SML (Topology and Homotopy), Summary Notes

Chris Bourne[1]

Spring Semester, 2026

The following are work-in-progress notes that summarise the lectures for the 'Special Mathematics Lecture (Topology and Homotopy)'. The content is mostly taken from the texts [4, 6]. These notes should not be considered as complete lecture notes as many details are still missing. Please also be wary of typos and other mistakes. Any corrections or suggestions are greatly appreciated ☺.

Contents

---

[1] cbourne@nagoya-u.jp

1 Preface

We generally first learn calculus by considering functions on Euclidean space, $f : \mathbb{R} \to \mathbb{R}$ (or $f : \mathbb{R}^n \to \mathbb{R}^d$ more generally). The vector space $\mathbb{R}^n$ is certainly a nice place to do analysis, but for many shapes we regularly encounter (spheres, donuts, ...), the standard calculus theory can not be applied without some extra work. Furthermore, on the very small scale (quantum mechanics) or the very large scale (black holes), we encounter spaces and objects for which we do not have a good geometric understanding or intuition. Nevertheless, we would still like to be able to talk about functions and continuity on these more complicated spaces. Roughly speaking, topology is the mathematical theory that allows us to discuss what it means to be continuous in the most general setting.

Not all topological spaces can be geometrically visualised, but let us assume that we are given two shapes X and Y that can be nicely embedded into Euclidean space. Topology only cares about whether functions are continuous and is not concerned with derivatives or smoothness. So we say that X and Y are the same topologically if we can continuously deform one space to another. This means that deformation may add corners and non-smooth points, but we can not tear or make a hole in a shape. This means that a sphere and a doughnut are 'topologically distinct' as a sphere has no holes whereas a has a hole. On the other hand, a coffee cup has a hole and a refrigerator has no holes. So a sphere is topologically the same as a refrigerator and a doughnut is topologically the same as a coffee cup.

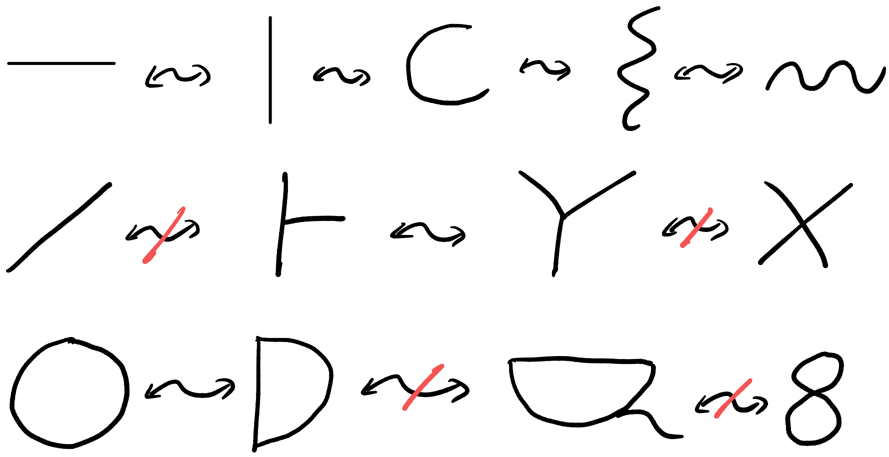
Of course, we need to make these notions mathematically precise. Our aim is to give mathematical meaning to the notion of continuity in the broadest possible sense. This will allow us to then say what if means for different shapes and spaces to be equivalent. Of particularly interest are the properties of these abstract spaces that are preserved under topological equivalence. We call these properties *topological invariants*.

A topological invariant is not just a nice mathematically property, but increasingly has found applications in physics and elsewhere. Changing and deforming a shape does not change its topological invariants. So if a physical property or phenomena can be described by these invariants, then this property must be very stable and robust. Indeed, even if we deform or perturb the system, so long as we do not break continuity via a tear or cut, the topological invariant and therefore physical property will stay the same.

Warm up – topological classification of capital letters

As a simple warm up exercise, let us consider (bounded) lines and curves, which we consider as one-dimensional shapes embedded in $\mathbb{R}^2$. Recall that two shapes are the same if we can deform one into another. In this case, this means that we can't make any cuts of our lines.

Let's first consider a single line. If the line never touches itself, then it can be deformed into any other line that does not touch itself. On the other hand, if it becomes a loop (e.g. a circle), then we cannot deform this shape into a straight line without making a cut (which is not allowed).



More generally, we could consider multiple lines which are joined at a vertex/branching points. Let's consider the example of capital letters. The capital letter I can be continuously deformed into the letter L or N, but it can not be deformed into the letter Y, which has a 3-vertex. Similarly, the letter Y can not be deformed into the letter X, which has a 4-vertex. We can continue this process for all the letters of the alphabet to arrive at a 'topological classification' of capital letters.

No holes or branching points

$$\Big\{\mathsf{C,\ G,\ I,\ J,\ L,\ M,\ N,\ S,\ U,\ V,\ W,\ Z}\Big\}.$$

One hole, no branching points

$$\Big\{\mathsf{D,\ O}\Big\}.$$

No holes, one 3-vertex

$$\Big\{\mathsf{E,\ F,\ T,\ Y}\Big\}.$$

One hole, one 3-vertex

$$\Big\{\mathsf{P}\Big\}.$$

No holes, two 3-vertices

$$\Big\{\mathsf{H,\ K}\Big\}.$$

One hole, two 3-vertices

$$\Big\{\mathsf{A,\ R}\Big\}.$$

Two holes, two 3-vertices

$$\Big\{\mathsf{B}\Big\}.$$

No holes, one 4-vertex

$$\Big\{\mathsf{X}\Big\}.$$

One hole, one 4-vertex

$$\Big\{\mathsf{Q}\Big\}.$$

So we have sorted the 26 letters into 9 topological classes. Of course, we have not been precise about how we were able to arrive at this conclusion. In what follows we develop the theory to make the above result mathematically rigorous.

2 Metric spaces

Before we consider more abstract topological spaces, we will first consider the special case of metric spaces, which roughly are sets with with an appropriate notion of distance. This structure is sufficient to import many of the ideas and constructions that we have seen in Calculus on $\mathbb{R}^n$ or $\mathbb{C}$.

2.1 Basic definitions and examples

We start with a motivating example.

Example 2.1.1 (Norms on $\mathbb{R}^n$)**.** We first recall the vector space

$$\mathbb{R}^n = \big\{\mathbf{x} = (x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\big\}.$$

Elements $\mathbf{x} \in \mathbb{R}^n$ are vectors for which we can define a *norm*, a notion of the size of a vector. Actually, we have many choices of norms for $\mathbb{R}^n$, e.g.

$$\|\mathbf{x}\|_1 = |x_1| + \ldots + |x_n|, \qquad \|\mathbf{x}\|_2 = \big(x_1^2 + \ldots + x_n^2\big)^{1/2}$$
$$\|\mathbf{x}\|_p = \big(|x_1|^p + \ldots + |x_n|^p\big)^{1/p}, \qquad \|\mathbf{x}\|_\infty = \max_{j=1,\ldots,n} |x_j|.$$

For $\|\cdot\|$ to be a norm, it must satisfy the following three properties:

1. $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$,
2. $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

The norm also allows us to consider the quantity $\|\mathbf{x} - \mathbf{y}\|$ as measuring the distance from $\mathbf{x}$ to $\mathbf{y}$ *with respect to our choice of norm*. Indeed,

$$\|(1,0) - (0,1)\|_1 = 2, \quad \|(1,0) - (0,1)\|_2 = \sqrt{2}, \quad \|(1,0) - (0,1)\|_p = 2^{1/p}, \quad \|(1,0) - (0,1)\|_\infty = 1$$

and the distance between points changes depending on how me measure size.

Exercise 2.1. Suppose that $\|\cdot\|$ is a norm on $\mathbb{R}^n$ Define the function $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Show that

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,
2. $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$,
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,
4. $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.

The above discussion can be generalised to norms on and distances on more general vector spaces V. On the other hand, many spaces we care about are not vector spaces, e.g. $X = (0,1) \subset \mathbb{R}$, the open unit interval, but still have come with a reasonable notion of distance. So we would like to understand the distance between points in sets that may not be closed under addition / scalar multiplication (or where these operations may not even be defined).

Definition 2.1.2 (Metric space)**.** A metric space (X, d) is a set X with a *metric*, a function $d : X \times X \to \mathbb{R}$ that satisfies the following properties

1. (Positivity) $d(x, y) \geq 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$,
2. (Symmetry) $d(x, y) = d(y, x)$ for all $x, y \in X$,
3. (Triangle inequality) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

Exercise 2.1 shows that any norm on $\mathbb{R}^n$ also makes $(\mathbb{R}^n, d)$ a metric space. This is also true for more general vector spaces with a norm. We will sometimes be sloppy and simply state that X is a metric space, where the choice of metric d is implicit.

Example 2.1.3 (Discrete metric)**.** Let X be any set. Then we can define the following discrete metric

$$d_{\text{disc}}(x, y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

By construction $d_{\text{disc}}(x, y) \geq 0$ and $d_{\text{disc}}(x, y) = 0$ only when $x = y$. The symmetry of d_{disc} is also easy to see. Considering the triangle inequality with $x, y, z \in X$, if $x = z$, then the $d_{\text{disc}}(x, z) = 0$ and the relation holds as $d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$. If $x \neq z$ and $x = y$ (or $y = z$), then we have

$$d_{\text{disc}}(x, z) = 1 \leq 1 = d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z).$$

If all are distinct, $x \neq y \neq z$, then $1 \leq 2$ gives the result. Finally if $x = y = z$, then $0 \leq 0$ and the triangle inequality holds. Hence d_{disc} is a metric.

Exercise 2.2. Let X be a set and let

$$B(X) = \big\{ f : X \to \mathbb{R} \mid \text{there exists some } C > 0 \text{ such that } |f(x)| < C \text{ for all } x \in X \big\}.$$

Show that $B(X)$ is a metric space when endowed with the metric

$$d(f, g) = \sup_{x \in X} \big| f(x) - g(x) \big|.$$

Example 2.1.4. Suppose that (X, d) is a metric space and $Y \subset X$ is a non-empty subset. Then (Y, d) is a metric space, where d is the restriction of the metric on X to Y. We say that (Y, d) is a (metric) subspace of (X, d).

Example 2.1.5. If $X = \mathbb{R}^2$, the function $\tilde{d}\big((x_1, x_2), (y_1, y_2)\big) = |x_1 - y_2|$ does **not** define a metric space. Indeed $\tilde{d}\big((0, 1), (1, 1)\big) = 0$, but $(0, 1) \neq (1, 1)$.

The set X in a metric space (X, d) might not have much structure or may be difficult to visualise. On the other hand, we know many things about $\mathbb{R}$ and can use this knowledge along with the metric d to infer properties of the space X.

Definition 2.1.6 (Open ball and open set). Let (X, d) be a metric space.

1. Given $x \in X$ and $r > 0$, the open ball centered at x with radius r is the set
$$B(x; r) = \{y \in X \mid d(x, y) < r\}.$$
2. We say that a subset $U \subset X$ is open if for all $x \in U$ there is some $\epsilon > 0$ such that $B(x; \epsilon) \subset U$.

Lemma 2.1.7. *Any open ball $B(x; r)$ is open.*

Proof. Take $y \in B(x; r)$, and $\epsilon = r - d(x, y) > 0$. If $z \in B(y, \epsilon)$, then
$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (r - d(x, y)) = r.$$
Hence $z \in B(x; r)$ and so $B(y; \epsilon) \subset B(x; r)$. $\square$

For a more general subset $Y \subset X$, we define
$$\operatorname{Int}(Y) = \{y \in Y \mid \text{there exists } \epsilon > 0 \text{ such that } B(y, \epsilon) \subset Y\},$$
where by construction Y is open if and only if $Y = \operatorname{Int}(Y)$.

Example 2.1.8. Let X be a metric space with the discrete metric. We claim that for any $x \in X$, the singleton set $\{x\}$ is open. Indeed, for any $0 < r < 1$, $B(x; r) = \{x\}$.

Proposition 2.1.9. *Let (X, d) be a metric space.*

1. *The union of a family of open subsets $\{U_\alpha\}_{\alpha \in I}$ is open.*
2. *The intersection of a finite family of open subsets $\{U_j\}_{j=1}^n$ is open.*

Proof. We take $x \in \bigcup_\alpha U_\alpha$. Then by assumption $x \in U_{\alpha'}$ for some $\alpha' \in I$. Because $U_{\alpha'}$ is open, there is some $\epsilon > 0$ such that $B(x; \epsilon) \subset U_{\alpha'}$. By the definition of the union, this also means that $B(x; \epsilon) \subset \bigcup_\alpha U_\alpha$. Hence the union is open.
We leave part (2) as an exercise. $\square$

Example 2.1.10. If (X, d) is a metric space, the empty set $\emptyset = \{\} \subset X$ is tautologically open as it contains no elements. The total space X is open as $B(x; r) \subset X$ for any $x \in X$.

Example 2.1.11. Take $\mathbb{R}^2$ as a metric space with the Euclidean metric (using the norm $\|\cdot\|_2$) The following subsets are **not** open. (Why?)

1. The y-axis $\{(0, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\} \subset \mathbb{R}^2$, with Euclidean metric $\|\cdot\|_2$.
2. The closed unit disc $\{\mathbf{x} \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\}$,
3. $\{(x, y) \mid xy \geq 0\}$.

Definition 2.1.12 (Closure and closed sets)**.** Let (X,d) be a metric space and $Y \subset X$.

1. The closure of Y is the set

$$\overline{Y} = \big\{x \in X \mid B(x;r) \cap Y \neq \emptyset \text{ for all } r > 0\big\}.$$

2. We say that Y is closed if $\overline{Y} = Y$.

Note that $Y \subset \overline{Y}$. We think of the closure of a set $Y \subset X$ as the set Y along with all of its boundary points in X.

Example 2.1.13. Take $\mathbb{R}^2$ with $\|\cdot\|_2$ and $B(\mathbf{0};1)$. Then $\overline{B(\mathbf{0};1)}$ is the closed unit disc $\big\{\mathbf{x} \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\big\}$ (exercise).

Caution: It does not always hold that $\overline{B(x;r)} = \{y \mid d(x,y) \leq r\}$. For example, take d_{disc}, then

$$\overline{B(x;\tfrac{1}{2})} = \{x\} = B(x;\tfrac{1}{2}).$$

Exercise 2.3. Show that the closure of a set $Y \subset X$ is closed, i.e. $\overline{\overline{Y}} = \overline{Y}$.

Example 2.1.14. The empty set $\emptyset$ and X are closed sets.

If a set is not open, then it is not necessarily closed (and vice verca). For example $[a,b) \subset \mathbb{R}$ is neither closed nor open. On the other hand, sets can be closed *and* open. The singleton $\{x\} \subset X$ is closed and open when X has the discrete metric.

To further relate open and closed sets, we first review some basic set theory.

Lemma 2.1.15 (De Morgan's laws for sets)**.** *Let A and B be subsets of X with $A^c = X \setminus A$ and $B^c = X \setminus B$. Then*

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c.$$

More generally, if $\{A_\alpha\}_{\alpha \in I}$ are a family of subsets of X, then

$$\Big(\bigcup_{\alpha \in I} A_\alpha\Big)^c = \bigcap_{\alpha \in I} (A_\alpha)^c, \qquad \Big(\bigcap_{\alpha \in I} A_\alpha\Big)^c = \bigcup_{\alpha \in I} (A_\alpha)^c.$$

Theorem 2.1.16. *Let (X,d) be a metric space. A subset $Y \subset X$ is closed if and only if $Y^c = X \setminus Y$ is open.*

Proof. Suppose $Y \subset X$ is closed and take $x \in Y^c$. Because $Y = \overline{Y}$, it follows that there is some $r > 0$ such that $B(x;r) \cap Y = \emptyset$. We then see that $B(x;r) \subset Y^c$ and so Y^c is open.

If Y^c is open, then for any $x \in Y^c$ there is some $r > 0$ such that $B(x;r) \subset Y^c$, which in particular implies that $B(x;r) \cap Y^c = \emptyset$. Therefore $x \notin \overline{Y}$ and we have shown that $Y^c \subset \overline{Y}^c$. But $Y \subset \overline{Y}$ implies that $\overline{Y}^c \subset Y^c$. So it follows that $Y^c = \overline{Y}^c$ and therefore $Y = \overline{Y}$. $\square$

Proposition 2.1.17. *Let (X, d) be a metric space.*

1. *The union of a finite family of closed subsets $\{C_j\}_{j=1}^n$ is closed.*
2. *The intersection of a family of closed subsets $\{C_\alpha\}_{\alpha \in I}$ is closed.*

Proof. Using Theorem 2.1.16 and Lemma 2.1.15, this becomes a restatement of Proposition 2.1.9. □

2.2 Continuity

We have seen that a metric lets us talk about open balls and sets. This gives us the local structure needed to discuss neighbourhoods and continuity of functions between different metric spaces $f : (X, d_Y) \to (Y, d_Y)$.

Definition 2.2.1. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f : X \to Y$ is continuous at $x \in X$ if for all $\epsilon > 0$ there is some $\delta > 0$ such that for all $z \in X$ with $d_X(x, z) < \delta$, we have that $d_Y\big(f(x), f(z)\big) < \epsilon$. We call f continuous if it is continuous at x for all $x \in X$.

Written in terms of open balls, f is continuous at x if for all $\epsilon > 0$, there is some $\delta > 0$ such that $z \in B(x; \delta)$ implies that $f(z) \in B(f(x); \epsilon)$.

Examples 2.2.2.

1. If $X = Y = \mathbb{R}$, then the condition for continuity can be written as $|x - z| < \delta$ implies $|f(x) - f(z)| < \epsilon$, which is the standard ϵ-δ definition one finds in real analysis.
2. If X has the discrete metric, then *any* function $f : X \to Y$ is continuous. If $\epsilon > 0$, then for any $\delta \in (0, 1)$, $B(x; \delta) = \{x\}$ and $d_Y(f(x), f(x)) = 0 < \epsilon$.

In this section we will show that continuity can also purely in terms of open sets. This will help motivate our more abstract definition of continuity in topological spaces.

Theorem 2.2.3. *A function $f : (X, d_X) \to (Y, d_Y)$ is continuous if and only if $f^{-1}(V) = \{x \in X \mid f(x) \in V\} \subset X$ is an open set for every open subset $V \subset Y$.*

Proof. Suppose f is continuous and take $V \subset Y$ open with $x \in f^{-1}(V)$. We need to show that $B(x; r) \subset f^{-1}(V)$ for some r. By assumption $f(x) \in V$, which is open so there is some ϵ such that $B(f(x); \epsilon) \subset V$. We also know that f is continuous so there is some δ such that $f(z) \in B(f(x); \epsilon) \subset V$ for any $z \in B(x; \delta)$. This shows that $B(x; \delta) \subset f^{-1}(V)$ and so $f^{-1}(V)$ is open.

Now syppose that $f^{-1}(V)$ is open for any open $V \subset Y$. Then for any $x \in X$ and $\epsilon > 0$, $B(f(x); \epsilon) \subset Y$ is open and therefore $f^{-1}\big(B(f(x); \epsilon)\big) \subset X$ is an open neighbourhood of x. Hence there is some $\delta > 0$ such that $B(x; \delta) \subset f^{-1}\big(B(f(x); \epsilon)\big)$. It then follows that any $z \in B(x; \delta)$ is such that $f(z) \in B(f(x); \epsilon)$, which implies that f is continuous at x. Because x was arbitrary, f is continous. □

Definition 2.2.4. A function $f : (X, d_X) \to (Y, d_Y)$ is *uniformly* continuous if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for any $x, z \in X$ with $d_X(x, z) < \delta$, it holds that $d_Y\big(f(x), f(z)\big) < \epsilon$.

We see that, unlike a continuous function, the constant $\delta > 0$ is independent of the point $x \in X$.

Example 2.2.5. The function $f : (0,1) \to \mathbb{R}$, $f(x) = x^{-1}$ is continuous but not uniformly continuous as it blows up when $x \to 0^+$. Indeed, take $\epsilon = 1$ and $\delta > 0$. We take $x = \min\{\delta, \frac{1}{2}\}$ and $z = \frac{x}{2}$. Then $d(x,z) = \frac{x}{2} < \delta$ but

$$\left|x^{-1} - z^{-1}\right| = \left|\frac{1}{x} - \frac{2}{x}\right| = \frac{1}{x} > 1.$$

Exercise 2.4. Let (X,d) be a metric space and fix $x_0 \in X$. Show that the function $f : X \to \mathbb{R}$, $f(x) = d(x, x_0)$ is uniformly continuous.

2.3 Sequences and completeness

In real analysis we can understand the closure and continuity of spaces uses sequences in $\mathbb{R}^n$. Many of these ideas can be generalised to metric spaces.

Definition 2.3.1. We say that a sequence $\{x_n\}_{n\geq 0} \subset X$ converges to $x \in X$ if for all $\epsilon > 0$ there is some $N > 0$ such that $d(x_n, x) < \epsilon$ for all $n > N$. In this case we write $x_n \to x$ in X.

Exercise 2.5. Show that $f : (X, d_X) \to (Y, d_Y)$ is continuous at $x \in X$ if and only if $x_n \to x$ in X implies that $f(x_n) \to f(x)$ in Y.

Proposition 2.3.2. *Let (X,d) be a metric space and $Y \subset X$. Then $x \in \overline{Y}$ if and only if there is a sequence $\{y_n\}_{n\geq 0} \subset Y$ such that $y_n \to x$.*

Proof. If $y_n \to x$, then for any $r > 0$, the ball $B(x;r)$ will contain elements y_n for n sufficiently large. Hence $x \in \overline{Y}$. Conversely, if $x \in \overline{Y}$, then for every $n \geq 1$, there is at least one point $y_n' \in B(x; \frac{1}{n}) \cap Y$. We use these points to build a sequence $\{y_n'\} \subset Y$ where $d(x, y_n') < \frac{1}{n} \to 0$ as $n \to \infty$. $\square$

Definition 2.3.3.
1. We say that a sequence $\{x_n\}_{n\geq 0} \subset (X,d)$ is Cauchy if for all $\epsilon > 0$ there exists $N > 0$ such that $d(x_n, x_m) < \epsilon$ for all $n, m > N$.
2. A metric space (X,d) is complete if every Cauchy sequence converges in X.

Put another way, $\{x_n\}$ is Cauchy if the double limit

$$\lim_{n,m\to\infty} d(x_n, x_m) = 0.$$

Examples 2.3.4. The real line $\mathbb{R}$ is complete.

Higher dimensional real and complex spaces $\mathbb{R}^n$ and $\mathbb{C}^m$ with Euclidean norm $\|\cdot\|_2$ are complete.

The set of rational numbers $\mathbb{Q}$ with metric induced from $\mathbb{R}$ is not complete as we can take a sequence $\{x_n\} \subset \mathbb{Q}$ such that $x_n \to \pi$, for example.

If X is a set (X, d_{disc}) is a complete metric space (exercise).

The open interval $(0,1)$ is not complete. Indeed $x_n = \frac{1}{n}$ is a Cauchy sequence in $(0,1)$, but the limit $x_n \to 0$ and $0 \notin (0,1)$.

The proof of the following is an exercise.

Proposition 2.3.5. *Any closed subspace of a complete metric space is complete.*

Definition 2.3.6. Let (X,d) be a metric space and $Y \subset X$ a subspace. We say that Y is dense if *dense* if $\overline{Y} = X$. We say that X is separable if there is a dense subset Y and an injective function $g : Y \to \mathbb{N}$ (Y is countable).

Example 2.3.7. The rational numbers $\mathbb{Q}$ is a dense and countable subset of $\mathbb{R}$. We will prove neither of these facts, though density follows from the property that if $a < b \in \mathbb{R}$, then there exists some $q \in \mathbb{Q}$ such that $a < q < b$.

The following result is a key element used to prove other important results in functional analysis (uniform boundedness principle, open mapping theorem, ...).

Theorem 2.3.8 (Baire Category Theorem)**.** *Let $\{U_n\}_{n\geq 1}$ be a (countable) sequence of open dense subsets of a complete metric space X. Then the intersection $\bigcap_{n\geq 1} U_n$ is also dense in X.*

Proof. We take $x \in X$ and $\epsilon > 0$. The result will follow if we can find some $y \in B(x;\epsilon) \cap \bigcap_n U_n$. We will find this y via a limit of elements $y_n \in U_n$. Indeed, U_1 is dense, so there is some $y_1 \in U_1 \cap B(x;\epsilon)$. Furthermore, because it is open, we know that $B(y_1, r_1) \subset U_1$ for some r_1. Shrinking r_1 if necessary, we can assume $r_1 < 1$ and $\overline{B(y_1;r_1)} \subset U_1 \cap B(x;\epsilon)$. Now looking at $y_1 \in U_1 \subset X$, by density of U_2, there is some $y_2 \in U_2 \cap B(y_1;r_1)$. By the same argument we can also take $r_2 < 1/2$ such that $\overline{B(y_2;r_2)} \subset U_2 \cap B(y_1;r_1)$. We can continue to iterate this argument to obtain a sequence $\{y_n\}_{n\geq 1}$ in X and a sequence $\{r_n\}_{n\geq 1}$ in $(0,\infty)$ such that $r_n < 1/n$ and

$$\begin{aligned} &\overline{B(y_n;r_n)} \subset U_{n-1} \cap B(y_{n-1};r_{n-1}), \\ &\overline{B(y_n;r_n)} \subset B(y_{n-1};r_{n-1}) \subset \ldots \subset B(y_1;r_1) \subset B(x;\epsilon). \end{aligned}$$

We see that $y_m \in B(y_n, r_n)$ for all $m > n$ and furthermore

$$d(y_n, y_m) < r_n < \frac{1}{n} \xrightarrow{n,m\to\infty} 0$$

and so $\{y_n\}$ is a Cauchy sequence. Because X is complete, there is some $y \in X$ with $y_n \to y$. By the construction of the sequence, it follows that $y \in \overline{B(y_n;r_n)} \subset U_n \cap B(y_{n-1}, r_{n-1})$ for all $n \geq 1$. So $y \in \bigcap_n U_n$ and furthermore $y \in \overline{B(y_n;r_n)} \subset B(x;\epsilon)$, which finishes the proof. $\square$

Related to dense subspaces are *nowhere dense* subspaces, where Y is nowhere dense if $\mathrm{Int}(\overline{Y}) = \emptyset$. A subset $Y \subset X$ is nowhere dense if $X \setminus \overline{Y}$ is a dense open subset of X. The integers $\mathbb{Z}$ is a nowhere dense subset of $\mathbb{R}$.

Example 2.3.9 (The Cantor set)**.** The Cantor set is an example of an uncountably infinite but nowhere dense set. We start with the closed unit inteval $C_0 = [0,1] \subset \mathbb{R}$. We now remove the middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$ to obtain $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$, a union of two intervals. We then remove the middle third of each part,

$$C_2 = \left[0, \tfrac{1}{9}\right] \cup \left[\tfrac{2}{9}, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, \tfrac{7}{9}\right] \cup \left[\tfrac{7}{9}, 1\right].$$

We can continue to remove the middle third from the remaining line segments and take the limit to infinity. At the step $n \geq 1$, we there are $3^n - 1$ line segments from which we remove the middle third. The limiting set as $n \to \infty$ is the cantor set C, where explicitly

$$C = [0, 1] \setminus \bigcup_{n=0}^{\infty} \bigcup_{k=0}^{3^n - 1} \left(\frac{3k+1}{3^{n+1}}, \frac{3k+2}{3^{n+1}} \right),$$

where $\left(\frac{3k+1}{3^{n+1}}, \frac{3k+2}{3^{n+1}}\right)$ is removed from $\left[\frac{k}{3^n}, \frac{k+1}{3^n}\right]$. We see that $C = [0,1] \setminus U$ with U open. Hence C is closed and therefore complete (Proposition 2.3.5).

Let us now consider $\mathrm{Int}(C) = \mathrm{Int}(\overline{C})$. The open balls of $\mathbb{R}$ are open intervals, $B(x; r) = (x-r, x+r)$ and $c \in \mathrm{Int}(C)$ if there exists some $\epsilon > 0$ such that $(c - \epsilon, c + \epsilon) \subset C$. However, for *any* $\epsilon > 0$ and $c \in C$, $(c-\epsilon, c+\epsilon)$ will intersect an interval which has previously been removed in the construction of C. Therefore

$$\mathrm{Int}(C) = \left\{ c \in C \mid \text{ there exists } \epsilon > 0 \text{ such that } B(c; \epsilon) \subset Y \right\} = \emptyset.$$

and C is nowhere dense.

2.4 Compact metric spaces

The real line $\mathbb{R}$ is uncountably infinite and so is any interval $[a, b) \subset \mathbb{R}$. On the other hand, intuitively there is a qualitative difference between the 'uncountable but bounded set' $[a, b]$ and the 'uncountable and unbounded set' $\mathbb{R}$. Here we take a *closed* interval as the map $(0,1) \ni x \mapsto x^{-1} \in (1, \infty)$ is a bijection. So open intervals can be stretched to infinitely long intervals. But (bounded) closed intervals can not.

The notion of compactness makes precise the notion that the real line $\mathbb{R}$ and the bounded closed interval $[a, b]$ have distinct properties, where the latter is much easier to control. Compactness will also appear in our studies of topological spaces and we give a definition that is amenable to this generalisation.

Definition 2.4.1. Let (X, d) be a metric space.

1. An open cover of X is a family of open sets $\{U_\alpha\}_{\alpha \in I}$ such that $\bigcup_{\alpha \in I} U_\alpha = X$.
2. We say an open cover $\{U_\alpha\}_{\alpha \in I}$ has a finite subcover if there is a finite sub-family $\{U_{\alpha_1}, \ldots, U_{\alpha_n}\}$ such that $X = \bigcup_{i=1}^{n} U_{\alpha_i}$.
3. We say that (X, d) is compact if every open cover of X has a finite subcover.

Note that for X to be compact, we are not saying that a finite cover exists (which can often be constructed). Rather, given *any* cover of X (including those with uncountably many elements), we can throw away all but finitely many elements and still cover the entire space.

Example 2.4.2. Let $X = \mathbb{R}$. We can cover $\mathbb{R}$ by open intervals centered around each integer,

$$\mathbb{R} = \bigcup_{n \in \mathbb{Z}} (n - 1, n + 1).$$

On the other hand. We are not able to cover all of $\mathbb{R}$ by taking a finite number of the intervals

$U_n = (n-1, n+1)$. Indeed, if we take finitely many U_ns, then there will be some $n_{\max}$ and $(n_{\max}+1, n_{\max}+2) \cap \bigcup_n U_n = \emptyset$. So $\mathbb{R}$ is not compact.

Example 2.4.3. The metric space (X, d_{disc}) is compact if and only if X is finite. This follows because the family of singletons $\big\{\{x\}\big\}_{x\in X}$ is an open cover of X.

Lemma 2.4.4. *The closed unit interval* $[0,1] \subset \mathbb{R}$ *is compact.*

Proof. Take an open cover $\{U_\alpha\}_{\alpha\in I}$ of $[0,1]$. We will assume that there is no finite subcover and derive a contradiction. We can restrict the open cover to $\left[0, \frac{1}{2}\right]$ and $\left[\frac{1}{2}, 1\right]$, which will give an open cover of each sub-interval. If both covers of these sub-intervals has a finite subcover, then we can combine them to obtain a finite subcover of $[0,1]$. So we assume that at least one sub-intervals I_1 is such that the open cover has no finite subcover. We can now divide I_1 into two sub-intervals (e.g. if $I_1 = \left[0, \frac{1}{2}\right]$, then we take $\left[0, \frac{1}{4}\right]$ and $\left[\frac{1}{4}, \frac{1}{2}\right]$). Arguing as before, one of these sub-intervals I_2 must not have finite subcover. Continuing this pattern, we get a nested sequence

$$[0,1] \supset I_1 \supset I_2 \supset \cdots$$

where I_n is an interval of length 2^{-n}. By assumption, no interval I_n admits a finite subcover. We consider the intersection $\bigcap_n I_n$, which will consist of a single point $x \in [0,1]$. Because $\{U_\alpha\}_{\alpha\in I}$ is a cover of $[0,1]$, $x \in U_{\alpha'}$ for some $\alpha' \in I$. Because $U_{\alpha'}$ is open, $(x-\delta, x+\delta) \cap [0,1] \subset U_{\alpha'}$ for some $\delta > 0$. Now for n sufficiently large $I_n \subset (x-\delta, x+\delta) \cap [0,1] \subset U_{\alpha'}$. But this implies that I_n is wholly contained in $U_{\alpha'}$ and so admits a finite subcover (the single element $U_{\alpha'}$). So we derive a contradiction, which implies our assumption of $\{U_\alpha\}_{\alpha\in I}$ having no finite subcover was false. Thus $[0,1]$ is compact. $\square$

Definition 2.4.5. Let (X,d) be a metric space.

1. We say (X,d) is bounded if there is some $B > 0$ such that $d(x,y) < B$ for all $x, y \in X$.
2. We say that (X,d) is totally bounded if for all $\epsilon > 0$ there are finitely many points $x_1, \ldots, x_n \in X$ such that $\{B(x_j;\epsilon)\}_{j=1}^n$ is an open cover of X.

Example 2.4.6. Any closed interval $[a,b]$ is bounded and totally bounded. To see this, for any $\epsilon > 0$ we can sub-divide the interval into $\lceil \frac{b-a}{\epsilon} \rceil$ parts (rounding up to the next integer) and take ϵ-balls to then cover the space.

Theorem 2.4.7. *Let (X,d) be a metric space. Then the following are equivalent.*

1. *X is compact.*
2. *Every sequence in X has a convergent subsequence.*
3. *X is totally bounded and complete.*

We will only prove (1) $\Rightarrow$ (2) $\Rightarrow$ (3). See [4, Section 1.5] for the full proof.

Proof. (1) $\Rightarrow$ (2). Suppose X is compact and let $\{y_n\}_{n\geq 1}$ be a sequence in X. We claim that there is some $x \in X$ such that for all $\epsilon > 0$, $B(x;\epsilon)$ contains infinitely many terms of the sequence $\{y_j\}$. We prove this claim by contradiction: so suppose that for each $x \in X$ there is some $\epsilon_x > 0$ such that $B(x;\epsilon_x)$ contains finitely many elements of the sequence $\{y_j\}$. We can use these balls to cover X, which must have a finite subcover,

$$X = \bigcup_{x\in X} B(x;\epsilon_x) = \bigcup_{j=1}^{m} B(x_j,\epsilon_j).$$

However, we have assumed that each $B(x_j,\epsilon_j)$ has only finitely many elements of the sequence $\{y_j\}$. This gives a contradiction and proves the claim.

We now construct a convergent subsequence for $\{y_j\}$. Taking $x \in X$, there are infinitely many elements of the sequence $\{y_j\}$ in $B(x,1)$, so we take some $y_{j_1} \in B(x;1)$. There are similarly infinitely many elements in $B(x,\frac{1}{2})$ and we take $y_{j_2} \in B(x;\frac{1}{2})$ with $j_2 > j_1$. We can continue this process and take $y_{j_n} \in B(x;\frac{1}{n})$ with $j_n > j_{n-1}$. By construction $\{y_{j_n}\}_{n=1}^{\infty}$ is a subsequence of $\{y_j\}$ with

$$d(x,y_{j_n}) < \frac{1}{n} \xrightarrow{n\to\infty} 0.$$

Hence the subsequence converges to $x \in X$.

(2) $\Rightarrow$ (3). We first show X is complete so suppose that $\{x_n\}$ is a Cauchy sequence in X. By assumption there is a convergent subsequence $x_{n_j} \to x$ as $j \to \infty$. For $\epsilon > 0$, we can take N sufficiently large so that $d(x_n,x_m) < \frac{\epsilon}{2}$ and $d(x_{n_j},x) < \frac{\epsilon}{2}$ for $n,m,n_j > N$. Then

$$d(x_n,x) \leq d(x_n,x_{n_j}) + d(x_{n_j},x) < \epsilon$$

which shows that $x_n \to x$ and so X is complete.

Taking $\epsilon > 0$, we wish to cover X by open balls of radius ϵ. Take some point $y_1 \in X$. If $B(y_1,\epsilon) = X$ we are done so assume $B(y_1,\epsilon)^c \neq \emptyset$ and take $y_2 \in B(y_1,\epsilon)^c$. If $B(y_1;\epsilon) \cup B(y_2,\epsilon) = X$, we stop otherwise take $y_3 \in \big(B(y_1;\epsilon) \cup B(y_2,\epsilon)\big)^c$. Continuing this procedure, we have a sequence $\{y_j\}_{j\geq 1}$ such that

$$y_n \in X \setminus \Big(\bigcup_{j=1}^{n-1} B(y_j,\epsilon)\Big), \qquad d(y_j,y_k) \geq \epsilon,\ j \neq k.$$

If the construction does not terminate, then we obtain an infinite sequence $\{y_j\}_{j=1}^{\infty}$ with each distinct element of distance $\geq \epsilon$ from each other. So $\{y_j\}$ can not have a convergent subsequence, a contradiction. Hence the construction of $y_1, y_2, \ldots$ must terminate after finite steps and so X is totally bounded. $\square$

We will particularly consider compactness in the context of $\mathbb{R}^n$ (with standard Euclidean metric).

Lemma 2.4.8 ([4, Lemma 5.4]). *A subset $Y \in \mathbb{R}^n$ is bounded if and only if it is totally bounded.*

Theorem 2.4.9 (Heine–Borel Theorem). *Let Y be a subspace of $\mathbb{R}^n$ (with standard metric). The following are equivalent.*

1. *Y is compact.*
2. *Every sequence in Y has a convergent subsequence.*
3. *Y is closed and bounded.*

Proof (Sketch). If E is closed and bounded, it is complete and totally bounded. So the result is a special case of Theorem 2.4.7. We will show (3) $\Rightarrow$ (1) in the case of $\mathbb{R}$ (with the idea in $\mathbb{R}^n$ being similar).

Suppose that $Y \subset \mathbb{R}$ is closed and bounded. Because Y is bounded, there are $a < b \in \mathbb{R}$ such that $a < y < b$ for all $y \in Y$. Also, because Y is closed $Y^c = \mathbb{R} \setminus Y$ is open. So if we take an open cover $\{U_\alpha\}_{\alpha \in I}$ of Y, $\big\{\{U_\alpha\}_\alpha, Y^c\big\}$ is an open cover of $\mathbb{R}$. We can then restrict this open cover of $\mathbb{R}$ to $[a,b]$, which is compact (cf. Lemma 2.4.4) and so has a finite subcover. Because Y is bounded by the interval $[a,b]$, restricting this finite subcover to Y also gives a finite subcover of Y. Note that while we added the set Y^c to cover $\mathbb{R}$, which was not part of the original cover of Y, it will not enter the finite subcover of Y as $Y \cap Y^c = \emptyset$. So we indeed have a finite subcover. $\square$

Exercise 2.6. Extend the previous proof to closed and bounded subsets in $\mathbb{R}^n$.

Compact metric spaces have many desirable properties. We list a few.

Theorem 2.4.10. *Let (X,d) be a compact metric space.*

1. *X is* separable*, meaning that there is a countable dense subset $X_0 \subset X$.*
2. *Any continuous function $f : X \to Y$ is uniformly continuous.*

Proof. (1) Let n be a positive integer. Because X is totally bounded, there are elements $x_{n,1}, \dots, x_{n,m} \in X$ such $\big\{B(x_{n,j}; \frac{1}{n})\big\}_{j=1}^m$ is a cover of X. By construction

$$X_0 = \big\{x_{n,j} \mid 1 \le j \le m,\, 1 \le n < \infty\big\} \subset X$$

is a countable set such that for any $x \in X$ and $1 \le n < \infty$, there is some $x_{n,j}$ such that $d(x, x_{n,j}) < \frac{1}{n}$. It follows that $\overline{X_0} = X$ and so X is separable.

(2) Suppose that $f : X \to Y$ is not uniformly continuous. Then for $\delta = \frac{1}{k}$ there is some $\epsilon > 0$ and $x_k, z_k \in X$ such that $d(x_k, z_k) < \frac{1}{k}$ but where $d\big(f(x_k), f(z_k)\big) \ge \epsilon$. We can do this for any positive integer k, which gives us a sequence $\{x_k\}$ in X. Because X is compact, there is a subsequence $x_{k_l} \to x$. Because $d(x_k, z_k) \to 0$ as $k \to \infty$, we can also assume that a subsequence of $\{z_k\}$ converges to x. Therefore $f(x_{k_l}) \to f(x)$ and $f(z_{k_m}) \to f(x)$ as f is continuous. Taking $0 < \epsilon' < \epsilon$, then for k sufficiently large,

$$d\big(f(x_k), f(z_k)\big) \le d\big(f(x_k), f(x)\big) + d\big(f(x), f(z_k)\big) < \epsilon' < \epsilon,$$

which contradicts our assumption and gives the result. $\square$

2.5 Fixed point theorems and some applications

We have been studying function $f : (X, d_X) \to (Y, d_Y)$, but we can also consider the case of maps from X to itself. In various applications (we will consider a few below), it is useful to know whether a continuous map $f : (X,d) \to (X,d)$ has a *fixed point*, $x^* \in X$ such that $f(x^*) = x^*$. Certainly not all maps will have fixed points (e.g. $(1,\infty) \ni x \mapsto x^2 \in (1,\infty)$). But we can apply conditions on the metric space that guarantee such properties.

Definition 2.5.1. We say $f : (X,d) \to (X,d)$ is a contraction if there some $c \in (0,1)$ such that

$$d\big(f(x), f(y)\big) \le c\, d(x,y), \quad \text{for all } x, y \in X.$$

Exercise 2.7. Show that any contraction is (uniformly) continuous.

Theorem 2.5.2 (Contraction mapping theorem). *If (X, d) is a complete metric space and $f : X \to X$ is a contraction, then f has a unique fixed point $x^* \in X$.*

Proof. Take $x \in X$ and define the sequence $y_0 = x$, $y_n = f(y_{n-1}) = f^n(x)$ (meaning we take the composition of f n times). Note that

$$d(y_m, y_{m+1}) = d\big(f^m(x), f^{m+1}(x)\big) \leq c\, d\big(f^{m-1}(x), f^m(x)\big) \leq \ldots \leq c^m\, d\big(x, f(x)\big)$$

and more generally

$$\begin{aligned} d(y_m, y_{m+k}) &= d\big(f^m(x), f^{m+k}(x)\big) \leq d\big(f^m(x), f^{m+1}(x)\big) + \ldots + d\big(f^{m+k-1}(x), f^{m+k}(x)\big) \\ &\leq c^m\, d\big(x, f(x)\big) + \ldots + c^{m+k-1} d\big(x, f(x)\big) \\ &= c^m (1 + c + \ldots + c^{k-1})\, d\big(x, f(x)\big) \\ &\leq \frac{c^m}{1-c} d\big(x, f(x)\big) \xrightarrow{m,k\to\infty} 0, \end{aligned}$$

which shows that $\{y_n\}$ is a Cauchy sequence and so has a limit y. We also note that

$$d(y, f(y)) = \lim_{m\to\infty} d(f^m(y), f(y)) \leq \lim_{m\to\infty} c\, d(f^{m-1}(y), y) = c\, d(y, y) = 0,$$

which shows that $y = f(y)$ and so y is a fixed point. Because we have constructed $y = \lim_{n\to\infty} y_n$ with $y_0 = x$ *any* point in X, then the limit must be unique. $\square$

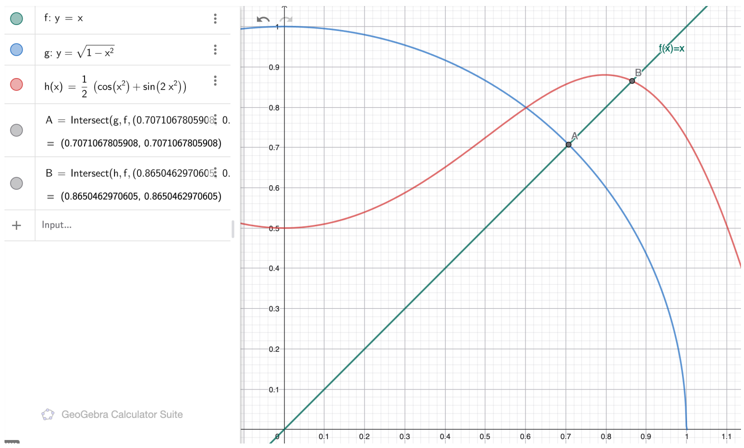
We say a subset $E \subset \mathbb{R}^n$ is convex if for any points $x, y \in E$, the straight line path $(1-t)x + ty \in E$ for all $t \in [0, 1]$. Using the usual metric in $\mathbb{R}^n$, open/closed balls are certainly convex. In two dimensions, standard polygons (pentagon, octagon, ...) are convex, but a star shape is not, for example.

Theorem 2.5.3 (Brouwer fixed point theorem). *Let $E \subset \mathbb{R}^n$ be convex and compact. Then any continuous function $f : E \to E$ has a fixed point.*

We will not prove this theorem (we may prove it in a special setting later in the course). Instead we will consider some of the many ways this result can be used. Further expanding these examples or presenting other applications of Brouwer's fixed point theorem may be a fruitful topic for a self-study report.

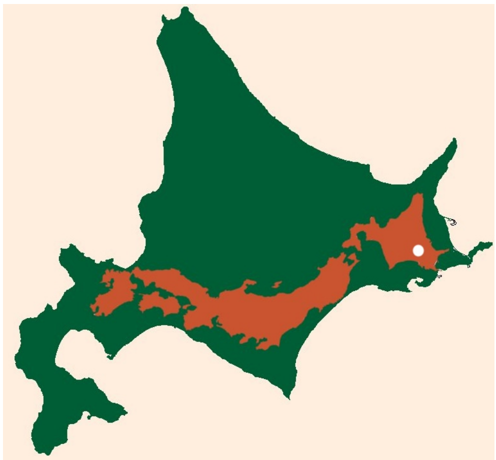
Examples 2.5.4.

1. As a simple application if $f : [0, 1] \to [0, 1]$ is continuous, then there is at least one point $x^* \in [0, 1]$ such that $f(x^*) = x^*$.



Fixed points for continuous functions $f : [0, 1] \to [0, 1]$. Plotted using GeoGebra.

2. Take a two-dimensional image, shrink, rotate it and then place it on top of the original image. At least one point on the new image sits directly above its original point.



2D fixed point. Original image by `@c_oi` on Twitter.

3. If the coffee in a cup is stirred, at least one atom will end up in the same place it started.

4. Suppose I hike up a mountain, rest for the evening and then come back the next day (leaving the same time). Then there is some point along the trail where I passed by at the exact same time on both days.

5. In economics and game theory, the Brouwer fixed point theorem is used to prove the existence of a Nash equilibrium. To roughly explain, players in a game have a set of strategies that they may use with an associated payoff for each strategy. The game is at Nash equilibrium when no player benefits from changing their strategy. As the name suggests, the existence of such an equilibrium was proved by John Nash in his PhD thesis.

6. Suppose we have $\mathbf{0} \in E$ and want to solve the equation $f(\mathbf{x}) = \mathbf{0}$ for some function $f : E \to E$. Provided the altered function $f(\mathbf{x}) + \mathbf{x}$ is still a continuous function $E \to E$, then we know that there is some $\mathbf{x}^*$ such that $f(\mathbf{x}^*) + \mathbf{x}^* = \mathbf{x}^*$ and hence $f(\mathbf{x}^*) = \mathbf{0}$.

Exercise 2.8. Let A be a $n \times n$ real matrix whose entries are strictly positive. Show that A has at least one eigenvalue $\lambda > 0$.

Exercise 2.9. Write a self-study report about the applications of fixed point theorem to ordinary differential equations and dynamical systems.

3 Topological spaces

In the previous section, we considered abstract spaces that come with a notion of distance. This metric allowed us to define neighbourhoods and open sets, which then can be used to give meaning to continuity, closure and other properties in this more general setting. In this section, we go a step further in our abstraction and instead consider sets with a prescribed collection of open sets. While this formalism is certainly abstract and not necessarily intuitive, in various settings and applications we can not always work in $\mathbb{R}^n$ or some other 'nice' space. So it is useful that we can still understand continuity and other properties in this setting.

3.1 Basic definitions

Definition 3.1.1. A topology on $\mathcal{T}$ on a set X is a collection of subsets of X with the following properties.

1. The sets $\emptyset$ and X are elements of $\mathcal{T}$,
2. If $\{U_\alpha\}_{\alpha\in I} \subset \mathcal{T}$, then $\bigcup_{\alpha\in I} U_\alpha \in \mathcal{T}$,
3. If $U_1, \dots, U_n \in \mathcal{T}$, then $\bigcap_{j=1}^{n} U_j \in \mathcal{T}$.

A topological space is a pair $(X, \mathcal{T})$ with X a set and $\mathcal{T}$ a topology on X. We call the elements of $\mathcal{T}$ the open subsets of X.

We remark that the definition of topology says that we can take unions of possibly uncountably many open sets and the result is still an open set. In contrast, we only allow for intersections of a finite number of open sets.

Before we dig into this definition too much, we should first check that the definition of open sets in a topology is compatible with the definition of open sets in a metric space.

Theorem 3.1.2. *Let (X, d) be a metric space and define*

$$\mathcal{T}_d = \big\{U \subset X \mid \textit{for all } x \in U \textit{ there exist } \delta > 0 \textit{ such that } B(x;\delta) \subset U\big\}.$$

Then $(X, \mathcal{T}_d)$ is a topological space. We call $\mathcal{T}_d$ the topology induced by the metric d.

Proof. It follows from the definition of $\mathcal{T}_d$ that $\emptyset, X \in \mathcal{T}_d$. The other statements follow from Proposition 2.1.9. □

Examples 3.1.3.
1. Let X be any set and define $\mathcal{T}_{\text{triv}} = \{\emptyset, X\}$. Then it is easy to check that $\mathcal{T}_{\text{triv}}$ is a topology, which we call the *trivial* topology.
2. Let X be any set and $\mathcal{T}_{\text{disc}} = \mathcal{P}(X) = \{Y \mid Y \subset X\}$. That is, we take $\mathcal{T}_{\text{disc}}$ to be all subsets of X. Then certainly $\emptyset, X \in \mathcal{T}_{\text{disc}}$ and any union or intersection of subsets of X will again be a subset of X. We call this topology the *discrete topology.*
3. Let $X = \{a, b, c\}$ be a set with three elements. There are many possible choices of a topology on X (see Figure 1). For example

$$\mathcal{T}_1 = \{\emptyset, X, \{a\}\}, \quad \mathcal{T}_2 = \{\emptyset, X, \{b\}, \{a,b\}, \{b,c\}\}, \quad \mathcal{T}_3 = \{\emptyset, X, \{b\}, \{c\}, \{a,b\}, \{b,c\}\}$$

are all topologies on X (check for yourself). On the other hand, $\mathcal{S} = \{\emptyset, X, \{a,b\}, \{b,c\}\}$

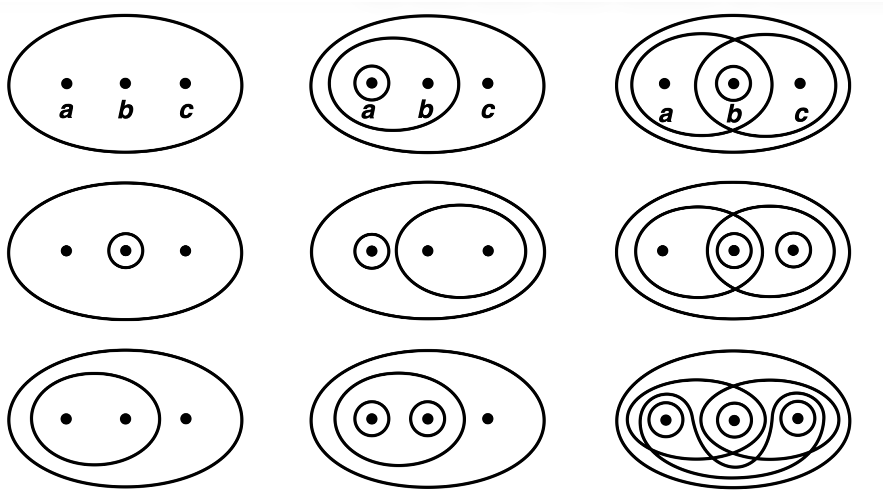


Figure 1: Topologies on $X = \{a, b, c\}$ [6, Figure 12.1].

is *not* a topology on X as $\{a, b\} \cap \{b, c\} = \{b\}$, but $\{b\} \notin \mathcal{S}$, so the set is not closed under intersections.

Definition 3.1.4. If $\mathcal{T}$ and $\mathcal{T}'$ are topologies on a set X. We say that $\mathcal{T}$ is finer than $\mathcal{T}'$ if $\mathcal{T}' \subset \mathcal{T}$. We say $\mathcal{T}$ is coarser than $\mathcal{T}'$ if $\mathcal{T} \subset \mathcal{T}'$.

Finer topologies have more open sets. We always that that $\mathcal{T}_{\text{triv}} \subset \mathcal{T} \subset \mathcal{T}_{\text{disc}}$, so $\mathcal{T}$ is finer than the trivial topology annd coarser than the discrete topology. It may also occur that two topologies $\mathcal{T}$ and $\mathcal{T}'$ are incomparable, where one does not fit inside the other. Figure 1 contains many incomparable topologies, for example.

Definition 3.1.5. Let $(X, \mathcal{T})$ be a topological space. A subset $C \subset X$ is closed if $C^c = X \setminus C$ is open.

Noting that $\emptyset^c = X$ and $X^c = \emptyset$, we see that $\emptyset$ and X are always closed sets. Applying de Morgan's laws (Lemma 2.1.15):

1. If $C_1, \ldots, C_n$ are closed, then $\bigcup_{j=1}^{n} C_j$ is closed,
2. If $\{C_\alpha\}_{\alpha \in I}$ is a family of closed sets, then $\bigcap_{\alpha \in I} C_\alpha$ is closed.

In the setting of metric spaces, we were able to characterise open and closed sets via the interior and closure of a set respectively.

Definition 3.1.6. Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$ a subset. The interior of Y is defined as

$$\operatorname{Int}(Y) := \bigcup \{U \in \mathcal{T} \mid U \subset Y\}.$$

The closure of Y is defined as

$$\overline{Y} = \bigcap \{C \text{ closed} \mid Y \subset C\}.$$

Exercise 3.1. Show that $\text{Int}(Y) \subset Y \subset \overline{Y} = \overline{\overline{Y}}$.

Proposition 3.1.7. *Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$.*

1. *Y is open if and only if $Y = \text{Int}(Y)$.*
2. *Y is closed if and only if $\overline{Y} = Y$.*

Proof. We will prove part (1) and leave part (2) as an exercise. If $Y = \text{Int}(Y)$, then Y is a union of open sets, hence open. If $x \in \text{Int}(Y)$, then $x \in U \subset Y$ for some open $U \in \mathcal{T}$, which implies $\text{Int}(Y) \subset Y$. If Y is open, then Y is an open subset of Y and $Y \subset \text{Int}(Y)$. Hence $Y = \text{Int}(Y)$ if Y is open. □

For a given subset $Y \subset X$, we would like to more concretely understand the closure $\overline{Y}$. In the metric space setting, we can understand $\overline{Y}$ as including the limit points of sequences in Y. For topological spaces, we can say something similar using open neighbourhoods of a point.

Definition 3.1.8. Let $(X, \mathcal{T})$ be a topological space and $x \in X$. An open neighbourhood of x is an open set $U \in \mathcal{T}$ containing x.

Proposition 3.1.9. *Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$.*

1. *A point $x \in \overline{Y}$ if and only if $U \cap Y \neq \emptyset$ for every open neighbourhood U of x.*
2. *$\overline{Y} = Y \cup Y'$, where*

$$Y' = \{x \in X \mid Y \cap (U \setminus \{x\}) \neq \emptyset \textit{ for any open neighbourhood } U \textit{ of } x\}$$

is the set of limit points of Y.

Proof. (1) We will prove via the contrapositive, $x \notin \overline{Y}$ if and only if there is an open neighbourhood U of x with $U \cap Y = \emptyset$. If $x \notin \overline{Y}$, then $x \in U = X \setminus \overline{Y}$, which is an open neighbourhood of x with $U \cap Y = \emptyset$. Conversely, if U is a neighbourhood of x with $U \cap Y = \emptyset$, then $X \setminus U$ is a closed set containing Y. By the definition of the closure $\overline{Y} \subset X \setminus U$ and so $x \notin \overline{Y}$.

(2) It follows from part (1) that $Y' \subset \overline{Y}$ and therefore $Y \cup Y' \subset \overline{Y}$. For the reverse inclusion, let $x \in \overline{Y}$. If $x \in Y$, then $x \in Y \cup Y'$ so we assume $x \notin Y$. If U is an open neighbourhood of x, then $U \cap Y \neq \emptyset$. On the other hand $x \neq Y$, so $x' \in U \cap Y$ for some $x' \neq x$. This implies $x \in Y'$ and so $x \in Y \cup Y'$. □

Examples 3.1.10.

1. If $X = \mathbb{R}$ and $Y = (a, b]$, then $\overline{Y} = [a, b]$ as $(a - \delta, a + \delta)$ intersects Y for all $\delta > 0$. If $x \notin [a, b]$, then $x \in (-\infty, a) \cup (b, \infty)$ and we can take $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subset (-\infty, a) \cup (b, \infty)$.
2. If X has the trivial topology, $\mathcal{T} = \{\emptyset, X\}$, then for any non-empty $Y \subset X$, $\overline{Y} = X$.

Definition 3.1.11. Let $(X, \mathcal{T})$ be a topological space and $\{x_n\}_{n\geq 1}$ a sequence in X. We say that x_n converges to $x \in X$ if for any open neighbourhood $U \in \mathcal{T}$ there is some N such that $x_n \in U$ for all $n \geq N$.

Warning. General limits in topological spaces might not be unique. (For example, consider a space with the trivial topology $\{\emptyset, X\}$.)

Exercise 3.2. Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$. Define the boundary of Y as the set $\partial Y = \overline{Y} \cap \overline{X \setminus Y}$. Show that

$$\mathrm{Int}(Y) \cap \partial Y = \emptyset, \qquad \mathrm{Int}(Y) \cup \partial Y = \overline{Y}.$$

3.2 Subspaces

If (X, d) is a metric space and $Y \subset X$ is a subset, then restricting the metric to Y will turn (Y, d) into a metric space. We would like to do same for topological spaces.

Proposition 3.2.1. *Let $(X, \mathcal{T})$ be a topological space and $Y \subset X$. Then*

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\}$$

is a topology for Y.

Proof. We can easily check that $\emptyset = Y \cap \emptyset$, $Y = Y \cap X \in \mathcal{T}_Y$. Similarly

$$\bigcup_{\alpha}(Y \cap U_\alpha) = Y \cap \bigcup_{\alpha} U_\alpha, \qquad \bigcap_{j=1}^{n}(Y \cap U_j) = Y \cap \bigcap_{j=1}^{n} U_j.$$

□

We call $\mathcal{T}_Y$ the subspace topology and Y is a *subspace* of X when equipped with this topology.

Theorem 3.2.2. *Let Y be a subspace of X and $A \subset Y$.*

1. *A is closed in Y if and only if $A = Y \cap C$ for some closed set C in X.*
2. *Let $\overline{A}$ be the closure of A in X. Then the closure of A in Y is the same as $Y \cap \overline{A}$.*

Proof. (1) If $A = Y \cap C$, then $X \setminus C$ is open in X and $Y \cap (X \setminus C) \in \mathcal{T}_Y$. On the other hand, $Y \cap (X \setminus C) = Y \setminus A$. Therefore $Y \setminus A$ is open in A and hence A is closed in Y. Conversely, if A is closed in Y, then $Y \setminus A = Y \cap U$ for some $U \in \mathcal{T}_X$. Then $X \setminus U$ is closed in X and $A = Y \cap (X \setminus U) =: Y \cap C$.

(2) We consider $\overline{A}^X$ and $\overline{A}^Y$. By part (1) $Y \cap \overline{A}^X$ is a closed in Y with $A \subset Y \cap \overline{A}^X$. By the definition of the closure $\overline{A}^Y \subset Y \cap \overline{A}^X$. On the other hand $\overline{A}^Y = Y \cap C$ for some closed set C in X. As $A \subset \overline{A}^Y$, it follows that $A \subset C$ and therefore $\overline{A}^X \subset C$ as $\overline{A}^X$ is the intersection of all closed sets containing A. Therefore $Y \cap \overline{A}^X \subset Y \cap C = \overline{A}^Y$. □

Note that part (2) of the above theorem is *false* if A is not a subset of Y. As we will see, there are many statements in topology that seem intuitively correct but do not hold in general.

Example 3.2.3. Let $X = \mathbb{R}$ and $Y = [0, 1) \cup \{2\}$. The following sets are open in Y:

$$(0, 1) = Y \cap (0, 1), \qquad [0, 1) = Y \cap (-1, 1), \qquad \{2\} = Y \cap \left(\tfrac{3}{2}, \tfrac{5}{2}\right).$$

In particular, if $A \in \mathcal{T}_Y$, then it does not necessarily follow that $A \in \mathcal{T}_X$.

3.3 Continuity and homeomorphism

We again use our knowledge of metric spaces (particularly Theorem 2.2.3) to extend the definition of continuity to more general topological spaces.

Definition 3.3.1. Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces and $f : X \to Y$ a function. We say that f is continuous if $f^{-1}(V)$ is open in X for all V open in Y.

Remark 3.3.2. At this point, we will begin to be sloppy and not always denote the set $\mathcal{T}$ when describing a topological space. When we say 'X is a topological space', then a choice of topology $\mathcal{T}$ is implicit and will be clear from the context.

We remark that our definition of continuity is global. We can also talk about continuity at a point, where $f : X \to Y$ is continuous at $x \in X$ if for any open neighbourhood V of $f(x) \in Y$, there is an open neighbourhood U of $x \in X$ such that $f(U) \subset V$. We leave it as an exercise to show that $f : X \to Y$ is continuous if and only if it is continuous at every $x \in X$.

Proposition 3.3.3. *1. Let $f : X \to Y$ and $g : Y \to Z$ be continuous maps between topological spaces. Then $g \circ f : X \to Z$ is continuous.*

2. If $f : X \to Y$ is continuous and $A \subset X$ is a subspace, then $f|_A : A \to Y$ is continuous (in the subspace topology).

Proof. 1) The proof is much easier than the ϵ-δ definition of continuity. Let $W \subset Z$ be open, then

$$(g \circ f)^{-1}(W) = f^{-1}\big(g^{-1}(W)\big) = f^{-1}(V) = U,$$

where V is some open set in Y and U is some open set in X.

2) If $V \subset Y$ is open, $f^{-1}(V) \cap A = (f|_A)^{-1}(V) \subset A$ is open by definition. $\square$

Definition 3.3.4. Let X and Y be topological spaces. A homeomorphism between X and Y is a continuous bijection $f : X \to Y$ such that the inverse function $f^{-1} : Y \to X$ is continuous. If a homeomorphism exists between X and Y, we say that the spaces X and Y are homeomorphic.

Exercise 3.3. Show that (a, b) and (c, d) are homeomorphic for $a, b, c, d \in \mathbb{R}$. Does this also hold if any of the endpoints are $\pm\infty$?

Exercise 3.4. Let $I \subset \mathbb{R}$ be any interval (open, closed, half-open, half-closed, bounded, unbounded). Show that I is homeomorphic to one of the following: $[0, 1]$, $(0, 1)$, $[0, 1)$, $(0, 1]$.

Example 3.3.5. Any open ball $B(\mathbf{x}_0, r) \subset \mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$. We define a map

$$f : B(\mathbf{x}_0, r) \to \mathbb{R}^n, \qquad f(\mathbf{x}) = \frac{d(\mathbf{x}, \mathbf{x}_0)}{r - d(\mathbf{x}, \mathbf{x}_0)}(\mathbf{x} - \mathbf{x}_0).$$

Note first that the map $[0, r) \ni t \overset{g}{\mapsto} \frac{t}{r-t} \in [0, \infty)$ gives a homeomorphism of half-open intervals. The function is continuous with a continuous inverse $g^{-1}(t) = \frac{tr}{1+t}$. Now, the function $f(\mathbf{x}) = g\big(d(\mathbf{x}, \mathbf{x}_0)\big)(\mathbf{x} - \mathbf{x}_0)$. The composition $g \circ d(\mathbf{x_0}, \cdot) : [0, r) \to [0, \infty)$ will scale the ball to $\mathbb{R}^n$ and the map $\mathbf{x} \mapsto \mathbf{x} - \mathbf{x}_0$ is a homeomorphism from $B(\mathbf{x_0}; r)$ to $B(\mathbf{0}, r)$. Combining these facts, we see that f is indeed a homeomorphism.

Example 3.3.6 (Stereographic projection (1D)). We consider a circle with the north pole removed. In particular we write $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + (y-\frac{1}{2})^2 = \frac{1}{4}\}$, which describes a circle in $\mathbb{R}^2$ with north pole $N = (0,1)$ and south pole $S = (0,0)$. We then define

$$f : S^1 \setminus \{N\} \to \mathbb{R}, \qquad f(x,y) = \frac{2x}{1-y},$$

which maps (x,y) to the point on $\mathbb{R} = \{(x',0) \mid x' \in \mathbb{R}\}$ via the straight line from N to $(x',0)$ that passes through (x,y). The map is well defined as we have excluded the point $N = (0,1)$. The inverse map is given by

$$f^{-1} : \mathbb{R} \to S^1 \setminus \{N\}, \qquad f^{-1}(x) = \Big(\frac{4x}{x^2+4}, \frac{x^2-4}{x^2+4}\Big).$$

We leave it as an exercise to check that both maps are continuous bijections.

Exercise 3.5. Extend the stereographic projection to find a homeomorphism $S^2 \setminus \{N\} \to \mathbb{R}^2$.

Whether a function $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is continuous or not crucially depends on the topology. As a simple example

$$(\mathbb{R}, |\cdot|) \ni x \mapsto x \in (\mathbb{R}, \mathcal{T}_{\text{disc}})$$

is not continuous as $f^{-1}(\{a\}) = \{a\}$, which is not open in the standard topology of $\mathbb{R}$ but is open in the discrete topology.

Remark 3.3.7 (Continuity via nets*). For the case of metric spaces and a function $f : (X, d_X) \to (Y, d_Y)$, we have the equivalence

$$f : X \to Y \text{ is continuous} \quad \Longleftrightarrow \quad \text{For all sequences } x_n \to x,\ f(x_n) \to f(x).$$

Such a condition no longer holds for general maps between topological spaces. However, we can recover an analogous result if we generalise our notion of sequences and convergence to that of *nets*. We will not provide a precise definition, see [6, Section 29, supplementary material] for example. Roughly speaking we consider $(x_\alpha)_{\alpha \in I}$, where I is a *directed* index set. The example to keep in mind is where $I = \mathcal{P}(X) = \{Y \subset X\}$, the set of subsets of a set X, where there is a partial order relation $\preceq$ given by $A \preceq B$ if $A \subset B$. Not every subset can be compared, but we have that if $A \preceq B$ and $B \preceq C$, then $A \preceq C$. If $(x_\alpha)_{\alpha \in I} \subset X$ is a net in a topological space $(X, \mathcal{T})$, we say that (x_α) converges to $x \in X$ if for every open neighbourhood $U \in \mathcal{T}_X$ of x, there is some $\alpha \in I$ such that $x_\beta \in U$ for all $\beta \in I$ such that $\alpha \preceq \beta$. We then write $(x_\alpha)_{\alpha \in I} \to x$ in X. Given a function $f : X \to Y$ between topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$, we have that

$$f : X \to Y \text{ is continuous} \quad \Longleftrightarrow \quad \text{For all nets } (x_\alpha)_{\alpha \in I} \to x,\ (f(x_\alpha))_{\alpha \in I} \to f(x).$$

We leave it as a possible report topic to provide a more detailed explanation of nets as well as the proof of the the above statement.

3.4 Basis of a topology

Suppose that (X, d) is a metric space. Then the sets $B(x; r)$ are open for all $r > 0$ and $U \subset X$ is open if for every $x \in U$ there is some r_x such that $B(x; r_x) \subset U$. So while general open sets $U \subset X$

are more general than open balls, we can characterise them via open balls. That is, the topology of the space can be determined via the set $B(x;r)$ for $x \in X$ and $r > 0$. Roughly speaking, a basis for a topology lets us understand the topology $\mathcal{T}$ of a space X in terms of simpler parts.

Definition 3.4.1. Let X be a set. A basis for a topology on X is a collection of subsets $\mathcal{B} \subset \mathcal{P}(X)$ such that

1. If $x \in X$, there is at least one $B \in \mathcal{B}$ such that $x \in B$,
2. If $x \in B_1 \cap B_2$ with $B_1, B_2 \in \mathcal{B}$, then there is some $B_3 \in \mathcal{B}$ such that $B_3 \subset B_1 \cap B_2$ and $x \in B_3$.

If $\mathcal{B}$ is a basis for a topology on X, then

$$\mathcal{T}_{\mathcal{B}} = \{U \subset X \mid \text{for all } x \in U \text{ there is some } B \in \mathcal{B} \text{ such that } x \in B \subset U\}.$$

is the topology generated by $\mathcal{B}$.

We should first check that this definition is reasonable.

Lemma 3.4.2. *If $\mathcal{B}$ is a basis for a topology on X, then $\mathcal{T}_{\mathcal{B}}$ is a topology.*

Proof. Supposing that $\mathcal{B}$ is a basis, if $U = \emptyset$, then $\emptyset \in \mathcal{T}_{\mathcal{B}}$ trivially. Similarly if $U = X$, then for any $x \in X$, by condition (1) for a basis, there is some $B \in \mathcal{B}$ with $x \in B$ and $B \subset X$. If $\{U_\alpha\}$ is a family of open sets, then we consider $U = \bigcup_\alpha U_\alpha$ and $x \in U$. By assumption $x \in U_{\alpha'}$ for some α' and so there is some $B_{\alpha'}$ with $x \in B_{\alpha'} \subset U_{\alpha'} \subset U$. Thus $U \in \mathcal{T}_{\mathcal{B}}$. For finite intersections, we use mathematical induction. If $U_1, U_2 \in \mathcal{T}_{\mathcal{B}}$, then if $x \in U_1 \cap U_2$, then $x \in B_1 \cap B_2$ and by condition (2) there is some $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2 \subset U_1 \cap U_2$. Thus $U_1 \cap U_2 \in \mathcal{T}_{\mathcal{B}}$. We now assume the result holds for $U_1, \ldots, U_n$ and consider $U_1 \cap \ldots \cap U_n \cap U_{n+1}$. But we can let $U' = U_1 \cap \ldots \cap U_n \in \mathcal{T}_{\mathcal{B}}$ and apply the same arguement to show that $U' \cap U_{n+1} = \bigcap_{j=1}^{n+1} U_n \in \mathcal{T}_{\mathcal{B}}$. □

Example 3.4.3. The set $\{(a,b) \subset \mathbb{R} \mid a, b \in \mathbb{R}\}$ is a basis for a topology on $\mathbb{R}$. The topology it generates is indeed the standard topology that can also be described by the metric $d(x,y) = |x - y|$. More generally, if (X,d) is a metric space, then the open balls

$$\{B(x;r) \subset X \mid x \in X, \, r > 0\}$$

are a basis for the topology of X.

A common way to construct topologies on a set is to specify a basis and then take the topology it generates. This is generally easier to do than to specify the whole set $\mathcal{T}$.

Example 3.4.4. The set $\{[a,b) \subset \mathbb{R} \mid a, b \in \mathbb{R}\}$ is also a basis for a topology on $\mathbb{R}$ (called the lower-limit topology on $\mathbb{R}$).

Exercise 3.6. If $\mathcal{T}_{(a,b)}$ and $\mathcal{T}_{[a,b)}$ denote the standard and lower-limit topology on $\mathbb{R}$, show that $\mathcal{T}_{(a,b)} \subsetneq \mathcal{T}_{[a,b)}$.

Lemma 3.4.5. *Let $\mathcal{B}$ be a basis for a topology on X. Then*

$$\mathcal{T}_{\mathcal{B}} = \Big\{U \subset X \,\Big|\, U = \bigcup_{\alpha} B_\alpha,\ B_\alpha \in \mathcal{B}\Big\}.$$

Proof. If $U \in \mathcal{T}_{\mathcal{B}}$, then for each $x \in U$ we have some $B_x \in \mathcal{B}$ such that $x \in B_x \subset U$. Therefore we may write $U = \bigcup_{x\in U} B_x$ and U is a union of elements in $\mathcal{B}$. Conversely, if $U = \bigcup_\alpha B_\alpha$, then because $\{B_\alpha\} \subset \mathcal{T}_{\mathcal{B}}$, $\bigcup_\alpha B_\alpha = U \in \mathcal{T}_{\mathcal{B}}$. So the sets are equal. □

A useful property of a basis is the following (which is sometimes used as the definition of a basis).

Theorem 3.4.6. *Let $(X, \mathcal{T})$ be a topological space. Then $\mathcal{B} \subset \mathcal{P}(X)$ is a basis for $\mathcal{T}$, meaning that $\mathcal{T}_{\mathcal{B}} = \mathcal{T}$, if and only if every $U \in \mathcal{T}$ can be written as a union of elements in $\mathcal{B}$.*

Proof. If $\mathcal{T}_{\mathcal{B}} = \mathcal{T}$ and $U \in \mathcal{T}$, then for each $x \in U$ we have some $B_x \in \mathcal{B}$ such that $x \in B_x \subset U$. Therefore we may write $U = \bigcup_{x\in U} B_x$ and U is a union of elements in $\mathcal{B}$. Conversely, if any U can be written as a union of elements in $\mathcal{B}$, then we can write

$$\mathcal{T} = \Big\{U \subset X \mid U = \bigcup_{\alpha} B_\alpha,\ B_\alpha \in \mathcal{B}\Big\} = \mathcal{T}_{\mathcal{B}}$$

by Lemma 3.4.5. □

We see that to understand a topology $\mathcal{T}$ on X, it suffices to find a colleciton of open sets such that any $U \in \mathcal{T}$ can be decomposed as a union of elements of these sets.

The following exercise gives a useful criteria to determine if a collection of open sets form a basis.

Exercise 3.7. Let $(X, \mathcal{T})$ be a topological space and $\mathcal{B} \subset \mathcal{T}$ a collection of open sets such that for each $U \in \mathcal{T}$ and $x \in U$, there is some $B \in \mathcal{B}$ with $x \in B \subset U$. Show that $\mathcal{B}$ is a basis for $\mathcal{T}$.

Remark 3.4.7. Any open set can be decomposed as a union of elements in a basis. On the other hand, this decomposition is **not** unique. So a basis of a topology is different to a basis of a vector space V, where the decomposition of $v \in V$ as a linear combination of basis elements is unique. Similalry, if $\mathcal{B}$ and $\mathcal{B}'$ are basis for a topology with $\mathcal{B} \neq \mathcal{B}'$, it may still occur that $\mathcal{T}_{\mathcal{B}} = \mathcal{T}_{\mathcal{B}'}$.

Exercise 3.8. Show that $f : X \to Y$ is continuous if $f^{-1}(V) \subset X$ is open for all $V \in \mathcal{B}_Y$, a basis of Y.

Definition 3.4.8. We say that $(X, \mathcal{T})$ is second countable if there is a countable family $\mathcal{B} \subset \mathcal{T}$ such that that forms a basis for $\mathcal{T}$, $\mathcal{T} = \mathcal{T}_{\mathcal{B}}$.

The real line $\mathbb{R}$ with the standard topology is second countable as we can take the basis

$$\mathcal{B} = \big\{(a, b) \subset \mathbb{R} \mid a, b \in \mathbb{Q}\big\},$$

which is countable.

3.5 Product topological spaces

Lemma 3.5.1. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, then the set*

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X,\, V \in \mathcal{T}_Y\}$$

is a basis for a topology on $X \times Y$.

Proof. We note that $X \times Y \in \mathcal{B}_{X\times Y}$, so certainly for any $(x, y) \in X \times Y$ there is a basis element containing (x, y). Now if $(x, y) \in B_1 \cap B_2 = (U_1 \times V_1) \cap (U_2 \times V_2)$, then $x \in U_1 \cap U_2$ and $y \in U_1 \cap U_2$. But $U_1 \cap U_2 \in \mathcal{T}_Y$ and $V_1 \cap V_2 \in \mathcal{T}_Y$. So we let $B_3 = (U_1 \cap U_2) \times (V_1 \cap V_2)$ which contains (x, y) and is contained in $B_1 \cap B_2$. □

Definition 3.5.2. If $(X, \mathcal{T}_Y)$ and $(Y, \mathcal{T}_Y)$ are topological spaces, the product topology on $X \times Y$ is the topology $\mathcal{T}_{X\times Y} = \mathcal{T}_{\mathcal{B}_{X\times Y}}$ with $\mathcal{B}_{X\times Y}$ the basis from Lemma 3.5.1.

Examples 3.5.3.

1. The plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ with the standard topology is the product of $(\mathbb{R}, |\cdot|)$ with itself. We can iterate this procedure to obtain $\mathbb{R}^n$.
2. The 2-torus $\mathbb{T}^2 = S^1 \times S^1$ with the product topology. More generally $\mathbb{T}^n = \underbrace{S^1 \times \ldots \times S^1}_{n \text{ times}}$.

To characterise the product topology on $X \times Y$, it is easier still to work with bases for X and Y.

Theorem 3.5.4. *If $\mathcal{B}_X$ is a basis for a topology on X and $\mathcal{B}_Y$ is a basis for a topology on Y, then*

$$\mathcal{B}_X \times \mathcal{B}_Y = \{B \times C \mid B \in \mathcal{B}_X,\, C \in \mathcal{B}_Y\}$$

is a basis for the product topology $\mathcal{T}_{X\times Y}$.

Proof. Let $(x, y) \in W \subset X \times Y$ with W open. Using the standard basis for $\mathcal{T}_{X\times Y}$, there is some $U \times V \in \mathcal{T}_X \times \mathcal{T}_Y$ with $(x, y) \in U \times V \subset W$, i.e. $x \in U$ and $y \in V$. Because $\mathcal{B}_X$ and $\mathcal{B}_Y$ are basis of $\mathcal{T}_X$ and $\mathcal{T}_Y$, there is some $B \in \mathcal{B}_X$ and $C \in \mathcal{B}_Y$ such that $x \in B$, $y \in C$ and hence $(x, y) \in B \times C \subset U \times V \subset W$. Applying the result of Exercise 3.7, this suffices to show that $\mathcal{B}_{X\times Y}$ is a basis for $\mathcal{T}_{X\times Y}$. □

The above result shows that we can characterise the standard topology of $\mathbb{R}^2$ to be generated by open rectangles

$$\mathcal{B}_{\mathbb{R}^2} = \{(a, b) \times (c, d) \mid a, b, c, d \in \mathbb{R}\}.$$

Proposition 3.5.5. *Let $X \times Y$ be a product space. Then the projection maps*

$$p_1 : X \times Y \to X, \quad p(x, y) = x, \qquad p_2 : X \times Y \to Y, \quad p_2(x, y) = y$$

are continuous surjections.

Proof. The maps p_1 and p_2 are clearly surjective. If $U \subset X$ is open, then

$$p_1^{-1}(U) = U \times Y \in \mathcal{B}_{X\times Y} \subset \mathcal{T}_{X\times Y}$$

and we see that p_1 is continuous. The proof for p_2 is analogous. $\square$

Using mathematical induction, we can extend the product topology to finite Cartesian products.

Definition/Theorem 3.5.6. *Let $n \in \mathbb{N}$ and $(X_1, \mathcal{T}_1), \ldots, (X_n, \mathcal{T}_n)$ be topological spaces. The product topology on $X_1 \times \ldots X_n$ is generated by the basis*

$$\begin{aligned}\mathcal{B}_{\text{prod}} &= \big\{U_1 \times \ldots \times U_n \subset X_1 \times \ldots \times X_n \mid U_j \in \mathcal{T}_j, \text{ for all } j = 1, \ldots, n\big\} \\ &= \mathcal{B}_1 \times \ldots \times \mathcal{B}_n\end{aligned}$$

with $\mathcal{B}_j$ a basis for $\mathcal{T}_j$ for all $j = 1, \ldots, n$.

We leave the proof as an exercise.

Infinite products

There are times where we need to consider not just finite but also infinite products of topological spaces. The question of what topology to put on this space is not necessarily just one of pure mathematical interest. Infinite products of spaces often occur in dynamical systems and in physics as an approximation of systems with uncontrolled degrees of freedom.

If I is an index set and $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha \in I}$ a family of topological spaces, the obvious choice for the product topology is to take a basis of products of open sets. We take a slightly different definition and then explain its advantages.

Definition 3.5.7. Let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha \in I}$ be a family of topological spaces. The product topology on $\prod_{\alpha \in I} X_\alpha$ is the topology determined by the basis

$$\mathcal{B}_{\text{prod}} = \Big\{ \prod_{\alpha \in I} U_\alpha \,\Big|\, U_\alpha \in \mathcal{T}_\alpha \text{ and } U_\alpha = X_\alpha \text{ for all but finitely many } \alpha \in I \Big\}.$$

If I is a finite set, then this topology agrees with our previous definition of the product topology.

Exercise 3.9. Show that for all $\alpha' \in I$, the projection functions

$$p_{\alpha'} : \prod_{\alpha \in I} X_\alpha \to X_{\alpha'}, \qquad p_{\alpha'}\big((x_\alpha)_{\alpha \in I}\big) = x_{\alpha'}$$

are continuous.

Theorem 3.5.8. *Let $f_\alpha : Y \to X_\alpha$ be a family of functions and define*

$$f : Y \to \prod_{\alpha \in I} X_\alpha, \qquad f(y) = \big(f_\alpha(y)\big)_{\alpha \in I}.$$

Then f is continuous if and only if f_α is continuous for all $\alpha \in I$.

Proof. If f is continuous, then we can write $f_\alpha = f \circ p_\alpha$, which is a composition of continuous functions (by the previous exercise). Hence f_α is continuous for all $\alpha \in I$.

Conversely, suppose that $\{f_\alpha\}_{\alpha \in I}$ are continuous. To check continuity of f it suffices to check that $f^{-1}(U)$ is open for $U \in \mathcal{B}_{\text{prod}}$. So we suppose that $U = \prod_{\alpha \in I} U_\alpha$ with $U_\alpha = X_\alpha$ for all but finitely many α. Let us label the open subsets $\{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ that are *not* equal to the entire space. Then

$$\begin{aligned} f^{-1}(U) &= \left\{ y \in Y \mid f_{\alpha_j}(y) \in U_{\alpha_j} \text{ for all } j = 1, \dots, n \right\} \\ &= \bigcap_{j=1}^{n} f^{-1}\left(U_{\alpha_j}\right) \in \mathcal{T}_Y \end{aligned}$$

as f_{α_j} is continuous for all j. □

If we were to take a basis of consisting of infinite products of proper open subsets, then the previous proof would involve an infinite intersection of open sets, which might not be open.

Exercise 3.10. Write a report further examining the possible topologies on infinite products $\prod_{\alpha \in I} X_\alpha$.

3.6 Hausdorff and normal topological spaces

Recall that a a homeomorphism between topological spaces is a continuous bijection $f : X \to Y$ with continuous inverse. We can define a relation on pairs of topological spaces where $X \sim Y$ if there exists a homeomorphism $f : X \to Y$. It is a simple exercise to show that $\sim$ is an equivalence relation and we can therefore consider the equivalence *class* $[X] = \left\{ Y \mid Y \sim X \right\}$. An important question is what properties of a topological space are preserved under homeomorphism (and therefore are also well-defined at the level of equivalence class).

The first such properties we consider are so-called separation axioms, though we will only consider a few special cases.

Definition 3.6.1.
1. A topological space $(X, \mathcal{T})$ is called Hausdorff if for every $x, y \in X$ such that $x \neq y$, there are open neighbourhoods $x \in U \in \mathcal{T}$, $y \in V \in \mathcal{T}$ such that $U \cap V = \emptyset$.
2. A topological space $(X, \mathcal{T})$ is called normal if for every disjoint pair of closed sets $A, B \subset X$, $A \cap B = \emptyset$, there are open sets $A \subset U \in \mathcal{T}$, $B \subset V \in \mathcal{T}$ such that $U \cap V = \emptyset$.

A Hausdorff space lets us separate points by open sets. A normal space lets us separate disjoint closed sets by open sets.

Exercise 3.11. Show that limits of sequences in Hausdorff topological spaces are unique. For those who know the definition of convergence of *nets*, show that the same result is true for limits of nets in X.

Exercise 3.12. If $\{X_\alpha\}_{\alpha \in I}$ is a family of Hausdorff spaces, show that $\prod_{\alpha \in I} X_\alpha$ is Hausdorff.

Lemma 3.6.2.
1. *If X is Hausdorff, then $\{x\} \subset X$ is a closed set for all $x \in X$.*
2. *If X is normal* and *$\{x\} \subset X$ is closed for all $x \in X$, then X is Hausdorff.*

Proof. (1) For any $x \in X$ and $y \neq X$, we can find an open set U such that $x \notin U$, i.e. $U \subset X \backslash \{x\}$. This holds for any point $y \in X \setminus \{x\}$ and so $X \setminus \{x\}$ is open. Part (2) is immediate. □

Remark 3.6.3. Hausdorff spaces are sometimes called T_2-spaces. Normal spaces such that $\{x\} \subset X$ is closed for all x are sometimes called T_4-spaces.

Exercise 3.13. Let $f : X \to Y$ be a homeomorphism. If X is normal and Hausdorff, show that Y is normal and Hausdorff.

Exercise 3.14. Show that any metric space (X, d) is Hausdorff and normal.

Certainly not every topological space is Hausdorff or normal, but almost all of the topological spaces that one encounters in the natural sciences are normal and Hausdorff. Thankfully, these spaces have particularly nice properties

Lemma 3.6.4. *A topological space $(X, \mathcal{T})$ is normal if and only if for every closed $A \subset X$ with $A \subset V \in \mathcal{T}$, there is some $U \in \mathcal{T}$ such that $A \subset U \subset \overline{U} \subset V$.*

Proof. If X is normal, then we consider A and V^c, which are disjoint open sets. Hence there are open sets $A \subset U_1$ and $V^c \subset U_2$ such that $U_1 \cap U_2 = \emptyset$. Then $\overline{U_1} \subset U_2^c \subset V$, which gives the desired open set. Conversely, take closed and disjoint sets A and B in X. Then $A \subset B^c$ and B^c is open. So by the assumption there is some $U \in \mathcal{T}$ such that $A \subset U \subset \overline{U} \subset B^c$. Now U and $V = (\overline{U})^c$ are disjoint open sets such that $A \subset U$, $B \subset V$ and $U \cap V = \emptyset$. Therefore X is normal. □

In addition to topological spaces themselves, we are often interested in continuous functions $f : X \to \mathbb{R}$. If the space X is normal, then this is just enough structure to allow us to:

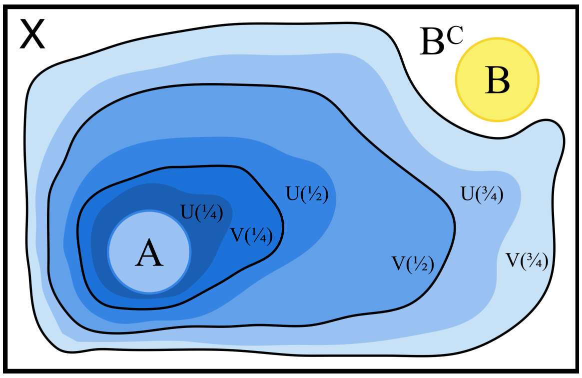
1. Restrict or 'localise' functions around a particular area by multiplying with a 'bump function' supported on a closed subset.
2. Extend functions defined on closed subsets $A \subset X$ to all of X in a reasonable way.

The following two results explain how this can be done. They also constitute the first results in abstract topological spaces that require a more substantial proof (that is, not just some basic rearrangement of the definitions).

Theorem 3.6.5 (Urysohn's lemma)**.** *Let $(X, \mathcal{T})$ be a normal topological space and A and B closed subsets with $A \cap B = \emptyset$. There exists a continuous function $f : X \to [0, 1]$ such that $f|_A = 0$ and $f|_B = 1$.*

Proof. Let $V = B^c$, which is an open set such that $A \subset V$. Applying Lemma 3.6.4, there is some $U_{1/2}$ such that $A \subset U_{1/2} \subset \overline{U_{1/2}} \subset V$. We can apply this lemma again to the open sets $U_{1/2}$ containing A and V containing $\overline{U_{1/2}}$ to get open sets $U_{1/4}$ and $U_{3/4}$ such that

$$A \subset U_{1/4} \subset \overline{U_{1/4}} \subset U_{1/2} \subset \overline{U_{1/2}} \subset U_{3/4} \subset \overline{U_{3/4}} \subset V.$$



The first few steps of the proof the construction of f. Here $V(a/b) = \overline{U_{a/b}}$ is a closed set (taken from Wikipedia).

We can continue this process for each dyadic fraction < 1, a number of the form $\frac{j}{2^n}$ for some $n \in \mathbb{N}$ and $j \in \{1, \ldots, 2^n - 1\}$. Namely, for each $n \geq 1$ and $j \in \{1, \ldots, 2^n - 1\}$, we have

$$\overline{U_{\frac{j}{2^n}}} \subset U_{\frac{2j+1}{2^{n+1}}} \subset \overline{U_{\frac{2j+1}{2^{n+1}}}} \subset U_{\frac{j+1}{2^n}}$$

with $A \subset U_{\frac{1}{2^n}}$ and $\overline{U_{\frac{2^n-1}{2^n}}} \subset V = B^c$. We now define the function

$$f : X \to [0,1], \qquad f(x) = \begin{cases} 0, & x \in U_r,\ r > 0, \\ \sup\{r : x \notin U_r\}, & x \notin U_r, \end{cases}$$

where $r \in (0,1)$ denotes a dyadic fraction. The function is well-defined and is such that $f(x) = 0$ for $x \in A$ and $f(x) = 1$ for $x \in B$. So we just have to show f is continuous. Take $x \in X$ and assume $f(x) \in (0,1)$ (the extreme points $f(x) = 0, 1$ are an exercise). For $\epsilon > 0$, we can find dyadic fractions $r, s \in (0,1)$ such that

$$f(x) \in (r,s) \subset (f(x) - \epsilon, f(x) + \epsilon)$$

as the dyadic fractions are dense in $(0,1)$. Because $r < f(x)$, $x \notin \overline{U_r}$ while, on the other hand, $f(x) < s$ so $x \in U_s$. So $W = U_s \setminus \overline{U_r}$ is an open neighbourhood of x such that $|f(y) - f(x)| < \epsilon$ for all $y \in W$. Thus f is continuous at x. $\square$

Exercise 3.15. Let $(X, \mathcal{T})$ be a normal topological space and A and B closed subsets with $A \cap B = \emptyset$. For any $a, b \in \mathbb{R}$ with $a < b$, show that exists a continuous function $f : X \to [a,b]$ such that $f|_A = a$ and $f|_B = b$.

Example 3.6.6 (Finite partition of unity)**.** Suppose that $(X, \mathcal{T})$ is a topological space and $\{U_1, \ldots, U_n\}$ is an open cover of X. We say that a collection of continuous functions $\{\chi_j\}_{j=1}^n \subset C(X, \mathbb{R})$ is a partition of unity subordinate to $\{U_1, \ldots, U_n\}$ if

1. For all $j = 1, \ldots, n$, $\chi_j(x) \in [0,1]$ and $\chi_j(x) = 0$ for all $x \in X \setminus U_j$ ($\mathrm{supp}(\chi_j) \subset U_j$),
2. For all $x \in X$, $\sum_{j=1}^{n} \chi(x) = 1$.

If $(X, \mathcal{T})$ is normal and has a finite cover, then a partition subordinate to this cover always exists. The proof can be found in [6, Theorem 36.1], for example, and relies on Urysohn's Lemma.

The advantage of partitions of unity is that it allows us to 'localise' a function $f : X \to \mathbb{R}$ by instead considering $\chi_j f$, which will only be non-zero for $x \in U_j$. We do not lose any information by localising functions as the full picture can be recovered,

$$f(x) = \sum_{j=1}^{n} \chi_j(x) f(x), \quad x \in X, \quad \operatorname{supp}(\chi_j f) \subset U_j.$$

Partitions of unity play a crucial role in the theory of topological and differentiable manifolds. More general notions of partitions of unity can be considered for locally finite covers of topological spaces. We leave it as a possible report topic to further study partitions of unity, their properties and applications.

Theorem 3.6.7 (Tietze extension theorem). *Let $(X, \mathcal{T})$ be a normal topological space, $Y \subset X$ a closed subset and $f : Y \to \mathbb{R}$ a bounded and continuous function. Then there is a bounded and continuous function $\tilde{f} : X \to \mathbb{R}$ such that $\tilde{f}|_Y = f$.*

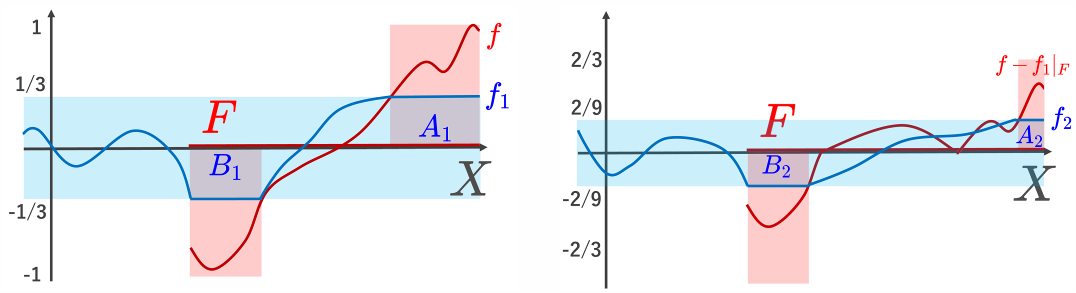
Proof. Because $f : Y \to \mathbb{R}$ is bounded, there is some value $C = \sup_{x \in Y} |f(x)| \in [0, \infty)$. Using C, we define

$$A_1 = \{y \in Y \mid f(y) \geq \tfrac{1}{3}C\}, \qquad B_1 = \{y \in Y \mid f(y) \leq -\tfrac{1}{3}C\}.$$

Applying Urysohn's Lemma (and the exercise that follows it), there is a continuous function $f_1 : X \to \left[-\frac{C}{3}, \frac{C}{3}\right]$ such that $f_1|_{A_1} = \frac{1}{3}C$ and $f_1|_{B_1} = -\frac{1}{3}C$. We also note that $\left|f_1(y) - f(y)\right| \leq \frac{2}{3}C$ by construction. We can repeat this construction replacing f with $f_1 - f$, where we have

$$A_2 = \{y \in Y \mid f(y) - f_1(y) \geq \tfrac{2}{9}C\}, \qquad B_2 = \{y \in Y \mid f(y) - f_1(y) \leq -\tfrac{2}{9}C\}$$

and $f_2 : X \to \left[-\frac{2}{9}C, \frac{2}{9}C\right]$.



The first few steps of the proof the approximation of f via Urysohn's Lemma (here F is the subset Y; taken from mathlandscape.com).

We can continue this pattern to obtain a sequence of functions $\{f_n\}_{n=1}^{\infty}$ such that

$$f_n : X \to \left[-\frac{2^{n-1}}{3^n}C, \frac{2^{n-1}}{3^n}C\right], \qquad \left|f - f_1 - f_2 - \ldots - f_n\right| \leq \frac{2^n}{3^n}C \text{ on } Y.$$

Indeed, supposing that we have constructed $f_1, \ldots, f_n$, we let

$$C_n = \sup_{y \in Y} \left|f(y) - f_1(y) - \ldots - f_n(y)\right|$$

and repeat the above argument replacing C with C_n and f with $f - f_1 \ldots - f_n$. We therefore obtain $f_{n+1} : X \to \mathbb{R}$ such that

$$\sup_{x \in X} \left|f_{n+1}(x)\right| \leq \frac{C_n}{3} = \frac{2^n}{3^{n+1}}C, \qquad \sup_{y \in Y} \left|f(y) - f_1(y) - \ldots - f_n(y)\right| \leq \frac{2C_n}{3} = \frac{2^{n+1}}{3^{n+1}}C.$$

We now set

$$\tilde{f} : X \to \mathbb{R}, \qquad f(x) = \sum_{j=1}^{\infty} f_j(x) = \lim_{n\to\infty} h_n(x), \qquad h_n = \sum_{j=1}^{n} f_j.$$

To see that $\tilde{f}$ is well-defined, recall that $C_b(X,\mathbb{R})$, the set of bounded continuous functions is a complete metric space with respect to the metric $d(f,g) = \sup_{x\in X} \big|f(x)-g(x)\big|$ (see the homework questions). By construction, f_j and h_n are elements of $C_b(X,\mathbb{R})$ for any (finite) j, n. Then for $n \geq m$,

$$\begin{aligned} d(h_n, h_m) &= \sup_{x\in X} \big|h_n(x) - h_m(x)\big| = \sup_{x\in X} \big|f_{m+1}(x) + \ldots + f_n(x)\big| \\ &\leq \left(\left(\tfrac{2}{3}\right)^{m+1} + \ldots \left(\tfrac{2}{3}\right)^{n}\right) \frac{C}{3} \\ &< \frac{C}{3} \sum_{j=m}^{\infty} \left(\frac{2}{3}\right)^j = \left(\frac{2}{3}\right)^m C \xrightarrow{m\to\infty} 0, \end{aligned}$$

which shows that $\{h_n\}$ is a Cauchy sequence. Therefore h_n so converges to the limit $\tilde{f}$ in the supremum norm, which will also be continuous and bounded. A similar argument will also show that for any $y \in Y$,

$$\big|f(y) - h_n(y)\big| \leq \left(\tfrac{2}{3}\right)^{n+1} C \xrightarrow{n\to\infty} 0,$$

which implies that $f(y) = \tilde{f}(y)$ for any $y \in Y$. □

3.7 Compactness

In the previous subsection we defined Hausdorff and normal spaces, which are properties of topological spaces that are preserved under homeomorphism. We now consider compact spaces, which we have already encountered in the context of metric spaces. As we will see, compactness is also preserved under homeomorphisms and has many desirable properties.

We use basically the same definition for compactness as we encountered in the metric space setting.

Definition 3.7.1. Let $(X, \mathcal{T})$ be a topological space.

1. An open cover of X is a family of open sets $\{U_\alpha\}_{\alpha\in I}$ such that $\bigcup_{\alpha\in I} U_\alpha = X$.
2. We say an open cover $\{U_\alpha\}_{\alpha\in I}$ has a finite subcover if there is a finite sub-family $\{U_{\alpha_1}, \ldots, U_{\alpha_n}\}$ such that $X = \bigcup_{i=1}^{n} U_{\alpha_i}$.
3. We say that $(X, \mathcal{T})$ is compact if every open cover of X has a finite subcover.

It follows immediately that any compact metric space is a compact topological space. In particular, closed and bounded subsets of $\mathbb{R}^n$ (with the standard topology) are compact.

Examples 3.7.2.

1. Any set X with the trival topology $\mathcal{T}_{\text{triv}} = \{X, \emptyset\}$ is compact as any open cover is already finite.
2. A set with the discrete topology $(X, \mathcal{T}_{\text{disc}})$ is compact if and only if X is finite.
3. The subspace $Y = \{0\} \cup \{1/n \mid n = 1, 2, \ldots\} \subset \mathbb{R}$ is compact in the subspace topology. To see this, take a cover $\{U_\alpha\}_{\alpha\in I}$ of Y. There is at least one $U_{\alpha'}$ with $0 \in U_{\alpha'}$. Under the standard topology of $\mathbb{R}$, then there is some $N > 0$ such that $U_{\alpha'}$ also contains all of the points $1/n$ for $n > N$. We then let U_j denote an element of the cover such that $1/j \in U_j$

for all $j = 1, \ldots, N$. Thus we can form a finite subcover

$$\{U_1, \ldots, U_N, U_{\alpha'}\} \subset \{U_\alpha\}_{\alpha \in I}$$

and so the space is compact.

Lemma 3.7.3. *Let $(Y, \mathcal{T}_Y)$ be a subspace of $(X, \mathcal{T})$. Then Y is compact if and only if every covering of Y by open sets in X contains a finite subcollection that covers Y.*

Proof. Suppose Y is compact and $\{U_\alpha\} \subset \mathcal{T}$ is a covering of Y. Then $\{U_\alpha \cap Y\}_{\alpha \in I}$ is an open cover of Y and has a finite subcover $\{U_{\alpha_1} \cap Y, \ldots, U_{\alpha_n} \cap Y\}$. Then $\{U_{\alpha_j}\}_{j=1}^n$ is a finite subcollection of open sets in X that covers Y.

Conversely, suppose every covering of Y by open sets in X contains a finite subcollection that covers Y. Take an open cover $\{V_\alpha\}_{\alpha \in I}$ of Y. By definition $V_\alpha = U_\alpha \cap Y$ for some $U_\alpha \in \mathcal{T}$. So $\{U_\alpha\}_{\alpha \in I}$ is a covering of Y by open sets in X. Applying the assumption, there is a finite subcollection $\{U_{\alpha_1}, \ldots, U_{\alpha_n}\}$ that covers Y. Thus $\{U_{\alpha_j} \cap Y\}_{j=1}^n$ is a finite subcover of Y. $\square$

We can say more about subspaces $(Y, \mathcal{T}_Y)$ if the larger space X is compact. The proof of the following (important) result is an exercise.

Theorem 3.7.4. *If $(X, \mathcal{T})$ is compact and $Y \subset X$ is closed, then $(Y, \mathcal{T}_Y)$ is compact (in the subspace topology).*

We obtain nicer properties still if the space X is Hausdorff.

Theorem 3.7.5. *Let $(X, \mathcal{T})$ be a Hausdorff space.*

1. *Every compact subset Y of X is closed.*
2. *If X is compact, X is normal.*

Proof. (1) We will prove a slightly stronger statement: if $Y \subset X$ is compact, then for every $x \in X \setminus Y$ there are open neighbourhoods U of x and V of Y such that $U \cap V = \emptyset$. Take $x \in X \setminus Y$ and any $y \in Y$. There are disjoint open neighbourhoods U_x and V_y of x and y respectively. Then $\{V_y \mid y \in Y\}$ is a covering of Y by open sets in X. So applying Lemma 3.7.3 there is a finite subcollection $\{V_{y_1}, \ldots, V_{y_n}\}$ that covers Y. We then define

$$V = V_{y_1} \cup \ldots \cup V_{y_n} \supset Y, \qquad U = U_{y_1} \cap \ldots \cap U_{y_n} \ni x$$

and where $U \cap V = \emptyset$. This proves the statement. It also follows that $X \setminus Y$ is open and so Y is closed.

(2) Now suppose that A and B are disjoint closed subsets in X, which are compact by Theorem 3.7.4. Applying the statement from part (1), for each $x \in A$ there are disjoint open neighbourhoods U_x and V_x of x and B respectively. Then $\{U_x\}_{x \in X}$ is a cover of A by open sets in X and we take a finite subcover $\{U_{x_1}, \ldots, U_{x_n}\}$. Then

$$U = U_{x_1} \cup \ldots \cup U_{x_n} \supset A, \qquad V = V_{x_1} \cap \ldots \cap V_{x_n}$$

are open neighbourhoods of A and B with $U \cap V = \emptyset$. Thus $(X, \mathcal{T})$ is normal. □

Compact Hausdorff spaces are very common in analysis. Theorem 3.7.5 shows that, for example, we can apply Urysohn's Lemma and the Tietze Extension Theorem to such spaces.

Proposition 3.7.6. *Let $(X, \mathcal{T})$ be compact and $f : (X, \mathcal{T}) \to (Y, \mathcal{T}')$ a continuous function, then the subspace $f(X) \subset Y$ is compact.*

Proof. Take a covering $\{V_\alpha\}_{\alpha \in I} \subset \mathcal{T}'$ of $f(X)$ by open sets in Y. Then $\{f^{-1}(V_\alpha)\}_{\alpha \in I}$ is an open cover of X with finite subcover $\{f^{-1}(V_{\alpha_1}), \ldots, f^{-1}(V_{\alpha_n})\}$. Then $\{V_{\alpha_1}, \ldots, V_{\alpha_n}\}$ is a cover of $f(X)$ and so $f(X) \subset Y$ is compact by Lemma 3.7.3. □

Exercise 3.16. Show that compactness is preserved under homeomorphisms of topological spaces.

Theorem 3.7.7. *Let $(X, \mathcal{T})$ be compact, $(Y, \mathcal{T}')$ Hausdorff and $f : X \to Y$ a continuous bijection. Then f is a homeomorphism and Y is compact.*

Proof. If $A \subset X$ is closed, then A is compact and $f(A) \subset Y$ is compact and therefore closed by Proposition 3.7.6 and Theorem 3.7.5. Because f is a bijection, this implies that $(f^{-1})^{-1}(A) = f(A)$ is closed for every closed $A \subset X$. This implies that f^{-1} is continuous and therefore f is a homeomorphism. The second statement follows from the previous exercise. □

Theorem 3.7.7 has some surprising consequences. For example, we can consider curves in space via continuous functions $f : [0,1] \to \mathbb{R}^d$. If this function is injective, then $[0,1]$ is homeomorphic to its image $f([0,1]) \subset \mathbb{R}^d$. On the other hand, space-filling curves exist, whose range is the entire unit square of higher dimensional space for example. So the homeomorphism takes a one dimensional interval and fills a two-dimensional subspace.

Exercise 3.17. Write a report about space filling curves.

Recall that for metric spaces (X, d) is compact if and only if every sequence $\{x_n\}_{n \geq 1}$ in X has a convergent subsequence. This result is no longer true for topological spaces, but can be amended by replacing sequences with nets.

Theorem 3.7.8. *A topological space $(X, \mathcal{T})$ is compact if and only if every net in X has a convergent subnet.*

Proof. (Advanced) exercise. □

The Tychonoff Theorem

We now investigate the compactness of product spaces. We will do this in steps and first consider finite products.

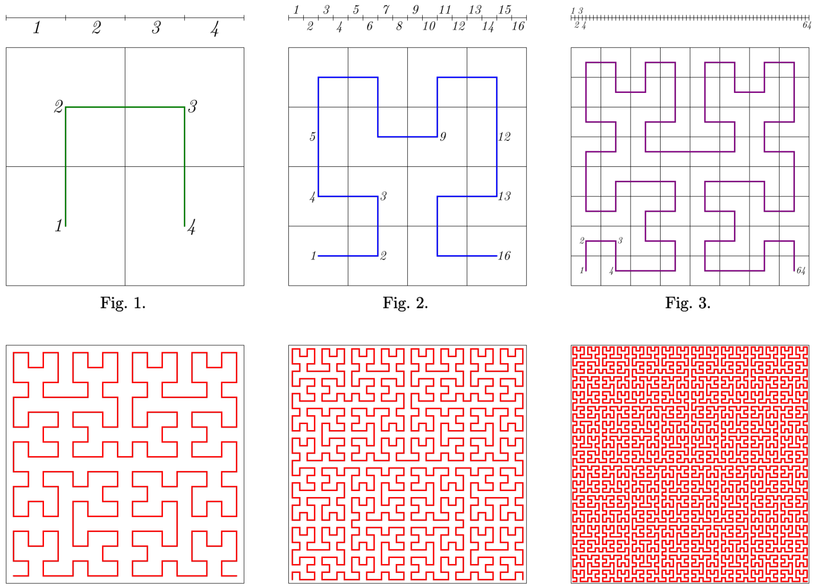


Figure 2: Construction of the Hilbert curve, whose limit fills the unit square in $\mathbb{R}^2$. Image from Wikipedia.

Exercise 3.18. Let $(X, \mathcal{T})$ be a topological space and $\mathcal{B}$ a base for the topology on X. If every cover of X by elements in $\mathcal{B}$ has a finite subcover, show that X is compact.

Lemma 3.7.9 (Tube lemma). *Let X and Y be topological spaces with Y compact. If $N \subset X \times Y$ is an open set containing the slice $\{x_0\} \times Y \subset X \times Y$, then N contains a tube $W \times Y$, where W is an open neighbourhood of x_0 in X.*

Proof. We can take a cover of $\{x_0\} \times Y$ by basis elements $\{U_\alpha \times V_\alpha \subset N\}_{\alpha \in I}$. Because $\{x_0\} \times Y$ is homeomorphic to Y, it is compact and so there is a finite subcover $\{U_1 \times V_1, \dots, U_n \times V_n\} \subset N$. Then $W = U_1 \cap \dots \cap U_n$ is an open neighbourhood of x_0 in X. By construction $\{U_1 \times V_1, \dots, U_n \times V_n\} \subset N$ is a cover of $W \times Y$. Hence $W \subset \bigcup_{j=1}^n U_j \times V_j \subset N$. $\square$

Theorem 3.7.10 (Tychonoff Theorem, finite case). *The product of finitely many compact spaces is compact.*

Proof. By induction, it will suffice to prove the theorem for the product of two compact spaces. Let X and Y be compact and $\{V_\alpha\}_{\alpha \in I}$ an open covering of $X \times Y$. Take $x_0 \in X$ and the slice $\{x_0\} \times Y$, which is compact and can be covered by finitely many $\{V_{\alpha_1}, \dots, V_{\alpha_m}\}$. The union $N = \bigcup_{j=1}^m V_{\alpha_j}$ is an open set containing $\{x_0\} \times Y$. Applying the tube lemma, N contains a tube $W \times Y$ with $W \subset X$ open. In particular $\{V_{\alpha_1}, \dots, V_{\alpha_m}\}$ is a covering of $W \times Y$ by open sets in $X \times Y$.

For each $x \in X$, we can therefore choose an open neighbourhood W_x of x such that the tube $W_x \times Y$ can be covered by finitely many elements from $\{V_\alpha\}_{\alpha \in I}$. Then $\{W_x\}_{x \in X}$ is an open cover

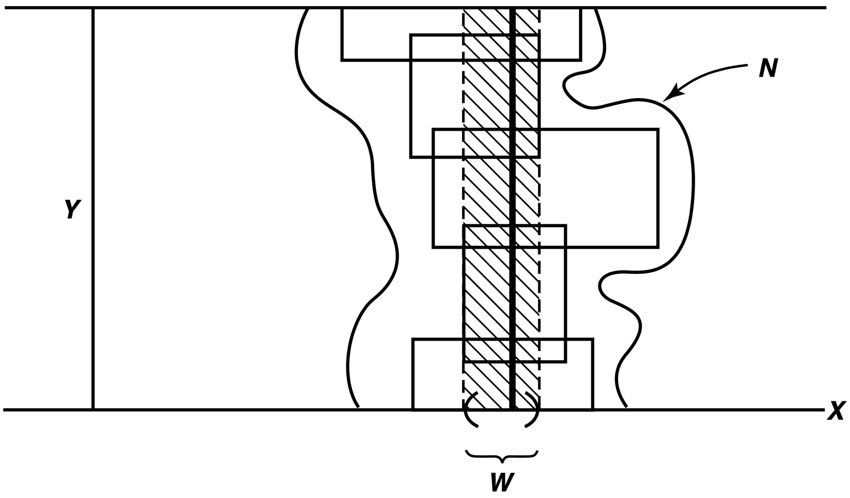


Figure 3: Visualisation of the tube lemma [6, Figure 26.2].

of X with finite subcover $\{W_{x_1}, \dots, W_{x_n}\}$. Hence

$$\bigcup_{k=1}^{n} W_{x_k} \times Y = X \times Y$$

and each $W_{x_k} \times Y$ is covered by finitely many elements from $\{V_\alpha\}_{\alpha \in I}$. Combining all these finite covers, we obtain a finite subcover of $X \times Y$. $\square$

The case of arbitrary products of compact spaces is a much harder and deeper result. The proof relies on Zorn's Lemma, which is a not-very-intuitive result that turns out to be equivalent to the axiom of choice, which is more intuitive and taken to be tautologically true.

Definition 3.7.11. A partial order on a set X is a relation $R \subset X^2$ that is reflexive (xRx for all $x \in X$), transitive (xRy and yRz implies that xRz) and anti-symmetric (xRy and yRx implies that $x = y$).
If R is a partial order on X such that for all $(x, y) \in X^2$, either xRy or yRx, then we say that X is linearly ordered by R.

Example 3.7.12. Let Y be a set and $\mathcal{P}(Y)$ the power set (set of all subsets). We can define a partial order on $\mathcal{P}(Y)$ where $A \in \mathcal{P}(Y)$ is in relation to $B \in \mathcal{P}(Y)$ if $A \subset B$.

Definition 3.7.13. Let R be a partial order on X. If there is some $M \in X$ such that MRx implies that $x = M$, then M is called a maximal element. If $Y \subset X$ and there is some $p \in X$ such that yRp for all $y \in Y$, then p is called an upper bound for Y.

Lemma 3.7.14 (Zorn's Lemma)**.** *If X is a nonempty partially ordered set such that every linearly ordered subset has an upper bound in X, then each linearly ordered set has some upper bound which is also a maximal element of X.*

Exercise 3.19. Use Zorn's Lemma to prove that every vector space has a basis.

Theorem 3.7.15 (Tychonoff Theorem). *Let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha \in I}$ be a family of compact topological spaces. Then $\prod_{\alpha \in I} X_\alpha$ is compact.*

Very roughly, one way to prove the Tychonoff theorem is to assume it is not true and construct a partially ordered set, apply Zorn's lemma to obtain a maximal element and use this element to derive a contradiction. We will omit the details. The interested reader can consult [4, Chapter 12] or [6, Section 37]. It turns out that Tychonoff implies and therefore is equivalent to the axiom of choice. Like infinite product spaces, the Tychonoff theorem appears in many different fields and is an extremely useful analytic result.

3.8 Locally compact spaces and compactification

We have already seen that $\mathbb{R}$ and $\mathbb{R}^n$ are not compact spaces, but for any point $x \in \mathbb{R}^n$, we can take a closed ball $\overline{B(x;r)}$ which is compact. That is, every point has a compact neighbourhood. We generalise this notion to other topological spaces.

Definition 3.8.1. A topological space $(X, \mathcal{T})$ is locally compact if for all $x \in X$, there is an open neighbourhood U of X and compact subspace C such that $x \in U \subset C$.

Examples 3.8.2.

1. Any compact space is locally compact.
2. Euclidean space $\mathbb{R}^n$ is locally compact as is any open or closed subset (see the exercise below).
3. The rational numbers $\mathbb{Q}$ as a subspace of $\mathbb{R}$ is *not* locally compact. To see this, consider a closed neighbourhood of $x \in \mathbb{Q}$ of the form $[a,b] \cap \mathbb{Q}$. On the other hand, we can take a Cauchy sequence $\{y_n\}$ in $[a,b] \cap \mathbb{Q}$ that converges to an irrational number $y \notin \mathbb{Q}$. So $[a,b] \cap \mathbb{Q}$ is not complete and therefore not compact by Theorem 2.4.7 (as $\mathbb{Q}$ is a metric space). Because $\mathbb{Q}$ is Hausdorff, this is enough to show it is not locally compact (see Theorem 3.8.6 below).

Exercise 3.20. Show that any open or closed subset of $\mathbb{R}^n$ is locally compact.

It is easy to think of compact spaces as 'small' and locally compact spaces as 'large'. But this is not always the case.

Exercise 3.21. Suppose that Y is a compact Hausdorff space and take $y_0 \in Y$. Show that the subspace $X = Y \setminus \{y_0\}$ is locally compact in the subspace topology.

It turns out that any locally compact Hausdorff space can be put in the form of the example given in Exercise 3.21.

Theorem 3.8.3. *Let $(X, \mathcal{T})$ be locally compact and Hausdorff. Define the set $X^+ = X \cup \{\infty\}$ with*

$$\mathcal{T}_{X^+} = \mathcal{T} \cup \left\{ X^+ \setminus C \mid C \subset X \text{ a compact subspace} \right\}.$$

Then $(X^+, \mathcal{T}_{X^+})$ is a compact Hausdorff topological space such that the subspace topology of $X \subset X^+$ coincides with $\mathcal{T}$. Furthermore, if X is not compact, then X is dense in X^+.

Proof. We first need to check that $\mathcal{T}_{X^+}$ is a topological space. We have $\emptyset \in \mathcal{T} \subset \mathcal{T}_{X^+}$ and $X^+ = X^+ \setminus \emptyset \in \mathcal{T}_{X^+}$. For unions and intersections, we check the cases that $U \in \mathcal{T}$ or $V = X^+ \setminus C$. Namely

$$
\begin{aligned}
&\bigcup_{\alpha \in I} U_\alpha = U \in \mathcal{T}, \\
&\bigcup_{\beta \in J} X^+ \setminus C_\beta = X^+ \setminus (\bigcap_\beta C_\beta) = X^+ \setminus C, \\
&(\bigcup_\alpha U_\alpha) \cup (\bigcup_\beta X^+ \setminus C_\beta) = U \cup (X^+ \setminus C) = X^+ \setminus (C \setminus U) = X^+ \setminus C',
\end{aligned}
$$

where $C' = C \setminus U$ is compact as it is a closed subspace of C. Similarly,

$$
\begin{aligned}
&U_1 \cap U_2 = U \in \mathcal{T}, \\
&(X^+ \setminus C_1) \cap (X^+ \setminus C_2) = X^+ \setminus (C_1 \cup C_2) = X^+ \setminus C, \\
&U \cap (X^+ \setminus C) = U \cap (X \setminus C) = U \cap U' \in \mathcal{T}.
\end{aligned}
$$

So $(X^+, \mathcal{T}_{X^+})$ is indeed a topological space.

Consider X with the relative topology as a subset $X \subset X^+$. Then $X \cap U = U$ if $U \in \mathcal{T}$ and $X \cap (X^+ \setminus C) = X \setminus C \in \mathcal{T}$ as $C \subset X$ is closed. So the subspace topology is contained in $\mathcal{T}$ and certainly the reverse inclusion holds. Therefore $X \subset X^+$ coincides with $(X, \mathcal{T})$.

We now show that X^+ is compact. Take an open cover $\{V_\alpha\}_{\alpha \in I}$ of X^+ there is at least one $\alpha' \in I$ such that $V_{\alpha'} = X^+ \setminus C$ as sets in $\mathcal{T}$ will not contain the point ∞. Then

$$
\{X \cap V_\alpha\}_{\alpha \in I \setminus \{\alpha'\}}
$$

is a collection of open sets in X that will cover C. Because C is compact, there is a finite subcollection $\{X \cap V_1, \ldots, X \cap V_n\}$ that will cover C. Then if we include the set $V_{\alpha'}$, the collection $\{V_1, \ldots, V_n, V_{\alpha'}\}$ will be a finite subcover of X^+. Hence X^+ is compact.

Next we show that X^+ is Hausdorff. Taking $x \neq y \in X^+$, if $x, y \in X$, then there are disjoint neighbourhoods U and V in $\mathcal{T}$ as X is Hausdorff. Suppose then that $x \in X$ and $y = \infty$. Because X is locally compact, there is a neighbourhood U of x and compact set C such that $x \in U \subset C$. Then take $V = X^+ \setminus C$. We see that $U \cap V = U \cap (X^+ \setminus C) = \emptyset$ as $U \subset C$.

Lastly to show that X is dense in X^+ it suffices to show that any open neighbourhood of ∞ has a non-trivial intersection with X. Any open neighbourhood of ∞ is of the form $X^+ \setminus C$ and $X \cap (X^+ \setminus C) = X \setminus C$. If X is not compact, then $C \neq X$ and so $X \setminus C \neq \emptyset$. Hence $\overline{X} = X^+$ as required. □

Definition 3.8.4. Let $(X, \mathcal{T})$ be a topological space. A compactification of X is a continuous injection (embedding) $\iota : X \to Y$ such that Y is compact and $\iota(X)$ is dense in Y.

Often we identify X with $\iota(X)$ and think of a compactification as a compact topological space Y such that $\overline{X} = Y$. We have seen that for the case of X locally compact and Hausdorff, a compactification $X^+ = X \cup \{\infty\}$ always exists, which we call the *one-point compactification.* It is the smallest possible compactification of X.

Example 3.8.5. If $X = (0, 1)$, then $[0, 1]$ is a compactification of X. This compactification is different from the one-point compactification. Indeed, $X^+ \cong S^1$, the circle.

Exercise 3.22. Show that $(\mathbb{R}^n)^+ \cong S^n = \{(x_1, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \ldots + x_{n+1}^2 = 1\}$.

Theorem 3.8.6. *Let $(X, \mathcal{T})$ be a Hausdorff space.*

1. *X is locally compact if and only if for any $x \in X$ and open neighbourhood U of x, there is an open neighbourhood V of x such that $\overline{V} \subset U$ and $\overline{V}$ is compact.*

2. *If X is locally compact, then any open or closed subset is locally compact in the subspace topology.*

Proof. (1) If the new condition holds, then x has a compact neighbourhood and X is locally compact. Suppose then that X is locally compact and take $x \in X$ and a neighbourhood $U \in \mathcal{T}$ with $x \in U$. Now consider $X \subset X^+$ and let $C = X^+ \setminus U$, which is a closed and therefore compact subspace of X^+. Recalling Theorem 3.7.5, we can take open neighbourhoods V of x and W of C such that $V \cap W = \emptyset$. Then $\overline{V}$ is a compact subset of X^+ with $\overline{V} \cap C = \emptyset$. Furthermore $\infty \notin \overline{V}$ and so $\overline{V}$ is compact in X with $\overline{V} \subset U$ as required.

(2) Let $A \subset X$ be closed and take $x \in A$ with compact neighborhood $x \in U \subset C$. Then $C \cap A$ is closed in C and therefore a compact set containing $U \cap A$, an open neighbourhood of x in A. Hence A is locally compact (the Hausdorff property was not used here).

Suppose $A \subset X$ is open and $x \in A$. By part (1) there is an open neighbourhood V of x such that $x \in \overline{V} \subset A$ and $\overline{V}$ compact. Hence $\overline{V} = A \cap \overline{V}$ is a compact neighbourhood of x that contains the open neighbourhood $A \cap V$. So A is locally compact. ☐

3.9 Connected spaces

If $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ are topological spaces with $X \cap Y = \emptyset$, we can easily make the new topological space $(X \cup Y, \mathcal{T}'')$, where $U \subset X \cup Y$ is open if $U \cap X$ and $U \cap Y$ are open. In this case, the spaces X and Y act independently of each other in $X \cup Y$ and we do not gain additional information by considering the union $X \cup Y$. Put another way, we could decompose $(X \cup Y, \mathcal{T}'')$ into two separate topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ without losing any information.

In this section we formalise this idea to consider connected/disconnected topological spaces and the decomposition of a space into connected components.

Definition 3.9.1. A topological space $(X, \mathcal{T})$ is disconnected if there are non-empty open and closed sets $U, V \subset X$ such that $U \cap V = \emptyset$ and $U \cup V = X$. A space $(X, \mathcal{T})$ is connected if it is not disconnected. We say that $Y \subset X$ is a connected subspace if Y is connected in the subspace topology.

Note that to be disconnected, it suffices to find disjoint and non-empty open sets U and V such that $U \cup V = X$. Then $U = X \setminus V$ and $V = X \setminus U$, so both must also be closed.

Examples 3.9.2.

1. Any set X with the trivial topology $\{\emptyset, X\}$ is connected.

2. Any set X with two or more elements and the discrete topology $\mathcal{T}_{\text{disc}} = \mathcal{P}(X)$ is disconnected.

3. The set $X = [-1, 0) \cup (0, 1] \subset \mathbb{R}$ is disconnected as $[-1, 0)$ and $(0, 1]$ are disjoint open sets (in the subspace topology) that cover X.

Exercise 3.23. Show that the rationals $\mathbb{Q} \subset \mathbb{R}$ are disconnected.

The following lemma is an immediate consequence of the definition (proof as a simple exercise).

Lemma 3.9.3. *A topological space $(X, \mathcal{T})$ is connected if and only if the only closed and open subsets of X are $\emptyset$ and X.*

Proposition 3.9.4. *If $(X, \mathcal{T})$ is connected and $f : X \to Y$ is continuous, then $f(X) \subset Y$ is connected.*

Proof. Let $A \subset f(X)$ be closed and open. Then $f^{-1}(A) \subset X$ is closed and open, hence $f^{-1}(A) = \emptyset$ or $f^{-1}(A) = X$. In the first case, $A = \emptyset$. In the second case, $A = f(X)$. $\square$

If X and Y are connected and disjoint, then certainly $X \cup Y$ is disconnected. However, if $X \cap Y \neq \emptyset$, then connectedness is preserved, as the following shows.

Theorem 3.9.5. *Let $\{A_\alpha\}_{\alpha \in I}$ be a family of connected subsets of a topological space $(X, \mathcal{T})$ such that $A_\alpha \cap A_\beta \neq \emptyset$ for any pair of indices $\alpha, \beta \in I$. Then the union $\bigcup_\alpha A_\alpha$ is a connected subset.*

Proof. Let $A = \bigcup_\alpha A_\alpha$ and $B \subset A$ closed and open. If $B = \emptyset$, there is nothing to show so we assume $B \neq \emptyset$. Take $x \in B \subset A$. Hence $x \in A_{\alpha'}$ for some $\alpha' \in I$. Then $B \cap A_{\alpha'}$ will be a non-empty closed and open subset of $A_{\alpha'}$. Since $A_{\alpha'}$ is connected, it follows that $B \cap A_{\alpha'} = A_{\alpha'}$ and so $A_{\alpha'} \subset B$. Now for any $\alpha \in I$, $B \cap A_\alpha$ is a closed and open subset that contains $A_{\alpha'} \cap A_\alpha \neq \emptyset$. By the same argument, $B \cap A_\alpha = A_\alpha$ and $A_\alpha \subset B$ for all $\alpha \in I$. It therefore follows that $B = \bigcup_\alpha A$, which gives the result. $\square$

Theorem 3.9.6. *Any interval $I \subset \mathbb{R}$ is connected.*

Proof. We will show that $[a, b]$ is connected and leave the rest as an exercise. (Recall that connectedness is preserved under homeomorphism and, from Exercise 3.4, any interval is homeomorphic to one of the following: $[0, 1]$, $(0, 1)$, $[0, 1)$, $(0, 1]$.) Suppose that $E \subset [a, b]$ is closed, open and non-empty. We also assume $b \notin E$ (otherwise, take $[a, b] \setminus E$, which is also closed, open and non-empty) and let $c = \sup\{x \in E\}$. Since E is closed, $c \in E$ and $c \neq b$. Since E is open, $[c, c + \epsilon) \subset E$ for some $\epsilon > 0$. But this contradicts the fact that c is the supremum of elements in E. Hence such a set $E \subset [a, b]$ cannot exist and therefore $[a, b]$ is connected. $\square$

Exercise 3.24. If $Y \subset X$ is a connected subspace, show that the closure $\overline{Y}$ is a connected subspace.

Let us come back to the question of decomposing a topological space into connected subspaces.

Definition 3.9.7. Let $(X, \mathcal{T})$ be a topological space and $x \in X$. The connected component of x in X is the union

$$C_x = \bigcup \{Y \subset X \mid Y \text{ connected and } x \in Y\}$$

It follows from Theorem 3.9.5 that C_x is connected for any $x \in X$ (as each connected subset Y containing x has a non-empty intersection, $\{x\}$). Therefore we can decompose our space

$$X = \bigcup_{x \in X} C_x$$

into a union of connected subspaces. As the following shows, this gives a *partition* of the space X.

Theorem 3.9.8. *Let $(X, \mathcal{T})$ be a topological space and $x, y \in X$. Then either $C_x \cap C_y = \emptyset$ or $C_x = C_y$.*

Proof. Suppose that $C_x \cap C_y \neq \emptyset$. Then by Theorem 3.9.5, $C_x \cup C_y$ is connected. But then $C_x \cup C_y$ is a connected set that contains x and y. Therefore $C_x \cup C_y \subset C_x$ and $C_x \cup C_y \subset C_y$. It follows that $C_x = C_x \cup C_y = C_y$. $\square$

Any partition of a space comes from an equivalence relation (see Appendix A). So we also have the following.

Corollary 3.9.9. *Let $(X, \mathcal{T})$ be a topological space. Say that $x \sim_c y$ if $y \in C_x$. Then $\sim_c$ is an equvalence relation on X.*

Examples 3.9.10.

1. Suppose that X is given the discrete topology, then $X = \bigcup_{x \in X} \{x\}$ is a decomposition into connected components. More generally, we say that a topological space is *totally disconnected* if the only connected components are singletons. For example, the Cantor set is totally disconnected.

2. Consider $X = \mathbb{Q} \subset \mathbb{R}$. If $q_1, q_2 \in \mathbb{Q}$ are distinct, say $q_1 < q_2$, then there is an irrational number $r \in \mathbb{R}$ such that $q_1 < r < q_2$. Then $U = \mathbb{Q} \cap (-\infty, r)$ and $V = \mathbb{Q} \cap (r, \infty)$ are disjoint open subsets of $\mathbb{Q}$ with $q_1 \in U$ and $q_2 \in V$. It follows that $q_2 \notin C_{q_1}$. This shows that the connected components of $\mathbb{Q}$ are singletons. So $\mathbb{Q}$ is also totally disconnected, but the connected components $\{q\}$ are *not* open.

3.10 Path connected spaces

We intuitively think of disconnected spaces as those which can be broken up into pieces. Another way we can consider the connectedness of a space is whether we can draw a curve between points that does not leave the space.

Definition 3.10.1. Let $(X, \mathcal{T})$ be a topological space with $x_0, x_1 \in X$. A path in X from x_0 to x_1 is a continuous function $\gamma : [0, 1] \to X$ such that $\gamma(0) = x_0$, $\gamma(1) = x_1$. We say that X is path connected if for any pair $x_0, x_1 \in X$, there is a path in X from x_0 to x_1.

Examples 3.10.2.

1. The circle $S^1 \subset \mathbb{R}^2$ is path connected as we can move along the circle $(r\cos(\theta), r\sin(\theta))$ to go between any two points.

2. Recall that a subset $E \subset \mathbb{R}^n$ is convex if for any $\mathbf{x}, \mathbf{y} \in E$, the straight line $(1-t)\mathbf{x} + t\mathbf{y} \in E$ for all $t \in [0, 1]$. Hence any convex set is path connected. In particular, the ball $B(\mathbf{x}, r)$ is path-connected for all $\mathbf{x} \in \mathbb{R}^n$, $r > 0$.

3. The set $X = [-1, 0) \cup (0, 1]$ is not path connected as there is no continuous path from -1

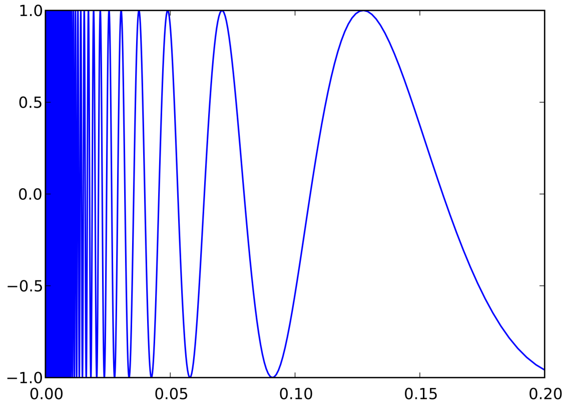


Figure 4: The topolgist's sine curve. Image from Wikipedia.

to 1 that stays within X.

Exercise 3.25. Show that the (finite or infinite) product of path connected spaces is path connected.

Theorem 3.10.3. *Any path connected space is connected.*

Proof. Fix a point $x_0 \in X$ and take $x \in X$. By assumption there is a path $\gamma_x : [0,1] \to X$ from x_0 to x in X. Because $[0,1]$ is connected and γ_x is continuous $\gamma_x([0,1]) \subset X$ connected. Furthermore $x_0 \in \gamma_x([0,1])$ for all $x \in X$. So $\{\gamma_x([0,1])\}_{x\in X}$ is a family of connected subsets of X with non-empty intersection. Hence by Theorem 3.9.5, $\bigcup_x \gamma_x([0,1]) = X$ is connected. $\square$

Unfortunately, the converse does not hold.

Example 3.10.4 (The topologist's sine curve)**.** First consider the space

$$S_0 = \big\{(x, \sin(x^{-1}) \mid x \in (0,1]\big\} \subset \mathbb{R}^2.$$

Now the interval $(0,1]$ is connected and we can write $S_0 = [0,1) \times f\big((0,1]\big)$ for $f(x) = \sin(x^{-1})$, which is continuous on $(0,1]$. Hence S_0 is connected and therefore so is $S = \overline{S_0} = S_0 \cup \{0\} \times [-1,1]$. Suppose now that S is path-connected. Then there is a continuous path $\gamma : [0,1] \to S$ from the point $(0,0)$ to $(1, \sin(1))$. Now $\gamma^{-1}(\{0\} \times [-1,1])) \subset [0,1]$ must be closed, so it has a maximum element $b \in [0,1]$. Then $\gamma|_{[b,1]}$ is a path such that $\gamma(b) \in \{0\} \times [-1,1]$ and $\gamma(t) \in S_0$ for all $t > b$. If we write $\gamma(t) = (x(t), y(t)) \in S$, then $x(0) = 0$ and $x(t) > 0$ for $t > b$ and $y(t) = \sin(x(t)^{-1})$ for $t > b$. Now for any $n \in \mathbb{N}$ we can take $b < u < x(b + \frac{1}{n})$ such that $\sin(u^{-1}) = (-1)^n$. Then by the intermediate value theorem there must be some $b < t_n < b + \frac{1}{n}$ such that $x(t_n) = u$. Then t_n is a sequence such that $t_n \to b$, but where $\sin(t_n^{-1}) = (-1)^n$ does not converge. This contradicts the fact that γ is a continuous function. Thus our assumption was incorrect and so S is not path connected.

Like we did with connected subspaces, we can also consider the path-connected components of a space X.

Lemma 3.10.5. *Let $(X, \mathcal{T})$ be a topological space. The relation $x \sim_\gamma y$ if there is a path $\gamma : [0,1] \to X$ such that $\gamma(0) = x$, $\gamma(1) = y$ is an equivalence relation.*

Proof. For reflexivity, take the constant path $\gamma(t) = x$. If $x \sim_\gamma y$, then we can define the reverse path $\gamma^- : [0,1] \to X$ such that $\gamma^-(t) = \gamma(1-t)$. Then $\gamma^-(0) = y$ and $\gamma^-(1) = x$. So the relation is symmetric. For transitivity, we suppose $x \sim_\alpha y$ and $y \sim_\beta z$ and concatenate the paths,

$$\gamma : [0,1] \to X, \qquad \gamma(t) = \begin{cases} \alpha(2t), & t \in [0, \frac{1}{2}], \\ \beta(2t-1), & t \in [\frac{1}{2}, 1]. \end{cases}$$

Then γ is continuous (as α and β are continuous with $\alpha(1) = \beta(0)$) with $\gamma(0) = x$ and $\gamma(1) = z$. □

Hence we can partition any topological space

$$X = \bigcup_{x \in X} \{y \in X \mid x \sim_\gamma y\} = \bigcup_{x \in X} C_x^\gamma$$

Note that by Theorem 3.10.3, the path connected component at $x \in X$ is such that $C_x^\gamma \subset C_x$, the connected component. The converse need not hold. Indeed, the topologist sine curve S has one connected component, S, and two path-connected components, S_0 and $\{0\} \times [-1, 1]$.

Exercise 3.26. We say that $(X, \mathcal{T})$ is *locally* path connected if for all $x \in X$ and open neighbourhood $x \in U \in \mathcal{T}$, there is a path-connected open neighbourhood $x \in V \in \mathcal{T}$ with $V \subset U$. Show that if X is locally path connected, then $C_x = C_x^\gamma$ for all $x \in X$.

3.11 Quotient spaces

So far we have seen a few ways of constructing new topological spaces from existing ones. The subspace topology is one example, product spaces are another. In this section we introduce a new method for building a new space from an existing one, where we quotient by an equivalence relation.

Definition 3.11.1. Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces and $q : X \to Y$ a surjective function. We say that q is a quotient map if $V \subset Y$ is open if and only if $q^{-1}(V) \in X$ is open.

Every quotient map is continuous, but not every continuous surjection is a quotient map.

Proposition 3.11.2. *If $(X, \mathcal{T})$ is a topological space, Y is a set and $q : X \to Y$ a surjective map, then there is exactly one topology on Y such that q is a quotient map, called the quotient topology.*

Proof. We define the topology as follows,

$$\mathcal{T}_q = \{U \subset Y \mid q^{-1}(U) \subset X \text{ is open}\}.$$

It is a simple exercise to check that $\mathcal{T}_q$ is a topology on Y. By construction, $\mathcal{T}_q$ is such that $q : X \to Y$ is a quotient map. For uniqueness, any other topology $\mathcal{T}'$ on Y such that q is a quotient map must be such that $\mathcal{T}' \subset \mathcal{T}_q$. □

Hence if $(X, \mathcal{T})$ is a topological space and $q : X \to Y$ is surjective, we can put a topology on Y such that q is a quotient map. Of particular interest to us will be the case where Y is the set of equivalence classes of an equivalence relation on X.

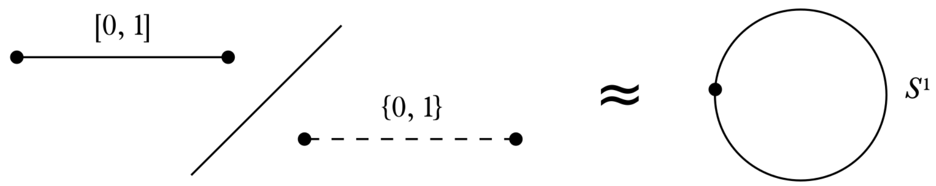
Definition 3.11.3. Let X be a set and $\sim$ an equivalence relation on X. We define the set

$$X/\!\sim \;=\; \big\{[x] \mid x \in X\big\}, \qquad [x] = \{y \in X \mid x \sim y\}.$$

If X is a topological space, then $X/\!\sim$ with the quotient topology under the surjection $q : X \to X/\!\sim$, $q(x) = [x]$ is called a quotient space.

Example 3.11.4. Let $X = [0,1] \subset \mathbb{R}$ with subspace topology. If we take $\sim$ to be the usual $=$, then $x \sim y$ if and only if $x = y$ and so $X/\!\sim \,\cong X$. To obtain a non-trivial quotient space, we are interested in equivalence relations such that $[x] \neq \{x\}$ in general.

Let us therefore consider the equivalence relation $\sim$ on $[0,1]$, where $\sim$ is $=$ on $(0,1)$ and we say that $0 \sim 1$. That is, we say the endpoints of the interval are equivalent. Hence $[0] = [1] = \{1, 0\}$ as equivalence classes. Geometrically, we consider the map $[0,1] \to [0,1]/\!\sim$ as identifying the points 0 and 1 and we attach or glue them together. If we consider the interval as living inside $\mathbb{R}^2$, then we turn the interval $[0,1]$ into the circle S^1. We will make this equivalence more precise below.



Quotient of the interval $[0,1]/\!\sim$. Image from Wikipedia.

We list some useful properties of quotient spaces.

Theorem 3.11.5. *Let $(X, \mathcal{T})$ be a topological space, $\sim$ an equivalence relation on X and $q : X \to X/\!\sim$.*

1. *A function $g : X/\!\sim \,\to Y$ is continuous if and only if $g \circ q : X \to Y$ is continuous.*
2. *If $f : X \to Y$ is continuous and such that f is constant on each equivalence class, then there is a continuous function $g : X/\!\sim \,\to Y$ such that $f = g \circ q$.*

Proof. (1) The map q is continuous by construction, so continuity of g implies continuity of $g \circ q$. Conversely, suppose that $g \circ q$ is continuous and let $V \subset Y$ be open. Then $(g \circ q)^{-1}(V) = q^{-1}(g^{-1}(V))$ is open in X. By the definition of the quotient topology, $g^{-1}(V)$ must be open in $X/\!\sim$ and so g is continuous.

(2) Because $x \sim y$ implies that $f(x) = f(y)$, the map $g([x]) = f(x)$ is well defined and we see that $f(x) = g \circ q(x)$ by construction. The continuity of g then follows by part (1). $\square$

We also note the following useful result.

Proposition 3.11.6. *Let X be compact, Y Hausdorff and $f : X \to Y$ continuous. Define an equivalence relation on X by declaring $x_0 \sim x_1$ if and only if $f(x_0) = f(x_1)$. Then $X/\sim$ is homeomorphic to $f(X) \subset Y$.*

Proof. Since $X/\sim = q(X)$ and q is continuous, $X/\sim$ is compact. By construction, we have that $x_0 \sim x_1$ implies that $f(x_0) = f(x_1)$ and we satisfy the conditions of part (2) of Theorem 3.11.5. Let $g : X/\sim \to f(X) \subset Y$ be the induced map such that $g = q \circ f$. Then g is a continuous bijection by construction. As $X/\sim$ is compact and Hausdorff, this is enough to guarantee that g is a homeomorphism. □

Example 3.11.7. Consider the circle $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ and define a function $f : [0, 1] \to S^1$, $f(x) = e^{2\pi i x}$. The function f is clearly a continuous surjection that is injective on $(0, 1)$ and has $f(0) = f(1)$. Because f is Hausdorff, we can apply Proposition 3.11.6 to conclude that the induced map $g : [0, 1]/\{0, 1\} \to S^1$ is a homeomorphism.

Exercise 3.27. Extend the example above to construct an homeomorphism between a quotient of $[0, 1]^n$ with the n-torus $\mathbb{T}^n$.

Exercise 3.28. Show that if X is compact / connected / path connected, then $X/\sim$ is compact / connected / path connected.

Example 3.11.8 (Gluing along a function)**.** Let X and Y be distinct topological spaces, i.e. $X \cap Y = \emptyset$. We also take $Z \subset X$ and $f : Z \to Y$ a continuous function. The union $X \cup Y = X \sqcup Y$ is a topological space (where we take unions of open sets in X and Y). The then define an equivalence relation, where we identify points $z \in Z$ with $f(z) \in Y$. The resulting space

$$X \cup_{Z_f} Y = X \sqcup Y/\{z \sim f(z)\}$$

is the gluing of X to Y by on the subset Z by the function f.

More generally still, we can consider A, X and Y topological spaces with $f : A \to X$ and $g : A \to Y$ continuous functions. We can then define

$$X \cup_{A_{f,g}} Y = X \sqcup Y/\{f(a) \sim g(a)\}$$

which is the gluing of X to Y via the space A and the functions f and g.

Exercise 3.29. Let $i_X : X \to X \cup_{A_{f,g}} Y$ and $i_Y : Y \to X \cup_{A_{f,g}} Y$ be the natural inclusion maps. Show that $i_X \circ f(a) = i_Y \circ g(a)$ for all $a \in A$. In other words, the following diagram commutes:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"] & X \arrow[d, "i_X"] \\
Y \arrow[r, "i_Y"] & X \cup_{A_{f,g}} Y.
\end{tikzcd}$$

Example 3.11.9. Let $D^2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1\}$ be the unit disc with boundary $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$. We can consider the space D^2/S^1, meaning that we identify all point $z \sim z'$ if $z \in S^1$. Another way to think of this quotient is to take $X = D^2$, $Y = \{*\}$ (a point) and then we glue D^2 to the point by the function $f : S^1 \to \{*\}$ (where $f(z) = *$ for all $z \in S^1$). We claim the quotient $D^2/S^1 \simeq S^2$. One way to see this equivalence is to note that D^2 is compact, so D^2/S^1 is compact. So D^2/S^1 is a compact set that comes from adding a point to $D^2 \setminus S^1 = B(\mathbf{0}, 1) \simeq \mathbb{R}^2$. So we can think of D^2/S^1 as the one-point compactification of $B(\mathbf{0}, 1) \simeq \mathbb{R}^2$, which is homeomorphic to S^2 via the stereographic projection.

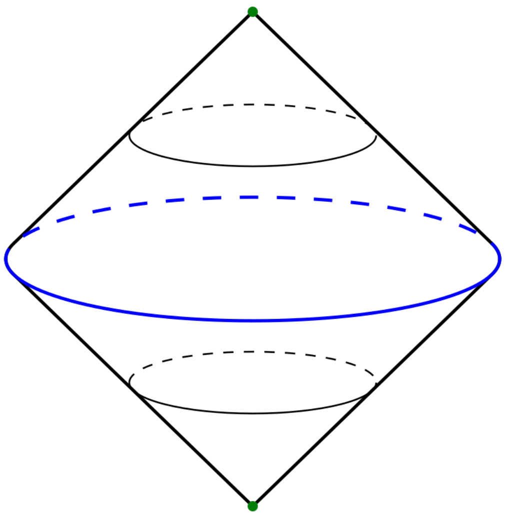
Example 3.11.10 (Cones and (unreduced) suspension)**.** Let I denote the interval $[0, 1]$. If X is a topological space, we can consider the product $X \times [0, 1]$ as the 'cyclinder' of X. The cone of a topological space is defined via a quotient of the cylinder,

$$CX := X \times I/(X \times \{0\}).$$

The (unreduced) suspension of a topological space is given by gluing two cones C_+X and C_-X along X. Put another way, we consider the cylinder $X \times I$ and collapse both ends.

$$\Sigma X = \frac{X \times I}{\{(x, 0) \sim (y, 0)\} \cup \{(x, 1) \sim (y, 1)\}}.$$

We see that $\Sigma\{\text{pt}\} \simeq I$



The suspension of S^1, $\Sigma S^1 \simeq S^2$, [Wikipedia]

Exercise 3.30. Show that there is a homeomorphism $\phi : \Sigma S^n \to S^{n+1}$.

Example 3.11.11 (The projective plane $\mathbb{R}P^2$)**.** We consider the space $\mathbb{R}^3$, where for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $\ell_{\mathbf{y},\mathbf{x}} = \{\mathbf{y} + t\mathbf{x} \mid t \in \mathbb{R}\}$ denotes the line that passes through $\mathbf{y} \in \mathbb{R}^3$ and has direction parallel to $\mathbf{x}$. We restrict our attention to those lines that pass through the origin,

$$\mathbb{R}P^2 = \{\ell_{\mathbf{0},\mathbf{x}} \mid \mathbf{x} \in \mathbb{R}^3\}.$$

We can give $\mathbb{R}P^2$ a metric by defining $d(\ell_{\mathbf{0},\mathbf{x}}, \ell_{\mathbf{0},\mathbf{y}})$ to be the angle of $\ell_{\mathbf{0},\mathbf{x}}$ and $\ell_{\mathbf{0},\mathbf{y}}$ at $\mathbf{0}$ (taking the smaller of the angles so $d(\ell_{\mathbf{0},\mathbf{x}}, \ell_{\mathbf{0},\mathbf{y}}) \in [0, \frac{\pi}{2}]$). We leave it as an exercise to show that this is indeed a metric.

Because we only consider lines that pass through $\mathbf{0}$, all lines will pass through the unit sphere $S^2 \subset \mathbb{R}^3$. So it suffices to consider $\ell_{\mathbf{x}} = \ell_{\mathbf{0},\mathbf{x}}$ with $\mathbf{x} \in S^2$. Furthermore, $\ell_{\mathbf{x}} = \ell_{-\mathbf{x}}$ as the line will connect antipodal points. In particular, we can identify

$$\mathbb{R}P^2 \simeq S^2/\{\mathbf{x} \sim -\mathbf{x}\}.$$

To see this more explicitly, there is obvious map $f : S^2 \to \mathbb{R}P^2$ given by $f(\mathbf{x}) = \ell_{\mathbf{x}}$, which is a two-to-one map, $f(\mathbf{x}) = f(-\mathbf{x})$. Furthermore, $f^{-1}(B(\ell_{\mathbf{x}}; \epsilon)) = B(\mathbf{x}; \epsilon) \cup B(-\mathbf{x}; \epsilon)$ and so f is a continuous and surjective map from a compact space to a Hausdorff space that is constant on the equvalence classes $\mathbf{x} \sim -\mathbf{x}$. It follows that we can identify $S^2/\sim \simeq f(S^2) \simeq \mathbb{R}P^1$ by Proposition 3.11.6.

Exercise 3.31. Write a report about the more general projective space $\mathbb{R}P^n$ (or the complex projective space $\mathbb{C}P^n$), which are defined similarly.

4 Homotopy and the fundamendal group

Now that we have studied the basics of topological spaces, let us come to a more practical question.

How to tell when two spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ are homeomorphic?

For general spaces, particularly those which can not easily be visualised or embedded into Euclidean space (and even if they can), this is a surprisingly difficult question. For example, we saw in in Section 3.7 that the closed interval $[0, 1]$ is homeomorphic to space filling curves $\gamma([0, 1]) \subset \mathbb{R}^n$, something that is certainly not obvious by visual inspection only. In other cases, even if we suspect that two spaces are homeomorphic, writing down an explicit function $f : X \to Y$ may be challenging.

We can instead consider a slightly different question: how to tell when $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ are *not* homeomoprhic? For example, if X is compact / connected / Hausdorff etc. and Y is not, then we know X and Y are topologically distinct. If we return to our geometric intuition, we distinguish a sphere S^2 from a torus $\mathbb{T}^2$ because $\mathbb{T}^2$ has a hole while S^2 does not. That is, we distinguish spaces via combinatorial data (number of holes, number of connected components etc.). This combinatorial data has a more rigid structure which may make concrete computations easier than working within abstract point-set topology.

In this section, we start to consider *algebraic* topology, which takes the information contained in a topological space $(X, \mathcal{T})$ and represents it via an algebraic object such as a group $G(X)$, which is often easier to work and compute with. Perhaps some information of the topological space is lost in this translation, but we should at least expect that $X \simeq Y$ implies that $G(X) \simeq G(Y)$ in an appropriate sense. Put another way, if $G(X)$ is *different* to $G(Y)$, then we can conclude that X and Y are not homeomorphic. This basic idea has proven to be very useful and is still an active area of research today.

For this first section, we focus on the fundamental group group constructed from loops in a topological space. Time permitting, we will consider other constructions in Section 5.

4.1 Group theory basics

In the context of this course, our aim is to use and work with group and not to develop their general theory. So we will only provide a very brief overview of the basics of group theory. A widely used English text is [3], for example.

Definition 4.1.1. A group is a set G with a group operation $* : G \times G \ni (a, b) \mapsto a * b \in G$ such that

1. The operation is associative, $a * (b * c) = (a * b) * c$,
2. There is an identity element, $e \in G$ such that $a * e = e * a = a$ for all $a \in G$,
3. For every $a \in G$, there is an inverse element $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

We say that a group G is abelian if $a * b = b * a$ for all $a, b \in G$.

Lemma 4.1.2. *The group identity and inverse elements are unique.*

Proof. Suppose that e' is also such that $a * e' = e' * a = a$ for all $a \in G$, then $e' = e' * e = e$. If $\tilde{a}$ is such that $\tilde{a} * a = e = a * \tilde{a}$, then

$$\tilde{a} = \tilde{a} * e = \tilde{a} * (a * a^{-1}) = (\tilde{a} * a) * a^{-1} = e * a^{-1} = a^{-1}. \qquad \square$$

Examples 4.1.3.

1. $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{C}$ are all groups under addition. Here the group identity is 0 and the inverse of x is $-x$.
2. $\mathbb{R}$ is *not* a group if we take the operation as multiplication. This is because a multiplicative inverse to 0 does not exist. However $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ is a group under multiplication with $e = 1$. Similarly, $\mathbb{Q}^\times$ and $\mathbb{C}^\times$ are also multiplicative groups. On the other hand, $\mathbb{Z} \setminus \{0\}$ is not a multiplicative group as, for example, $2^{-1} = \frac{1}{2} \notin \mathbb{Z} \setminus \{0\}$.
3. The unit circle $\mathbb{T} = \left\{e^{i\theta} \in \mathbb{C} \mid \theta \in \mathbb{R}\right\}$ is a group under multiplication.
4. The singleton $\{e\}$ is a group, often called the trivial group.
5. Any vector space V is a group with operation given by addition. Put another way, we can turn a vector space into a group by simply forgetting about scalar multiplication.
6. For any n, the set $\mathbb{Z}_n = \left\{0, 1, \ldots, n-1\right\}$ is a group with operation $x * y = x + y \mod n$.
7. The set $\mathrm{Mat}_{n\times m}(\mathbb{R})$ of $(n \times m)$-matrices with real coefficients is a group under addition (indeed, it is a vector space). We can also consider

$$GL_n(\mathbb{R}) = \left\{T \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid \det(T) \neq 0\right\}$$

which is a group under multiplication and group identity as I_n. Indeed, all elements of $GL_n(\mathbb{R})$ are invertible by definition. Because $AB \neq BA$ in general for $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, we see that $GL_n(\mathbb{R})$ is a non-abelian group.
8. Let X be a set and define $\mathrm{Sym}(X) = \left\{f : X \to X \mid f \text{ a bijection}\right\}$. We can give $\mathrm{Sym}(X)$ an operation via composition $f * g = f \circ g$ for $f, g \in \mathrm{Sym}(X)$. Because the composition of bijections and the inverse of a bijection is also a bijection, it is easy to show that $\mathrm{Sym}(X)$ is indeed a group, where the identity $e : X \to X$ is such that $e(x) = x$ for all $x \in X$. We call $\mathrm{Sym}(X)$ the symmetry group, which is not abelian in general.
9. We can refine the previous example slightly to the setting of topological spaces. Namely, if $(X, \mathcal{T})$ is a topological space, then

$$\mathrm{Homeo}(X) = \left\{f : X \to X \mid f \text{ a homeomorphism}\right\}$$

is also a group.

Let us give one more example of a nonabelian group

Example 4.1.4 (Permutation group S_3)**.** We consider the set with n elements $X = \{1, 2, \ldots, n\}$ and let S_n denote the *permutations* of X. Namely, elements of S_n are bijections of X where we also care about the order of the elements. A permutation $\sigma : X \to X$ is often expressed via the notation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}.$$

Let us consider the case of $X = \{1, 2, 3\}$. In this case there are six distinct permutations,

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

The group operation is given by composition of permutations, for example

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Similarly, every element in S_3 is invertible, for example

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

We leave the remaining details as an exercise. Also note that the group S_3 is non-abelian.

Exercise 4.1. Write a report about the the permutation group S_n for $n > 3$.

The permutation group $S_n \subset \mathrm{Sym}(\{1, \ldots, n\})$ is an example of a *subgroup*.

Definition 4.1.5. A subset H of a group G is a subgroup if H is closed under the multiplication and inverse of the group operation of G.

We see that $\mathbb{Z}$ is a subgroup of $\mathbb{Q}$, which is a subgroup of $\mathbb{R}$, which is a subgroup of $\mathbb{C}$.

Definition 4.1.6. Let G and H be groups. A group homomorphism is a function $\phi : G \to H$ such that $\phi(a * b) = \phi(a) * \phi(b)$ for all $a, b \in G$. If a homomorphism ϕ is a bijection, we say that ϕ is an isomorphism. In such a case, we say that G and H isomorphic.

The key feature of a homomorphism is that it respects the group operation, namely $\phi(a * b)$ involves the operation in G whereas $\phi(a) * \phi(b)$ involves the group operation in H.

Examples 4.1.7.

1. Take $\mathbb{R}$ as the group via addition and $\mathbb{T}$ the circle group. We can define

$$\phi : \mathbb{R} \to \mathbb{T}, \qquad \phi(x) = e^{ix} \in \mathbb{T},$$

where $\phi(x + y) = e^{i(x+y)} = e^{ix}e^{iy}$ and so is indeed a group homomorphism. The homomorphism ϕ is surjective but not injective, $\phi(x) = \phi(x + 2\pi k)$ for any $k \in \mathbb{Z}$.

2. Define a map $\phi : \mathbb{Z} \to \mathbb{Z}_n$, where $\phi(m) = m \bmod n$. For any $m_1, m_2 \in \mathbb{Z}$, there are $K_1, K_2 \in \mathbb{N}$ such that

$$m_1 = a + K_1 n, \qquad m_2 = b + K_2 n, \quad a, b \in \{0, 1, \ldots, n-1\}.$$

Therefore

$$\begin{aligned} m_1 + m_2 &= a + b + (K_1 + K_2)n \\ &= (a + b) \bmod n + (K_1 + K_2 + R)n, \quad R = \begin{cases} 0, & a + b \leq n - 1 \\ 1, & a + b \geq n \end{cases} \end{aligned}$$

and we see that

$$\phi(m_1 + m_2) = a + b = a + b \bmod n = (a + b) + (K_1 + K_2)n \bmod n = a \bmod n + b \bmod n$$
$$= \phi(m_1) + \phi(m_2).$$

3. There are many possible homomorphisms from $\mathbb{R}$ to matrices over $\mathbb{R}$. We list a few

$$\phi_1 : \mathbb{R} \to \mathrm{Mat}_{2\times 2}(\mathbb{R}), \qquad \phi_1(x) = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix},$$
$$\phi_2 : \mathbb{R} \to \mathrm{Mat}_{2\times 2}(\mathbb{R}), \qquad \phi_2(x) = \begin{pmatrix} x & x \\ x & x \end{pmatrix},$$
$$\phi_3 : \mathbb{R}^\times \to GL_2(\mathbb{R}) \qquad \phi_3(x) = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix},$$
$$\phi_4 : \mathbb{R}^\times \to GL_2(\mathbb{R}) \qquad \phi_4(x) = \begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}.$$

4. Take $\mathbb{R}$ with addition and $\mathbb{R}_+ = (0, \infty)$ with multiplication. Then

$$\phi : \mathbb{R} \to \mathbb{R}_+, \qquad \phi(x) = e^x$$

is a group homomorphism. Because ϕ is a bijection, we see that $(\mathbb{R}, +)$ and $(\mathbb{R}_+, \times)$ are isomorphic as groups.

We can combine groups with topological spaces to study so-called topological groups.

Definition 4.1.8. A topological group is a topological space $(G, \mathcal{T})$ such that G is also a group and the operations

$$G \times G \ni (a, b) \mapsto a * b \in G, \qquad G \ni a \mapsto a^{-1} \in G$$

are continuous maps for all $a, b \in G$.

Examples 4.1.9. 1. Any normed vector space $(V, \|\cdot\|)$ is a topological group.

2. The group $GL_n(\mathbb{R})$ can be given a topology via the subspace topology of $\mathbb{R}^{n\times n}$. Equipped with this topology, $GL_n(\mathbb{R})$ is a topological group. Actually, group multiplication and inversion in $GL_n(\mathbb{R})$ are *smooth* and so $GL_n(\mathbb{R})$ is a *Lie group*.

Exercise 4.2. Write a report about topological (or Lie) groups.

4.2 Homotopy

In what follows we fix the topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ and consider the set of continuous functions between these spaces.

$$C(X, Y) = \{f : X \to Y \mid f \text{ continuous}\}.$$

Remark 4.2.1. The set $C(X, Y)$ can be made into a topological space via the so-called compact-open topology. Given a compact set $K \subset X$ and an open set $V \subset Y$, then we can consider the subset

$$C(K, V) = \{f \in C(X, Y) \mid f(K) \subset V\} \subset C(X, Y).$$

Then the open sets in $C(X, Y)$ are given by finite intersections and arbitrary unions of sets of the form $C(K, V)$. The resulting topology is called the compact-open topology. We leave the details as an exercise/report.

Our aim is make precise the notion of being able to deform one space into another. We will do this by first considering continuous functions $f, g : X \to Y$ and whether it is possible to deform/interpolate one function into another. In order to interpolate, we need an extra deformation parameter, so we should instead consider functions $X \times [0,1] \to Y$.

Definition 4.2.2. The functions $f, g \in C(X, Y)$ are homotopic if if there is a continuous function

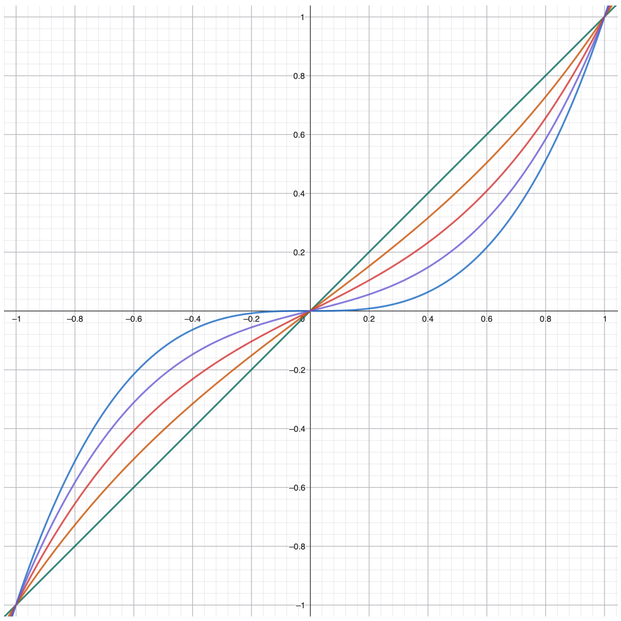
$$F : X \times [0,1] \to Y, \qquad F(x,0) = f(x), \quad F(x,1) = g(x), \quad \text{for all } x \in X.$$

We then say that F is a homotopy between f and g.

We can think of F as representing a family of functions $\{f_t\}_{t\in[0,1]}$, where $f_t(x) = F(x,t)$, that continuously deforms between f and g. In particular $f_t : C(X, Y)$ for all $t \in [0,1]$. So the homotopy F gives us a (continuous) path of *functions*, $[0,1] \ni t \mapsto f_t \in C(X, Y)$.

Example 4.2.3. Take $X = [-1,1]$ and the functions $f(x) = x^3$ and $g(x) = x$. We can define a homotopy between these functions by taking the straight line path between them,

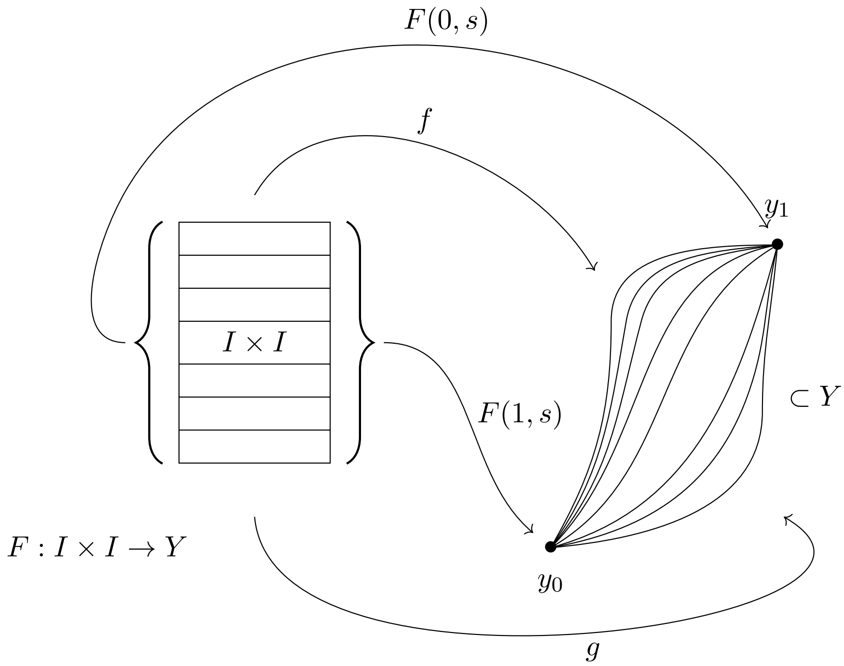
$$F : [-1,1] \times [0,1] \to \mathbb{R}, \qquad F(x,t) = (1-t)f(x) + tg(x) = (1-t)x^3 + tx$$



Example 4.2.4. Suppose we're in the case that $X = [0,1]$, then $C(X, Y) = C([0,1], Y)$ describes the set of paths in Y. We can furthermore consider paths with fixed endpoints $f(0) = g(0) = y_0$ and $f(1) = g(1) = y_1$. Then a homotopy with fixed endpoints is a continuous function $F : [0,1] \times [0,1] \to Y$ such that

$$F(t,0) = f(t), \qquad F(t,1) = g(t), \qquad F(0,s) = y_0, \qquad F(1,s) = y_1,$$

see the diagram below.



We will return to such examples in the next section.

Recall that $E \subset \mathbb{R}^n$ is convex if for all $x, y \in E$, then $(1-t)x + ty \in E$ for all $t \in [0,1]$.

Proposition 4.2.5. *If $E \subset \mathbb{R}^n$ is convex, then any two maps $f, g \in C(X, E)$ are homotopic.*

Proof. We define the homotopy as follows,

$$F : X \times [0,1] \to E, \qquad F(x,t) = (1-t)f(x) + tg(x).$$

It is clear that F is continuous (as it is a linear combination of two continuous functions), $F(x,t) \in E$ for all (x,t) by convexity of E, $F(x,0) = f(x)$ and $F(x,1) = g(x)$ for all $x \in E$. □

In particular, this means that any continuous function $f : \mathbb{R}^n \to \mathbb{R}^m$ can be continuously deformed to the function $g : \mathbb{R}^n \to \mathbb{R}^m$ such that $g(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$.

Lemma 4.2.6. *The relation $f \sim_h g$ if f is homotopic to g defines an equivalence relation on $C(X,Y)$.*

Proof. The proof is basically identical to Lemma 3.10.5. Firstly $f \sim_h f$ as we can take $F : X \times [0,1] \to Y$ given by $F(x,t) = f(x)$ for all (x,t). Next, if $f \sim_h g$, via $F : X \times [0,1] \to Y$, we define

$$F^- : X \times [0,1] \to Y, \quad F^-(x,t) = F(x, 1-t), \quad F^-(x,0) = F(x,1) = g(x), \quad F^-(x,1) = f(x).$$

Lastly, if $f \sim_h g$ via F and $g \sim_h h$ via G, then we take

$$H : X \times [0,1] \to Y, \qquad H(x,t) = \begin{cases} F(x,2t), & t \in [0, \frac{1}{2}], \\ G(x, 2t-1), & t \in [\frac{1}{2}, 1]. \end{cases}$$ □

Definition 4.2.7. Given topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$,

$$[X, Y] = C(X, Y)/\sim_h = \bigcup_{f \in C(X,Y)} [f] = \bigcup_{f \in C(X,Y)} \{g \in C(X, Y) \mid f \sim_h g\}$$

denotes the set of homotopy equivalence classes of functions from X to Y.

Because $C(X, Y)$ is also a topological space via the compact open topology, we can also consider $[X, Y]$ as a topological space via the quotient topology.

Example 4.2.8. For Euclidean space $[\mathbb{R}^n, \mathbb{R}^m] \cong \{\mathrm{pt}\}$ as any two functions are homotopic.

We will see that there are many examples where $[X, Y]$ is non-trivial.

Using homotopy of functions, we can consider homotopy of spaces.

Definition 4.2.9. Two topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ are *homotopic* or *homotopy equivalent* if there are continuous functions $f \in C(X, Y)$ and $g \in C(Y, X)$ such that $g \circ f \sim_h \mathrm{id}_X$ and $f \circ g \sim_h \mathrm{id}_Y$, where id is the identity function ($\mathrm{id}_X(x) = x$ for all $x \in X$).

We write $X \sim_h Y$ if they are homotopy equivalent, which is justified by the following exercise.

Exercise 4.3. Show that homotopy equivalence of topological spaces is an equivalence relation.

Proposition 4.2.10. *If X is homeomorphic to Y, then X is homotopy equivalent to Y.*

Proof. By assumption, there is a homeomorphism $f \in C(X, Y)$ so that $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. So it certainly holds that $f^{-1} \circ f \sim_h \mathrm{id}_X$ and $f \circ f^{-1} \sim_h \mathrm{id}_Y$. □

We see that homotopy of spaces is a much weaker property than homeomorphism. We only require $f \circ g$ and $g \circ f$ to be deformable to the identity, rather than obtain an explicit inverse function. As such, there are many spaces that are homotopic, but not homeomorphic.

Lemma 4.2.11. *If $X \sim_h Y$, then $[X, Z] = [Y, Z]$ and $[Z, X] = [Z, Y]$.*

Proof. Let $f \in C(X, Y)$ and $g \in C(Y, X)$ that implement the homotopy equivalence. Then for any $h \in C(X, Z)$ and $j \in C(Y, Z)$, we have $h \circ g \in C(Y, Z)$ and $j \circ f \in C(X, Z)$. Then we find

$$\begin{aligned} h &= h \circ \mathrm{id}_X \sim_h h \circ (g \circ f) = (h \circ g) \circ f \\ j &= j \circ \mathrm{id}_Y \sim_h j \circ (f \circ g) = (j \circ f) \circ g, \end{aligned}$$

which shows that $[X, Z] \ni [h] \mapsto [h \circ g] \in [Y, Z]$ will give a bijection. The proof $[Z, X] = [Z, Y]$ is similar. □

Example 4.2.12. Let $X = \{\mathrm{pt}\}$. We claim that $\{\mathrm{pt}\} \sim_h \mathbb{R}^n$. We define $f : \mathbb{R}^n \to \{\mathrm{pt}\}$ by $f(\mathbf{x}) = \mathrm{pt}$ (indeed, this is our only option) and $g : \{\mathrm{pt}\} \to \mathbb{R}^n$ by $g(\mathrm{pt}) = \mathbf{0}$. Then $f \circ g(\mathrm{pt}) = \mathrm{pt} = \mathrm{id}(\mathrm{pt})$ and $g \circ f(\mathbf{x}) = \mathbf{0}$, which we have shown is homotopic to $\mathrm{id}_{\mathbb{R}^n}$ as any two functions are homotopic.

Definition 4.2.13. We say that a space is contractible if it is homotopy equivalent to a point.

The same argument as Example 4.2.12 shows that any convex subspace $E \subset \mathbb{R}^n$ is contractible. We will eventually show that S^1 is not contractible.

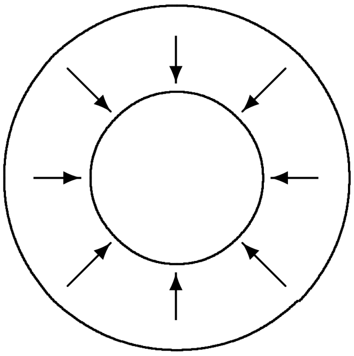
Example 4.2.14. Take the unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ and the annulus

$$A = \{(x, y) \in \mathbb{R}^2, \mid 1 \leq \sqrt{x^2 + y^2} \leq 2\}.$$

We claim that $A \sim_h S^1$. Firstly we take $f : S^1 \to A$ by $f(x, y) = (x, y)$ and $g : A \to S^1$ via $g(x, y) = (x^2 + y^2)^{-1/2}(x, y)$. Then $g \circ f = \mathrm{id}_{S^1}$ and $f \circ g(x, y) = (x^2 + y^2)^{-1/2}(x, y)$. We then define

$$F : A \times [0, 1] \to A, \qquad F((x, y), t) = \frac{t\sqrt{x^2 + y^2} + (1 - t)}{\sqrt{x^2 + y^2}}(x, y),$$

where $F((x, y), 0) = f \circ g(x, y)$ and $F((x, y), 0) = (x, y) = \mathrm{id}_A(x, y)$ as required.



Homotopy $S^1 \sim_h A$ [2, p99].

Exercise 4.4. Show that $\mathbb{C}^\times \sim_h \mathbb{T}$.

Theorem 4.2.15. *If X is connected and Y is disconnected, then X and Y are not homotopy equivalent.*

Proof. Suppose that there is a homotopy equivalence via maps $f : X \to Y$ and $g : Y \to X$. Hence there is $F \in C(Y \times [0, 1], Y)$ such that $F(y, 0) = f(g(y))$ and $F(y, 1) = y$ for all $y \in Y$. Because Y is disconnected, $Y = U \cup V$ and there is a continuous function $p : Y \to \{-1, 1\}$ such that $p(y) = 1$ if $p \in U$ and $p(y) = -1$ if $y \in V$. Next, because X is connected, so is $f(X)$ and so $f(X) \subset U$ or $f(X) \subset V$, so we assume $f(X) \subset U$. Then taking some $v \in V$, we define $h : [0, 1] \to \{-1, 1\}$ such that $h(t) = p(F(v, t))$, which is continuous as F and p are continuous. Then we find that

$$F(v, 0) = f(g(0)) \in U, \qquad f(v, 1) = v \in V, \qquad \Longrightarrow \qquad h(0) = 1, \qquad h(1) = -1$$

Hence we have a continuous surjection $h : [0, 1] \to \{-1, 1\}$, which is impossible. Hence the map F cannot exist and, thus, X and Y are not homotopy equivalent. $\square$

We do *not* have an analogue of Theorem 4.2.15 for compact spaces. Indeed both $\overline{B(\mathbf{0},r)}$ and $B(\mathbf{0},r)$ are homotopy equivalent to a point, but one is compact and the other is not.

Exercise 4.5. Find a Hausdorff topological space X which is homotopic to a non-Hausdorff space Y.

4.3 The fundamental group

In the previous section we looked at general homotopy of functions and spaces. Here we restrict our attention to *loops* in a topological spaces, where a loop is a path $\gamma \in C([0,1],X)$ such that $\gamma(0)=\gamma(1)$. Because $\gamma(0)=\gamma(1)$, we can equivalently think of a loops as a map from $[0,1]/\{0\sim 1\} \simeq S^1$ to X.

In order to be precise where our loops start and finish, we work with so-called *pointed* spaces (X,x_0), which is a topological space with a specific choice of $x_0 \in X$, which we call the base point. For pointed spaces, we can define pointed (continuous) maps

$$C\big((X,x_0),(Y,y_0)\big) = \big\{f \in C(X,Y) \mid f(x_0)=y_0\big\}$$

and pointed homotopies

$$F \in C(X\times[0,1],Y), \qquad F(x_0,t)=y_0 \text{ for all } t\in[0,1].$$

For the case of $S^1 = \big\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2=1\big\}$, we will take $(1,0)\in S^1$ to be the base point. We can then define the sets

$$\mathrm{Loop}(X,x_0) = C\big((S^1,(1,0)),(X,x_0)\big), \qquad \pi_1(X,x_0) = \mathrm{Loop}(X,x_0)/\sim_h^{x_0}$$

where $\sim_h^{x_0}$ denotes homotopies that are pointed at x_0.

Implicit in our defnition is that pointed homotopy is an equivalence relation, which follows from the same proof as Lemma 3.10.5. Given $\alpha\in\mathrm{Loop}(X,x_0)$, we write $[\alpha]\in\pi_1(X,x_0)$ the equivalence class.

Remark 4.3.1. If (X,x_0) and (Y,y_0) are pointed topological spaces.

$$\big[(X,x_0),(Y,y_0)\big] = C(X,Y)/\{f\sim_h g \text{ pointed homotopy }\}.$$

As a set $\pi_1(X,x_0) = \big[(S^1,(1,0)),(X,x_0)\big]$. By taking the circle (and higher-dimensional spheres), $\big[(S^1,(1,0)),(X,x_0)\big]$ can be given a group structure.

In what follows we will use the presentation $S^1\sim[0,1]/\{0\sim 1\}$ with base point $0=1$. Our aim is to give $\mathrm{Loop}(X,x_0)$ additional algebraic structure. First, if $\alpha,\beta\in\mathrm{Loop}(X,x_0)$ we define the concatenation

$$\alpha\#\beta\in\mathrm{Loop}(X,x_0), \qquad \alpha\#\beta(t) = \begin{cases}\alpha(2t), & t\in\big[0,\frac{1}{2}\big],\\ \beta(2t-1), & t\in\big[\frac{1}{2},1\big].\end{cases}$$

Hence, the loop $\alpha\#\beta$ goes along α, returns to x_0, then goes along β. Similarly, if $\alpha\in\mathrm{Loop}(X,x_0)$, we define the reverse loop

$$\alpha^-\in\mathrm{Loop}(X,x_0), \qquad \alpha^-(t)=\alpha(1-t),$$

which travels along α in the opposite direction.

Lemma 4.3.2. *If $\alpha\sim_h^{x_0}\tilde\alpha$ and $\beta\sim_h^{x_0}\tilde\beta$, then $\alpha\#\beta\sim_h^{x_0}\tilde\alpha\#\tilde\beta$ and $\alpha^-\sim_h^{x_0}\tilde\alpha^-$.*

Proof. Let $F,G:[0,1]\times[0,1]\to X$ be the pointed homotopies from α to $\tilde\alpha$ and β to $\tilde\beta$ respectively. We take $(t,s)\in[0,1]$ where t is the loop parameter and s is the deformation parameter. Then

$$F\#G:[0,1]\times[0,1]\to X, \qquad F\#G(t,s) = \begin{cases}F(2t,s), & t\in\big[0,\frac{1}{2}\big],\\ G(2t-1,s), & t\in\big[\frac{1}{2},1\big].\end{cases}$$

gives a homotopy from $\alpha\#\beta$ to $\tilde{\alpha}\#\tilde{\beta}$. Similarly,

$$F^- : [0,1]\times[0,1] \to X, \qquad F^-(t,s) = F(1-t,s)$$

will give a homotopy from α^- to $\tilde{\alpha}^-$. □

Our aim is to define a group structure on the base pointed loops in X. To do this, we need associativity of the product. However, we immediately run into a problem as

$$(\alpha\#\beta)\#\gamma(t) = \begin{cases} \alpha(4t), & t\in\left[0,\frac{1}{4}\right], \\ \beta(4t-1), & t\in\left[\frac{1}{4},\frac{1}{2}\right], \\ \gamma(2t-1), & t\in\left[\frac{1}{2},1\right], \end{cases} \qquad \alpha\#(\beta\#\gamma)(t) = \begin{cases} \alpha(2t), & t\in\left[0,\frac{1}{2}\right], \\ \beta(4t-2), & t\in\left[\frac{1}{2},\frac{3}{4}\right], \\ \gamma(4t-3), & t\in\left[\frac{3}{4},1\right], \end{cases}$$

and there will be many cases where $(\alpha\#\beta)\#\gamma \neq \alpha\#(\beta\#\gamma)$. On the other hand, we can recover associativity *up to homotopy*.

Lemma 4.3.3. *Let $\alpha\in \mathrm{Loop}(X,x_0)$ and $g\in C([0,1],[0,1])$ such that $g(0)=0$ and $g(1)=1$. Then $\alpha\circ g \sim_h^{x_0} \alpha$.*

Proof. Because $[0,1]$ is convex, g is homotopic to $\mathrm{id}_{[0,1]}$ (with endpoints fixed). Therefore we have that

$$\alpha\circ g \sim_h^{x_0} \alpha\circ \mathrm{id}_{[0,1]} = \alpha.$$ □

Lemma 4.3.4. *As elements in $\pi_1(X,x_0)$, $[(\alpha\#\beta)\#\gamma] = [\alpha\#(\beta\#\gamma)]$.*

Proof. We can write

$$\alpha\#(\beta\#\gamma)(t) = (\alpha\#\beta)\#\gamma(g(t)), \qquad g:[0,1]\to[0,1], \quad g(t) = \begin{cases} \frac{1}{2}t, & t\in\left[0,\frac{1}{2}\right], \\ t-\frac{1}{4}, & t\in\left[\frac{1}{2},\frac{3}{4}\right], \\ 2t-1, & t\in\left[\frac{3}{4},1\right]. \end{cases}$$

We see that g is continuous and such that $g(0)=0$ and $g(1)=1$. Therefore we obtain a homotopy $(\alpha\#\beta)\#\gamma \sim_h^{x_0} \alpha\#(\beta\#\gamma)$ by Lemma 4.3.3. □

Theorem 4.3.5. *Let $(X,\mathcal{T})$ be a topological space. Then for any $x_0\in X$, $\pi_1(X,x_0)$ is a group such that*

$$[\alpha]*[\beta] = [\alpha\#\beta], \qquad [\alpha]^{-1} = [\alpha^-].$$

Proof. The group operation is well defined by Lemma 4.3.2 and is associative by Lemma 4.3.4. To obtain a group, we need an identity and inverses. For an identity, we take the loop $\alpha_{x_0}\in \mathrm{Loop}(X,x_0)$, where $\alpha_{x_0}(t)=x_0$ for all $t\in[0,1]$. We need to show that $[\alpha_{x_0}\#\beta] = [\beta\#\alpha_{x_0}] = [\beta]$ for any $\beta\in\mathrm{Loop}(X,x_0)$, We can write

$$\alpha_{x_0}\#\beta(t) = \beta\circ g(t), \qquad g(t) = \begin{cases} 0, & t\in\left[0,\frac{1}{2}\right], \\ 2t-1, & t\in\left[\frac{1}{2},1\right], \end{cases}$$

where $g:[0,1]\to[0,1]$ is continuous with $g(0)=0$, $g(1)=1$. Hence $[\alpha_{x_0}\#\beta]=[\beta]$ by Lemma 4.3.3. A similar argument will show that $[\beta\#\alpha_{x_0}]=[\beta]$. Hence $[\alpha_{x_0}]$ acts as the group identity under concatenation of loops.

by Lemma 4.3.2, if $[\alpha] = [\tilde{\alpha}]$, then $[\alpha^-] = [\tilde{\alpha}^-]$. So the operation $[\alpha]^{-1} = [\alpha^-]$ is well-defined in $\pi_1(X, x_0)$. We just need to check that $[\alpha^-]$ is indeed the inverse under concatenation in $\pi_1(X, x_0)$. That is, we check that $[\alpha] * [\alpha^-] = [\alpha \# \alpha^-] = [\alpha_{x_0}]$. To do this we define the homotopy

$$F : [0,1] \times [0,1] \to X, \qquad F(t,s) = \begin{cases} \alpha(2t), & t \in \left[0, \frac{s}{2}\right], \\ \alpha(s), & t \in \left[\frac{s}{2}, 1 - \frac{s}{2}\right], \\ \alpha(2-2t), & t \in \left[1 - \frac{s}{2}, 1\right]. \end{cases}$$

That is, $\alpha_s = F(\cdot, s) \in \mathrm{Loop}(X, x_0)$ describes a path where we move along α at double speed until time $s/2$, rest until time $1 - s/2$ and then return along α^- at double speed. By construction F is continuous and $\alpha_0(t) = \alpha(0) = x_0 = \alpha_{x_0}(t)$. We also find that

$$\alpha_1(t) = F(t,1) = \begin{cases} \alpha(2t), & t \in \left[0, \frac{1}{2}\right], \\ \alpha(2-2t), & t \in \left[\frac{1}{2}, 1\right], \end{cases} = \alpha \# \alpha^-(t),$$

which shows $[\alpha] * [\alpha^-]$ is the identity. A similar argument shows $[\alpha^-] * [\alpha] = [\alpha_{x_0}]$. $\square$

Example 4.3.6. If $E \subset \mathbb{R}^n$ is convex, then $\pi(E, x_0)$ isomorphic to the trivial group $\{e\}$ for any $x_0 \in E$.

It will take us some effort to show that there are topological spaces with non-trivial fundamental groups. The simplest example is the circle, which we study in detail in the next section. In the meantime, we consider some basic properties of $\pi_1(X, x_0)$.

The first basic question is the dependence of $\pi_1(X, x_0)$ on the choice of basepoint $x_0 \in X$. If there is a path γ from x_0 to x_1, then we can translate any loop $\alpha \in \mathrm{Loop}(X, x_0)$ to a loop $\gamma^- \# \alpha \# \gamma \in \mathrm{Loop}(X, x_1)$. As the following shows does not change the algebraic structure of the fundamental group.

Theorem 4.3.7. *Let γ be a path from x_0 to x_1 in X. Then the induced map*

$$\gamma_* : \pi_1(X, x_0) \to \pi_1(X, x_1), \qquad \gamma_*([\alpha]) = [\gamma^- \# \alpha \# \gamma]$$

is an isomorphism of groups.

Proof. We first check γ_* is a homomorphism,

$$\begin{aligned} \gamma_*\big([\alpha] * [\beta]\big) = \gamma_*\big([\alpha \# \beta]\big) &= \big[\gamma^- \# \alpha \# \beta \# \gamma\big] \\ &= \big[\gamma^- \# \alpha \# \gamma \# \gamma^- \# \beta \# \gamma\big] = \big[\gamma^- \# \alpha \# \gamma\big] * \big[\gamma^- \# \beta \# \gamma\big] \\ &= \gamma_*([\alpha]) * \gamma_*([\beta]), \end{aligned}$$

which holds for any $\alpha, \beta \in \mathrm{Loop}(X, x_0)$. To show injectivity, we assume that $\gamma_*([\alpha]) = \gamma_*([\beta])$. Then

$$\begin{aligned} [\alpha] &= [\gamma \# \gamma^- \# \alpha \# \gamma \# \gamma^-] = \gamma_*^-\big([\gamma^- \# \alpha \# \gamma]\big) = \gamma_*^-\big(\gamma_*([\alpha])\big) \\ &= \gamma_*^-\big(\gamma_*([\beta])\big) = [\gamma \# \gamma^- \# \beta \# \gamma \# \gamma^-] = [\beta]. \end{aligned}$$

To show surjectivity, we take $[\xi] \in \pi_1(X, x_1)$. Then $[\gamma \# \xi \# \gamma^-] \in \pi_1(X, x_0)$ and

$$\gamma_*\big([\gamma \# \xi \# \gamma^-]\big) = \big[\gamma^- \# \gamma \# \xi \# \gamma^- \# \gamma\big] = [\xi]. \qquad \square$$

Corollary 4.3.8. *If $(X, \mathcal{T})$ is path connected, then the groups $\pi_1(X, x_0)$, $x_0 \in X$ are all isomorphic.*

When the space is path connected, we often use the simpler notation $\pi_1(X)$ to denote $\pi_1(X, x_0)$ for any $x_0 \in X$. We say that a space X is **simply connected** if X is path connected and $\pi_1(X) \cong \{e\}$, the trivial group. We have already seen that any convex space is simply connected. As we will show, $S^1 \sim_h \mathbb{C} \setminus \{0\}$ is path connected, but has non-trivial fundamental group.

Exercise 4.6. Show that any contractible space is simply connected.

On the other hand, if x_0 and x_1 are in different path connected components, there is no reason to expect that $\pi(X, x_0)$ and $\pi(X, x_1)$ are related. Indeed, $C^\gamma_{x_0}$ might be contractible while $C^\gamma_{x_1}$ a much more complicated space.

Now that we have defined the fundamental group, we wish to use it to consider the question considered at the start of this chapter: can we use $\pi_1(X, x_0)$ and $\pi_1(Y, y_0)$ to detect if X and Y are homeomorphic or not? To do this, we need to be able to pass between fundamental groups of different spaces.

Lemma 4.3.9. *Let $f \in C(X, Y)$. If α, $\beta \in \operatorname{Loop}(X, x_0)$ are homotopic, then $f \circ \alpha$, $f \circ \beta \in \operatorname{Loop}(Y, y_0)$ are homotopic.*

Proof. If $F \in C([0,1] \times [0,1], X)$ is the homotopy in X, then $G : [0,1] \times [0,1] \to Y$, $G(t,s) = f\big(F(t,s)\big)$ is the homotopy in Y. □

Proposition 4.3.10. *Let (X, x_0) and (Y, y_0) be pointed spaces with $f : X \to Y$ continuous (so $f(x_0) = y_0$). Then the induced map*

$$f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0), \qquad f_*([\alpha]) = [f \circ \alpha]$$

is a homomorphism of groups. Furthermore, if $f : (X, x_0) \to (Y, y_0)$ and $g : (Y, y_0) \to (Z, z_0)$ are continuous pointed maps, then $(g \circ f)_* = g_* \circ f_*$.

Proof. The map f_* is well defined by Lemma 4.3.9 and we check that

$$f_*([\alpha] * [\beta]) = \big[f \circ (\alpha \# \beta)\big] = \big[(f \circ \alpha) \# (f \circ \beta)\big] = f_*([\alpha]) * f_*([\beta])$$

so f_* is a group homomorphism. The proof $(g \circ f)_* = g_* \circ f_*$ is a simple check. □

Remark 4.3.11. Proposition 4.3.10 shows that the fundamental group gives a **covariant functor** from the category of pointed topological spaces Top_* to the category of grous Gp. We will not emphasise the role of category theory and functors in this course, but it provides the mathematical structure needed to understand algebraic topology and homological algebra more generally.

Proposition 4.3.12. *If $f : X \to Y$ is a homeomorphism of topological spaces, then the induced map $f_* : \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is a group isomorphism.*

Proof. The existence of a homeomorphism gives maps $f_* : \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ and $(f^{-1})_* : \pi_1(Y, f(x_0)) \to \pi_1(X, x_0)$ such that

$$(f^{-1})_* \circ f_* = (f^{-1} \circ f)_* = (\mathrm{id}_X)_*, \qquad f_* \circ (f^{-1})_* = (\mathrm{id}_Y)_*.$$

It is then a simple exercise to show that $(\mathrm{id}_X)([\alpha]) = [\alpha]$ and $(\mathrm{id}_Y)_*$ give an isomorphism of $\pi_1(X, x_0)$ and $\pi(Y, y_0)$ respectively. $\square$

Therefore, we can say that if X and Y are path connected topological spaces with $\pi_1(X) \neq \pi_1(Y)$, then we can conclude that X and Y are not homeomorphic. Hence we can use the fundamental group to distinguish between spaces that are topologically distinct.

We can actually show something much stronger, *homotopic* spaces have the same fundamental group.

Lemma 4.3.13. *Suppose that $f, g \in C(X, Y)$ are homotopic maps via $F \in C(X \times [0,1], Y)$. Then for $x_0 \in X$, $y_0 = f(x_0)$ and $y_1 = g(x_0)$, define a path $\gamma(t) = F(x_0, t)$ from y_0 to y_1. Then $g_* = \gamma_* \circ f_*$ with $\gamma_* : \pi_1(Y, y_0) \xrightarrow{\simeq} \pi_1(Y, y_1)$ the isomorphism from Theorem 4.3.7. Put another way, the following diagram commutes:*

$$\begin{tikzcd}
\pi_1(X, x_0) \arrow[r, "f_*"] \arrow[dr, "g_*"'] & \pi_1(Y, y_0) \arrow[d, "\gamma_*"] \\
& \pi_1(Y, y_1).
\end{tikzcd}$$

In particular f_* is a group isomorphism if and only if g_* is a group isomorphism.

Proof. The relation $g_* = \gamma_* \circ f_*$ follows if we can show that

$$\left[\gamma^- \#(f \circ \alpha)\#\gamma\right] = \left[(g \circ \alpha)\right]$$

for any $\alpha \in \mathrm{Loop}(X, x_0)$. Fixing such an α, we consider the function

$$G \in C([0,1] \times [0,1], Y), \qquad G(t,s) = F\big(\alpha(t), s\big)$$

with F the homotopy between f and g. We also define four functions that trace along the edges of the unit square $[0,1] \times [0,1]$

$$\beta_\downarrow(t) = (t, 0), \qquad \beta_\leftarrow(s) = (0, s), \qquad \beta_\uparrow(t) = (t, 1), \qquad \beta_\rightarrow(t) = (1, s)$$

with $t, s \in [0,1]$. Then by construction we have

$$G \circ \beta_\downarrow(t) = f(\alpha(t)), \qquad G \circ \beta_\leftarrow = \gamma, \qquad G \circ \beta_\uparrow = g \circ \alpha, \qquad G \circ \beta_\rightarrow = \gamma.$$

Using that $[0,1] \times [0,1]$ is convex, the paths $\beta_\uparrow$ and $(\beta_\leftarrow^- \#\beta_\downarrow)\#\beta_\rightarrow$ both describe paths from $(0,1)$ to $(1,1)$ and hence are homotopic with endpoints fixed. So we can apply Lemma 4.3.9

$$g \circ \alpha = G \circ \beta_\uparrow \sim_h G \circ (\beta_\leftarrow^- \#\beta_\downarrow)\#\beta_\rightarrow = \gamma^- \#(f \circ \alpha)\#\gamma,$$

which gives the relation $\left[\gamma^- \#(f \circ \alpha)\#\gamma\right] = \left[(g \circ \alpha)\right]$. Because γ_* is an isomorphism, it's clear that f_* is an isomorphism if and only if g_* is an isomorphism. $\square$

Theorem 4.3.14. *If $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ are homotopy equivalent via functions $f \in C(X, Y)$ and $g \in C(Y, X)$, then $f_* : \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is an isomorphism of groups.*

Proof. Recall that to be homotopy equivalent, $g \circ f \sim_h \mathrm{id}_X$ and $f \circ g \sim_h \mathrm{id}_Y$. Because $(\mathrm{id}_X)_*$ and $(\mathrm{id}_Y)_*$ are isomorphisms of $\pi_1(X, x_0)$ and $\pi_1(Y, f(x_0))$, if follows from Lemma 4.3.13 and Proposition 4.3.10 that $(g \circ f)_* = g_* \circ f_*$ and $(f \circ g)_* = f_* \circ g_*$ are group isomorphism of

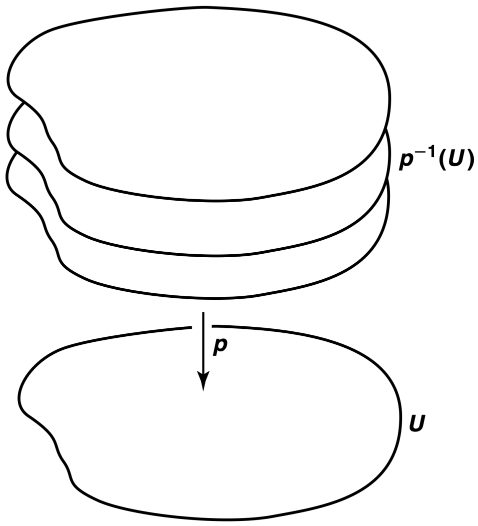


Figure 5: A covering map [6, Figure 53.1].

$\pi_1(X, x_0)$ and $\pi_1(Y, f(x_0))$ respectively. We can then conclude that f_* and g_* must be group isomorphisms. $\square$

We therefore see that the fundamental group does not distinguish between homotopic spaces. Once again, this is useful as it is very difficult in general to show that two spaces are homotopic. On the other hand, if their fundamental groups are different then we know they can not be homotopic, a stronger statement than spaces not being homeomorphic.

Exercise 4.7. Show that S^n is simply connected for $n \geq 2$.

4.4 Covering spaces

In order to show that there are topological spaces with $\pi_1(X, x_0)$ *not* the trivial group $\{e\}$, we will need some additional machinery which will allow us to 'lift' paths/loops in a space X to paths/loops in a 'simpler' space E where we can more easily analyse their properties.

Definition 4.4.1. Let $(X, \mathcal{T})$ and $(\widetilde{X}, \widetilde{\mathcal{T}})$ be topological spaces. A *covering map* is a continuous surjection $p : \widetilde{X} \to X$ such that for each $x \in X$ there is a neighbourhood $x \in U_x \in \mathcal{T}$ and discrete space D_x such that

$$p^{-1}(U_x) = \bigsqcup_{d \in D_x} V_d, \quad V_d \in \widetilde{\mathcal{T}}, \qquad p|_{V_d} : V_d \xrightarrow{\simeq} U_x \text{ a homeomorphism for all } d \in D_x.$$

The space $(\widetilde{X}, \widetilde{\mathcal{T}})$ is called a *covering space*.

The discrete space $D_x = p^{-1}(x) \subset \widetilde{X}$ is called the fiber over x. In particular, for each $x \in X$, $p^{-1}(U_x)$ is homeomorphic to $D_x \times U_x$. We also call the subsets $V_d \subset p^{-1}(U_x)$ the sheets of $p^{-1}(U)$.

Examples 4.4.2.

1. The identity map id_X is a covering map, though it is not very interesting.

2. Identifying $S^1 \simeq \mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$, the exponential map $p : \mathbb{R} \to S^1$, $p(t) = e^{2\pi i t}$ is a

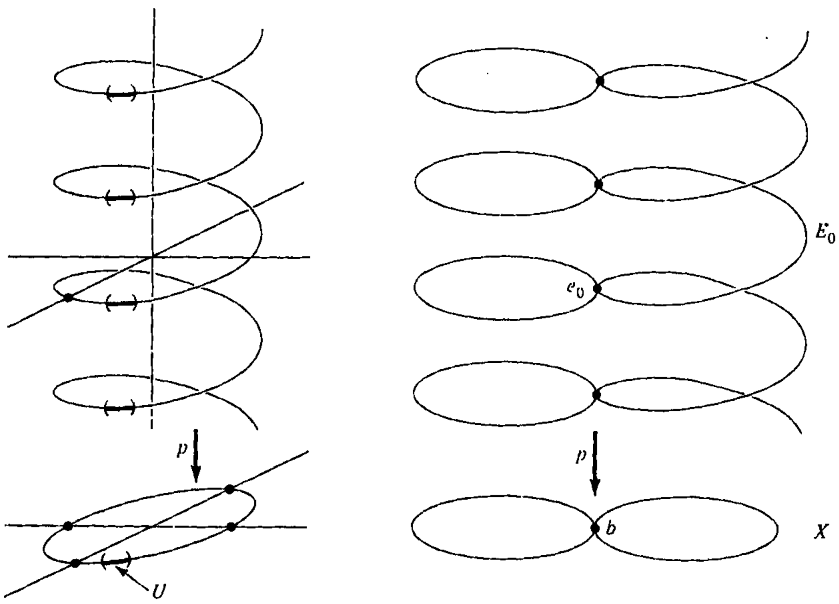


Figure 6: Coverings of S^1 (left) and the figure eight (right) [4, p124].

covering map. Indeed if we take $e^{2\pi i t_0} \in S^1$ and an open neighbourhood

$$U = \{e^{2\pi i t} \mid |t - t_0| < \epsilon\}, \qquad p^{-1}(U) = \bigsqcup_{m \in \mathbb{Z}} (m - t_0 - \epsilon, m + t_0 + \epsilon)$$

where $p|_{(m-t_0-\epsilon, m+t_0+\epsilon)}$ is a homeomorphism. To help our geometric intuition we can identify $\mathbb{R} \simeq \{(\cos(2\pi t), \sin(2\pi t), t) \mid t \in \mathbb{R}\}$, a helix in $\mathbb{R}^3$, with the covering map projecting onto the xy-plane (see Figure 6). We also see that the fibre $p^{-1}(z) \cong \mathbb{Z}$ for any $z \in \mathbb{T}$.

3. Take X to the figure eight space, which comes from joining $S^1 \sqcup S^1$ at a single point (a quotient space). A covering sapce for X comes from unwiding one of the circles in the figure eight and considering the helix with projection map like the circle (see Figure 6).

Exercise 4.8. Show that the map $f : S^2 \to \mathbb{R}P^2$, $f(\mathbf{x}) = \ell_{\mathbf{x}}$ from Example 3.11.11 is a covering map.

Exercise 4.9. Show that any covering map is open (maps open sets to open sets).

If $p : \widetilde{X} \to X$ is a covering map and $f \in C(Y, X)$ a map, we would like to know if we can 'factorise' f via a map $\widetilde{f} \in C(Y, \widetilde{X})$ such that $f(y) = p \circ \widetilde{f}(y)$ for all $y \in Y$. We then call $\widetilde{f}$ a lift of f to the covering space.

$$\begin{tikzcd} & \widetilde{X} \arrow[d, "p"] \\ Y \arrow[ur, dotted, "\widetilde{f}"] \arrow[r, "f"] & X \end{tikzcd}$$

If a lift exists, the following shows that it is unique.

Lemma 4.4.3. *Let $p : \widetilde{X} \to X$ be a covering map and Y a connected space. Suppose $f \in$*

$C(Y,X)$ and $g,h \in C(Y,\widetilde{X})$ are two lifts of f. If $g(y_0) = h(y_0)$ for some $y_0 \in Y$, then $g = h$.

Proof. We take the subsets

$$S = \{y \in Y \mid g(y) = h(y)\}, \qquad T = \{y \in Y \mid g(y) \neq h(y)\},$$

where $S \cup T = Y$ and $S \cap T = \emptyset$. Because Y is connected, if we can show that both S and T are open, it follows that $T = \emptyset$ and $S = Y$, hence $g = h$.

Let $y \in Y$ and take a neighbourhood U of $f(y) \in X$ such that $p^{-1}(U) = \bigsqcup_d V_d \in \widetilde{\mathcal{T}}$. Because $g \circ p$ and $h \circ p$ compute f, $g(y) \in V_j$ and $h(y) \in V_{j'}$. If $g(y) = h(y)$, then $V_j = V_{j'}$. Otherwise V_j and $V_{j'}$ are disjoint. Because g and h are continuous, there is an open neighbourhood W of Y such that $g(W) \subset V_j$ and $h(W) \subset V_{j'}$. If $y \in T$, then $V_j \cap V_{j'} = \emptyset$ and so $g(z) \neq h(z)$ for all $z \in W \cap T$. Hence T is open. On the other hand, if $y \in S$, then $y \in S$ and $V_j = V_{j'}$. Because $p|_{g(W)} \to X$ and $p|_{h(W)}$ are homeomorphisms on to their image and both g and h are lifts of f, it must follow that $g = h$ on W. Hence $W \subset S$ and so S is open. □

Of particular interest to us is the case when $Y = [0,1]$, where maps $C([0,1],X)$ are the paths in X.

Theorem 4.4.4 (Path Lifting Theorem). *Let $p : \widetilde{X} \to X$ be a covering map, $\gamma : [0,1] \to X$ a path and $e_0 \in \widetilde{X}$ such that $p(e_0) = \gamma(0)$. Then there is a unique path $\widetilde{\gamma} : [0,1] \to \widetilde{X}$ that is a lift of γ.*

Proof. For each $x \in X$, we can take a neighbourhood U_x such that $p^{-1}(U_x)$ is a discrete union of open sets in $\widetilde{X}$. The collection $\{\gamma^{-1}(U_x)\}_{x \in X}$ give an open cover of $[0,1]$, which is compact. So we can take a finite subcollection $\{U_1, \ldots, U_m\} \subset X$ such that the inverse image covers $[0,1]$. Given that γ is a path, we therefore have a partion $0 = t_0 < t_1 \ldots < t_m = 1$ such that $\gamma([t_{j-1}, t_j]) \subset U_j$ and $p^{-1}(U_j) = \bigsqcup_d V_{j,d}$.

With this setup, we construct the lift $\widetilde{\gamma}$ as follows. Because $p(e_0) = \gamma(0) \in U_1$, there is some open neighbourhood $V_1 \in \widetilde{\mathcal{T}}$ of e_0 with $p|_{V_1} \xrightarrow{\simeq} U_1$ a homeomorphism. We can therefore define $\widetilde{\gamma}(t) = p^{-1}(\gamma(t))$ for $t \in [0,t_1]$. By construction, we see that $p \circ \widetilde{\gamma}(t) = \gamma(t)$ for all $t \in [0,t_1]$. We can do the same procedure for each subinterval $[t_{j-1}, t_j]$ and U_j to obtain the complete lift for γ. Uniqueness of the lift follows from Lemma 4.4.3. □

We can improve upon the previous result, where not just paths but homotopies of paths can be lifted to the covering space.

Theorem 4.4.5 (Homotopy lifting). *Let $p : \widetilde{X} \to X$ be a covering map, $F \in C([0,1] \times [0,1], X)$ and $e_0 \in \widetilde{X}$ such that $p(e_0) = F(0,0)$. Then there exists a unique lift $\widetilde{F} \in C([0,1] \times [0,1], \widetilde{X})$.*

Proof. Fixing $s = 0$, there is a path $\gamma_F^0(t) = F(t,0)$ such that $\gamma_F^0(0) = F(0,0) = p(e_0)$. Applying the Path Lifting Theorem, there is a path $\widetilde{\gamma}_F^0(t) = e_t \in \widetilde{X}$ such that $p(e_t) = F(t,0)$ for all $t \in [0,1]$. Similarly, for any fixed $t \in [0,1]$, $\gamma_F^t(s) = F(t,s)$ defines a path in X with $\gamma_F^t(0) = F(t,0) = p(e_t)$. Hence we can again apply the Path Lifting Theorem to define $\widetilde{F}(t,s) = \widetilde{\gamma}_F^t(s)$ such that $\widetilde{F}(t,0) = e_t$ and $p \circ \widetilde{F}(t,s) = F(s,t)$ for all $(t,s) \in [0,1] \times [0,1]$. Hence $\widetilde{F}$ is a lift of F as a function. We just need to show that $\widetilde{F} : [0,1] \times [0,1]$ is continuous.

We can also take the path $\alpha : [0,1] \to X$, $\alpha(s) = F(0,s)$, which has a lift $\widetilde{\alpha}(s) = \widetilde{F}(0,s)$. Applying the same construction as in the proof of the Path Lifting Theorem, there are open sets

$\{U_1, \ldots, U_n\} \subset X$ with $\alpha^{-1}(U_j)$ an open neighbourhood of $[s_{j-1}, s_j]$. Because F is continuous $[0,\epsilon] \times [s_{j-1}, s_j] \subset F^{-1}(U_j)$ for some $\epsilon > 0$. Hence $F([0,\epsilon] \times [s_{j-1}, s_j]) \subset U_j$ for all $j = 1, \ldots, n$. Now $\widetilde{F}(0,0) \in V_1$ for some sheet $V_1 \subset \widetilde{\mathcal{T}}$, $e_t = \widetilde{F}(t,0) \in V_1$ for all $t \in [0,\epsilon]$. This shows that

$$\big(p|_{V_1}^{-1} \circ F\big)(t,s) = \widetilde{F}(t,s), \quad (t,s) \in [0,\epsilon] \times [0, s_1],$$

which shows that $\widetilde{F}$ is continuous on the rectangle $[0,\epsilon] \times [0, s_1]$. We can apply this argument for each sub-interval $[s_{j-1}, s_j]$ in the partition of $[0,1]$ to obtain that $\widetilde{F}$ is continuous on the rectangle $[0,\epsilon] \times [s_{j-1}, s_j]$ for all $j = 1, \ldots, n$. Hence $\widetilde{F}$ is continuous on $[0,\epsilon] \times [0,1]$.

The same proof can be used to show that for each $t_0 \in (0,1)$, there is some $\epsilon > 0$ such that $\widetilde{F}$ is continuous on the rectangle $[t_0 - \epsilon, t_0 + \epsilon] \times [0,1]$ and $[1-\epsilon, 1] \times [0,1]$. Putting this all together, $\widetilde{F}$ is continuous on $[0,1] \times [0,1]$. □

Now that we have showed some useful properties of covering spaces, we can apply these results to understand the fundamental group of some spaces.

Take a pointed space (X, x_0) and a pointed covering map $p : (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$. If $\gamma \in \mathrm{Loop}(X, x_0)$, then we have a lift $\widetilde{\gamma} : [0,1] \to \widetilde{X}$. We certainly have that $\widetilde{\gamma}(0) = \widetilde{x}_0$. But $\widetilde{\gamma}$ need not be a loop in $\widetilde{X}$ and it may occur that $\widetilde{\gamma}(1) \neq \widetilde{x}_0$. We can only say for sure that $p(\widetilde{\gamma}(1)) = x_0$ and so $\widetilde{\gamma}(1) \in p^{-1}(x_0) \cong D$, the fiber over x_0.

Lemma 4.4.6. *Let $p : (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ be a pointed covering map and suppose that $[\alpha] = [\beta] \in \pi_1(X, x_0)$. Then the lifts $\widetilde{\alpha}, \widetilde{\beta} \in C([0,1], \widetilde{X})$ of α and β are such that $\widetilde{\alpha}(0) = \widetilde{x}_0 = \widetilde{\beta}(0)$ and $\widetilde{\alpha}(1) = \widetilde{\beta}(1)$.*

Proof. Because $[\alpha] = [\beta]$, there is a homotopy $F : [0,1] \times [0,1] \to X$, which lifts to $\widetilde{F} : [0,1] \times [0,1] \to \widetilde{X}$. Because $\widetilde{F}$ is unique, it follows that $\widetilde{F}(t,0) = \widetilde{\alpha}(t)$ and $\widetilde{F}(t,1) = \widetilde{\beta}(t)$. The path $s \mapsto \widetilde{F}(0,s)$ is a lift of the constant loop γ_{x_0} starting at $\widetilde{x}_0$. By uniqueness of the lift $\widetilde{x}_0 = \widetilde{F}(0,s)$ for all $s \in [0,1]$. Similarly, $s \mapsto \widetilde{F}(1,s)$ is a lift of the constant loop γ_{x_0} starting at $\widetilde{F}(1,0) = \widetilde{\alpha}(1)$. It follows by uniqueness of the lift that

$$\widetilde{\alpha}(1) = \widetilde{F}(1,0) = \widetilde{F}(1,t) = \widetilde{F}(1,1) = \widetilde{\beta}(1).$$ □

A corollary of the above result is that we can define a function

$$\Phi : \pi_1(X, x_0) \to p^{-1}(x_0), \qquad \Phi([\alpha]) = \widetilde{\alpha}(1),$$

where $\widetilde{\alpha} : [0,1] \to \widetilde{X}$ is any lift of $\alpha \in \mathrm{Loop}(X, x_0)$ such that $\widetilde{\alpha}(0) = \widetilde{x}_0$. The map Φ is currently just a map of *sets* and the image $p^{-1}(x_0)$ might not have the structure of a group. But in some cases, we can extend Φ as a group homomorphism or isomorphism.

Proposition 4.4.7. *Let $\widetilde{X}$ be simply connected and $p : (\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ a pointed covering map. Then the map $\Phi : \pi_1(X, x_0) \to p^{-1}(x_0)$ is an bijection of sets.*

Put another way, $\Phi : \big[(S^1, (1,0)), (X, x_0)\big] \to p^{-1}(x_0)$ is a bijection.

Proof. Take $y \in p^{-1}(x_0)$. Because $\widetilde{X}$ is path connected, there is some path $\alpha : [0,1] \to \widetilde{X}$ such that $\alpha(0) = \widetilde{x}_0$ and $\alpha(1) = y$. Then $p \circ \alpha \in \mathrm{Loop}(X, x_0)$ and $\Phi([p \circ \alpha]) = \alpha(1) = y$. Hence Φ is surjective.

Now take $\gamma_0, \gamma_1 \in \mathrm{Loop}(X, x_0)$ such that $\Phi([\gamma_0]) = \Phi([\gamma_1])$. We then take lifts $\widetilde{\gamma}_0$ and $\widetilde{\gamma}_1$ where by assumption $\widetilde{\gamma}_0(0) = \widetilde{\gamma}_1(0) = \widetilde{x}_0$ and $\widetilde{\gamma}_0(1) = \widetilde{\gamma}_1(1)$. Hence the concatenation $\widetilde{\gamma}_0 \# \widetilde{\gamma}_1^- \in \mathrm{Loop}(\widetilde{X}, \widetilde{x}_0)$. Because $\pi_1(\widetilde{X})$ is trivial, there is a homotopy $G : [0,1] \times [0,1] \to \widetilde{X}$ from $\widetilde{\gamma}_0 \# \widetilde{\gamma}_1^-$

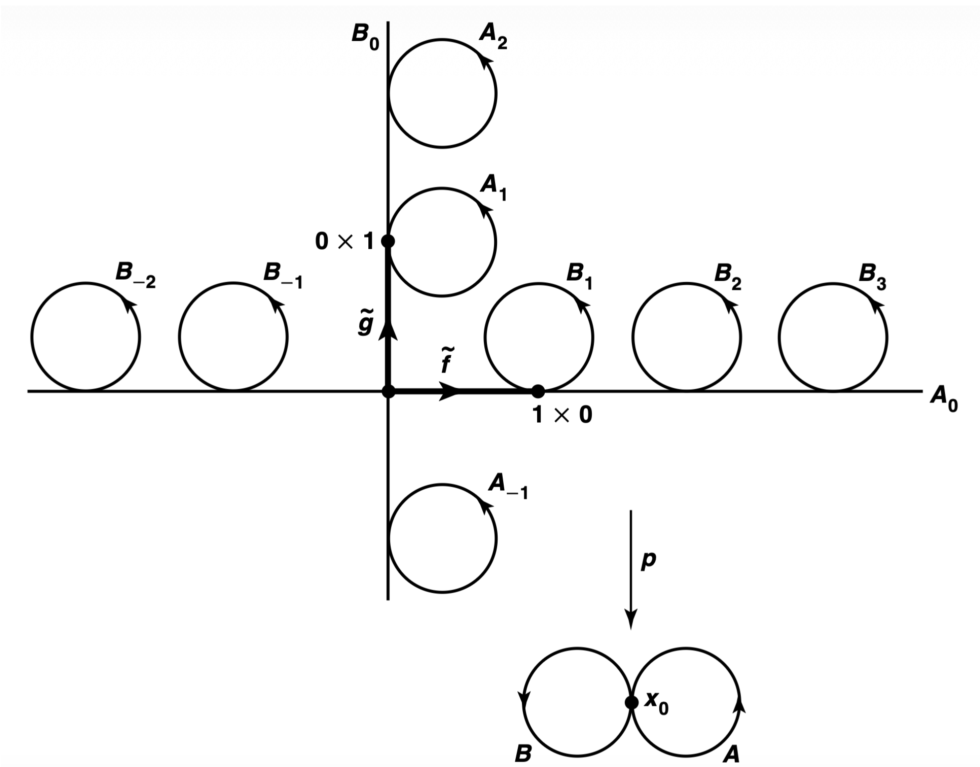
to the constant loop $\gamma_{\tilde{x}_0}$. Taking the covering map $p \circ G : [0,1] \times [0,1] \to X$ is a homotopy in X from $\gamma_0 \# \gamma_1^-$ to γ_{x_0}. It follows that $[\gamma_1]^{-1} = [\gamma_1^-] = [\gamma_0]^{-1}$ and so $[\gamma_0] = [\gamma_1]$. Hence the map Φ is injective. $\square$

We now show how to use the map Φ to better understand the fundamental groups of a few spaces.

Examples 4.4.8. 1. Using the results of Exercises 4.7 and 4.8, we have that $p : S^2 \to \mathbb{R}P^2$ is a covering map and S^2 is simply connected. Taking $N = (0,0,1) \in S^2$ the north pole, we take $p(N) \in \mathbb{R}P^2$ as the base point. Recall that $p(N)$ is the line that passes through the north and south pole. Hence $p^{-1}(p(N)) = \{N, S\}$. Apply Proposition 4.4.7, $\pi_1(\mathbb{R}P^2, p(N))$ has two elements as a set. So as a group, the identity in $\pi_1(\mathbb{R}P^2, p(N))$ is the homotopy class of the constant loop $[\gamma_{p(N)}]$ and the other element is $[p \circ \alpha]$ where $\alpha : [0,1] \to S^2$ is any path from N to S. We can therefore algebraically extend Φ to a group isomorphism and conclude that $\pi_1(\mathbb{R}P^2, p(N)) \cong \mathbb{Z}_2$.

2. Take X to be the figure eight space where we label the circles A and B with basepoint x_0 at the point where the two circles attach. We can define loops $f, g \in \text{Loop}(X, x_0$, where f travels once around A, g travels once around B. We claim that $[f \# g] \neq [g \# f]$ and so $\pi_1(X, x_0)$ is a non-abelian group.

   Because $\mathbb{R}$ is a covering space for S^1, we can construct a covering space for the figure eight by taking two lines (one for each circle), which we put on the x and y axis, and then attaching circles to every integer. The lifts $\widetilde{g}$ and $\widetilde{f}$ of f and g start at the same intersecting point of the two $\mathbb{R}$ lines, but then go in different directions. Hence the lifts $\widetilde{(f \# g)}(1) \neq \widetilde{(g \# f)}(1)$. Because homotopies in X lift to the covering space, it must be the case that $f \# g$ and $g \# f$ are *not* homotopic. Thus $\pi_1(X, x_0)$ is a non-abelian group.



Lifting a loop on the figure eight space [6, Figure 60.1].

Exercise 4.10. Show that if X is the figure eight space, then $\pi_1(X, x_0) \cong \mathbb{F}_2$, the free group with two generators.

Example 4.4.9 (The circle)**.** Using the covering map $p : (\mathbb{R}, 0) \to (\mathbb{T}, 1)$, we have seen that $p^{-1}(1) \cong \mathbb{Z} \cong \pi_1(\mathbb{T}, 1)$ as a set. To show this is a group isomorphism, we need to show that Φ is a group homomorphism. In this particular case, the map $\Phi([\alpha]) = \widetilde{\alpha}(1) = \mathrm{Wind}(\alpha) \in \mathbb{Z}$, the **winding number** of $\alpha \in \mathrm{Loop}(\mathbb{T}, 1)$. To complete the proof that $\pi_1(\mathbb{T}) \cong \mathbb{Z}$, it suffices to show that $\mathrm{Wind}(\alpha \# \beta) = \mathrm{Wind}(\alpha) + \mathrm{Wind}(\beta)$ for any $\alpha, \beta \in \mathrm{Loop}(\mathbb{T}, 1)$. We take $g, h : [0, 1] \to \mathbb{R}$ such that $g(0) = h(0) = 0$ and $\alpha(t) = e^{2\pi i g(t)}$ and $\beta(t) = e^{2\pi i h(t)}$. Then $\mathrm{Wind}(\alpha) = g(1)$, $\mathrm{Wind}(\beta) = g(1)$. Now combine these real-valued functions

$$f :\in C([0,1], \mathbb{R}), \qquad f(t) = \begin{cases} g(2t), & t \in \left[0, \frac{1}{2}\right], \\ g(1) + h(2t-1), & t \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Then by construction, $\alpha \# \beta(t) = e^{2\pi i f(t)}$, $\mathrm{Wind}(\alpha \# \beta) = f(1)$ and

$$\mathrm{Wind}(\alpha \# \beta) = f(1) = g(1) + h(1) = \mathrm{Wind}(\alpha) + \mathrm{Wind}(\beta).$$

Hence

$$\pi_1(\mathbb{T}, 1) \cong \pi_1(S^1, (1, 0)) \ni [\alpha] \mapsto \mathrm{Wind}(\alpha) \in \mathbb{Z}$$

is an isomorphism of groups.

Exercise 4.11. Extend the winding number to be defined for any map $\alpha \in \mathrm{Loop}(\mathbb{C} \setminus \{0\}, 1)$. Furthermore, show that $\pi_1(\mathbb{C} \setminus \{0\}, 1) \cong \mathbb{Z}$.

Exercise 4.12. Show that the n-torus is path connected and $\pi_1(\mathbb{T}^n) \cong \mathbb{Z}^n$.

4.5 Applications

We have shown that $\pi_1(S^1) \cong \mathbb{Z}$ and, hence, the circle is not contractible. While the proof was rather long, we can use this fact to prove many other interesting results.

We let $D^2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\} \simeq \{z \in \mathbb{C} \mid |z| \leq 1\}$ denote the unit disc with boundary $\partial D^2 = S^1 \cong \mathbb{T}$.

Proposition 4.5.1. *Let $f : D^2 \to \mathbb{T}$ be a continuous map such that $f(1) = 1$. Then the loop $\alpha \in \mathrm{Loop}(\mathbb{T}, 1)$ such that $\alpha(t) = f(e^{2\pi t})$ is such that $\mathrm{Wind}(\alpha) = 0$.*

Proof. If we define $\beta \in \mathrm{Loop}(D^2, 1)$ such that $\beta(t) = e^{2\pi i t}$, then $\alpha = f \circ \beta$. On the other hand, because D^2 is simply connected, $\pi_1(D^2) = \{e\}$ and the induced homomorphism $f_* : \pi_1(D^2, 1) \to \pi_1(\mathbb{T}, 1)$ must be such that $[\alpha] = f_*([\beta]) = [\gamma_1]$. Hence the winding number is trivial. □

Theorem 4.5.2. *There is no continuous map $f : D^2 \to \mathbb{T}$ such that $f(z) = z$ for all $z \in \mathbb{T}$.*

Proof. We assume such a map exists, then we can factorise $\mathrm{id}_{\mathbb{T}}(z) = f \circ \iota(z)$ where $\iota : \mathbb{T} \hookrightarrow D^2$

is the inclusion map. That is, the following diagram commutes

$$\begin{tikzcd}
(\mathbb{T},1) \arrow[r, "\iota"] \arrow[dr, "\mathrm{id}_{\mathbb{T}}"'] & (D^2,1) \arrow[d, "f"] \\
& (\mathbb{T},1).
\end{tikzcd}$$

We also have the induced maps $\iota_* : \pi_1(\mathbb{T},1) \to \pi_1(D^2,1)$ and $f_* : \pi_1(D^2,1) \to \pi_1(\mathbb{T},1)$ where $(f \circ \iota)_* = f_* \circ \iota_*$. So the following diagram also commutes

$$\begin{tikzcd}
\pi_1(\mathbb{T},1) \cong \mathbb{Z} \arrow[r, "\iota_*"] \arrow[dr, "(\mathrm{id}_{\mathbb{T}})_*"'] & \pi_1(D^2,1) \cong \{e\} \arrow[d, "f_*"] \\
& \pi_1(\mathbb{T},1) \cong \mathbb{Z}.
\end{tikzcd}$$

Therefore $f_* \circ \iota_*$ computes the identity isomorphism $\mathbb{Z} \cong \mathbb{Z}$. But both ι_* and f_* are zero homomorphism as D^2 is simply connected. So there is a contradiction. □

We can use the previous result to give a simple proof of Brower's fixed point theorem (Theorem 2.5.3) for the complex disc.

Theorem 4.5.3 (Brouwer's fixed point therem on the disc)**.** *Any continuous map $\phi : D^2 \to D^2$ has a fixed point.*

Proof. Assume $\phi : D \to D$ has no fixed point. We define $f : D \to S^1$ such that $f(z)$ is the point on $\mathbb{T}$ coming from the ray from $\phi(z)$ to z. That is, $|z + t(z - \phi(z))| = 1$ has a solution t_z and $f(z) = z + t_z(z - \phi(z))$ is continuous with $f(z) = z$ for all $z \in \mathbb{T}$. But this contradicts Theorem 4.5.2 and so ϕ can not exist. □

Vector fields

A vector field is a continuous function $F : \Omega \to \mathbb{R}^n$ where $S \subset \mathbb{R}^n$ is a subset. If $f : S \to \mathbb{R}$ is a differentiable function with continuous partial derivatives, then $\nabla f = (\partial_1 f, \ldots, \partial_n f)$ is a vector field, for example. A point $\mathbf{x}_0 \in S$ is a zero of a vector field F if $F(\mathbf{x}_0) = \mathbf{0}$.

Suppose that $n = 2$ and $F : S \to \mathbb{R}^2$ is a non-vanishing vector field, $F(x,y) \neq (0,0)$ for all $(x,y) \in S$. Then taking a base point $(x_0,y_0) \in S$ and $\gamma \in \mathrm{Loop}(S,(x_0,y_0))$, then $F \circ \gamma \in \mathrm{Loop}\left(\mathbb{R}^2 \setminus \{0\}, (x_0,y_0)\right)$. We can identify $\mathbb{R}^2 \simeq \mathbb{C}$ and $S \subset \mathbb{C}$ to obtain that $F \circ \gamma \in \mathrm{Loop}\left(\mathbb{C} \setminus \{0\}, z_0\right)$. Recalling Exercise 4.11, we can define $\mathrm{Wind}(F \circ \gamma)$ for any non-vanishing F and $\gamma \in \mathrm{Loop}(S,(x_0,y_0))$.

The proof of the following is an exercise (or see [4, Lemma 9.1]).

Lemma 4.5.4. *Let $F : \overline{B(\mathbf{0};R)} \to \mathbb{R}^2$ be a vector field such that F is non-vanishing on $\{(x,y) \mid x^2 + y^2 = R^2\}$. Consider the loop $\gamma \in \mathrm{Loop}\left(\overline{B(\mathbf{0};R)}, (1,0\right)$ given by $\gamma(t) = R(\cos(2\pi t), \sin(2\pi))$. If $\mathrm{Wind}(F \circ \gamma) \neq 0$, then there is some point $(x,y) \in B(\mathbf{0},R)$ such that $F(x,y) = 0$.*

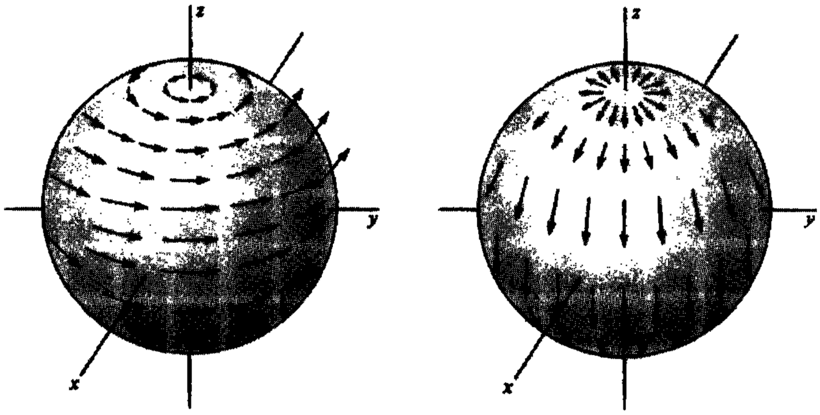
Let us now consider three dimensional vector fields and $S = S^2 = \left\{(x,y,z) \in \mathbb{R} \mid x^2 + y^2 + z^2 = 1\right\}$. We say that a vector field $F : S^2 \to \mathbb{R}^3$ is tangent to S^2 if $F(\mathbf{x})$ is orthogonal to $\mathbf{x}$ for all $\mathbf{x} \in S^2$.

Examples 4.5.5. 1. The vector field $F : S^2 \to \mathbb{R}^3$, $F(x, y, z) = (-y, x, 0)$ is tangent to S^2 and vanishes at the north and south poles $(0, 0, \pm 1)$. Geometrically, the vectors point west to east along the sphere away from the north and south poles.

2. Define $G : S^2 \to \mathbb{R}^3$ such that $G(x, y, z) = (xz, yz, z^2 - 1)$, where

$$\langle \mathbf{x}, G(\mathbf{x})\rangle = (x^2 + y^2 + z^2 - 1)z = 0$$

and G is tangent to S^2. Geometrically, the vectors point from north to south and, again, vanish at the north/south poles.



$F(x, y, z) = (-y, x, 0)$ $G(x, y, z) = (xz, yz, z^2 - 1)$

Vector fields tangent to the sphere [4, p149].

In both examples above, the tangent vector field had zeros at the north and south poles. This is unavoidable as a consequence of the the hairy ball/hedgehog theorem: you can't comb a hairy ball flat without some part sticking up.

Theorem 4.5.6 (Hairy ball theorem). *Every vector field tangent to S^2 has a zero.*

The idea of the proof is that observing a constant vector field near the north pole will look different depending on whether we are at the north pole or south pole (namely one has winding number 0 and the other winding number 2). To make this precise, we need to use the stereographic projection $\sigma : S^2 \setminus \{N\} \to \mathbb{R}^2$.

Proof. We first review the stereographic projection and inverse,

$$\sigma : S^2 \setminus \{N\} \to \mathbb{R}^2, \qquad \sigma(x, y, z) = \Big(\frac{x}{1-z}, \frac{y}{1-z}\Big),$$

$$\sigma^{-1} : \mathbb{R}^2 \to S^2 \setminus \{N\}, \qquad \sigma^{-1}(u, v) = \Big(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1}\Big).$$

Now take $F : S^2 \to R^3$ a vector field tangent to S^2, $F(\mathbf{x}) = \big(F_1(\mathbf{x}), F_2(\mathbf{x}), F_3(\mathbf{x})\big)$, and define $\tilde{F} : S^2 \to \mathbb{R}^3$ and $G : \mathbb{R}^2 \to \mathbb{R}^2$, where $G(u, v) = \tilde{F}\big(\sigma^{-1}(u, v)\big)$ and

$$\tilde{F}(x, y, z) = \big((1-z)F_1(x, y, z) + xF_3(x, y, z),\ (1-z)F_2(x, y, z) + yF_3(x, y, z)\big).$$

If $G(u, v) = \mathbf{0}$ for some $(u, v) \in \mathbb{R}^2$, then we obtain the equations

$$(1-z)F_1 + xF_3 = 0 = (1-z)F_2 + yF_2.$$

Because F is tangent to S^2, $xF_1 + yF_2 = -zF_3$ and $x^2 + y^2 = 1 - z^2$, we do some algebra

$$\begin{aligned} 0 &= x(1-z)F_1 + x^2F_3 + y(1-z)F_2 + y^2F_3 = (1-z)(xF_1 + yF_2) + (1-z^2)F_3 \\ &= \big((1-z^2) - z(1-z)\big)F_3 = (1-z)F_3. \end{aligned}$$

Because $(x, y, z) \in \sigma^{-1}(\mathbb{R}^2) = S^2 \setminus \{N\}$, $z \neq 1$. Hence $F_3(x, y, z) = 0$, which then implies that $F_1 = F_2 = 0$. Conversely, if $F(x, y, z) = 0$, then $G(u, v) = 0$. Therefore (x, y, z) is a zero of F if and only if $\sigma(x, y, z)$ is a zero of G.

Now take $\widehat{F} : S^2 \to \mathbb{R}^3$ such that $\widehat{F}(x, y, z) = (z, 0, -x)$, which is tangent to S^2. Applying the procedure above, we get a two-dimensional vector field $\widehat{G} : \mathbb{R}^2 \to \mathbb{R}^2$,

$$\widehat{G}(\sigma(x, y, z)) = \big(z - z^2 - x^2, -xy\big),$$

which can be written in polar coordinates as

$$\widehat{G}(r\cos(\theta), r\sin(\theta)) = \frac{-2r^2}{(r^2+1)^2}\Big(\frac{1}{r^2} + \cos(2\theta),\, \sin(2\theta)\Big).$$

If $r > 1$, $G \neq \mathbf{0}$ and so we can take the loop $\gamma_r(\theta) = (r\cos(\theta), r\sin(\theta))$, where

$$\operatorname{Wind}\big(\widehat{G} \circ \gamma_r\big) = \operatorname{Wind}\Big(\frac{1}{r^2} + \cos(2\theta),\, \sin(2\theta)\Big) = 2.$$

Now suppose $F : S^2 \to \mathbb{R}^3$ is a vector field tangent to S^2 such that $F(0, 0, 1) = \widehat{F}(0, 0, 1) = (1, 0, 0)$. Then $(F \cdot \widehat{F}) : S^2 \to \mathbb{R}$ is a continuous function such that $(F \cdot \widehat{F})(0, 0, 1) = 1$. So by continuity there is a neighbourhood U of $(0, 0, 1)$ such that $F \cdot \widehat{F} > 0$ on U. We can also take U such that $F \cdot F$ and $\widehat{F} \cdot \widehat{F}$ are strictly positive on U. Hence the convex combination $tF + (1-t)\widehat{F}$ will have no zeros on U for all $t \in [0, 1]$.

Considering the transformed vector fields G and $\widehat{G}$ on $\mathbb{R}^2$, $tG + (1-t)\widehat{G}$ is a vector field on $\mathbb{R}^2$ that contains no zeros in $\sigma(U) \subset \mathbb{R}^2$. Because U is a neighbourhood of the north pole, $\sigma(U)$ is a neighbourhood of infinity (in the one-point compactification) and we can take a sufficiently large R such that $(u, v) \in \sigma(U)$ for all $u^2 + v^2 \geq R$. In particular $G(R\cos(\theta), R\sin(\theta))$ and $\widehat{G}(R\cos(\theta), R\sin(\theta))$ are non-vanishing loops in $\mathbb{R}^2 \setminus \{0\}$. We have already seen that $\operatorname{Wind}(\widehat{G}) = 2$. Furthermore, $(1-t)G + t\widehat{G}$ gives a homotopy between G and $\widehat{G}$, hence $\operatorname{Wind}(G) = 2$. Applying Lemma 4.5.4, G has a zero inside $B(\mathbf{0}; R)$ and so F must have a zero in $\sigma^{-1}(B(\mathbf{0}; R))$.

Now take an arbitrary tangent vector field $H : S^2 \to \mathbb{R}^3$. If $H(0, 0, 1) = \mathbf{0}$, then H has a zero. Otherwise, take a rotation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $T\big(H(0, 0, 1)\big) = (c, 0, 0)$ with $c \neq 0$. Applying the previous paragraph, the transformed vector field $\frac{1}{c}T \circ F \circ T^{-1}$ has a zero. Because rotations are homeomorphisms, it follows that F must have a zero. □

What is interesting about the proof of the above result is that it mostly relied on our knowledge that $\pi_1(S^1) \cong \mathbb{Z}$ via the winding number, despite being a result about higher dimensional spaces.

5 Elements of algebraic topology

In the previous section, we associated a group $\pi_1(X)$ to a topological space. We also showed that if we can compute $\pi_1(X)$ in a few cases, we can potentially learn many non-trivial facts about topological spaces. In this section, we continue this train of thought and consider algebraic objects that can be associated to topological spaces.

5.1 Homotopy groups

In the previous chapter, we investigated the set $\big[(S^1,(1,0),(X,x_0)\big] = \pi_1(X,x_0)$, which can be given the structure of a group. There is an obvious generalisation, where for $n \geq 2$ we let N be the north pole of S^n and define $\pi_n(X,x_0) = \big[(S^n,N),(X,x_0)\big]$. We can import many of the ideas from our study of $\pi_1(X,x_0)$ to the higher dimensional setting.

Let us first consider $n = 0$ and $S^0 = \{-1,1\}$, which has different behaviour to the other degrees. In particular any element

$$[f] \in \big[(S^0,1),(X,x_0)\big] = \big\{f : \{-1,1\} \to X \mid f(1) = x_0\big\}/\sim_h$$

is entirely determined by $f(-1)$ as $f(1)$ is fixed. Note that because S^0 is discrete, any function $f : S^0 \to X$ is continuous.

Proposition 5.1.1. *Let $(X,\mathcal{T})$ be a topological space. There is a bijection between the set of path components of X and $\pi_0(X)$.*

Proof. Two functions $f,g : \{-1,1\} \to X$ are pointed homotopic with $f(1) = g(1) = x_0$ if and only if $f(-1) \sim_\gamma g(-1)$. Hence $[f] = [g]$ if and only if $g(-1) \in C^\gamma_{f(-1)}$, the path connected component of $f(-1)$. Hence we have a well-defined and injective map

$$\pi_0(X) \ni [f] \mapsto C^\gamma_{f(-1)} \in X/\sim_\gamma.$$

The map is surjective as given $x \in X$, we can define $f(-1) = x$, $f(1) = x_0$ which will map to C^γ_x. □

On the other hand, because S^0 is disconnected, we can *not* give $\pi_0(X)$ a group structure in general. Indeed, there is no obvious way to compose or invert path connected components of a space.

Let us now move on to $\pi_n(X,x_0)$ for $n \geq 2$. Like we did for π_1, we will use a presentation of S^n as a quotient of a closed product of intervals,

$$I^n := [0,1]^n, \qquad S^n \simeq I^n/\partial I^n.$$

Hence we think of an element $(t_1,\dots,t_n) \in S^n$ as $t_1,\dots,t_n \in [0,1]$ and where we identify all boundary points

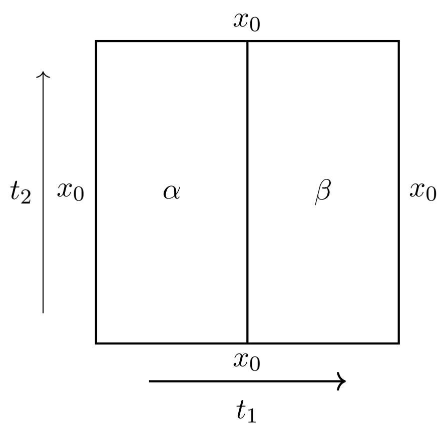
$$(0,t_2,\dots,t_n) \sim (1,t_2,\dots,t_n) \sim (t_0,0,\dots,t_n) \sim (t_0,1,\dots,t_n) \sim \dots \sim (t_0,\dots,t_{n-1},0) \sim (t_1,\dots,t_{n-1},1).$$

Note that because I^n and S^n is path connected, any two paths in I^n and S^n with fixed endpoints will be homotopic (with fixed endpoints). This will give us flexibility to construct products and inverses in $\pi_n(X,x_0)$. Indeed, it will allow us to focus on one variable t_1 for concatenation and reverse paths.

If $\alpha,\beta \in \pi_n(X,x_0)$, we therefore define

$$\alpha\#\beta(t_1,\dots,t_n) = \begin{cases} \alpha(2t_1,t_2,\dots,t_n), & t_1 \in [0,\frac{1}{2}], \\ \beta(2t_1-1,t_2,\dots,t_n), & t_1 \in [\frac{1}{2},1], \end{cases} \qquad \alpha^-(t_1,\dots,t_n) = \alpha(1-t_1,t_2,\dots,t_n).$$

That is, we concatenate and reverse in the first variable and leave the others fixed. This will allow us to mimic what we did for the case of π_1 for more general n. For $n = 2$, we can visually represent the product by considering a function $\alpha, \beta : [0, 1] \times [0, 1] \to X$ with all boundary points sent to x. Then, the concatenation can be geometrically seen as follows:



Note that the concatenation takes value x_0 along the center vertical line as well.

Theorem 5.1.2. *For $n \geq 1$, $\pi_1(X, x_0)$ can be given the structure of a group with identity $e = \gamma_{x_0}$ with $\gamma_{x_0}(t_1, \ldots, t_n) = x_0$ and operations*

$$[\alpha] * [\beta] = [\alpha\#\beta], \qquad [\alpha]^{-1} = [\alpha^-].$$

Proof. By working with t_1 and keeping $(t_2, \ldots, t_n)$ fixed, we can basically repeat the proof of Theorem 4.3.5 without any problems. For example, The proof that $\alpha \sim_h^{x_0} \tilde{\alpha}$ and $\beta \sim_h^{x_0} \tilde{\beta}$ implies $\alpha\#\beta \sim_h^{x_0} \tilde{\alpha}\#\tilde{\beta}$ and $\alpha^- \sim_h^{x_0} \tilde{\alpha}^-$ is the same as π_1. Namely, if F and G are the homotopies, we take

$$(F\#G)(t_1, \ldots, t_n, s) = \begin{cases} F(2t_1, t_2, \ldots, t_n, s), & t_1 \in [0, \frac{1}{2}], \\ G(2t_1 - 1, t_2, \ldots, t_n, s), & t_1 \in [\frac{1}{2}, 1], \end{cases}$$
$$F^-(t_1, \ldots, t_n, s) = \alpha(1 - t_1, t_2, \ldots, t_n, s).$$

Similarly for the product, $\alpha\#(\beta\#\gamma)(t_1, \ldots, t_n) = (\alpha\#\beta)\#\gamma(g(t_1, \ldots, t_n))$, where

$$g : I^n \to I^n, \quad g(t) = \begin{cases} (\frac{1}{2}t_1, t_2, \ldots, t_n), & t_1 \in \left[0, \frac{1}{2}\right], \\ (t_1 - \frac{1}{4}, t_2, \ldots, t_n), & t_1 \in \left[\frac{1}{2}, \frac{3}{4}\right], \\ (2t_1 - 1, t_2, \ldots, t_n), & t_1 \in \left[\frac{3}{4}, 1\right]. \end{cases}$$

and $\alpha \circ g \sim_h^{x_0} \alpha$ for such a g. Thus the product is well-defined.
Still following Theorem 4.3.5, we can write

$$\gamma_{x_0}\#\beta(t_1, \ldots, t_n) = \beta \circ g(t_1, \ldots, t_n), \qquad g(t) = \begin{cases} 0, & t_1 \in \left[0, \frac{1}{2}\right], \\ (2t_1 - 1, t_2, \ldots, t_n), & t_1 \in \left[\frac{1}{2}, 1\right], \end{cases}$$

and $[\gamma_{x_0}\#\beta] = [\beta]$. A similar argument will show that $[\beta\#\gamma_{x_0}] = [\beta]$. Hence $[\gamma_{x_0}]$ acts as the group identity under concatenation of loops.

Lastly, we check that $[\alpha] * [\alpha^-] = [\alpha \# \alpha^-] = [\gamma_{x_0}]$. To do this we define the homotopy

$$F : I^n \times [0,1] \to X, \qquad F(t_1, \dots, t_n, s) = \begin{cases} \alpha(2t_1, t_2, \dots, t_n), & t_1 \in \left[0, \frac{s}{2}\right], \\ \alpha(s, t_2, \dots, t_n), & t_1 \in \left[\frac{s}{2}, 1 - \frac{s}{2}\right], \\ \alpha(2 - 2t_1, t_2, \dots, t_n), & t_1 \in \left[1 - \frac{s}{2}, 1\right]. \end{cases}$$

That is, $\alpha_s = F(\cdot, s)$ describes a function where we move along α at double speed until time $s/2$, rest until time $1 - s/2$ and then return along α^- at double speed. By construction F is continuous and $\alpha_0(t) = \alpha(0) = x_0 = \gamma_{x_0}(t)$. We also find that

$$\alpha_1(t_1, \dots, t_n) = F(t_1, \dots, t_n, 1) = \begin{cases} \alpha(2t_1, t_2, \dots, t_n), & t_1 \in \left[0, \frac{1}{2}\right], \\ \alpha(2 - 2t_1, t_2, \dots, t_n)), & t_2 \in \left[\frac{1}{2}, 1\right], \end{cases} = \alpha \# \alpha^-(t_1, \dots, t_n),$$

which shows $[\alpha] * [\alpha^-]$ is the identity. A similar argument shows $[\alpha^-] * [\alpha] = [\gamma_{x_0}]$. □

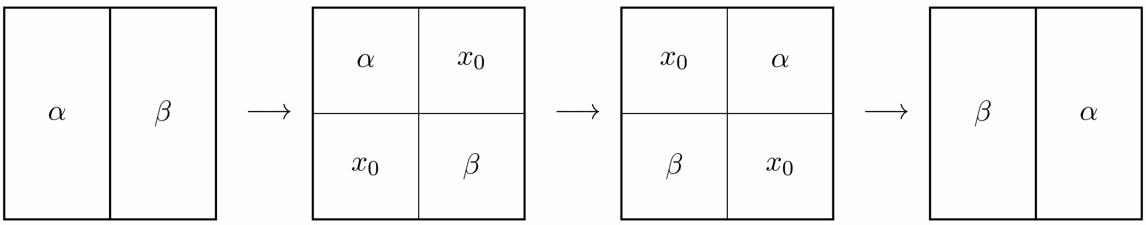
The groups $\{\pi_n(X, x_0)\}_{n \geq 1}$ are called the **homotopy groups** of the space X. While their construction is very analogous to the fundamental group, the higher homotopy groups have some properties that the fundamental group does not.

Theorem 5.1.3. *The homotopy group $\pi_n(X, x_0)$ is abelian when $n \geq 2$.*

Proof. We first consider the case $n = 2$ and the 'geometric idea'. Namely, the second variable t_2 will no longer be passive and we can use it to slide α and β past each other. The basic idea is as follows:

1. Replace $\alpha \# \beta$ with $(\alpha \# \gamma_0) \# (\gamma_0 \# \beta)$ where we concatenate with γ_0 *in the second variable* (we can do this via a homotopy),
2. Slide $\alpha \# \gamma_0$ and $\gamma_0 \# \beta$ past each other (using the first variable),
3. Return $(\gamma_0 \# \beta) \# (\alpha \# \gamma_0)$ to $\beta \# \alpha$.

Diagramatically, this can be seen by the following steps



For step 2 to step 3, we shrink x_0 on the top-left corner, while adding x_0 in the top right corner. We also shrink x_0 in the bottom-right corner while adding x_0 in the bottom-left corner.

To prove the result in $n \geq 3$ we do the same steps as $n = 2$ and leave the remaining variables $(t_3, \dots, t_n)$ fixed. □

We also consider the functorial and homotopy invariance properties of the homotopy groups. The proof is the same as the case of the fundamental group.

Theorem 5.1.4. *Let $(X, \mathcal{T})$ be a topological space and $x_0 \in X$. Then for all $n \geq 1$:*

1. *If $x_1 \in C^{\gamma}_{x_0}$, the path connected component of x_1, then $\pi_n(X, x_0) \cong \pi_n(X, x_1)$.*
2. *If $f : C\big((X, x_0), (Y, y_0)\big)$, then there is an induced group homomorphism*
$$f_* : \pi_n(X, x_0) \to \pi_n(X, x_0), \qquad \pi_n([\alpha]) = [f \circ \alpha].$$
If $g \in C\big((Y, y_0), (Z, z_0)\big)$, then $(g \circ f)_* = g_* \circ f_*$.
3. *If $f : C(X, Y)$ is a homotopy equivalence, then the induced map $f_* : \pi_n(X, x_0) \to \pi_n(Y, f(x_0))$ is a group isomorphism.*

Exercise 5.1. If $f, g : C\big((X, x_0), (Y, y_0)\big)$ are homotopic with basepoint fixed, show that $f_* = g_*$.

Example 5.1.5 (Homtotopy groups of S^1)**.** We have already seen that $\pi_0(S^1) \cong \{x_0\}$ (path connected) and $\pi_1(S^1) \cong \mathbb{Z}$ by the winding number. We now claim that $\pi_n(S^1) = \{e\}$ if $n \geq 2$. First consider $n = 2$ and a continuous map $f : I^2 \to S^1$ such that $f(t_1, t_2) = (1, 0)$ for all $(t_1, t_2) \in \partial I^2$. By the homotopy lifting theorem (Theorem 4.4.5), we have a map $\tilde{f} : I^2 \to \mathbb{R}$. Considering the fibre at the base point $f(t_1, t_2) \in \mathbb{Z}$ for all $(t_1, t_2) \in \partial I^2$. On the other hand, ∂I^2 is connected, so $\tilde{f}(t_1, t_2)$ must be constant for all $(t_1, t_2) \in \partial I^2$. Because $\mathbb{R}$ is convex, we can homotopy $\tilde{f}$ to $g(t_1, t_2) = f(0, 0) \in \mathbb{Z}$. Composing this homotopy with the covering map $p : \mathbb{R} \to S^1$, we get a homotopy from f to a constant map, which must therefore be the map γ_{x_0}. This shows that $\pi_2(S^1)$ is trivial. For higher dimensions, we can apply the same argument and leave $(t_3, \ldots, t_n)$ fixed to show that $\pi_n(S^1) = \{e\}$ for all $n \geq 2$.

Let us now consider the homotopy groups of higher dimensional spheres. We will not prove the following result, though will make some comments.

Theorem 5.1.6. *For any $n \geq 1$, $\pi_i(S^n) = \{e\}$ for $i < n$ and $\pi_n(S^n) \cong \mathbb{Z}$.*

Example 5.1.5 may lead one to think that $\pi_{n+k}(S^n)$ is trivial for $n \geq 2$ and $k \geq 1$, but this is not true.

Example 5.1.7 (The Hopf fibration)**.** We construct a continuous and non-contractible map $H : S^3 \to S^2$. We use the presentation of S^3 as unit vectors $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. We can then identify $\mathbb{R}^4 \cong \mathbb{C}^2$ via $z_1 = x_1 + ix_2$ and $z_2 = x_3 + ix_4$. Hence $S^3 \simeq \big\{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\big\}$. We then define the map
$$H : S^3 \to \mathbb{C} \cup \{\infty\}, \qquad H(z_1, z_2) = \begin{cases} \frac{z_1}{z_2}, & z_2 \neq 0, \\ \infty, & z_2 = 0. \end{cases}$$
Because $\mathbb{C} \cup \{\infty\} \simeq (\mathbb{C})^+ \simeq (\mathbb{R}^2)^+ \simeq S^2$, we therefore have a (continuous) map $H : S^3 \to S^2$. It is an exercise to show that, for $z \in \mathbb{C} \cup \{\infty\}$, $H^{-1}(z) \cong \mathbb{Z}$. This can then be used to show that $[H] \neq [\gamma_N] \in \pi_3(S^2)$ and, furthermore, $\pi_3(S^2) \cong \mathbb{Z}$.

The question of computing $\pi_{n+k}(S^n)$ for $k \geq 1$ is in general extremely difficult and many higher order homotopy groups of spheres are still unknown. So while the homotopy groups of a space $\pi_n(X, x_0)$ are relatively easy to define, they are not always so easy to compute (and often are quite difficult). This is one reason why we often consider other so-called homology theories of topological spaces, which also associate (generally abelian) groups to topological spaces. The definition of these alternative homology theories are more involved, but often come with more properties that allow us to compute the groups associated to the topological space of interest.

Lastly, we will say a few words about the isomorphism $\pi_n(S^n) \cong \mathbb{Z}$. Given any function $f \in C(S^n, S^n)$, the integer $\deg(f) \in \mathbb{Z}$ that corresponds to $[f] \in \pi_n(S^n) \cong \mathbb{Z}$ is called the *degree* of the map f.

We would like an explicit expression for $\deg(f)$. This is not so easy for an arbitrary continuous function $f : S^n \to S^n$, but it is possible if we instead consider *differentiable* or smooth (infinitely differentiable) functions. This can be done if we import a little technology from differential geometry.

Lemma 5.1.8. *The set $C^\infty(S^n, S^n)$ of smooth functions is dense in $C(S^n, S^n)$. Furthermore for any $f \in C(S^n, S^n)$, there is some $g \in C^\infty(S^n, S^n)$ such that $f \sim_h g$ in $C(S^n, S^n)$.*

Proof (Sketch). Because S^n is compact and Hausdorff, $C(S^n, S^n)$ is a metric space with $d(f_1, f_2) = \sup_{x\in S^n} d_{S^n}\big(f_1(x), f_2(x)\big)$. Then a more involved version of the Stone–Weierstrass Approximation Theorem will give a sequence $g_n \in C^\infty(S^n, S^n)$ with $g_n \to f$. Given the smooth function g_n with n sufficiently large, we can then take the homotopy

$$F \in C(S^n \times [0,1], S^n), \qquad F(x,t) = \frac{(1-t)f(x) + tg_n(x)}{\|(1-t)f(x) + tg_n(x)\|}. \qquad \square$$

Using the previous lemma, let us now assume that $f : S^n \to S^n$ is smooth. We can naturally embed $S^n \subset \mathbb{R}^{n+1}$.

Treating f as a vector field in $\mathbb{R}^{n+1}$, we can compute the Jacobian matrix $\mathcal{J}_f(\mathbf{p})$ for any $\mathbf{x} \in S^n$. We say that $\mathbf{p}$ is non-singular if $\det\big(\mathcal{J}_f(\mathbf{p})\big) \neq 0$ and orientation preserving/reversing if $\det\big(\mathcal{J}_f(\mathbf{p})\big)$ is strictly positive/strictly negative. We say that $\mathbf{p}$ is a singular point if $\det\big(\mathcal{J}_f(\mathbf{p})\big) = 0$. Considering the image, $\mathbf{q} \in S^n$ is a non-critical point if every point in $f^{-1}(\mathbf{q})$ is non-singular (otherwise we say that $\mathbf{q}$ is a critical point).

Lemma 5.1.9 (Sard's Theorem). *If $f \in C^\infty(S^n, S^n)$, then the set of non-critical points for f is open and dense in S^n.*

Theorem 5.1.10. *Let $f \in C^\infty(S^n, S^n)$ and $X_f = \{\mathbf{q} \in S^n \text{ non-critical for } f\}$. Then, for $\mathbf{q} \in X_f$, the integer*

$$\big|\{\mathbf{p} \in f^{-1}(\mathbf{q}) \mid \det(\mathcal{J}_f(\mathbf{p})) > 0\}\big| - \big|\{\mathbf{p} \in f^{-1}(\mathbf{q}) \mid \det(\mathcal{J}_f(\mathbf{p})) < 0\}\big| \in \mathbb{Z}$$

is constant on X_f. Furthermore, this integer computes the value $\deg(f)$ under the isomorphism $\pi_n(S^n) \cong \mathbb{Z}$.

To more easily access tools from several variable calculus, let us extend f to a function

$$\hat{f} : S^n \to \mathbb{R}^{n+} \setminus \{\mathbf{0}\}, \qquad \hat{f}(\mathbf{x}) = \|\mathbf{x}\|\, f\Big(\frac{\mathbf{x}}{\|\mathbf{x}\|}\Big).$$

Because f is smooth, so is $\hat{f}$. The extension $\hat{f}$ will map the sphere of radius R to itself. We can also compute the degree via the extended function $\hat{f} : S^n \to \mathbb{R}^{n+1} \setminus \{\mathbf{0}\}$.

Theorem 5.1.11. *Let $f \in C^\infty(S^n, S^n)$. Then the integral*

$$\frac{1}{\mathrm{Vol}(S^n)} \int_{S^n} \det\big(\mathcal{J}_{\hat{f}}(\mathbf{x})\big)\, \mathrm{d}V = \deg(f),$$

where $\mathrm{d}V$ is the n-dimensional Euclidean volume and $\mathrm{Vol}(S^n) = \int_{S^n} \mathrm{d}V$.

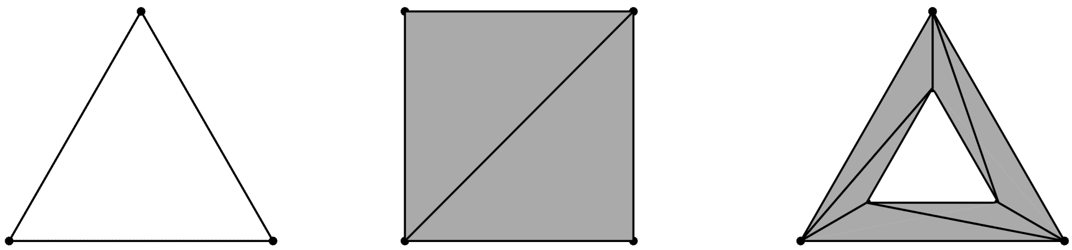


Figure 7: Simplicial complexes that are homeomorphic to S^1, a solid square and an annulus [2, p117].

We will not prove either of the two theorems above (though it is not so difficult to prove in the case $d = 1$ at least). Unlike the abstract definition of $\deg(f)$ in terms of homotopy groups, the formulas above are much easier to compute in practice as we can use techniques from differential calculus. This idea persists in more general algebraic topology: we are interested in defining purely topological quantities, but often want to find representatives involving differentiable/smooth quantities which are amenable to computation (e.g. de Rham cohomology). A classic (advanced) book on this subject is [1].

5.2 Simiplicial Complexes and the Euler characteristic

The Euler characteristic plays an analogous role to 'counting the number of holes' in a topological space. So we expect to be able to use it to distinguish between different topological spaces. The Euler characteristic is not a group, but is simply a number (an integer). This means that finding the Euler characteristic of spaces X and Y is in principle much easier than finding $[X, Y]$ for example.

On the other hand, to compute the Euler characteristics, we need to be able to decompose a space into 0-dimensional, 1-dimensional, 2-dimensional pieces et cetera. We also need a systematic procedure to construct spaces from this basic data. Hence we first restrict our attention to a particularly nice sub-class of topological spaces, simplicial complexes, which are embedded within Euclidean space $\mathbb{R}^n$.

Roughly speaking, a simplicial complex is a space that is built from points, lines and triangles (and higher-dimensional versions). For example, to build something one-dimensional, we take two points v_0 and v_1 (0-dimensional) and the straight line between them (convex set containing v_0 and v_1). To build something two dimensional, we combine one-dimensional lines and enclose a region.

On the other hand, some care is needed to make this procedure systematic. If we take points v_0, v_1, v_2, v_3 along the same line, then joining this line will not increase our dimension. Hence we need to choose points $v_0, v_1, \ldots, v_k \in \mathbb{R}^n$ such that

$$v_1 - v_0, \quad v_2 - v_1, \quad \ldots, \quad v_k - v_{k-1}$$

are all linearly independent. Then the convex subspace V_k that contains $v_0, v_1, \ldots, v_k$ will be k-dimensional.

Definition 5.2.1 (k-simplex). Let $v_0, \ldots, v_k \in \mathbb{R}^n$ be such that $v_1 - v_0, v_2 - v_1, \ldots, v_k - v_{k-1}$ are linearly independent. A k-simplex is the smallest convex subset containing $v_0, v_1, \ldots, v_k$. Specifically, we write

$$[v_0, \ldots, v_k] = \big\{t_0 v_0 + \ldots + t_k v_k \in \mathbb{R}^n \mid t_0, \ldots, t_k \in [0, 1] \text{ such that } t_0 + \ldots t_k = 1\big\}.$$

The points $v_0, \ldots v_k$ are the called vertices of the k-simplex. The coefficients $t_0, \ldots, t_k \in [0, 1]$ are called the barycentric coordinates of the point $\sum_{j=0}^{k} t_j v_j \in [v_0, \ldots, v_k]$.

Examples 5.2.2. 1. A 0-simplex $[v_0]$ is the singleton $\{v_0\} \subset \mathbb{R}^n$,

2. A 1-simplex $[v_0, v_1]$ is the (closed) line segment that joints v_0 and v_1.

3. A 2-simplex $[v_0, v_1, v_2]$ is a (filled-in) triangle with vertices v_0, v_1, v_2.

4. A 3-simplex $[v_0, v_1, v_2, v_3]$ is a tetrahedron with vertices v_0, v_1, v_2, v_3.

If $[v_0, \dots, v_k]$ is a k-simplex, then we can take $v_{j_0}, \dots, v_{j_l}$ from this generating set of vertices to form an l-simplex, which we call a subsimplex. If $l = (k-1)$, then we call the subsimplex $[v_{j_0}, \dots, v_{j_{k-1}}]$ a face. Any k-simplex has $k+1$ faces and $2^{k+1} - 1$ subsimplices.

Example 5.2.3. A 2-simplex (triangle) $[v_0, v_1, v_2]$ has three faces: $[v_0, v_1]$, $[v_0, v_2]$ and $[v_1, v_2]$ (the sides), and three 0-subsimplices $[v_0]$, $[v_1]$ and $[v_2]$ (the corners)

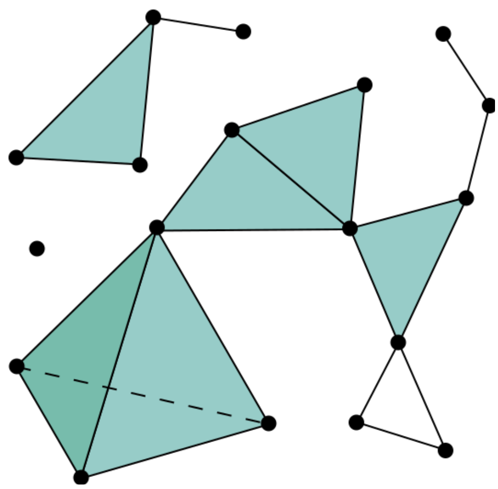
The boundary of a k-simplex is the union of its faces and, for $k > 0$, the interior is the complement of the boundary. The interior of a 0-simplex $[v]$ is the whole 0-simplex (which is a point). We also use the terminology simplex to denote a k-simplex of any k.

Definition 5.2.4 (Simplicial complex)**.** A simplicial complex is a subset $K \subset \mathbb{R}^n$ and finite collection of simplices such that:

1. The set K the union of the collection of simplices. Furthermore, each point in K lies in the interior of only one simplex.

2. The face of every simplex in the collection is also in the collection.

If K has a k-simplex and no n-simplices for $n > k$, we say that K is a k-dimensional simplicial complex.



A 3-dimensional simplicial complex, from Wikipedia.

The written definition of a simplicial complex is a little difficult to absorb. But the key point is that a simplicial complex is a union of simplices which are disjoint or intersect along a face or subsimplex. Note that any simplex is compact and hence any simplicial complex is compact.

Proposition 5.2.5. *Let K be a simplicial complex which includes the simplices S and T. Then $S \cap T$ is either empty or a subsimplex of S and T.*

Proof. We suppose that S is an i-simplex and T is a j-simplex. We will prove the result via induction on $i+j$. If $i+j=0$, then both S and T are 0-simplices, i.e. points. Either these points coincide and $S=T$ or they are different points and $S\cap T=\emptyset$.

Suppose now the result holds for $i+j<n$. If $S\cap T$ is non-empty, then any point $x\in S\cap T$ must be in on the boundary of S or T as $x\in K$ and can not lie in the interior of two simplices. If $x,y\in S\cap T$, then $(1-t)x+ty\in S\cap T$ as both S and T are convex. So if, for example, x is in the interior of S and y is in the interior of T, then the midpoint $\frac{1}{2}(x+y)$ will be in the interior of both S and T, which contradicts the definition of a simplicial complex. Hence $S\cap T$ is either in the boundary of S or the boundary of T. We suppose $S\cap T\subset\partial T$ with ∂T a union of the faces of T.

We claim that $S\cap T$ must lie within one of the faces of T. To see this, suppose $x,y\in S\cap T$ with $x\in[v_0,\cdots,v_k]$, $y\in[w_0,\ldots,w_m]$ and $v_1,\ldots,v_k,w_0,\ldots,w_m$ vertices of T. If T has a vertex outside of the set $\{v_1,\ldots,v_k,w_0,\ldots,w_m\}$, then x and y are in the face of T that omits this vertex. On the other hand, if $\{v_1,\ldots,v_k,w_0,\ldots,w_m\}$ are all the vertices of T, then we relabel as $z_0,\ldots,z_j$ and $x=t_0z_0+\ldots+t_jz_j$ and $y=s_0z_0+\ldots s_jz_j$. Then the midpoint

$$\frac{x+y}{2}=\frac{t_0+s_0}{2}z_0+\ldots+\frac{t_j+s_j}{2}z_j$$

is non-zero and so lies in the interior of T. But $\frac{1}{2}(x+y)$ also lies in $S\cap T$, which is a contradiction. This proves the claim.

Because $S\cap T$ is contained in a face of T, it is an intersection of an i-simplex with a $(j-1)$-simplex. By the inductive hypothesis, $S\cap T$ is a subsimplex of S and the the relevant face of T, which is already a subsimplex of T by the definition of simplicial complexes. This shows the result for $i+j=n$. $\square$

Now that we have defined simplicial complexes, spaces built from k-simplices, we can associate a number that comes from the data used to construct the space.

Definition 5.2.6. Let T be a n-dimensional simplicial complex. For each $0\le k\le n$, let i_k denote the number of k-simplices in T. The Euler characteristic of T is then defined as

$$\chi(T)=i_0-i_1+i_2-\ldots+(-1)^n i_n=\sum_{k=0}^{n}(-1)^k i_k\in\mathbb{Z}.$$

Examples 5.2.7. Trivially, we see that any point $\{v_0\}$ has Euler characteristic 1. Let's consider the simplicial complexes from Figure 7.

1. The simplicial circle (triangle) as three 0-simplices and three 1-simplices. Hence the Euler characteristic is $3-3=0$.

2. The simplicial square has four 0-simplicies, five 1-simplices and two 2-simplices. Hence $\chi=4-5+2=1$.

3. The simplicial annulus has six 0-simplices, twelve 1-simplices and six 2-simplices. Therefore $\chi=6-12+6=0$.

We would like to extend the definition of Euler characteristic to more general topological spaces. We do this by considering spaces which topologically admit a description as a simplicial complex.

Definition 5.2.8. A triangulation of a space $(X, \mathcal{T})$ is a simplicial complex K and a homeomorphism $f : X \xrightarrow{\simeq} K$. If X has a triangulation, then we define $\chi(X) = \chi(K)$.

Examples 5.2.9.

1. We can triangulate S^1 via three 1-simplices joined at the three 0-simplices (a triangle). Then the Euler number $\chi(S^1) = 3 - 3 = 0$.
2. The circle S^1 is also homeomorphic to the unit square with vertices $(\pm 1, \pm 1)$ and $(\pm 1, \mp 1)$. This is a simplicial complex with four 1-simplices and four 0-simplices. Hence $\chi(S^1) = 4 - 4 = 0$.
3. We can triangulate S^2 as a hollow tetrahedron. Hence it is determined by four 2-simplices, six 1-simplices and four 0-simplices. Thus $\chi(S^2) = 4 - 6 + 4 = 2$.

We saw in the examples above that different triangulations give the same Euler characteristic. Certainly this would be a problem if this did not happen. This is a very special case of a much more general result, whose proof is beyond the scope of this course.

Theorem 5.2.10. *If X and Y are homotopy equivalent topological spaces which admit triangulations, then $\chi(X) = \chi(Y)$.*

Corollary 5.2.11. *If X and Y are topological spaces which admit triangulations with $\chi(X) \neq \chi(Y)$, then X is not homotopic to Y.*

We can use this result to rigorously show that S^2 is not homotopic (let alone homeomorphic) to a torus $\mathbb{T}^2$. Similarly, any space X with $\chi(X) \neq 1$ is not contractible. Hence we can learn a lot about a space by taking a triangulation and then computing its Euler characteristic.

The Euler characteristic also plays a key role in the classification of surfaces, Hausdorff topological spaces X such that for every $x \in X$ there is an open neighbourhood $x \in U \in \mathcal{T}$ and a homeomorphism from U to an open ball in $\mathbb{R}^2$. The sphere, torus and Klein bottle are all examples of surfaces.

Theorem 5.2.12 (Classification of surfaces)**.** *Let S and T be surfaces that admit triangulations. Then S is homeomorphic to T if and only if $\chi(S) = \chi(T)$ and S and T are both orientable/non-orientable.*

A proof can be found in [5, Chapter 1], for example.

A Relations and equivalence relations

Given $x, y \in \mathbb{R}$, we have various ways to understand x *in relation to* y. For example $x = y$, $x \leq y$, $x > y$ etc. Relations abstractify this process to allow us to compare elements in arbitrary sets.

Definition A.0.1. Let A be a set. A relation R on A is a subset $R \subset A \times A$.

We write xRy to denote the element $(x, y) \in R \subset A \times A$. Note that the order is important. As xRy does not imply that yRx.

Example A.0.2.
1. The equality $=$ defines a relation $R_= \{(x, y) \in A \times A \mid x = y\}$. So $R_= = \{(x, x) \mid x \in A\}$.
2. For $A = \mathbb{R}$, $>$ gives the relation $R_> = \{(x, y) \in \mathbb{R}^2 \mid x > y\}$. In particular $x > y$ guarantees that $y < x$ is false.

Example A.0.3. Let A denote the set of all people on Earth. We define three relations on A,

$$\begin{aligned} D &= \big\{(x, y) \in A \times A \mid x \text{ is a descendent of } y\big\} \subset A \times A, \\ B &= \big\{(x, y) \in A \times A \mid x \text{ has an ancestor who is also an ancestor of } y\big\}, \\ S &= \big\{(x, y) \in A \times A \mid x \text{ and } y \text{ have the same parents}\big\}. \end{aligned}$$

Given people x and y, we have that xDy if x is a descendent of y, xBy if x and y are blood-related and xSy if x and y are siblings. Note that xDy implies that xBy and xSy implies that xBy. On the other hand, xSy implies that xDy is false.

Example A.0.4. Let G be the set of groceries that I buy from the supermarket. Different groceries are stored in different rooms (food is stored in the kitchen, detergent is stored in the laundry, shampoo is stored in the bathroom, ...). So we define a relation

$$R = \big\{(x, y) \in G \times G \mid x \text{ and } y \text{ are stored in the same room}\big\}.$$

We see that xRy implies that yRx are they are stored in the same room.

Example A.0.5. Let A be a set and $\mathcal{P}(A) = \{B \subset A\}$ the set of all subsets. We define a relation via inclusion,

$$R = \big\{(B, C) \in \mathcal{P}(A) \mid B \subset C\big\} \subset \mathcal{P}(A) \times \mathcal{P}(A).$$

Definition A.0.6. Let A be a set and $R \subset A \times A$ a relation on A.

1. We say R is reflexive if xRx for all $x \in A$.
2. We say that R is symmetric if xRy implies that yRx.
3. We say that R is antisymmetric if xRy and yRx implies that $x = y$.
4. We that that R is transitive if xRy and yRz implies that xRz.

On the set $\mathbb{R}$, $=$ is reflexive, symmetric and transitive. On the other hand $\leq$ is reflexive, antisymmetric and transitive. The relation $R \subset \mathcal{P}(A) \times \mathcal{P}(A)$ is reflexive, antisymmetric and transitive.

Example A.0.7. Let us return to the relations D, B and S on $A \times A$ from Example A.0.3. Then D is transitive but not symmetric or reflexive, B is reflexive and symmetric but not transitive (for example, a father is blood-related to their children, who are blood-related to their mother, but the father is (ideally) not blood-related to their mother). The sibling relation S is reflexive, symmetric and transitive.

Definition A.0.8. An *equivalence relation* on A is a relation $R \subset A \times A$ that is reflexive, symmetric and transitive. If R is an equivalence relation and $x \in A$,

$$[x] = \{y \in A \mid xRy\}$$

denotes the *equivalence class* at x.

From our examples $S \subset A \times A$ and $R \subset G \times G$ are equivalence relations. Considering the sibling relation and $x \in A$, $[x]$ denotes the set of all siblings of x. If $g \in G$, then $[g]$ will be the set of all groceries that are stored in the same room as g.

One reason we like equivalence relations is the following.

Proposition A.0.9. *Let A be a set and R an equivalence relation on A, then there is a partition of A into equivalence classes of R,*

$$A = \bigcup_{x \in A} [x], \qquad [x] \cap [y] = \begin{cases} [x], & xRy, \\ \emptyset, & \textit{otherwise}. \end{cases}$$

Proof. Because R is transitive, $x \in [x]$ for all $x \in A$ and so $A = \bigcup_x [x]$. Suppose now that $z \in [x] \cap [y]$. By definition zRx and zRy, so xRy by transitivity. So if $x' \in [x]$, then $x'Rx$ and $x'Ry$ (as xRy), which implies that $x' \in [y]$. Hence $[x] \subset [y]$ and the same argument will show $[y] \subset [x]$. Thus if $[x] \cap [y] \neq 0$, it must follow that $[x] = [y]$ as required. $\square$

The equality $A = \bigcup_{x \in A} [x]$ is an example of a partition, a decomposition of a set into pairwise disjoint subsets. The benefit is that it takes a potentially large set A and reduces it into smaller components $[x]$.

Example A.0.10. The grocery equivalence relation $R \subset G \times G$ will decompose G into a union of products that all go in the same room (bathroom products, kitchen products, bedroom products etc.). The sibling relation $S \subset A \times A$ decomposes the total population into groups of siblings.

Definition A.0.11. A reflexive and transitive relation $R \subset A \times A$ is called a preorder of a set. A directed set is a set I with a preorder $R \subset I \times I$ such that for any $x, y \in I$ there is some $c \in I$ such that xRc and yRc.

Examples A.0.12.
1. For any set A, the power set $\mathcal{P}(A)$ is a directed set via the relation by partial inclusion BRC if $B \subset C$ if $B \subset C$. Indeed, $B \subset A$ and $C \subset A$ for all $B, C \in \mathcal{P}(A)$.
2. If $X \subset \mathbb{R}$ is closed, we can define a directed set via the relation xRy if $x \leq y$. Furthermore, x has some maximal element $x_{\max} \in X$ with $x \leq x_{\max}$ for all $x \in X$ (note that we needed X closed to guarantee this).

If X is a set and I is a directed set, then a net in X is a function $f : I \to X$, where we typically write $f(\alpha) = x_\alpha$ for $\alpha \in I$. See Remark 3.3.7 for more about nets.

References

[1] R. Bott and L. W. Tu. *Differential forms in algebraic topology.* Graduate Texts in Mathematics, 82. Springer-Verlag, New York-Berlin, 1982. xiv+331

[2] M. D. Crossley. *Essential topology.* Springer Undergraduate Mathematics Series. Springer-Verlag London, Ltd., London, 2005. x+224 pp.

[3] D. S. Dummit, R. M. Foote. *Abstract Algbera.* Third Edition. Wiley, 2003.

[4] T. W. Gamelin and R. E. Greene. *Introduction to Topology.* Second Edition. Dover Publications, New York, 1999.

[5] W. S. Massey. *A basic course in algebraic topology.* Graduate Texts in Mathematics, 127. Springer-Verlag, New York, 1991. xvi+428 pp.

[6] J. R. Munkres. *Topology.* Second Edition. Pearson, New York, 2018.