18.901 — Lecture Notes

Sanjana Das

Spring 2023

Lecture notes from the MIT class **18.901** (Introduction to Topology), taught by Professor Anthony Conway. All errors are my responsibility.

Contents

§1 First Definitions

§1.1 Definition of a Topology

Definition 1.1. A *topology* on a set X is a family $\mathcal{T}$ of subsets of X, called *open sets*, such that:

(1) Both $\emptyset$ and X are elements of $\mathcal{T}$.

(2) For any collection $\{U_i\}_{i \in I}$ of elements of $\mathcal{T}$, their union $\bigcup_{i \in I} U_i$ is also an element of $\mathcal{T}$.

(3) For any *two* elements U_1 and U_2 of $\mathcal{T}$, their intersection $U_1 \cap U_2$ is also an element of $\mathcal{T}$.

A *topological space* is a set X together with a topology $\mathcal{T}$ on X.

(The third condition is equivalent to requiring that any intersection of a *finite* number of open sets is open. However, this doesn't have to be true for *infinite* intersections.)

First we'll look at a few examples of topologies.

Example 1.2
If $X = \{a, b, c, d\}$, then $\mathcal{T} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, X\}$ is a topology on X.

Definition 1.3. For any set X, the *trivial topology* is the topology $\mathcal{T}_{\text{triv}} = \{\emptyset, X\}$, and the *discrete topology* is the topology consisting of all subsets of X.

The trivial and discrete topologies are both examples of topologies.

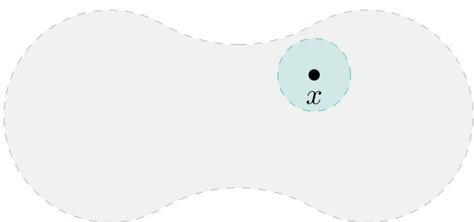
Example 1.4
If $X = \mathbb{R}^n$, let $d(x, y)$ denote the Euclidean distance between x and y, and let

$$B(x, \varepsilon) = \{y \in \mathbb{R}^n \mid d(x, y) < \varepsilon\}$$

be the *ball* of radius ε centered at x. The family of subsets

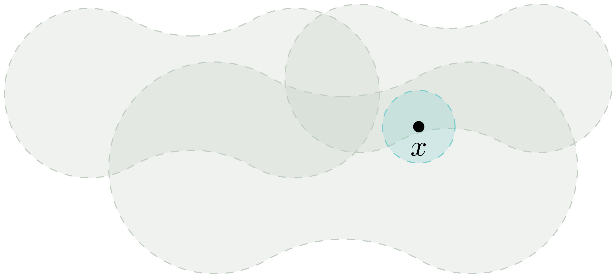
$$\mathcal{T} = \{U \subseteq \mathbb{R}^n \mid \text{for all } x \in U \text{ there exists } \varepsilon > 0 \text{ with } B(x, \varepsilon) \subseteq U\}$$

forms a topology on $\mathbb{R}^n$, called the *standard topology*.

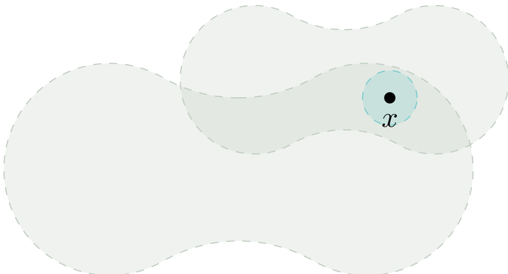


Proof. First, it's clear that $\emptyset$ and $\mathbb{R}^n$ belong to $\mathcal{T}$.

For the second condition, suppose $\{U_i\}$ is a collection of elements of $\mathcal{T}$. To show that their union is as well, for any $x \in \bigcup_i U_i$, there must exist some i with $x \in U_i$. Then there is some ball $B(x, \varepsilon) \subseteq U_i$, and for this ε we have $B(x, \varepsilon) \subseteq \bigcup_i U_i$ as well.



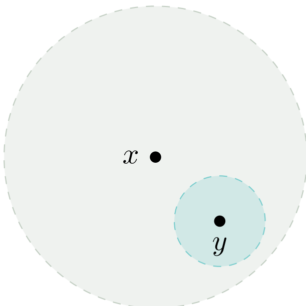
For the third condition, suppose that U_1 and U_2 are in $\mathcal{T}$. To show their intersection is as well, suppose $x \in U_1 \cap U_2$. Then since $x \in U_1$ there must exist ε_1 such that $B(x, \varepsilon_1) \subseteq U_1$, and since $x \in U_2$ there must exist ε_2 such that $B(x, \varepsilon_2) \subseteq U_2$. Then take $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$.



Then $B(x, \varepsilon)$ is contained in both U_1 and U_2, and therefore in their intersection. □

Fact 1.5 — In the standard topology on $\mathbb{R}^n$, the balls $B(x, r)$ (with $r > 0$) are open.

Proof. Given any point $y \in B(x, r)$, we can find ε with $0 < \varepsilon < r - d(x, y)$. Then we claim $B(y, \varepsilon) \subseteq B(x, r)$.



To show this, if $z \in B(y, \varepsilon)$ we have

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \varepsilon < r,$$

so $z \in B(x, r)$ as well. This means $B(x, r)$ contains a ball around each of its points y, so it is open. □

In particular, when $n = 1$, we get that the intervals $(x - r, x + r)$ are open. More generally, the interval (a, b) is open for any $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$.

Definition 1.6. For a set X with two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$ then we say $\mathcal{T}_1$ is *coarser* than $\mathcal{T}_2$, and $\mathcal{T}_2$ is *finer* than $\mathcal{T}_1$.

Example 1.7

Of the four topologies on $X = \{1, 2\}$, the trivial topology $\mathcal{T}_{\text{triv}} = \{\emptyset, \{1, 2\}\}$ is coarsest and the discrete topology $\mathcal{T}_{\text{disc}} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ is finest. The topologies $\mathcal{T}_1 = \{\emptyset, \{1\}, \{1, 2\}\}$ and $\mathcal{T}_2 = \{\emptyset, \{2\}, \{1, 2\}\}$ are both finer than $\mathcal{T}_{\text{triv}}$ and coarser than $\mathcal{T}$disc, but cannot be compared with each other.

§1.1.1 Closed Sets

Definition 1.8. A set $C \subseteq X$ is *closed* in X if $X \setminus C$ is open in X.

The axioms of a topology can be equivalently stated in terms of closed sets, as follows.

(1) Both $\emptyset$ and X are closed.

(2) The intersection of any (not necessarily finite) collection of closed sets is closed.

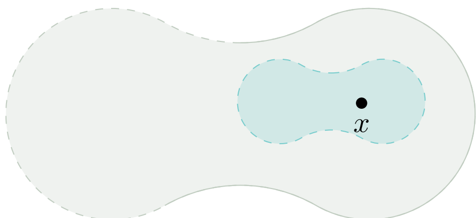
(3) The union of any *finite* collection of closed sets is closed.

Example 1.9

The interval $[a, b]$ is closed in $\mathbb{R}$, as its complement $(-\infty, a) \cup (b, \infty)$ is open (we've seen that $(-\infty, a)$ and (b, ∞) are both open, so their union is as well).

§1.1.2 Neighborhoods

Definition 1.10. A set $N \subseteq X$ is a *neighborhood* of an element $x \in X$ if there exists an open set $U \subset X$ such that $x \in U \subseteq N$.



Example 1.11

The closed interval $[x - \varepsilon, x + \varepsilon] \subseteq \mathbb{R}$ is a neighborhood of x — it contains $(x - \varepsilon, x + \varepsilon)$, which is an open set containing x.

We'll now consider two useful concepts related to neighborhoods — the *interior* and *closure* of a set.

Definition 1.12. For a subset $A \subseteq X$, the *interior* of A is the set

$$\mathring{A} = \{x \in X \mid A \text{ is a neighborhood of } x\}.$$

Proposition 1.13

The interior of A is the union of all open sets U such that $U \subseteq A$.

Proof. First we'll show that $\mathring{A}$ is contained in the union of all open sets $U \subseteq A$. Consider any point $x \in \mathring{A}$. Then since A is a neighborhood of x, there must exist an open set $U \subseteq A$ which contains x, and therefore x is in the union of all open sets contained in A.

Conversely, if x is in any open set U contained in A, then by definition A is a neighborhood of x, and therefore $x \in \mathring{A}$. So the union of all open sets $U \subseteq A$ is also contained in $\mathring{A}$, which means they are equal. ☐

Corollary 1.14

The interior of A is the largest open set contained in A.

In particular, this means $\mathring{A}$ is open and contained in A.

Proof. First, $\mathring{A}$ is a union of open sets, so it is itself open; it is also a union of sets contained in A, so it is itself contained in A. Meanwhile, any open set $U \subseteq A$ is included in this union, and is therefore contained in $\mathring{A}$. So $\mathring{A}$ is the largest open set contained in $\mathring{A}$. ☐

This gives the following characterization of open sets, which is often very useful.

Corollary 1.15

A set A is open if and only if it is a neighborhood of each of its points.

Proof. We've seen that $\mathring{A}$ is the largest open set contained in A, so A is open if and only if $\mathring{A} = A$, which by definition means A is a neighborhood of each of its points. ☐

Definition 1.16. For a subset $A \subseteq X$, the *closure* of A is the set

$$\overline{A} = \{x \in X \mid X \setminus A \text{ is not a neighborhood of } x\}.$$

Similarly to the interior, there is an alternate characterization of the closure in terms of *closed* sets.

Proposition 1.17

The closure of A is the intersection of all closed sets C such that $C \supseteq A$.

Proof. First we'll show that $\overline{A}$ must be contained in this intersection. Suppose that x is *not* in the intersection of closed sets $C \supseteq A$, so there exists a closed set $C \supseteq A$ not containing x. Then $X \setminus C$ is an open set contained in $X \setminus A$ which contains x, so $X \setminus A$ is a neighborhood of x; this means x is not in $\overline{A}$ either.

Conversely, if x is not in $\overline{A}$, then $X \setminus A$ is a neighborhood of x, so there must exist an open set $U \subseteq X \setminus A$ containing x. Then $X \setminus U$ is a closed set which contains A (and is therefore present in the intersection of all closed $C \supseteq A$) and does not contain x; so x cannot be in this intersection. ☐

Corollary 1.18

The closure of A is the smallest closed set in X containing A.

In particular, this means $\overline{A}$ is closed and contains A. So we have $\mathring{A} \subseteq A \subseteq \overline{A}$, giving rise to the following definition.

Definition 1.19. The *boundary* of A is the set $\partial A = \overline{A} \setminus \mathring{A}$.

Example 1.20

If $X = \mathbb{R}$ and $A = [0, 1)$, we have $\mathring{A} = (0, 1)$ and $\overline{A} = [0, 1]$, which means $\partial A = \{0, 1\}$.

§1.2 Subspaces

We'll now see how given a topological space, we can define a topology on a subset.

Definition 1.21. Let $(X, \mathcal{T}_X)$ be a topological space. The *induced topology* on a subset $A \subseteq X$ is

$$\mathcal{T}_A = \{U \subseteq A \mid \text{there exists an open set } V \subseteq X \text{ with } U = A \cap V\}.$$

We say that A (under this topology) is a *subspace* of X.

Note that a set $U \subseteq A$ may be open in the subspace topology on A but not in the original topology on X — so it only makes sense to declare a set open in a certain topological space (i.e., if the ambient topological space is not clear, we should state that a set U is open in X rather than just that U is open).

Example 1.22

Let X be $\mathbb{R}$ with the standard topology, and let $A = [0, \infty)$. Then the set $U = [0, 1)$ is open in A, as we can write $[0, 1) = (-1, 1) \cap A$ and $(-1, 1)$ is open in $\mathbb{R}$. However, U is *not* open in $\mathbb{R}$, as U is not a neighborhood of 0.

§2 Maps Between Topological Spaces

§2.1 Continuous Maps

Definition 2.1. Let X and Y be topological spaces. A map $f\colon X \to Y$ is *continuous* if for every open set $U \subseteq Y$, the set $f^{-1}(U) \subseteq X$ is open.

Example 2.2

The identity map $\mathrm{id}_X\colon (X, \mathcal{T}) \to (X, \mathcal{T})$ is open — if $U \subseteq X$ is open, then $\mathrm{id}_X^{-1}(U) = U$ is also open. More generally, $\mathrm{id}_X\colon (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is continuous if and only if $\mathcal{T}_1$ is finer than $\mathcal{T}_2$.

Fact 2.3 — If $f\colon X \to Y$ and $g\colon Y \to Z$ are continuous, then $g \circ f\colon X \to Z$ is also continuous.

Proof. If U is open in Z, then $g^{-1}(U)$ is open in Y, so $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$ is open in X. □

There are a few equivalent ways of defining continuity. For example, it would be equivalent to require that for every *closed* $C \subseteq Y$, the set $f^{-1}(C) \subseteq X$ is closed; this equivalence follows from the fact that

$$f^{-1}(Y \setminus U) = f^{-1}(Y) \setminus f^{-1}(U) = X \setminus U.$$

A different and often useful way of showing continuity is by *continuity at a point.*

Definition 2.4. A map $f\colon X \to Y$ is *continuous at* $x \in X$ if for every neighborhood $V \subseteq Y$ of $f(x)$ in Y, the set $f^{-1}(V) \subseteq X$ is a neighborhood of x in X.

Proposition 2.5

A map $f\colon X \to Y$ is continuous if and only if it is continuous at every $x \in X$.

Proof. First we'll show that continuity implies continuity at every point. Suppose f is continuous, and consider a point $x \in X$ and a neighborhood V of $f(x)$. Then there must exist an open set $U \subseteq Y$ such that $f(x) \in U \subseteq V$. But then we have $x \in f^{-1}(U) \subseteq f^{-1}(V)$, and $f^{-1}(U)$ is open because f is continuous; this means $f^{-1}(V)$ is a neighborhood of x, as desired.

Now we'll show that continuity at every point implies continuity. Suppose f is continuous at every $x \in X$. Let $U \subseteq Y$ be any open set, so we wish to prove that $f^{-1}(U)$ is open as well. We'll use the characterization from 1.15 that a set is open if and only if it's a neighborhood of each of its points.

Since U is open, then U is a neighborhood of all its points. In particular, for all $x \in f^{-1}(U)$ we have $f(x) \in U$, so U must be a neighborhood of $f(x)$. But since f is continuous at x, then $f^{-1}(U)$ must be a neighborhood of x. So $f^{-1}(U)$ is a neighborhood of each of its points, which means it is open, and therefore f is continuous. $\square$

Now we'll see a few more examples of continuous maps.

Example 2.6

Any constant map is continuous.

Proof. Let $f\colon X \to Y$ be the constant map $x \mapsto y$ (for some fixed $y \in Y$). Then for *any* set $U \subseteq Y$, its inverse image $f^{-1}(U)$ is X if $y \in U$ and $\emptyset$ otherwise; both of these sets are open in X. $\square$

Example 2.7

If X has the discrete topology, then every map $f\colon X \to Y$ is continuous, as *all* subsets of X are open.

Example 2.8

If Y has the trivial topology, then every map $f\colon X \to Y$ is continuous — the only open sets $U \subseteq Y$ are $\emptyset$ and Y, whose inverse images are $\emptyset$ and X respectively (both of which are necessarily open in X).

In the next section (on metric spaces), we'll prove that if X and Y are metric spaces, then the topological definition of continuity coincides with the ε–δ definition from analysis. (In particular, this is true if X and Y are $\mathbb{R}^n$ and $\mathbb{R}^m$ under the standard topology.)

Continuity behaves well with respect to restriction of our spaces, as seen by the following proposition.

Proposition 2.9

Let X and Y be topological spaces, and $f\colon X \to Y$ a continuous map.

(1) If $A \subseteq X$ is a subspace, the restriction $f|_A\colon A \to Y$ is continuous.

(2) If $B \subseteq Y$ is a subspace with $\operatorname{im}(f) \subseteq B$, then the function $h\colon X \to B$ obtained by restricting the target of f is also continuous.

(3) If $Y \subseteq Z$ is a subspace, then the function $h\colon X \to Z$ obtained by extending the target of f is also continuous.

In particular, taking f to be the identity id_X in (1) gives that the inclusion map $A \hookrightarrow X$ is continuous.

Proof. For (1), for any $U \subseteq Y$ we have $(f|_A)^{-1}(U) = f^{-1}(U) \cap A$. If U is open, then since f is continuous, $f^{-1}(U)$ is open in X, so by the definition of the subspace topology then $f^{-1}(U) \cap A$ is open in A.

For (2), for any open subset $U \subseteq B$ we can write $U = V \cap B$ for an open subset $V \subseteq Y$. Then $h^{-1}(U) = f^{-1}(U) = f^{-1}(V)$ (where the latter equality is because $\operatorname{im}(f) \subseteq B$, so $f^{-1}(V) = f^{-1}(V \cap B)$ for any set V). Since $f\colon X \to Y$ is continuous, then $f^{-1}(V)$ must be open, so $h^{-1}(U)$ is open as well.

Finally, (3) can be proven using the same reasoning as (2) in reverse. $\square$

§2.2 Homeomorphisms

Definition 2.10. A map $f\colon X \to Y$ between two topological spaces is a *homeomorphism* if f is continuous, f is bijective, and $f^{-1}\colon Y \to X$ is continuous.

Definition 2.11. If there exists a homeomorphism $f\colon X \to Y$, then we say X and Y are *homeomorphic*, denoted by $X \cong Y$.

Intuitively, homeomorphism captures the notion of two topological spaces being 'essentially the same.'

It is necessary to require that f^{-1} is continuous as well, as this is not implied by f being bijective and continuous — for example, the map $\mathrm{id}_X\colon (X, \mathcal{T}_{\text{disc}}) \to (X, \mathcal{T}_{\text{triv}})$ is continuous and bijective, but its inverse is not continuous.

We'll now see a few examples of homeomorphisms.

Example 2.12

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined as $x \mapsto ax + b$ (for $a \neq 0$) is a homeomorphism — it is continuous and bijective, and its inverse $y \mapsto \frac{y-b}{a}$ is also continuous (the continuity of both maps can be proved using the fact that continuity in $\mathbb{R}$ coincides with the ε–δ definition of continuity from real analysis).

Example 2.13

The function $f\colon \mathbb{R} \to (-1, 1)$ defined as $x \mapsto \frac{x}{1+|x|}$ is a homeomorphism — f is bijective with inverse $g\colon (-1, 1) \to \mathbb{R}$ given by $y \mapsto \frac{y}{1-|y|}$, and both f and g are continuous by real analysis.

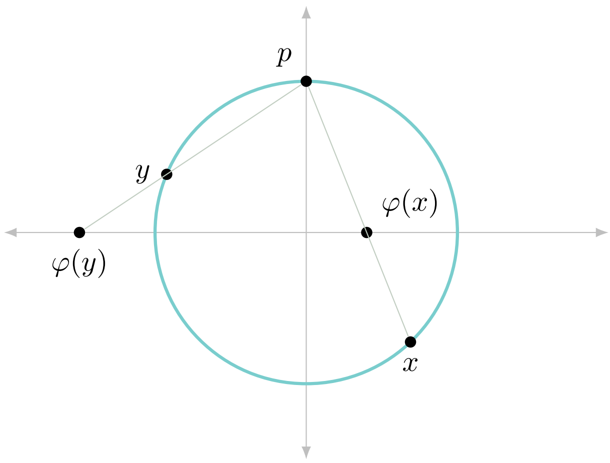
Example 2.14

For each $n \geq 1$, let $\mathbb{S}^n$ denote the n-sphere $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid |x| = 1\}$. Then $\mathbb{S}^n \setminus \{(0, \ldots, 0, 1)\}$ is homeomorphic to $\mathbb{R}^n$.

Proof. One explicit homeomorphism is the *stereographic projection*

$$\varphi\colon (x_1, x_2, \ldots, x_n) \mapsto \left(\frac{x_1}{1-x_n}, \frac{x_2}{1-x_n}, \ldots, \frac{x_{n-1}}{1-x_n}\right).$$

Geometrically, given a point $x \in \mathbb{S}^n$ other than the north pole $p = (0, \ldots, 0, 1)$, this map sends x to the intersection of line px with the plane $x_n = 0$.



It's easy to see geometrically that φ is bijective, as for every point q on the plane $x_n = 0$ there is a unique intersection of line pq with $\mathbb{S}^n$ (other than p). The fact that φ is continuous can be checked using analysis, and it is possible to explicitly write down its inverse and check its continuity using analysis as well. $\square$

Finally, here is a less trivial example of a function that is continuous and bijective, but *not* a homeomorphism.

Example 2.15

Endow $\mathbb{R}$ and $\mathbb{R}^2$ with the standard topology, and consider the subspaces $[0, 1) \subseteq \mathbb{R}$ and

$$\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$$

and the map $f\colon [0, 1) \to \mathbb{S}^1$ defined as $t \mapsto (\cos(2\pi t), \sin(2\pi t))$. Then f is continuous and bijective, but f^{-1} is not continuous.

Proof. The continuity of f follows from first considering the same map $\mathbb{R} \to \mathbb{R}^2$, which is continuous by analysis, and restricting its domain and range. It is easy to see that f is bijective.

However, its inverse $g\colon \mathbb{S}^1 \to [0, 1)$ is not continuous at $(1, 0)$ — for example, the set $U = [0, 1/4)$ is a neighborhood of $g(1, 0) = 0$ in $[0, 1)$, but $g^{-1}(U) = f(U)$ is not a neighborhood of $(1, 0)$ in $\mathbb{S}^1$. $\square$

It is sometimes useful to rephrase the final condition of a homeomorphism (that f^{-1} is continuous) using the following terminology.

Definition 2.16. For topological spaces X and Y, a map $f\colon X \to Y$ is *open* if for every open $U \subseteq X$, its image $f(U) \subseteq Y$ is open as well.

In particular, if f is bijective, then f^{-1} is continuous if and only if f is open.

§3 Metric Spaces

One common way in which topological spaces naturally arise is from *metric spaces*. (This is a generalization of the way we defined a topology on $\mathbb{R}^n$ to general metric spaces.)

Definition 3.1. For a set X, a map $d: X \times X \to [0, \infty)$ is called a *metric* (or *distance*) if it satisfies the following three properties (for all x, y, $z \in X$):

- *Nondegeneracy* — $d(x, y) = 0$ if and only if $x = y$.
- *Symmetry* — $d(x, y) = d(y, x)$.
- The *triangle inequality* — $d(x, z) \leq d(x, y) + d(y, z)$.

One especially useful object in metric spaces is an (open) ball.

Definition 3.2. Let (X, d) be a metric space. For any $x \in X$ and $r > 0$, the *open ball* of radius r centered at x is defined as

$$B_d(x, r) = \{y \in X \mid d(x, y) < r\}.$$

When the metric is clear from context, we write $B(x, r)$ in place of $B_d(x, r)$.

Example 3.3

The set $\mathbb{R}$ can be equipped with the metric $d(x, y) = |x - y|$, in which case the open balls are $B(x, r) = (x - r, x + r)$.

Example 3.4

For any $p \geq 1$, the set $\mathbb{R}^n$ can be equipped with the metric

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}.$$

(The fact that d_p satisfies the triangle inequality is nontrivial, and follows from *Minkowski's inequality*.) The case $p = 2$ corresponds to the standard Euclidean metric; we can also take $p = \infty$ in the above definition, which gives the metric

$$d_\infty(x, y) = \max_{1 \leq i \leq n} \{|x_i - y_i|\}.$$

Example 3.5

Any nonempty set X can be equipped with the *discrete metric*

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise.} \end{cases}$$

A metric gives rise to a topology in the following way.

Definition 3.6. Given a metric space (X, d), the topology induced by d is

$$\mathcal{T}_d = \{U \subseteq X \mid \text{for all } x \in U, \text{ there exists } \varepsilon > 0 \text{ with } B_d(x, \varepsilon) \subseteq U\}.$$

The fact that this is a valid topology can be proven in the same way as in Example 1.4. Note that open balls are open in this topology (the proof is the same as in Fact 1.5). In particular, this means a set $V \subseteq X$ is a neighborhood of a point x if and only if there exists $\varepsilon > 0$ with $B(x, \varepsilon) \subseteq V$ — first, if such ε exists then $B(x, \varepsilon)$ is an open set containing x and contained in V. Conversely, if V is a neighborhood of x then there must exist an open set U with $x \in U \subseteq V$, and since U is open it must contain a ball $B(x, \varepsilon)$.

Example 3.7

The Euclidean metric d_2 induces the standard topology on $\mathbb{R}^n$. (This is true by definition.)

Example 3.8

For any set X, the discrete metric induces the discrete topology — for every set U, for all points $x \in U$ we have $B(x, \varepsilon) = \{x\} \subseteq U$ for all $0 < \varepsilon < 1$, so U is open.

§3.1 Equivalent Metrics

Question 3.9. Do the different metrics d_p on $\mathbb{R}^n$ induce the same topology?

To answer this question, we'll introduce some terminology.

Definition 3.10. Two metrics d and d' on a set X are *equivalent* if there exist positive constants c and c' such that for all x, $y \in X$ we have

$$c \cdot d(x, y) \le d'(x, y) \le c' \cdot d(x, y).$$

Proposition 3.11

Equivalent metrics induce the same topology.

Proof. Note that we have

$$B_{d'}(x, cr) \subseteq B_d(x, r) \subseteq B_{d'}(x, c'r)$$

(the first inclusion follows from the fact that if $d'(x, y) < cr$ then $cd(x, y) \le d'(x, y) < cr$, so $d(x, y) < r$; the second follows similarly).

Now if U is open in $\mathcal{T}_d$, then for all $x \in U$ there exists r with $B_d(x, r) \subseteq U$, which means $B_{d'}(x, cr) \subseteq B_d(x, r) \subseteq U$ as well, and therefore U is open in $\mathcal{T}_{d'}$. The converse follows from the same argument. $\square$

Remark 3.12. The converse is not true — it is possible for non-equivalent metrics to induce the same topology. For example, for any metric space (X, d), we can define

$$\overline{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)}.$$

Then $\overline{d}$ is a metric and induces the same topology as d, but d and d' are not equivalent in general.

This answers our original question about the different metrics on $\mathbb{R}^n$.

Example 3.13

The different metrics d_p on $\mathbb{R}^n$ are all equivalent, as

$$d_\infty(x, y) \le d_p(x, y) \le n^{1/p} d_\infty(x, y)$$

for every p; it then follows that all induce the standard topology.

§3.2 Continuity in Metric Spaces

Finally, we'll show (as mentioned earlier) that for a metric space, the topological definition of continuity is the same as the definition from analysis.

Proposition 3.14

Let (X, d_X) and (Y, d_Y) be metric spaces, and consider any point $x \in X$. Then the following two statements are equivalent:

(1) f is continuous at x.

(2) For every $\varepsilon > 0$, there exists $\delta > 0$ such that for all points $x' \in X$ with $d_X(x, x') < \varepsilon$, we have $d_Y(f(x), f(x')) < \varepsilon$.

(Here (1) refers to the topological definition of continuity, and (2) is the ε–δ definition from real analysis.)

Proof. By definition, f is continuous if and only if for every neighborhood $V \subseteq Y$ of $f(x)$, the set $f^{-1}(V)$ is a neighborhood of x. We've seen that a set is a neighborhood of a point x if and only if it contains a ball of positive radius around x, so this is true if and only if for every $\varepsilon > 0$, the set $f^{-1}(B_{d_Y}(f(x), \varepsilon))$ contains $B_{d_X}(x, \delta)$ for some $\delta > 0$ (the backwards implication follows from the fact that any neighborhood $V \subseteq Y$ of $f(x)$ contains a ball $B_{d_Y}(f(x), \varepsilon)$, while the forwards direction follows from the fact that any such ball is itself a neighborhood of $f(x)$). But this is precisely the statement in (2), so we are done. $\square$

§4 More Topological Spaces

§4.1 Bases and Subbases

We'll soon describe ways to build more topological spaces. But in some of these cases (especially the product topology), it will be clunky to describe *all* the open sets. So before we do so, we'll see an easier way to describe a topology that requires us to specify fewer sets, by writing down a *basis* or *subbasis* instead.(This is similar to in linear algebra, where in order to fully describe a vector space, it's enough to describe a basis.)

Definition 4.1. Let $(X, \mathcal{T})$ be a topological space.

- A subset $\mathcal{B} \subseteq \mathcal{T}$ is a *basis* if every $U \in \mathcal{T}$ can be written as a union of elements of $\mathcal{B}$.
- A subset $\mathcal{S} \subseteq \mathcal{T}$ is a *subbasis* if the collection of *finite* intersections of elements of elements of $\mathcal{S}$ forms a basis.

We say that such a collection $\mathcal{B}$ or $\mathcal{S}$ *generates* the topology $\mathcal{T}$.

Note that every basis is also a subbasis. (In the above definition, we consider $\emptyset$ to be the union of an empty collection of sets, so $\mathcal{B}$ does not have to contain $\emptyset$.)

Example 4.2

Let $X = \{0, 1, 2\}$ and $\mathcal{T} = \{\emptyset, \{0\}, \{0, 1\}, \{0, 2\}, \{0, 1, 2\}\}$. Then the set $\mathcal{B} = \{\{0\}, \{0, 1\}, \{0, 2\}\}$ is a basis, as we can write $\{0, 1, 2\} = \{0, 1\} \cup \{0, 2\}$. Meanwhile $\mathcal{S} = \{\{0, 1\}, \{0, 2\}\}$ is a subbasis, as we can write $\{0\} = \{0, 1\} \cap \{0, 2\}$.

Example 4.3

Let X be any set with the discrete topology. Then $\mathcal{B} = \{\{x\} \mid x \in X\}$ is a basis, since we can write any $U \subseteq X$ as $U = \bigcup_{x \in U} \{x\}$.

Example 4.4

Let X be any set with the trivial topology. Then $\mathcal{B} = \{X\}$ is a basis.

Example 4.5

Let (X, d) be a metric space. Then the set of all balls — i.e., $\mathcal{B} = \{B(x, r) \mid x \in X, r > 0\}$ — is a basis for the induced topology $\mathcal{T}_d$.

This follows immediately from the following characterization of bases.

Proposition 4.6

A collection of sets $\mathcal{B} \subseteq \mathcal{T}$ is a basis for $\mathcal{T}$ if and only if for all $U \in \mathcal{T}$ and $x \in U$, there exists $B \in \mathcal{B}$ with $x \in B \subseteq U$.

Proof. To show that any basis has this property, suppose that $U \in \mathcal{T}$ and $x \in U$. By the definition of a basis, we can write $U = \bigcup_i B_i$ as a union of sets $B_i \in \mathcal{B}$. Then $B_i \subseteq U$ for all i, and since $x \in U$ we must have $x \in B_i$ for some i.

Meanwhile to show that any $\mathcal{B}$ with this property is a basis, given any $U \in \mathcal{T}$, for every $x \in U$ fix some set $B(x) \in \mathcal{B}$ with $x \in B(x) \subseteq U$. Then we can write $U = \bigcup_{x \in U} B(x)$, so $\mathcal{B}$ is indeed a basis. $\square$

Remark 4.7. Unlike in linear algebra, a basis does *not* have to be minimal.

§4.1.1 Conditions for Bases and Subbases

We'll often want to create a topology by specifying a basis or subbasis; so it will be useful to have a simple criterion to check whether a collection is a valid basis or subbasis.

Proposition 4.8

Let X be a set, and $\mathcal{B}$ a collection of subsets of X. Then $\mathcal{B}$ generates a topology on X if and only if the following two conditions hold:

- $X = \bigcup_{B \in \mathcal{B}} B$.
- For all B_1, $B_2 \in B$ and all $x \in B_1 \cap B_2$, there exists $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

Proof. First we'll show that if $\mathcal{B}$ generates a valid topology then both conditions must hold.

- For the first condition, X must be an open set, so it must be possible to write $X = \bigcup_i B_i$ for some sets $B_i \in \mathcal{B}$. Then adding in the remaining sets in $\mathcal{B}$ to the union has no effect, as all such B are subsets of X, so X must also be the union of *all* the sets in B.
- For the second condition, since B_1 and B_2 are in $\mathcal{B}$ they must be open, so $B_1 \cap B_2$ must be open as well. The condition then follows from Proposition 4.6.

Now we'll show that if $\mathcal{B}$ satisfies both conditions, then $\mathcal{B}$ generates a valid topology — i.e., that the set $\mathcal{T} = \{\bigcup_{i \in I} B_i \mid B_i \in \mathcal{B}\}$ is a topology on X. We can check the three axioms:

- We can write $\emptyset$ as an empty union, so $\emptyset \in \mathcal{T}$. Meanwhile $X \in \mathcal{T}$ by the first condition.
- The fact that a union of elements of $\mathcal{T}$ is also in $\mathcal{T}$ is obvious, since the union of two unions of elements of $\mathcal{B}$ is still a union of elements of $\mathcal{B}$.
- Finally, we need to check that if U_1 and U_2 are in $\mathcal{T}$, then so is $U_1 \cap U_2$. We can write $U_1 = \bigcup_{i \in I} B_i$ and $U_2 = \bigcup_{j \in J} B_j$, so that $U_1 \cap U_2 = \bigcup_{i,j} (B_i \cap B_j)$. Then it suffices to show that $B_i \cap B_j$ is in $\mathcal{T}$ for all $B_i, B_j \in \mathcal{B}$.

  But this follows from the second condition — for all $x \in B_i \cap B_j$ there must exist $B(x) \in \mathcal{B}$ with $x \in B(x) \subseteq B_i \cap B_j$. Then $B_i \cap B_j = \bigcup_{x \in B_i \cap B_j} B(x)$ is a union of elements of $\mathcal{B}$, so is in $\mathcal{T}$. ☐

This also gives an even easier criterion to check that a set is a *subbasis*.

Corollary 4.9

Let X be a set, and $\mathcal{S}$ a collection of subsets of X. Then $\mathcal{S}$ is a subbasis for some topology on X if and only if $X = \bigcup_{V \in \mathcal{S}} V$.

Proof. First if $\mathcal{S}$ is a subbasis, then X is a union of finite intersections of elements of $\mathcal{S}$. But any such set is contained in the union of *all* sets $V \in \mathcal{S}$ (and all sets $V \in \mathcal{S}$ are contained in X), so we must have $X = \bigcup_{V \in \mathcal{S}} V$.

For the reverse direction, let $\mathcal{B}$ be the set of finite intersections of elements of $\mathcal{S}$. We can check that $\mathcal{B}$ is a valid basis using Proposition 4.8 — the first condition is immediate (as $\mathcal{S} \subseteq \mathcal{B}$), and the second condition is satisfied because if B_1 and B_2 are both finite intersections of sets in $\mathcal{S}$, then $B_1 \cap B_2$ is as well (so we can simply take $B = B_1 \cap B_2$ in the condition). ☐

§4.1.2 Checking Continuity

It turns out that it is easy to check continuity using a subbasis. (This will be useful when discussing the product topology.)

Lemma 4.10

Let X and Y be topological spaces, and let $\mathcal{S}$ be a subbasis of Y. Then a map $f\colon X \to Y$ is continuous if and only if $f^{-1}(U)$ is open for every $U \in \mathcal{S}$.

Proof. First, if f is continuous, then $f^{-1}(U)$ is open for every open $U \subseteq Y$, so in particular this is true for every $U \in \mathcal{S}$. To prove the converse, since $\mathcal{S}$ is a subbasis, every open $V \subseteq Y$ can be written as a union of finite intersections of sets $U_i \in \mathcal{S}$. Using the facts that $f^{-1}(\bigcup_i U_i) = \bigcup_i f^{-1}(U_i)$ and $f^{-1}(\bigcap_i U_i) = \bigcap_i f^{-1}(U_i)$, it then follows that $f^{-1}(V) \subseteq X$ is a union of finite intersections of sets $f^{-1}(U_i) \subseteq X$ for $U_i \in \mathcal{S}$. But each such set is open by assumption, so $f^{-1}(V)$ is open as well. ☐

A similar statement holds for checking whether a map is open.

Lemma 4.11

Let X and Y be topological spaces, and let $\mathcal{B}$ be a basis of X. Then a map $f\colon X \to Y$ is open if and only if $f(U)$ is open for every $U \in \mathcal{B}$.

Proof. Similarly to in the previous lemma, the 'only if' direction is obvious, and the 'if' direction follows from the fact that every open $V \subseteq X$ can be written as a union of sets $U_i \in \mathcal{B}$, and $f(\bigcup_i U_i) = \bigcup_i f(U_i)$. So if each $f(U_i)$ is open, then so is $f(V)$. $\square$

Remark 4.12. In Lemma 4.11, if f is injective then it is enough to check openness on a *subbasis*, as for f injective we have $f(\bigcap_i U_i) = \bigcap_i f(U_i)$ as well; however, this is not true in general.

§4.2 Box and Product Topologies

We'll now describe how to put a topology on a product of topological spaces. We'll begin by defining the product of *sets*.

Definition 4.13. For sets $X_1, \ldots, X_n$, their *product*, denoted by $\prod_{i=1}^n X_i$ or $X_1 \times \cdots \times X_n$, is the set

$$\prod_{i=1}^n X_i = \{(x_1, \ldots, x_n) \mid x_i \in X_i \text{ for all } i\}.$$

This definition extends to infinite products as well.

Definition 4.14. For any collection $\{X_i\}_{i\in I}$ of sets, their *product*, denoted $\prod_{i\in I} X_i$, is the set

$$\prod_{i\in I} X_i = \{(x_i)_{i\in I} \mid x_i \in X_i \text{ for all } i\}.$$

More precisely, $\prod_{i\in I} X_i$ is the set of *functions* $x\colon I \to \bigcup_{i\in I} X_i$ such that $x_i \in X_i$ for each i (where $x(i)$ corresponds to x_i in the above definition). However, we generally view elements of $\prod_{i\in I} X_i$ as in the above definition.

There are two ways of placing a topology on a product — the *box* topology and the *product* topology. These are the same for *finite* products, but differ for *infinite* products. We'l begin with the box topology — the box topology is easier to describe and perhaps matches what we would naively expect, but it turns out to have somewhat undesirable behavior on infinite products; we'll see later that the product topology fixes these issues.

§4.2.1 The Box Topology

Definition 4.15. Let $(X_i)_{i\in I}$ be a collection of topological spaces. The *box topology* on $\prod_{i\in I} X_i$ is the topology generated by the basis

$$\mathcal{B} = \left\{\prod_{i\in I} U_i \mid U_i \subseteq X_i \text{ open}\right\}.$$

For this to make sense, we need to check that $\mathcal{B}$ is a valid basis for some topology. This follows from Proposition 4.8 — the first condition is satisfied because $X = \prod_{i\in I} X_i$ is itself in $\mathcal{B}$, and the second is satsified because the intersection of any two elements of $\mathcal{B}$ is in fact in $\mathcal{B}$ as well — we have

$$\prod_{i\in I} U_i \cap \prod_{i\in I} V_i = \prod_{i\in I} U_i \cap V_i,$$

and if each U_i and V_i is open then so is each $U_i \cap V_i$.

Example 4.16

If $\mathbb{R}$ is endowed with the standard topology, then the box topology on $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ is the same as the standard topology.

(This is an exercise on a problem set.)

Example 4.17

For a collection of spaces $\{X_i\}_{i\in I}$ where each X_i has the discrete topology, the box topology on $\prod_{i\in I} X_i$ is discrete as well.

Proof. For each $x = (x_i)_{i\in I}$ in the product, we have $\{x\} = \prod_{i\in I}\{x_i\} \in \mathcal{B}$ (since each $\{x_i\} \subseteq X_i$ is open). This means every set $U \subseteq \prod_{i\in I} X_i$ is open, as we can write $U = \bigcup_{x\in U}\{x\}$. $\square$

Unfortunately, the box topology does not behave well with infinite products, as seen in the following example.

Example 4.18

Consider the product $\mathbb{R}^{\mathbb{N}} = \mathbb{R} \times \mathbb{R} \times \cdots$ (where $I = \mathbb{N}$ and each X_i is $\mathbb{R}$ with the standard topology). Then the map $f\colon \mathbb{R} \to \mathbb{R}^{\mathbb{N}}$ sending $t \mapsto (t, t, t, \ldots)$ is not continuous.

Proof. For each $i \in \mathbb{N}$, let $U_i = (-\frac{1}{i}, \frac{1}{i}) \subseteq X_i$. Then $U = \prod_{i\in\mathbb{N}} U_i$ is open, but

$$f^{-1}(U) = \{t \in \mathbb{R} \mid t \in U_i \text{ for all } i \in \mathbb{N}\} = \bigcap_{i\in\mathbb{N}} U_i = \{0\},$$

which is not open. So f is not continuous. $\square$

This is undesirable because f is a fairly nice function — in particular, each of its components is continuous — so we would like a topology on $\mathbb{R}^{\mathbb{N}}$ in which f is continuous. The problem with the box topology is roughly that it contains too many open sets — for example, the above example shows that we don't want the set $\prod_{i\in I}(-\frac{1}{i}, \frac{1}{i})$ to be open. We'll now see a different topology — the *product* topology — that works the same way as the box topology for finite products, but fixes this issue for infinite products.

§4.2.2 The Product Topology

Definition 4.19. Given a product of sets $X = \prod_{i\in I} X_i$, for each $j \in I$ the *jth projection map*, denoted π_j, is the map $\pi_j\colon X \to X_j$ sending $x \mapsto x_j$ for each $x = (x_i)_{i\in I} \in X$.

Definition 4.20. Let $\{X_i\}_{i\in I}$ be a collection of topological spaces. The *product topology* on $\prod_{i\in I} X_i$ is the topology with subbasis

$$\mathcal{S} = \{\pi_j^{-1}(U_j) \mid j \in I,\, U_j \subseteq X_j \text{ open}\}.$$

In other words, the subbasis $\mathcal{S}$ of the product topology consists of sets of the form $U_j \times \prod_{i\neq j} X_i$ for open sets $U_j \subseteq X_j$ (since $\pi_j^{-1}(U_j)$ consists precisely of the points whose jth coordinate is in U_j). So the corresponding basis $\mathcal{B}$ (i.e., the collection of finite intersections of sets in $\mathcal{S}$) consists of sets of the form $\prod_{i\in I} U_i$ where each $U_i \subseteq X_i$ is open, and there are only finitely many i for which $U_i \neq X_i$ (namely, the i which appear as j in one of the sets in the intersection). In particular, the product topology is coarser than the box topology (as its basis is contained in the basis of the box topology), and if I is finite then the two topologies are the same (as they then have the same basis).

Remark 4.21. One way to view the product topology is as the coarsest topology such that each projection π_j is continuous (since π_j is continuous if and only if $\pi^{-1}(U_j)$ is open for all $U_j \subseteq X_j$).

Example 4.22

The product topology on $X = \prod_{i\in I}\{0, 1\}$ (where each set $\{0, 1\}$ has the discrete topology) is the discrete topology if and only if I is finite.

One of the main issues we had with the box topology was that a componentwise continuous map such as $t \mapsto (t, t, \ldots)$ need not be continuous. The product topology fixes this issue — in the product topology it *is* true that any componentwise continuous map is continuous, as seen in the following proposition.

Proposition 4.23

Let $X = \prod_{i\in I} X_i$ be endowed with the product topology. Then a function $f\colon Y \to X$ is continuous if and only if its ith coordinate $\pi_i \circ f\colon Y \to X_i$ is continuous for each $i \in I$.

Proof. First π_i is continuous for each i, so if f is continuous, then so is $\pi_i \circ f$. Conversely, in order to check that such a function f is continuous, by Lemma 4.10 it suffices to check that $f^{-1}(U)$ is open for every $U \in \mathcal{S}$. But every set U in our subbasis can be written as $\pi_j^{-1}(U_j)$ for some $j \in I$ and some open set $U_j \subseteq X_j$. Then $f^{-1}(U) = (\pi_j \circ f)^{-1}(U_j)$, which is open by the continuity of $\pi_j \circ f$. □

§4.3 The Quotient Topology

Now we'll see how to define a topology on the *quotient* of a set.

§4.3.1 Definitions

First we'll define the quotient of a set by an equivalence relation.

Definition 4.24. An *equivalence relation* on a set X is a binary relation $\sim$ on X satisfying the following three properties:

- *Reflexivity* — $x \sim x$ for all $x \in X$.
- *Symmetry* — $x \sim y$ if and only if $y \sim x$ for all $x, y \in X$.
- *Transitivity* — if $x \sim y$ and $y \sim z$, then $x \sim z$ (for any $x, y, z \in X$).

Example 4.25

The relation $x \sim y$ if and only if $x = y$ is an equivalence relation (on any set).

Example 4.26

On the set $\mathbb{R}$, the relation $x \sim y$ if and only if $x - y \in n\mathbb{Z}$ (for fixed n) is an equivalence relation.

Definition 4.27. The *equivalence class* of an element $x \in X$, denoted $[x]$, is defined as

$$[x] = \{y \in X \mid x \sim y\}.$$

The properties of an equivalence relation imply that $x \in [x]$ and that the different equivalence classes form a partition of X.

Definition 4.28. Given a set X and equivalence relation $\sim$, the *quotient* set, denoted $X/\sim$, is defined as the set of equivalence classes in X.

Definition 4.29. Given a quotient $X/\sim$, the *projection map* $\pi\colon X \to X/\sim$ is the map $x \mapsto [x]$.

We are now ready to place a topology on the quotient of a set.

Definition 4.30. For a topological space X with equivalence relation $\sim$ and projection map $\pi\colon X \to X/\sim$, the *quotient topology* on $X/\sim$ is defined as

$$\{U \subseteq X/\sim \mid \pi^{-1}(U) \subseteq X \text{ open}\}.$$

It can be checked that this defines a valid topology. Note that the projection map π is continuous in the quotient topology — in fact, the quotient topology is the finest topology on $X/\sim$ for which this is true.

§4.3.2 Maps on a Quotient Space

Definition 4.31. Given a quotient space $X/\sim$ with projection map $\pi\colon X \to X/\sim$, we say a map $f\colon X \to Y$ *descends* to a map $\widetilde{f}\colon X/\sim \to Y$ if $f = \widetilde{f} \circ \pi$.

(We may also say that the map $\widetilde{f}$ is *induced* by f.) A map $f\colon X \to Y$ descends to the quotient $X/\sim$ if and only if whenever $x \sim y$ we have $f(x) = f(y)$; in that case $\widetilde{f}$ is uniquely defined as $\widetilde{f}([x]) = f(x)$ for all $x \in X$ (the above condition implies that this is well-defined).

Note that if f descends to $\widetilde{f}$, then $\operatorname{im}(\widetilde{f}) = \operatorname{im}(f)$ — in particular, $\widetilde{f}$ is surjective if and only if f is. Meanwhile, $\widetilde{f}$ is injective if and only if whenever $f(x) = f(y)$ we have $x \sim y$.

Lemma 4.32

If a map $f\colon X \to Y$ descends to $\widetilde{f}\colon X/\sim \to Y$, then $\widetilde{f}$ is continuous if and only if f is continuous.

Proof. For each open set $U \subseteq Y$, by the definition of the quotient topology $\widetilde{f}^{-1}(U) \subseteq X/\sim$ is open if and only if $\pi^{-1}(\widetilde{f}^{-1}(U)) = (\widetilde{f} \circ \pi)^{-1}(U) = f^{-1}(U)$ is open. $\square$

Lemma 4.33

If a map $f\colon X \to Y$ descends to $\widetilde{f}\colon X/{\sim} \to Y$ and f is open, then $\widetilde{f}$ is open.

(Unlike in the previous lemma, the converse is not true.)

Proof. Given an open set $U \subseteq X/{\sim}$, we have $\widetilde{f}(U) = f(\pi^{-1}(U))$. (Explicitly, this is because we have $\widetilde{f}(U) = \{\widetilde{f}([x]) \mid [x] \in U\} = \{f(x) \mid x \in \pi^{-1}(U)\}$.) But $\pi^{-1}(U)$ is open by the definition of the quotient topology. Since f is open, this means $f(\pi^{-1}(U))$ is open, so $\widetilde{f}(U)$ is open. □

§4.3.3 Examples of Quotient Spaces

Now we'll look at a few examples of quotient spaces.

Example 4.34

Consider the space $\mathbb{R}^2$ with the equivalence relation $(x_1, x_2) \sim (y_1, y_2)$ if and only if $x_1^2+x_2^2 = y_1^2+y_2^2$ (in other words, $x \sim y$ if and only if $|x| = |y|$). Then $\mathbb{R}^2/{\sim}$ is homeomorphic to $[0, \infty)$ (with the subspace topology of the standard topology on $\mathbb{R}$).

Intuitively, this makes sense because quotienting $\mathbb{R}^2$ by $\sim$ means that we only care about the radius of each point, and we can think of $[0, \infty)$ as recording this radius.

Proof. Consider the map $f\colon \mathbb{R}^2/{\sim} \to [0, \infty)$ mapping $x \mapsto |x|$. By definition $f(x) = f(y)$ if and only if $x \sim y$, so f descends to an injective map $\widetilde{f}\colon \mathbb{R}^2/{\sim} \to [0, \infty)$. Additionally, f is clearly surjective, as for each $r \in [0, \infty)$ we have $f((0, r)) = r$; this means $\widetilde{f}$ is surjective as well, and is therefore bijective.

We claim that $\widetilde{f}$ is in fact a homeomorphism. First, in order to check that $\widetilde{f}$ is continuous, it suffices (by Lemma 4.32) to check that f is continuous; this can be easily checked using real analysis.

Similarly, in order to check that $\widetilde{f}^{-1}$ is continuous — i.e., that $\widetilde{f}$ is open — by Lemma 4.33 it suffices to check that f is open; and by Lemma 4.11 it suffices to check that $f(B(x, r))$ is open for every $x \in \mathbb{R}^2$ and $r > 0$ (since the open balls $B(x, r)$ form a basis for the topology on $\mathbb{R}^2$). But we have

$$f(B(x, r)) = \begin{cases} (|x| - r, |x| + r) & \text{if } |x| \leq r \\ [0, |x| + r) & \text{otherwise,} \end{cases}$$

which is open in either case. So f and therefore $\widetilde{f}$ is open, which means $\widetilde{f}^{-1}$ is continuous.

So $\widetilde{f}$ is indeed a homeomorphism between $\mathbb{R}^2/{\sim}$ and $[0, \infty)$, as desired. □

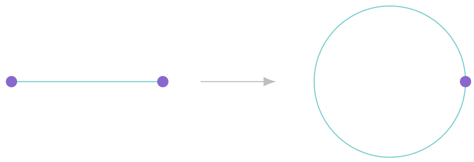
One common use of quotient spaces is to squash part of our space into a single point — this can be formalized in the following way.

Definition 4.35. Let X be a space and $A \subseteq X$ a subspace. Then X/A is defined as the quotient of X by the equivalence relation $x \sim y$ if and only if $x = y$ or $x, y \in A$.

Example 4.36

The quotient $[0, 1]/\{0, 1\}$ is homeomorphic to $\mathbb{S}^1$.

Intuitively, this makes sense because if we take the interval $[0, 1]$ and identify its two endpoints by gluing them together, we obtain a circle.



Proof. View $\mathbb{S}^1$ as $\{z \in \mathbb{C} \mid |z| = 1\}$ (with the induced topology), and consider the map $f\colon [0,1] \to \mathbb{S}^1$ sending $t \mapsto e^{2\pi i t}$. Then $e^{2\pi i \cdot 0} = e^{2\pi i \cdot 1} = 1$, so f descends to a map $\widetilde{f}\colon [0,1]/\{0,1\} \to \mathbb{S}^1$. We will show that $\widetilde{f}$ is a homeomorphism; as before, we need to check that it is continuous, bijective, and open.

First, the fact that $\widetilde{f}$ is continuous follows from the fact that f is continuous, which can again be proven using analysis (the map $t \mapsto e^{2\pi i t}$ is continuous as a map $\mathbb{R} \to \mathbb{R}^2$, so it remains continuous when we restrict its domain and range).

Next, f is clearly surjective, which means $\widetilde{f}$ is surjective. Meanwhile if $e^{2\pi i s} = e^{2\pi i t}$ then $s - t \in \mathbb{Z}$, which implies $s \sim t$ (as either $s = t$ or $\{s,t\} = \{0,1\}$); this means $\widetilde{f}$ is injective. So $\widetilde{f}$ is bijective.

Finally, it remains to prove that $\widetilde{f}^{-1}$ is continuous (i.e., that $\widetilde{f}$ is open); we need to check that for every open set $U \subseteq [0,1]/\{0,1\}$, the set $\widetilde{f}(U) = f(\pi^{-1}(U)) \subseteq \mathbb{S}^1$ is also open. We'll do this by checking that it is the neighborhood of each of its points — i.e., given any open set $U \subseteq [0,1]/\{0,1\}$, for every $t \in \pi^{-1}(U)$ we have that $f(\pi^{-1}(U))$ is a neighborhood of $f(t)$.

Case 1 ($t \notin \{0,1\}$). Since $U \subseteq [0,1]/\{0,1\}$ is open, then $\pi^{-1}(U)$ is open, so there exists $\varepsilon > 0$ such that $(t-\varepsilon, t+\varepsilon) \subseteq \pi^{-1}(U) \cap (0,1)$. It then suffices to show that $f((t-\varepsilon, t+\varepsilon))$ is open. But if we define $\log\colon \mathbb{S}^1 \to [0,1)$ as the inverse of $t \mapsto e^{2\pi i t}$, then log is continuous on $\mathbb{S}^1/\{1\}$; then we can write

$$f((t-\varepsilon, t+\varepsilon)) = \frac{1}{2\pi i}\log^{-1}((t-\varepsilon, t+\varepsilon)),$$

which means it must be open.

Case 2 ($t \in \{0,1\}$). This means both 0 and 1 are in $\pi^{-1}(U)$, so we can find $\varepsilon > 0$ such that $[0,\varepsilon) \cup (1-\varepsilon, 1] \subseteq \pi^{-1}(U)$. We can again check that $f([0,\varepsilon) \cup (1-\varepsilon, 1]) \subseteq \mathbb{S}^1$ is open, so $f(\pi^{-1}(U))$ is again a neighborhood of $f(t) = 1$.

This shows that $\widetilde{f}$ is open, and therefore $\widetilde{f}^{-1}$ is continuous, as desired. $\square$

Remark 4.37. In both the examples we've seen here, proving that $\widetilde{f}$ is continuous and bijective was fairly easy, but proving its inverse is continuous was surprisingly painful. Later, we'll see that if our two spaces are 'nice' enough (we'll define the specific criterion later), then *any* map which is continuous and bijective also has continuous inverse; this will make our lives a lot easier when trying to prove that certain maps are homeomorphisms.

Example 4.36 generalizes to higher dimensions in the following way.

Definition 4.38. For every n, the n-sphere $\mathbb{S}^n$ is defined as

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid |x| = 1\},$$

and the n-disk $\mathbb{D}^n$ is defined as

$$\mathbb{D}^n = \{x \in \mathbb{R}^n \mid |x| \le 1\}.$$

Note that the n-dimensional sphere $\mathbb{S}^n$ lives in the $(n+1)$-dimensional space $\mathbb{R}^{n+1}$, and the *boundary* of $\mathbb{D}^n$ is $\partial\mathbb{D}^n = \mathbb{S}^{n-1}$.

Example 4.39

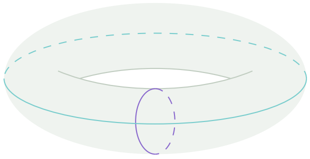
For any n, the quotient $\mathbb{D}^n/\partial\mathbb{D}^n$ is homeomorphic to $\mathbb{S}^n$.

The case $n = 1$ is Example 4.36 (as $\mathbb{D}^n = [-1, 1]$ is a closed interval). We won't prove this in general, but it makes intuitive sense — for example, when $n = 2$, if we take a circular disk and glue together all the points on its boundary, the resulting object is essentially a (hollow) sphere.

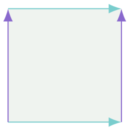
Example 4.40

Consider the space $[0, 1]^2$ with the equivalence relation $(x, y) \sim (x', y')$ if and only if $x, x' \in \{0, 1\}$ and $y = y'$, or $y, y' \in \{0, 1\}$ and $x = x'$. Then $[0, 1]^2/\sim \,\cong \mathbb{S}^1 \times \mathbb{S}^1$.

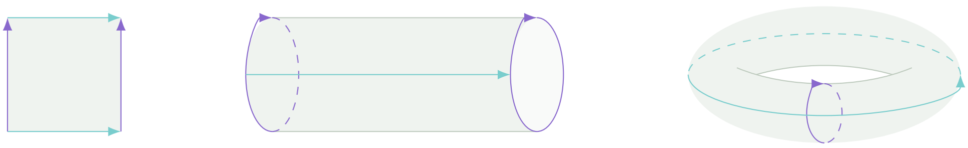
The space $\mathbb{S}^1 \times \mathbb{S}^1$ is called the *torus*. It can be drawn in the following way, where the two circles shown represent the two factors of $\mathbb{S}^1$.



Meanwhile, to visualize $[0, 1]^2/\sim$, we can draw it in the following way (where the two blue arrows represent that we identify the two horizontal sides of the square in the same orientation, and the two purple arrows represent that we identify the vertical sides).



To intuitively see why these two spaces are homeomorphic, imagine taking the above figure and gluing together the blue edges to produce a cylinder, and then gluing together the purple edges to produce a torus.



This homeomorphism can be rigorously proven in the following way.

Proof. We'll instead show that $[0, 1]^2/\sim$ is homeomorphic to $[0, 1]/\{0, 1\} \times [0, 1]/\{0, 1\}$ — the result then follows from the fact that $[0, 1]/\{0, 1\} \cong \mathbb{S}^1$, which we've shown earlier.

To see this, consider the map $f\colon [0, 1] \times [0, 1] \to [0, 1]/\{0, 1\} \times [0, 1]/\{0, 1\}$ sending $(x, y) \mapsto ([x], [y])$. (Here $[x]$ denotes the equivalence class of x in the quotient of $[0, 1]$ by $\{0, 1\}$.) We can check that f descends to a map $\widetilde{f}\colon [0, 1]^2/\sim \,\to [0, 1]/\{0, 1\} \times [0, 1]/\{0, 1\}$, and that $\widetilde{f}$ is bijective. Then f is continuous (as each of its components is continuous), so $\widetilde{f}$ is continuous as well; and we can also check that $\widetilde{f}^{-1}$ is continuous (this can be checked by hand, but it'll also follow from the criterion we'll see in a few weeks). So $\widetilde{f}$ is a homeomorphism, as desired. $\square$

§5 Sequences and Convergence

We'll now see how the notion of convergence in analysis can be generalized to any topological space — it turns out that we can define convergence even without a metric.

§5.1 Neighborhood Bases

Before we begin discussing sequences and convergence, we'll define a notion that will be useful in this discussion.

Definition 5.1. Let X be a topologicla space. A *neighborhood basis* $\mathcal{B}_x$ for a point $x \in X$ is a collection of neighborhoods of x such that for every $V \subseteq X$ which is a neighborhood of x, there exists a neighborhood $B \in \mathcal{B}_x$ such that $B \subseteq V$.

Example 5.2

If (X, d) is a metric space, for each $x \in X$ the set $\mathcal{B}_x = \{B(x, r) \mid r > 0\}$ is a neighborhood basis of x.

Proof. Given any neighborhood V of x, by the definition of a neighborhood there must exist an open set U with $x \in U \subseteq V$, and by the definition of open sets in the metric topology there must exist $r > 0$ such that $B(x, r) \subseteq U$, so then $B(x, r) \subseteq V$ as well. □

Note that we could have taken a much smaller neighborhood basis in the above example.

Example 5.3

If (X, d) is a metric space, for each $x \in X$ the set $\mathcal{B}_x = \{B(x, \frac{1}{n}) \mid n \in \mathbb{N}\}$ is a neighborhood basis of x.

Proof. This follows from Example 5.2 and the fact that for any $r > 0$, there exists $n \in \mathbb{N}$ with $\frac{1}{n} \leq r$, and therefore $B(x, \frac{1}{n}) \subseteq B(x, r)$. □

Example 5.4

If X has the discrete topology, for every $x \in X$ the set $\mathcal{B}_x = \{\{x\}\}$ is a neighborhood basis of x.

Proof. First, $\{x\}$ is open in the discrete topology, so it is indeed a neighborhood of x. Meanwhile for any neighborhood V of x, since $x \in V$ we must have $\{x\} \subseteq V$, as desired. □

Example 5.5

If X has the discrete topology, for every $x \in X$ the set $\mathcal{B}_x = \{X\}$ is a neighborhood basis of x.

§5.2 Definition of Convergence

We'll soon define what it means for a sequence to converge; first we'll formally define a sequence.

Definition 5.6. A *sequence* in a set X is a map $x\colon \mathbb{N} \to X$.

We generally use x_n to denote the nth element of a sequence, and $(x_n)_{n\in\mathbb{N}}$ (or (x_n) as shorthand) to denote the entire sequence $x_1, x_2, x_3, \ldots$.

Definition 5.7. Given a sequence (x_n) in a topological space X, we say that (x_n) *converges* to a point $x \in X$ — or that x is a *limit* of the sequence (x_n) — if for every open set $U \subseteq X$ containing x, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in U$.

In other words, (x_n) converges to x if for each open set U around x, all terms sufficiently far out in the sequence lie in U (where the meaning of 'sufficiently far out' may depend on U).

Remark 5.8. It is sometimes useful to rephrase the above definition in terms of neighborhoods — a sequence (x_n) converges to x if and only if for every *neighborhood* $V \subseteq X$ of x, there exists N such that for all $n \geq N$ we have $x_n \in V$. (The backwards direction is clear as every open set U containing x is a neighborhood of x, while the forwards direction follows from the fact that every neighborhood V of x contains an open set U containing x, so there exists N such that for $n \geq N$ we have $x_n \in U \subseteq V$.)

Notation 5.9. We use $x_n \to x$ to denote that the sequence (x_n) converges to x.

It turns out that we often don't have to check the above condition on *every* open set U — it's enough to check it on a neighborhood basis or a subbasis, as the following two lemmas state. (This will often be simpler to check.)

Lemma 5.10

If $\mathcal{B}_x$ is a neighborhood basis of $x \in X$, then a sequence (x_n) converges to $x \in X$ if and only if for all neighborhoods $B \in \mathcal{B}_x$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in B$.

Proof. The forwards direction follows from Remark 5.8, as if (x_n) converges to x then *every* neighborhood of x must have this property. The backwards direction similarly follows from the fact that every open set U with $x \in U$ is a neighborhood of x, and therefore contains some $B \in \mathcal{B}_x$. $\square$

Lemma 5.11

If $\mathcal{S}$ is a subbasis of the topology on X, then a sequence (x_n) converges to a point x if and only if for every $U \in \mathcal{S}$ containing x, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in U$.

Proof. The forwards direction is clear, as every $U \in \mathcal{S}$ is open. To prove the backwards direction, let $\mathcal{B}$ be the corresponding basis (consisting of finite intersections of sets in $\mathcal{S}$).

First we'll show that every $B \in \mathcal{B}$ containing x has the same property — i.e., that there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in B$. To do so, write $B = U_1 \cap \cdots \cap U_r$ for sets $U_i \in \mathcal{S}$. Then since $x \in B$, for each $i \in \{1, \ldots, r\}$ we must have $x \in U_i$, and therefore there exists $N_i \in \mathbb{N}$ such that for all $n \geq N_i$ we have $x_n \in U_i$. Then taking $N = \max(N_1, \ldots, N_r)$ gives the desired result.

Now we'll show that having this property for all $B \in \mathcal{B}$ containing x implies that the same property holds for *all* open sets U containing x. Write $U = \bigcup_i B_i$ for some collection of sets $B_i \in \mathcal{B}$. Then since $x \in U$, there must exist some i with $x \in B_i$; then there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in B_i$, and since $B_i \subseteq U$ this implies $x_n \in U$ as well. $\square$

§5.3 Some Examples of Convergence

Now we'll see a few examples of convergence. First, we'll see that for metric spaces, the topological definition of convergence is equivalent to the one from analysis.

Example 5.12

If (X, d) is a metric space, a sequence (x_n) converges to x if and only if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$.

Proof. By Lemma 5.10, in order to check convergence at x, it suffices to consider all elements of a neighborhood basis of x (rather than all open sets containing x); and by Example 5.2 the balls $B(x, \varepsilon)$ form a neighborhood basis of x. So (x_n) converges to x if and only if for every $\varepsilon > 0$ there exists N such that $x_n \in B(x, \varepsilon)$, or in other words, such that $d(x_n, x) < \varepsilon$. □

Example 5.13

If X is discrete, then any convergent sequence must be eventually constant.

(Sequences which are eventually constant are also called *stationary*.)

Proof. Suppose that (x_n) converges to x. Then since $\{x\}$ is open, there must exist N such that for $n \geq N$ we have $x_n \in \{x\}$, and therefore $x_n = x$. □

Example 5.14

If X has the *cocountable* topology

$$\mathcal{T}_c = \{U \subseteq X \mid X \setminus U \text{ is countable or } U = \emptyset\},$$

then again any convergent sequence must be eventually constant.

Proof. Assume that a sequence (x_n) converges to x, and consider the set

$$U = (X \setminus \{x_n \mid n \in \mathbb{N}\}) \cup \{x\},$$

which is open as $\{x_n \mid n \in \mathbb{N}\}$ is countable. Then there must exist N such that for all $n \geq N$ we have $x_n \in U$, and since we cannot have $x_n \in X \setminus \{x_n \mid n \in \mathbb{N}\}$, this implies $x_n = x$. □

Example 5.15

If X has the trivial topology, then every sequence converges to every point.

This is because then the only open set is X itself, in which case the condition for convergence is trivially satisfied.

In particular, a sequence may have *multiple* limits, as seen in the last example. This behavior is perhaps somewhat unsettling, as in analysis a sequence can only have one limit; we'll now see a condition on our space that does guarantee the uniqueness of limits.

§5.4 Hausdorff Spaces

Definition 5.16. A space X is *Hausdorff* if for all distinct points $x, y \in X$, there exist disjoint open sets $U, V \subseteq X$ with $x \in U$ and $y \in V$.

In other words, in a Hausdorff space we can separate any two different points using open sets.

Remark 5.17. Being Hausdorff is a *topological property*, meaning that it is preserved by homeomorphism (i.e., if X and Y are homeomorphic, then X is Hausdorff if and only if Y is) — this is true of basically any property defined purely in terms of open sets.

Proposition 5.18

Let X be a Hausdorff space, and (x_n) a sequence in X. If (x_n) converges to two points x and y, then we must have $x = y$.

Proof. Assume $x \neq y$. Since X is Hausdorff we can find disjoint open sets $U, V \subseteq X$ with $x \in U$ and $y \in V$. Then by the definition of convergence, there must exist N_1 and N_2 such that for all $n \geq N_1$ we have $x_n \in U$, and for all $n \geq N_2$ we have $x_n \in V$. But this is impossible, as then for all $n \geq \max(N_1, N_2)$ we have $x_n \in U \cap V = \emptyset$. $\square$

§5.4.1 Some Examples

Now we'll consider a few examples of Hausdorff spaces.

Example 5.19

Every discrete space is Hausdorff — for any two points $x \neq y$, we can take the disjoint open sets $U = \{x\}$ and $V = \{y\}$.

Example 5.20

Any space with the trivial topology (and with at least two elements) is not Hausdorff.

Example 5.21

Every metric space is Hausdorff.

Proof. Given two points $x \neq y$ in our metric space, let $r = \frac{1}{2}d(x, y) > 0$ and take the open sets $U = B(x, r)$ and $V = B(y, r)$. These are disjoint by the triangle inequality — if a point z were in their intersection, we would have

$$d(x, z) + d(z, y) < r + r = d(x, y),$$

contradicting the triangle inequality. $\square$

Lemma 5.22

If X is Hausdorff and $Y \subseteq X$ a subspace, then Y is Hausdorff as well.

Proof. Given two points $x \neq y$ in X, we can choose disjoint open sets U and V in X which contain x and y respectively. Then $U \cap Y$ and $V \cap Y$ are disjoint open sets in Y which contain x and y respectively. □

Lemma 5.23

If $\{X_i\}_{i \in I}$ is a family of Hausdorff spaces, then their product $\prod_{i \in I} X_i$ is Hausdorff as well (with the product topology).

Proof. Given two points $x \neq y$ in the product space, fix a coordinate i at which they differ, and let U and V be disjoint open subsets of X_i containing x_i and y_i respectively. Then $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are disjoint open subsets of $\prod_{i \in I} X_i$ containing x and y respectively. □

Remark 5.24. This also implies that a product of Hausdorff spaces with the *box* topology is Hausdorff, since the box topology is finer than the product topology.

However, it is *not* true in general that a quotient of a Hausdorff space is Hausdorff.

§5.4.2 A Partial Converse

Question 5.25. Is the converse of Proposition 5.18 true — i.e., if all limits in X are unique, must X be Hausdorff?

The answer in general is no. For example, any uncountably infinite set with the cocountable topology is not Hausdorff, but has the property that limits are unique — we saw in Example 5.14 that in the cocountable topology, a sequence (x_n) converges to x if and only if the sequence x_n becomes eventually constant at x.

However, it turns out that there is a fairly reasonable condition we can impose on X under which the converse *is* true.

Definition 5.26. A space X is *first countable* if for every $x \in X$, there exists a countable neighborhood basis $\mathcal{B}_x$ of x.

Example 5.27

Every metric space (X, d) is first countable, since as seen in Example 5.3, each $x \in X$ has a countable neighborhood basis $\mathcal{B}_x = \{B(x, \frac{1}{n}) \mid n \in \mathbb{N}\}$.

Example 5.28

The *line with two origins*, defined as the quotient of $\mathbb{R} \times \{\pm 1\}$ by the relation $(x, -1) \sim (x, 1)$ for all $x \neq 0$, is first countable but not Hausdorff.

(The 'two origins' are the points corresponding to $(0, \pm 1)$ — we can visualize the above space by taking two lines and gluing them together at all points except their origins.)

Proof. Let X denote our space, and let π denote the projection map $\mathbb{R} \times \{\pm 1\} \to X$.

First, to see that X is not Hausdorff, note that the sequence (x_n) defined as $x_n = \pi(\frac{1}{n}, 1) = \pi(\frac{1}{n}, -1)$ converges to both $\pi(0,1)$ and $\pi(0,-1)$ — this implies there cannot exist disjoint open sets separating $\pi(0,1)$ and $\pi(0,-1)$.

Meanwhile, to see that X is first countable, for every $x \neq \pi(0,-1)$ we can take the countable neighborhood basis $\{\pi((x - \frac{1}{n}, x + \frac{1}{n}) \times \{1\}) \mid n \in \mathbb{N}\}$, and for $x = \pi(0,-1)$ we can take the countable neighborhood basis $\{\pi((-\frac{1}{n}, \frac{1}{n}) \times \{-1\}) \mid n \in \mathbb{N}\}$. So every point has a countable neighborhood basis, and therefore X is first countable. $\square$

We'll soon prove that under the assumption of first countability, the converse of Proposition 5.18 is indeed true. Before we do so, we'll establish the following lemma (which will be useful for many similar proofs as well).

Lemma 5.29

If X is first countable, then for every $x \in X$ we can find a sequence $U_1 \supseteq U_2 \supseteq U_3 \supseteq \cdots$ of open sets containing x, such that any sequence (x_n) with $x_n \in U_n$ for each n must converge to x.

Proof. Let $\mathcal{B}_x = \{B_n\}_{n \in \mathbb{N}}$ be a countable neighborhood basis of x, and for each $n \in \mathbb{N}$, define V_n to be an open set such that $x \in V_n \subseteq B_n$ (such a set V_n must exist because B_n is a neighborhood of x). Then define $U_n = V_1 \cap \cdots \cap V_n$. It's clear that the sets U_n are nested open sets containing x. To show that they have the final property, consider any sequence (x_n) with $x_n \in U_n$ for all n, and let U be any open set containing x. Then since U is a neighborhood of x and $\mathcal{B}_x$ is a neighborhood basis, there must exist some N such that $V_N \subseteq B_N \subseteq U$, which means $U_n \subseteq U$ for all $n \geq N$. So then $x_n \in U_n \subseteq U$ for all $n \geq N$, as desired. $\square$

Now using this lemma, we can prove the converse of Proposition 5.18 for first countable spaces.

Theorem 5.30

Suppose that X is a first countable topological space. Then X is Hausdorff if and only if every convergent sequence has a *unique* limit.

Proof. We've already proven the forwards direction in Proposition 5.18 (in particular, it holds for *all* spaces X, not just first countable ones). To prove the backwards direction, assume that X is *not* Hausdorff. Choose x and y such that there do not exist disjoint open sets separating x and y; we will construct a sequence (x_n) which converges to both x and y.

Consider the chains of open sets $U_1 \supseteq U_2 \supseteq \cdots$ and $V_1 \supseteq V_2 \supseteq \cdots$ obtained from applying Lemma 5.29 to x and y respectively. Then for each n, by assumption U_n and V_n are not disjoint, so we can choose some $x_n \in U_n \cap V_n$. Then (by Lemma 5.29) the sequence (x_n) converges to both x and y. $\square$

§5.5 Sequential Properties

It turns out that several of the topological properties we've seen earlier — in particular closedness and continuity — have sequential versions. These versions will be related but not equivalent to the usual notions; but under the additional assumption of first countability, they will in fact be equivalent.

§5.5.1 Sequential Closedness

Definition 5.31. Given a topological space X, a subset $A \subseteq X$ is *sequentially closed* if whenever (a_n) is a sequence in A converging to some $x \in X$, we have $x \in A$.

Example 5.32

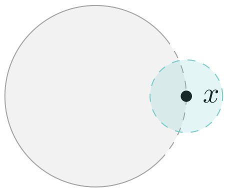
The subset $(0, 1) \subseteq \mathbb{R}$ is *not* sequentially closed, since the sequence (x_n) defined as $x_n = \frac{1}{n} \in (0, 1)$ converges to $0 \notin (0, 1)$.

Theorem 5.33

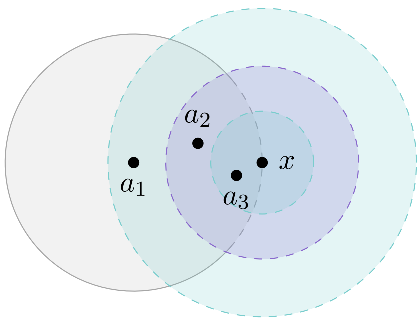
In any topological space X, every closed subset $A \subseteq X$ is sequentially closed. Furthermore, if X is first countable, then every sequetially closed $A \subseteq X$ is also closed.

Proof. First we'll show that if $A \subseteq X$ is closed, then it is sequentially closed. Assume for contradiction that there exists a sequence (a_n) in A converging to a point $x \notin A$. Since A is closed, $X \setminus A$ is an open set containing x, so there must exist N such that for all $n \geq N$ we have $a_n \in X \setminus A$, which is a contradiction as $a_n \in A$ for all n.

Now we'll prove the converse for first countable spaces X — we'll show that if $A \subseteq X$ is *not* closed, then it is not sequentially closed. Since $X \setminus A$ is not open, it is not a neighborhood of each of its points, so we can choose some $x \in X \setminus A$ such that $X \setminus A$ is not a neighborhood of x — this means every open set U containing x must intersect A.



Use Lemma 5.29 to choose a chain $U_1 \supseteq U_2 \supseteq \cdots$ of open sets containing x such that any sequence with $x_n \in U_n$ for all n converges to x. Now since each U_n must intersect A, we can define a sequence (a_n) such that $a_n \in U_n \cap A$ for all n.



Then (a_n) is a sequence in A which converges to $x \notin A$, so A is not sequentially closed. □

§5.5.2 Sequential Continuity

A very similar story holds for continuity and sequential continuity.

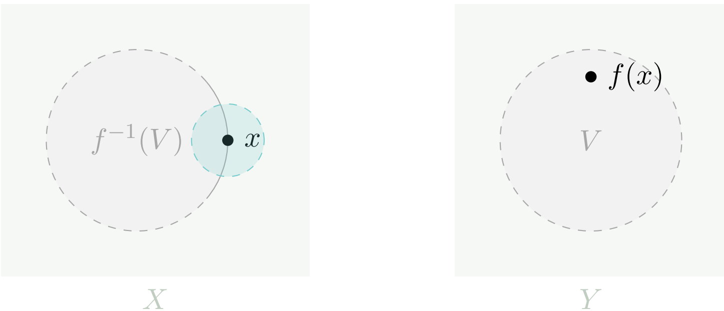
Definition 5.34. A map $f\colon X \to Y$ is *sequentially continuous* at a point $x \in X$ if for every sequence (x_n) in X with $x_n \to x$, we have $f(x_n) \to f(x)$. (We say f is *sequentially continuous* if it is sequentially continuous at every point.)

Theorem 5.35

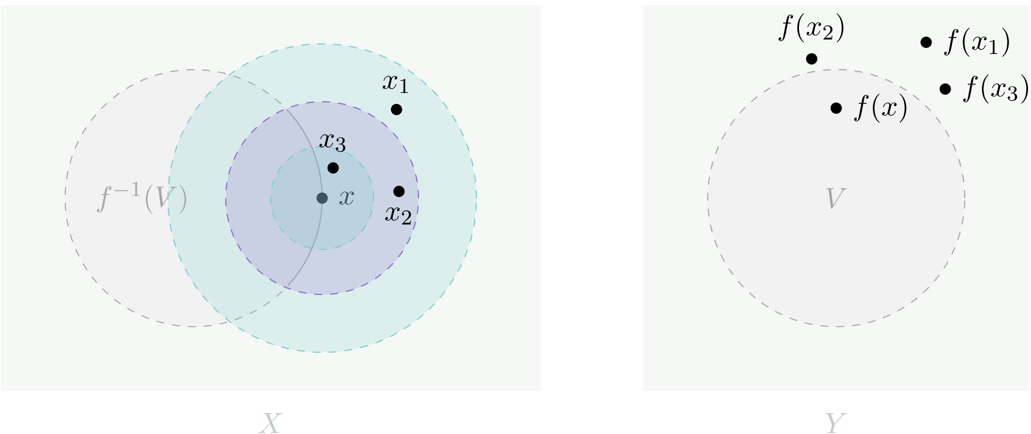
If a map $f\colon X \to Y$ is continuous at a point $x \in X$, then f is also sequentially continuous at x. If X is first countable then the converse holds as well.

Proof. First we'll prove that if f is continuous at x, then it is sequentially continuous at x. Consider any sequence (x_n) in X with $x_n \to x$. We wish to show that $f(x_n) \to f(x)$ in Y, meaning that for every neighborhood $V \subseteq Y$ of $f(x)$, for all large n we have $f(x_n) \in V$. But given any neighborhood V of $f(x)$, by the continuity of f we know that $f^{-1}(V)$ is a neighborhood of x, and therefore there exists an open set $U \subseteq X$ with $x \in U \subseteq f^{-1}(V)$. Then since $x_n \to x$, there must exist N such that for all $n \geq N$ we have $x_n \in U \subseteq f^{-1}(V)$, and therefore $f(x_n) \in f(U) \subseteq V$.

Now we'll prove the converse when X is first countable — we'll show that if f is not continuous at x, then f is not sequentially continuous at x either. Since f is not continuous at x, there exists a neighborhood $V \subseteq Y$ of $f(x)$ such that $f^{-1}(V) \subseteq X$ is not a neighborhood of x, which means that any open set $U \subseteq X$ containing x cannot be contained in $f^{-1}(V)$.



Now use Lemma 5.29 to again produce a chain $U_1 \supseteq U_2 \supseteq \cdots$ of open sets in X containing x such that any sequence (x_n) with $x_n \in U_n$ for each n converges to x. Since none of the sets U_n can be contained in $f^{-1}(V)$, we can construct a sequence (x_n) with $x_n \in U_n \cap (X \setminus f^{-1}(V))$ for each n.



Then $x_n \in U_n$ for each n, so $x_n \to x$. However, x_n is not in $f^{-1}(V)$ for any n, and therefore $f(x_n)$ is not in V; since V is a neighborhood of $f(x)$, this means we cannot have $f(x_n) \to f(x)$. $\square$

§6 Connectedness

We'll now see the notion of connectedness, which gives one useful adjective to describe a topological space.

§6.1 Definition and First Examples

Definition 6.1. A space X is *connected* if for any disjoint open sets U and V such that $X = U \cup V$, either U or V is empty.

In other words, a space X is *not* connected if and only if we can write X as a union of two nonempty disjoint open sets U and V. This definition captures what we'd intuitively expect — a space is disconnected if it splits into two separate pieces.

Remark 6.2. From the definition, it is often fairly easy to show that a space *isn't* connected — we can simply provide a decomposition $X = U \cup V$ — but it can be much harder to prove that a space *is* connected (i.e., that no such decomposition exists). Later we'll prove that $[a, b]$ is connected; this makes intuitive sense, but is not easy to prove.

Remark 6.3. Connectedness is a topological notion (i.e., it is preserved under homeomorphism), since it is defined purely in terms of open sets. This means it gives us a tool to tell that two spaces are *not* homeomorphic — if one space is connected and the other isn't, then they can't be homeomorphic.

The definition of connectedness can be reformulated in terms of *clopen* sets — sets which are both open and closed — in the following way.

Lemma 6.4

A space X is connected if and only if $\emptyset$ and X are the only clopen sets.

Proof. If there exists a clopen set $U \subsetneq X$ other than $\emptyset$ and X, then we can write X as the disjoint union of U and $X \setminus U$. Both are nonempty (as U is not $\emptyset$ or X) and open (the fact that U is closed means $X \setminus U$ is open), so X is not connected.

Conversely, if X is not connected, then we can write X as a disjoint union of nonempty open sets U and V. Then $V = X \setminus U$ is open, so U is also closed, and is therefore a clopen set other than $\emptyset$ and X. $\square$

We'll also often use the following rephrasing of connectedness for subspaces.

Lemma 6.5

If $Y \subseteq X$ is a subspace, then Y is connected if and only if for all open $U, V \subseteq X$ such that $Y \subseteq U \cup V$ and $Y \cap U \cap V = \emptyset$, we have $Y \subseteq U$ or $Y \subseteq V$.

This can be easily proven using the definition of the subspace topology — the decomposition of Y corresponding to such sets U and V is $(Y \cap U) \cup (Y \cap V)$.

We'll now see a few examples of connectedness.

Example 6.6

Any space X with the trivial topology is connected, as the only open sets are X and $\emptyset$.

Example 6.7

Any space X with the discrete topology (and with $|X| \geq 2$) is *not* connected, as every set is clopen.

Alternatively, we can write X as the disjoint union of $\{x\}$ and $X \setminus \{x\}$ (both of which are open in the discrete topology) for any point $x \in X$.

Example 6.8

The space $[-1, 0) \cup (0, 1)$ is not connected, as we can write it as the disjoint union of $[-1, 0)$ and $(0, 1)$, both of which are open in the subspace topology.

Alternatively, we could apply Lemma 6.5 with $U = (-2, 0)$ and $V = (0, 1)$.

Example 6.9

All intervals $[a, b]$ and (a, b) are connected — we will prove this next class.

§6.2 Some Properties of Connectedness

Now we'll see some facts that let us deduce the connectedness of one space from that of another.

Proposition 6.10

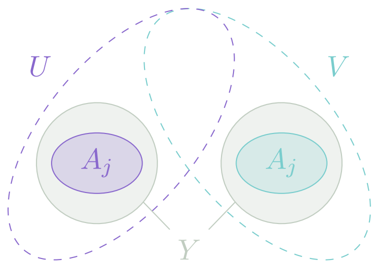
Let X be a space and let $\{A_i\}_{i \in I}$ be a collection of connected subspaces of X with $\bigcap_{i \in I} A_i \neq \emptyset$. Then $\bigcup_{i \in I} A_i$ is connected as well.

Intuitively this makes sense — if we have a bunch of connected blobs which overlap, their union should be connected as well.

Proof. Let $\bigcup_{i \in I} A_i = Y$, and assume that $U, V \subseteq X$ are open with $Y \subseteq U \cup V$ and $Y \cap U \cap V = \emptyset$; we want to show that $Y \subseteq U$ or $Y \subseteq V$.

First, for every $i \in I$, since $A_i \subseteq Y$ we must have $A_i \subseteq U \cup V$ and $A_i \cap U \cap V = \emptyset$; since A_i is connected, this means $A_i \subseteq U$ or $A_i \subseteq V$.

But we cannot have $A_i \subseteq U$ and $A_j \subseteq V$ for any i and j — if this were the case, then since $Y \cap U \cap V = \emptyset$ (and A_i and A_j are both subsets of Y) we would have $A_i \cap A_j = \emptyset$, which is impossible since we assumed the intersection of *all* the sets A_i was nonempty.



So then either $A_i \subseteq U$ for all i or $A_i \subseteq V$ for all i. In the first case we have $Y \subseteq U$, while in the second we have $Y \subseteq V$, so we're done. □

Remark 6.11. Note that this proof only requires the *pairwise* intersections $A_i \cap A_j$ to be nonempty, so the proposition is true in slightly greater generality than stated.

Proposition 6.12

Let X be a space and let $A, B \subseteq X$. If $A \subseteq B \subseteq \overline{A}$ and A is connected, then B is connected.

Recall that $\overline{A}$ denotes the *closure* of A, defined as the set of points $x \in X$ for which $X \setminus A$ is not a neighborhood of x; equivalently, we saw in Corollary 1.18 that $\overline{A}$ is the smallest closed set containing A.

Proof. We want to show that for any open sets $U, V \subseteq X$ with $B \subseteq U \cup V$ and $B \cap U \cap V = \emptyset$, we must have $B \subseteq U$ or $B \subseteq V$. Let U and V be two such sets.

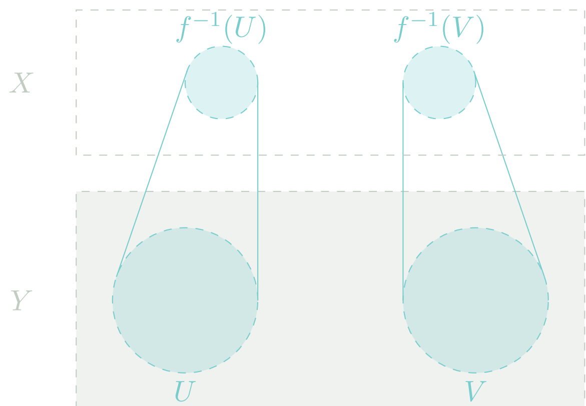
First, since $A \subseteq B$, we must have $A \subseteq U \cup V$ and $A \cap U \cap V = \emptyset$ as well; since A is connected, this means $A \subseteq U$ or $A \subseteq V$. Assume that $A \subseteq U$, so that $A \cap V = \emptyset$. We will show that $B \subseteq U$ as well.

Assume for the sake of contradiction that there is some $x \in B$ with $x \notin U$, which means $x \in V$. Then since $V \subseteq X \setminus A$ is open and contains x, then $X \setminus A$ must be a neighborhood of x, contradicting the fact that $x \in B \subseteq \overline{A}$ (and $\overline{A}$ is defined as the points $x \in X$ for which $X \setminus A$ is *not* a neighborhood of x). □

Proposition 6.13

If $f\colon X \to Y$ is continuous and X is connected, then $f(X)$ is connected.

Proof. We'll show the contrapositive — that if $f(X)$ is *not* connected, then neither is X. Since $f(X)$ is not connected, we can find nonempty disjoint open sets U and V in $f(X)$ with union $f(X)$. We claim that then $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint open sets in X with union X.



The sets $f^{-1}(U)$ and $f^{-1}(V)$ are nonempty because U and V are nonempty subsets of $\text{im}(f)$, and are disjoint because U and V are disjoint. Furthermore, U and V are open in $f(X)$ and f is continuous when viewed as a map $X \to f(X)$ (since restricting the target preserves continuity), so $f^{-1}(U)$ and $f^{-1}(V)$ are open. Finally we have $f(X) = U \cup V$, so $X = f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$.

This means X is not connected, as desired. ∎

Corollary 6.14

If a space X is connected, then any quotient $X/\sim$ is connected as well.

Proof. Let $\pi\colon X \to X/\sim$ denote the corresponding projection map; then $\pi(X) = X/\sim$, and the result follows from the above proposition. ∎

Proposition 6.15

Given spaces $X_1, \ldots, X_n$, their product $X_1 \times \cdots \times X_n$ is connected if and only if each of the spaces $X_1, \ldots, X_n$ is connected.

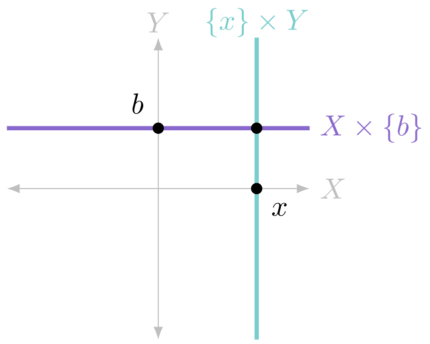
(This is true for both the box and the product topology, as the two are the same for finite products.)

Proof. For the forwards direction, suppose that $X_1 \times \cdots \times X_n$ is connected. For each index i, the projection map $\pi_i\colon X_1 \times \cdots \times X_n \to X_i$ is continuous and has image X_i, so X_i is connected by Proposition 6.13.

To prove the converse, it suffices to show that if X and Y are connected, then $X \times Y$ is connected (the general case then follows from the fact that we can write $X_1 \times \cdots \times X_n \cong (X_1 \times \cdots \times X_{n-1}) \times X_n$).

In order to prove this, we'll use Proposition 6.10 — fix some $b \in Y$, and for each $x \in X$ define

$$T_x = (\{x\} \times Y) \cup (X \times \{b\}).$$



Then each set T_x is connected by Proposition 6.10, as both $\{x\} \times Y$ and $X \times \{b\}$ are connected (they are homeomorphic to Y and X respectively) and their intersection is $\{(x, b)\} \neq \emptyset$.

Meanwhile we have $X \times Y = \bigcup_{x \in X} T_x$, and $\bigcap_{x \in X} T_x = X \times \{b\}$ is nonempty. So then $X \times Y$ is connected by another application of Proposition 6.10. (The purpose of including the set $X \times \{b\}$ in each T_x was so that they would have nonempty intersection — the statement $X \times Y = \bigcup_{x \in X} T_x$ would remain true even if we replaced T_x with just $\{x\} \times Y$.) ∎

Remark 6.16. The above proposition shows that a *finite* product of connected spaces is connected. The same is in fact true for *infinite* products of connected spaces, using the product topology. However, it's *not* true in general for infinite products with the *box* topology — for example, in $\mathbb{R}^{\mathbb{N}}$ with the box topology, the set of bounded sequences

$$\{(x_1, x_2, \ldots) \mid \text{exists } c > 0 \text{ such that } |x_i| \leq c \text{ for all } i\}$$

is clopen, and therefore $\mathbb{R}^{\mathbb{N}}$ is not connected.

§6.3 Connectedness in $\mathbb{R}$

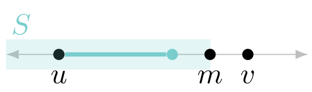
Theorem 6.17

Every open interval $(a, b) \subseteq \mathbb{R}$ is connected.

Proof. Assume for contradiction that (a, b) is not connected, so there exist nonempty disjoint open sets U and V in (a, b) with $U \cup V = (a, b)$. Choose some $u \in U$ and $v \in V$, and without loss of generality assume $u < v$. Now consider the set

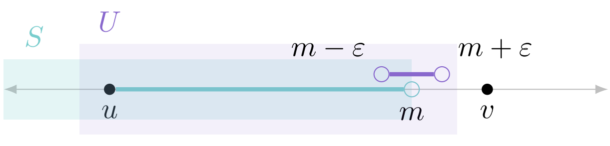
$$S = \{s \in (a, b) \mid [u, s] \subseteq U\}.$$

Note that S is nonempty (as it contains u) and bounded, so S has a supremum $m = \sup S$. We will show that m is in (a, b) but cannot be in either U or V, which will give the desired contradiction.



First note that for all $w \in V$ with $w > u$, we have $m \leq w$, as w is an upper bound for S (if $s \geq w$ then w is in $[u, s]$, so $[u, s]$ cannot be contained in U). In particular this means $u \leq m \leq v$, so m is in (a, b).

Now if m is in U, then since U is open, there must exist some $\varepsilon > 0$ such that $(m - \varepsilon, m + \varepsilon) \subseteq U$. But we have $[u, m) \subseteq U$ (if some $x \in [u, m)$ were not in U, then x would be an upper bound for S), so combining these gives that $[u, m + \varepsilon) \subseteq U$. Then S must contain $m + \frac{1}{2}\varepsilon$ (for example), contradicting the fact that m is an upper bound for S. So m cannot be in U.



Meanwhile, if m is in V, then there must exist some $\varepsilon > 0$ such that $(m - \varepsilon, m + \varepsilon) \subseteq V$. But then $m - \varepsilon$ is an upper bound for S, contradicting the fact that m is the least upper bound.

So m is in (a, b) but not U or V, contradicting the fact that $(a, b) = U \cup V$. $\square$

Corollary 6.18

All intervals of the form (a, b), $[a, b)$, $(a, b]$, $[a, b]$, $(-\infty, a)$, $(-\infty, a]$, (a, ∞), $[a, \infty)$, and $\mathbb{R}$ are connected.

Proof. First, we've seen that all open intervals are homeomorphic, so since (a, b) is connected, the same is true for (a, ∞), $(-\infty, a)$, and $\mathbb{R}$. The connectedness of the remaining intervals follows from Proposition 6.12, which states that if A is connected and $A \subseteq B \subseteq \overline{A}$ then B is connected as well (for example, we have $(a, b) \subseteq [a, b) \subseteq \overline{(a, b)} = [a, b]$, so $[a, b)$ is connected). $\square$

§6.3.1 The Intermediate Value Theorem

In analysis, we've seen the *intermediate value theorem*, which states that a continuous function $f\colon [a, b] \to \mathbb{R}$ takes on all values between $f(a)$ and $f(b)$. It turns out that the relevant property of $[a, b]$ is that it is connected, and this generalizes to continuous maps from *any* connected space.

Theorem 6.19

Suppose X is connected and $f\colon X \to \mathbb{R}$ is continuous. Then for any $a, b \in X$ and any $r \in \mathbb{R}$ with $f(a) < r < f(b)$, there exists $c \in X$ with $f(c) = r$.

Proof. Assume for contradiction that there is no such c. Since X is connected and f is continuous, we know that $f(X)$ is connected as well. But since $r \notin f(X)$, we can split $f(X)$ as

$$f(X) = (f(X) \cap (-\infty, r)) \cup (f(X) \cap (r, \infty)).$$

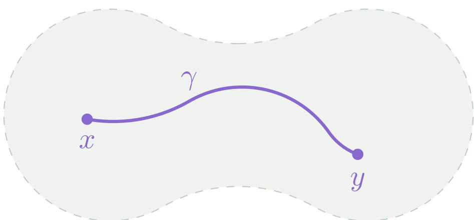
These two sets are disjoint and nonempty (as the first set contains $f(a)$ and the second contains $f(b)$), and both are open (as $(-\infty, r)$ is open in $\mathbb{R}$, so $f(X) \cap (-\infty, r)$ is open in $f(X)$ under the subspace topology). So $f(X)$ is a union of two nonempty disjoint open sets, contradicting its connectedness. $\square$

§6.4 Path-connectedness

We'll now define another notion of connectedness, called *path-connectedness*. This definition will be useful because we'll see later that it implies connectedness, and it's often easier to check.

§6.4.1 Definition of a Path

Definition 6.20. Let X be a topological space, and let x and y be points in X. A *path* from x to y is a continuous map $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$.



Note that we could have used any interval $[a, b]$ instead of $[0, 1]$, since all such intervals are homeomorphic.

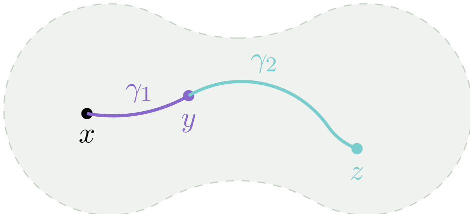
The following fact will frequently be useful — it essentially states that the concatenation of two paths is again a path.

Proposition 6.21

Suppose that $\gamma_1\colon [0,1] \to X$ is a path from x to y, and $\gamma_2\colon [0,1] \to X$ is a path from y to z. Then the function $\gamma\colon [0,1] \to X$ defined as

$$\gamma(t) = \begin{cases} \gamma_1(2t) & \text{if } t \in [0, \frac{1}{2}] \\ \gamma_2(2t-1) & \text{if } t \in [\frac{1}{2}, 1] \end{cases}$$

is a path from x to z.



Intuitively, γ is the path where we first go along γ_1 twice as fast (so we're essentially using $[0, \frac{1}{2}]$ rather than $[0, 1]$ as the input to γ_1), and then go along γ_2 twice as fast (using $[\frac{1}{2}, 1]$ rather than $[0, 1]$ as the input). Note that we've defined $\gamma(\frac{1}{2})$ in both cases above, but these definitions agree (both define it as y).

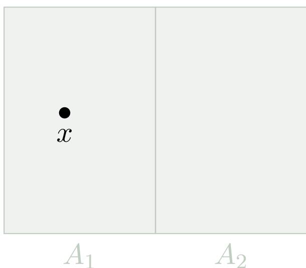
In order to prove this, it's enough to show that γ is continuous. This follows immediately from the following more general lemma (applied to the closed sets $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$).

Lemma 6.22

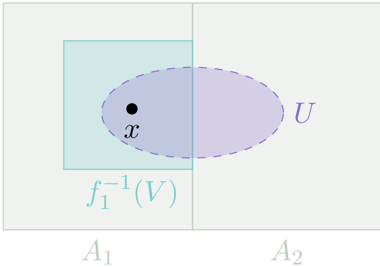
Let X be a space and let A_1 and A_2 be closed subsets of X. Then given a map $f\colon X \to Y$, if the restrictions of f to A_1 and A_2 are both continuous, then f is continuous as well.

Proof. We'll check that f is continuous at every point — i.e., that for all points $x \in X$ and all neighborhoods $V \subseteq Y$ of $f(x)$, the set $f^{-1}(V)$ is a neighborhood of x. Let f_1 and f_2 denote the restrictions of f to A_1 and A_2 respectively, and note that $f_1^{-1}(V) \subseteq f^{-1}(V)$ and $f_2^{-1}(V) \subseteq f^{-1}(V)$ for any $V \subseteq Y$.

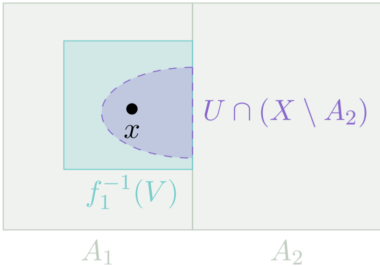
First consider the case where x is in A_1 but not A_2. (The case where x is in A_2 but not A_1 is the same.)



Then given any neighborhood $V \subseteq Y$ of $f(x)$, since f_1 is continuous we know that $f_1^{-1}(V)$ is a neighborhood of x in A_1, so there exists an open set (here 'open' refers to the subspace topology on A_1) which contains x and is contained in $f_1^{-1}(V)$. By the definition of the subspace topology, we can write this set as $U \cap A_1$ for an open set $U \subseteq X$ (here 'open' refers to the topology on X).

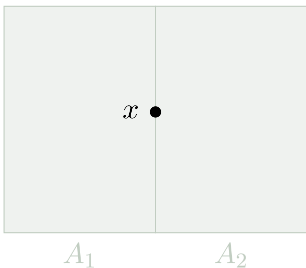


Now consider the set $U \cap (X \setminus A_2)$. Since both U and $X \setminus A_2$ are open, this set is open as well; meanwhile since $X = A_1 \cup A_2$, we have $X \setminus A_2 \subseteq A_1$, and therefore $U \cap (X \setminus A_2) \subseteq U \cap A_1 \subseteq f_1^{-1}(V) \subseteq f^{-1}(V)$.

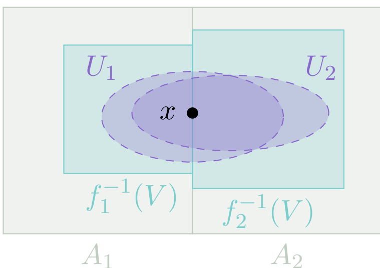


Finally, since $x \notin A_2$ we have $x \in X \setminus A_2$. So $U \cap (X \setminus A_2)$ is an open set which contains x and is contained in $f^{-1}(V)$, and therefore $f^{-1}(V)$ is indeed a neighborhood of x.

Now consider the case where x is in both A_1 and A_2.



Then by the same reasoning, we can find open sets U_1 and U_2 in X, both of which contain x, such that $U_1 \cap A_1 \subseteq f_1^{-1}(V)$ and $U_2 \cap A_2 \subseteq f_2^{-1}(V)$.



Now consider the set $U_1 \cap U_2$, which is open and contains x (since both U_1 and U_2 are open and contain x). We claim that $U_1 \cap U_2$ is contained in $f^{-1}(V)$ — this is because every point in $U_1 \cap U_2$ must be in either A_1 or A_2, and $U_1 \cap U_2 \cap A_1 \subseteq U_1 \cap A_1 \subseteq f_1^{-1}(V) \subseteq f(V)$, while $U_1 \cap U_2 \cap A_2 \subseteq U_2 \cap A_2 \subseteq f_2^{-1}(V) \subseteq f(V)$.

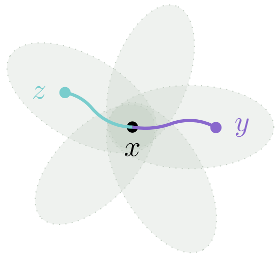
So $U_1 \cap U_2$ is an open set in X which contains x and is contained in $f^{-1}(V)$, which means that $f^{-1}(V)$ is again a neighborhood of x. $\square$

§6.4.2 Definition of Path-Connectedness

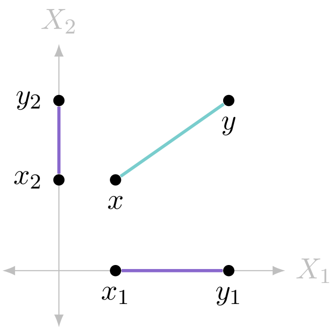
Definition 6.23. A space X is *path-connected* if for any two points x and y in X, there exists a path from x to y in X.

Path-connectedness satisfies many of the properties we saw for connectedness in Subsection 6.2.

- If $\{A_i\}_{i\in I}$ is a collection of path-connected spaces with nonempty intersection, their union $\bigcup_{i\in I} A_i$ is path-connected as well — given any points $y \in A_i$ and $z \in A_j$, fix a point $x \in A_i \cap A_j$. Then we can find a path from y to x in A_i and a path from z to x in A_j, and concatenating these paths gives a path from y to z in $\bigcup_{i\in I} A_j$.



- Any product of path-connected spaces is path-connected (this is true for both finite and infinite products with the product topology) — given a product $\prod_{i\in I} X_i$ of path-connected spaces X_i and two points $x = (x_i)_{i\in I}$ and $y = (y_i)_{i\in I}$, for each i we can find a path γ_i from x_i to y_i in X_i, and combining these (i.e., letting $\gamma\colon [0,1] \to \prod_{i\in I} X_i$ be the function whose ith coordinate is γ_i) produces a path γ from x to y.



- If $f\colon X \to Y$ is continuous and X is path-connected, then $f(X)$ is path-connected — if $\gamma\colon [0,1] \to X$ is a path from x to y, then $f \circ \gamma\colon [0,1] \to f(X)$ is a path from $f(x)$ to $f(y)$. In particular, path-connectedness is preserved under homeomorphism (i.e., it is a topological property), and any quotient of a path-connected space is path-connected.

§6.4.3 Path-Connectedness vs. Connectedness

Proposition 6.24

If X is path-connected, then X is also connected.

Proof. Assume for contradiction that X is path-connected but not connected, so $X = U \cup V$ for some disjoint nonempty open sets U and V. Fix $u \in U$ and $v \in V$, and consider a path $\gamma\colon [0,1] \to X$ from u to v. Since γ is continuous, both $\gamma^{-1}(U)$ and $\gamma^{-1}(V)$ must be open. However, they must also be disjoint and have union $[0,1]$ (since U and V are disjoint and have union X), and both are nonempty as $0 \in \gamma^{-1}(U)$ and $1 \in \gamma^{-1}(V)$. This contradicts the fact that $[0,1]$ is connected. $\square$

The converse is not true, however, as the following counterexample shows.

Example 6.25

The *topologist's sine curve* $\{(t, \sin \frac{1}{t}) \mid t > 0\} \cup \{(0,0)\}$ is connected but not path-connected.

(We won't prove this, but proving that it is connected is not that hard.)

However, there's a fairly mild additional assumption we can make under which the converse *does* become true. (As with many of the other results we've seen, the general converse isn't true, but it becomes true once we impose a reasonable condition.)

Definition 6.26. A space X is *locally path-connected* if for every $x \in X$, there is a path-connected neighborhood of x.

Proposition 6.27

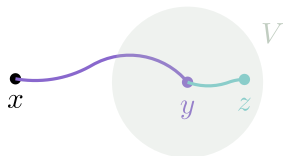
If X is connected and locally path-connected, then X is path-connected.

Proof Sketch. Given a point $x \in X$, consider the set

$$U = \{y \in X \mid \text{exists a path from } y \text{ to } x\}.$$

We'll show that $U = X$, which will imply that X is path-connected. In order to do this, we'll show that U is clopen — this suffices because since X is connected the only clopen sets are $\emptyset$ and X, and U is nonempty as it contains x, so then it must be X.

In order to show that U is open, we'll show that U is a neighborhood of each of its points. Fix any point $y \in U$. Then since X is locally path-connected, there exists a path-connected neighborhood V of y. But then we must have $V \subseteq U$, since for every $z \in V$, we can find a path from z to y (since V is path-connected) and a path from y to x (since y is in U), and concatenating them gives a path from z to x.



Then since V is a neighborhood of y and $V \subseteq U$, this implies U is a neighborhood of y as well.

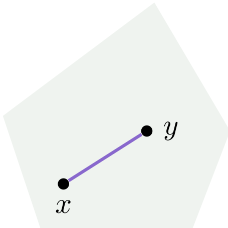
Similarly, to show that U is closed, we'll show that $X \setminus U$ is a neighborhood of each of its points. Fix any point $y \in X \setminus U$, and consider a path-connected neighborhood V of y. Then we must have $V \subseteq X \setminus U$ — if there were any point $z \in U \cap V$, then there would exist a path from y to z and from z to x, so concatenating them would give a path from y to x, contradicting the fact that $y \notin U$.

So U is both open and closed, as desired. $\square$

§6.4.4 Some Examples of Path-Connectedness

We'll now see a few examples of path-connectedness.

Definition 6.28. A set $X \subseteq \mathbb{R}^n$ is *convex* if for any two points x and y in X, the line segment $[x, y] = \{(1-t)x + ty \mid t \in [0, 1]\}$ is contained in X.



Example 6.29

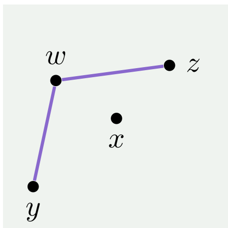
Any convex subspace $X \subseteq \mathbb{R}^n$ is path-connected, and therefore connected.

Proof. Given any two points x and y in X, the map $\gamma\colon [0,1] \to X$ defined as $\gamma(t) = (1-t)x + ty$ is a path from x to y — this map is clearly continuous, and its image is contained in X by convexity. □

Example 6.30

For every $n > 1$ and any $x \in \mathbb{R}^n$, the space $\mathbb{R}^n \setminus \{x\}$ is path-connected (and therefore connected).

Proof. Given any two points y and z in $\mathbb{R}^n \setminus \{x\}$, we wish to find a path between y and z in $\mathbb{R}^n$ not passing through x. To do so, choose a point w which does not lie on the lines xy or xz.



Then the line segment yw gives a path from y to w in $\mathbb{R}^n \setminus \{x\}$ (given by the equation $t \mapsto (1-t)y + tw$), and the line segment wz gives a path from w to z. Concatenating them gives a path from y to z, as desired. □

Example 6.31

For every $n \geq 1$, the sphere $\mathbb{S}^n$ is path-connected, and therefore connected.

This is intuitively clear, as by walking along the sphere we can reach any point from another.

Proof. Given two points x and y in $\mathbb{S}^n$, we wish to show that there exists a path between them. To do so, fix a point z on the sphere other than x and y; then we've seen (in Example 2.14) that $\mathbb{S}^n \setminus \{z\}$ is homeomorphic to $\mathbb{R}^n$. Since $\mathbb{R}^n$ is path-connected, then $\mathbb{S}^n \setminus \{z\}$ is path-connected as well. So there exists a path from x to y in $\mathbb{S}^n \setminus \{z\}$, which must also be a path from x to y in $\mathbb{S}^n$. □

These examples illustrate one reason the notion of path-connectedness is useful — proving an interval $[a, b]$ is connected was quite hard, but now path-connectedness allows us to find much simpler proofs that more complicated spaces are connected. (Note that the fact path-connectedness implies connectedness relies on the fact that $[a, b]$ is connected.)

Connectedness allows us to tell certain spaces apart, as seen in the following simple example.

Example 6.32

The space $\mathbb{R}$ is not homeomorphic to $\mathbb{R}^n$ for any $n \geq 2$.

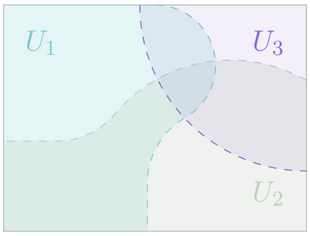
Proof. Assume for contradiction that there exists a homeomorphism $\varphi\colon \mathbb{R} \to \mathbb{R}^n$, and let $x = \varphi(0)$. Then φ restricts to a homeomorphism $\mathbb{R} \setminus \{0\} \to \mathbb{R}^n \setminus \{x\}$. But $\mathbb{R} \setminus \{0\}$ is not connected — we can write it as $(-\infty, 0) \cup (0, \infty)$ — while $\mathbb{R}^n \setminus \{x\}$ is connected, contradiction. □

§7 Compactness

We'll now see another adjective that we can use to describe topological spaces, the notion of *compactness*. Compactness will be another tool we can use to distinguish between topological spaces — for example, we'll be able to show that $\mathbb{S}^2$ is not homeomorphic to $\mathbb{R}^2$.

§7.1 Definition of Compactness

Definition 7.1. Given a topological space X, we say that a collection of open sets $\mathcal{U} = \{U_i\}_{i \in I}$ is an *open cover* of X if $X = \bigcup_{i \in I} U_i$.



Definition 7.2. A space X is *compact* if every open cover of X admits a finite subcover.

A finite subcover of an open cover $\mathcal{U}$ is a finite subset of $\mathcal{U}$ which still forms an open cover of X — so explicitly, a space X is compact if for every family $\mathcal{U} = \{U_i\}_{i \in I}$ of open sets with $X = \bigcup_{i \in I} U_i$, there exist finitely many indices $i_1, \ldots, i_n$ such that $X = U_{i_1} \cup \cdots \cup U_{i_n}$.

Note that compactness is again a topological property — if X and Y are homeomorphic, then X is compact if and only if Y is. (As in the case of connectedness, this follows from the fact that compactness is defined purely in terms of open sets.)

We can describe compactness of a subspace $Y \subseteq X$ in the following way — Y is compact if and only if for every family $\{U_i\}_{i \in I}$ of open sets in X such that $Y \subseteq \bigcup_{i \in I} U_i$ — which we will again refer to as an open cover of Y — there exists a finite subcover. (This can be easily checked using the definition of the subspace topology.)

§7.2 Some Examples

Example 7.3

If X is a topological space with only finitely many open sets, then X is compact. In particular, any space with the trivial topology is compact.

This is immediate, as if there are only finitely many open sets, then every open cover is *already* finite.

Example 7.4

A space X with the discrete topology is compact if and only if X is finite.

Proof. First, if X is finite, then there are only finitely many open sets (as X only has finitely many subsets), so X is compact. Conversely, if X is infinite, then the open cover $\mathcal{U} = \{\{x\} \mid x \in X\}$ consisting of all singletons does not admit a finite subcover (since we need to include each $\{x\}$ to cover the element x). $\square$

Example 7.5

The space $\mathbb{R}$ with the standard topology is not compact.

Proof. The open cover $\mathcal{U} = \{(-n, n) \mid n \in \mathbb{N}\}$ does not have a finite subcover — the union of finitely many sets of the form $(-n, n)$ is simply $(-m, m)$ where m is the largest integer n present in the union, and is therefore not $\mathbb{R}$. $\square$

Intuitively, $\mathbb{R}$ fails to be compact because it is 'big' in some sense.

Example 7.6

The intervals (a, b), $(-\infty, a)$, and (b, ∞) are not compact (for any real numbers a and b).

This follows from the fact that each of these intervals is homeomorphic to $\mathbb{R}$.

Example 7.7

The *closed* interval $[a, b]$ is compact.

This is difficult to prove, and we will see the proof next lecture.

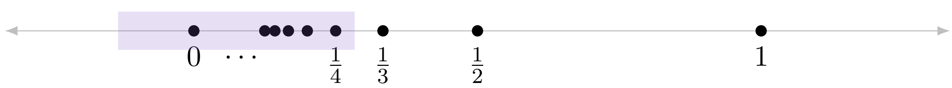
Example 7.8

The space $X = \{0\} \cup \{\frac{1}{n} \mid n \in \mathbb{N}\}$ is compact (as a subspace of $\mathbb{R}$ with the standard topology).

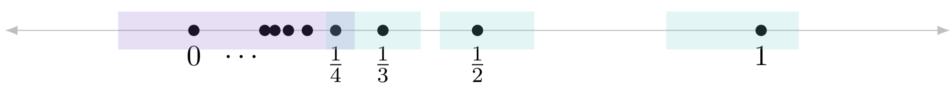
Intuitively, this makes sense because the sequence is very 'localized' — it doesn't wander off in any way (in particular, it doesn't go to ∞, and the point that it does approach is in our set).

Proof. Suppose we have an open cover $X \subseteq \bigcup_{i \in I} U_i$ (for open sets $U_i \subseteq \mathbb{R}$); we wish to find a finite subcover of this open cover.

First, since the U_i cover 0, we can find some set U among the U_i which contains 0.



Then since $\frac{1}{n} \to 0$, all but finitely many of the points $\frac{1}{n}$ must lie in U, and therefore if we place U in our finite subcover, there are only finitely many points remaining that we still have to cover.



For each of these remaining points, we can choose one set U_i containing our point; then taking U along with each of these U_i (of which there are finitely many, since we have finitely many points) gives a finite subcover. □

Example 7.9

The set $\mathbb{Q} \cap [0, 1]$ (consisting of all rational numbers between 0 and 1) is not compact.

(Of course, $\mathbb{Q}$ itself is not compact for the same reason that $\mathbb{R}$ is not compact; this statement is less obvious, but still true.)

Proof. Fix an irrational number $x \in [0, 1]$, and consider the collection $\mathcal{U} = \{[0, 1] \setminus [x - \varepsilon, x + \varepsilon] \mid \varepsilon > 0\}$. These sets are each open in $[0, 1]$ (as their complements $[x - \varepsilon, x + \varepsilon]$ are closed), and they cover $\mathbb{Q} \cap [0, 1]$ because their union is $[0, 1] \setminus x$, and $x \notin \mathbb{Q}$, so they form an open cover of $\mathbb{Q} \cap [0, 1]$. But they do not admit a finite subcover — if we only took finitely many sets corresponding to ε_1, ..., ε_n, then letting $\varepsilon = \min(\varepsilon_1, \ldots, \varepsilon_n)$ the union of our sets would not touch $[x - \varepsilon, x + \varepsilon]$, and there must exist a rational number in this interval. □

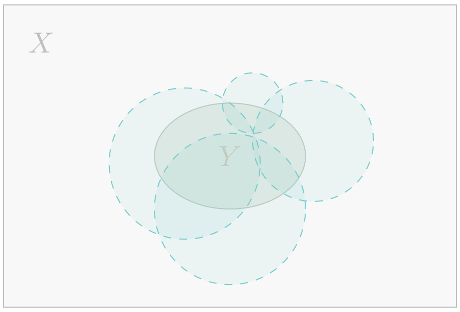
§7.3 Some Properties of Compactness

Now we'll prove a few properties regarding compactness.

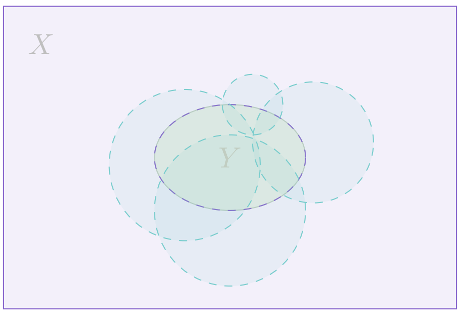
Proposition 7.10

If X is compact and $Y \subseteq X$ is closed, then Y is compact.

Proof. Suppose we are given an open cover $\mathcal{U} = \{U_i\}_{i \in I}$ of Y, i.e., open sets $U_i \subseteq X$ such that $Y \subseteq \bigcup_{i \in I} U_i$.



Then the set $X \setminus Y$ is open (since Y is closed), so $\mathcal{U} \cup \{(X \setminus Y)\}$ forms an open cover of X.



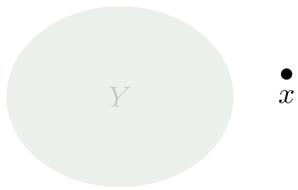
Since X is compact, we can find a finite subcover of this cover of X. Then we can simply remove the set $X \setminus Y$ to obtain a finite subcover of our original cover of Y (if $X \setminus Y$ is not present in the subcover of X, we do not need to perform any modifications). $\square$

The converse is not true in general — for example, in the trivial topology, every subspace is compact but no nontrivial subspace is closed. However, under the condition that X is *Hausdorff*, it *is* true that every compact subspace is closed (this does not require X to be compact).

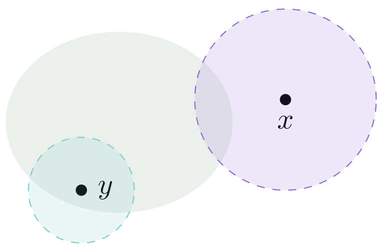
Proposition 7.11

If X is Hausdorff and $Y \subseteq X$ is compact, then Y must be closed.

Proof. We will show that $X \setminus Y$ is open by showing that it is a neighborhood of each of its points. Fix a point $x \in X \setminus Y$, so we wish to find an open set $V \subseteq X \setminus Y$ such that $x \in V$.

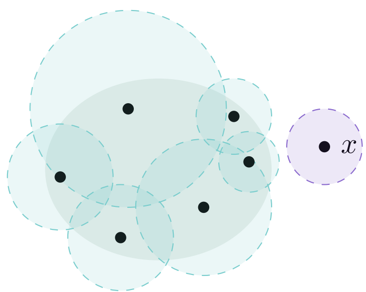


First, given any point $y \in Y$, since $x \neq y$ and X is Hausdorff, we can find disjoint open sets U_y and V_y such that $y \in U_y$ and $x \in V_y$.



Now the collection $\mathcal{U} = \{U_y \mid y \in Y\}$ forms an open cover of Y, since each point $y \in Y$ is covered by its corresponding set U_y.

Since Y is compact, this open cover must have a finite subcover, so we can find points $y_1, \ldots, y_n$ such that $Y \subseteq U_{y_1} \cup \cdots \cup U_{y_n}$. Now set $V = V_{y_1} \cap \cdots \cap V_{y_n}$. We will show that V has the desired properties.



First, we have $x \in V$ because by definition $x \in V_y$ for all y. Meanwhile, V and Y must be disjoint — for each $y \in Y$, since the sets U_{y_i} cover Y there must exist i with $y \in U_{y_i}$, and since U_{y_i} and V_{y_i} are disjoint by construction, y cannot be in V_{y_i}, and therefore cannot be in V. Finally, V is open because it is a finite intersection of open sets.

So we have shown that V is an open set with $x \in V \subseteq X \setminus Y$. This means $X \setminus Y$ is a neighborhood of each of its points and is therefore open, so Y is closed. $\square$

Proposition 7.12

If X is compact and $f\colon X \to Y$ is continuous, then $f(X)$ is compact.

Proof. Suppose that we have an open cover $\{U_i\}_{i\in I}$ of $f(X)$, i.e., a family of open sets $U_i \subseteq Y$ such that $f(X) \subseteq \bigcup_{i\in I} U_i$. Then the sets $f^{-1}(U_i) \subseteq X$ form an open cover of X — they are open because f is continuous, and they cover X because $\bigcup_{i\in I} f^{-1}(U_i) = f^{-1}(\bigcup_{i\in I} U_i) = X$.

Then since X is compact, we can find a finite subcover of our open cover of X. Taking the corresponding sets in our open cover of Y (i.e., if our finite subcover in X contains $f^{-1}(U_i)$ then we place U_i in our finite subcover of Y) gives a finite subcover of Y as well. $\square$

§7.3.1 A Criterion for Homeomorphisms

Using these properties, we can obtain a condition that guarantees a continuous and bijective map automatically is a homeomorphism. This is very practically useful — in many examples we saw earlier, proving our functions were continuous and bijective was easy, but proving their inverses were continuous was difficult; and now under the right conditions, we can simply get this for free.

Corollary 7.13

If X is compact and Y is Hausdorff, then any continuous bijective map $f\colon X \to Y$ is a homeomorphism.

Proof. In order to show that f^{-1} is continuous, it suffices to show that if $C \subseteq Y$ is closed, then $f(C) \subseteq Y$ is closed as well.

First, since X is compact and $C \subseteq X$ is closed, we know C is compact. This means $f(C)$ is compact as well. But since Y is Hausdorff, this implies $f(C)$ is closed. $\square$

§7.3.2 Compactness of Products

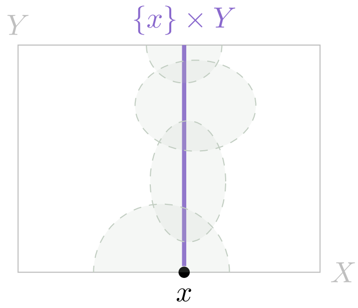
Theorem 7.14

Given spaces $X_1, \ldots, X_n$, the product $X_1 \times \cdots \times X_n$ is compact if and only if each of the spaces $X_1, \ldots, X_n$ is compact.

Proof. For the forwards direction, for each $i \in \{1, \ldots, n\}$ the projection map $\pi_i\colon X_1 \times \cdots \times X_n \to X_i$ is continuous and has image X_i, so if $X_1 \times \cdots \times X_n$ is compact then X_i must be compact as well.

For the reverse direction, by induction it suffices to prove that if X and Y are compact, then $X \times Y$ is compact as well. Suppose we have an open cover $\mathcal{W} = \{W_i\}_{i\in I}$ of $X \times Y$; we wish to find a finite subcover.

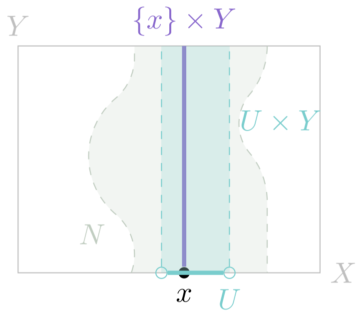
First fix a point $x \in X$. Then since $\{W_i\}_{i\in I}$ covers all of $X \times Y$, it in particular covers $\{x\} \times Y$. But $\{x\} \times Y$ is compact (as it is homeomorphic to Y), so we can find a finite subcover $\mathcal{W}(x) \subseteq \mathcal{W}$ of $\{x\} \times Y$ (i.e., we can find finitely many indices $i_1, \ldots, i_n$ such that $\{x\} \times Y \subseteq W_{i_1} \cup \cdots \cup W_{i_n}$).



Let $N(x)$ be the union of the open sets in $\mathcal{W}(x)$, and note that $N(x)$ is open. We'll attempt to show that we can choose finitely many points $x \in X$ so that the union of their corresponding sets $N(x)$ covers $X \times Y$ — this suffices because each set $N(x)$ is the union of finitely many sets in our original open cover $\mathcal{W}$. In order to do so, we need the following lemma.

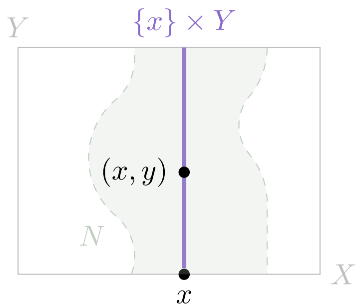
Lemma 7.15 (Tube Lemma)

Let X be any space and let Y be compact, and fix $x \in X$. If N is any open subset of $X \times Y$ which contains $\{x\} \times Y$, then there exists an open set $U \subseteq X$ containing x such that N contains $U \times Y$.

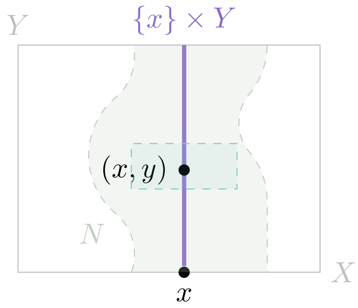


Intuitively, this lemma states that if we have an open set in $X \times Y$ which contains a slice (i.e., a set $\{x\} \times Y$), then it also contains a slight thickening of that slice.

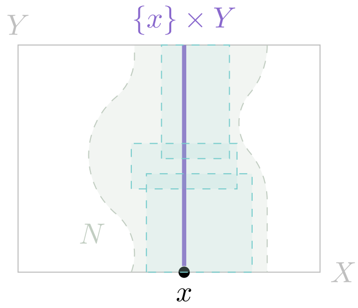
Proof. First fix a point $y \in Y$, so that N contains (x, y).



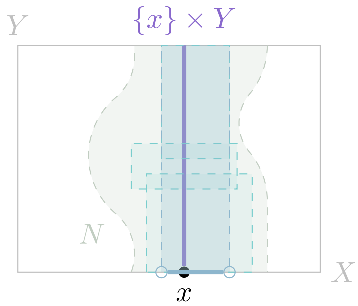
The product topology on $X \times Y$ has a basis given by the sets $U \times V$ for open subsets $U \subseteq X$ and $V \subseteq Y$. Since N is open, it can be written as the union of basis elements. Since (x, y) is in N it must be in one of the basis elements present in this union, so we can find open subsets $U_y \subseteq X$ and $V_y \subseteq Y$ such that $(x, y) \in U_y \times V_y \subseteq N$.



Now the collection $\mathcal{V} = \{V_y \mid y \in Y\}$ forms an open cover of Y (since V_y contains y for each $y \in Y$). So since Y is compact, we can find a finite subcover — i.e., we can find finitely many points $y_1, \ldots, y_n$ such that $Y = V_{y_1} \cup \cdots \cup V_{y_n}$.



Now define $U = U_{y_1} \cap \cdots \cap U_{y_n}$. We will show this choice of U has the claimed properties.



First, U contains x because each of the sets U_y contains x, and U is open because it is an intersection of finitely many open sets. Finally, we have $U \times Y = (U \times V_{y_1}) \cup \cdots \cup (U \times V_{y_n})$, and for each i we have $U \times V_{y_i} \subseteq U_{y_i} \times V_{y_i} \subseteq N$, so $U \times Y \subseteq N$. So U has all the necessary properties, and we're done. $\square$

Now to finish our proof that the product of two compact spaces is compact, recall that for each point $x \in X$, we have defined an open set $N(x)$ which contains $\{x\} \times Y$ (by taking a finite subcover of $\{x\} \times Y$ in our original cover), and we wish to find finitely many points $x_1, \ldots, x_m$ such that the sets $N(x_i)$ cover $X \times Y$. By the above lemma, for each point $x \in X$, we can find an open set $U_x \subseteq X$ containing x such that $U_x \times Y \subseteq N(x)$. Then the collection $\mathcal{U} = \{U_x \mid x \in X\}$ forms an open cover of X, so since X is compact, we can find a finite subcover, consisting of the finitely many sets $U_{x_1}, \ldots, U_{x_m}$. Then the sets $U_{x_1} \times Y, \ldots, U_{x_m} \times Y$ form a finite cover of $X \times Y$, and since $U_x \times Y \subseteq N(x)$ for each $x \in X$, the sets $N(x_1), \ldots, N(x_m)$ form a finite cover of $X \times Y$ as well. $\square$

Remark 7.16. This theorem is still true for *infinite* products as well — the product of any collection of compact spaces, with the product topology, is compact. This is known as Tychonoff's theorem, and the proof is much harder.

However, it is *not* true for infinite products with the *box* topology — for example, if we place the discrete topology on $\{0, 1\}$ (which is finite and therefore compact), then the box topology on $\{0, 1\}^{\mathbb{N}}$ is also the discrete topology; but $\{0, 1\}^{\mathbb{N}}$ is not finite, so it cannot be compact.

§7.4 Compactness in $\mathbb{R}^n$

We'll now prove the statement mentioned earlier that closed intervals are compact, and use it to obtain a characterization of compactness in $\mathbb{R}^n$ we may have seen in analysis — that a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

§7.4.1 Compactness of Closed Intervals

Theorem 7.17

The closed interval $[a, b]$ is compact.

Proof. Suppose we have an open cover of $[a, b]$, i.e., a collection $\mathcal{U} = \{U_i\}_{i\in I}$ of open subsets of $\mathbb{R}$ such that $[a, b] \subseteq \bigcup_{i\in I} U_i$. We want to find a finite open subcover.

Consider the set
$$S = \{x \in (a, b] \mid [a, x] \text{ has a finite subcover of } \mathcal{U}\}.$$

Our proof will consist of three steps:

(1) We'll show that S is nonempty and bounded, and therefore we can define $m = \sup S$.

(2) We'll show that m must be in S.

(3) Finally, we'll show that $m = b$.

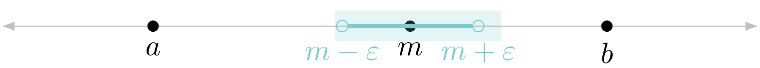
Together, these steps will imply that $b \in S$, and therefore $[a, b]$ admits a finite subcover, as desired.

For (1), it is clear that S is bounded, as $S \subseteq (a, b]$. In order to see that S is nonempty, there must exist some set $U \in \mathcal{U}$ which contains a, and since U is open there must exist some ε such that $(a - \varepsilon, a + \varepsilon) \subseteq U$.

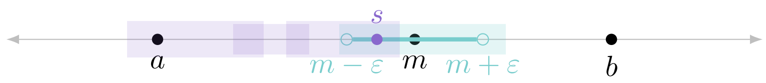
Then in particular $[a, a + \frac{1}{2}\varepsilon] \subseteq U$, so $a + \frac{1}{2}\varepsilon$ must be in S (as this set U alone forms a finite subcover of the interval $[a, a + \frac{1}{2}\varepsilon]$).

So S is bounded and nonempty, which lets us define $m = \sup S$. First note that we must have $m \in (a, b]$ — we must have $m > a$ because $m \geq s$ for all $s \in S$ and $S \subseteq (a, b]$ is nonempty, while we must have $m \leq b$ because b is an upper bound for S (and m is the *least* upper bound).

Then for (2), we need to show that $[a, m]$ admits a finite subcover. First, since $m \in [a, b]$, we can find an open set $U \in \mathcal{U}$ which contains m, and therefore which contains $(m - \varepsilon, m + \varepsilon)$ for some $\varepsilon > 0$.



Now since m is the least upper bound of S, there must exist $s \in S$ with $s > m - \varepsilon$ — otherwise $m - \varepsilon$ would be an upper bound of S as well. By the definition of S, this means we can find a finite subcover of $[a, s]$.



Then adding U to this finite subcover of $[a, s]$ produces a finite subcover of $[a, m] = [a, s] \cup (m - \varepsilon, m]$. So we've found a finite subcover of $[a, m]$, which means $m \in S$.

Finally for (3), assume for contradiction that $m < b$. Then similarly to in the previous step, we can find a set $U \in \mathcal{U}$ and some $\varepsilon > 0$ such that $(m - \varepsilon, m + \varepsilon) \subseteq U$, and such that $m + \varepsilon < b$ (by shrinking ε if necessary). Now consider the interval $[a, m + \frac{1}{2}\varepsilon]$. We can find a finite subcover of $[a, m]$ since $m \in S$, and since $[m, m + \frac{1}{2}\varepsilon] \subseteq U$, then adding U to this finite subcover produces a finite subcover of $[a, m + \frac{1}{2}\varepsilon]$. This implies $m + \frac{1}{2}\varepsilon \in S$, contradicting the fact that m is an upper bound for S.

So we must have $m = b$, and therefore $b \in S$; so we are done. $\square$

§7.4.2 Extreme Value Theorem

A few lectures ago, we saw that the intermediate value theorem from analysis generalizes to continuous maps from all *connected* spaces. Today we'll see that another theorem from analysis, the *extreme value theorem* (which states that any continuous function on a closed interval has a minimum and maximum) generalizes to all *compact* spaces.

Theorem 7.18 (Extreme Value Theorem)

If X is compact and $f\colon X \to \mathbb{R}$ is continuous, then there exist $x_{\min}$ and $x_{\max}$ in X such that for all $x \in X$ we have $f(x_{\min}) \le f(x) \le f(x_{\max})$.

Proof. We wish to show that $f(X)$ has a minimum and maximum (then taking $x_{\min}$ and $x_{\max}$ to be the values of x producing the minimum and maximum gives the desired statement). We will prove that it has a maximum, i.e., that there exists $m \in f(X)$ such that $y \le m$ for all $y \in f(X)$; the proof that it has a minimum is identical.

First note that since X is compact and f is continuous, $f(X)$ is compact. Now assume for contradiction that $f(X)$ does not have a maximum, so for every $m \in f(X)$ we can find a point $y \in f(X)$ with $y > m$. This means $f(X) \subseteq \bigcup_{y \in f(X)} (-\infty, y)$ — for each $m \in f(X)$ there exists some y such that $m \in (-\infty, y)$, and therefore this union covers each $m \in f(X)$.

Since $f(X)$ is compact and $\{(-\infty, y) \mid y \in f(X)\}$ forms an open cover, we can find a finite subcover — so there exist finitely many points $y_1, \ldots, y_n$ in $f(X)$ such that $f(X) \subseteq (-\infty, y_1) \cup \cdots \cup (-\infty, y_n)$. Without loss of generality assume y_n is the largest of $y_1, \ldots, y_n$; then this union is simply $(-\infty, y_n)$. But we cannot have $f(X) \subseteq (-\infty, y_n)$ because y_n is in $f(X)$ by definition; so this is a contradiction. $\square$

§7.4.3 Compactness in $\mathbb{R}^n$

We'll now prove the characterization of compact sets in $\mathbb{R}^n$ (with the standard topology) that we may have seen in analysis.

Definition 7.19. A set $A \subseteq \mathbb{R}^n$ is *bounded* if there exists $r > 0$ such that $A \subseteq B(0, r)$.

Note that boundedness depends on which metric we use; here we use the standard Euclidean metric d_2 (though we could have used any metric d_p — equivalent metrics produce the same notion of boundedness).

Theorem 7.20

A set $A \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proof. First we'll prove the forwards direction — suppose that $A \subseteq \mathbb{R}^n$ is compact. First, since $\mathbb{R}^n$ is Hausdorff, this implies A must be closed in $\mathbb{R}^n$. To show that A must be bounded, consider the collection $\{B(0, r) \mid r > 0\}$ (consisting of all balls of positive radius centered at the origin). These balls form an open cover of $\mathbb{R}^n$, and therefore of A. Since A is compact, we can find a finite subcover — so we can find finitely many radii $r_1 < \cdots < r_k$ such that $A \subseteq B(0, r_1) \cup \cdots \cup B(0, r_k) = B(0, r_k)$, and therefore A is bounded.

Now we'll prove the reverse direction — that if A is closed and bounded, then it is compact. First, fix $r > 0$ such that $A \subseteq B(0, r)$; then since $B(0, r) \subseteq [-r, r]^n$ we have $A \subseteq [-r, r]^n$ as well. We've seen that the closed interval $[-r, r]$ is compact and that products of compact spaces are compact, and combining these gives that $[-r, r]^n$ is compact. Then $A \subseteq [-r, r]^n$ is a closed subset of a compact space, so it is compact as well. $\square$

Example 7.21

The space $\mathbb{S}^n$ is compact for each n, as $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ is closed and bounded.

In particular, note that this means $\mathbb{S}^n$ is not homeomorphic to $\mathbb{R}^m$ for any n and m, since $\mathbb{S}^n$ is compact and $\mathbb{R}^m$ is not.

§7.5 Sequential Compactness

As with many other properties (e.g. closedness and continuity), there exists a sequential version of the definition of compactness; the two definitions will be equivalent in $\mathbb{R}^n$ (and in fact in any metric space) but not in full generality.

Definition 7.22. A space X is *sequentially compact* if every sequence in X has a convergent subsequence.

Given a sequence $(x_n)_{n \in \mathbb{N}}$, a *subsequence* is a sequence $(x_{n_k})_{k \in \mathbb{N}}$ for indices $1 \leq n_1 < n_2 < \cdots$ — intuitively, a subsequence of (x_n) is a sequence consisting of only some of the terms of (x_n), in their original order.

Example 7.23

Any space X with the trivial topology is sequentially compact, as *every* sequence is convergent.

Example 7.24

A space X with the discrete topology is sequentially compact if and only if it is finite.

Proof. We've seen that in the discrete topology, a sequence is convergent if and only if it is eventually constant. If X is finite, then any sequence (x_n) in X will have some point x which appears infinitely many times; then taking all occurrences of x gives a convergent subsequence of (x_n). Conversely, if X is infinite, then we can construct a sequence (x_n) all of whose terms are distinct (by first choosing x_1, then choosing x_2 different from x_1, then choosing x_3 different from x_1 and x_2, and so on). Then any subsequence of (x_n) will also have the property that all its terms are distinct, so it cannot converge. $\square$

Example 7.25

In $\mathbb{R}$ with the standard topology:

- $\mathbb{R}$ is not sequentially compact, as the sequence $x_n = n$ does not have a convergent subsequence.
- $(0, 1)$ is not sequentially compact, as the sequence $x_n = \frac{1}{n}$ does not have a convergent subsequence.
- Closed intervals $[a, b]$ are sequentially compact — this is known as the Bolzano–Weirstrass theorem.

(The proof of the Bolzano–Weirstrass theorem is covered in an analysis class — the idea is to repeatedly split our interval in half, and consider one half which contains infinitely many points.)

§7.5.1 Sequential Compactness vs. Compactness

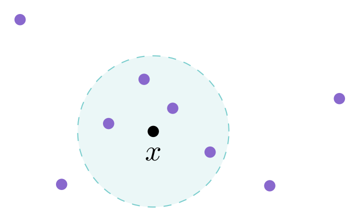
In all the examples we've seen, sequential compactness is the same as compactness. This isn't true in general, but we'll now see some conditions under which it is.

Theorem 7.26

If X is compact and first countable, then X is sequentially compact.

The proof uses the notion of *accumulation points* — we'll show that every sequence has an accumulation point, and then that we can find a subsequence converging to this accumulation point.

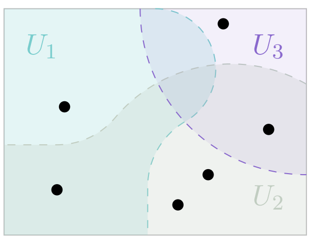
Definition 7.27. Given a sequence (x_n) in X, we say that a point $x \in X$ is an *accumulation point* of (x_n) if for every open set U containing x, there are infinitely many n for which $x_n \in U$.



Lemma 7.28

If X is compact, then every sequence (x_n) in X has an accumulation point.

Proof. Assume that (x_n) does not have an accumulation point. Then for every point $x \in X$, since x is not an accumulation point of (x_n), we can find an open set U_x containing x such that there are only finitely many n with $x_n \in U_x$. Then $\{U_x \mid x \in X\}$ is an open cover of X, so since X is compact, we can find a finite subcover — i.e., we can find finitely many sets U_x which together cover X.



Then since our finitely many sets U_x cover X, each point x_n must be in at least one of these sets. But there are finitely many sets U_x and each contains only finitely many points x_n; this is a contradiction, as there are infinitely many points x_n in total. □

Lemma 7.29

Let X be a first countable space, and suppose that x is an accumulation point of a sequence (x_n) in X. Then there exists a subsequence of (x_n) which converges to x.

Proof. Since X is first countable, by Lemma 5.29 we can find a chain $U_1 \supseteq U_2 \supseteq \cdots$ of open sets containing x such that any sequence (y_k) with $y_k \in U_k$ for each $k \in \mathbb{N}$ converges to x. Now since x is an accumulation point of (x_n), for each k there are infinitely many n for which $x_n \in U_k$. So we can choose n_1 such that $x_{n_1} \in U_1$, then $n_2 > n_1$ such that $x_{n_2} \in U_2$, then $n_3 > n_2$ such that $x_{n_3} \in U_3$, and so on. This produces a subsequence (x_{n_k}) which has the property that $x_{n_k} \in U_k$ for each k, and therefore which converges to x. □

Combining Lemmas 7.28 and 7.29 immediately implies Theorem 7.26.

Now we'll prove the converse for *metric spaces.*

Theorem 7.30

If X is a sequentially compact *metric space*, then X is compact.

Note that all metric spaces are first countable, and therefore Theorem 7.26 applies to them — so together, the two theorems imply that compactness and sequential compactness are equivalent in any metric space.

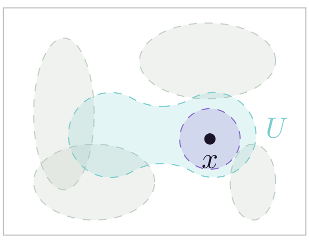
Proof. Suppose that we are given an open cover $\mathcal{U}$ of X. We will perform a process that attempts to find a finite subcover — we will construct points $x_1, x_2, \ldots$ in X and open sets $U_1, U_2, \ldots$ in $\mathcal{U}$ (where we construct x_n and U_n on the nth step for each $n \in \mathbb{N}$) in such a way that if the process terminates after n steps, then the sets $U_1, \ldots, U_n$ we have constructed form a finite subcover. If the process *doesn't* terminate after finitely many steps, then we will have produced a sequence $(x_n)_{n\in\mathbb{N}}$, and we will use the fact that this sequence must have a convergent subsequence to obtain a contradiction.

We define our process in the following way. For each $n \geq 1$, on the nth step, suppose that we have already constructed points $x_1, \ldots, x_{n-1}$ and open sets $U_1, \ldots, U_{n-1}$. Then:

- If $U_1 \cup \cdots \cup U_{n-1} = X$, then terminate the process — then $U_1, \ldots, U_{n-1}$ form a finite subcover of $\mathcal{U}$, so we are done.
- Otherwise, we'll intuitively choose x_n to be a point not covered by $U_1 \cup \cdots \cup U_{n-1}$ and U_n to be a reasonably large set containing x_n. To formalize this, consider the set

$$S_n = \{\varepsilon \in (0, 1] \mid \text{exists } x \in X \setminus (U_1 \cup \cdots \cup U_{n-1}) \text{ and } U \in \mathcal{U} \text{ with } B(x, \varepsilon) \subseteq U\}.$$

  Note that S_n is nonempty — there must exist some point $x \in X \setminus (U_1 \cup \cdots \cup U_{n-1})$, or else the process would have terminated. For each such point, there must exist a set $U \in \mathcal{U}$ containing x (since $\mathcal{U}$ covers X), and since U is open there must exist $\varepsilon \in (0, 1]$ with $B(x, \varepsilon) \subseteq U$.



Meanwhile, clearly S_n is bounded, as $S_n \subseteq (0, 1]$. So we can define $c_n = \sup S_n$. Then there must exist some element of S_n which is at least $\frac{1}{2}c_n$ (or else $\frac{1}{2}c_n$ would be a smaller upper bound for S_n); let ε_n be any such element.

Define x_n and U_n to be the sets x and U corresponding to our chosen ε_n — so x_n is a point in $X \setminus (U_1 \cup \cdots \cup U_{n-1})$ and U_n a set in $\mathcal{U}$ such that $B(x_n, \varepsilon_n) \subseteq U_n$.

We wish to show that this process must terminate; assume not. Then the process produces a sequence $(x_n)_{n\in\mathbb{N}}$, and since X is compact, this sequence must have a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges to some point $x \in X$.

First, we claim that x cannot be in any of our chosen sets U_m — assume for contradiction that x is in U_m for some m. Since in each step x_n is chosen from $X \setminus (U_1 \cup \cdots \cup U_{n-1})$, we cannot have $x_n \in U_m$ for any $m \geq n$. But U_m is an open set containing x, so this contradicts the fact that (x_n) has a subsequence (x_{n_k}) converging to x (which requires that $x_{n_k} \in U_m$ for all large k).

Now since $\mathcal{U}$ covers X there must exist some $U \in \mathcal{U}$ containing x, and since U is open there must exist $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. Then at each step we must have $\varepsilon \in S_n$ (since x is not in any of the sets U_m, and is therefore always one of the points considered when defining S_n), and therefore $\varepsilon_n \geq \frac{1}{2}\varepsilon$.

But since our subsequence (x_{n_k}) converges to x, for all large k we must have $d(x_{n_k}, x) < \frac{1}{2}\varepsilon$, and therefore x must be contained in $B(x_{n_k}, \varepsilon_{n_k})$. But by construction $B(x_{n_k}, \varepsilon_{n_k}) \subseteq U_{n_k}$, so then x must be contained in U_{n_k}, contradicting the fact that x is not in U_m for any m. □

§8 The Fundamental Group

So far, we've discussed point-set topology. We'll now turn to algebraic topology and associate to each topological space X a *group* $\pi_1(X)$ which stores certain data about X. As some motivation, with the tools we've seen so far (connectedness and compactness) we can tell *some* spaces apart, but there are many spaces we *can't* tell apart — for example, we can't yet prove that the torus $\mathbb{S}^1 \times \mathbb{S}^1$ and the sphere $\mathbb{S}^2$ are not homeomorphic (which does *seems* true). It'll turn out that we can do this using the fundamental group — we'll see that if two spaces are homeomorphic then their fundamental groups are isomorphic. We'll also see that $\pi_1(\mathbb{S}^1 \times \mathbb{S}^1) = \mathbb{Z}^2$ and $\pi_1(\mathbb{S}^2) = \{1\}$, so the two spaces cannot be homeomorphic.

§8.1 Definition of the Fundamental Group

Informally, the fundamental group will consist of loops in X up to deformation. We'll now see how to make this intuition into a precise definition.

§8.1.1 Loops and Path Composition

Definition 8.1. A *loop* in X is a path $\gamma\colon [0, 1] \to X$ such that $\gamma(0) = \gamma(1)$. We call the point $\gamma(0) = \gamma(1)$ the *basepoint* of the loop γ.

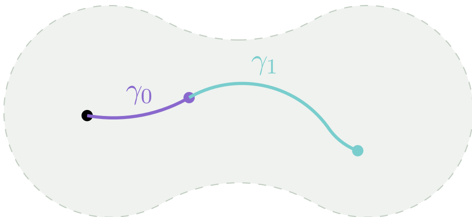
Intuitively, a loop is a path that returns to where it starts, as we would expect.

Since we're trying to turn these loops into a group, we'll need a group law (i.e., a way to compose two loops). This will essentially be given by concatenating the two loops.

Definition 8.2. Given two paths $\gamma_0\colon [0,1] \to X$ and $\gamma_1\colon [0,1] \to X$ with $\gamma_0(1) = \gamma_1(0)$, their *composition*, denoted $\gamma_0 \cdot \gamma_1$, is the path defined as

$$\gamma_0 \cdot \gamma_1(t) = \begin{cases} \gamma_0(2t) & \text{if } t \in [0, \frac{1}{2}] \\ \gamma_1(2t-1) & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

Note that the definition of composition applies to paths in general, not just loops. We proved in Proposition 6.21 that the composition $\gamma_0 \cdot \gamma_1$ is indeed a path (i.e., that the piecewise function defined in this way is continuous).



§8.1.2 Homotopy of Paths

We mentioned earlier that we want to define the fundamental group to consist of loops 'up to deformation.' We'll use the notion of *homotopy* to make precise what it means for two loops to be the same up to deformation.

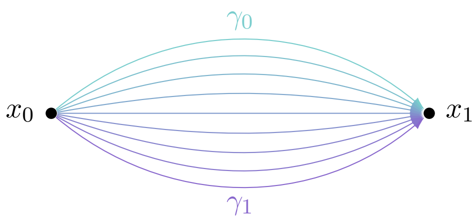
Definition 8.3. Let X be a space, let x_0 and x_1 be points in X, and let γ_0 and γ_1 be paths from x_0 to x_1 (i.e., each γ_i is a continuous map $[0,1] \to X$ with $\gamma_i(0) = x_0$ and $\gamma_i(1) = x_1$). Then a *homotopy* between γ_0 and γ_1 is a continuous function $F\colon [0,1] \times [0,1] \to X$ such that:

- $F(0,t) = x_0$ and $F(1,t) = x_1$ for all $t \in [0,1]$.
- $F(x,0) = \gamma_0(x)$ and $F(x,1) = \gamma_1(x)$ for all $x \in [0,1]$.

We say γ_0 and γ_1 are *homotopic* (denoted $\gamma_0 \simeq \gamma_1$) if there exists a homotopy between them.

Notation 8.4. We'll often denote a homotopy by $f_t(x)$ instead of $F(x,t)$.

Intuitively, a homotopy between γ_0 and γ_1 is a collection of paths f_t from x_0 to x_1 (with one path for each $t \in [0,1]$) which interpolate continuously between γ_0 and γ_1.



Example 8.5

In $\mathbb{R}$ (or more generally, in any convex subset of $\mathbb{R}^n$), any two paths γ_0 and γ_1 with the same starting and ending points are homotopic, with the homotopy given by

$$f_t(x) = (1-t)\gamma_0(x) + t\gamma_1(x).$$

The following observation will frequently be useful.

Fact 8.6 — Let γ_0 and γ_1 be two paths in a space X, and let $\varphi\colon X \to Y$ be a continuous map. If $\gamma_0 \simeq \gamma_1$ via a homotopy f_t, then $\varphi \circ \gamma_0 \simeq \varphi \circ \gamma_1$ via the homotopy $\varphi \circ f_t$.

(Note that $\cdot$ denotes path composition, while $\circ$ denotes function composition.) We've seen earlier (when discussing path-connectedness) that if $\gamma\colon [0,1] \to X$ is a path from x_0 to x_1 in X and $\varphi\colon X \to Y$ a continuous map, then $\varphi \circ \gamma$ is a path from $\varphi(x_0)$ to $\varphi(x_1)$ in Y — this follows from the fact that a composition of continuous functions is continuous. Here a similar argument can be used to show that the map $(x,t) \mapsto \varphi \circ f_t(x)$ is continuous, and it is straightforward to check that the other conditions on a homotopy are satisfied.

§8.1.3 The Fundamental Group

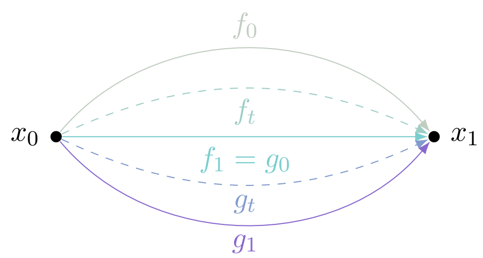
We'll eventually define the fundamental group of X to consist of all loops in X (with a fixed basepoint) up to homotopy; for this to make sense, we need to check that homotopy is an *equivalence relation*.

Lemma 8.7

Let X be a space, and fix two points x_0 and x_1 in X. Then homotopy of paths is an equivalence relation on the set of paths from x_0 to x_1.

Proof. We'll check that the three properties of an equivalence relation all hold.

- *Reflexivity* (i.e., $f \simeq f$) — we can simply take the homotopy $f_t = f$ for all $t \in [0,1]$.
- *Symmetry* (i.e., if $f_0 \simeq f_1$ then $f_1 \simeq f_0$) — if f_t is a homotopy from f_0 to f_1, then f_{1-t} is a homotopy from f_1 to f_0. (Intuitively, given a way to continuously deform f_0 into f_1, we can run this deformation backwards to continuously deform f_1 into f_0.)
- *Transitivity* (i.e., if $f_0 \simeq f_1$ and $f_1 \simeq f_2$, then $f_0 \simeq f_2$) — for notational convenience, let $g_0 = f_1$ and $g_1 = f_2$, and suppose that $f_0 \simeq f_1$ via the homotopy f_t and $g_0 \simeq g_1$ by g_t. Then we claim $f_0 \simeq g_1$ via the homotopy
$$h_t(x) = \begin{cases} f_{2t}(x) & \text{if } t \in [0, \frac{1}{2}] \\ g_{2t-1}(x) & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$
Intuitively, h_t first deforms our path f_0 into $f_1 = g_0$, then deforms this path into g_1.



Note that $(x,t) \mapsto h_t(x)$ is continuous by Lemma 6.22 (which states that a function which is continuous on two closed subsets which together cover our entire space must be continuous on the entire space) applied to the closed subsets $[0,1] \times [0, \frac{1}{2}]$ and $[0,1] \times [\frac{1}{2}, 1]$ of $[0,1] \times [0,1]$, and is therefore a valid homotopy. $\square$

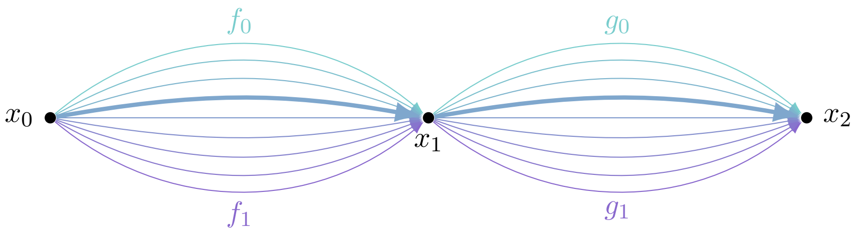
This allows us to define the *set* $\pi_1(X, x_0)$ (for a fixed basepoint $x_0 \in X$) as the set of all loops based at x_0, quotiented by the equivalence relation $\simeq$ (i.e., two loops are equivalent if there exists a homotopy between them). We'll soon define a group law on this set given by path composition; for this to make sense (i.e., to be well-defined), we need the following proposition.

Proposition 8.8

Let f_0, f_1, g_0, and g_1 be paths in X such that $f_i(1) = g_i(0)$ for each $i \in \{0, 1\}$, and $f_0 \simeq f_1$ and $g_0 \simeq g_1$. Then $f_0 \cdot g_0 \simeq f_1 \cdot g_1$.

Proof. Suppose that $f_0 \simeq f_1$ via a homotopy f_t, and $g_0 \simeq g_1$ via a homotopy g_t. Then we claim $f_0 \cdot g_0 \simeq f_1 \cdot g_1$ via the homotopy

$$h_t = f_t \cdot g_t = \begin{cases} f_t(2x) & \text{if } x \in [0, \dfrac{1}{2}] \\ g_t(2x-1) & \text{if } x \in [\dfrac{1}{2}, 1]. \end{cases}$$



The continuity of the map $(x,t) \mapsto h_t(x)$ again follows from Lemma 6.22, as it is continuous on both $[0, \frac{1}{2}] \times [0,1]$ and $[\frac{1}{2}, 1] \times [0,1]$. So h_t is again a valid homotopy. $\square$

This means path composition gives a well-defined operation on *homotopy classes* of loops based at a fixed point x_0 — using $[f]$ to denote the class of a loop f, we can define $[f] \cdot [g]$ as $[f \cdot g]$, as the above proposition implies that $[f \cdot g]$ is independent of the choice of representatives of $[f]$ and $[g]$.

Definition 8.9. The *fundamental group* of a space X with basepoint x_0, denoted $\pi_1(X, x_0)$, is the set

$$\pi_1(X, x_0) = \{\text{loops based at } x_0\}/\simeq$$

with group operation $[f] \cdot [g] = [f \cdot g]$.

(Note that to define the fundamental group of a space we must first fix a basepoint; we'll say more about the dependence on the basepoint later.)

It is not obvious that $\pi_1(X, x_0)$ is a group; to prove this, we need to check the three group axioms — that composition is associative, there is an identity, and every element has an inverse. (We'll state these claims in slightly more generality, for paths instead of just loops.)

First we'll construct an identity element.

Notation 8.10. For any point $y \in X$, we use c_y to denote the constant loop at y — i.e., the path $c_y \colon [0,1] \to X$ which maps $x \mapsto y$ for all $x \in [0,1]$.

Lemma 8.11

Suppose that f is a path from x_0 to x_1. Then $c_{x_0} \cdot f \simeq f \simeq f \cdot c_{x_1}$.

Proof. We'll just show that $c_{x_0} \cdot f \simeq f$; the proof that $f \cdot c_{x_1} \simeq f$ is essentially identical.

Intuitively, $c_{x_0} \cdot f$ is the path that stands still at x_0 for the interval $[0, \frac{1}{2}]$ and then performs f; and we can continuously deform it into f alone by gradually decreasing the amount of time we spend at x_0 at the start. It's not hard to write down an explicit homotopy formalizing this, but we'll instead use a slightly more indirect argument that can be used to cleanly prove many similar statements.

Consider the function $\varphi \colon [0,1] \to [0,1]$ defined as

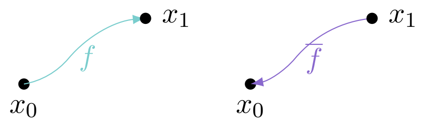
$$\varphi(x) = \begin{cases} 0 & \text{if } x \in [0, \frac{1}{2}] \\ 2x - 1 & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

Then we can think of $c_{x_0} \cdot f$ as the path $f \circ \varphi$, and f as the path $f \circ \mathrm{id}$. But φ and id can both be thought of as paths from 0 to 1 in $[0,1]$, and since $[0,1]$ is convex, we've seen in Example 8.5 that any two such paths are homotopic (one explicit homotopy between φ and id is $\varphi_t(x) = (1-t)\varphi(x) + tx$). So $\varphi \simeq \mathrm{id}$ in $[0,1]$, and since f is a continuous map $[0,1] \to X$, by Fact 8.6 we have $f \circ \varphi \simeq f \circ \mathrm{id}$ as well, i.e., $c_{x_0} \cdot f \simeq f$. □

In particular, this implies that $[c_{x_0}]$ is an identity element of $\pi_1(X, x_0)$.

Now we'll construct inverses.

Notation 8.12. For any path f from x_0 to x_1, we use $\overline{f}$ to denote the path from x_1 to x_0 defined as $\overline{f}(x) = f(1-x)$ for each $x \in [0,1]$.



Lemma 8.13

For any path f from x_0 to x_1, we have $f \cdot \overline{f} \simeq c_{x_0}$ and $\overline{f} \cdot f \simeq c_{x_1}$.

Proof. We'll only prove the first statement; the proof of the second is very similar. We'll use the same idea as in the previous proof — write $f \cdot \overline{f}$ as $f \circ \varphi_0$, where $\varphi_0 \colon [0,1] \to [0,1]$ is the map

$$\varphi_0(x) = \begin{cases} 2x & \text{if } x \in [0, \frac{1}{2}] \\ 2 - 2x & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

and write c_{x_0} as $f \circ \varphi_1$, where $\varphi_1 \colon [0,1] \to [0,1]$ is the map $x \mapsto 0$. Then both φ_0 and φ_1 are loops in $[0,1]$ based at 0, so since $[0,1]$ is convex there exists a homotopy φ_t between them; then by Fact 8.6 we have that $f \circ \varphi_t$ is a homotopy between $f \cdot \overline{f} = f \circ \varphi_0$ and $c_{x_0} = f \circ \varphi_1$. □

In particular, this implies that each $[f] \in \pi_1(X, x_0)$ has inverse $[\overline{f}]$.

Lemma 8.14

Let f, g, and h be paths such that $f(1) = g(0)$ and $g(1) = h(0)$. Then $(f \cdot g) \cdot h \simeq f \cdot (g \cdot h)$.

Proof. The main idea is that similarly to before, we can define a map $\varphi\colon [0,1] \to [0,1]$ such that $f \cdot (g \cdot h) = ((f \cdot g) \cdot h) \circ \varphi$ (since the two paths essentially trace the same route at different speeds), and use the fact that $\varphi \simeq \text{id}$ to conclude that $f \cdot (g \cdot h) \circ \text{id} \simeq ((f \cdot g) \cdot h) \circ \varphi$.

Explicitly, we have

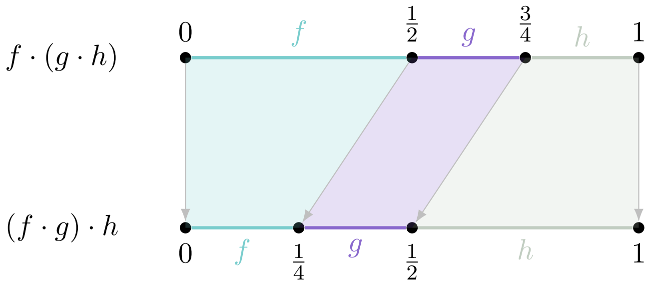
$$(f \cdot g) \cdot h = \begin{cases} f(4x) & \text{if } x \in [0, \frac{1}{4}] \\ g(4x-1) & \text{if } x \in [\frac{1}{4}, \frac{1}{2}] \\ h(2x-1) & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

and similarly we have

$$f \cdot (g \cdot h) = \begin{cases} f(2x) & \text{if } x \in [0, \frac{1}{2}] \\ g(4x-2) & \text{if } x \in [\frac{1}{2}, \frac{3}{4}] \\ h(4x-3) & \text{if } x \in [\frac{3}{4}, 1]. \end{cases}$$

This means we have $f \cdot (g \cdot h) = ((f \cdot g) \cdot h) \circ \varphi$, where $\varphi\colon [0,1] \to [0,1]$ is defined as

$$\varphi(x) = \begin{cases} \frac{1}{2}x & \text{if } x \in [0, \frac{1}{2}] \\ x - \frac{1}{4} & \text{if } x \in [\frac{1}{2}, \frac{3}{4}] \\ 2x-1 & \text{if } x \in [\frac{3}{4}, 1]. \end{cases}$$



Then φ is a path in $[0,1]$ from 0 to 1, and since all such paths are homotopic, as before we have $\varphi \simeq \text{id}$, and therefore $f \cdot (g \cdot h) = ((f \cdot g) \cdot h) \circ \varphi \simeq ((f \cdot g) \cdot h) \circ \text{id} = (f \cdot g) \cdot h$. $\square$

Combining these statements gives the following result.

Theorem 8.15

For any space X and basepoint $x_0 \in X$, the fundamental group $\pi_1(X, x_0)$ is a group.

§8.1.4 Dependence on the Basepoint

Note that our definition of the fundamental group requires us to choose a basepoint $x_0 \in X$. However, the choice of basepoint turns out to not be so important.

First, we'll see that when considering the fundamental group, it essentially suffices to consider only path-connected spaces.

Definition 8.16. For a space X and a point $x \in X$, the *path component* of x, denoted $C(x)$, is the set of points $y \in X$ such that there exists a path from x to y.

Note that the relation $x \sim y$ if there exists a path from x to y is an equivalence relation; we can think of the path-components of X as its equivalence classes. In particular, the different path-components form a partition of X.

Fact 8.17 — For any space X and basepoint $x_0 \in X$, we have $\pi_1(X, x_0) \cong \pi_1(C(x_0), x_0)$.

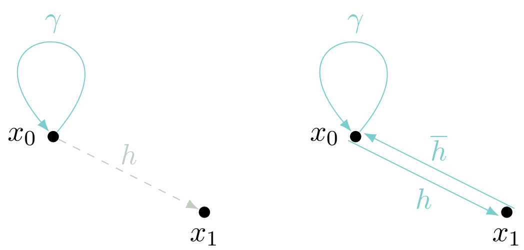
This is essentially because any loop in X based at x_0 must in fact be a loop in $C(x_0)$ (if a loop starting at x_0 passes through a point y, then there must exist a path from x_0 to y), and two such loops are homotopic in X if and only if they are homotopic in $C(x_0)$ (since all 'intermediate' loops in a homotopy in X must lie in $C(x_0)$ for the same reason).

This means when we're considering $\pi_1(X, x_0)$ we only need to focus on the path-component of x_0 — and if x_0 and x_1 are two points in different path-components, then $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$ don't interact with each other.

On the other hand, if x_0 and x_1 *are* in the same path-component, it turns out the corresponding groups are isomorphic.

Lemma 8.18

Let X be a space, and let x_0 and x_1 be points in X. Suppose there exists a path h from x_0 to x_1. Then the map $\beta_h \colon \pi_1(X, x_0) \to \pi_1(X, x_1)$ sending $[\gamma] \mapsto [\overline{h} \cdot \gamma \cdot h]$ is a group isomorphism.



Proof. First note that for any $[\gamma_1], [\gamma_2] \in \pi_1(X, x_0)$ we have

$$[(\overline{h} \cdot \gamma_1 \cdot h) \cdot (\overline{h} \cdot \gamma_2 \cdot h)] = [\overline{h} \cdot \gamma_1 \cdot (h \cdot \overline{h}) \cdot \gamma_2 \cdot h] = [\overline{h} \cdot (\gamma_1 \cdot \gamma_2) \cdot h],$$

which means that β_h is a group *homomorphism*. Meanwhile the map $\beta_{\overline{h}}\colon \pi_1(X, x_1) \to \pi_1(X, x_0)$ sending $[\gamma] \mapsto [h \cdot \gamma \cdot \overline{h}]$ (where γ is now a loop based at x_1) is its inverse, so it is in fact a group isomorphism. □

In particular, if X is path-connected, then the fundamental group $\pi_1(X, x_0)$ does not depend on the choice of basepoint (up to isomorphism).

Definition 8.19. A space X is *simply-connected* if X is path-connected and $\pi_1(X)$ is trivial.

§8.2 Some Examples

Now we'll calculate the fundamental group of a few spaces. First we'll see a very simple example.

Example 8.20

We have $\pi_1(\mathbb{R}^n) = \{1\}$ — any two loops in $\mathbb{R}^n$ (with fixed basepoint) are homotopic, so there is only one equivalence class of loops. For the same reason, for any convex subset $C \subseteq \mathbb{R}^n$ we have $\pi_1(C) = \{1\}$.

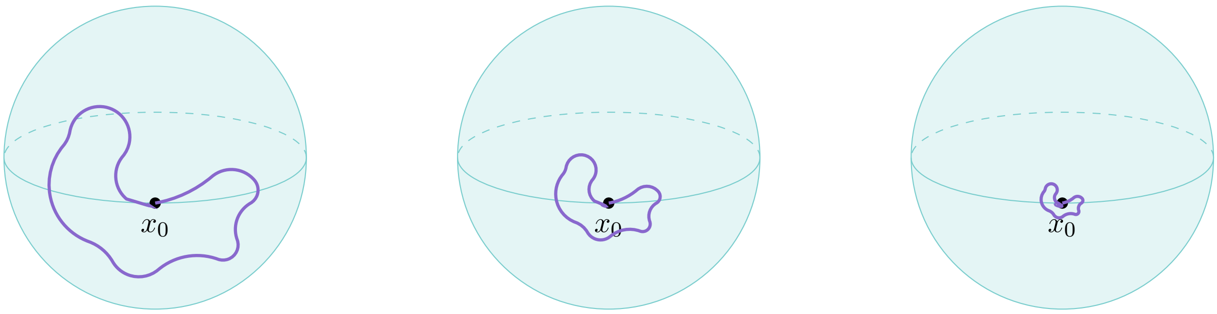
Now we'll see a few less trivial examples — we'll find the fundamental group of the sphere $\mathbb{S}^n$ for each n.

§8.2.1 Fundamental Group of a Sphere

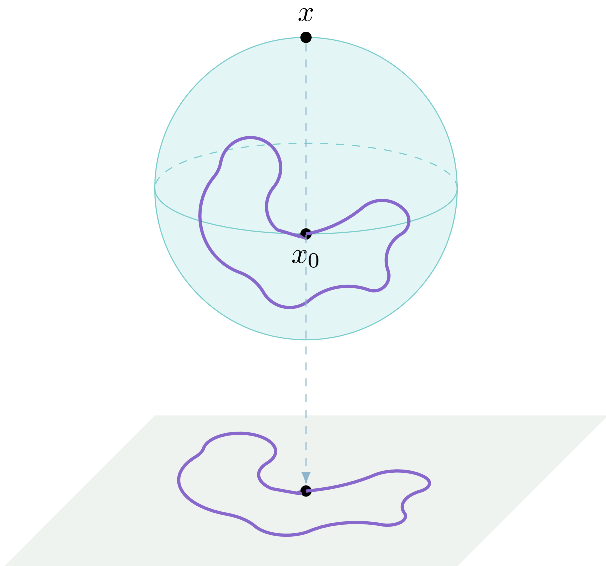
Theorem 8.21

For all $n \geq 2$, we have $\pi_1(\mathbb{S}^n) = \{1\}$.

In order to prove this, fix a basepoint x_0; we need to show that every loop $f\colon [0,1] \to \mathbb{S}^n$ is homotopic to the constant loop c_{x_0}. Intuitively, this makes sense as we can 'squish' any loop to a single point; however, the details are fairly difficult. (Note the assumption that $n \geq 2$ — this isn't true for the circle $\mathbb{S}^1$, where a loop winding around the circle cannot be squished down to a point.)



The first idea in the proof is that if there is some point x not in our path f, then the proof is easy — we can then think of f as a loop in $\mathbb{S}^n \setminus \{x\}$, and we've seen (in Example 2.14) that $\mathbb{S}^n \setminus \{x\}$ is homeomorphic to $\mathbb{R}^n$ via the stereographic projection at x.



But $\mathbb{R}^n$ is simply connected, so every loop in $\mathbb{R}^n$ is homotopic to the trivial loop. This means f is homotopic to c_{x_0}, as desired.

So in the rest of the proof, our goal will essentially be to show that any path f is homotopic to one which misses some point. (Intuitively, we can think of homotopy as 'wiggling' our path around; so we're trying to wiggle our path to ensure it misses a point.)

Define two open sets B_0 and B_1 by taking two opposite hemispheres and enlarging them a bit, so that we have $x_0 \in B_0 \cap B_1$, $\mathbb{S}^n = B_0 \cup B_1$, and $B_0 \cap B_1 \cong \mathbb{S}^1 \times (-\varepsilon, \varepsilon)$ for $\varepsilon > 0$.