connected allows us to find a path $\gamma\colon [a, b] \to X$ with $\gamma(a) \in U$ and $\gamma(b) \in V$. Then we obtain the decomposition

$$[a, b] = f^{-1}(U) \cup f^{-1}(V)$$

of the interval $[a, b]$ as the disjoint union of open subsets. These are non-empty since $a \in f^{-1}(U)$ and $b \in f^{-1}(V)$. This implies that $[a, b]$ is not connected, the desired contradiction. □

For typical topological spaces we will consider, the properties "connected" and "path connected" are equivalent. But here is an example known as the *topologist's sine curve* which is connected, but not path connected, see [Mu, Example 7, p. 156]. It is the following subspace of $\mathbb{R}^2$:

$$X = \{(x, \sin\frac{1}{x}) \in \mathbb{R}^2 \mid 0 < x < 1\} \cup \{(0, y) \in \mathbb{R}^2 \mid -1 \leq y \leq 1\}.$$

References

[Mu] Munkres, James R. *Topology: a first course*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975. xvi+413 pp.

2 Topological manifolds

The purpose of this section is to provide interesting examples of topological spaces and homeomorphisms between them. There are many examples of "weird" topological spaces. There are non-Hausdorff spaces (they don't have well-defined limits) or the topologist's sine curve, which is connected, but not path connected. While there is a huge literature concerning pathological topological spaces, I must admit that I find those examples most interesting that "show up in nature". For example, topological spaces that appear as "configuration spaces" or "phase spaces" of physical systems. Often these are a particularly nice kind of topological space known as *manifold*.

There is much to say about manifolds. For example, you can find the text books *Introduction to topological manifolds* and *Introduction to smooth manifolds* on the reserved book shelf for this course. For this section, our focus is to discuss manifolds of dimension 2. Unlike higher dimensional manifolds, we can represent manifolds of dimension 2 by pictures, which greatly helps the intuition about these objects.

2.1 Definition and basic examples of manifolds

Definition 2.1. A *manifold of dimension n* or *n-manifold* is a topological space X which is locally homeomorphic to $\mathbb{R}^n$, that is, every point $x \in X$ has an open neighborhood U which is homeomorphic to an open subset V of $\mathbb{R}^n$. Moreover, it is useful and customary to require that X is Hausdorff (see Definition 1.30) and *second countable*, which means that the topology of X has a countable basis.

In most examples, the technical conditions of being Hausdorff and second countable are easy to check, since these properties are inherited by subspaces.

Homework 2.2. Show that a subspace of a Hausdorff space is Hausdorff. Show that a subspace of a second countable space is second countable.

Examples of manifolds.

1. Any open subset $U \subset \mathbb{R}^n$ is an n-manifold. The technical condition of being a second countable Hausdorff space is satisfied for U as a subspace of the second countable Hausdorff space $\mathbb{R}^n$; a countable basis for the topology on $\mathbb{R}^n$ is provided by the collection of balls $B_r(x)$, for which the radius r as well as all components of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ are rational numbers.

2. The n-sphere $S^n := \{x \in \mathbb{R}^n \mid ||x|| = 1\}$ is an n-manifold. To prove this, let us look at the subsets

$$U_i^+ := \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid x_i > 0\} \subset S^n$$
$$U_i^- := \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid x_i < 0\} \subset S^n$$

We want to argue that the map

$$\phi_i^\pm \colon U_i^\pm \longrightarrow \mathring{D}^n \qquad \text{given by} \qquad \phi_i^\pm(x_0, \dots, x_n) := (x_0, \dots, x_{i-1}, \widehat{x}_i, x_{i+1}, \dots, x_n)$$

is a homeomorphism, where $\mathring{D}^n := \{(v_1, \dots, v_n) \in D^n \mid v_1^2 + \cdots + v_n^2 < 1\}$ is the open n-disk. It is easy to verify that the map

$$\mathring{D}^n \longrightarrow U_i^\pm \qquad v = (v_1, \dots, v_n) \mapsto (v_1, \dots, v_i, \pm\sqrt{1 - ||v||^2}, v_{i+1}, \dots, v_n)$$

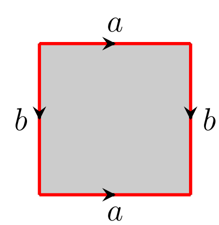
is in fact the inverse to $\phi_i^\pm$. Here $||v||^2 = v_1^2 + \cdots + v_n^2$ is norm squared of $v \in \mathring{D}^n$. Both maps, $\phi_i^\pm$ and its inverse, are continuous since all their components are continuous. This shows that $\phi_i^\pm$ is in fact a homeomorphism, and hence the n-sphere S^n is a manifold of dimension n.

Homework 2.3. Show that the product $X \times Y$ of manifold X of dimension m and a manifold Y of dimension n is a manifold of dimension $m+n$. Make sure to prove that $X \times Y$ is second countable and Hausdorff.

Homework 2.4. Show that the real projective space $\mathbb{RP}^n$ is manifold of dimension n. Make sure to prove that $\mathbb{RP}^n$ is second countable (we know it is Hausdorff by Lemma 1.38.

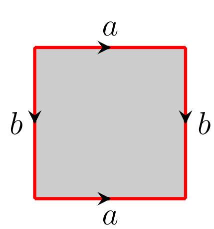
Examples of manifolds of dimension 2.

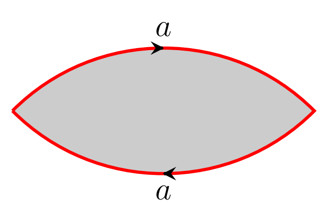
1. The 2-torus T can be described as subspace of $\mathbb{R}^3$, as the product $S^1 \times S^1$ and as the quotient of the square $[0,1] \times [0,1]$ by the identifying its edges as indicated in the following picture (see Example 1.25(4)):



   From the description of the torus as the product $S^1 \times S^1$ and Lemma **??** it follows that the torus is a manifold of dimension 2.

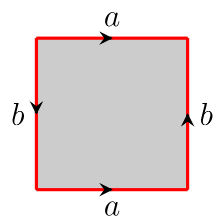
2. The real projective plane $\mathbb{RP}^2$. We recall from Example 1.25 (8) and Lemma 1.26 that $\mathbb{RP}^2$ is homeomorphic to the quotient spaces of the square resp. disk by identifying edges as indicated by the following pictures.





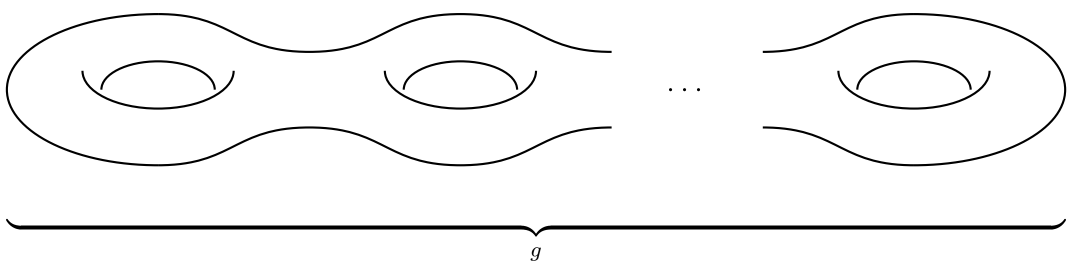
We prefer to draw the disk D^2 as a bigon here, since our goal is to describe *all* compact connected 2-manifolds as quotients of polygons by suitably identifying edges. We think of the bigon as a polygon with two vertices and two edges.

3. In Example 1.25(7) we defined the Klein bottle K as the quotient of the square with the identification of edges given by the following picture.



It is not hard to verify directly that K is a manifold of dimension 2 (draw open neighborhoods of a point in the interior of the square, on an edge of the square and of the one point of K represented by the vertices to convince yourself). Alternatively, we will see in Lemma **??** that the Klein bottle is homeomorphic to the connected sum $\mathbb{RP}^2 \# \mathbb{RP}^2$ of two copies of the projective plane $\mathbb{RP}^2$, which implies in particular that K is a 2-manifold.

4. The surface Σ_g of genus g is the subspace of $\mathbb{R}^3$ given by the following picture:



Here g is the number of "holes" of Σ_g. In particular Σ_1, the surface of genus 1, is the torus. By convention, the surface Σ_0 of genus 0 is the 2-sphere S^2. Since we have described the surface of genus g as a subspace of $\mathbb{R}^3$ given by a picture rather than a formula, it is impossible to give a precise argument that this subspace is locally homeomorphic to $\mathbb{R}^2$, but hopefully the picture makes this obvious at a heuristic level.

2.2 The connected sum construction

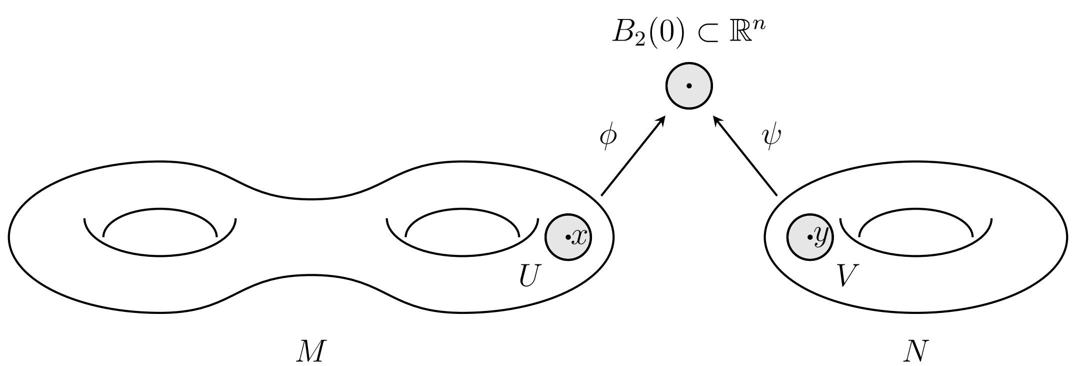
This construction produces a new manifold $M \# N$ of dimension n from two given manifolds M and N of dimension n. The manifold $M \# N$ is called the *connected sum* of M and N. The construction proceeds as follows. First we make some choices:

- We pick points $x \in M$ and $y \in N$.
- We pick a homeomorphism ϕ between an open neighborhood U of x and the open ball $B_2(0)$ of radius 2 around the origin $0 \in \mathbb{R}^n$. Similarly, we pick a homeomorphism $\psi\colon V \xrightarrow{\approx} B_2(0)$ where $V \subset N$ is an open neighborhood of $y \in N$.

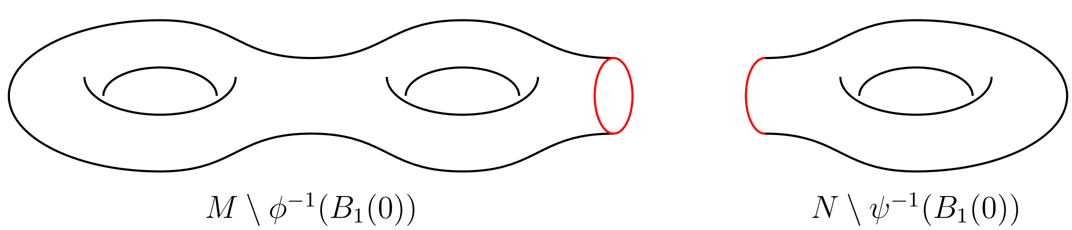
The existence of homeomorphisms ϕ, ψ with these properties follows from the assumption that M, N are manifolds of dimension n. This implies that there is an open neighborhood $U' \subset M$ of x and a homeomorphism ϕ' between U' and an open subset $V' \subset \mathbb{R}^n$. Composing ϕ by a translation in $\mathbb{R}^n$ we can assume that $\phi(x) = 0 \in \mathbb{R}^n$. Since V' is open, there is some $\epsilon > 0$ such that the open ball $B_\epsilon(0)$ of radius ϵ around $0 \in \mathbb{R}^n$ is contained in V'. Then restricting ϕ' to $U := (\phi')^{-1}(B_\epsilon(0)) \subset M$ gives a homeomorphism between U and $B_\epsilon(0)$. Then the composition

$$U \xrightarrow[\approx]{\phi'_U} B_\epsilon(0) \xrightarrow[\approx]{\text{multiplication by } 2/\epsilon} B_2(0)$$

is the desired homeomorphism ϕ between a neighborhood U of $x \in M$ and $B_2(0) \subset \mathbb{R}^n$. Analogously, we construct the homeomorphism ψ. Here is a picture illustrating the situation.



The next step is to remove the open disc $\phi^{-1}(B_1(0))$ from the manifold M and the open disc $\psi^{-1}(B_1(0))$ from the manifold N. The following picture shows the resulting topological spaces $M\backslash\phi^{-1}(B_1(0))$ and $N\backslash\psi^{-1}(B_1(0))$. Here the red circles mark the points corresponding to the sphere $S^{n-1} \subset B_2(0)$ via the homeomorphisms ϕ and ψ, respectively.



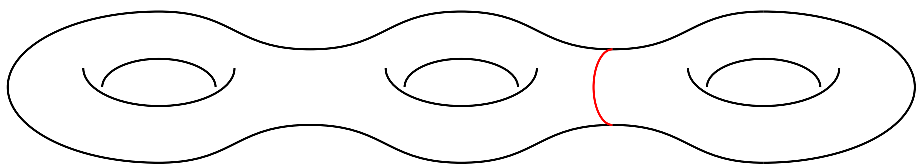
The final step is to pass to a quotient space of the union

$$M \setminus \phi^{-1}(B_1(0)) \quad \cup \quad N \setminus \psi^{-1}(B_1(0))$$

given by identifying points in $\phi^{-1}(S^{n-1})$ with their images under the homeomorphism

$$\phi^{-1}(S^{n-1}) \xrightarrow{\approx} \psi^{-1}(S^{n-1}) \qquad z \mapsto \psi^{-1}(\phi(z)).$$

The connected sum $M\#N$ is this quotient space. In terms of our pictures, the manifold $M\#N$ is obtained by gluing the two red circles, and is given by the following picture.



$M\#N$

Question: Is $M\#N$ independent of the choices made in its construction? A crucial ingredient of the construction of the connected sum $M\#N$ are the homeomorphisms $M \supset U \xrightarrow{\phi} B_2(0) \subset \mathbb{R}^n$ and $N \supset V \xrightarrow{\psi} B_2(0) \subset \mathbb{R}^n$. Since we remove in the first step of the construction the open disks $\phi^{-1}(B_1(0) \subset M$ and $\psi^{-1}(B_1(0) \subset N$, the set $M\#N$ will be different if we remove different disks.

Fact: Up to homeomorphism, the topological space $M\#N$ does not depend on these choices if "we are careful with orientations". Fortunately, for 2-dimensional manifolds, it is always independent of the choices.

Later this semester we will define what an orientation for a smooth manifold is (which is easier than defining an orientation for a topological manifold). We will restrict us to 2-manifolds, so orientations don't play a role, and we use the fact above for 2-manifolds without proof.

Example 2.5. (Examples of connected sums).

1. Our pictures above show that the connected sum $\Sigma_2\#T$ of the surface of genus two and the torus is homeomorphic to the surface of genus 3. More generally, it is clear from drawing appropriate pictures that the connected sum $\Sigma_g\#\Sigma_{g'}$ is homeomorphic to $\Sigma_{g+g'}$ of genus $g+g'$. It follows that

$$\underbrace{T\#T\#\dots\#T}_{g} \approx \Sigma_g.$$

Strictly speaking, we have mathematically defined what we mean by a surface of genus g only for $g = 1$ (it is the torus T) and for $g = 0$ (it is the sphere S^2). For $g > 1$, we have only drawn a picture of what we mean by a surface of genus g, and hence we can prove the statement $\Sigma_g \# \Sigma_{g'} \approx \Sigma_{g+g'}$ only at that level of precision: by drawing pictures. From mathematical point of view, we can (and will) view the above homeomorphism now as the *definition* of the surface of genus g.

2. The connected sum $X_k := \underbrace{\mathbb{RP}^2 \# \dots \# \mathbb{RP}^2}_{k}$ is a 2-manifold that, together with the surface of genus g, plays an important role in the Classification Theorem for compact connected 2-manifolds 2.6. Munkres refers to X_k as the *k-fold projective plane* [Mu, Definition on p. 462].

2.3 Classification of compact connected 2-manifolds

Theorem 2.6. (Classification of compact connected 2-manifolds.) *Every compact connected manifold of dimension 2 is homeomorphic to exactly one of the following manifolds:*

- *The surface of genus g, denoted Σ_g which is the connected sum $\underbrace{T \# \dots \# T}_{g}$ of g copies of the torus T, for $g > 0$, and the 2-sphere S^2 for $g = 0$.*
- *The connected sum $X_k = \underbrace{\mathbb{RP}^2 \# \dots \# \mathbb{RP}^2}_{k}$ of k copies of the real projective plane $\mathbb{RP}^2$, $k \geq 1$;*

In this class, we won't give a complete proof of this classification result, but we will introduce the techniques used for the proof of this theorem (see e.g., [Mu]), and we prove partial results. Like any classification result, the classification of 2-manifolds involves two quite distinct aspects:

(1) the proof that the 2-manifolds Σ_0, Σ_1, Σ_2,..., X_1, X_2,... are pairwise non-homeomorphic.

(2) the proof that any compact connected 2-manifold Σ is homeomorphic to a manifold on this list.

We will prove part (1) by introducing the Euler characteristic for compact 2-manifolds in section 2.4 and showing that this invariant can be used to show that Σ_g is not homeomorphic to $\Sigma_{g'}$ for $g \neq g'$ and that Σ_k is not homeomorphic to $\Sigma_{k'}$ for $k \neq k'$, see 2.14. But the Euler characteristic can't be used to show that Σ_g and X_{2g} are not homeomorphic, since they have the same Euler characteristic. In section **??** we define what it means for a 2-manifold to be *orientable* and we distinguish these manifolds by showing that Σ_g is orientable for all $g \geq 0$, while none of the manifolds X_k is orientable.

At first glance some manifolds are conspicuously absent from the Classification Theorem, for example, what about the Klein bottle K? What about connected sum that involve tori *and* projective planes, e.g., $T\#\mathbb{RP}^2$? In section 2.5 we will prove the following results.

Lemma 2.7. *1. The connected sum $\mathbb{RP}^2\#\mathbb{RP}^2$ of two copies of the projective plane is homeomorphic to the Klein bottle K.*

2. The connected sum $\mathbb{RP}^2\#T$ of a projective plane and the torus is homeomorphic to $\mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2$.

More precisely in section 2.5, we will develop the technique used to verify these statements and will prove part (2), leaving the easier part (1) as a homework problem. Then we will outline how these techniques are used to show that any compact connected 2-manifold is homeomorphic to Σ_g or X_k.

2.4 The Euler characteristic of compact 2-manifolds

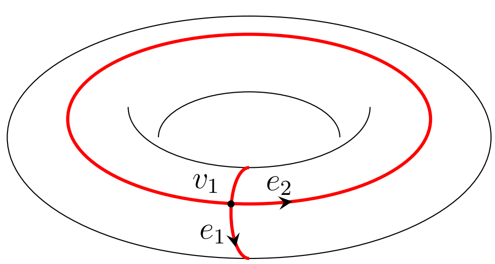
In this section we introduce the Euler characteristic of compact 2-manifolds. This invariant will allows us to show that some compact 2-manifolds are not homeomorphic. Our definition of the Euler characteristic is very geometric (and not particularly precise).

Definition 2.8. Let Σ be a compact 2-manifold. A *graph* Γ on Σ is a collection of finitely many points $v_1, \dots, v_k \in \Sigma$ (called *vertices*) and finitely many paths $e_i \colon [0,1] \to \Sigma$, $i = 1, \dots, \ell$ (called *edges*) such that

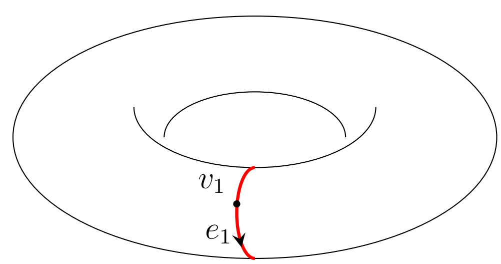
- the endpoints of e_i belong to the set of vertices $V := \{v_1, \dots, v_k\}$.
- the only intersection points of paths occur at their endpoints.

We call a graph Γ a *pattern of polygons* if the complement of all vertices and edges in Σ is a disjoint union of subspaces homeomorphic to open 2-disks.

Example 2.9. Both pictures below show examples of graphs Γ, Γ' on the torus T. The complement of Γ in T is the open square, and hence Γ is a pattern of polygons. The complement of Γ' is a cylinder, and so Γ' is not a pattern of polygons.



Γ

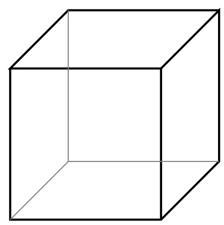


Γ'

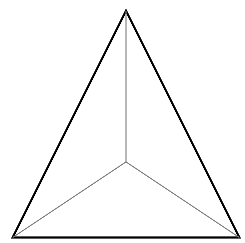
Let Γ be a pattern of polygons on a compact 2-manifold Σ.

$$\chi(\Sigma;\Gamma) := \#\{vertices\} - \#\{edges\} + \#\{polygons\}.$$

For example, the surface of a cube is homeomorphic to the sphere S^2. Via this homeomorphisms, the vertices, edges and faces of the cube can be interpreted as a pattern of polygons Γ_{cube} on S^2. More physically, think of the edges of the cube as a wireframe inside of a translucent sphere equipped with a light source at its center. Then the shadows of the edges give pattern of polygons (in this case quadrilaterals) on the sphere. Similarly, the tetrahedron can be interpreted as giving a pattern of polygons Γ_{tetra} on the sphere.



Γ_{cube}



Γ_{tetra}

We observe that

$$\chi(S^2;\Gamma_{\text{cube}}) = 8 - 12 + 6 = 2$$
$$\chi(S^2;\Gamma_{\text{cube}}) = 4 - 6 + 4 = 3$$

give the same number, independent whether we choose the pattern Γ_{cube} or Γ_{tetra} on S^2. This is in fact true generally:

Lemma 2.10. *Let Γ, Γ' be two patterns of polygons on a compact 2-manifold Σ. Then $\chi(\Sigma;\Gamma) = \chi(\Sigma;\Gamma')$.*

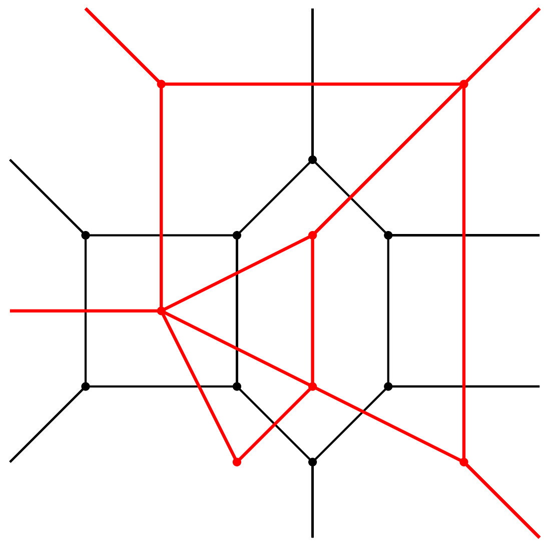
Proof. **Step 1.** By moving the vertices and edges of the graph Γ' a little bit, we can assume that the vertex sets of Γ and Γ' are disjoint, and that that there are only finitely many intersection points between edges of Γ and edges of Γ'. We claim that then there is a pattern of polygons Γ'' which is a *refinement* of both, Γ and Γ'. This means that Γ'' can be obtained from Γ (resp. Γ') by inductively adding new vertices on the interior of existing edges, and adding new edges between two vertices of a polygon.

The graph Γ'' is constructed as follows:

- The vertices of Γ'' are the vertices of Γ, the vertices of Γ' and all intersection points of edges of Γ and edges of Γ'.

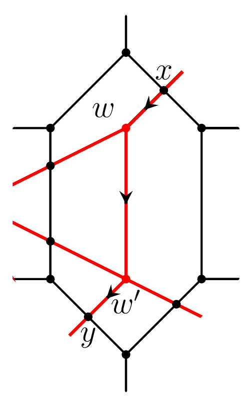
- The edges of Γ'' are segments of edges of Γ or Γ' whose endpoints are vertices of Γ''.

The following picture shows (part of) the graph Γ on some surface Σ with black vertices and edges and (part of) the graph Γ' colored red.



The graph Γ'' is simply the graph you see when we you ignore the color (and indicate that every intersection point is a vertex by drawing a little dot). It is clear that the complement of the graph Γ'' in Σ is again a disjoint union of open balls, since each connected component of the complement is an (open) polygon obtained by subdividing a polygon of Γ by edges.

To show that Γ'' is a refinement of Γ we first add all intersection points of edges of Γ and Γ' as new vertices (which subdivide the existing edges of Γ). Before we can add vertices of Γ' we need to add new edges: if w is a vertex of Γ' in the interior of some polygon P of Γ (e.g., the top red vertex in the black hexagon in the center of the picture above), there is a path through w along red edges that starts at some intersection point x of a red edge with an edge of P and ends at an intersection point y of some red edge with an edge of P. The following picture shows the vertex w, and the path along red edge segments (indicated by arrows) starting at x and ending at y.



We add this path as a new edge to our graph. Then we can add the red vertices w and w' to our graph (thus subdividing our new edge). Finally, we add the three additional red edges that connect w resp. w' to intersection points on the boundary of P. Doing this for all polygons of Γ we see that Γ'' is a refinement of Γ.

Step 2. Let Γ_1 be a pattern of polygons on Σ, and let Γ_2 be obtained by adding a new vertex to the interior of an edge of a graph Γ_1. We claim that $\chi(\Sigma;\Gamma_2) = \chi(\Sigma;\Gamma_1)$. To prove this, let $V(\Gamma_i)$ be the number of vertices, $E(\Gamma_i)$ the number of edges and $F(\Gamma_i)$ the number of faces of Γ_i. We note that $V(\Gamma_2) = V(\Gamma_1)$, due to the additional vertex, and $E(\Gamma_2) = E(\Gamma_1)$, since the creation of the new vertex on an edge subdivides that edge in two edges. The number of faces is unchanged and hence

$$\chi(\Sigma;\Gamma_2) = V(\Gamma_2) - E(\Gamma_2) + F(\Gamma_2) = (V(\Gamma_1)+1) - (E(\Gamma_1)+1) + F(\Gamma_1) = \chi(\Sigma;\Gamma_1).$$

Step 3. Let Γ_1 be a pattern of polygons on Σ, and let Γ_2 be obtained by introducing a new edge which connects two vertices of some polygon in Γ_1. Then the number of edges and faces goes up by one while the number of vertices is unchanged. Hence again, $\chi(\Sigma,\Gamma_2) = \chi(\Sigma,\Gamma_1)$.

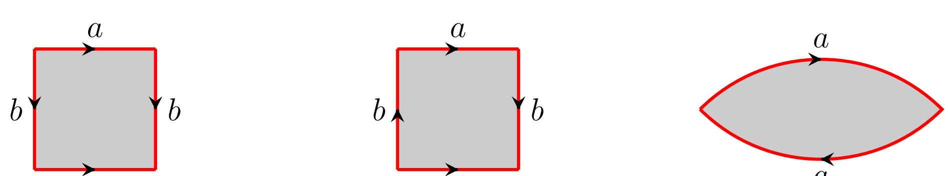
Steps 2 and 3 show that the alternating sum $\chi(\Sigma;\Gamma)$ doesn't change when we refine the graph Γ by adding vertices or edges. In particular, due to the existence of a common refinement Γ'' of graphs Γ and Γ' we conclude that

$$\chi(\Sigma,\Gamma) = \chi(\Sigma,\Gamma'') = \chi(\Sigma,\Gamma').$$

□

Definition 2.11. Let Σ be a compact 2-manifold. The *Euler characteristic of* Σ is defined to be the integer $\chi(\Sigma) := \chi(\Sigma;\Gamma)$.

To calculate the Euler characteristic of the torus T, the Klein bottle K and the real projective plane $\mathbb{RP}^2$ we use the fact that all three spaces can be described as quotients of polygons by identifying edges equipped with the same label.



T K $\mathbb{RP}^2$

The square from which the torus and the Klein bottle is built has four vertices, four edges and one face. However, we need to count vertices, edges and faces not for the square, but for the quotient space. The edges labeled a (resp. b) map to the *same* edge in the quotient under the projection map. Similarly, all four vertices (of the square) and the two vertices (of the bigon) map to the same vertex in the quotient. This shows that

$$\begin{aligned}\chi(T) =&1-2+1=0\\ \chi(K) =&1-2+1=0\\ \chi(\mathbb{RP}^2) =&1-1+1=-1\end{aligned}$$

We note that a homeomorphism $f\colon \Sigma \xrightarrow{\approx} \Sigma'$ between two compact 2-manifolds allows us to interpret a pattern of polygons Γ on Σ as a pattern of polygons on Σ'. This shows that the Euler characteristic of homeomorphic manifolds agrees. In others words, the Euler characteristic is an invariant that allows us to show that some 2-manifolds are not homeomorphic. In particular, our calculations above imply:

Corollary 2.12. *The compact 2-manifolds S^2, T and $\mathbb{RP}^2$ are pairwise non homeomorphic to each other.*

Lemma 2.13. *Let Σ, Σ' be compact 2-manifolds. Then $\chi(\Sigma\#\Sigma') = \chi(\Sigma) + \chi(\Sigma') - 2$.*

The proof is a homework problem.

Applying this inductively to the connected sums

$$\Sigma_g = \underbrace{T\#\dots\#T}_{g} \qquad \text{and} \qquad X_k = \underbrace{\mathbb{RP}^2\#\dots\#\mathbb{RP}^2}_{k}$$

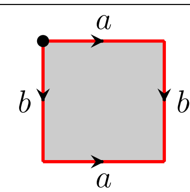
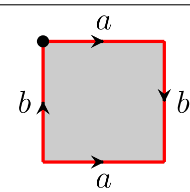
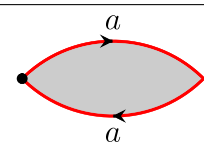
leads to the following result.

Corollary 2.14. *$\chi(\Sigma_g) = 2-2g$ and $\chi(X_k) = 2-k$. In particular, Σ_g is homeomorphic to $\Sigma_{g'}$ if and only if $g = g'$, and $X_k \approx X_{k'}$ if and only if $k = k'$.*

2.5 A combinatorial description of compact connected 2-manifolds

The Euler characteristic is an invariant which is very useful to show that two compact 2-manifolds are *not homeomorphic*. As advertised earlier, the goal of this section is introduce the technique used to *construct homeomorphisms* between 2-manifolds, notably to show that the Klein bottle $K$is homeomorphic to $\mathbb{RP}^2\#\mathbb{RP}^2$ and $T\#\mathbb{RP}^2 \approx \mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2$.

We recall that all three manifolds T, K and $\mathbb{RP}^2$ can be described as polygons (squares resp. bigons) with edge identifications as shown in the following table.

space	combinatorial picture	word
T = torus		$aba^{-1}b^{-1}$
K = Klein bottle		$aba^{-1}b$
$\mathbb{RP}^2$ = projective plane		aa

If we choose a distinguished vertex for these polygons, indicated by a black dot in the picture above, then the labeling of the edges by letters a, b and arrows can be encoded as follows. Going along the edges of the polygon clockwise, starting at the distinguished vertex, we write down for each edge

- the letter a if the edge has label a and the arrow of the edge points in the clockwise direction, or
- the letter a^{-1} if the edge has label a and the arrow of the edge points in the counter-clockwise direction.

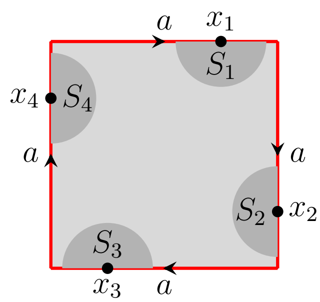
Doing this in order for all of the edges of the polygon, we obtain a string of symbols, that is, a *word* whose *letters* are the edge labels and their inverses. The words obtained this way for our examples are shown in the third column of the table above. This process can be reversed by interpreting a word W consisting of letters $a, a^{-1}, b, b^{-1}, \dots$ as giving the edges of a polygon P a label and a direction. This in turn determines an equivalence relation $\sim_W$ on P according to which corresponding points on edges with the same label are identified, and hence a quotient space $\Sigma(W) := P/\sim_W$. Here is the formal definition.

Definition 2.15. Let L be a set whose elements we refer to as *labels*, typically $a, b, \dots \in L$. Let $W = x_1 x_2 \dots x_n$ be an n letter word with letters x_i belonging to the *alphabet*, which is the set consisting of the symbols ℓ and ℓ^{-1} for $\ell \in L$. So typically, our alphabet is the set $A = \{a, a^{-1}, b, b^{-1}, \dots\}$.

Let P_n be an n-gon (i.e., the polygon with n edges) with a distinguished vertex. Going around P_n clockwise, starting at the distinguished vertex, label the edges of P_n by

$x_1, x_2, \ldots, x_n$. More precisely, if $x_i = \ell$ or $x_i = \ell^{-1}$ label the i-th edge by the label ℓ and equip it with an arrow according to the convention explained above. Let $\sim_W$ be the equivalence relation on P_n which identifies any point on an edge labeled ℓ with the corresponding point on any other edge with the same label. Then the *topological space associated to* W, denoted $\Sigma(W)$ or $\Sigma(x_1x_2\ldots x_n)$ is defined to be the quotient space $P_n/\sim_W$.

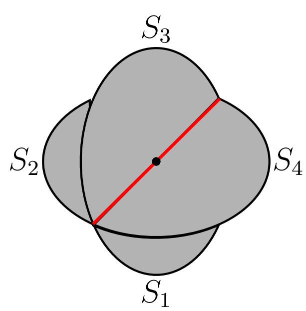
Warning. While all the spaces $\Sigma(W)$ mentioned as examples above were *manifolds*, this is not generally the case. For example $\Sigma(W) = \Sigma(aaaa)$ is *not* a manifold. To see this, consider a point x_1 in the interior of an edge of the square P_4. The equivalence class $[x_1] \in\in \Sigma(W) = P_4/\sim_W$ consists of the four points x_1, x_2, x_3, x_4, one on each edge as shown in the picture below. An open neighborhood of a point x_i consists of the dark semi-disk S_i containing x_i.

$\Sigma(aaaa) =$ 

It follows that an open neighborhood of $[x_1] \in \Sigma(W)$ has the form

$$(S_1 \cup S_2 \cup S_3 \cup S_4) / \sim,$$

where the equivalence relation is the restriction of $\sim_W$ to the union of these semi-circles. More geometrically, this is obtained by gluing these four semi-disks along there straight edge. Here is a picture of that quotient space; the red line is the line where the semi-disks are glued together and the marked point is the point $[x_1] \in \Sigma(W)$.

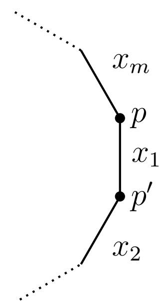


The following properties are immediate consequences of the construction of $\Sigma(W)$. We state it as a lemma to reference it later.

Lemma 2.16. *(1) Let W be a word built from labels in a set L. Let $L \leftrightarrow L'$ be a bijection of sets, and let W' be the word obtained by replacing each occurrence of the letter $\ell^{\pm 1}$ by $(\ell')^{\pm 1}$ where ℓ' corresponds to ℓ via the bijection. Then $\Sigma(W) \approx \Sigma(W')$.*

(2) If $x_i \in A = \{a, a^{-1}, b, b^{-1}, \dots\}$, then $\Sigma(x_1 x_2 \dots x_m) \approx \Sigma(x_2 x_3 \dots x_m x_1)$. More generally, for words W_1, W_2 with letters in A there is a homeomorphism $\Sigma(W_1 W_2) \approx \Sigma(W_2 W_1)$. Here $W_1 W_2$ is the concatenation of the words W_1 and W_2. More explicitly, if $W_1 = x_1 \dots x_m$ and $W_2 = y_1 \dots y_n$, then $W_1 W_2 = x_1 \dots x_m y_1 \dots y_n$.

Proof. Part (1) is evident. To prove part (2) consider the following figure showing a polygon with edges marked by the letters x_i.



The labeling of the edges determines the equivalence relation and hence the quotient space. Just in order to read off a word, we need to pick a vertex. If we pick the vertex labeled p in the picture above, that word is $x_1 x_2 \dots x_m$; if we choose p' instead, the resulting word is $x_2 \dots x_m x_1$. This proves that $\Sigma(x_1 \dots x_m)$ is homeomorphic to $\Sigma(x_2 \dots x_m x_1)$. Moving one letter of the word W_1 to right at a time we see that

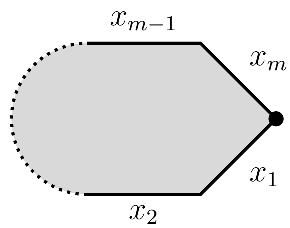
$$\begin{aligned}\Sigma(W_1 W_2) &= \Sigma(x_1 \dots x_m y_1 \dots y_n) \approx \Sigma(x_2 \dots x_m y_1 \dots y_n x_1) \approx \Sigma(x_3 \dots x_m y_1 \dots y_n x_1 x_2) \\ &\approx \cdots \approx \\ &\approx \Sigma(x_m y_1 \dots y_n x_1 \dots x_{m-1}) \approx \Sigma(y_1 \dots y_n x_1 \dots x_m) = \Sigma(W_2 W_1)\end{aligned}$$

□

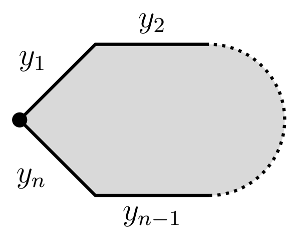
Proposition 2.17. *Let M, N be two compact connected 2-manifolds which are described combinatorially as $M = \Sigma(W_1)$, $N = \Sigma(W_2)$, where W_1 and W_2 are words from disjoint alphabets. Then the connected sum $M \# N$ is homeomorphic to $\Sigma(W_1 W_2)$.*

Proof. Let $W_1 = x_1 \dots x_m$ and $W_2 = y_1 \dots y_n$. Then M and N are described quotient spaces by the combinatorial pictures

$M = \Sigma(x_1 \ldots x_m) =$

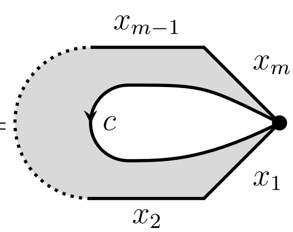


$N = \Sigma(y_1 \ldots y_n) =$

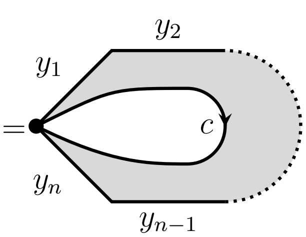


Here the dot marks the distinguished vertex. Now we remove an open disk $\mathring{D}^2$ from M and N. In the pictures below, this is the disk enclosed by the curve labeled c. So after removing the open disk bounded by the curve c, this curve is the boundary of the resulting manifold with boundary.

$M \setminus \mathring{D}^2 =$

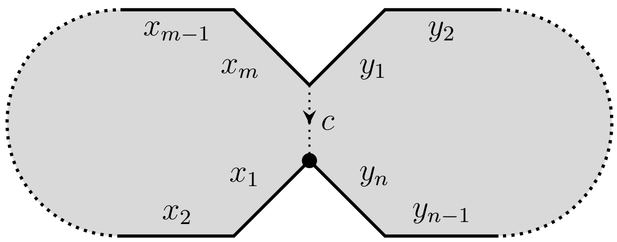


$N \setminus \mathring{D}^2 =$



Finally gluing these two spaces along the boundary circle c we obtain the connected sum

$M\#N$, which looks as follows:

$M\#N =$ 

This shows that the connected sum $M\#N$ is homeomorphic to

$$\Sigma(x_1 \dots x_m y_1 \dots y_n) = \Sigma(W_1 W_2)$$

as claimed. $\square$

Corollary 2.18. *(1)* $\Sigma_g = \underbrace{T\# \dots \#T}_{g}$ *is homeomorphic to* $\Sigma(a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1}\dots a_gb_ga_g^{-1}b_g^{-1})$.

(2) $\underbrace{\mathbb{RP}^2\# \dots \#\mathbb{RP}^2}_{k}$ *is homeomorphic to* $\Sigma(a_1a_1a_2a_2 \dots a_ka_k)$.

Proof. To prove part (1), we recall $T \approx \Sigma(aba^{-1}b^{-1})$. Then

$$\begin{aligned}\underbrace{T\# \dots \#T}_{g} &\approx \Sigma(a_1b_1a_1^{-1}b_1^{-1})\# \dots \#\Sigma(a_gb_ga_g^{-1}b_g^{-1})\\ &\approx \Sigma(a_1b_1a_1^{-1}b_1^{-1}\dots a_gb_ga_g^{-1}b_g^{-1}),\end{aligned}$$

where the last homeomorphism follows from the proposition. Similarly, to prove part (2), we use that $\mathbb{RP}^2 \approx \Sigma(aa)$ and hence

$$\begin{aligned}\underbrace{\mathbb{RP}^2\# \dots \#\mathbb{RP}^2}_{k} &\approx \Sigma(a_1a_1)\# \dots \#\Sigma(a_ka_k)\\ &\approx \Sigma(a_1a_1a_2a_2 \dots a_ka_k)\end{aligned}$$

$\square$

Proposition 2.19. *Let W_1, W_2, W_3 be words, and let a be a letter which does not occur in these words. Then there are homeomorphisms*

$$\Sigma(W_1aW_2aW_3) \approx \Sigma(W_1aaW_2^{-1}W_3) \tag{2.20}$$
$$\Sigma(W_1aW_2aW_3) \approx \Sigma(W_1W_2^{-1}aaW_3) \tag{2.21}$$

Here W_2^{-1} is the inverse of the word $W_2 = x_1 \dots x_n$, given explicitly by $W_2^{-1} = x_n^{-1}\dots x_1^{-1}$ (as for products of elements of a group).

Proof. By part (2) of Lemma 2.16 there are homeomorphisms

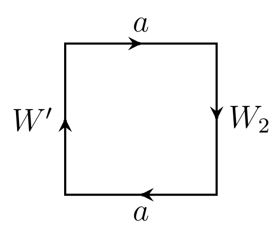
$$\Sigma(W_1aW_2aW_3) \approx \Sigma(aW_2aW') \qquad \text{and} \qquad \Sigma(W_1aaW_2^{-1}W_3) \approx \Sigma(aaW_2^{-1}W')$$

where $W' = W_3W_1$. Hence it suffice to produce homeomorphisms

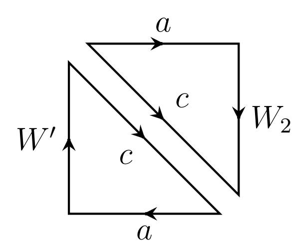
$$\Sigma(aW_2aW') \approx \Sigma(aaW_2^{-1}W') \tag{2.22}$$
$$\Sigma(aW_2aW') \approx \Sigma(W_2^{-1}aaW') \tag{2.23}$$

The homeomorphism (2.22) is given by the composition of the following homeomorphisms

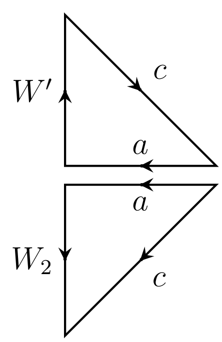
$\Sigma(aW_2aW') \approx$



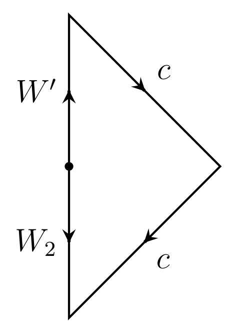
$\approx$



$\approx$



$\approx$



$$\approx \qquad \Sigma(ccW_2^{-1}W') \qquad \approx \qquad \Sigma(aaW_2^{-1}W')$$

1. Here the first homeomorphism is an equality, by definition of the quotient space $\Sigma(aW_2aW')$ associated to the word aW_2aW';

2. The second homeomorphism arises by cutting the square along the diagonal. (Strictly speaking, this "square" is a polygon which may have many many more than four edges: the number of edges is the length of the word aW_2aW'. However, if we draw

the edges corresponding to the words W' and W_2 vertically, and the two edges labeled a horizontally, then this polygon very much looks like a square, and so we prefer to use that terminology.) This results in two triangles (again, a slight abuse of language). We label the two new edges by the same label c (a new label distinct from all the other labels used so far) and the indicated direction. In Definition 2.15 we interpreted the labeling of the edges of one polygon as giving an equivalence relation on the polygon and hence an associated quotient space. Generalizing from one polygon with edge labeling to a disjoint union of polygons with edge labeling, we again interpret these pictures as giving us a quotient of the disjoint union of polygons by identifying all edges with the same label. Note that the order in which we glue the edges is irrelevant, and hence first gluing along the edge c gives back the previous quotient space.

3. The third homeomorphism is tautological, since the picture shows the same two polygons with the same edge labeling – we only moved the polygon drawn on the top right in the second picture to be below the other polygon (and we flipped it), so that the two edges labeled a in the two polygons are lined up.

4. The argument for the fourth homeomorphism is the same as for the second homeomorphism: first gluing along the edge labeled a, and then along the other edges gives the *same quotient* as identifying all edges with the same label simultaneously.

5. The fifth homeomorphism holds by definition of $\Sigma(ccW_2^{-1}W')$.

6. The sixth homeomorphism holds, since we may rename edges without changing the quotient they describe (see Lemma 2.16).

The homeomorphism (2.23) is constructed completely analogously by a sequence of pictures. The difference comes from using the *other diagonal* to cut the square in the first picture (going from the top right to the bottom left corner). □

Proof of Lemma 2.7(2). The desired homeomorphism is given by the following composition of homeomorphism. The numbers below these homeomorphism indicate the reference to the appropriate Lemma/Proposition/Definition.

$$\begin{aligned} T\#\mathbb{RP}^2 &\underset{(2.5)}{\approx} \Sigma(aba^{-1}b^{-1})\#\Sigma(cc) \underset{(2.19)}{\approx} \Sigma(aba^{-1}b^{-1}cc) \underset{(2.21)}{\approx} \Sigma(abcbac) \\ &\underset{(2.20)}{\approx} \Sigma(abbc^{-1}ac) \underset{(2.16)(2)}{\approx} \Sigma(bbc^{-1}aca) \underset{(2.21)}{\approx} \Sigma(bbc^{-1}c^{-1}aa) \\ &\underset{(2.19)}{\approx} \Sigma(bb)\#\Sigma(c^{-1}c^{-1})\#\Sigma(aa) \underset{(2.5)}{\approx} \mathbb{RP}^2\#\mathbb{RP}^2\#\mathbb{RP}^2 \end{aligned}$$

□

Outline of the constructive part of the Classification Theorem 2.6. Here by "constructive part" we mean the statement that every compact connected 2-manifold is homeomorphic to either Σ_g or X_k.

1. Show that every compact surface Σ admits a pattern of polygons Γ. Usually, this is stated as the stronger statement that every compact surface can be *triangulated*, meaning that it admits a pattern of triangles. Labeling all edges of Γ with a different letter and an arrow, and then cutting Σ along all edges gives a disjoint union of labeled polygons. By construction, Σ is the homeomorphic to the quotient space of this disjoint union by gluing along the pair of edges with the same label (see [Mu, Thm. 78.1]).

2. The number of polygons involved can be reduced by one by gluing pairs of edges with the same label belonging to different polygons. Inductively, this shows that Σ can be obtained by edge identifications of *one* polygon (see [Mu, Thm. 78.2]).

3. Use moves of the type described in Lemma 2.16 or Proposition 2.19 to show that the labeling of the edges of the polygon can be modified without changing the homeomorphism type of the quotient space to obtain the standard labeling for the surface of genus g or the k-fold projective space X_k (see [Mu, Thm. 77.5]).

□

2.6 Orientable 2-manifolds

Goal of this section is to define what "orientable" means for 2-manifolds (Definition 2.24) and to show that Σ_g, the surface of genus g is orientable, and that X_k, the k-fold projective plane, is not orientable (Proposition 2.25).

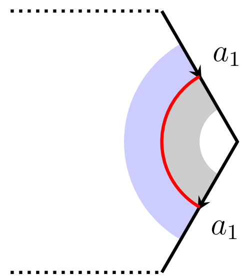
Definition 2.24. A 2-manifold Σ is *non-orientable* if it contains a subspace homeomorphic to the Möbius band. Otherwise Σ is called *orientable*.

Proposition 2.25. *(i) The k-fold projective plane $X_k = \mathbb{RP}^1 \# \dots \# \mathbb{RP}^2$ is not orientable*

(ii) The surface Σ_g of genus g is orientable.

If $f\colon \Sigma \to \Sigma'$ is a homeomorphism of 2-manifolds, then either both are orientable, or both are non-orientable, since if Σ contains a subspace M homeomorphic to the Möbius band, then $f(M)$ is a subspace of Σ' homeomorphic to the Möbius band. This in particular shows that Σ_g is not homeomorphic to X_k for any k. We recall that for $k = 2g$ these two manifolds have the same Euler characteristic, meaning that we needed a different invariant to show that these two are not homeomorphic.

Proof. **Proof of part (i).** We use our standard description of X_k as $\Sigma(a_1a_1a_2a_2\dots a_ka_k)$. The picture below shows the polygon, emphasizing the part of it involving the first two edges labeled a_1. When the two edges labeled a_1 are identified, the bicolored strip inside the polygon turns into a Möbius strip, since the blue part of the strip gets identified with the gray part and vice versa. This shows that X_k contains a Möbius band and hence X_k is not orientable.

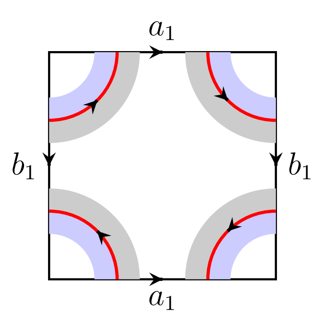


For the proof of part (ii) it will be useful to have the following *recognition principle for the Möbius band.*

Lemma 2.26. *Let M be the (open) Möbius band $M = ([0,1]\times(-1,1))/\sim$, where the equivalence relation is given by $(0,t)\sim(1,-t)$ for all $t\in(0,1)$. Let $C\subset M$ be the central circle of M, given by the image of the loop $\gamma\colon[0,1]\to M$ defined by $\gamma(s):=[s,0]$. Let U be an open neighborhood of C, i.e., U is an open subset of M which contains C. Then there exists a neighborhood V of C with $V\subset U$ such that the complement $V\setminus C$ is path connected.*

Proof of part (ii). Suppose that Σ_g is non-orientable, i.e., there is a map f from the Möbius band M to Σ_g that is a homeomorphism onto its image. More precisely, M is the open Möbius band that we describe as the quotient space $M = ([0,1]\times(-1,1))/\sim$, where the equivalence relation is given by $(0,t)\sim(1,-t)$ for all $t\in(0,1)$. Let $\gamma\colon[0,1]\to M$ be the central loop in the Möbius band, defined by $\gamma(s):=[s,0]$, and let$\delta\colon I\to\Sigma_g$ be the loop in Σ_g given by the composition $f\circ\gamma$.

As usual, we describe Σ_g as the quotient space of the $4g$-gon P_{4g} modulo edge identifications described by the word $a_1b_1a_1^{-1}b_1^{-1}\dots a_gb_1a_1^{-1}b_1^{-1}$. The picture below shows the the torus $T=\Sigma_1$ as the quotient of the square $= P_4$.



The four red path segments determine a *loop* in the quotient space $\Sigma_1 = P_4/\sim$, which is an example of the loop δ we are considering. The colored strip S around the loop δ is mathematically speaking an open neighborhood of image(δ). The complement $S \setminus \text{image}(\delta)$ is the disjoint union of the two open half-strips S^+ (colored blue) and S^- (colored black). We observe that under the edge identification points belonging to the blue (resp. black) part of the strip are identified with points lying in a region of the *same color*. This is the crucial difference between this picture and the previous one.

We will show that the image of δ is contained in small strip S around it which is two-sided, meaning that the complement $S \setminus \text{image}(\delta)$ is the union of two open disjoint subsets $S^+, S^- \subset S$

As usual, we describe Σ_g as the quotient space of the $4g$-gon P_{4g} modulo edge identifications described by the word $a_1b_1a_1^{-1}b_1^{-1}\dots a_gb_1a_1^{-1}b_1^{-1}$. Let $p\colon P_{4g} \to \Sigma_g$ be the projection map, let $x_0 \in \Sigma_g$ the image under p of all vertices of P_{4g}, and let $\alpha_1,\dots,\alpha_g,\beta_1,\dots,\beta_g$ be the $2g$ loops in Σ_g given by the images of the edges of P_{4g}. By moving the point x_0 a little bit, and a deformation of the paths α_i, β_i, we can assume that

- the path δ does not go through the point x_0;
- the point $\delta(0) = \delta(1)$ does not lie on any of the loops $\alpha_1,\dots,\alpha_g,\beta_1,\dots,\beta_g$;
- the path δ intersects these loops in only finitely many points;
- at any intersection point the path δ *crosses* the loop it intersects with.

Let $0 < s_1 < \cdots < s_k < 1$ be the only values of s for which $\delta(s)$ is an intersection point. It will be convenient to reparametrize the path δ to obtain a new path

$$\delta'\colon [s_1,\dots,1+s_1] \longrightarrow \Sigma_g \qquad \text{given by} \qquad \delta'(s) := \begin{cases} \delta(s) & s_1 \le s \le 1 \\ \delta(s-1) & 1 \le s \le 1+s_1 \end{cases}$$

Identifying the complement of the loops $\alpha_1,\dots,\alpha_g,\beta_1,\dots,\beta_g$ in Σ_g with the interior of the polygon P_{4g}, we obtain paths

$$\delta_i\colon [s_i, s_{i+1}] \to P_{4g}$$

for $i = 1,\dots,k$, defining $s_{k+1} := 1 + s_1$, such that for $s \in [s_i, s_{i+1}$ $\delta_i(s) \in P_{4g}$ maps to $\delta'(s) \in \Sigma_g$ under the projection map $p\colon P_{4g} \to \Sigma_g$ for $i = 1,\dots,k-1$, and

$$p(\delta_k(s)) = \begin{cases} \delta(s) & s \in [s_k, 1] \\ \delta(s-1) & s \in [1, 1+s_1] \end{cases}$$

We note that the paths $\delta_1,\dots,\delta_k$ are *disjoint*. Since the path $\delta\colon [0,1] \to \Sigma_g$ is injective, except for mapping both endpoints to the same points, it is clear that the only intersection points are possibly given by $\delta_i(s_{i+1})$ and $\delta_{i+1}(s_{i+1})$. Both of these points map to the same point $\delta(s_{i+1})$ under the projection map $p\colon P_{4g} \to \Sigma_g$. However, our requirement that the path δ *crosses* the loops at an intersection point implies that the points $\delta_i(s_{i+1})$ and $\delta_{i+1}(s_{i+1})$ *must belong to different edges of* P_{4g} (that map to the same loop under the projection map p).

It follows that we can thicken each path δ_i into a narrow open strip S_i around the path such that

- the intersection of S_i with the boundary ∂P_{4g} consists of two little intervals around the endpoints of δ_i, contained in the edges these endpoints lie on.
- the strips $S_1, \dots, S_k$ are mutually disjoint.

Let S The complement of the path δ_i inside the strip has two connected components. We can use the right hand rule to distinguish these components: pointing the thumb of our right hand in the direction of the path δ_i, our index finger points to one connected component of $S_i \setminus$ image of δ_i that we denote S_i^+; we write S_i^- for the other connected component. The crucial observation is that for each pair of edges of P_{4g} that are identified in Σ_g, the arrows of these edges are pointing in opposite direction (i.e, if one arrow points clockwise, the other points counterclockwise and vice versa). It follows that S_i^+ get identif

these strips are still disjoint, and

that it contains a subspace M which is homeomorphic to the open Möbius band. Let $C \subset \Sigma_g$ be the image of the closed central circle of the Möbius band. Composing a parametrization of the central circle with the homeomorphism then gives a loop $\gamma\colon I \to C \subset M \subset \Sigma_g$. Identifying Σ_g as usual with the quotient of the $4g$-gon P_{4g} modulo edge identifications, we can assume after deforming γ slightly that it misses the one point of Σ_g that is the image of all vertices of P_{4g}.

The path γ does not lift to a path in P_{4g}; rather

□

Proof of Lemma 2.26. First we will prove that for sufficiently small $\epsilon > 0$ the subset

$$V := \{[s,t] \mid -\epsilon < t < \epsilon\} \subset M$$

is an open neighborhood of C contained in U. Then we will show that $V \setminus C$ is path connected.

Let $p\colon [0,1] \times (-1,1) \to M$ be the projection map onto the quotient space M. Then $p^{-1}(U)$ is an open neighborhood of $[0,1] \times \{0\}$, and hence for each $s \in [0,1]$ there are $\delta_s, \epsilon_s > 0$ such that the product of intervals

$$B_{\delta_s}(s) \times B_{\epsilon_s}(0) = (s - \delta_s, s + \delta_s) \times (-\epsilon_s, \epsilon_s)$$

is contained in $p^{-1}(U)$. The intersections of these products with $[0,1]$ form an open cover of $[0,1]$. Hence by compactness of $[0,1]$, there are a finitely many points $s_1, \dots, s_k \in [0,1]$

such that $[0,1]\times\{0\}$ is contained in the union of $B_{\delta_{s_i}}(s_i)\times B_{\epsilon_{s_i}}(0)$ for $i=1,\dots,k$. Defining $\epsilon := \min\{\epsilon_{s_i} \mid i=1,\dots,k\}$, it follows that $\widetilde{V} := [0,1]\times(-\epsilon,\epsilon)$ is contained in $p^{-1}(U)$. Then $V := p(\widetilde{V})$ is an open neighborhood of C; it is open since $p^{-1}(V)=\widetilde{V}$ is open in $[0,1]\times(-1,1)$.

To show that $V\setminus C$ is path connected, we consider the map

$$f\colon [0,2]\times(0,\epsilon)\longrightarrow V\setminus C \qquad \text{given by} \qquad f(s,t) := \begin{cases} [s,t] & 0\le s\le 1 \\ [s-1,-t] & 1\le s\le 2 \end{cases}$$

The map f is well-defined due to the equivalence relation defining the Möbius band. The restriction of f to $[0,1]\times(0,\epsilon)$ and $[1,2]\times(0,\epsilon)$ is map is continuous, since these maps are compositions $p\circ g_i$ of the projection map p with maps

$$\begin{aligned} g_1\colon [0,1]\times(0,\epsilon) &\longrightarrow [0,1] \qquad g_1(s,t) := (s,t) \\ g_2\colon [1,2]\times(0,\epsilon) &\longrightarrow [0,1] \qquad g_2(s,t) := (s-1,-t) \end{aligned}$$

which are continuous since their component maps are. It follows that f is continuous.

We observe that the map f is surjective, since points of the form $[s,t]$ with $t>0$ are in the image of f restricted to $[0,1]\times(0,\epsilon)$, while for $t<0$, they are in the image of f restricted to $[1,2]\times(0,\epsilon)$.

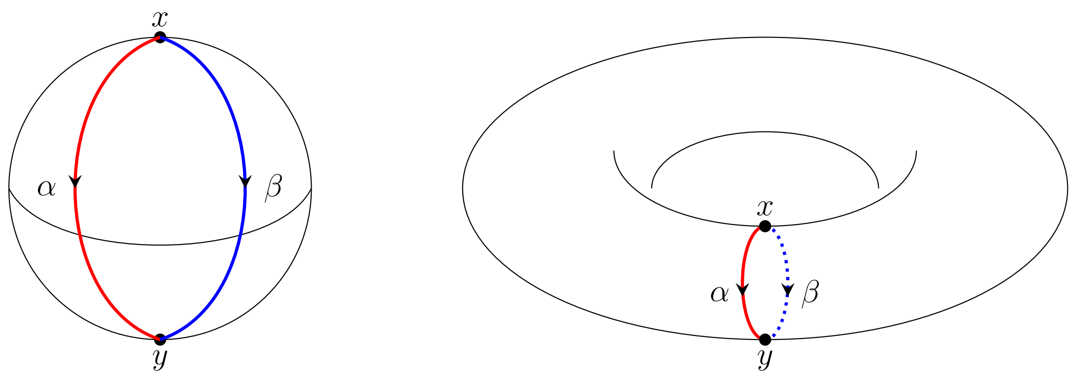
To show that any two points $x,y\in V$ can be connected by a path, pick $\widetilde{x},\widetilde{y}\in[0,2]\times(0,\epsilon)$ with $f(\widetilde{x})=x$ and $f(\widetilde{y})=y$. Since $[0,2]\times(0,\epsilon)$ is is a convex subset of $\mathbb{R}^2$, these two points can be connected by the straight line path from $\widetilde{x}$ to $\widetilde{y}$. Composing this path with f then yields a path in V connecting x and y. □

3 The fundamental group

In this section we define the fundamental group of a topological space X. This is an invariant which can be used to distinguish topological spaces. For example, we will see that all the compact connected manifolds can be distinguished by their fundamental group. This is done using the Seifert van Kampen Theorem, a powerful tool to calculate fundamental groups.

3.1 The definition of the fundamental group

The basic idea of the fundamental group is that paths in different topological spaces might have different behavior. For example, the picture below shows two paths α, β with the same starting point and end point in the sphere and the torus. The difference between the two situations is that for the sphere the path α can be deformed to give the path β, while the paths on the torus cannot be deformed into one another.



Definition 3.1. A *path* in a topological space X is a continuous map $\gamma\colon [0,1] \to X$. The point $\gamma(0) \in X$ is the *starting point*, the point $\gamma(1) \in X$ is the *endpoint* of the path γ. With a slight abuse of language, both point $\gamma(1)$ and $\gamma(0)$ might be referred to as *endpoints* of the path γ. If $\gamma(0) = x$ and $\gamma(1) = y$, we say that γ is a path from x to y.

Let γ, δ be two paths in X from x to y. These paths are *homotopic relative endpoints* or *path homotopic* or simply *homotopic* if for every $t \in [0,1]$ there is a path γ_t from x to y such that

- $\gamma_0 = \gamma$ and $\gamma_1 = \delta$;
- The map $H\colon [0,1] \times [0,1] \to X$, $(s,t) \mapsto \gamma_t(s)$ is continuous. This condition expresses the idea that the family of paths γ_t depends continuously on the parameter t.

The map H is called a *homotopy from* γ *to* δ, and we write $\gamma \sim \delta$ to say that γ is homotopic to δ. It is easy to show that *homotopic* is an equivalence relation (we leave the proof to the reader). We use the notation $[\gamma]$ for the homotopy class of a path γ.

Let $U \subset \mathbb{R}^n$ be a *convex* subset, i.e., for any points $x, y \in U$ the straight line segment between x and y is contained in U. Explicitly, the straight line segment is the set

$$\{(1-t)x + ty \in \mathbb{R}^n \mid 0 \le t \le 1\}.$$

Examples of convex subspaces of $\mathbb{R}^n$:

- $\mathbb{R}^n$;
- an open ball $B_r(x)$ of radius r around some point $x \in \mathbb{R}^n$;
- a closed ball $D_r(x) := \{y \in \mathbb{R}^n \mid ||y - x|| \le r\}$ of radius r around some point $x \in \mathbb{R}^n$;

The punctured space $\mathbb{R}^n \setminus \{v\}$ is not convex, since for any nonzero $w \in \mathbb{R}^n$ the straight line segment between $x = v + w$ and $y = v - w$ contains the point v.

Lemma 3.2. *Let U be a convex subset of $\mathbb{R}^n$, and let α, β be paths in U with the same endpoints (i.e., $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$). Then α and β are homotopic (relative endpoints). An explicit homotopy, called* linear homotopy *is given by the formula*

$$H\colon [0,1] \times [0,1] \longrightarrow U \qquad \textit{is given by} \qquad H(s,t) := (1-t)\gamma(s) + t\delta(s).$$

We note that for fixed $s \in [0,1]$ the path $t \mapsto H(s,t) = (1-t)\gamma(s) + t\delta(s)$ is the straight line path from $\gamma(s)$ to $\delta(s)$.

Definition 3.3. Let $\alpha, \beta\colon I \to X$, be paths in a topological space X. If $\alpha(1) = \beta(0)$, i.e., if the endpoint of α matches the starting point of β, then we can form a new path $\alpha * \beta$ called the *concatenation of* α *and* β by first following the path α and then following the path β. Explicitly, the path

$$\alpha * \beta\colon I \to X \qquad \text{is given by} \qquad (\alpha * \beta)(s) = \begin{cases} \alpha(2s) & 0 \le s \le \frac{1}{2} \\ \beta(2s-1) & \frac{1}{2} \le s \le 1 \end{cases}$$

It has starting point $\alpha(0)$ and endpoint $\beta(1)$.

Let α, β and γ be paths in X with $\alpha(1) = \beta(0)$ and $\beta(1) = \gamma(0)$, then we can form the concatenated paths $\alpha * \beta$ and $\beta * \gamma$. Since the endpoint of $\alpha * \beta$ is $\beta(1)$, it can further be concatenated with γ, forming the path $(\alpha * \beta) * \gamma$. Similarly, we can form the concatenation $\alpha * (\beta * \gamma)$. We want to point out that these paths are typically *not* equal to each other:

- $\alpha * (\beta * \gamma)(s)$ is a point on the path α for $0 \le s \le 1/2$, on the path β for $1/2 \le s \le 3/4$ and on the path γ for $3/4 \le s \le 1$, while
- $((\alpha * \beta) * \gamma)(s)$ on the path α for $0 \le s \le 1/4$, on the path β for $1/4 \le s \le 1/2$, and on the path γ for $1/2 \le s \le 1$.

Lemma 3.4. *Concatenation is associative up to homotopy, that is, if α, β, γ are paths in X with $\alpha(1) = \beta(0)$ and $\beta(1) = \gamma(0)$, then the paths*

$$\alpha * (\beta * \gamma) \qquad \textit{and} \qquad (\alpha * \beta) * \gamma \qquad\qquad \textit{are homotopic};$$

in other words, $[\alpha * (\beta * \gamma)] = [(\alpha * \beta) * \gamma]$.

The associativity of the concatenation of paths up to homotopy suggests that we might be able to construct a *group* associated to a topological space X by taking the elements of this group to be homotopy classes of paths in X. The problem with this is that paths can only be concatenated it the endpoint of the first path matches the starting point of the second path, while *any* two elements of a group can be multiplied with each other. There are two ways to deal with this issue:

- We pick a point $x_0 \in X$ and only consider paths that start and end at x_0; this is what we will do in the definition below of the fundamental group of a topological space X.
- We give up the idea of constructing a *group*, but instead construct a *groupoid* which is called the *fundamental groupoid* of X.

Definition 3.5. Let X be a topological space and let $x_0 \in X$ be a point of X, usually referred to as *base point*. Such a pair (X, x_0) is called a *pointed topological space*. A *based loop* in (X, x_0) is a path $\gamma\colon I \to X$ with $\gamma(0) = x_0 = \gamma(1)$. Let

$$\pi_1(X, x_0) := \{\text{based loops in } (X, x_0)\}/\text{homotopy}.$$

Proposition 3.6. *The set $\pi_1(X, x_0)$ is a group, the* fundamental group of (X, x_0), *with*

- *multiplication given by concatenation of based loops, i.e., $[\alpha] \cdot [\beta] := [\alpha * \beta]$ for based loops α, β;*
- *the identity element of $\pi_1(X, x_0)$ is given by the homotopy class of the constant path c_{x_0} (i.e., $c_{x_0}(s) = x_0$ for all $s \in I$);*
- *the inverse of an element $[\gamma] \in \pi_1(X, x_0)$ is given by $[\bar{\gamma}]$, where $\bar{\gamma}\colon I \to X$ is the path γ run backwards, i.e., $\bar{\gamma}(s) = \gamma(1 - s)$.*

The proof of this statement is pretty straightforward. The associativity of the product is a consequence of Lemma 3.4, which is a more general since it is a statement for composable paths rather than just based loops. Similarly, the claim that the homotopy class of constant path c_{x_0} is the identity element of $\pi_1(X, x_0)$ is a consequence of the first two homotopies of the following lemma, while the last two homotopies imply that $[\bar{\gamma}]$ is the inverse to $[\gamma]$ for a based loop in (X, x_0). Again, it will be useful for us to state these homotopies for paths, rather than just based loops.

Lemma 3.7. *Let $\gamma\colon I \to X$ be a path in a topological space X. Let $\bar{\gamma}\colon I \to X$ be the path defined by $\bar{\gamma}(s) := \gamma(1 - s)$ and let $c_x\colon I \to X$ be the constant path at a point $x \in X$. Then there are homotopies*

$$\gamma * c_{\gamma(1)} \sim \gamma \qquad c_{\gamma(0)} * \gamma \sim \gamma \qquad \gamma * \bar{\gamma} \sim c_{\gamma(0)} \qquad \bar{\gamma} * \gamma \sim c_{\gamma(1)}$$

Example 3.8. Let X be a convex subset of $\mathbb{R}^n$, and $x_0 \in X$. Then by Lemma 3.2 any based loop in (X, x_0) is homotopic to the constant loop c_{x_0}. Hence the fundamental group $\pi_1(X, x_0)$ is trivial.

Lemma 3.9. *Let X be a topological space and let β be a path from x_0 to x_1. Then the map*

$$\Phi_\beta\colon \pi_1(X, x_0) \longrightarrow \pi_1(X, x_1) \qquad [\gamma] \mapsto [\bar{\beta} * \gamma * \beta]$$

is an isomorphism of groups. In particular, the isomorphism class of the fundamental group $\pi(X, x_0)$ of a path connected space does not depend on the choice of the base point $x_0 \in X$.

3.2 Fundamental group of the circle

Proposition 3.10. *The fundamental group of the circle $\pi_1(S^1, x_0)$ is isomorphic to $\mathbb{Z}$.*

This isomorphism is given by associating to each loop γ based at x_0 its *winding number* $W(\gamma) \in \mathbb{Z}$, which intuitively measures how often (and in which direction) the loop γ winds around the circle. For example, consider the

Let us choose $x_0 = 1 \in S^1 \subset \mathbb{C}$ as the basepoint for the circle S^1.

The goal of this section is to determine the fundamental group of the circle. Let us choose $x_0 = 1 \in S^1 \subset \mathbb{C}$ as the basepoint.

Let $f\colon X \to Y$ be a continuous map. Then for every path $\gamma\colon I \to X$ in X the composition

$$I \xrightarrow{\ \gamma\ } X \xrightarrow{\ f\ } Y$$

is a path in Y. The following lemma shows that this construction is compatible with homotopies and concatenation of paths.

Lemma 3.11. *Let $f\colon X \to Y$ be a continuous map, and let γ, δ be paths in X.*

Compatibility with homotopies: *If γ, δ have the same endpoints (i.e., $\gamma(0) = \delta(0)$ and $\gamma(1) = \delta(1)$) and are homotopic (relative endpoints), then also $f \circ \gamma$ and $f \circ \delta$ are homotopic (relative endpoints).*

Compatibility with concatenation: *If $\gamma(1) = \delta(0)$, i.e., if the concatenation $\gamma * \delta$ is defined, then*

$$f \circ (\gamma * \delta) = (f \circ \gamma) * (f \circ \delta).$$

Definition 3.12. Let $f\colon X \to Y$ be a continuous map, let $x_0 \in X$, and $y_0 = f(x_0) \in Y$. Then the map

$$f_*\colon \pi_1(X, x_0) \longrightarrow \pi_1(Y, y_0) \qquad \text{defined by} \qquad [\gamma] \mapsto [f \circ \gamma]$$

is called the *map of fundamental groups induced by f*. The map f_* is well-define by part (i) of the above lemma. It is a group homomorphism since by part (ii) of the lemma composition with f is compatible with concatenation of paths.

We have defined the winding number $W(\gamma) \in \mathbb{Z}$ for loops in S^1 based at $1 \in S^1$. We would like to generalize this definition by defining the winding number $W(\gamma, w)$ of loops $\gamma\colon I \to \mathbb{C} \setminus \{w\}$ in the complement of $w \in \mathbb{C}$ based at some point z_0 (i.e., $\gamma(0) = \gamma(1) = z_0$). Such a loop γ represents an element in the fundamental group $\pi_1(\mathbb{C} \setminus \{w\}, z_0)$. So the idea is to find a map $\mathbb{C} \setminus \{w\} \longrightarrow S^1$ that maps the basepoint $z_0 \in \mathbb{C} \setminus \{w\}$ to the basepoint $1 \in S^1$ and hence induces a homomorphism between the corresponding fundamental groups. The following map does the job

$$f_{z_0}\colon \mathbb{C} \setminus \{w\} \longrightarrow S^1 \qquad \text{be defined by} \qquad z \mapsto \frac{z - w}{|z - w|} \frac{|z_0 - w|}{z_0 - w}$$

Definition 3.13. Let $\gamma\colon I \to \mathbb{C} \setminus \{w\}$ be a loop based at z_0. Then the *winding number* $W(\gamma, w)$ *of* γ *with respect to* $w \in \mathbb{C}$ is defined to be the image of $[\gamma] \in \pi_1(\mathbb{C} \setminus \{w\}, z_0)$ under the homomorphism

$$(f_{z_0})_*\colon \pi_1(\mathbb{C} \setminus \{w\}, z_0) \xrightarrow{(f_{z_0})_*} \pi_1(S^1, 1) \xrightarrow[\cong]{W} \mathbb{Z}.$$

Definition 3.14. Two continuous maps $f, g\colon X \to Y$ are *homotopic* if and only if there is a continuous map $H\colon X \times [0,1] \to Y$ such that $H(x, 0) = f(x)$, $H(x, 1) = g(x)$. This is an equivalence relation. Notation $f \sim g$;

$$[X, Y] := \{\text{homotopy classes of maps } f\colon X \to Y\}$$

Note:

1. Fixing $t \in I$ leads to a map $f_t\colon X \to Y$ defined by $f_t(x) = H(x, t)$. Lets us think of a homotopy as a family of maps interpolating between f and g; in other words, g is a *deformation* of f.

2. Fixing $x \in I$ leads to path $f_x\colon I \to Y$ defined by $f_x(t) = H(x, t)$. This lets us think of H as a family of paths (f_x goes from $f(x)$ to $g(x)$).

Indicate both points of view using different colors in the example of maps $f, g\colon S^1 \to T$.

Theorem 3.15. The Fundamental Theorem of Algebra. *Let*

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

be a polynomial of degree $n > 0$. Then p has a zero, that is there is some $z \in \mathbb{C}$ such that $p(z) = 0$.

Proof. Aiming for a proof by contradiction, we assume that $p(z)$ belongs to $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ for all $z \in \mathbb{C}$. This allows us to talk about the winding number $W(\gamma) \in \mathbb{Z}$ of the closed curve

$$\gamma\colon S^1 \longrightarrow \mathbb{C}^\times \qquad \text{given by} \qquad z \mapsto p(z).$$

We will calculate $W(\gamma)$ in two different ways, the first one resulting in $W(\gamma) = 0$, the other one resulting in $W(\gamma) = n$. This is the desired contradiction.

The loop γ is obtained by restricting the polynomial $p(z)$ to the unit circle. Restricting $p(z)$ instead to the circle of radius r, we obtain the loop

$$\gamma_r\colon S^1 \longrightarrow \mathbb{C}^\times \qquad \text{defined by} \qquad z \mapsto p(rz).$$

We note that the loop γ_r is homotopic to $\gamma = \gamma_1$. A homotopy is given by

$$H\colon S^1 \times I \longrightarrow \mathbb{C}^\times \qquad \text{given by} \qquad H(z, t) = p((tr + (t-1))z).$$

This implies that $W(\gamma) = W(\gamma_r)$ for all r, including 0.
First calculation. The loop γ_0 is the constant loop, and hence $W(\gamma_0) = 0$. It follows that $W(\gamma) = 0$.
Second calculation. Instead of shrinking r to 0, we will now consider the loop γ_r for large radius r. Writing $\gamma_r(z)$ in the form

$$\gamma_r(z) = a_n r^n z^n + a_{n-1} r^{n-1} z^{n-1} + \cdots + a_0 = r^n \left(a_n z^n + \frac{a_{n-1}}{r} z^n + \cdots + \frac{a_0}{r^n} \right),$$

we see that the term $b_r(z) := \frac{a_{n-1}}{r} z^n + \cdots + \frac{a_0}{r^n}$ converges to 0 uniformly for $z \in S^1$. In particular, for sufficiently large r we have

$$|b_r(z)| < |a_n z^n| \qquad \text{for all } z \in S^1.$$

It follows that $a_n z^n + t b_r(z)$ belongs to $\mathbb{C}^\times$ for all $z \in S^1$ and $t \in [0, 1]$. Then

$$H\colon S^1 \times I \longrightarrow \mathbb{C}^\times \qquad \text{defined by} \qquad H(z, t) := r^n \left(a_n z^n + t b_r(z) \right)$$

is a homotopy between $\gamma_r(z)$ and the loop $\gamma'(z) := r^n a_n z^n$. The loop γ' in turn is homotopic to the loop ω_n defined by $\omega_n(z) := z^n$; a homotopy $H'\colon S^1 \times I \to \mathbb{C}^\times$ is given by chosing a path δ in $\mathbb{C}^\times$ with $\delta(0) = r^n a_n$, $\delta = 1$, and defining

$$H'(z, t) := \delta(t) z^n.$$

Since homotopic loops have the same winding number, it follows that

$$W(\gamma_r) = W(\gamma') = W(\omega_n) = n.$$

□

3.3 The fundamental group as a functor

In the previous sections we have discussed how we can associate to any pointed topological space (X, x_0) a group $\pi_1(X, x_0)$ (the fundamental group, Definition 3.5) and how to associate to a base point preserving map $f\colon (X, x_0) \to (Y, y_0)$ between pointed topological spaces a group homomorphism $f_*\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$ (the induced homomorphism on fundamental groups, Definition **??**). In other words, this construction associates to one kind of mathematical object (a pointed topological space) a different kind of mathematical object (a group), and to appropriate maps between the first kind of objects (basepoint preserving continuous maps) appropriate maps between the second kind of objects (group homomorphisms).

Such a construction is called a *functor between categories*. The goal of this section is to provide a quick introduction to categories and functors. Even if you haven't seen the formal definition of a category, it is likely that you already know many examples of categories. So it seems appropriate to mention some mathematical objects and appropriate maps between them that will then motivate the definition of a category.

When studying various mathematical objects, we usually also talk about the appropriate kind of maps between these objects. The following table lists some well known examples.

mathematical objects	appropriate maps
sets	maps
groups	group homomorphisms
vector spaces	linear maps
topological spaces	continuous maps

What is the structure that is common to all of these four types of mathematical objects and the maps between them? There isn't too much there, but we observe that composing "appropriate maps" leads again to "appropriate maps" (assuming the domain/source of one map matches the codomain/target of the other map), and that there is an "identity map" for every object. The following definition captures this structure, which is called a *category*. The four kinds of mathematical objects and the maps between them are then examples of categories.

Remark 3.16. Let X, Y, Z be sets and let $g\colon X \to Y$ and $f\colon Y \to Z$ be maps. Then there are *two* usual ways to write the composition, namely as

$$g \circ f \qquad \text{or as} \qquad X \xrightarrow{f} Y \xrightarrow{g} Z$$

Using the first way to write compositions, it is natural to think of composition as the map given by $(g, f) \mapsto g \circ f$. Writing $\mathrm{Maps}(X, Y)$ for the set of maps from X to Y, this is the map

$$\begin{array}{rcl} \mathrm{Maps}(Y, Z) \times \mathrm{Maps}(X, Y) & \longrightarrow & \mathrm{Maps}(X, Z) \\ (g, f) & \mapsto & g \circ f \end{array}$$

However, thinking about it the second way, it is more natural to think of composition as the map

$$\begin{array}{rcl} \mathrm{Maps}(X, Y) \times \mathrm{Maps}(Y, Z) & \longrightarrow & \mathrm{Maps}(X, Z) \\ (X \xrightarrow{f} Y, Y \xrightarrow{g} Z) & \mapsto & X \xrightarrow{f} Y \xrightarrow{g} Z \end{array}$$

Both ways have their advantages and disadvantages; to me, the second one seems more elegant, but alas, the first way is probably too deeply entrenched in mathematics to be thrown out. The sad effect is that there is there is no general consensus of how to write compositions in categories. I will follow the first convention.

Definition 3.17. A *category* $\mathcal{C}$ consists of the following data:

- A class of objects, denoted $\mathrm{ob}(\mathcal{C})$; the elements of $\mathrm{ob}(\mathcal{C})$ are called the *objects* of the category $\mathcal{C}$.
- For each pair of objects $X, Y \in \mathrm{ob}(\mathcal{C})$ a set $\mathrm{mor}_{\mathcal{C}}(X, Y)$. The elements of $\mathrm{mor}_{\mathcal{C}}(X, Y)$ are called *morphisms in* $\mathcal{C}$ *from* X *to* Y or morphisms with domain (or source) X and codomain (or target) Y. Alternative notations for this set include $\mathrm{mor}(X, Y)$ or $\mathcal{C}(X, Y)$.
- For objects $X, Y, Z \in \mathcal{C}$ there is a composition map
$$\circ\colon \mathcal{C}(Y, Z) \times \mathcal{C}(X, Y) \longrightarrow \mathcal{C}(X, Z) \qquad (g, f) \mapsto g \circ f.$$
- For each object $X \in \mathrm{ob}(\mathcal{C})$ a morphism $\mathrm{id}_X \in \mathcal{C}(X, X)$ called *identity morphism.*

These data are subject to the following requirements:

associativity For morphisms $f \in \mathcal{C}(U, X)$, $g \in \mathcal{C}(X, Y)$, $h \in \mathcal{C}(Y, Z)$ we have
$$(h \circ g) \circ f = h \circ (g \circ f) \in \mathcal{C}(U, Z).$$

identity property For $f \in \mathcal{C}(X, Y)$ we have $f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f \ \in \mathcal{C}(X, Y)$.

Remark 3.18. For a morphism $f \in \mathcal{C}(X, Y)$ we often write $X \xrightarrow{f} Y$ to indicate the domain and codomain of f. For $f \in \mathcal{C}(Y, Z)$ and $g \in \mathcal{C}(X, Y)$ we often write $X \xrightarrow{g} Y \xrightarrow{f} Z$ for the composition $f \circ g$.

Definition 3.19. Let $\mathcal{C}$ be a category. A morphism $f \in \mathcal{C}(X, Y)$ is called an *isomorphism* if there exists a morphism $g \in \mathcal{C}(Y, X)$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

We can now recast our motivating examples of sets, groups, vector spaces and topological spaces as categories.

Example 3.20.

category $\mathcal{C}$	objects	morphisms	isomorphisms
Set	sets	maps	bijections
Grp	groups	group homomorphisms	group isomorphisms
Vect	vector spaces	linear maps	linear isomorphisms
Top	topological spaces	continuous maps	homeomorphisms

Our previous examples of category might suggest that morphisms are always maps of sets compatible with additional structure these sets might have. In the following examples of categories, this is not the case.

Example 3.21. (Examples of categories whose morphisms are not maps between sets).

- To any group G we can associate a category $\mathcal{C}$ as follows. The category $\mathcal{C}$ has one object denoted $*$, and the set of morphisms $\mathcal{C}(*,*)$ from $*$ to $*$ is the set G of group elements. The composition map

$$\circ\colon \mathcal{C}(*,*) \times \mathcal{C}(*,*) \longrightarrow \mathcal{C}(*,*)$$

  is given by the map $m\colon G \times G \to G$ that describes multiplication of elements of G. The identity morphism id_* is defined to be the identity element $1 \in G$ of the group G. Associativity and the identity property hold for the category $\mathcal{C}$, since the group multiplication is associative and $1 \in G$ is the identity element of the group G.

  We note that every morphism $g \in \mathcal{C}(*,*)$ is an isomorphism (its inverse is given by the group element g^{-1}), and hence $\mathcal{C}$ is a groupoid.

- To any topological space X we can associate a groupoid $\Pi_1(X)$, called the *fundamental groupoid of* X. As the name suggests, this is a generalization of the fundamental group of X. The objects of $\Pi_1(X)$ are the points of X. For x, y, the set of morphisms $\mathrm{mor}(x,y)$ is defined to be

$$\mathrm{mor}(x,y) := \{\text{paths } \gamma\colon I \to X \text{ with } \gamma(1) = x,\, \gamma(0) = y\}/\text{homotopy}.$$

  For $x, y, z \in X$, the composition in this category is induced by concatenation of paths:

$$\circ\colon\ \mathrm{mor}(y,z) \times \mathrm{mor}(x,y) \longrightarrow \mathrm{mor}(x,z) \qquad \text{is given by} \qquad ([\alpha],[\beta]) \mapsto [\alpha * \beta]$$

Definition 3.22. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A functor $F\colon \mathcal{C} \to \mathcal{D}$

- associates to every object X of $\mathcal{C}$ an object $F(X)$ of $\mathcal{D}$;
- associates to every morphism $f \in \mathcal{C}(X,Y)$ a morphism $F(f) \in \mathcal{D}(F(X), F(Y))$,

subject to the following requirements:

compatibility with composition: $F(g \circ f) = F(g) \circ F(f)$ for $f \in \mathcal{C}(X,Y)$ and $g \in \mathcal{C}(Y,Z)$;

compatibility with identities: $F(\mathrm{id}_X) = id_{F(X)}$ for any object X of $\mathcal{C}$.

3.4 Product, coproducts and pushouts

So far, we've calculated the fundamental groups for very few spaces: for convex subspaces of $\mathbb{R}^n$, for the circle S^1 and products of the circle. The main technique for calculating the fundamental group of more complicated spaces X is to write X as a union of open subspaces X_1 and X_2 such that the fundamental groups of X_1, X_2 and the intersection $X_1 \cap X_2$ are already known. The Seifert van Kampen Theorem then gives a formula for the fundamental of X in terms of the fundamental groups of X_1, X_2, $X_1 \cap X_2$ and the group homomorphisms

$$\begin{tikzcd} \pi_1(X_1 \cap X_2) \arrow[r, "(j_1)_*"] \arrow[d, "(j_2)_*"'] & \pi_1(X_1) \\ \pi_1(X_2) & \end{tikzcd}$$

induced by the inclusion maps of $X_1 \cap X_2$ into X_1 resp. X_2.

Perhaps surprisingly, the statement of the Seifert van Kampen Theorem can be stated conceptually by saying that "the fundamental group functor preserves pushouts". The goal of this section is define what a "pushout" is and what "preserving pushouts" means. Before discussing pushouts we define the more basic notions "categorical product" and "categorical coproduct".

3.4.1 Products

Earlier this semester we have defined the Cartesian product $X_1 \times X_2$ of topological spaces. We noticed that it is easy to construct continuous maps $f\colon Y \to X_1 \times X_2$ from some topological space Y to the product: neglecting continuity for a minute, a map f to the Cartesian product $X_1 \times X_2$ (considered just as a set) is determined by its component maps $f_1\colon Y \to X_1$ and $f_2\colon Y \to X_2$. The product topology on $X_1 \times X_2$ had the cool feature that a map $f\colon Y \to X_1 \times X_2$ is continuous if and only if both component maps f_1, f_2 are continuous. Summarizing, we can say that a continuous map $f\colon Y \to X_1 \times X_2$ is uniquely determined by a pair of continuous maps $f_1\colon Y \to X_1$ and $f_2\colon Y \to X_2$. Moreover, the component map f_i is given by the composition $p_i \circ f$ of f with the projection map $p_i\colon X_1 \times X_2 \to X_i$.

It turns out that this property can neatly expressed in terms of the commutative diagram:

$$\begin{tikzcd} & X_1 \\ Y \arrow[ur, "f_1"] \arrow[r, dashed, "\exists! f"] \arrow[dr, "f_2"'] & X_1 \times X_2 \arrow[u, "p_1"'] \arrow[d, "p_2"] \\ & X_2 \end{tikzcd}$$

This diagram should be interpreted as the statement that given the commutative diagram consisting of all the maps represent by *solid arrow*, there is a unique map f indicated by the dashed arrow which makes the whole diagram commutative. In view of the fact that the commutativity of the bottom and top triangle amounts to saying that f_1 and f_2 are the component maps of f, this diagram expresses concisely the result of our discussion above.

The benefit of describing a property in terms of a commutative diagram is that the same statement can be made *in any category*. This motivates the following definition.

Definition 3.23. Let X_1, X_2 be objects in a category $\mathcal{C}$. An object X in $\mathcal{C}$ is called the *categorical product* (often denoted $X_1 \times X_2$) if there are morphism $p_1\colon X \to X_1$ and $p_2\colon X \to X_2$ such that the diagram $X_1 \overset{p_1}{\leftarrow} X \overset{p_2}{\rightarrow} X_2$ has the property expressed by the commutative diagram

$$\begin{tikzcd}
 & X_1 \\
Y \arrow[ur, "f_1"] \arrow[r, dashed, "\exists! f"] \arrow[dr, "f_2"'] & X \arrow[u, "p_1"'] \arrow[d, "p_2"] \\
 & X_2
\end{tikzcd} \tag{3.24}$$

Expressed in words, this is: for any pair of morphisms $f_1\colon Y \to X_1$, $f_2\colon Y \to X_2$, there is a unique morphism $f\colon Y \to X$ making the diagram commutative.

Remark 3.25. The adjective "the" in front of "categorical product" has to be taken with a grain of salt: the object X is determined by this property *up to isomorphism*. To see this, suppose that $X_1 \overset{p_1'}{\leftarrow} X' \overset{p_2'}{\rightarrow} X_2$ is another object with morphisms to X_1 and X_2 which also satisfies that property expressed by the diagram above. Then consider the diagram

$$\begin{tikzcd}
 & X_1 & \\
X' \arrow[ur, "p_1'"] \arrow[rr, shift right, "f"'] \arrow[dr, "p_2'"'] & & X \arrow[ul, "p_1"'] \arrow[ll, shift right, "g"'] \arrow[dl, "p_2"] \\
 & X_2 &
\end{tikzcd}$$

The morphism f exists by the property (3.24) for X (applied to $Y = X'$), g exists by the property (3.24) for X' (applied to $Y = X$), and the composition $f \circ g$ is the identity morphism id_X by the uniqueness statement of the property (3.24) for X (applied to $Y = X$).

Lemma 3.26. *Let $\mathcal{C}$ be the category* Set, Grp, Vect, Top *or* Top_*. *Then the categorical product of objects X_1 and X_2 is given by the usual Cartesian product $X_1 \times X_2$ equipped with the usual projection maps $p_i\colon X_1 \times X_2 \to X_i$ for $i = 1, 2$.*

Proof. This was already proved in the category Top at the beginning of this section, which motivated Definition 3.23 of the categorical product. The argument is completely analogous in the other cases: a set map $f\colon Y \to X_1 \times X_2$ to the Cartesian product of two sets X_1, X_2 is uniquely determined by the pair of component maps $f_1\colon Y \to X_1$ and $f_2\colon Y \to X_2$. This proves the statement for $\mathcal{C} = \mathsf{Set}$. For $\mathcal{C} = \mathsf{Top}$, it follows from the fact that f is continuous if and only if both component maps f_1, f_2 are continuous. Similarly, for $\mathcal{C} = \mathsf{Top}_*$ the map f is continuous and basepoint preserving if and only if f_1 and f_2 are; for $\mathcal{C} = \mathrm{Grp}$ (resp. $\mathcal{C} = \mathsf{Vect}$) the map f is a group homomorphism (resp. linear) if both component maps are. □

3.4.2 Coproducts

Before characterizing the coproduct of objects X_1, X_2 in a category $\mathcal{C}$ by a universal property in Definition 3.28, we discuss disjoint unions of sets as a motivating example.

Definition 3.27. Let X_1, X_2 be sets. The *disjoint union* of X_1 and X_2, denoted $X_1 \amalg X_2$ is defined to be the set

$$X_1 \amalg X_2 := \{(x,1) \mid x \in X_1\} \cup \{(x,2) \mid x \in X_2\} \subset (X_1 \cup X_2) \times \{1,2\}.$$

Let $X_1 \xrightarrow{i_1} X_1 \amalg \xleftarrow{i_2} X_2$ be the maps defined by $i_1(x) := (x,1)$ and $i_2(x) := (x,2)$.

Question: Can the disjoint union of two sets be characterized up to isomorphism by a universal property, similar to the universal property (3.24) for the product of sets?

We observe that the images of i_1 and i_2 are *disjoint* subsets of $X_1 \amalg X_2$ whose union is all of $X_1 \amalg X_2$. Hence any map f from $X_1 \amalg X_2$ to some set Y is uniquely determined by its restriction to the image of i_1 resp. i_2. Since the maps i_1, i_2 are injective, this means that f is uniquely determined by the compositions $f_1 := f \circ i_1$ and $f_2 := f \circ i_2$. As in the case of the Cartesian product, this property of the disjoint union of sets can neatly expressed by the following commutative diagram.

$$\begin{tikzcd}
X_1 \arrow[d, "i_1"'] \arrow[dr, "f_1"] & \\
X_1 \amalg X_2 \arrow[r, dashed, "\exists! f"] & Y \\
X_2 \arrow[u, "i_2"] \arrow[ur, "f_2"'] &
\end{tikzcd}$$

Definition 3.28. Let X_1, X_2 be objects of a category $\mathcal{C}$. An object X in $\mathcal{C}$ is called the *coproduct* of X_1 and X_2 (often denoted $X_1 \amalg X_2$) if there are morphisms

$$X_1 \xrightarrow{i_1} X \xleftarrow{i_2} X_2$$

such that this pair of maps satisfies the universal property expressed by the following commutative diagram

$$\begin{tikzcd}
X_1 \arrow[d, "i_1"'] \arrow[dr, "f_1"] & \\
X \arrow[r, dashed, "\exists! f"] & Y \\
X_2 \arrow[u, "i_2"] \arrow[ur, "f_2"'] &
\end{tikzcd} \tag{3.29}$$

While the universal property (3.29) determines the object X up to isomorphism, the category $\mathcal{C}$ might not have an object with this property. For example, the coproduct of two sets is given by their disjoint union. However, if we consider the category $\mathcal{C}$ whose objects are sets of cardinality 3 and whose morphisms are maps between these sets, then there are no coproducts in $\mathcal{C}$. So it requires an explicit *construction* to show that coproducts exists in a given category $\mathcal{C}$.

Theorem 3.30. *Coproducts (of two arbitrary objects X_1, X_2) exist in the categories* Set*,* Top*,* Top_**,* Vect *and* Grp*. The following table shows the usual notation and terminology for the coproduct in these categories.*

category	*coproduct of objects*
Set	$X_1 \amalg X_2$, *the disjoint union of sets* X_1, X_2
Top	$X_1 \amalg X_2$, *the disjoint union of topological spaces* X_1, X_2
Top_*	$X_1 \vee X_2$, *the wedge product of topological spaces* X_1, X_2
Grp	$X_1 * X_2$, *the free product of groups* X_1, X_2

The disjoint union of sets and its universal property was for us the example motivating the general definition of the coproduct, and hence the theorem holds in case of the category Set. To proof the result for the other categories, we will go through them one by one, first giving in each case an explicit construction of a candidate for the coproduct (including the morphisms i_1 and i_2), namely the disjoint union of topological spaces, the wedge product of pointed topological spaces, and the free product of groups. Then we show in each case that this object satisfies the universal property of the coproduct.

Definition 3.31. (The disjoint union of topological spaces). Let X_1, X_2 be topological spaces. The *disjoint union* $X_1 \amalg X_2$ is the topological space whose underlying set is the disjoint union of X_1 and X_2, considered as sets. The topology on the set $X_1 \amalg X_2$ is defined by declaring a subset $U \subset X_1 \amalg X_2$ to be open if and only if $i_1^{-1}(U)$ is an open subset of X_1 and $i_2^{-1}(U)$ is an open subset of X_2. With this topology on $X_1 \amalg X_2$ the maps

$$X_1 \xrightarrow{i_1} X_1 \amalg X_2 \xleftarrow{i_2} X_2 \tag{3.32}$$

are both continuous.

Lemma 3.33. *The diagram* (3.32) *satisfies the universal property* (3.29)*. In particular, the disjoint union of topological spaces is the coproduct in the category* Top*.*

Proof. It is clear that there is at most one continuous map $f\colon X_1 \amalg X_2 \to Y$ making the diagram (3.29) for given maps f_1, f_2, since the underlying set $X_1 \amalg X_2$ is the coproduct of the sets X_1, X_2, and hence there is exactly one map $f\colon X_1 \amalg X_2 \to Y$ making the diagram commutative (without insisting on its continuity).

So it remains to show that f is continuous if f_1 and f_2 are. So let V be an open subset of Y. Then $f^{-1}(V)$ is open, since $i_\ell^{-1}(f^{-1}(V)) = (f \circ i_\ell)^{-1}(V) = f_\ell^{-1}(V)$ is open, since f_ℓ is continuous for $\ell = 1, 2$. □

Let (X_1, x_1), (X_2, x_2) be pointed topological spaces. We need to come up with a candidate for the coproduct of these pointed spaces. Note that the disjoint union $X_1 \amalg X_2$ is not a good candidate, since we would like the maps (3.32)

$$X_1 \xrightarrow{i_1} X_1 \amalg X_2 \xleftarrow{i_2} X_2$$

to be basepoint preserving, but $i_1(x_1) \neq i_2(x_2)$. The way to fix this is to pass to a quotient space of $X_1 \amalg X_2$ where we identify these two points.

Definition 3.34. Let (X_1, x_1), (X_2, x_2) be pointed topological spaces. The quotient space

$$X_1 \vee X_2 := (X_1 \amalg X_2)/\{i_1(x_1), i_2(x_2)\}$$

equipped with the base point $*$ given by the equivalence class represented by these two points $i_1(x_1)$, $i_2(x_2)$ is the *wedge product* of the pointed spaces X_1, X_2.

Lemma 3.35. *For* $\ell = 1, 2$ *let* $k_\ell\colon X_\ell \to X_1 \amalg X_2$ *be the composition of* $i_\ell\colon X_\ell \to X_1 \amalg X_2$ *and the projection map* $p\colon X_1 \amalg X_2 \to X_1 \vee X_2$*. Then the diagram*

$$X_1 \xrightarrow{k_1} X_1 \vee X_2 \xleftarrow{k_2} X_2 \tag{3.36}$$

satisfies the universal property (3.29)*. In particular, the wedge product* $X_1 \vee X_2$ *of pointed topological spaces* X_1, X_2 *is the coproduct in the category* Top_**.*

Proof. The universal property we need to check is expressed by the commutative diagram

$$\begin{tikzcd}
X_1 \arrow[d, "k_1"'] \arrow[dr, "f_1"] & \\
X_1 \vee X_2 \arrow[r, dashed, "\exists! f"] & Y \\
X_2 \arrow[u, "k_2"] \arrow[ur, "f_2"'] &
\end{tikzcd} \tag{3.37}$$

By the universal property of the disjoint union $X_1 \amalg X_2$, there is a unique continuous map $\widetilde{f}\colon X_1 \amalg X_2 \to Y$ such that $\widetilde{f} \circ i_1 = f_1$ and $\widetilde{f} \circ i_2 = f_2$. Since f_1, f_2 are basepoint preserving, $\widetilde{f}(i_1(x_1)) = \widetilde{f}(x_1) = y_0$ and $\widetilde{f}(i_2(x_2)) = \widetilde{f}(x_2) = y_0$, where $y_0 \in Y$ is the basepoint in Y. This implies that $\widetilde{f}$ factors through the quotient space $X_1 \vee X_2 = (X_1 \amalg X_1)/\{i_1(x_1), i_2(x_2)\}$, i.e., $\widetilde{f}$ can be written as composition

$$X_1 \amalg X_2 \xrightarrow{p} X_1 \vee X_2 \xrightarrow{f} Y$$

for a unique map f. This map is basepoint preserving. It also is continuous: by the continuity criterion for maps out of a quotient space 1.24, the map f is continuous if and only if the composition $f \circ p = \widetilde{f}$ is continuous. $\square$

Definition 3.38. Let X_1, X_2 be groups. Their *free product* $X_1 * X_2$ is the group whose elements are equivalence classes of words $s_1 \dots s_k$ whose letter s_i belong to X_1 or X_2 (we assume that X_1, X_2 are disjoint as sets). The equivalence relation $\sim$ on these words is generated by

(1) $s_1 \dots s_i \dots s_k \sim s_1 \dots \widehat{s_i} \dots s_k$ if s_i is the identity element of X_1 or X_2, and $s_1 \dots \widehat{s_i} \dots s_k$ is the word obtained by deleting the letter s_i.

(2) $s_1 \dots s_i s_{i+1} \dots s_k \sim s_1 \dots (s_i \cdot s_{i+1}) \dots s_k$ if s_i and s_{i+1} both belong to X_1 or to X_2 and $s_i \cdot s_{i+1} \in X_j$ denotes their product in that group.

The multiplication in $X_1 * X_2$ is induced by concatenation of words. The identity element is represented by the empty word, and the inverse of the element represented by the word $s_1 \dots s_k$ is given by $s_k^{-1} \dots s_1^{-1}$. Let

$$X_1 \xrightarrow{i_1} X_1 * X_2 \xleftarrow{i_2} X_2 \tag{3.39}$$

be the group homomorphisms given by sending an element s of X_1 or X_2 to the element $[s] \in X_1 * X_2$ represented by the one-letter-word s.

Lemma 3.40. *The diagram* (3.39) *satisfies the universal property* (3.29). *In particular, the free product $X_1 * X_2$ of groups X_1, X_2 is the coproduct in the category* Grp.

The proof of this lemma is left as a homework problem.

3.4.3 Pushouts

Before defining what a pushout is in a category in Definition 3.42, we consider a motivating example of a pushout in the category of topological spaces.

Example 3.41. Let X_1, X_2 be open subsets of a topological space X. Then considering the inclusion maps relating X_1, X_2, X and $X_1 \cap X_2$ we have the following commutative square in the category Top of topological spaces:

$$\begin{tikzcd} X_1 \cap X_2 \arrow[r, "j_1"] \arrow[d, "j_2"'] & X_1 \arrow[d, "k_1"] \\ X_2 \arrow[r, "k_2"] & X \end{tikzcd}$$

Let $f_1 \colon X_1 \to Y$ and $f_2 \colon X_2 \to Y$ be continuous maps which agree on the subspace $X_1 \cap X_2$. Then there is well-defined map $f \colon X \to Y$ whose restriction to X_1 is the map f_1 and whose restriction to X_2 is the map f_2. Moreover, by an earlier homework problem, the continuity of f_1 and f_2 imply the continuity of the map f (here we use the assumption that X_1, X_2 are open subsets of X).

Definition 3.42. Let $\mathcal{C}$ be a category, and let

$$\begin{tikzcd} A \arrow[r, "j_1"] \arrow[d, "j_2"'] & X_1 \arrow[d, "k_1"] \\ X_2 \arrow[r, "k_2"] & X \end{tikzcd}$$

be a commutative diagram of objects and morphisms in $\mathcal{C}$. This diagram is a *pushout diagram* or *pushout square* if it satisfies the universal property expressed by the diagram

$$\begin{tikzcd} A \arrow[r, "j_1"] \arrow[d, "j_2"'] & X_1 \arrow[d, "k_1"] \arrow[ddr, bend left, "f_1"] & \\ X_2 \arrow[r, "k_2"] \arrow[drr, bend right, "f_2"'] & X \arrow[dr, dashed, "\exists! f"] & \\ & & Y \end{tikzcd} \tag{3.43}$$

The object X is called the *pushout* of the diagram

$$\begin{tikzcd} A \arrow[r, "j_1"] \arrow[d, "j_2"'] & X_1 \\ X_2 & \end{tikzcd} \tag{3.44}$$

Theorem 3.45. *The pushout of any diagram* (3.44) *exists in the categories* Top, Top$_*$ *and* Grp*. The following table shows the usual notation and terminology for the pushout in these categories.*

category	*coproduct of objects*
Set	$X_1 \cup_A X_2$
Top	$X_1 \cup_A X_2$
Top$_*$	$X_1 \cup_A X_2$
Grp	$X_1 *_A X_2$*, the amalgamated free product*

There seems to be no standard terminology for $X_1 \cup_A X_2$. Both of the notations $X_1 \cup_A X_2$ and $X_1 *_A X_2$ suppress the dependence of this object on the morphisms j_1, j_2. To indicate the dependence, some people use the notation $X_1 \cup_{j_1,A,j_2} X_2$ and $X_1 *_{j_1,A,j_2} X_2$.

Definition 3.46. Let $j_1\colon A \to X_1$ and $j_2\colon A \to X_2$ be maps of sets. Then we define

$$X_1 \cup_A X_2 := (X_1 \amalg X_2)/\sim,$$

where the equivalence relation $\sim$ is generated by $i_1(j_1(a)) \sim i_2(j_2(a))$ for $a \in A$. Here $i_\ell\colon X_\ell \to X_1 \amalg X_2$ are the inclusion maps featured in the definition of the disjoint union **??**.

3.5 The Seifert van Kampen Theorem

Theorem 3.47. *Let U, V be open subsets of topological space X such that $U \cup V = X$ and $U \cap V$ is path connected. Let*

$$\begin{tikzcd} U \cap V \arrow[r, "j^U"] \arrow[d, "j^V"] & U \arrow[d, "i^U"] \\ V \arrow[r, "i^V"'] & X \end{tikzcd}$$

be the pushout diagram of topological spaces given by the inclusion maps (see Example **??** *and Definition 3.42). Then for $x_0 \in U \cap V$ the induced commutative diagram of fundamental groups*

$$\begin{tikzcd} \pi_1(U, x_0) \cap V \arrow[r, "j^U"] \arrow[d, "j^V"] & \pi_1(U, x_0) \arrow[d, "i^U"] \\ \pi_1(V, x_0) \arrow[r, "i^V"'] & \pi_1(X, x_0) \end{tikzcd}$$

is a pushout diagram in the category of groups.

The Seifert van Kampen Theorem says in particular that the fundamental group functor preserves pushouts (under suitable additional assumptions).

Corollary 3.48. *With the assumptions of the theorem, the fundamental group $\pi_1(X, x_0)$ is isomorphic to the free amalgamated product $\pi_1(U, x_0) *_{\pi_1(U\cap V, x_0)} \pi_1(V, x_0)$.*

Corollary 3.49. *Let (X_1, x_1), (X_2, x_2) be pointed topological spaces such that x_1, x_2 have contractible open neighborhoods U_1, U_2. Let $j_1 \colon X_1 \to X_1 \vee X_2$ and $j_2 \colon X_2 \to X_1 \vee X_2$ be the inclusion maps. Then the map*

$$\pi_1(X_1) * \pi_1(X_2) \longrightarrow \pi_1(X_1 \vee X_2)$$

given by

$$\begin{aligned} \pi_1(X_1) \ni c_1 &\mapsto (j_1)_*(c) \in \pi_1(X_1 \vee X_2) \\ \pi_1(X_2) \ni c_2 &\mapsto (j_2)_*(c_2) \in \pi_1(X_1 \vee X_2) \end{aligned}$$

is an isomorphism of groups.

Proposition 3.50. $\pi_1(K) = \langle a, b \mid aba^{-1}b \rangle$.

Similar propositions for Σ_g, X_k.

4 Covering spaces

4.1 Homotopy lifting property for covering spaces

Definition 4.1. A map $p \colon \widetilde{X} \to X$ is a *covering map* if for every $x \in X$ there is an open neighborhood U such that $p^{-1}(U)$ is the union of disjoint subsets $U_i \subset \widetilde{X}$ such that the map $p_{U_i} \colon U_i \to U$ is a homeomorphisms for each U_i (here p_{U_i} denotes the restriction of p to $U_i \subset \widetilde{X}$). A subset $U \subset X$ with this property is called *evenly covered*. The space $\widetilde{X}$ is called a *covering space of* X.

Proposition 4.2. (Unique path lifting for covering spaces). *Let $f \colon I \to X$ be a path and let $\widetilde{x}_0 \in \widetilde{X}$ be a point with $p(\widetilde{x}_0) = f(0)$. Then there is a unique path $\widetilde{f} \colon I \to \widetilde{X}$ such that*

1. *$p \circ f = f$, and*
2. *$\widetilde{f}(0) = \widetilde{x}_0$.*

A path $\widetilde{f} \colon I \to \widetilde{X}$ with the first property is called a *lift* of the path $f \colon I \to X$. Using that terminology, the proposition says that every path $f \colon I \to X$ with starting point $f(0) = x_0$ has a unique lift $\widetilde{f}$ with prescribed starting point $\widetilde{x}_0 \in p^{-1}(x_0)$.

Proof. We begin by the constructing the lift $\widetilde{f}$. The idea is to find points $0 = t_0 < t_2 < \cdots < t_k = 1$ such that each subinterval $[t_i, t_{i+1}]$ via f maps to an evenly covered open set $U \subset X$. Then the lift $\widetilde{f}\colon I \to \widetilde{X}$ is constructed first over $[0, t_1]$, then $[t_1, t_2]$, e.t.c, using in each step the homeomorphism $p_{U_i}\colon U_i \xrightarrow{\approx} U$.

To construct the points $0 = t_0 < t_2 < \cdots < t_k = 1$, we note that for each $s \in I$, the point $f(s) \in X$ is contained in some evenly covered open neighborhood $U \subset X$. Then $f^{-1}(U)$ is an open neighborhood of s, and hence there is some interval (a_s, b_s) such that $s \in (a_s, b_s) \cap I \subset f^{-1}(U)$ (since these form a basis for the subspace topology of $I \subset \mathbb{R}$). The open subsets $(a_s, b_s) \cap I$, $s \in I$, form an open cover of I, and hence by compactness of I, there is a finite set $S = \{s_1, \ldots, s_\ell\}$ such that the intervals $(a_s, b_s) \cap I$ for $s \in S$ still cover I. Without loss of generality we may assume that none of these intervals is contained in another, and that $b_{s_i} < b_{s_j}$ for $i < j$. We note that

1. 0 belongs to $(a_{s_1}, b_{s_1}) \cap I$ (if not, $0 \in (a_{s_i}, b_{s_i}) \cap I$ for some $i > 1$, but since $b_{s_i} \geq b_{s_1}$, this implies $(a_{s_1}, b_{s_1}) \cap I \subset (a_{s_i}, b_{s_i}) \cap I$ contradicting our assumptions).

2.

is a finite set $T \subset I$ *finite* collection of these intervals that still cover I. This allows us to pick elements finitely many points $0 = t_0 < t_2 < \cdots < t_k = 1$ such that each $[t_{i-1}, t_i]$ is contained in one of these

subdivide the interval I into □

$$
\begin{tikzcd}
Y \times \{0\} \arrow[r, "\widetilde{f}_0"] \arrow[d, "i"] & \widetilde{X} \arrow[d, "p"] \\
Y \times I \arrow[r, "f"] \arrow[ur, "\widetilde{f}"] & X
\end{tikzcd}
$$

Proposition 4.3. (Unique homotopy lifting property for covering spaces). *Let $p\colon \widetilde{X} \to X$ be a covering space, $f_t\colon Y \to X$ a homotopy, and $\widetilde{f}_0\colon Y \to \widetilde{X}$ a map lifting f_0. Then there is a unique homotopy $\widetilde{f}_t$ of $\widetilde{f}_0$ that lifts f_t.*

For $Y = I$, the map f_0 is a path in X and $\widetilde{f}_0$ is a lift to $\widetilde{X}$. The maps f_t and $\widetilde{f}_t$ are homotopies, not necessarily relative endpoints. We recall that f_t is a *homotopy relative endpoints* or, equivalently, f_t *preserves endpoints* if $f_t(0) = f_0(0)$ and $f_t(1) = f_0(1)$ for all $t \in [0, 1]$. Since for paths we are primarily interested in homotopies relative endpoints, we will frequently use the following observation.

Observation 4.4. *The homotopy $\widetilde{f}_t$ preserves endpoints if and only if f_t preserves endpoints.*

It is clear that if $\widetilde{f}_t$ preserves endpoints, then so does $f_t = p \circ \widetilde{f}_t$. Conversely, assume that f_t preserves endpoints. Let $x := f_t(0) \in X$ be the common starting point of the paths f_t. Then $p(\widetilde{f}_t(0)) = x$ and hence $\widetilde{f}_t(0) \in p^{-1}(x)$. Since $\widetilde{f}_t(0)$ depends continuously on $t \in I$, the image of this map is connected, but that forces $\widetilde{f}_t(0)$ to be constant. The same arguments shows $\widetilde{f}_t(1)$ is constant and hence $\widetilde{f}_t$ is a homotopy relative endpoints.

Proposition 4.5. *Let $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a covering map. Then*

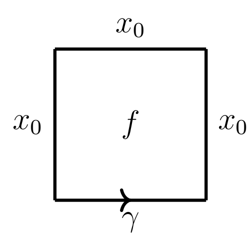
(i) the induced homomorphism $p_*\colon \pi_1(\widetilde{X}, \widetilde{x}_0) \to \pi_1(X, x_0)$ is injective.

(ii) A based loop γ in (X, x_0) represents an element in the image of p_* if and only if its unique lift $\widetilde{\gamma}\colon I \to \widetilde{X}$ with $\widetilde{\gamma}(0) = \widetilde{x}_0$ is a loop, i.e., $\widetilde{\gamma}(1) = \widetilde{x}_0$.

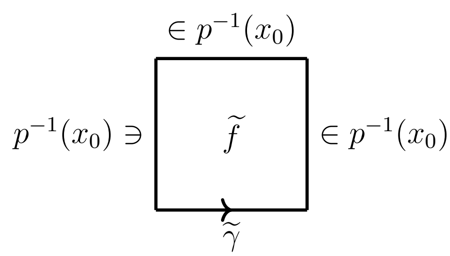
Proof. The crucial tool is the homotopy lifting property for $Y = I$. To prove part (i), let $\widetilde{\gamma}$ be a based loop in $(\widetilde{X}, \widetilde{x}_0)$ in the kernel of p_*. Let $f\colon I \times I \to X$ be a homotopy from $p \circ \widetilde{\gamma}$ to the constant loop at the base point x_0, i.e.,

$$f(s,0) = p \circ \widetilde{\gamma}(s) \qquad f(s,1) = x_0 \qquad f(0,t) = f(1,t) = x_0$$

The following picture might be helpful to visualize what we know about the map f from the square $I \times I$ to X: restricted to the bottom edge, it is the map $\gamma := p \circ \widetilde{\gamma}$, while it maps the three other edges to the base point x_0. We indicate this by labeling the edges by x_0 resp. γ.

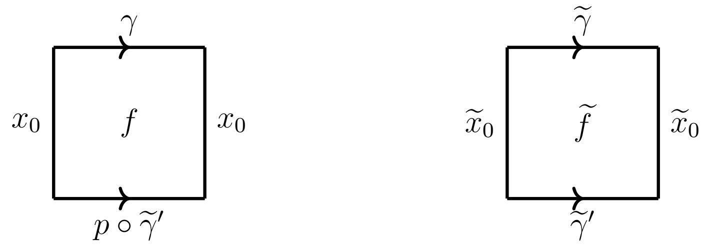


Using the homotopy lifting property (see Proposition 4.3), let $\widetilde{f}\colon I \times I \to \widetilde{X}$ be a lift of the homotopy f with $f(s,0) = \widetilde{\gamma}(s)$. The following picture represents what we know about the value of $\widetilde{f}$ on the edges of the square $I \times I$.



The union of the left, top and right edge of $I \times I$ are a path-connected subspace S of $I \times I$, and hence the restriction $\widetilde{f}_{|S} \colon S \to p^{-1}(x_0)$ to this subspace must be constant by Lemma **??**. Since $\widetilde{f}(0,0) = \widetilde{\gamma}(0) = \widetilde{x}_0$, then $\widetilde{f}$ maps all of this subspace to the base point $\widetilde{x}_0$. In particular, $\widetilde{\gamma}(1) = \widetilde{f}(1,0) = \widetilde{x}_0$ and hence $\widetilde{\gamma}$ is a loop based a $\widetilde{x}_0$ as claimed.

To prove part (ii), suppose that γ is a loop in X based at x_0 such that $[\gamma] \in \pi_1(X, x_0)$ is in the image of p_*, i.e., there is some based loop $\widetilde{\gamma}'$ in $(\widetilde{X}, \widetilde{x}_0)$ such that $p_*[\widetilde{\gamma}'] = [p \circ \gamma']$ is equal to $[\gamma] \in \pi_1(X, x_0)$. Let f be a homotopy (relative endpoints) from $p \circ \gamma'$ to γ, and let $\widetilde{f}$ be a lift of f with $\widetilde{f}(s, 0) = \widetilde{\gamma}'$ (which exists by the homotopy lifting property 4.3). The following pictures represent what we know about f and $\widetilde{f}$.



In more detail: the map $\widetilde{f}$ maps the left edge to $p^{-1}(x_0)$ and the point $(0,0)$ to $\widetilde{\gamma}(0) = \widetilde{x}_0$, and hence the whole edge must map to $\widetilde{x}_0$. Similarly, $\widetilde{f}$ must map to right edge to $\widetilde{x}_0$. The restriction of $\widetilde{f}$ to the top edge is a lift of γ with starting point $\widetilde{f}(0,1) = \widetilde{x}_0$ and hence by the uniqueness of lifts of paths, this is the path $\widetilde{\gamma}$. It follows that $\widetilde{\gamma}(1) = \widetilde{f}(1,1) = \widetilde{x}_0$, i.e., $\widetilde{\gamma}$ is a loop in $\widetilde{X}$ based at $\widetilde{x}_0$ as claimed. □

We recall that we defined the winding number $W(\gamma) \in \mathbb{Z}$ of a based loop γ in (S^1, x_0), x_0 by making use of the covering map $p \colon \mathbb{R} \to S^1$, $t \mapsto e^{2\pi i t}$. More precisely, if γ is a based loop in (S^1, x_0), x_0 we defined

$$W(\gamma) := \widetilde{\gamma}(1) \in \mathbb{Z}$$

where $\widetilde{\gamma} \colon I \to \mathbb{R}$ is a lift with $\widetilde{\gamma}(0) = \widetilde{x}_0 = 0 \in \mathbb{R}$, and hence $\widetilde{\gamma}(1) \in p^{-1}(x_0) = \mathbb{Z} \subset \mathbb{R}$.

This construction can be generalized to covering maps $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$. If γ is a based loop in (X, x_0) we define

$$W(\gamma) := \widetilde{\gamma}(1) \in p^{-1}(x_0),$$

where $\widetilde{\gamma} \colon I \to \widetilde{X}$ is the unique lift of γ with $\widetilde{\gamma}(0) = x_0$.

In the case of the covering map $\mathbb{R} \to S^1$ we showed that $W(\gamma)$ depends only on the homotopy class $[\gamma] \in \pi_1(S1, x_0)$, and that the map $\pi_1(S^1, x_0) \to \mathbb{Z}$, $[\gamma] \mapsto W(\gamma)$ is a bijection. This statement generalizes to the following result.

Proposition 4.6. *Let $\widetilde{X}$ be path-connected and let $p\colon (\widetilde{X},\widetilde{x}_0)\to(X,x_0)$ be a covering map. Let $G:=\pi_1(X,x_0)$ and $H\subset G$ the subgroup $H:=p_*\pi_1(\widetilde{X},\widetilde{x}_0)$. Let $H\backslash G:=\{Hg\mid g\in H\}$ be the set of left H-cosets. Then the map*

$$\Psi\colon H\backslash G\longrightarrow p^{-1}(x_0)\qquad \textit{given by}\qquad [\gamma]\mapsto W(\gamma)$$

is a well-defined bijection. In particular, the number of sheets of $\widetilde{X}\to X$ (the cardinality of $p^{-1}(x_0)$) is equal to the index $[G:H]$ of the subgroup $H\subset G$ (the cardinality of $H\backslash G$).

Proof. To show that Ψ is well-defined, let $g\in G=\pi_1(X,x_0)$ be represented by a based loop γ in (X,x_0) and let $h\in H=p_*\pi_1(\widetilde{X},\widetilde{x}_0)$ be represented by a based loop δ in (X,x_0). We let $\widetilde{\gamma},\widetilde{\delta}\colon I\to\widetilde{X}$ be the unique lifts of γ resp. δ starting at $\widetilde{x}_0$. Since $[\delta]$ is in $p_*\pi_1(\widetilde{X},\widetilde{x}_0)$ its lift $\widetilde{\delta}$ is a based loop in $(\widetilde{X},\widetilde{x}_0)$ by Proposition 4.5. It follows the endpoint of $\widetilde{\delta}$ is the starting point of $\widetilde{\gamma}$ and hence the concatenation $\widetilde{\delta}*\widetilde{\gamma}$ is defined. Moreover,

$$p\circ(\widetilde{\delta}*\widetilde{\gamma})=(p\circ\widetilde{\delta})*(p\circ\widetilde{\gamma})=\delta*\gamma,$$

and hence $\widetilde{\delta}*\widetilde{\gamma}$ is the unique lift of $\delta*\gamma$ with starting point $\widetilde{x}_0$. This implies that

$$\Psi(hg)=\Psi([\delta][\gamma])=\Psi([\delta*\gamma])=W(\delta*\gamma)=(\widetilde{\delta}*\widetilde{\gamma})(1)=\widetilde{\gamma}(1)=W(\gamma)=\Psi([\gamma])=\Psi(g)$$

and hence Ψ is well-defined.

It is clear that Ψ is surjective, since due the assumption that $\widetilde{X}$ is path-connected for any $\widetilde{x}\in p^{-1}(x_0)$ there is a path $\widetilde{\gamma}\colon I\to\widetilde{X}$ from $\widetilde{x}_0$ to x. Then $\gamma:=p\circ\widetilde{\gamma}$ is a loop in (X,x_0), and hence $\Psi([\gamma])=W(\gamma)=\widetilde{\gamma}(1)=\widetilde{x}$.

To see that Ψ is injective, let $g_1,g_2\in\pi_1(X,x_0)$ be elements with $\Psi(g_1)=\Psi(g_2)$. Let γ_i be a based loop in (X,x_0) representing g_i and let $\widetilde{\gamma}_i\colon I\to\widetilde{X}$ be its unique lift with $\widetilde{\gamma}_i(0)=\widetilde{x}_0$. Then $\Psi(g_i)=\widetilde{\gamma}_i(1)$, and hence our assumption $\Psi(g_1)=\Psi(g_2)$ implies that the paths $\widetilde{\gamma}_1$, $\widetilde{\gamma}_1$ have the same endpoints. In particular, the concatenation $\widetilde{\gamma}_1*\overline{\widetilde{\gamma}}_2$ is defined and is a based loop in $(\widetilde{X},\widetilde{x}_0)$. It follows that

$$p_*[\widetilde{\gamma}_1*\overline{\widetilde{\gamma}}_2]=[p\circ(\widetilde{\gamma}_1*\overline{\widetilde{\gamma}}_2)]=[\gamma_1*\bar{\gamma}_2]=[\gamma_1][\gamma_2]^{-1}=g_1g_2^{-1},$$

showing that $g_1g_2^{-1}\in H$ and hence the left cosets Hg_1 and Hg_2 agree. □

Give a covering space $p\colon\widetilde{X}\to X$, it will be important for us to lift not just homotopies, but more general maps $f\colon Y\to X$. In other words, we are looking for a base point preserving map $\widetilde{f}$ making the diagram

$$\begin{tikzcd} & (\widetilde{X},\widetilde{x}_0) \arrow[d, "p"] \\ (Y,y_0) \arrow[ur, dashed, "\widetilde{f}"] \arrow[r, "f"] & (X,x_0) \end{tikzcd} \tag{4.7}$$

commutative.

There is an obvious necessary condition for the existence of a lift $\widetilde{f}$: such a lift induces a commutative diagram of fundamental groups

$$\begin{tikzcd} & \pi_1(\widetilde{X}, \widetilde{x}_0) \arrow[d, "p_*"] \\ \pi_1(Y, y_0) \arrow[ur, dashed, "\widetilde{f}_*"] \arrow[r, "f_*"] & \pi_1(X, x_0) \end{tikzcd}, \tag{4.8}$$

and hence $f_*\pi_1(Y, y_0)$ is contained in $p_*\pi_1(\widetilde{X}, \widetilde{x}_0)$. We will show that this is also a sufficient condition for the existence of a lift $\widetilde{f}$, provided the topological space Y isn't too crazy.

Definition 4.9. A topological space Y is *locally path-connected* if for any point $y \in Y$ and any neighborhood U of y there is an open neighborhood $V \subset U$ which is path-connected. More generally, if P is any property of a topological space (e.g., compact, connected,...), then Y is *locally* P if for any point $y \in Y$ and any neighborhood U of y there is an open neighborhood $V \subset U$ such that V has property P.

Example 4.10. (Path-connected versus locally path-connected). There are many examples of spaces which are locally path-connected, but non path-connected, for example the disjoint union of path-connected spaces is locally path-connected. An example of a space which is not locally path-connected is provided by the *topologist's sine curve* **??**, consisting of the union of the graph of the function $(0, \infty) \to \mathbb{R}$, $x \mapsto \sin(1/x)$ and the vertical line segment $\{0\} \times [-1, +1]$. As discussed then, the topologist's sine curve is connected, but not path-connected. The same argument shows in fact that any open neighborhood V of a point y on the vertical line segment is not path-connected (since it always contains points on the graph of $\sin(1/x)$; those cannot be reached by paths starting at y).

Even more interesting is that there are spaces which is path-connected, but not locally path-connected, for example, the *Warsaw circle*. This is a variant of the topologist's sine curve obtained by restricting the graph of $\sin(1/x)$ to some finite interval $(0, a)$ and connecting the point $(a, \sin(1/a))$ on the graph with the point $(0, 0)$ via an arc in $\mathbb{R}^2$ which intersects the topologist's sine curve only in those two points as shown in the figure below.

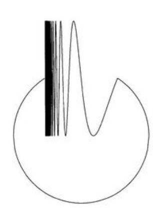


Figure 1: The Warsaw circle

The Warsaw circle is path-connected, since any point on the vertical line can be connected via a path running along the added arc to any point of graph. Adding that arc does not change the fact that any open neighborhood of a point on the vertical line segment is not path-connected, provided it is small enough that it doesn't contain the arc. In particular, the Warsaw circle is not locally path-connected.

Proposition 4.11. (Lifting Criterion). *Let* $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ *be a covering map and let* $f\colon (Y, y_0) \to (X, x_0)$ *be a basepoint preserving map whose domain* Y *is path-connected and locally path-connected. Then a lift* $\widetilde{f}$ *in the diagram* (4.8) *exists if and only if* $f_*\pi_1(Y, y_0) \subset \pi_1(\widetilde{X}, \widetilde{x}_0)$. *There is at most one such lift.*

Remark 4.12. The hypothesis that Y is locally path-connected cannot be dropped. This follows by showing that the statement above does not hold if Y is the Warsaw circle W, which is proved by showing

(a) The fundamental group of W vanishes, but

(b) the map $f\colon W \to S^1$ which wraps the Warsaw circle once around the circle S^1 does not have a lift $\widetilde{f}\colon W \to \mathbb{R}$.

Proof. We have argued above that $f_*\pi_1(Y, y_0) \subset \pi_1(\widetilde{X}, \widetilde{x}_0)$ is a necessary condition for the existence of a lift $\widetilde{f}$. It is also easy to see that there is at most one such lift: if f, f' are two lifts of f and $y \in Y$, let $\gamma\colon I \to Y$ be a path in Y from y_0 to y. Then $\widetilde{f} \circ \gamma$ and $\widetilde{f}' \circ \gamma$ are two paths in $\widetilde{X}$ which are both lifts of the path $f \circ \gamma$ in X with starting point $\widetilde{x}_0$. The uniqueness of lifted paths then implies $\widetilde{f}(y) = \widetilde{f} \circ \gamma(1) = \widetilde{f}' \circ \gamma(1) = f'(y)$.

The idea for *constructing* the lift $\widetilde{f}\colon Y \to \widetilde{X}$ is to use the *existence* of lifts of paths, similar to the way we used *uniqueness of path-lifting* to prove the *uniqueness of* $\widetilde{f}$: we define the map

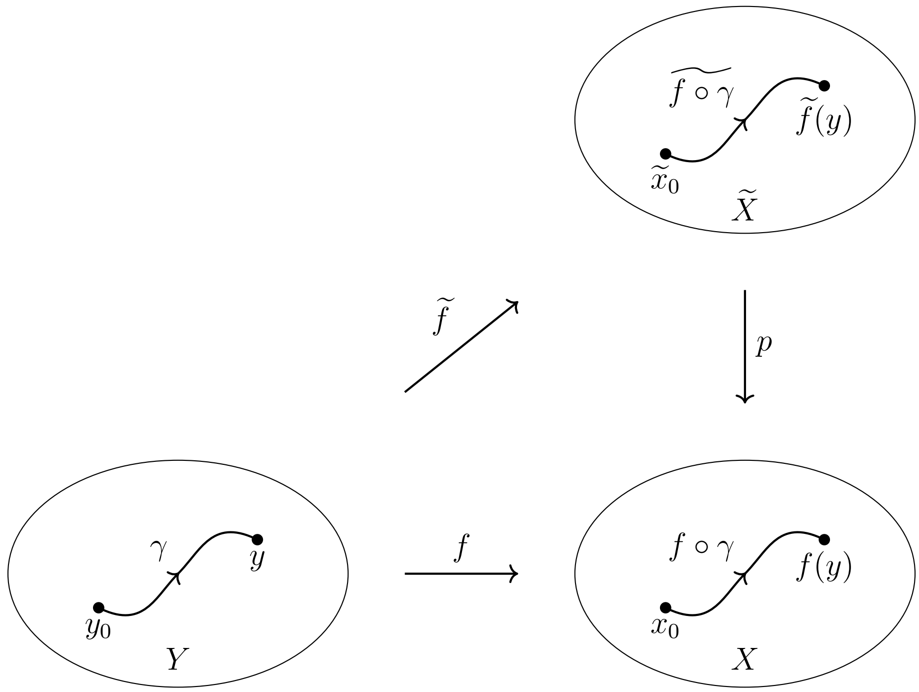
$$\widetilde{f}\colon Y \longrightarrow \widetilde{X} \qquad \text{by} \qquad f(y) := (\widetilde{f \circ \gamma})(1)$$

where

$\gamma\colon I \to Y$ is path from y_0 to y, and

$\widetilde{f \circ \gamma}\colon I \to \widetilde{X}$ is a lift of $f \circ \gamma\colon I \to X$ with starting point $\widetilde{x}_0$.

The following figure illustrates the various paths involved and their endpoints.



To show that $\widetilde{f}$ is well-defined, we need to verify that $\widetilde{f}(y)$ is independent of the choice of the path γ from the basepoint y_0 to y. So suppose that $\gamma' \colon I \to Y$ is another path from y_0 to y. The concatenation $\gamma * \bar{\gamma}'$ is then a based loop in (Y, y_0) and hence the loop $f \circ (\gamma * \bar{\gamma}') = (f \circ \gamma) * (f \circ \bar{\gamma}')$ is a based loop in (X, x_0) which represents an element in the image of $f_* \colon \pi_1(Y, y_0) \to \pi_1(X, x_0)$. By assumption, this element is then in the image of $p_* \colon \pi_1(\widetilde{X}, \widetilde{x}_0) \to \pi_1(X, x_0)$, which by Proposition 4.5 implies that the loop $(f \circ \gamma) * (f \circ \bar{\gamma}')$ lifts to a based loop $\widetilde{\delta} \colon I \to \widetilde{X}$ in $(\widetilde{X}, \widetilde{x}_0)$. By uniqueness of lifted paths, the first half of $\widetilde{\delta}$ is $\widetilde{f \circ \gamma}$ and the second half is $\widetilde{f \circ \gamma'}$ traversed backwards, with the common midpoint $\widetilde{f \circ \gamma}(1) = \widetilde{f \circ \gamma'}(1)$. This shows that $\widetilde{f}$ is well-defined.

To prove that $\widetilde{f}$ is continuous, it suffices to show that the restriction of $\widetilde{f}$ to a suitable open neighborhood of $y \in Y$ is continuous. A convenient choice in this context is to choose $f^{-1}(U)$ where U is an evenly covered neighborhood of $f(y) \in X$. Using the assumption that X is locally path-connected we can pass to a smaller open neighborhood $V \subset f^{-1}(U)$ of y, which is path-connected, making it possible to connect y_0 to any point $y' \in V$ via a path of the form $\gamma * \delta$, where γ is a path from y_0 to y and δ is a path in V from y to y'. The fact that the image path $f \circ \delta$ is contained in the evenly covered subset $U \subset X$ makes it easy to control the lifts of δ, and hence to evaluate $\widetilde{f}$ at any point $y \in V$. The details are left as a homework problem.

□

Lemma 4.13. *Let $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a covering map. Let $\widetilde{x}_1 \in \widetilde{X}$, let $\widetilde{\beta} \colon I \to \widetilde{X}$ be* *a path from* $\widetilde{x}_0$ *to* $\widetilde{x}_1$*, and let* $g := [p \circ \beta] \in \pi_1(X, x_0)$ *be the element of the fundamental group represented by the based loop* $p \circ \beta$ *in* (X, x_0)*. Let* $H \subset G := \pi_1(X, x_0)$ *be the image of* $p_* \colon \pi_1(\widetilde{X}, \widetilde{x}_0) \to \pi_1(X, x_0)$*. Then the image of* $p_* \colon \pi_1(\widetilde{X}, \widetilde{x}_1) \to \pi_1(X, x_0)$ *is the subgroup* $g^{-1}Hg \subset G$*.*

Proof. Earlier in the semester we have proved that the homomorphism

$$\Phi_\beta \colon \pi_1(\widetilde{X}, \widetilde{x}_0) \longrightarrow \pi_1(\widetilde{X}, \widetilde{x}_1) \qquad \text{given by} \qquad [\gamma] \mapsto [\bar{\beta} * \gamma * \beta]$$

is an isomorphism. Hence it suffices to calculate the image of the composition

$$\begin{array}{ccccc} \pi_1(\widetilde{X}, \widetilde{x}_0) & \xrightarrow{\Phi_\beta} & \pi_1(\widetilde{X}, \widetilde{x}_1) & \xrightarrow{p_*} & \pi_1(X, x_0) \\ & & & & \\ [\gamma] & \longmapsto & [\bar{\beta} * \gamma * \beta] & \longmapsto & [p \circ (\bar{\beta} * \gamma * \beta)] \end{array}$$

$$[p \circ (\bar{\beta} * \gamma * \beta)] = [p \circ \bar{\beta}][p \circ \gamma][p \circ \beta] = b^{-1} p_*([\gamma]) b.$$

$\square$

Corollary 4.14. *Let* $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ *be a covering map with* $\widetilde{X}$ *path-connected and locally path-connected. Let* $\Psi\colon$ *Then there is a deck transformation* $f \colon \widetilde{X} \to \widetilde{X}$ *with* $\widetilde{f}$

4.2 Classification of coverings

Definition 4.15. A covering $p \colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ is called a *universal covering* if $\widetilde{X}$ is simply connected. In that case, $\widetilde{X}$ is called a *universal covering space* of X.

If X is locally path-connected, a universal covering indeed enjoys a universal property which can be expressed by the following commutative diagram

$$\begin{tikzcd} (\widetilde{X}, \widetilde{x}_0) \arrow[rr, dashed, "\exists!\ \widetilde{f}"] \arrow[dr, "p"'] & & (\widetilde{X}', \widetilde{x}_0') \arrow[dl, "p'"] \\ & (X, x_0) & \end{tikzcd}$$

for any covering $p' \colon (\widetilde{X}', \widetilde{x}_0') \to (X, x_0)$. This universal property is a consequence of the lifting criterion **??**, which requires that the domain, in this case $\widetilde{X}$, is path-connected and locally path-connected. This holds, due to our assumptions that $\widetilde{X}$ is simply connected (which in particular requires $\widetilde{X}$ to be path-connected) and that X is locally path-connected (which implies that $\widetilde{X}$ is locally path connected).

In particular, for locally path-connected X its universal coverings are unique up to isomorphism, resulting in the custom to talk about *the* universal covering of X.

Example 4.16. (Examples of universal coverings).

1. $p\colon \mathbb{R} \to S^1$, $t \mapsto e^{2\pi i t}$ is the universal covering of S^1.
2. The projection map $p\colon S^n \to \mathbb{RP}^n = S^n/\sim$ is the covering of the projective space $\mathbb{RP}^n$ for $n \geq 2$, since the sphere S^n is simply connected for $n \geq 2$.

Not every space X has a universal covering. The following condition turns out to be a necessary and sufficient condition for the existence of a universal covering of X.

Definition 4.17. A space X is *semilocally simply-connected* if every point $x \in X$ has an open neighborhood U such that the induced homomorphism $\pi_1(U, x) \to \pi_1(X, x)$ is the trivial map.

To see that this is in fact a necessary condition for the existence of a universal covering $\widetilde{X} \to X$ of a path-connected space X, assume that $p\colon \widetilde{X} \to X$ is a universal covering. For $x \in X$, let U be an evenly covered open neighborhood of x, and let $\widetilde{U} \xrightarrow{\approx} U$ be the homeomorphism given by restricting p, and let $\widetilde{x} \in \widetilde{U}$ be the unique point with $p(\widetilde{x}) = x$. Let $i^U : U \to X$ and $i^{\widetilde{U}} : \widetilde{U} \to \widetilde{X}$ be the inclusion maps. Then the commutative diagram

$$\begin{tikzcd}
\pi_1(\widetilde{U}, \widetilde{x}) \arrow[r, "i_*^{\widetilde{U}}"] \arrow[d, "p_*", "\cong"'] & \pi_1(\widetilde{X}, \widetilde{x}) = \{1\} \arrow[d, "p_*"] \\
\pi_1(U, x) \arrow[r, "i_*^U"] & \pi_1(X, x_0)
\end{tikzcd}$$

shows that the homomorphism i_*^U is trivial.

Example 4.18. An example of a space which is not semilocally simply connected is the *Hawaiian Earring*, the subspace of $\mathbb{R}^2$ given by the union of the circles of radius $1/n$ with center $(0, -1/n)$ for $n = 1, 2, \ldots$, shown in the picture below (for typographical reasons the figure does not show *all* circles).

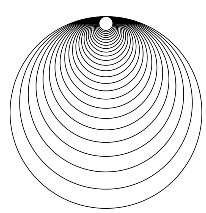


Figure 2: The Hawaiian Earring

We note that the condition *locally simply-connected* is stronger than *semilocally simply-connected*. For example, a locally simply connected space is in particular locally path-connected; the Warsaw circle is not locally path-connected, but it is semilocally simply-connected since its fundamental group is trivial. Another example is provided by the cone of the Hawaiian Earring, which has vanishing fundamental group, but not every point x has an open neighborhood with vanishing fundamental group contained in a prescribed neighborhood.

Theorem 4.19. *A path-connected space X has a universal covering if and only if X is semilocally simply-connected.*

For the proof we refer to Hatcher's book (p. 63-65) or Munkres' book (Theorem 82.1, p. 495).

Definition 4.20. Let G be a topological group.

(i) A *G-action on a space* X is a continuous map $\mu\colon G \times X \to X$ which is *associative* in the sense that $g(hx) = (gh)x$ for $g, h \in G$ and $x \in X$ (here we write $gh \in G$ for the product of g and h, and $hx \in X$ for $\mu(h, x)$). If G is a discrete group (i.e., the topology on G is the discrete topology), then the continuity of the action map $\mu\colon G \times X \to X$ is equivalent to the continuity of the map $\mu(g, \ )\colon X \to X$, $x \mapsto \mu(g, x) = gx$ for all $g \in G$.

(ii) The action is *free* if for all $x \in X$ the only $g \in G$ with $gx = x$ is the identity element.

(iii) The action is *transitive* if for all $x, y \in X$ there is some $g \in G$ with $gx = y$.

(iv) For $x \in X$ the subset $Gx := \{gx \mid g \in G\}$ is the *orbit through* x, and $G\backslash X = \{\text{orbits}\}$ the *orbit space* of the G-action of X. The topology on $G\backslash X$ is the quotient topology determined by the surjective map $p\colon X \to G\backslash X$, given by $x \mapsto Gx$.

Example 4.21. (Examples of group actions).

(1) The group $\mathbb{Z}$ acts on $\mathbb{R}$ via the action map $(n, t) \mapsto n + t$. The orbit space $\mathbb{Z}\backslash\mathbb{R}$ is homeomorphic to S^1 via the map $[t] \mapsto e^{2\pi it}$. Via this homeomorphism the projection map $\mathbb{R} \to \mathbb{Z}\backslash\mathbb{R}$ corresponds to our standard covering map $p\colon \mathbb{R} \to S^1$.

(2) The group $\mathbb{Z}/2 \cong \{\pm 1\}$ acts on S^n via the action map $\{\pm 1\} \times S^n \to S^n$, $(\pm 1, x) \mapsto \pm 1$. The orbit space $\{\pm 1\}\backslash S^n$ is equal to the projective space $\mathbb{RP}^n$.

We note that in these examples the projection maps $p\colon X \to G\backslash X$ are the universal coverings of the quotient spaces S^1 resp. $\mathbb{RP}^n$ ($n \geq 2$) mentioned in Example **??**. The group acting is the fundamental group of the quotient space.

Much more generally, if $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ is the universal covering of a path-connected space X, then for any point $\widetilde{x} \in \widetilde{X}$ there is a path $\widetilde{\alpha}$ in $\widetilde{X}$ from $\widetilde{x}_0$ to $\widetilde{x}$, which is unique up to homotopy since $\widetilde{X}$ is simply-connected.

Lemma 4.22. *Let $p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be the universal covering of a path-connected and locally path-connected space X. Then the fundamental group $\pi_1(X, x_0)$ acts freely on $\widetilde{X}$ such that the orbits are precisely the fibers of p.*

The proof is a homework problem.

Definition 4.23. Let X be a topological space. Let $\mathrm{Cov}(X)$ be the *category of path-connected coverings of the space* X defined as follows:

- The objects of $\mathrm{Cov}(X)$ are coverings $p\colon E \to X$ of X with path-connected total space E.
- The morphisms of $\mathrm{Cov}(X)$ from a covering $p\colon E \to X$ to a covering $p'\colon E' \to X$ are maps $\phi\colon E \to E'$ which make the diagram

$$\begin{tikzcd} E \arrow[rr, "\phi"] \arrow[dr, "p"'] & & E' \arrow[dl, "p'"] \\ & X & \end{tikzcd}$$

  commutative.
- the composition of morphisms in $\mathrm{Cov}(X)$ is given by composing the maps $\phi\colon E \to E'$ and $\phi'\colon E' \to E''$; the identity morphism of $E \to X$ is the identity map of E.

There is a variant of this category, namely the category $\mathrm{Cov}_*(X, x_0)$ of *based path-connected coverings* of the pointed space (X, x_0), where the objects are based coverings $p\colon (E, e_0) \to (X, x_0)$ with E path-connected, and the morphisms are maps $\phi\colon (E, e_0) \to (E', e_0')$ that are compatible with the projection maps to X.

While the definition of the category of coverings of a topological space X does not require assumptions on X, we need assumptions in order to use the tools at our disposal to analyze that category:

- the Lifting Criterion 4.11 involves assumes that the domain of the lift is path-connected and locally path-connected;
- the existence of the universal covering of a space requires it to be semilocally simply-connected.

So typically we will make the following requirement.

Definition 4.24. We will call a space X *reasonable* if X is path-connected, locally path-connected and semilocally simply connected.

Theorem 4.25. (Classification of path-connected based coverings). *Let (X, x_0) be a reasonable path-connected space. Then there is a bijection between*

$$\{\textit{based coverings } p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0 \textit{ with } \widetilde{X} \textit{ path-connected}\}/\textit{isomorphism}$$

and

$$\{\textit{subgroups of } \pi_1(X, x_0)\}.$$

It is given by sending a covering p to the subgroup $p_*\pi_1(\widetilde{X}, \widetilde{x}_0) \subset \pi_1(X, x_0)$.

4.3 G-coverings

We recall that the fundamental group $G := \pi_1(X, x_0)$ of a reasonable space X (in the sense of our convention 4.24) acts on the universal covering $p\colon \widetilde{X} \to X$ of a space X. This action has the following properties

(i) The action is compatible with the projection map p in the sense that $p(g\widetilde{x}) = p(\widetilde{x})$. In particular, the action restricts to an action on each fiber $p^{-1}(x) \subset \widetilde{X}$ for $x \in X$.

(ii) The action on each fiber is transitive, i.e., for any $\widetilde{x}, \widetilde{x}' \in p^{-1}(x)$, there is some $g \in G$ such that $g\widetilde{x} = \widetilde{x}'$.

(iii) The action is free, i.e., if if $g\widetilde{x} = \widetilde{x}$ for some $\widetilde{x} \in \widetilde{X}$, then g is the identity element of G.

Definition 4.26. Let G be a group, and let X be a space. A *G-covering of X* is a covering $p\colon E \to X$ together with a G-action on E with the properties (i), (ii) and (iii) above.

Here is an example of a G-covering of X that is not the universal cover of X. Let G be the cyclic group of three elements and let $p\colon S^1 \to S^1$ be the covering given by $z \mapsto z^3$. To describe the action, it will be convenient to think of G not as $\mathbb{Z}/3$, but as the group of third roots of unity, i.e., as $G = \{\zeta \in S^1 \subset \mathbb{C} \mid \zeta^3 = 1\}$. Let G act on S^1 via the map

$$G \times S^1 \longrightarrow S^1 \qquad (\zeta, z) \mapsto \zeta z.$$

This is an action map since for $\zeta, \zeta' \in G$ and $z \in S^1$ we have $\zeta(\zeta' z) = (\zeta\zeta')z$. It has the required properties:

(i) $p(\zeta z) = (\zeta z)^3 = \zeta^3 z^3 = z^3 = p(z)$.

(ii) G acts transitively on the fibers $p^{-1}(x)$ for any $x \in S^1$, since if $z, z' \in p^{-1}(x)$, then $z^3 = p(z) = p(z') = (z')^3$ and hence $1 = z^{-3}(z')^3 = (z^{-1}z')^3$. In other words, $\zeta := z^{-1}z'$ is a third root of unity and hence $\zeta z = z'$ shows that z' is in the G-orbit of z.

(iii) If $\zeta \in G$ fixes some point $z \in S^1$, then $\zeta z = z$, and hence $\zeta = zz^{-1} = 1$. In other words, G acts freely.

A slight variation of this example is given by the covering $p\colon \mathbb{C}^\times \to \mathbb{C}^\times$, $z \mapsto z^3$ (here $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$). As above, let G be the group of third roots of unity and let G act on the total space $\mathbb{C}^\times$ by $(\zeta, z) \mapsto \zeta z$. The same arguments as above show that this is a G-covering. In fact, restricting this G-covering of $\mathbb{C}^\times$ to the subspace $S^1 \subset \mathbb{C}^\times$, we obtain the first G-covering.

Theorem 4.27. *Let (X, x_0) be a reasonable path-connected space. Then there is a bijection between*

$$\{\textit{based } G\textit{-coverings } p\colon (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)\}/\textit{isomorphism}$$

and

$$\{\textit{group homomorphisms } \phi\colon \pi_1(X, x_0) \to G\}$$

It is given by sending a G-covering p to the group homomorphism $\phi\colon \pi_1(X, x_0) \to G$ which sends the element of $\pi_1(X, x_0)$ represented by a based loop γ in (X, x_0) to the unique $g \in G$ such that $\widetilde{\gamma}(1) = g\widetilde{x}_0$, where $\widetilde{\gamma}\colon I \to \widetilde{X}$ is the unique lift of γ with $\widetilde{\gamma}(0) = \widetilde{x}_0$.

Want addendum saying that this bijection is compatible with restrictions

Proof. Proof that ϕ is a homomorphism. Let γ, δ be based loops in (X, x_0) and let $\widetilde{\gamma}, \widetilde{\delta}\colon I \to \widetilde{X}$ be their lifts starting at $\widetilde{x}_0$. Let $g = \phi([\gamma]) \in G$, $h = \phi([\delta]) \in G$ be the unique elements such that $\widetilde{\gamma}(1) = g\widetilde{x}_0$ and $\widetilde{\delta}(1) = h\widetilde{x}_0$. To determine $\phi([\gamma][\delta]) = \phi([\gamma * \delta])$, we need to determine the unique lift of the concatenated path $\gamma * \delta$. Let $g\widetilde{\delta}\colon I \to \widetilde{X}$ be the path obtained by letting $g \in G$ act on the path $\widetilde{\delta}$, explicitly, $(g\widetilde{\delta})(s) := g(\widetilde{\delta}(s))$ for $s \in I$. Then $g\widetilde{\delta}$ has starting point $g\widetilde{\delta}(0) = g\widetilde{x}_0$, which is the endpoint of $\widetilde{\gamma}$, allowing us to form the concatenation $\widetilde{\gamma} * (g\widetilde{\delta})$. This is a lift of the loop $\gamma * \delta$ with endpoint of $(\widetilde{\gamma} * (g\widetilde{\delta})) = g\widetilde{\delta}(1) = g(h\widetilde{x}_0)$. This proves $\phi([\gamma][\delta]) = gh$ and hence ϕ is a homomorphism.

Proof incomplete □

4.3.1 Proof of the Seifert van Kampen Theorem for reasonable spaces

Proof. We need to show that the commutative diagram of groups

$$\begin{tikzcd} \pi_1(U \cap V, x_0) \arrow[r, "j_*^U"] \arrow[d, "j_*^V"'] & \pi_1(U, x_0) \arrow[d, "i_*^U"] \\ \pi_1(V, x_0) \arrow[r, "i_*^V"'] & \pi_1(X, x_0) \end{tikzcd} \tag{4.28}$$

is a pushout diagram. Let $U' \subset U$ be the *path component* of U containing the basepoint x_0, which consists of all points $x \in U$ for which there is a path starting in x_0 and ending in x. This subspace of U is a path-connected space and the inclusion map $i\colon U' \to U$ induces an isomorphism $i_*\colon \pi_1(U', x_0) \xrightarrow{\cong} \pi_1(U, x_0)$, since every based loop in (U, x_0) and homotopy between based loops is necessarily contained in U'.

To verify that diagram (4.28) is a pushout square, we need to check whether it has the universal property expressed by the following diagram of groups and group homomorphisms

$$\begin{tikzcd}
\pi_1(U\cap V, x_0) \arrow[r, "j^U_*"] \arrow[d, "j^V_*"'] & \pi_1(U, x_0) \arrow[d, "i^U_*"] \arrow[ddr, bend left, "f_1"] & \\
\pi_1(V, x_0) \arrow[r, "i^V_*"] \arrow[drr, bend right, "f_2"'] & \pi_1(X, x_0) \arrow[dr, dashed, "\exists! f"] & \\
 & & G
\end{tikzcd} \tag{4.29}$$

This diagram contains many group homomorphisms from fundamental groups of U, V, $U\cap V$ and X to the group G. Via Theorem 4.27 these homomorphisms can be interpreted geometrically as G-coverings over these topological spaces (up to isomorphism). In particular, the homomorphism $f_2 \in \mathrm{Hom}(\pi_1(U, x_0), G)$ corresponds to some G-covering

$$p_1\colon (\widetilde{U}, \widetilde{x}_1) \longrightarrow (U, x_0).$$

This G-covering restricts to a G-covering $(\widetilde{U}_{|U\cap V}, \widetilde{x}_1) \longrightarrow (U\cap V, x_0)$, which by the Addendum corresponds to the composition

$$\pi_1(U\cap V, x_0) \xrightarrow{j^U_*} \pi_1(U, x_0) \xrightarrow{f_1} G.$$

Similarly, $f_2\colon \pi_1(V, x_0) \to G$ corresponds to a G-covering $p_2\colon (\widetilde{V}, \widetilde{x}_2) \to (V, x_0)$. Its restriction to $U\cap V$ is a G-covering $(\widetilde{V}_{|U\cap V}, \widetilde{x}_2) \longrightarrow (U\cap V, x_0)$ which corresponds to the composition $f_2 \circ j^V_*\colon \pi_1(U\cap V, x_0) \to G$. By the commutativity of the outer square of the diagram (**??**), $f_1 \circ j^U_* = f_2 \circ j^V_*$. By the Classification Theorem for G-coverings, this implies that there is an isomorphism between the corresponding based covering spaces over $U\cap V$, i.e., there is a G-equivariant map basepoint preserving map ϕ making the following diagram commutative: □

Proof incomplete

5 Smooth manifolds

5.1 Smooth manifolds

Motivation: the possible states of a classical physical system corresponds to the points of a manifold called the *phase space* of the system. For example, the phase space of a particle moving in $\mathbb{R}^3$ is $\mathbb{R}^6$, with the first three coordinates describing the *position* of the particle at a fixed time, the last giving its *velocity*. According to the laws of classical physics, this information suffices to determine the position and velocity of the particle for any time by solving a differential equation.

If the particle is restricted to move in $S^2 \subset \mathbb{R}^3$, its velocity vector is restricted to be tangent to S^2. Hence the pair $(x(t), v(t))$ in this case is restricted to be a point in the 4-dimensional manifold consisting of the pairs $(x, v) \in \mathbb{R}^3 \times \mathbb{R}^3$ such that $||x|| = 1$ and v is perpendicular to x. The physical state $(x(t), v(t))$ can be determined from $(x(0), v(0))$ by solving a *differential equation*. This example makes it clear that we want to do

Calculus on manifolds

Let U be an open subset of $\mathbb{R}^n$, and let $f\colon U \to \mathbb{R}$ be a real valued function. The two main constructions of Calculus are

- **Differentiation**, which associates to f its (total) derivative
$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n \tag{5.1}$$
- **Integration**, which associates to f the integral
$$\int_U f(x) dx_1 \dots dx_n. \tag{5.2}$$

Needless to say, some conditions on the function f are needed to ensure that df exists (f should be differentiable) or that the integral over U exists (f should be integrable). For our geometric purposes here, we will assume that the functions we consider are *smooth*, i.e., can be differentiated as often as desired, which will ensure that all derivatives or integrals we will consider below exist. This is a condition much stronger than needed, but it will be pretty clear how the theory of smooth manifolds can be modified to get away with less amount of differentiability.

Definition 5.3. Let U be an open subset of $\mathbb{R}^n$. A function $f\colon U \to \mathbb{R}$ is *smooth* if for all n-tupels $(k_1, \dots, k_n)$, $k_i \in \mathbb{N}$, the corresponding partial derivative
$$\frac{\partial^k f}{\partial x^{k_1} \dots \partial x_n^{k_n}}(x)$$
exist for all points $x \in U$; here $k = \sum_{i=1}^n k_i$. A map $f = (f_1, \dots, f_m)\colon U \to V \subset \mathbb{R}^m$ is smooth if all its component functions f_i are smooth. This map f is a *diffeomorphism* if f is a bijection and its inverse map $f^{-1}\colon V \to U$ is smooth as well (it turns out that this can only happen if $m = n$).

Goals of this section.

(i) Define what a "smooth function" $f\colon M \to R$ on a manifold M is.

(ii) Define the total derivative *df* of a smooth function. What kind of object is a *df*?

(iii) Define suitable objects on an n-manifold M that can be integrated over M (if M is an open subset of $\mathbb{R}^n$, these are just of the form $f(x)dx_1 \dots dx_n$).

Before addressing (i) in the section, we remark that *df* is a *smooth section of the cotangent bundle of M*, while the object sought in (iii) will turn out to be a *smooth section of a vectorbundle* built from the cotangent bundle. This explains our interest in vector bundles and their sections which will be discussed in the following sections.

Definition 5.4. Let M be a topological manifold of dimension n. A *chart* for M is an open subset $U \subset M$ and a homeomorphism

$$M \underset{\text{open}}{\supset} U \xrightarrow{\phi} \phi(U) \underset{\text{open}}{\subset} \mathbb{R}^n.$$

A collection of charts $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in A}$ is an *atlas* for M if $\{U_\alpha\}_{\alpha\in A}$ is an open cover of M, i.e., if $M = \bigcup_{\alpha\in A} U_\alpha$.

Example 5.5. (Examples of charts and atlases).

1. Let $M = S^n$ be the sphere of dimension n. For $i = 0, \dots, n$ let $U_i^+ := \{(x_0, \dots, x_n) \in S^n \mid x_i > 0\}$ and $U_i^- := \{(x_0, \dots, x_n) \in S^n \mid x_i < 0\}$. Let $\phi_i^\pm\colon U_i^\pm \to B_1^n$ be the homeomorphism given by

   $$\phi_i^\pm(x_0, \dots, x_n) = (x_1, \dots, x_{i-1}, \widehat{x}_i, x_{i+1}, \dots, x_n).$$

   Then (U_i^+, ϕ_i^+) and (U_i^-, ϕ_i^-) are $2(n+1)$ charts for the manifold S^n. Since every point of S^n belongs to some hemisphere $U_i^\pm$, this collection of charts form an atlas for S^n.

2. A smaller atlas of S^n consisting of just two charts is obtained by using the homeomorphisms

   $$\phi^\pm\colon U^\pm := S^n \setminus \{(0, \dots, \pm 1)\} \xrightarrow{\approx} \mathbb{R}^n$$

   given by stereographic projections.

3. Let M be the projective space $\mathbb{RP}^n$, considered as the quotient space $\mathbb{R}^n \setminus \{0\}/x \sim \lambda x$ for $x \in \mathbb{R}^n \setminus \{0\}$ and $\lambda \in \mathbb{R} \setminus \{0\}$. Let $U_i \subset \mathbb{RP}^n$ be the open subset given by $U_i := \{([x_0, \dots, x_n] \in \mathbb{RP}^n \mid x_i \neq 0\}$. Then the map

   $$U_i \xrightarrow{\phi_i} \mathbb{R}^n \qquad [x_0, \dots, x_n] \mapsto (\frac{x_0}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{\widehat{x_i}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_n}{x_i})$$

   is a homeomorphism with inverse given by $\phi_i^{-1}(v_1, \dots, v_n) = [v_1, \dots, v_{i-1}, 1, v_i, \dots, v_n]$. In particular, the collection of charts $\{(U_i, \phi_i)\}_{i=0,\dots,n}$ is an atlas for $\mathbb{RP}^n$.

Smoothness of a function $\mathbb{R}^n \supset U \to \mathbb{R}$ is a *local* property in the sense that if U is the union of open subsets $U_\alpha \subset U$ with $\bigcup_{\alpha \in A} U_\alpha = U$, then f is smooth if and only if the restriction f_{U_α} is smooth for all $\alpha \in A$. Let M be an n-manifold with an atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A})$, and let $f\colon M \to \mathbb{R}$ be a map. The observation above suggests to define that $f_{|U_\alpha}\colon U_\alpha \to \mathbb{R}$ is *smooth* if the composition

$$\mathbb{R}^n \supset \phi_\alpha(U_\alpha) \xrightarrow{\phi_i^{-1}} U_\alpha \xrightarrow{f_{|U_\alpha}} \mathbb{R}$$

is smooth (unlike $U_\alpha \subset M$, the image $\phi_\alpha(U_\alpha)$ is an open subset of $\mathbb{R}^n$, and hence we already know by Definition 5.3 what a smooth function with domain $\phi_\alpha(U_\alpha)$ is). Then we define f to be *smooth* if its restriction $f_{|U_\alpha}\colon U_\alpha \to \mathbb{R}$ to each $U_\alpha \subset M$ is smooth.

The problem with the proposed definition. Suppose there are two charts (U, ϕ), (V, ψ) belonging to the atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in A}$ with $U \cap V \neq \emptyset$. Then according to the above definition, there would be *two* ways to determine whether the restriction $f_{|U \cap V}$ is smooth: we could use the chart (U, ϕ) and check whether the composition

$$\mathbb{R}^n \supset \phi(U \cap V) \xrightarrow[\approx]{\phi^{-1}} U \cap V \xrightarrow{f} \mathbb{R}$$

is smooth. Alternatively, we could use the chart (V, ψ) and check for smoothness of the composition

$$\mathbb{R}^n \supset \psi(U \cap V) \xrightarrow[\approx]{\psi^{-1}} U \cap V \xrightarrow{f} \mathbb{R}$$

The problem is that these two ways to test for smoothness of the function f on $U \cap V$ might **not** yield the same answer. We note that the second map $f \circ \psi^{-1}$ can be expressed as the composition

$$\begin{tikzcd}
\psi(U \cap V) \arrow[r, "\phi\circ\psi^{-1}", "\approx"'] \arrow[rr, bend right, "f\circ\psi^{-1}"'] & \phi(U \cap V) \arrow[r, "f\circ\phi^{-1}"] & \mathbb{R}
\end{tikzcd}$$

We also note that $\phi \circ \psi^{-1}$ is a map between open subsets of $\mathbb{R}^n$, and hence we can check by Definition 5.3 whether it is smooth. This map is a homeomorphism, but in general, there is no reason that this map should be smooth. It follows that if $f \circ \phi^{-1}$ is smooth, the map $f \circ \psi^{-1}$ in general won't be smooth. In other words, the smoothness test for f restricted to $U \cap V$ using the chart (U, ϕ) in general won't give the same answer as the smoothness test using the chart (V, ψ).

Let M be a topological manifold. How can we define whether a map $f\colon M \to \mathbb{R}$ is smooth?

First try. We call a f smooth at a point x if

Still missing: transition map, smooth compatible, smooth atlas, smooth structure, smooth manifold, smooth function on a manifold, smooth maps between manifolds

5.2 Tangent space

We begin by reviewing the differential of a smooth map between open subsets of Euclidean spaces. Our goal is to extend this definition to smooth maps between manifolds.

Definition 5.6. Let $\mathbb{R}^m \supset U \xrightarrow{F} V \subset \mathbb{R}^n$ be a smooth map. For a point $p \in U$, the *differential of F at p* is the linear map

$$dF(p)\colon \mathbb{R}^m \longrightarrow \mathbb{R}^n \tag{5.7}$$

which corresponds to the Jacobi matrix

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \dots & \frac{\partial F_1}{\partial x_m}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(p) & \dots & \frac{\partial F_n}{\partial x_m}(p) \end{pmatrix}$$

via the usual mechanism, i.e., the i-th column vector of the Jacobi matrix is equal $dF(p)$ applied to $e_i \in \mathbb{R}^m$, where $\{e_i\}_{i=1,\dots,m}$ is the standard basis of $\mathbb{R}^m$.

Theorem 5.8. (Chain Rule). *Let $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^p$ be open subsets, and let*

$$U \xrightarrow{F} V \xrightarrow{G} V$$

be smooth maps. Then for $p \in U$ the differential $d(G \circ F)(p)$ is the composition

$$\mathbb{R}^m \xrightarrow{dF(p)} \mathbb{R}^n \xrightarrow{dG(F(p))} \mathbb{R}^p.$$

Our goal is to generalize the construction of the differential to smooth maps $F\colon M \to N$ between manifolds; i.e., for $p \in M$ we want to construct the *differential* $dF(p)$ which should be a linear map. One might suspect that the domain of $dF(p)$ is $\mathbb{R}^m$, where $m = \dim M$, and its codomain is $\mathbb{R}^n$, $n = \dim N$. It turns out to be more involved, namely $dF(p)$ is a linear map

$$dF(p)\colon T_pM \longrightarrow T_{F(p)}N, \tag{5.9}$$

where T_pM is an m-dimensional vector space associated to M and $p \in M$, called the *tangent space of M at the point p*.

So our next goal is to define the tangent space T_pM; in fact, we will provide *two* definitions, the "geometric" definition, denoted $T_p^{\text{geo}}M$ and the "algebraic" definition, denoted $T_p^{\text{alg}}M$. The reason for dealing with both, rather than settling on one of these is that both have their pros and cons, and hence it's good to know both of them.

5.2.1 The geometric tangent space

The strategy for for how to construct either $T_p^{\mathrm{geo}}M$ or $T_p^{\mathrm{alg}}M$ is basically the same: if M is an open subset of $\mathbb{R}^n$, the to be constructed tangent space $T_p^{\mathrm{geo}}M$ resp. $T_p^{\mathrm{alg}}M$ should be isomorphic to $\mathbb{R}^n$. So in both cases, we first observe that there is a very peculiar, roundabout way to think about a vector $v \in \mathbb{R}^n$. The only redeeming quality of this is that this very peculiar, roundabout way still makes sense when we are no longer looking at a point p of some open subset of $\mathbb{R}^n$, but for p a point in a smooth manifold M.

Observation. Let $M \subset \mathbb{R}^n$ be an open subset, and let p be a point of M. We define an equivalence relation $\sim$ on smooth paths $\gamma\colon (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ as follows

$$((-\epsilon_1, \epsilon_1) \xrightarrow{\gamma_1} M) \ \sim \ ((-\epsilon_2, \epsilon_2) \xrightarrow{\gamma_2} M) \qquad \text{if and only if} \qquad \gamma_1'(0) = \gamma_2'(0)$$

It is evident that the map

$$\{\gamma\colon (-\epsilon, \epsilon) \to M \mid \gamma \text{ is smooth}\}/\sim \quad \to \quad \mathbb{R}^n \qquad \text{given by} \qquad [\gamma] \mapsto \gamma'(0)$$

is a bijection. In other words, these equivalence classes of smooth paths is just a very complicated way to think about vectors in $\mathbb{R}^n$. The only redeeming quality of doing this is that this construction works in the more general case where M is a smooth manifold rather than an open subset of $\mathbb{R}^n$ and motivates the following definition.

Definition 5.10. Let M be a smooth n manifold and $p \in M$. We define the *geometric tangent space of M at p* to be

$$T_p^{\mathrm{geo}}M := \{\gamma\colon (-\epsilon, \epsilon) \to M \mid \gamma \text{ is smooth and } \gamma(0) = p\}/\sim,$$

where two such paths $\gamma_1\colon (-\epsilon_1, \epsilon_1) \to M$ and $\gamma_2\colon (-\epsilon_2, \epsilon_2) \to M$ are declared equivalent if for some smooth chart (U, ϕ) with $p \in U$ the tangent vectors of the paths $\phi \circ \gamma_1$ and $\phi \circ \gamma_2$ in $\mathbb{R}^n$ have the same tangent vector at $t = 0$ (we might restrict the domain to γ_i to a smaller interval around 0 such that the compositions $\phi \circ \gamma_i$ are defined. We observe that the seemingly stronger requirement that $(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$ for *all* smooth charts (U, ϕ) with $p \in U$ is actually equivalent to the condition above.

We note that if the manifold M is an open subset of $\mathbb{R}^n$, we can use the smooth chart (M, i) provided by the inclusion map $i\colon M \to \mathbb{R}^n$. So in this case, the geometric tangent space $T_p^{\mathrm{geo}}M$ is equal to the quotient space discussed above and the map

$$T_p^{\mathrm{geo}}M \xrightarrow{\cong} \mathbb{R}^n \qquad [\gamma] \mapsto \gamma'(0) \tag{5.11}$$

is a bijection. So in this case, the geometric tangent space can be identified with $\mathbb{R}^n$ via this bijection, which we will often do without further comment.

Example 5.12. The idea of extracting the more concrete object $\gamma'(0)$ from the "abstract beast" $[\gamma] \in T_p^{\mathrm{geo}} M$ works in more general situations, for example for $M = S^n \subset \mathbb{R}^{n+1}$ (or more generally for submanifolds $M \subset \mathbb{R}^{n+k}$, a concept we will define a little later). It is easy to show that if γ is a smooth path in S^n, then its composition $i \circ \gamma$ with the inclusion map $i\colon S^n \to \mathbb{R}^{n+1}$ is a smooth path in $\mathbb{R}^{n+1}$, and hence it has a tangent vector $(i \circ \gamma)'(0) \in \mathbb{R}^{n+1}$. It is not hard to show that

$$T_p^{\mathrm{geo}} S^n \longrightarrow \mathbb{R}^{n+1} \qquad \text{given by} \qquad [\gamma] \mapsto (i \circ \gamma)'(0)$$

is a well-defined injective map. From the geometric picture it is clear that the tangent vector $(i \circ \gamma)'(0)$ of the path $i \circ \gamma$ should be perpendicular to the vector $p \in \mathbb{R}^{n+1}$. This can also be verified by the following calculation. Let $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ be the function $f(x_0, \dots, x_n) = x_0^2 + \dots + x_n^2$, which can be used to describe S^n as $S^n = f^{-1}(1) \subset \mathbb{R}^{n+1}$. Hence if γ is a smooth path in S^n, then $f \circ \gamma$ is constant and so its derivative vanishes. On the other hand, we can calculate the derivative $(f \circ \gamma)'(0)$ via the chain rule and obtain the following equation:

$$0 = (f \circ \gamma)'(0) = \langle (\mathrm{grad} f)(\gamma(0)), \gamma'(0) \rangle.$$

For $\gamma(0) = p = (x_0, \dots, x_n)$ we have $\mathrm{grad} f = (\frac{\partial f}{\partial x_0}, \dots, \frac{\partial f}{\partial x_n}) = (2x_0, \dots, 2x_n) = 2p$. It follows that $0 = \langle p, \gamma'(0) \rangle$, i.e., the tangent vector $\gamma'(0)$ is perpendicular to p. Again, it is not difficult to show that every vector $v \in \mathbb{R}^{n+1}$ perpendicular to p is the tangent vector of some path γ in S^n. Summarizing, we obtain a bijection

$$T_p^{\mathrm{geo}} S^n \xrightarrow{\cong} \{v \in \mathbb{R}^{n+1} \mid v \text{ is perpendicular to } p\} \qquad \text{given by} \qquad [\gamma] \mapsto \gamma'(0)$$

Definition 5.13. Let M, N be smooth manifolds, and let $F\colon M \to N$ be a smooth map. Then for $p \in M$ the *induced map of geometric tangent spaces* or the *differential of F at p* is the map

$$T_p^{\mathrm{geo}} M \xrightarrow{F_*^{\mathrm{geo}}} T_{F(p)}^{\mathrm{geo}} N \qquad \text{given by} \qquad [\gamma] \mapsto [F \circ \gamma]$$

We might also write F_* if it is clear that we are using the geometric version of the tangent space, or $dF(p)$.

Of course, we should justify the notation $dF(p)$ by showing that the map F_*^{geo} agrees with the differential as defined via the Jacobian matrix if M, N are open subsets of Euclidean space. This is the content of the next Lemma.

Lemma 5.14. *Let $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ be open subsetsm, let $F\colon U \to V$ be a smooth map, and let $p \in U$. Then the diagram*

$$\begin{tikzcd}
T_p^{\mathrm{geo}} U \arrow[r, "F_*^{\mathrm{geo}}"] \arrow[d, "\cong"] & T_{F(p)}^{\mathrm{geo}} V \arrow[d, "\cong"] \\
\mathbb{R}^m \arrow[r, "dF(p)"] & \mathbb{R}^n
\end{tikzcd}$$

Here the vertical maps are the bijections (5.11)*, and the bottom horizontal map is the usual differential of F at p (see* (5.7)*) corresponding to the Jacobian matrix.*

Proof. Let $\gamma\colon (-\epsilon,\epsilon) \to U$ be a smooth path. Then the left vertical map sends $[\gamma] \in T_p^{\mathrm{geo}}U$ to $\gamma'(0) \in \mathbb{R}^m$, which via $dF(p)$ maps to $(dF(p))(\gamma'(0)) \in \mathbb{R}^n$.

Going the other way, $F_*\mathrm{geo}([\gamma]) = [F \circ \gamma]$, which via the right vertical map is send to $(F \circ \gamma)'(0)$. Using the chain rule,

$$(F \circ \gamma)'(0) = (dF(\gamma(0))(\gamma'(0)) = (dF(p))(\gamma'(0)),$$

which proves that the diagram is commutative. □

Lemma 5.15. (Chain rule for the induced map of geometric tangent spaces). *Let M, N, P be smooth manifolds and let $F\colon M \to N$, $G\colon N \to P$ be smooth maps. Then for $p \in M$ the following diagram is commutative:*

$$\begin{tikzcd} T_p^{\mathrm{geo}}M \arrow[r, "F_*^{\mathrm{geo}}"'] \arrow[rr, bend left, "(G\circ F)_*^{\mathrm{geo}}"] & T_{F(p)}^{\mathrm{geo}}N \arrow[r, "G_*^{\mathrm{geo}}"'] & T_{G(F(p))}^{\mathrm{geo}}P \end{tikzcd}$$

Proof.

$$G_*^{\mathrm{geo}}(F_*^{\mathrm{geo}}([\gamma])) = G_*^{\mathrm{geo}}([F \circ \gamma]) = [G \circ (F \circ \gamma)] = [(G \circ F) \circ \gamma] = (G \circ F)_*^{\mathrm{geo}}([\gamma]).$$

□

Corollary 5.16. *If F is a diffeomorphisms with inverse G, then F_*^{geo} is a bijection with inverse G_*^{geo}.*

In particular, if M is a smooth manifold with a smooth chart $M \supset U \xrightarrow{\phi} V := \phi(U) \subset \mathbb{R}^n$, then ϕ is a diffeomorphism, and hence for $p \in U$ we have bijections

$$T_p^{\mathrm{geo}}M = T_p^{\mathrm{geo}}U \xrightarrow[\cong]{\phi_*^{\mathrm{geo}}} T_{\phi(p)}^{\mathrm{geo}}V \xrightarrow{\cong} \mathbb{R}^n \tag{5.17}$$

Here the last map is the bijection (5.11), and the equality $T_p^{\mathrm{geo}}M = T_p^{\mathrm{geo}}U$ follows from immediately from the definition of the geometric tangent space.

Ok, the construction of the geometric tangent space is pretty straightforward and it agrees with what we want if the manifold is an open subset of Euclidean space. Moreover, the map F_*^{geo} of geometric tangent spaces induced by a smooth map F is easy to define, agrees with the usual differential for open subsets of Euclidean spaces, and is compatible with compositions (i.e., we have a chain rule). So, what's not to like about the geometric tangent space?

The biggest drawback is that the geometric tangent space $T_p^{\text{geo}}M$ is not a vector space in an obvious way. We can use the bijection (5.17) to *define* a vector space structure on $T_p^{\text{geo}}M$ (which turns out to be independent of the choice of the chart (U,ϕ)), but this vector space structure does not have a direct geometric description. For example, if γ and δ are smooth paths in M, it is not clear how to construct a path that represents the sum $[\gamma]+[\delta] \in T_p^{\text{geo}}M$.

5.2.2 The algebraic tangent space

The definition of the algebraic tangent space $T_p^{\text{alg}}M$ at a point p of a smooth manifold M is more involved and less intuitive than that of the geometric tangent space $T_p^{\text{geo}}M$. Its big advantage is that unlike the geometric tangent space, the algebraic tangent space has an evident vector space structure. To motivate the definition, we first provide a really complicated way to think about vectors in $\mathbb{R}^n$ as "derivations".

5.3 Smooth submanifolds

Definition 5.18. Let N be a smooth manifold of dimension $m+k$. A subset $M \subset N$ is a *smooth submanifold* of N of dimension n if the inclusion $M \hookrightarrow N$ locally is isomorphic to the inclusion $\mathbb{R}^n \hookrightarrow \mathbb{R}^{m+k}$, i.e., if for every $p \in M$ there is a smooth chart

$$N \underset{\text{open}}{\supset} U \xrightarrow{\phi} \mathbb{R}^{m+k} = \mathbb{R}^m \times \mathbb{R}^k$$

with $p \in U$ such that

$$\phi(M \cap U) = \phi(U) \cap (\mathbb{R}^m \times \{0\}).$$

A chart (U,ϕ) with this property will be called a *submanifold chart* for $M \subset N$.

Remark 5.19. The restriction of ϕ to $M \cap U$ gives a chart $\phi_| \colon M \cap U \to \mathbb{R}^m$ for M. In particular, M is a topological manifold of dimension m. We claim that if (U,ϕ), (V,ψ) are two smooth charts of the special form described in the definition above, then the corresponding charts $\phi_|$, $\psi_|$ for M are smoothly compatible. To check this, we need to show that the transition map

$$\mathbb{R}^m \supset \phi(M \cap U \cap V) \xrightarrow{\phi_|^{-1}} M \cap U \cap V \xrightarrow{\psi_|} \psi(M \cap U \cap V) \subset \mathbb{R}^m$$

is smooth. Note that this map is given by restricting the transition map

$$\mathbb{R}^{m+k} \supset \phi(U \cap V) \xrightarrow{\phi^{-1}} U \cap V \xrightarrow{\psi} \psi(U \cap V) \subset \mathbb{R}^{m+k}$$

to $\mathbb{R}^m \subset \mathbb{R}^{m+k}$. Since the charts (U,ϕ), (V,ψ) are smoothly compatible, the transition function $\psi \circ \phi^{-1}$ is smooth, and hence also its restriction to $\mathbb{R}^m$. This shows that $(M \cap U, \phi_|)$

and $(M \cap V, \psi_|)$ are smoothly compatible. Hence the atlas for M obtained from these special charts for N is in fact smooth, and hence determines a smooth structure on the topological manifold M. Summarizing: a submanifold M is in fact a *smooth manifold*.

Definition 5.20. regular/critical points, regular values: first for maps $\mathbb{R}^{m+1} \supset U \to \mathbb{R}$, then $\mathbb{R}^{m+k} \to \mathbb{R}^k$, then $N^{m+k} \to Q^k$.

Theorem 5.21. *Let $F\colon N \to Q$ be a smooth map.*

(i) If q is a regular value, then $M = F^{-1}(q)$ is a manifold of dimension $\dim M = \dim N - \dim Q$.

(ii) For $p \in M$ the tangent space $T_pM \subset T_pN$ is equal to $\ker(F_*\colon T_pN \to T_qQ)$.

Examples:

1. Sphere
2. $SL_n(\mathbb{R})$
3. $V_k(\mathbb{R}^n)$, maybe do special case $O(n)$ first.

The proof of Theorem 5.21 is based on the Inverse Function Theorem.

Proof. Finding a submanifold chart for $M \subset N$ at the point $p \in M$ is a *local* question; i.e., it suffices to find a submanifold chart for $U \cap M \subset U \cap N$ where U is an open neighborhood of p. Let (U, ϕ) be a smooth chart for M with $p \in U$, and let (V, ψ) be a smooth chart for Q with $q \in V$. Replacing U by $U \cap F^{-1}(V)$ and restricting ϕ to that smaller neighborhood of $p \in N$, we can assume $F(U) \subset V$. We may also assume (by composing ϕ resp. ψ by translations in $\mathbb{R}^{m+k}$ resp. $\mathbb{R}^k$) that $\phi(p) = 0 \in \mathbb{R}^{m+k}$ and $\psi(q) = 0 \in \mathbb{R}^k$. Let

$$\mathbb{R}^{m+k} \supset \phi(U) \xrightarrow{G} \psi(V) \subset \mathbb{R}^k$$

be the smooth map that makes the following diagram commutative:

$$\begin{array}{ccccccc} N & \supset & U & \xrightarrow{F} & V & \subset & Q \\ & & {\scriptstyle\phi}\big\downarrow{\scriptstyle\cong} & & {\scriptstyle\cong}\big\downarrow{\scriptstyle\psi} & & \\ \mathbb{R}^{m+k} & \supset & \phi(U) & \xrightarrow{G} & \psi(V) & \subset & \mathbb{R}^k \end{array}$$

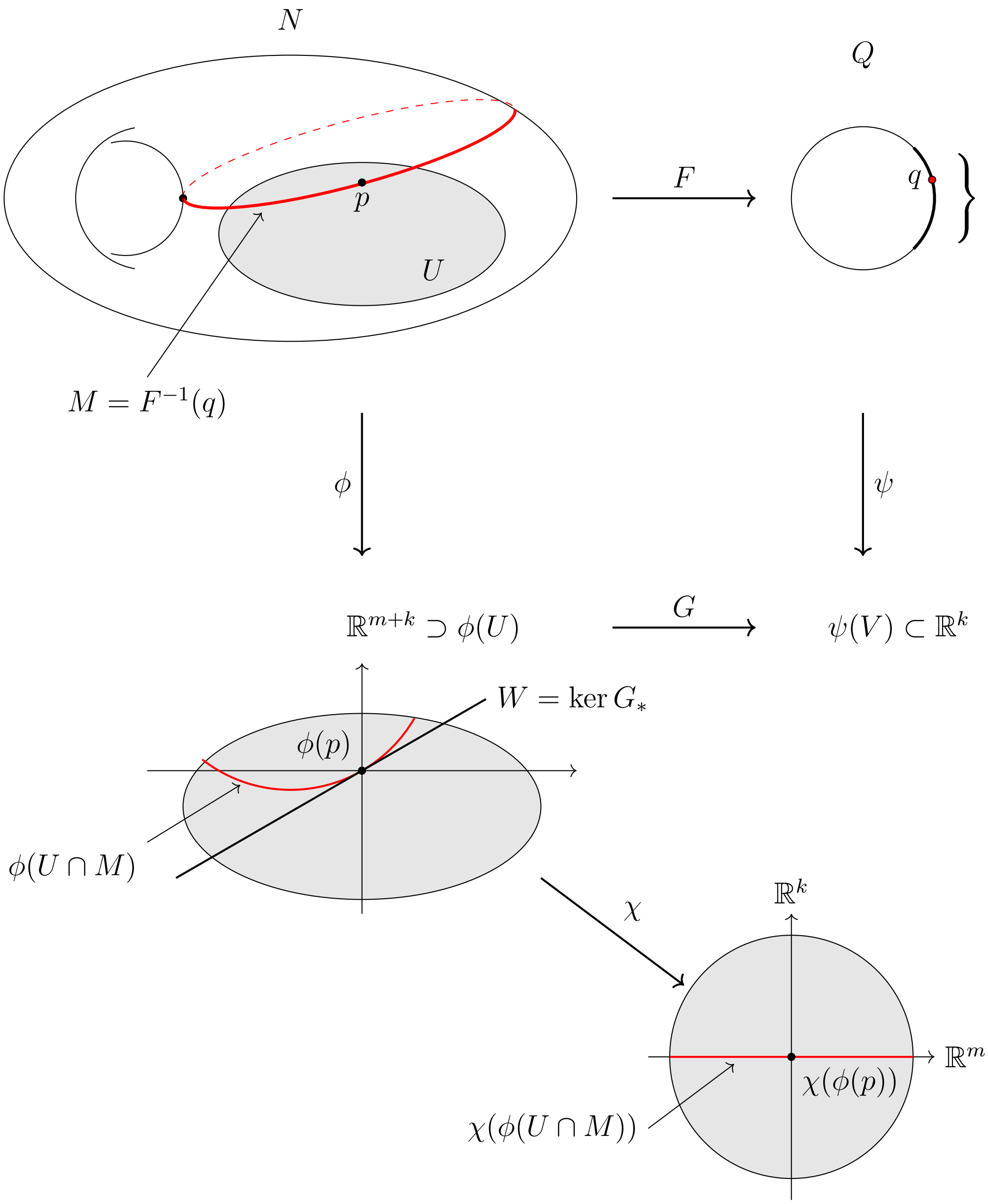
Here the symbol $\cong$ next to ϕ and ψ indicate that these maps are diffeomorphisms onto their image. Since p is a regular point of F, the commutative diagram implies that $\phi(p) = 0 \in \mathbb{R}^{m+k}$ is a regular point of G, i.e., the differential

$$G_*\colon T_0\phi(U) \longrightarrow T_0\psi(V)$$

is surjective. In particular, the vector space $W := \ker G_*$ has dimension

$$\dim T_0\phi(U) - \dim T_0\psi(V) = (m+k) - k = m.$$

The situation is depicted in the following figure.



In this picture the level set $M = F^{-1}(q)$ and $\phi(U \cap M)$ are drawn in red. To construct a submanifold chart of $M \subset N$ at $p \in M$ it suffices to find a submanifold chart for $\phi(U \cap M) \subset \phi(U)$ at $0 \in \phi(U) \subset \mathbb{R}^{m+k}$. In other words, we need to find a smooth map

$$\chi\colon \phi(U) \longrightarrow \mathbb{R}^m \times \mathbb{R}^k$$

satisfying the following properties:

(a) $\chi(\phi(U \cap M)) \subset \mathbb{R}^m = \mathbb{R}^m \times \{0\} \subset \mathbb{R}^m \times \mathbb{R}^k$, and

(b) the restriction of χ to a neighborhood of 0 is a diffeomorphism onto its image.

Let $\chi^1\colon \phi(U) \to \mathbb{R}^m$, $\chi^2\colon \phi(U) \to \mathbb{R}^k$ be the component maps of χ, i.e., $\chi(x) = (\chi^1(x), \chi^2(x)) \in \mathbb{R}^m \times \mathbb{R}^k$ for all $x \in \phi(U)$. Then the first condition can be satisfied by defining χ^2 to be the smooth function G, which by construction has the property $G(x) = 0$ if and only if $x \in \phi(U \cap M)$. For the construction of χ_1 we note that by the Inverse Function Theorem **??**, condition (ii) holds provided the differential of χ at 0 is an isomorphism. Let $\pi^1\colon \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m$ and $\pi^2\colon \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ be the projection maps and consider the following composition:

$$T_0\phi(U) \xrightarrow{\chi_*} T_{(0,0)}(\mathbb{R}^m \times \mathbb{R}^k) \xrightarrow[\cong]{\pi^1_* \times \pi^2_*} T_0\mathbb{R}^m \times T_0\mathbb{R}^k$$

By Lemma 1.4 the map $\pi^1_* \times \pi^2_*$ is an isomorphism of vector spaces, and hence

$$\ker \chi_* = \ker(\pi^1_* \circ \chi_*) \cap \ker(\pi^2_* \circ \chi_*) = \ker \chi^1_* \cap \ker \chi^2_* = \ker \chi^1_* \cap \ker G_*,$$

where all these differentials are taken at 0. In other words, it suffices to construct $\chi_1\colon \phi(U) \to \mathbb{R}^m$ in such a way that its differential (at $0 \in \phi(U)$ is injective on $W := \ker G_*$. Let χ_1 be the composition

$$\mathbb{R}^{m+k} \supset \phi(U) \xrightarrow{\pi} W \xrightarrow[\cong]{h} \mathbb{R}^m,$$

where π is the orthogonal projection from $\mathbb{R}^{m+k}$ onto its subspace W, and h is any choice of an isomorphism g between the vector spaces W and $\mathbb{R}^m$; such an isomorphism exists since W is the kernel of the surjective map $G_*\colon T_0\mathbb{R}^{m+k} \to T_0\mathbb{R}^k$, and hence $\dim W = m$. Since χ^1 is a linear maps, its differential $\chi^1_*\colon T_0\mathbb{R}^{m+k} = \mathbb{R}^{m+k} \to T_0\mathbb{R}^m = \mathbb{R}^m$ is equal to χ^1. In particular, χ^1_* restricted to W is given by the isomorphism g and hence $\ker \chi^1_* \cap W = \{0\}$. which finishes the proof of part (i).

Proof of part (ii). To show that T_pM is contained in the kernel of F_*, let $[\gamma] \in T^{\text{geo}}_p M$, where $\gamma\colon (-\epsilon, \epsilon) \to M$ is a smooth path with $\gamma(0) = p$. Then $F_*[\gamma] = [F \circ \gamma]$ is the zero element of $T^{\text{geo}}_q Q$, since $F \circ \gamma(t) = q$ is the constant path at q. This proves $T_pM \subset \ker F_*$. To prove the converse inclusion we note that both vector spaces have the same dimension, since

$$\dim T_pM = \dim M = \dim N - \dim Q$$

by part (i). The dimension of the kernel of the surjective map $F_*\colon T_pN \to T_qQ$ is $\dim T_pN - \dim T_qQ = \dim N - \dim Q$, which proves part (ii).

□

6 Smooth vector bundles

The goal of this section is to define the notion of smooth section of a vector bundle E over a smooth manifold M. We will begin with examples motivating the need for this notion, follow it up with a preliminary definition which captures some but not all the required features and end up with the technical Definition **??** of smooth vector bundles and their sections.

Example 6.1. Let M be a smooth manifold and let $\gamma\colon \mathbb{R}\times M \to M$ be a smooth action of the group $\mathbb{R}$. For $p \in M$ let $\gamma_p\colon \mathbb{R} \to M$ be the smooth path given by $\gamma_p(t) := \gamma(t,p)$. The path γ_p represents an element $[\gamma_p]$ in the geometrically defined tangent space. Conforming with the standard notation in the case of open subsets of Euclidean space, we will use the notation $\gamma_p'(0) := [\gamma_p] \in T_pM$. The assignment

$$M \ni p \;\mapsto\; \gamma_p'(0) \in T_pM$$

is an example of a *vector field* on M in the sense of the following definition.

Definition 6.2. (Preliminary!) A vector field on a smooth manifold M is an assignment X that assigns to any point $p \in M$ a tangent vector $X(p) \in T_pM$.

Example 6.3. Let $f\colon M \to \mathbb{R}$ be a smooth function on a smooth manifold M. For $p \in M$, let

$$df_p = f_*\colon T_pM \longrightarrow T_{f(p)}\mathbb{R} = \mathbb{R}$$

be the differential of f at the point p (as usual we identify here the tangent space $T_q\mathbb{R}^n$ at $q \in \mathbb{R}^n$ with the vector space $\mathbb{R}^n$). We recall that for a vector space V the space $\mathrm{Hom}(V,\mathbb{R})$ is called the *dual vector space*and is denoted by V^*. In particular, df_p is an element of $\mathrm{Hom}(T_pM,\mathbb{R}) = (T_pM)^*$, which is called the *cotangent space of M at p* and is denoted by T_p^*M. The assignment df given by

$$M \ni p \;\mapsto\; df_p \in T_p^*M,$$

is called the *differential of f*. It is an example of a 1-form, defined as follows.

Definition 6.4. (Preliminary!) A 1-*form* α on a smooth manifold M is an assignment

$$M \ni p \;\mapsto\; \alpha_p \in T_p^*M.$$

Extracting the commonality of these examples, we make the following (preliminary!) definition.

Definition 6.5. (Preliminary!) A *vector bundle E of rank k over a smooth manifold M* is a family $\{E_p\}_{p\in M}$ of vector space E_p of dimension k parametrized by points p in M. The vector space E_p is called the *fiber over* p. A *section* s of E is an assignment

$$M \ni p \;\mapsto\; s(p) \in E_p.$$

Example 6.6. Let V be a vector space of dimension k. Then the vector bundle E given by $E_p = V$ for all $p \in M$ is called the *trivial vector bundle over M with fiber V*. We note that a section of this bundle is simply a map $M \to V$ with values in the vector space M. So a section s of a general vector bundle E should be viewed as generalization of vector-valued function on M, whose value $s(p)$ at a point $p \in M$ is a vector $s(p)$ in a vector space E_p, which for a general vector bundle *depends on the point* p.

Question. What is missing in the above definitions of vector field, 1-form and section?

To see what is missing, we revisit the examples 6.1 and 6.3 in the special case where the smooth manifold M is an open subset of $\mathbb{R}^n$.

- $\mathbb{R}^n \supset M \ni p \;\mapsto \gamma_p'(0) \in T_pM = \mathbb{R}^n$ is an $\mathbb{R}^n$-valued function. This is a *smooth* map, as can be seen by rewriting $\gamma_p'(0)$ in the form
$$\gamma_p'(0) = \frac{\partial\gamma}{\partial t}(0, p_1, \dots, p_n)$$
- Similarly, $df_p \in T_p^*M = \mathrm{Hom}(T_pM, \mathbb{R}) = \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^* \cong \mathbb{R}^n$ using our standard identification $T_pM = \mathbb{R}^n$ for open subsets $M \subset \mathbb{R}^n$ and the isomorphism $(\mathbb{R}^n)^* \cong \mathbb{R}^n$, under which the standard basis vector $e_i \in \mathbb{R}^n$ corresponds to the i-th vector $e^i \in (\mathbb{R}^n)^*$ of the dual basis $\{e^i\}_{i=1,\dots,n}$ of the dual space $(\mathbb{R}^n)^*$. It is easy to check that via this isomorphism, the cotangent vector $df_p \in T_p^*M$ corresponds to $(\mathrm{grad} f)_p \in \mathbb{R}^n$, which is a *smooth function* of p.

Definition 6.7. Let M be a smooth manifold. A *smooth vector bundle of rank k over M* consists of the following data:

1. A smooth manifold E, called the *total space* and a smooth map $\pi\colon E \to M$.
2. For each $p \in M$ the set $E_p := \pi^{-1}(p)$, called *the fiber over* p, has the structure of a k-dimensional vector space.

It is required that E is *locally trivial* in the sense that for each point $p \in M$, there is an open neighborhood U and a diffeomorphism Φ making the diagram

$$\begin{tikzcd} E_{|U} := \pi^{-1}(U) \arrow[rr, "\Phi"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\ & M & \end{tikzcd}$$

commutative, such that the map $E_p \to \{p\} \times \mathbb{R}^k = \mathbb{R}^k$ given by restriction of Φ is a vector space isomorphism for each $p \in U$. The map Φ is called a *local trivialization* of E. We note that this implies in particular that $\dim E = \dim(U \times \mathbb{R}^k) = \dim U + \dim \mathbb{R}^k = \dim M + k$.

A *section* is a map $s\colon M \to E$ with $\pi \circ s = \mathrm{id}_M$; in other words, $s(p)$ belongs to the fiber E_p for every $p \in M$. A section s is *smooth* if $s\colon M \to E$ is a smooth map. The vector space of smooth sections of E will be denoted $\Gamma(M, E)$.

Remark 6.8. There is an analogous definition of a vector bundle E over a topological space M. Here E is required to be topological space, $\pi\colon E \to M$ is continuous map, and the local trivializations Φ are required to be homeomorphisms. Vector bundles over topological spaces are important objects in topology, but since our goal this part of the semester is to do calculus on manifolds, we will focus on smooth vector bundles.

Example 6.9. (Examples of smooth vector bundles).

1. Let M be a smooth manifold and let V be a finite dimensional vector space. Then $E = M \times V$ equipped with the projection map $\pi\colon E \to M$ is a smooth vector bundle over M called the *trivial vector bundle with fiber* V. It is clear that the product $E = M \times V$ is a smooth manifold, and that the projection map π is smooth. Each fiber $E_p = \pi^{-1}(p) = \{p\} \times V = V$ also has an obvious vector space structure. To show that E is locally trivial, we choose $U = M$, pick a vector space isomorphism $h\colon V \xrightarrow{\cong} \mathbb{R}^k$ (which always exists for $k = \dim V$) and define
$$\Phi\colon E_{|U} = M \times V \longrightarrow M \times \mathbb{R}^k \qquad (p, v) \mapsto (p, h(v))$$
This map satisfies all requirements of a locally trivialization.

2. Let U_1, U_2 be the open subsets of S^1 defined by $U_1 := S^1 \setminus \{-1\}$ and $U_2 := S^1 \setminus \{1\}$. Let E be the quotient of the disjoint union
$$U_1 \times \mathbb{R} \quad \amalg \quad U_2 \times \mathbb{R} \tag{6.10}$$
modulo the equivalence relation $\sim$ defined by
$$(1, z, x) \sim (2, z, \epsilon(z)x) \qquad \text{for } z \in U_1 \cap U_2 \text{ and} \quad \epsilon(z) := \begin{cases} +1 & \text{for } \mathrm{im}(z) > 0 \\ -1 & \text{for } \mathrm{im}(z) < 0 \end{cases}$$
Here $(1, z, x) \in U_1 \times \mathbb{R}$ and $(2, z, x) \in U_2 \times \mathbb{R}$, i.e., the number in the first component just indicates whether (z, x) is to be considered as an element of the first or the second summand in the disjoint union (6.10). Moreover, $\mathrm{im}(z)$ is the imaginary part of $z \in U_1 \cap U_2 \subset S^1 \subset \mathbb{C}$. The map $\pi\colon E \to S^1$ given by $[1, z, x] \mapsto z$ and $[2, z, x] \mapsto z$ is a well-defined continuous map (by the continuity property of maps out of quotients;

the pre-composition of π with the projection map from $(U_1 \times \mathbb{R}) \amalg (U_2 \times \mathbb{R})$ to the quotient E is clearly continuous).

We construct bundle charts (U_1, Φ_1) and (U_2, Φ_2) for E as follows:

$$\Phi_i \colon E_{|U_i} \longrightarrow U_i \times \mathbb{R} \quad \text{given by} \quad \Phi_i[i, z, x] = (z, x) \tag{6.11}$$

Obviously, the restriction of Φ_i to each fiber E_z for $z \in U_i$ is a vector space isomorphism, and it is not hard to check that the maps Φ_i and their inverses are continuous, and so Φ_i is a homeomorphism. The only thing not clear is why Φ is a diffeomorphism; in fact, it is not even clear what that would mean, since we haven't constructed a smooth structure on E!

So our goal is to construct a smooth atlas for E in such a way that the maps Φ_i are diffeomorphisms. We observe that the homeomorphisms $\Phi_i \colon E_{|U_i} \xrightarrow{\approx} U_i \times \mathbb{R}$ can essentially be thought of as *charts* for E. This is not literally true, since U_i is an open subset of S^1 rather than an open subset of $\mathbb{R}$. However, $U_i \subset S^1$ is diffeomorphic to an open subset of $\mathbb{R}$, e.g. the map $(-1, 1) \to U_1$, $t \mapsto e^{\pi i t}$ is a diffeomorphism; similarly for U_2. Secretly composing with these diffeomorphisms, we will allow ourselves to think of Φ_1, Φ_2 as charts for E. To show that $\{(E_{U_1}, \Phi_1), (E_{U_2}, \Phi_2)\}$ is a smooth atlas, we need to check that the transition maps are smooth. For example, $\Phi_2 \circ \Phi_1^{-1}$ is given explicitly as follows:

$$\begin{array}{ccccc}
(U_1 \cap U_2) \times \mathbb{R} & \xrightarrow{\quad \Phi_1^{-1} \quad} & E_{|U_1 \cap U_2} & \xrightarrow{\quad \Phi_2 \quad} & (U_1 \cap U_2) \times \mathbb{R} \\
\\
(z, x) & \longmapsto & [1, z, x] = [2, z, \epsilon(z)x] & \longmapsto & (z, \epsilon(z)x)
\end{array}$$

We note that this map is locally constant; in particular, it is smooth. It is equal to its own inverse inverse, and hence it is a diffeomorphism, which proves that $\{(E_{U_1}, \Phi_1), (E_{U_2}, \Phi_2)\}$ is a smooth atlas. As we have argued before, each chart of a smooth atlas is a diffeomorphisms between an open subset of the manifold and its image, which is an open subset of Euclidean space. In particular, the maps Φ_i of (6.11) are diffeomorphisms.

Lemma 6.12. (Vector Bundle Construction Lemma). *Let M be a smooth manifold of dimension n, and let $\{E_p\}$ be a collection of vector spaces parametrized by $p \in M$. Let E be the set given by the disjoint union of all these vector spaces, which we write as*

$$E := \coprod_{p \in M} E_p = \{(p, v) \mid p \in M,\ v \in E_p\}$$

and let $\pi \colon E \to M$ be the projection map defined by $\pi(p, v) = p$. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of M, and let for each $\alpha \in A$, let $\Phi_\alpha \colon \pi^{-1}(U_\alpha) \longrightarrow U_\alpha \times \mathbb{R}^k$ be maps with the following properties *(i) The diagram*

$$\begin{tikzcd} E_{|U_\alpha} := \pi^{-1}(U_\alpha) \arrow[rr, "\Phi_\alpha"] \arrow[dr, "\pi"'] & & U_\alpha \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\ & U_\alpha & \end{tikzcd}$$

is commutative, where π_1 is the projection onto the first factor.

(ii) For each $p \in U_\alpha$, the restriction of Φ_α to $E_p = \pi^{-1}(p)$ is a vector space isomorphism between E_p and $\{p\} \times \mathbb{R}^k = \mathbb{R}^k$ (which implies that Φ_α is a bijection).

(iii) For $\alpha, \beta \in A$, the composition

$$(U_\alpha \cap U_\beta) \times \mathbb{R}^k \xrightarrow{\Phi_\alpha^{-1}} \pi^{-1}(U_\alpha \cap U_\beta) \xrightarrow{\Phi_\beta} (U_\alpha \cap U_\beta) \times \mathbb{R}^k$$

is smooth.

Then the total space E can be equipped with the structure of a smooth manifold of dimension $n+k$ such that $\pi \colon E \to M$ is a smooth vector bundle of rank k with local trivializations Φ_α.

6.1 Measurements in manifolds

Let U be an open subset of $\mathbb{R}^n$ and let $\gamma \colon [a,b] \to U$ be a smooth path. Then the length of the path γ is given by

$$\text{length}(\gamma) = \int_1^b ||\gamma'(t)||dt, \tag{6.13}$$

where $||\gamma'(t)||$ is the norm of the tangent vector $\gamma'(t)$ of the path at the point $\gamma(t) \in U$. If γ is a smooth path in a manifold M, we would like to calculate the length of γ in a similar way. For each $t \in [a,b]$ the tangent vector $\gamma'(t)$ belongs to the tangent space $T_{\gamma(t)}M$, and so the question is how to make sense of the norm $||\gamma'(t)||$. We recall that the usual devise to make sense of the norm $||v|| \in [0,\infty)$ of a vector v of a vector space V is the following.

Definition 6.14. An *inner product* on a vector space V is a map $g \colon V \times V \to \mathbb{R}$ with the following properties:

(i) **multilinear:** g is a linear function in each of its two slots;

(ii) **symmetric:** $g(v,w) = g(w,v)$ for $v, w \in V$;

(iii) **positive definite:** $g(v,v) \geq 0$ for all $v \in V$ and $g(v,v) = 0$ if and only if $v = 0$.

A map $g \colon V \times V \to \mathbb{R}$ satisfying (i) and (ii) is called a *symmetric bilinear form on V*. The set of all symmetric bilinear forms is a vector space which is denoted $\mathsf{Sym}^2(V;\mathbb{R})$.

The usual scalar product on $\mathbb{R}^n$ given by $g(v,w) = v_1 w_1 + \cdots + v_n w_n$ for $v, w \in \mathbb{R}^n$, $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$, is an inner product on $\mathbb{R}^n$. The scalar product on $\mathbb{R}^n$ allows us to calculate the length $||v||$ of a vector $v \in \mathbb{R}^n$ or the angle $\alpha(v,w) \in [0,\pi]$ between vectors $v, w \in \mathbb{R}^n$. Similarly, an inner product g on a vector space V allows us to do the same for vectors $v, w \in V$ by defining:

$$||v|| := \sqrt{g(v,v)} \qquad \cos\alpha(v,w) := \frac{g(v,w)}{||v||||w||}.$$

So an inner product on a vector space V should be thought of as a "yard stick" making it possible to do measurements of lengths and angles in V. In particular in order to talk about the norm of tangent vectors of a smooth manifold M, we need an inner product g_p on the tangent space T_pM for all points $p \in M$. What we want to express is the desideratum that the inner product $g_p \in \mathsf{Sym}^2(T_pM;\mathbb{R})$ "depends smoothly on p". This is entirely analogous to asking how to make precise the statement that for a smooth function $f \in C^\infty(M)$ the differential $df_p \in T_p^*M = \mathrm{Hom}(T_pM, \mathbb{R})$ "depends smoothly on p".

Lemma 6.15. *Let E be a smooth vector bundle over a smooth manifold M. Then there is a smooth vector bundle $\mathsf{Sym}^2(E;\mathbb{R})$ whose fiber over $p \in M$ is the vector space $\mathsf{Sym}^2(E_p;\mathbb{R})$ of symmetric bilinear forms on the fiber E_p.*

The construction of the vector bundle $\mathsf{Sym}^2(E;\mathbb{R})$ is entirely analogous to the construction of the dual vector bundle E^*: from the local trivializations $E_{|U_\alpha} \cong U_\alpha \times \mathbb{R}^k$ of E we build maps

$$\coprod_{p\in U_\alpha} \mathsf{Sym}^2(E_p;\mathbb{R}) \longrightarrow U_\alpha \times \mathsf{Sym}^2(\mathbb{R}^k;\mathbb{R}) \cong U_\alpha \times \mathbb{R}^\ell, \qquad \ell = \dim \mathsf{Sym}^2(\mathbb{R}^k;\mathbb{R}).$$

which commute with the projection maps to U_α and are fiberwise isomorphisms of vector spaces. Then the Vector Bundle Construction Lemma 6.12 can be used to show that $\mathsf{Sym}^2(E;\mathbb{R}) = \{(p,v) \mid p \in M,\ v \in \mathsf{Sym}^2(E_p;\mathbb{R})\}$ has the structure of a smooth vector bundle.

Definition 6.16. Let M be a smooth manifold. A *Riemannian metric on M* is a smooth section $g\colon M \to \mathsf{Sym}^2(TM;\mathbb{R})$ of the vector bundle $\mathsf{Sym}^2(TM;\mathbb{R})$ such that for each $p \in M$ the symmetric bilinear form $g_p \in \mathsf{Sym}^2(T_pM;\mathbb{R})$ is positive definite (in particular, g_p is an inner product on the tangent space T_pM for every $p \in M$).

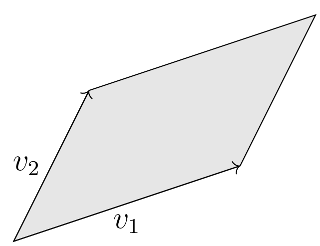
If M is a Riemannian manifold and $\gamma\colon [a,b] \to M$ is a smooth path in M, then the length of γ is defined by the formula (6.13), where the norm $||\gamma'(t)|| \in [0,\infty)$ of the tangent vector $\gamma'(t) \in T_pM$, $p = \gamma(t)$ is evaluated using the inner product g_p on T_pM.

6.1.1 Measuring volumes

Our eventual goal is to integrate over manifolds. When defining the Riemann integral of a function f over an open subset $U \subset \mathbb{R}^n$ we divide U into a bunch of small boxes and approximate the integral over f by the integral over a function which is constant on each small box, thus reducing the calculation of an integral to the calculation of the volume of rectangles. To define integration over manifolds we will use charts to reduce the calculation to open subsets of Euclidean space. However, it turns out to to be useful to not restrict ourselves to *rectangles* in $\mathbb{R}^n$, since the image of a rectangle in $\mathbb{R}^2$ under a linear map $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is typically no box, but a parallelogram. More generally, the image of the standard n-cube $\{(x_1, \dots, x_n) \in \mathbb{R}^n \mid 0 \le x_i \le 1\} \subset \mathbb{R}^n$ under a linear map T with $T(e_i) = v_i \in \mathbb{R}^n$ is the parallelepiped

$$P(v_1, \dots, v_n) := \{\sum_{i=1}^n x_i v_i \mid 0 \le x_i \le 1\} \subset \mathbb{R}^n.$$

For $n = 2$, a parallelepiped $P(v_1, v_2)$ is simply the parallelogram spanned by the vectors v_1, v_2 (which is a “degenerate” if v_1, v_2 are linearly dependent). Here is a picture of $P(v_1, v_2)$:



Lemma 6.17. *The volume of the (possibly degenerate) parallelepiped $P(v_1, \dots, v_n)$ spanned by $v_1, \dots, v_n \in \mathbb{R}^n$ is given by the formula*

$$\mathrm{vol}(P(v_1, \dots, v_n)) = |\det(v_1, \dots, v_n)|.$$

Here det *is interpreted as a map* $\det\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ *that sends an n-tupel $(v_1, \dots, v_n)$ of vectors $v_i \in \mathbb{R}^n$ to the determinant of the $n \times n$ matrix with column vectors v_1, ..., v_n.*

We will prove this statement since the techniques going into that proof will be useful for us. Before doing so, we recall properties of the determinant function $\det\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$.

1. The determinant is a *multilinear* map, i.e., it is linear in each slot; explicitly,

   $$\det(v_1, \dots, av_i + bv_i', \dots, v_n) = a\det(v_1, \dots, v_i, \dots, v_n) + b\det(v_1, \dots, v_i', \dots, v_n)$$

   for $v_1, \dots, v_n, v_i' \in \mathbb{R}^n$, $a, b \in \mathbb{R}$.

2. The determinant is *alternating*, i.e., for any permutation $\sigma \in S_k$

$$\det(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \operatorname{sign}(\sigma) \det(v_1, \dots, v_n),$$

where $\operatorname{sign}(\sigma) \in \{\pm 1\}$ is the sign of the permutation σ. We recall that $\operatorname{sign}(\sigma) = 1$ if σ is the composition of an even number of transpositions; otherwise $\operatorname{sign}(\sigma) = -1$.

Definition 6.18. Let V be a vector space. A map

$$\omega \colon \underbrace{V \times \cdots \times V}_{k} \longrightarrow \mathbb{R}$$

is called

1. **multilinear** if ω is linear in each slot;
2. **alternating** if $\omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \operatorname{sign}(\sigma)\omega(v_1, \dots, v_k)$ for all $v_1, \dots, v_k$ and $\sigma \in S_k$.

Let $\operatorname{Alt}^k(V; \mathbb{R})$ denote the set of multilinear alternating maps $\omega \colon V \times \cdots \times V \to \mathbb{R}$. This is a vector space, since the sum of two multilinear alternating maps is again a multilinear alternating map; multiplying such a map by a constant $c \in \mathbb{R}$ again such a map.

To calculate the dimension of $\operatorname{Alt}^k(V; \mathbb{R})$, we want to construct a basis for this vector space. Let $\{e_i\}_{i=1,\dots,n}$ be a basis for V, and let $\{e^i\}_{i=1,\dots,n}$ be the dual basis for V^*. Given a multi-index $I = (i_1, \dots, i_k)$ with $i_j \in \{1, \dots, n\}$, it is evident that the map

$$\underbrace{V \times \cdots \times V}_{k} \longrightarrow \mathbb{R} \qquad \text{given by} \qquad (v_1, \dots, v_k) \mapsto e^{i_1}(v_1)e^{i_2}(v_2) \cdots e^{i_k}(v_k)$$

is multilinear. However, in general it is not alternating, since the value of this function on a k-tupel $(v_1, \dots, v_k)$ is unrelated to the value on the permuted k-tupel $(v_{\sigma(1)}, \dots, v_{\sigma(k)})$. For example, if $k = n$, $I = (1, \dots, n)$ and $v_i = e_i$, then

$$e^1(v_1) \dots e^n(v_n) = 1 \qquad \text{but} \qquad e^1(v_{\sigma(1)}) \dots e^n(v_{\sigma(n)}) = 0 \quad \text{for } \sigma \neq \mathrm{id}.$$

However, out of this non-alternating multilinear map we can can manufacture an alternating map e^I by a signed sum over permutations of the v_i:

$$e^I(v_1, \dots, v_k) := \sum_{\sigma \in S_k} \operatorname{sign}(\sigma) e^{i_1}(v_{\sigma(1)}) e^{i_2}(v_{\sigma(2)}) \cdots e^{i_k}(v_{\sigma(k)})$$

for $I = (i_1, \dots, i_k)$, $v_1, \dots, v_k \in V$. It is not hard to check that the multilinear map

$$e^I \colon V \times \cdots \times V \to \mathbb{R}$$

is in fact alternating and so $e^I \in \operatorname{Alt}^k(V; \mathbb{R})$. It is also straightforward to show that if $J = (i_{\sigma(1)}, \dots, i_{\sigma(k)})$ is a permutation of the multi-index $I = (i_1, \dots, i_k)$, then $e^J = \operatorname{sign}(\sigma) e^I$.

Lemma 6.19. *The collection* $\{e^I \mid I = (i_1, \ldots, i_k)$ *with* $i_1 < i_2 < \cdots < i_k\}$ *is a basis for* $\mathrm{Alt}^k(V;\mathbb{R})$.

A proof of this fact can be found in Lee's book. It follows that the dimension of the vector space $\mathrm{Alt}^k(V;\mathbb{R})$ is equal to the number of multi-indices $I = (i_1, \ldots, i_k)$ which are strictly increasing in the sense that $i_1 < \cdots < i_k$. Mapping a strictly increasing multi-index $I = (i_1, \ldots, i_k)$ to the subset $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ yields a bijection between the set of strictly increasing multi-indices and the set of cardinality k subsets of $\{1, \ldots, n\}$. In particular, we conclude:

Corollary 6.20. *If V is a vector space of dimension n, then* $\dim \mathrm{Alt}^k(V;\mathbb{R}) = \binom{n}{k}$.

Proof of Lemma 6.17. Our strategy to prove $\mathrm{vol}(P(v_1, \ldots, v_n)) = |\det(v_1, \ldots, v_n)|$ is to use the fact that det is an alternating multilinear map, i.e., an element of $\mathrm{Alt}^n(\mathbb{R}^n;\mathbb{R})$, and that the dimension of $\mathrm{Alt}^n(\mathbb{R}^n;\mathbb{R})$ is $\binom{n}{n} = 1$. The idea is that while $\mathrm{vol}(P(v_1, \ldots, v_n))$ is *not* an alternating multilinear map (e.g., its values are non-negative), it is the *absolute value* of an alternating multilinear map

$$\mathrm{svol}\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n} \longrightarrow \mathbb{R}$$

called the *signed volume*, defined by $\mathrm{svol}(v_1, \ldots, v_n) := \epsilon(v_1, \ldots, v_n)\,\mathrm{vol}(P(v_1, \ldots, v_n))$, where

$$\epsilon(v_1, \ldots, v_n) := \begin{cases} +1 & \det(v_1, \ldots, v_n) > 0 \\ -1 & \det(v_1, \ldots, v_n) < 0 \\ 0 & \det(v_1, \ldots, v_n) = 0 \end{cases}$$

It is clear from the definition that $|\mathrm{svol}(v_1, \ldots, v_n)| = \mathrm{vol}(P(v_1, \ldots, v_n))$, and we claim that svol is an element of $\mathrm{Alt}^n(\mathbb{R}^n;\mathbb{R})$. Permuting the vectors $v_1, \ldots, v_n$ does not change the associated parallelepiped, but

$$\det(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sign}(\sigma)\det(v_1, \ldots, v_n)$$

and hence

$$\epsilon(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sign}(\sigma)\epsilon(v_1, \ldots, v_n).$$

It follows that svol is alternating. To show that svol is linear in each slot, let us first argue that

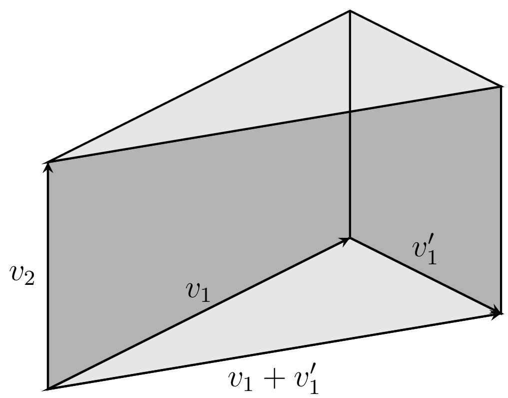
$$\mathrm{svol}(v_1, \ldots, cv_i, \ldots, v_n) = c\,\mathrm{svol}(v_1, \ldots, v_i, \ldots, v_n) \tag{6.21}$$

for $c \in \mathbb{R}$. If c is a positive integer, this is clear geometrically; for $c = -1$, again it is clear geometrically that the volume of the associated parallelepipeds $P(v_1, \ldots, -v_i, \ldots, v_n)$ and $P(v_1, \ldots, v_i, \ldots, v_n)$ is the same, but $\epsilon(v_1, \ldots, -v_i, \ldots, v_n) = -\epsilon(v_1, \ldots, v_i, \ldots, v_n)$. This implies equation (6.21) for $c \in \mathbb{Z}$ and hence for $c \in \mathbb{Q}$. Approximating a real number

$c \in \mathbb{R}$ by rational numbers c_ℓ and taking the limit of the equation $\operatorname{svol}(v_1, \dots, c_\ell v_i, \dots, v_n) = c_\ell \operatorname{svol}(v_1, \dots, v_i, \dots, v_n)$ as for $\ell \to \infty$ yields equation for a general $c \in \mathbb{R}$. The additivity property

$$\operatorname{svol}(v_1, \dots, v_i + v_i', \dots, v_n) = \operatorname{svol}(v_1, \dots, v_i, \dots, v_n) + \operatorname{svol}(v_1, \dots, v_i', \dots, v_n)$$

follows from a geometric argument which is illustrated by the following picture for $n = 2$: the area of the parallelogram $P(v_1 + v_1', v_2$ is equal to the sum of the areas of $P(v_1, v_2)$ and $P(v_1', v_2)$.



This shows that the signed volume svol is an alternating multilinear map, i.e., an element of $\operatorname{Alt}^n(\mathbb{R}^n; \mathbb{R})$. Since this vector space has dimension 1, the element svol must be a scalar multiple of the non-zero element $\det \in \operatorname{Alt}^n(\mathbb{R}^n; \mathbb{R})$, i.e., $\operatorname{svol} = c \det$ for some $c \in \mathbb{R}$. To determine c, we evaluate both sides on the n-tupel $(e_1, \dots, e_n)$, where $\{e_i\}_{i=1,\dots,n}$ is the standard basis of $\mathbb{R}^n$. The determinant $\det(e_1, \dots, e_n)$ is the determinant of the identity matrix and hence 1. The parallelepiped $P(e_1, \dots, e_n)$ is the standard cube which has volume 1 and hence $\operatorname{svol}(e_1, \dots, e_n) = \operatorname{vol}(e_1, \dots, e_n) = 1$. It follows that $c = 1$, and hence for every n-tupel $(v_, \dots, v_n)$ of vectors $v_1, \dots, v_n \in \mathbb{R}^n$ we have

$$\operatorname{svol}(v_1, \dots, v_n) = \det(v_1, \dots, v_n).$$

Taking the absolute value of both sides we conclude the statement of Lemma 6.17. □

I think of an element $\omega \in \operatorname{Alt}^k(V; \mathbb{R})$ as a little machine that takes an input of vectors $v_1, \dots, v_k \in V$ and produces as output the number $\omega(v_1, \dots, v_k) \in \mathbb{R}$; this output depends linearly on each v_i, and permuting the input vectors changes the output by a factor of ± 1, given by the sign of the permutation $\sigma \in S_k$. This suggests that we can multiply a machine ω with k inputs and η with ℓ inputs to obtain a machine typically denoted $\omega \otimes \eta$ with $k + \ell$ inputs by defining:

$$(\omega \otimes \eta)(v_1, \dots, v_{k+\ell}) := \omega(v_1, \dots, v_k)\eta(v_{k+1}, \dots, v_{k+\ell}).$$

It is clear that $\omega \otimes \eta$ is linear in each of its $k+\ell$ input slots, and changes by a factor of $\mathrm{sign}(\sigma) \in \{\pm 1\}$ when using a permutation $\sigma \in S_{k+\ell}$ to permute the $k+\ell$ input vectors, *provided* σ belongs to the subgroup $S_k \times S_\ell \subset S_{k+\ell}$. If we interchange one of the vectors $v_1, \dots, v_k$ with one of the vectors $v_{k+1}, \dots, v_{k+\ell}$, there is no reason that the output just is multiplied by -1. To produce an *alternating* multilinear map, we use the same method we applied before by using a signed sum over all permutations.

Definition 6.22. For $\omega \in \mathrm{Alt}^k(V;\mathbb{R})$ and $\eta \in \mathrm{Alt}^\ell(V;\mathbb{R})$ their *wedge product* is the alternating multilinear form $\omega \wedge \eta \in \mathrm{Alt}^{k+\ell}(V;\mathbb{R})$ defined by

$$(\omega \wedge \eta)(v_1, \dots, v_{k+\ell}) := \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \mathrm{sign}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}).$$

We note that for $\sigma \in S_k \times S_\ell \subset S_{k+\ell}$ the summand

$$\mathrm{sign}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)})$$

is equal to $\omega(v_1, \dots, v_k)\eta(v_{k+1}, \dots, v_{k+\ell})$. In particular, summing over this subgroup of order $k!\ell!$, we simply obtain $k!\ell!\omega(v_1, \dots, v_k)\eta(v_{k+1}, \dots, v_{k+\ell})$. This motivates the factor $\frac{1}{k!\ell!}$ in the definition of the wedge product.

Lemma 6.23. (Properties of the wedge product).

1. **Bilinearity:**
2. **Associativity:** $\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \xi$
3. **Graded Commutativity:** $\omega \wedge \eta = (-1)^{k\ell} \eta \wedge \omega$ *for* $\omega \in \mathrm{Alt}^k(V;\mathbb{R})$ *and* $\eta \in \mathrm{Alt}^\ell(V;\mathbb{R})$.
4. *For any multi-index* $I = (i_1, \dots, i_k)$, $e^I = e^{i_1} \wedge \dots \wedge e^{i_k} \in \mathrm{Alt}^k(V;\mathbb{R})$. *In particular, an element* $\omega \in \mathrm{Alt}^k(V;\mathbb{R})$ *can uniquely be written as a sum*

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1,\dots,i_k} e^{i_1} \wedge \dots \wedge e^{i_k} \qquad a_{i_1,\dots,i_k} \in \mathbb{R}.$$

Lemma 6.24. *Let M be a smooth manifold and $E \to M$ a smooth vector bundle over M. Then there is a smooth vector bundle* $\mathrm{Alt}^k(E;\mathbb{R}) \to M$ *whose fiber at a point $p \in M$ is* $\mathrm{Alt}^k(E_p;\mathbb{R})$.

We note that $\mathrm{Alt}^1(E_p;\mathbb{R})$ is the dual space E_p^*, and $\mathrm{Alt}^1(E;\mathbb{R}) \to M$ is the dual vector bundle $E* \to M$. Like the construction of the dual vector bundle, the proof of the above statement uses the Vector Bundle Construction Lemma 6.12.

We recall that a 1-form on a smooth manifold M is a section of the cotangent bundle T^*M. Noting that T^*M is equal to the vector bundle $\mathrm{Alt}^1(TM;\mathbb{R})$, this suggests the following generalization of 1-forms.

Definition 6.25. Let M be a smooth manifold. A *k-form* or *differential form of degree k* on M is a smooth section of the vector bundle $\mathrm{Alt}^k(TM;\mathbb{R})$. The usual notation for the vector space of k-forms on M is

$$\Omega^k(M) := \Gamma(M; \mathrm{Alt}^k(TM;\mathbb{R})).$$

How do we do explicit calculations with differential forms? To do calculations with linear maps between vector spaces, it is often useful to choose a basis for the vector spaces involved. Thinking of a vector bundle as a collection of vector spaces parametrized by points $p \in M$, it is natural to ask how to generalize the notion of "basis" from vector spaces to vector bundles.

Definition 6.26. Let $E \to M$ be a smooth vector bundle of rank k. If $U \subset M$ is an open subset, a *local frame* for E over U is a collection $\{b^i\}_{i=1,\dots,k}$ of smooth sections of $E_{|U}$ such that $\{b^1_p, \dots, b^k_p\}$ is a basis of E_p for each $p \in U$.

Let $\{b^i\}_{i=1,\dots,k}$ be a local frame for E over U and let $s \in \Gamma(U;E)$, i.e., $s\colon U \to E$ is a smooth section of $E_{|U}$. Then for any $p \in U$, the element $s(p) \in E_p$ can be expanded in terms of the basis $\{b^i(p)\}$ to obtain

$$s(p) = \sum_{i=1,\dots,k} s_i(p) b^i(p) \qquad \text{with } s_i(p) \in \mathbb{R}.$$

It is not hard to see that $s_i(p)$ is a *smooth* function of p, since s and b^i are smooth sections of E. Hence we can write the section s as a linear combination

$$s = \sum_{i=1,\dots,k} s_i b^i$$

of the sections b^i whose coefficients s_i are smooth functions $U \to \mathbb{R}$.

Example 6.27. Let M be a smooth manifold of dimension n, and let $M \supset U \xrightarrow{\phi} \mathbb{R}^n$ be a smooth chart (i.e., (U,ϕ) belongs to the maximal smooth atlas defining the smooth structure on M). Let $x^1, \dots, x^n \in C^\infty(U)$ be the component functions of ϕ, i.e., $\phi(p) = (x^1(p), \dots, x^n(p))$. Then the differentials $dx^1_p, \dots, dx^n_p \in T^*_pM$ form a basis

6.2 Algebraic structures on differential forms

The goal of this section is to discuss the various algebraic structures on differential forms and their compatibility.

Definition 6.28. (Pullback). Let M, N be smooth manifolds and $F\colon M \to N$ a smooth map. Given a differential form $\omega \in \Omega^k(N)$, its *pullback* $F^*\omega \in \Omega^k(M)$ is defined by

$$(F^*\omega)_p(v_1, \dots, v_k) := \omega_p(F_*v_1, \dots, F_*v_k) \qquad \text{for } p \in M,\ v_1, \dots, v_k \in T_pM.$$

In more detail: the k-form $F^*\omega$ is a section of the vector bundle $\mathrm{Alt}^k(TM;\mathbb{R})$, and hence it can be evaluated at $p \in M$ to obtain an element $(F^*\omega)_p$ in the fiber of that vector bundle over p, which is $\mathrm{Alt}^k(T_pM;\mathbb{R})$. In other words, $(F^*\omega)_p$ is an alternating multilinear map

$$(F^*\omega)_p\colon \underbrace{T_pM \times \cdots \times T_pM}_{k} \longrightarrow \mathbb{R},$$

and hence it can be evaluated on the k tangent vectors $v_1, \ldots, v_k \in T_pM$ to obtain a real number $(F^*\omega)_p(v_1, \ldots, v_k)$. On the right hand side to the equation defining $F^*\omega$, the map $F_*\colon T_pM \to T_{F(p)}N$ is the differential of F. Hence the alternating multilinear map $\omega_{F(p)} \in \mathrm{Alt}^k(T_{F(p)}N;\mathbb{R})$ can be evaluated on $F_*v_1, \ldots, F_*v_k$ to obtain the real number $\omega_p(F_*v_1, \ldots, F_*v_k)$.

For $k = 1$, $\omega \in \Omega^0(N) = C^\infty(N)$ is a smooth function, and its pullback $F^*\omega$ is the previously defined pullback of functions, simply given by $(F^*\omega)(p) = \omega(F(p))$ for $p \in M$. We also previously defined the pullback of a 1-form $\omega \in \Omega^1(N)$, and we showed that the differentials are compatible with pullbacks in the sense that

$$d(F^*f) = F^*(df) \qquad \text{for } f \in C^\infty(N).$$

Definition 6.29. (Wedge products of differential forms. For $\omega \in \Omega^k(M)$ and $\eta \in \Omega^\ell(M)$, their *wedge product* $\omega \wedge \eta \in \Omega^{k+\ell}(M)$ is defined by

$$(\omega \wedge \eta)_p := \omega_p \wedge \eta_p \in \mathrm{Alt}^{k+\ell}(T_pM;\mathbb{R}) \qquad \text{for } p \in M.$$

Lemma 6.30. *The wedge product of differential forms has the following properties.*

(i) **Bilinearity:**

(ii) **Associativity:**

$$\omega \wedge (\eta \wedge \xi) = (\omega \wedge \eta) \wedge \boldsymbol{x} \qquad \textit{for differential forms } \omega,\ \eta,\ \xi \textit{ on a smooth manifold } M.$$

(iii) **Graded Commutativity:**

$$\omega \wedge \eta = (-1)^{k\ell}\eta \wedge \omega \qquad \textit{for } \omega \in \Omega^k(M) \textit{ and } \eta \in \Omega^\ell(M).$$

(iv) **Compatibility with pullbacks:** *If $F\colon M \to N$ is a smooth map and ω, η differential forms on N, then*

$$F^*(\omega \wedge \eta) = (F^*\omega) \wedge (F^*\eta).$$

Proof. The wedge product of differential form is defined *pointwise*, i.e., the wedge product of differential forms $\omega \in \Omega^k(M)$, $\eta \in \Omega^\ell(M)$ is defined by declaring for any point $p \in M$ the element $(\omega \wedge \eta)_p \in \mathrm{Alt}^{k+\ell}(T_pM;\mathbb{R})$ to be the wedge product $\omega_p \wedge \eta_a$ of the alternating multilinear maps $\omega_p \in \mathrm{Alt}^k(T_pM;\mathbb{R})$ and $\eta_p \in \mathrm{Alt}^\ell(T_pM;\mathbb{R})$. It follows that bilinearity, associativity and graded commutativity of the wedge product for alternating multlinear maps stated in Lemma **??** immediately imply these properties for the wedge product of differential forms.

To prove compatibility with pullbacks, let $p \in M$ and $v_1, \dots, v_{k+\ell} \in T_pM$. Then

$$
\begin{aligned}
&(F^*(\omega \wedge \eta))_p(v_1, \dots, v_{k+\ell})\\
=&(\omega \wedge \eta)_{F(p)}(F_*(v_1), \dots, F_*(v_{k+\ell}))\\
=&(\omega_{F(p)} \wedge \eta_{F(p)})(F_*(v_1), \dots, F_*(v_{k+\ell}))\\
=&\frac{1}{k!\ell!}\sum_{\sigma \in S_{k+\ell}} (\omega_{F(p)}(F_*(v_{\sigma(1)}), \dots, F_*(v_{\sigma(\ell)}))(\eta_{F(p)}(F_*(v_{\sigma(k+1)}), \dots, F_*(v_{\sigma(k+\ell)}))\\
=&\frac{1}{k!\ell!}\sum_{\sigma \in S_{k+\ell}} (F^*\omega)_p(v_{\sigma(1)}, v_{\sigma(\ell)})(F^*\eta_p(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)})\\
=&((F^*\omega)_p \wedge (F^*\eta)_p)(v_1, \dots, v_{k+\ell})\\
=&(F^*\omega \wedge F^*\eta)_p(v_1, \dots, v_{k+\ell})
\end{aligned}
$$

□

Definition 6.31. (Proposition/Definition). Let M be a smooth manifold. Then there is a unique map $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$ called the *de Rham differential* with the following properties:

(i) d is linear;

(ii) for $f \in \Omega^0(M) = C^\infty(M)$, the de Rham differential $df \in \Omega^1(M)$ is the usual differential of f;

(iii) d is a *graded derivation*, i.e., it satisfies the following "product rule with signs":

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta. \tag{6.32}$$

(iv) $d^2 = 0$.

Remark 6.33. The signs appearing in the graded commutativity of the wedge product as well in the product rule for the de Rham differential are examples of the meta principle known as *Koszul sign rule*, according to which a good way to deal with objects with an integer degree (like differential forms) and signs, is to set up definitions such that *permuting* *objects of degree k and ℓ results in a sign of* $(-1)^{k\ell}$. This is satisfied for the wedge product (for alternating multilinear maps or for differential forms). This is also the case for the *graded derivation rule* (ii) above. We recall that a derivation of an algebra A is a linear map $D\colon A \to A$ satisfying the product rule

$$D(a \cdot b) = D(a) \cdot b + a \cdot D(b) \qquad \text{for } a, b \in A.$$

We note that on the left hand side of this equation, as well in the first term on the right hand side the symbols occur in the order D, a, b. By contrast, in the second summand the objects a and D switch occur in the opposite order, which according to the Koszul sign paradigm should involve the sign $(-1)^{\deg(D)\deg(a)}$ in a context where these objects have "degrees" $\deg(D), \deg(a) \in \mathbb{Z}$. For example, in equation (6.32), the differential forms ω, η have degrees $\deg(\omega) = k$, $\deg(\eta) = \ell$, and it is reasonable to declare the de Rham differential d to have degree $+1$, since applying it to a differential form of degree k results in a form of degree $k+1$. This shows that the "graded derivation property" (ii) conforms to the Koszul sign paradigm.

Lemma 6.34. (Additional Properties of the de Rham Differential).

1. **(Compatibility with pullbacks).** *If $F\colon M \to N$ is a smooth map, and $\omega \in \Omega^k(N)$, then $d(F^*\omega) = F^*(d\omega)$.*

2. **(Local Formula for d).** *Let (U, ϕ) be a smooth chart for an n-manifold M, and let $x^1, \dots, x^n \in C^\infty(U)$ be the local coordinate functions (the components of $\phi\colon U \to \mathbb{R}^n$). Then*

$$d(f dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

for $f \in C^\infty(U)$, $i_1, \dots, i_k \in \{1, \dots, n\}$.

We remark that the collection of sections $\{dx^{i_1} \wedge \cdots \wedge dx^{i_k}\}_{i_1 < \cdots < i_k}$ is a local frame for the vector bundle $\mathrm{Alt}^k(TM; \mathbb{R})$ restricted to $U \subset M$. Hence every k-form $\omega \in \Omega^k(U) = \Gamma(U; \mathrm{Alt}^k(TM; \mathbb{R}))$ can be written uniquely as linear combination

$$\omega = \sum_{i_1 < \cdots < i_k} f_{i_1, \dots, i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

for functions $f_{i_1, \dots, i_k} \in C^\infty(U)$. In particular, the local formula for the de Rham differential above allows us to calculate $d\omega$ for any k-form ω on $U \subset M$.

6.3 de Rham cohomology

Let M be a smooth manifold. Then the differential forms on M and the de Rham differentials relating forms of different degrees can be arranged in the following sequence of vector spaces and linear maps:

$$\Omega^0(M) \xrightarrow{d} \Omega^0(M) \xrightarrow{d} \dots \qquad \dots \xrightarrow{d} \Omega^{n-1}(M) \xrightarrow{d} \Omega^n(M).$$

This is known as the *de Rham complex* of M, where the word "complex" refers to the fact that the composition of any two consecutive maps in that sequence are zero. This innocuous property has the consequence that the kernel of the map $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$, known as the *closed k-forms* contains as a subspace the image of the map $d\colon \Omega^{k-1}(M) \to \Omega^k(M)$, known as the *exact k-forms*.

Definition 6.35. The quotient vector space

$$H^k_{\mathrm{dR}}(M) := \frac{\{\text{closed } k\text{-forms on } M\}}{\{\text{exact } k\text{-forms on } M\}} = \frac{\ker d\colon \Omega^k(M) \to \Omega^{k+1}(M)}{\operatorname{im} d\colon \Omega^{k-1}(M) \to \Omega^k(M)}$$

is called the *degree k de Rham cohomology of M*.

We note that $H^k_{\mathrm{dR}}(M) = 0$ for $k > \dim M$, since $\Omega^k(M) = 0$ for $k > \dim M$. The simplest interesting part of de Rham cohomology is $H^0_{\mathrm{dR}}(M)$:

$$H^0_{\mathrm{dR}}(M) = \ker(d\colon C^\infty(M) \to \Omega^1(M)) = \{f\colon M \to \mathbb{R} \mid df = 0\}$$

We observe that the vanishing of the differential df means that the function f is locally constant. In particular, f is constant on each path component of M, but can have different values on different path components. In particular,

$$\dim H^0_{\mathrm{dR}}(M) = \#\{\text{path components of } M\}.$$

This shows the topological nature of $H^0_{\mathrm{dR}}(M)$.

Proposition 6.36. *Let Σ_g be the surface of genus $g \geq 0$ (the connected sum of g copies of the torus $T = S^1 \times S^1$), and let $X_\ell = \mathbb{RP}^2 \# \dots \mathbb{RP}^2$ be the connected sum of $\ell \geq 1$ copies of the real projective plane $\mathbb{RP}^2$. Then the de Rham of Σ_g and X_ℓ is as follows.*

$$\dim H^k_{\mathrm{dR}}(\Sigma_g) = \begin{cases} 1 & k = 0 \\ 2g & k = 1 \\ 1 & k = 2 \end{cases} \qquad \dim H^k_{\mathrm{dR}}(X_\ell) = \begin{cases} 1 & k = 0 \\ \ell - 1 & k = 1 \\ 0 & k = 2 \end{cases}$$

We recall our calculations of the abelianized fundamental group of Σ_g resp. X_ℓ:

$$\pi_1^{\mathrm{ab}}(\Sigma_g) \cong \mathbb{Z}^{2g} \qquad \pi_1^{\mathrm{ab}}(X_\ell) \cong \mathbb{Z}^{\ell-1} \oplus \mathbb{Z}/2.$$

We observe for $M = \Sigma_g, X_\ell$ the dimension of $H^1_{\mathrm{dR}}(M)$ is equal to the $\mathbb{Z}$-summands in the abelianized fundamental group $\pi_1^{\mathrm{ab}}(M)$. This is true in generality.

Theorem 6.37. *Let M be a smooth connected manifold. Then the dimension of $H^1_{\mathrm{dR}}(M)$ is equal to the rank of $\pi_1^{\mathrm{ab}}(M)$, i.e., the number of $\mathbb{Z}$-summands in $\pi_1^{\mathrm{ab}}(M)$.*

We also recall that the manifolds Σ_g are orientable, while the manifolds X_ℓ are not. The observed relationship between orientability and the top dimensional de Rham cohomology for these 2-dimensional compact connected manifolds is again a completely general phenomenon.

Theorem 6.38. *Let M be a compact connected smooth n-manifold. Then*

$$\dim H^n_{\mathrm{dR}}(M) = \begin{cases} 1 & M \textit{ is orientable} \\ 0 & M \textit{ is not orientable} \end{cases}$$

The next theorem calculates the de Rham cohomology of some of manifolds we came across this semester.

Theorem 6.39.

$$\dim H^k_{\mathrm{dR}}(S^n) = \begin{cases} 1 & k = 0, n \\ 0 & \textit{otherwise} \end{cases}$$

$$\dim H^k_{\mathrm{dR}}(S^2 \times S^4) = \begin{cases} 1 & k = 0, 2, 4, 6 \\ 0 & \textit{otherwise} \end{cases}$$

$$\dim H^k_{\mathrm{dR}}(\mathbb{CP}^n) = \begin{cases} 1 & k = 0, 2, 4, \dots, 2n \\ 0 & \textit{otherwise} \end{cases}$$

The wedge product of forms is compatible with the de Rham differential d in the sense that d is a graded derivation (see **??**(iii)). It follows that the wedge product $\omega \wedge \eta$ of two closed forms $\omega \in \Omega^k(M)$, $\eta \in \Omega^\ell(M)$ is again closed, and that the de Rham cohomology class $[\omega \wedge \eta] \in H^{k+\ell}_{\mathrm{dR}}(M)$ depends only on $[\omega] \in H^k(M)$, $[\eta] \in H^\ell_{\mathrm{dR}}(M)$. In other words, the wedge product of forms induces a well defined product on de Rham cohomology

$$\cup\colon H^k_{\mathrm{dR}}(M) \times H^\ell_{\mathrm{dR}}(M) \longrightarrow H^{k+\ell}_{\mathrm{dR}}(M)$$

called the *cup product*. The cup product inherits the formal properties of the wedge product: it is bilinear, associative and graded commutative. It is usual and convenient to consider the direct sum

$$H^*_{\mathrm{dR}}(M) := \bigoplus_{k=0}^{\dim M} H^k_{\mathrm{dR}}(M)$$

of the de Rham cohomology spaces, and to extend the cup product by bilinearity to a product

$$\cup\colon H^*_{\mathrm{dR}}(M) \times H^*_{\mathrm{dR}}(M) \longrightarrow H^*_{\mathrm{dR}}(M).$$

This gives $H^*_{\mathrm{dR}}(M)$ the structure of a *graded associative algebra* which is graded commutative (the reader is invited to guess the definition of a graded associative algebra based on this example). The unit of this algebra is the de Rham cohomology class $[1] \in H^0_{\mathrm{dR}}(M)$ represented by the constant function $1 \in C^\infty(M) = \Omega^0(M)$.

In the case of the complex projective space $\mathbb{CP}^n$, if $x \in H^2_{\mathrm{dR}}(\mathbb{CP}^n)$ is a non-zero element, then

$$x^k := \underbrace{x \cup \cdots \cup x}_{k} \in H^{2k}_{\mathrm{dR}}(\mathbb{CP}^n) \qquad \text{is non-trivial for } 0 \le k \le n.$$

In other words, the de Rham cohomology algebra $H^*_{\mathrm{dR}}(\mathbb{CP}^n)$ is isomorphic to the quotient $\mathbb{R}[x]/(x^{n+1})$ of the polynomial ring $\mathbb{R}[x]$ by the ideal $(x^{n+1}) \subset \mathbb{R}[x]$ generated by $x^{n+1} \in \mathbb{R}[x]$.

Theorem 6.40. (Künneth Theorem for de Rham cohomology). *Let M, N be smooth manifolds. Then the map*

$$H^*_{\mathrm{dR}}(M) \otimes H^*_{\mathrm{dR}}(N) \longrightarrow H^*_{\mathrm{dR}}(M \times N) \qquad \textit{given by } [\omega] \otimes [\eta] \mapsto [(p_M^*\omega) \wedge (p_N^*\eta)]$$

is an isomorphism of graded associative algebras.

Example 6.41. Let $x \in H^n_{\mathrm{dR}}(S^n)$ be a non-zero element of this 1-dimensional vector space, and let $1 \in H^0_{\mathrm{dR}}(S^n)$ be the unit of $H^*(S^n)$ (represented by the constant function with value 1). Then

$$H^*_{\mathrm{dR}}(S^n) = \bigoplus_{k=0}^{n} H^k_{\mathrm{dR}}(S^n) = H^0_{\mathrm{dR}}(S^n) \oplus H^n_{\mathrm{dR}}(S^n) = \mathbb{R}1 \oplus \mathbb{R}x = \mathbb{R}[x]/(x^2).$$

By the Künneth Theorem it follows that

$$H^*_{\mathrm{dR}}(S^n \times S^m) = H^*_{\mathrm{dR}}(S^n) \otimes H^*_{\mathrm{dR}}(S^m) = \mathbb{R}[x]/(x^2) \otimes \mathbb{R}[y]/(y^2) = \mathbb{R}[x,y]/(x^2, y^2),$$

where the element x has degree n and y has degree m. More explicitly,

$$H^*_{\mathrm{dR}}(S^n \times S^m) = (\mathbb{R}1 \oplus \mathbb{R}x) \otimes (\mathbb{R}1 \oplus \mathbb{R}y) = \mathbb{R}1 \oplus \mathbb{R}x \oplus \mathbb{R}y \oplus \mathbb{R}xy.$$

This allows us to read off the dimension of the de Rham cohomology spaces of $S^n \times S^m$ in each degree k:

$$\dim H^k_{\mathrm{dR}}(S^n \times S^m) = \begin{cases} 1 & k = 1, n, m, n+m \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } n \neq m$$

$$\dim H^k_{\mathrm{dR}}(S^n \times S^n) = \begin{cases} 1 & k = 1, 2n \\ 2 & k = n \\ 0 & \text{otherwise} \end{cases}$$

We observe that the dimensions of the de Rham cohomology spaces of $S^2 \times S^4$ agrees with those of $\mathbb{CP}^3$, and so we can not rule out that these two 6-dimensional manifolds are diffeomorphic based on the dimensions of de Rham cohomology spaces. However, the *algebra structure* of $H^*_{\mathrm{dR}}(S^2 \times S^4)$ is different than that of $H^*(\mathbb{CP}^3)$, since the non-trivial element $x \in H^2(S^2 \times S^4)$ is represented by $p_1^*\omega$ for some $\omega \in \Omega^2(S^2)$, hence

$$x^2 = [p_1^*\omega] \cup [p_1^*\omega] = [(p_1^*\omega) \wedge (p_1^*\omega)] = [p_1^*(\omega \wedge \omega)] = 0,$$

since $\omega \wedge \omega \in \Omega^4(S^2)$ is trivial for dimensional reasons. By contrast, if $x \in H^2_{\mathrm{dR}}(\mathbb{C}^3)$ is a non-trivial element, then $x^2 \neq 0 \in H^4_{\mathrm{dR}}(\mathbb{CP}^3)$.

6.4 Integration on manifolds

A smooth n-manifold is locally diffeomorphic to open subsets V of $\mathbb{R}^n$, and we know to integrate a smooth function $f \in C^\infty(U)$. More precisely, if we assume that

$$\mathrm{supp}(f) = \text{closure of } \{x \in U \mid f(x) \neq 0\},$$

the *support of* f, is compact, the integral can be defined by

$$\int_U f := \lim \sum_i f(p_i)\,\mathrm{vol}(R_i),$$

where U is divided into suitable small regions R_i (e.g., small n-cubes) whose union covers U, p_i is some point chosen in R_i, and $\mathrm{vol}(R_i)$ denotes the volume of R_i. The limit is taken over increasingly fine subdivisions of U into regions R_i. To make sense of integration over manifolds, we need to understand how integration behaves under diffeomorphisms, namely those coming from transition maps between charts. What we need is the statement of the *Change of Variables Theorem*.

We won't be proving this theorem, but before stating it, we would like to make its statement plausible. We first consider the effect of a linear map $F\colon \mathbb{R}^n \to \mathbb{R}^n$ on the volumes of parallelepipeds. Let $\mathrm{vol}(P(v_1, \dots, v_n))$ be the volume of the parallelepiped spanned by the vectors $v_1, \dots, v_n \in \mathbb{R}^n$. We recall that

$$\mathrm{vol}(P(v_1, \dots, v_n)) = |\det(v_1, \dots, v_n)|.$$

The image of $P(v_1, \dots, v_n)$ under the linear map F is the parallelepiped $P(F(v_1), \dots, F(v_n))$, and hence

$$\mathrm{vol}(F(P(v_1, \dots, v_n))) = |\det(F(v_1), \dots, F(v_n))|.$$

We note that the matrix with column vectors $(F(v_1), \dots, F(v_n))$ is $F \cdots V$, where V is the matrix with column vectors $(v_1, \dots, v_n)$. It follows that

$$\det(F(v_1), \dots, F(v_n)) = \det(F \cdots V) = \det(F)\det(v_1, \dots, v_n),$$

and hence

$$\mathrm{vol}(F(P(v_1,\dots,v_n))) = |\det(F)|\,\mathrm{vol}(P(v_1,\dots,v_n)). \tag{6.42}$$

Now suppose that $\mathbb{R}^n \supset U \xrightarrow{F} V \subset \mathbb{R}^n$ is a diffeomorphism. Let $p \in U$, and let $P(p; v_1, \dots, v_n)$ be the *affine parallelepiped* spanned by the vectors $v_1, \dots, v_n$ viewed as starting at the point p; in other words,

$$P(p; v_1,\dots,v_n) = \{p + \sum_{i=1}^{n} x_i v_i \mid 0 \le x_i \le 1\}.$$

Then the image of $P(p; v_1, \dots, v_n)$ under F is no longer a parallelepiped, since F is no longer linear. However, near the point p the map F is well-approximated by the linear map given by its differential $dF_p \colon \mathbb{R}^n \to \mathbb{R}^n$. Hence for small $v_1, \dots, v_n$, the image of $P(p; v_1, \dots, v_n)$ under F is close to the parallelepiped $P(F(p); dF_p(v_1), \dots, dF_p(v_n))$ and hence

$$\begin{aligned}
\mathrm{vol}(F(P(p; v_1,\dots,v_n))) &\approx \mathrm{vol}(P(F(p); dF_p(v_1),\dots,dF_p(v_n))) \\
&= \mathrm{vol}(dF_p(v_1),\dots,dF_p(v_n))) \\
&= |\det(dF_p)|\,\mathrm{vol}(P(v_1,\dots,v_n))
\end{aligned}$$

Let $f \colon V \to \mathbb{R}$ be a smooth function with compact support. Let C_i, $i \in I$ be a collection of small cubes covering $U \subset \mathbb{R}^n$, and let $p_i \in C_i$ be the "lower left corner in C_i". Then the images $F(C_i)$ cover V and hence we can use this decomposition of V to approximate the integral of f over V as follows:

$$\begin{aligned}
\int_V f &\approx \sum_{i \in I} f(F(p_i))\,\mathrm{vol}(F(C_i)) \\
&\approx \sum_{i \in I} f(F(p_i))|\det(dF_{p_i})|\,\mathrm{vol}(C_i) \\
&\approx \int_U f(F(p))|\det(dF_p)|
\end{aligned}$$

Taking the limit as the size of the cubes approaches zero, these approximations are become better and hence we conclude:

Theorem 6.43. (Change of Variables Theorem). *Let $\mathbb{R}^n \supset U \xrightarrow{F} V \subset \mathbb{R}^n$ be a diffeomorphism and let $f \colon V \to \mathbb{R}$ be a function with compact support. Then*

$$\int_V f = \int_U F^*(f)|\det(dF)|.$$