Topology notes

mike miller eismeier

Contents

As you will see, example and theorem numbering continues sequentially through this document. That means you pretty quickly get to gigantic theorem numbers. *Do not try to remember theorems by the way I’ve numbered them here.* I always try to remember a theorem by what it says. This in mind, I will label every theorem/proposition/lemma below with a short name.

9/9: Review of metric spaces

Topology starts in real analysis. There, you first learn about the real line and its properties: you learn about how to define continuity with ϵs and δs; you learn about open and closed intervals in the real line; you learn about compact sets and connected sets.

The real line is lovely and a great place to do a lot of math, but it's just a starting place. If you want to think about objects in the plane or 3-dimensional space (and we've liked that for millenia), you need a new framework.

In Modern Analysis I you leaned the first iteration of this framework: *metric spaces*. Just like the real line, these have a notion of distance, and this is enough to build ideas of continuity.

Today, we'll review the notion of metric spaces and the corresponding notions of continuity and open sets.

It is important to have some intuition from metric spaces; after today, I am going to implicitly assume you picked some up from Modern Analysis I. When you approach a new problem in topology, it's often helpful to pretend everything is a metric space for your first pass at understanding. At first, topology is going to feel like it's all about taking ideas from analysis and attempting to abstract and clarify them (to make them qualitative instead of quantitative).

If your memory is hazy or you never picked up much of an intuition, I recommend going over Munkres' section on metric spaces andd having a PDF of "Metric spaces" by Micheál O'Searcoid onhand, which contains many examples.

Metric spaces and examples

Definition 1. *A* metric space *is a pair (X, d), where X is a set, and $d : X \times X \to [0, \infty)$ is a function ('the distance function', or 'the metric') satisfying the following properties.*

a) If $d(x, y) = 0$, then $x = y$.

b) For all $x, y \in X$ we have $d(x, y) = d(y, x)$.

c) For all $x, y, z \in X$ we have $d(x, z) \leqslant d(x, y) + d(y, z)$.

The first condition says that no two points are 'infinitesimally far apart from each other': if two points are different, there is some positive distance between them. The second condition says that distance starting from x and going to y is the same as the other way around. (I intuit this by thinking of a measuring stick: flip it around, it's still just as long.) The third condition is the *triangle inequality*. In the plane, it quite literally says that one leg of a triangle is not longer than the sum of the lengths of the other two. (I intuit this as: 'It is not faster to meander on the way from x to z.' Do you see the connection to the claim that a straight line is the shortest piecewise-linear path in the plane between two points?)

Definition 2. *Let (X, d) be a metric space. We say that the open r-ball around $x \in X$ is*

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

If the metric d is not clear from context we will write $B_r^d(x)$.

Notice the strict inequality. One may also define closed r-balls, but these are not nearly as important from our perspective.

Example 1. The first metric space you ever met was $(\mathbb{R}, d)$, where the distance is $d(x, y) = |x - y|$. When you were taught the idea of a limit in calculus, it used this distance.

Here, $B_r(x) = (x - r, x + r)$ is the interval of radius r around x.

As I mentioned before, metric spaces allow us to consider spaces like the Euclidean plane or Euclidean spaces more generally. However, there is no 'obvious' choice of metric. There are a number of options for well-behaved metrics on $\mathbb{R}^n$. I will write an infinite family below, but there are many, many more than you can imagine, even for the plane!

Example 2. Consider the pair $(\mathbb{R}^n, d_p)$, where $\mathbb{R}^n = \{x = (x_1, \cdots, x_n) \mid x_i \in \mathbb{R}\}$ is the Cartesian product of n copies of the real line, and

$$d_p(x, y) = (|x_1 - y_1|^p + \cdots + |x_n - y_n|^p)^{1/p}.$$

The function d_p gives a metric on $\mathbb{R}^n$. (This is not trivial, but it is not too hard to show, and doing so is a good exercise in the spirit of Modern Analysis I.) This is called the p-distance, and $d_p(x, 0) = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ is called the p-norm of the vector x.

For $p = \infty$, we write

$$d_\infty(x, y) = \max(|x_1 - y_1|, \cdots, |x_n - y_n|).$$

This too is a metric, and $\lim_{p\to\infty} d_p(x, y) = d_\infty(x, y)$.

The most well-known and important of these are the 1-distance

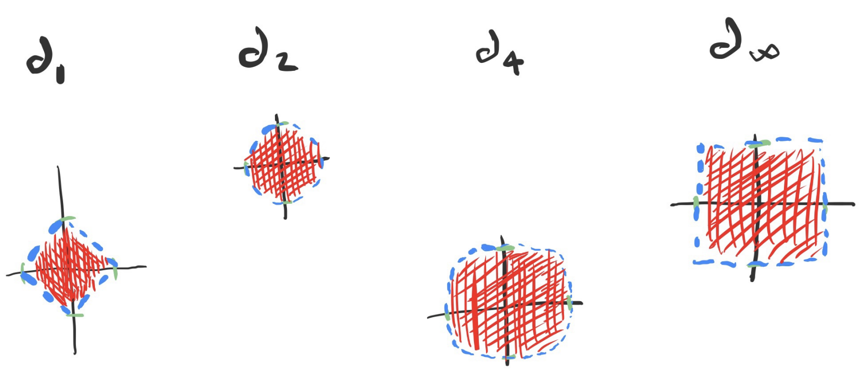
$$d_1(x, y) = |x_1 - y_1| + \cdots + |x_n - y_n|$$

the Euclidean distance

$$d_2(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

and the d_∞-distance described above.

To get a visual sense of what these distances 'look like', we can draw their unit balls $B_1^{d_p}(0)$:



There's something disquieting about the fact that in the land of metric spaces, there isn't an obvious special metric on $\mathbb{R}^2$, but many. You will show later this semester that in the land of topology, there *is* a canonical choice of topology on $\mathbb{R}^2$, and in fact that all of the above metrics give rise to the same topology.

Continuity of maps between metric spaces

In topology (at least at first), the notion we are most interested in is *continuity*, and the kind of functions we are most interested in are continuous functions. Metric spaces give the first home for a good definition of continuity.

Definition 3. *A function $f : X \to Y$ between metric spaces (X, d) and (Y, d') is* continuous at $x \in X$ *if, for every real number $\epsilon > 0$, there* exists *a $\delta = \delta(\epsilon) > 0$ so that*

$$d(x, y) < \delta \implies d'(f(x), f(y)) < \epsilon.$$

A function f is continuous if it is continuous at x for all $x \in X$.

Here the notation $\delta = \delta(\epsilon)$ mens that the number δ is chosen after ϵ is, and our choice of δ may depend (in fact, definitely does) on the previous choice of ϵ.

Intuitively, this says 'If we want $f(x)$ and $f(y)$ to be close, then this is guaranteed so long as x and y are really close.' Here 'close' is measured by being less than ϵ apart, while 'really close' is measured by being less than δ apart. How close 'really close' is depends on how close we want the outputs to be.

Example 3. The function $m : (\mathbb{R}^2, d_\infty) \to (\mathbb{R}, |\cdot|)$ is continuous (in the codomain we use the usual metric on $\mathbb{R}$).

To prove this, we need to show continuity at an arbitrary point (x_0, y_0) in the domain. I'm going to work backwards: I'm going to start by assuming that

$$d_\infty((x, y), (x_0, y_0)) < \delta,$$

and at the end of the argument we will see what δ has to be so that the output is less than ϵ.

By definition, this means that

$$\max(|x - x_0|, |y - y_0|) < \delta \quad \Longrightarrow \quad |x - x_0| < \delta \text{ and } |y - y_0| < \delta.$$

Now to say that $d(xy, x_0y_0) < \epsilon$ — what we want to show — means $|xy - x_0y_0| < \epsilon$, by definition. If we can bound $|xy - x_0y_0|$ from above by a nice function of δ, we can probably figure out what δ we need to make the output less than ϵ.

Now, note that

$$\begin{aligned} |xy - x_0y_0| &= |xy - x_0y + x_0y - x_0y_0| \leqslant |xy - x_0y| + |x_0y - x_0y_0| \\ &= |x - x_0||y| + |y - y_0||x_0| \leqslant \delta(|x_0| + |y|) \\ &\leqslant \delta(|x_0| + |y_0| + \delta) = \delta^2 + \delta|y_0| + \delta|x_0|. \end{aligned}$$

In the last line we used that $|y - y_0| < \delta$ to see that $|y| < |y_0| + \delta$.

We want to force this sum to be bounded by ϵ. My idea is that all three summands should be less than $\epsilon/3$, so that their sum is less than ϵ. So I choose

$$\delta < \min(\epsilon/3|x_0|, \epsilon/3|y_0|, \sqrt{\epsilon/3}).$$

Each of these terms are positive, so there exists $\delta > 0$ satisfying this for each $\epsilon > 0$.

We have thus shown that m is continuous at (x_0, y_0); because this point is arbitrary we've shown continuity at every point in the domain. It follows that m is continuous.

That was kind of a lot of work, but there's not much you can do about it. This is just what a proof of continuity between metric spaces looks like. Notice that δ depended on both ϵ and (x_0, y_0).

Open sets in metric spaces

The other fundamental concept at play in topology is *openness*. Intuitively, an open set is a subset $U \subset X$ so that U contains 'every point nearby to a given point in U' — where 'nearby' depends on the point and on U.

This intuitive phrasing is pretty easy to formalize for metric spaces.

Definition 4. *Let (X, d) be a metric space. A subset $U \subset X$ is called* open *if, for every $x \in U$, there exists an $r = r(x) > 0$ so that $B_r(x) \subset U$.*

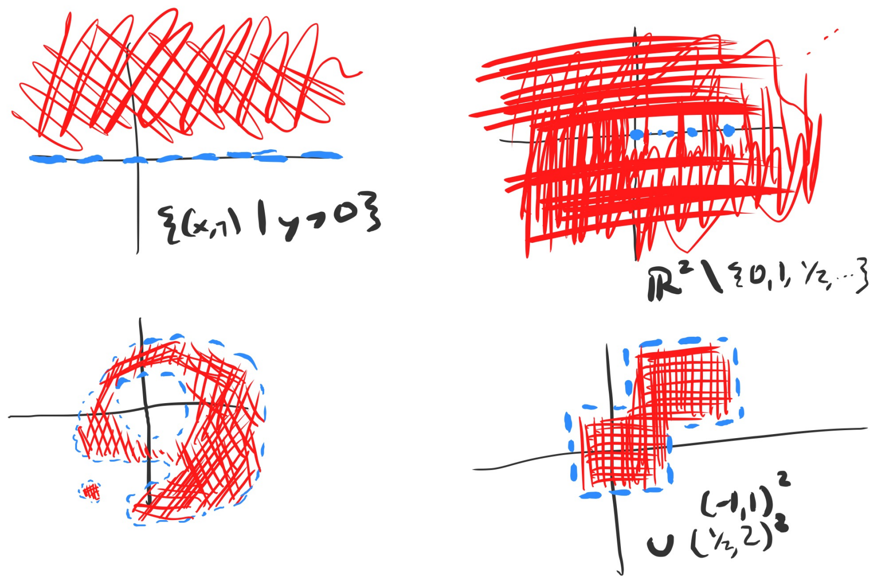
This strikes me as being a direct formal translation of the idea above. As you might hope from the name, the following theorem says that *open balls are open sets.*

Proposition 1 (Open balls are open sets)**.** *Let (X, d) be a metric space. For any $x \in X$ and any $r > 0$, the set $B_r(x)$ is an open set.*

Proof. If $y \in B_r(x)$, we need to show that there exists some $s > 0$ so that $B_s(y) \subset B_r(x)$. This means that for all z with $d(y, z) < s$, we also have $d(x, z) < r$. This smells like the triangle inequality to me. I choose $s = r - d(x, y) > 0$ (which is true by the assumption that $y \in B_r(x)$). Then if $d(y, z) < s$, we have

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

as desired. □



Remark 4. Convince yourself that the empty set is tautologically an open set. For $r \leqslant 0$, convince yourself that $B_r(x)$ is the empty set. It follows that we didn't need to make the restriction $r > 0$ in the above proposition. On the other hand, who cares bout balls of nonpositive radius?

Example 5. Here are some open sets in the plane.

You can perhaps see the real analysis idea 'an open set is a set which contains none of its boundary points' in these pictures.

Open sets in metric spaces satisfy some properties. I'll record and prove these here — it will be a good exercise in playing with definitions. There are, in fact, many properties of open sets in metric spaces that I *am not* listing here. (For instance, for any $x \in X$, the set $X \backslash \{x\}$ is open. Can you prove this?) Some of them will come up later in the course; this is what matters to us for now.

Theorem 2 (Open sets in metric spaces form a topology)**.** *Let (X, d) be a metric space. Open sets in X satisfy the following properties.*

a) The empty set $\varnothing$ and X itself are both open.

b) Suppose we have some collection $\{U_i\}_{i \in I}$ of open sets, indexed by a ***arbitrary*** *set I. Then the union*

$$U = \bigcup_{i \in I} U_i$$

is also open.

c) Suppose we have a ***finite*** *collection $V_1, \cdots, V_n$ of open sets. Then the intersection*

$$V = V_1 \cap \cdots \cap V_n$$

is an open set.

Proof. a) I mentioned this above for the empty set; it's just as vacuous for X. You should quickly confirm that this is true.

b) If $x \in U$, then by definition of the union $x \in U_i$ for some $i \in I$. Because U_i is open, by definition there is some > 0 so that $B_r(x) \subset U_i$. Because $U_i \subset U$, it follows that $B_r(x) \subset U$ for some $r > 0$, as desired.

c) If $x \in V$, then by definition of the intersection $x \in V_i$ for all $1 \leqslant i \leqslant n$. These are open sets, so by definition there are constants r_i for $1 \leqslant i \leqslant n$ so that $B_{r_i}(x) \subset V_i$.

We want this open ball to be contained in *all* of the V_i, so we shrink the radius. Set $r = \min(r_1, \cdots, r_n)$; then

$$B_r(x) \subset B_{r_i}(x) \subset V_i$$

for all i. It follows that $B_r(x) \subset \cap_{i=1}^n V_i = V$, as desired.

Draw a picture of this argument in the plane!

□

If you are not familiar with the idea of arbitrary unions, I recommend reviewing the relevant section in Munkres' Chapter 1. This will be very, very important to us, and taking unions over very big indexing sets I will be a standard technical trick.

As a relatively simple example, $(x-1, x+1)$ defines an open subset U_x of $\mathbb{R}$ for all $x \in \mathbb{R}$. Consider the set $I = \{x \in (0,1)\}$; then we have a family of open sets $\{U_x\}_{x\in(0,1)}$, and the union $\bigcup_{x\in(0,1)} U_x = (-1, 2)$. As expected, this is open.

It is traditional in topology to use U and V to denote open sets. If I ever see the sentence 'U is a closed set' I will get a mild stomachache.

Continuity and open sets

The following operation — which you should be familiar with from your review of set theory, say from Munkres chapter 1 — is going to be used pretty much every day for the rest of the semester.

Definition 5. *If $f : X \to Y$ is a function between two sets, and $S \subset Y$ is a subset, then the* inverse image *is*

$$f^{-1}(S) = \{x \in X \mid f(x) \in S\}.$$

This is **not the same** as applying the inverse of a function to a set. We will be using this operation for arbitrary functions, which are very rarely invertible. In particular, $f^{-1}(f(S))$ is usually not S, nor is $f(f^{-1}(S))$ usually S. The operation f^{-1} doesn't mean 'undo f', but rather, 'write down all the points which f maps into S.'

This is relevant because of the following definition and result, which you might consider the founding theorem of topology.

Definition 6. *A map $f : (X, d) \to (Y, d')$ between metric spaces is called* open-set continuous *when, for **all** open sets $U \subset Y$, the inverse image $f^{-1}(U)$ is an open subset of X.*

Theorem 3 (ϵ-δ continuity is open-set continuity)**.** *A map $f : (X, d) \to (Y, d')$ between metric spaces is open-set continuous if and only if it is continuous in the ϵ-δ sense.*

Proof. This will be a careeful definition chase. I think the easiest way to do this is to slowly translate the ϵ-δ definition. With quantifiers, this says

$$\forall x \forall \epsilon \exists \delta \big[\forall y [d(x, y) < \delta \implies d'(f(x), f(y)) < \epsilon\big].$$

Scary as this may look, it's just a symbolic writing of our previous definition of ϵ-δ continuity. The first thing I want to point out is that the expression $\forall y[d(x, y) < \delta \implies d'(f(x), f(y)) < \epsilon$ means precisely that

$$f(B_\delta(x)) \subset B_\epsilon(f(x)).$$

(Make sure you see why.) So we may rewrite our definition as

$$\forall x \forall \epsilon \exists \delta \big[f(B_\delta x) \subset B_\epsilon f(x)\big].$$

Next, as practice with the inverse image, I want you to convince yourself that

$$f(A) \subset B \iff A \subset f^{-1}(B).$$

Therefore, we can rewrite the definition once more as

$$\forall x \forall \epsilon \exists \delta \big[B_\delta x \subset f^{-1}\left(B_\epsilon f(x)\right) \big].$$

The presence of the inverse image here is promising! Let's show that *this* is equivalent to open-set continuity.

Assume f is ϵ-δ continuous. We want to prove it is open-set continuous. Pick $U \subset Y$ an open set. We want to show that $f^{-1}(U)$ is open. So pick $x \in f^{-1}(U)$; this means that $f(x) \in U$. Because U is open, that means (by definition) that there exists some $\epsilon > 0$ so that $B_\epsilon f(x) \subset U$. But by the reformulation of ϵ-δ continuity above, we see that there exists some $\delta > 0$ so that

$$x \in B_\delta x \subset f^{-1}(B_\epsilon f(x)) \subset f^{-1}(U),$$

so we've shown tht $f^{-1}(U)$ is open, as desired. Because U was an arbitrary open set wee see that f is open-set continuous.

Assume f is open-set continuous. We need to show that for all x and $\epsilon > 0$, there exists some $\delta > 0$ so that $B_\delta x \subset f^{-1}(B_\epsilon f(x))$.

Now, we poved earlier that $B_\epsilon f(x)$ is an open set. Because f is open-set continuous, $f^{-1}(B_\epsilon f(x))$ is an open set. Because $x \in f^{-1}(B_\epsilon f(x))$, by definition of open, we see that there exists some $\delta > 0$ so that

$$x \in B_\delta x \subset f^{-1}(B_\epsilon f(x)),$$

as desired. So we've shown f is ϵ-δ continuous. □

This abstract definition will not magically make it easier to prove that a function is continuous. If you start with a function between metric spaces you know nothing about except its definition, if you sit down to prove it is continuous using either definition, you will find yourself returning to the ϵ-δ definition. The equivalent and more abstact definition is great for proving theorems, but for checking examples, it is more or less worthless.

When we cover bases for topological spaces we will see a general form of this idea: if you want to check continuity, it usually suffices to check that the inverse image of a collection of nice sets are open. This is what the ϵ-δ definition of continuity is doing, and this idea is hidden in the above proof.

9/14: Topological spaces and continuity

Last time we saw the definition of a metric space: A set X equipped with a distance function d satisfying certain axioms.

We defined an open set in a metric space: it's a set U so that for any $x \in U$, there is also a ball of radius $r > 0$ with $B_r(x) \subset U$ as well. Intuitively: *an open set is a set for everything near (sufficiently close) to a point in the set is also in the set.*

We saw that open sets in a metric space satisfy certain properties. We also saw that the traditional $\epsilon - \delta$ definition of continuity in real analysis can be interpreted entirely in terms of open sets.

If topology is the abstract study of continuity, then we may be inspired by this to say: *continuity is about open sets; we should start by abstracting a notion of open set.*

Topological spaces

Definition 7. *If X is a set, a topology on X is a collection $\mathcal{T}$ of subsets of X (called open sets) — that is, a subset of the powerset, $\mathcal{T} \subset \mathcal{P}(X)$, whose elements are the open sets — satisfying the following axioms:*

- *Both the empty set $\varnothing$ and the entire space X are open sets. That is, $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$.*
- *Finite intersections of open sets are open. That is, if we have finitely many open sets $U_1, \cdots, U_n$, then their intersection $U_1 \cap \cdots \cap U_n$ is also an open set.*
- *Arbitrary unions of open sets are open. That is, if we have a set I of open sets U_i, one for each $i \in I$, then their union $\bigcup_{i \in I} U_i$ is also an open set.*

A topological space is a pair $(X, \mathcal{T})$, where X is a set and $\mathcal{T}$ a topology on X.

When you talk about a topology on a set, you will do so by stating a way to determine which sets are open. I would say 'a topology is a collection of sets which we deem to be open sets, which satisfy some properties.' The formulation as a subset of the power set is helpful to formulate this notion, but I do not suggest talking about it that way when you're trying to prove results or work with examples. It can be obscuring.

Example 6. The most familiar example is that of metric spaces. In each metric space, we have a notion of open sets. That specifies a topology:

Definition 8. *Let (X, d) be a metric space. There is a topology $\mathcal{T}_d$ on X, called the metric topology induced by d, defined as follows.*

A set U is open (with respect to $\mathcal{T}_d$; that is, $U \in \mathcal{T}_d$) if, for all $x \in U$, there exists an $r > 0$ with $B_r(x) \subset U$.

We proved that this is a topology on Tuesday, when we showed that the collection of open sets in a metric space satisfies the axioms of a topology (though we didn't use that language!)

Thus, perhaps the most famous topological space is $(\mathbb{R}, \mathcal{T}_d)$, where $d(x, y) = |x - y|$. This is called *the standard topology on* $\mathbb{R}$; in the homework I abbreviate this pair as $\mathbb{R}_{\text{std}}$. Following that, we have the Euclidean space $\mathbb{R}^n$ equipped with the metric

$$d_{\text{Euc}}\big((x_1, \cdots, x_n), (y_1, \cdots, y_n)\big) = \sqrt{(x_1 - y_1)^2 + \cdots (x_n - y_n)^2};$$

again the induced topology is called the *standard topology* on $\mathbb{R}^n$.

Notice that every subset S of a metric space X is *also a metric space*, by restricting the distance function to S. Important examples to us include intervals $[a, b), [a, b], (a, b], (a, b) \subset \mathbb{R}$, and the unit circle $S^1 \subset \mathbb{C}$.

Many different metrics may induce the same topology. We will see next week that all of the metrics d_p on $\mathbb{R}^n$, for $1 \leqslant p \leqslant \infty$, induce the same topology on $\mathbb{R}^n$ — the standard topology. (Later this semester you will prove something even more general, which shows that any metric which has properties similar to d_p induces the same topology.)

Example 7. Another important topology on a set X is the *discrete topology on* X, in which *every subset of* X *is open*; that is, $\mathcal{T}_{\text{disc}} = \mathcal{P}(X)$.

This is tautologically a topology. First, $\varnothing$ and X are open (because every subset of X is open); second, finite intersections of open sets are open (because every subset of X is open), and unions of open sets are open (because every subset of X is open).

The discrete topology is still a metric topology: it is the metric topology induced by the metric

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

We can get some intuition out of this. In the discrete topology, we think of every point of X as being isolated. That's literally true in terms of the metric d_{disc}; the ball

$$B_r(x) = \{y \in X \mid d_{\text{disc}}(x, y) < r\}$$

of radius $r \leqslant 1$ around *any* point x is just the single point $\{x\}$; every other point has distance at least 1 from x.

Intuitively: every set is open, because for any point in that set, all nearby points are in that set — because there are *no other points* nearby.

Example 8. Going in the opposite direction is the *indiscrete topology on* X, in which the only open sets are $\varnothing$ and X itself: the ones which are axiomatically required to be open sets in any topology. Phrased set-theoretically,

$$\mathcal{T}_{\text{indisc}} = \{\varnothing, X\} \subset \mathcal{P}(X).$$

This satisfies the axioms of a topology axiomatically again: if U_i is a collection of open sets, either one of them is X (in which case their union is X, which is open), or they're all empty, in which case their union is still empty — again open. Similarly, either one of them is empty (in which case their intersectioin is empty) or all of them are X (whose intersection is again X).

We will see in a moment that this is *not a metric topology*; imagine the indiscrete topology as a scenario in which all of the points of X were clumped together, so close as to be essentially indistinguishable.

We can gain some intuition by pretending it is a metric topology, induced by the "metric" $d(x, y) = 0$. (This is not a metric for $|X| > 1$, because of the axiom that $d(x, y) = 0 \implies x = y$.) With respect to this "metric", every ball $B_r(x)$ for $r > 0$ is the *whole space* — every point is within distance r of x, because everything is distance 0 from x. If one tried to apply the definition of the metric topology, then the only non-empty open set would be X itself.

Lemma 4 (Metric spaces are T_1 spaces)**.** *Suppose $\mathcal{T}_d$ is the metric topology on X, for (X, d) a metric space. Then for any point $x \in X$, the set $X \backslash \{x\}$ (the complement of the point x) is an open set.*

Proof. Pick a point $y \in X \backslash \{x\}$. Because $x \neq y$, we have $d(x, y) > 0$. Choose $0 < r \leqslant d(x, y)$; then

$$x \notin B_r(y) = \{z \in X \mid d(y, z) < r\}$$

because $d(x, y) \geqslant r$. Equivalently said, $B_r(y) \subset X \backslash \{x\}$. Because y was arbitrary, by the definition of open sets in the metric topology, $X \backslash \{x\}$ is an open set. □

Corollary 5 (The indiscrete topology is not metric)**.** *If $|X| > 1$ (that is, X has cardinality more than 1: it has more than one point), then the indiscrete topology is not the topology induced by any metric d.*

This follows immediately: If $|X| > 1$, then for any $x \in X$, the set $X \backslash \{x\}$ is neither X nor the empty set, so is not open in the indiscrete topology; but it is open in any metric topology. Thus the indiscrete topology is not a metric topology.

Example 9. The *Sierpinski space* is the set $\{0, 1\}$ equipped with the topology $\mathcal{T}_S = \{\varnothing, \{0\}, \{0, 1\}\}$ – that is, the empty set and the whole set are open, as is the set $\{0\}$.

This is neither a metric topology (in which $\{1\} = X \backslash \{0\}$ would be open) nor the indiscrete topology (in which $\{0\}$ would not be open).

It is hard to visualize this space! I might think of it as a space in which 0 is near 1, but 1 is not near 0; as a sort of 'asymmetric' metric space. You might get some more intuition when we look at continuous functions later, and even more when we much later cover quotient spaces.

As we go forward, I may begin to say 'a topological space X'; the fact that it is *equipped* with a topology $\mathcal{T}$ will be implicit.

Continuous functions

Now that we have an abstract notion of 'a space equipped with open sets', we can just apply the definition of a topology we learned while studying metric spaces.

Definition 9. *A map $(X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is continuous if, for all open subsets $U \subset Y$ (with respect to $\mathcal{T}_Y$; that is, $U \in \mathcal{T}_Y$), the inverse image $f^{-1}(U)$ is an open subset of X (with respect to $\mathcal{T}_X$; that is, $f^{-1}(U) \in \mathcal{T}_X$).*

Last time, we proved that if both of these topologies are metric topologies, then a map is continuous if and only if it's continuous in the $\epsilon - \delta$ sense. We can also try to understand what continuity means in the examples we just learned.

Example 10. Some functions from real (or a little bit of complex) analysis that you can check are continuous using the $\epsilon - \delta$ definition: $f : \mathbb{C} \to \mathbb{C}$, given by e^z, is continuous; $m : \mathbb{C}^2 \to \mathbb{C}$ given by $m(x, y) = xy$; $a : \mathbb{C}^2 \to \mathbb{C}$, given by $a(x, y) = x + y$; division $q : \mathbb{C}^2 \backslash (\mathbb{C} \times \{0\}) \to \mathbb{C}$, given by $q(x, y) = x/y$; their restriction to any subsets (in particular, to the reals, or to intervals).

Some less common ones that you certainly used a lot in calculus include trig functions and their inverses (on the correct domains!)

Proposition 6 (Any function from a discrete space is continuous)**.** *If $(X, \mathcal{T}_{disc})$ is a discrete topological space, and $(Y, \mathcal{T})$ is any topological space whatsoever, then any function $f : X \to Y$ is continuous.*

Proof. Given any open set $U \subset Y$, the inverse image $f^{-1}(U)$ is open with respect to the discrete topology — because *every set* is open with respect to the discrete topology. □

Functions *to* discrete spaces are less common. Can you characterize when they exist?

Proposition 7 (Any function to an indiscrete space is continuous)**.** *If $(X, \mathcal{T}_{indisc})$ is an indiscrete topological space, and $(Y, \mathcal{T})$ is any topological space, then any function $g : Y \to X$ is continuous. (That is: any function to an indiscrete space is continuous.)*

Proof. The only open subsets of an indiscrete space are $\varnothing$ and X. Then we just need to check that $g^{-1}(\varnothing) = \varnothing$ is open (axiomatic) and whether $g^{-1}(X) = Y$ is open (axiomatic). □

Again, functions *from* indiscrete spaces are less common; we may see how to characterize those in a week or so.

Proposition 8 (Open sets correspond to maps to Sierpinski space)**.** *Let $(X = \{0, 1\}, \mathcal{T}_S)$ be the Sierpinski space, and let $(Y, \mathcal{T})$ be a topological space. Then $f : Y \to X$ is continuous if and only if $f^{-1}(0)$ is open.*

Proof. We know axiomatically that $f^{-1}(\varnothing)$ and $f^{-1}(X)$ are open, for any function f. Thus f is continuous if and only if the inverse image of open sets *other* than $\varnothing$ and X are open. In the case of the Sierpinski space, there is only one other open set: $\{0\}$. Thus $f : Y \to X$ is continuous if and only if $f^{-1}(\{0\}) = f^{-1}(0)$ is open. □

There are also some ways to construct new continuous functions from old ones. We will use this all the time to reduce the amount of new work we have to put in to showing that a function is continuous.

Theorem 9 (Compositions of continuous functions are continuous)**.** *Let $(X, \mathcal{T}_1), (Y, \mathcal{T}_2), (Z, \mathcal{T}_3)$ be topological spaces. If $f : X \to Y$ is continuous, and $g : Y \to Z$ is continuous, then the composite $gf : X \to Z$ is continuous (all with respect to the given topologies on X, Y, and Z).*

Proof. If $U \subset Z$ is open (with respect to $\mathcal{T}_3$), we want to show $(gf)^{-1}(U)$ is open (with respect to $\mathcal{T}_1$).

Now we're given that g is continuous, so $g^{-1}(U)$ is open (w/r/t $\mathcal{T}_2$), by definition of continuity. And because f is continuous, we further have that $(gf)^{-1}(U) = f^{-1}(g^{-1}(U))$ is open (w/r/t $\mathcal{T}_1$), as desired. □

Proposition 10 (Constant functions are continuous)**.** *If $(X, \mathcal{T}_1)$ and $(Y, \mathcal{T}_2)$ are topological spaces, then for each $y \in Y$, the constant map $f_y(x) = y$ (which sends every point in X to the point y) is continuous.*

Proof. If $U \subset Y$ is open and $y \in U$, then $f_y^{-1}(U) = X$, which is axiomatically open. If, on the other hand, $y \notin U$, we have that $f_y^{-1}(U) = \varnothing$, which is also axiomatically open. □

Lastly, but not least, we have that the identity map is always continuous:

Proposition 11 (The identity function is continuous)**.** *If $(X, \mathcal{T}_1)$ is a topological space, the identity map $1_X : X \to X$, given by $1_X(x) = x$, is continuous (with respect to the topology $\mathcal{T}_1$ on both domain and codomain).*

Proof. If $U \subset X$ is open with respect to $\mathcal{T}_1$, then $1_X^{-1}(U) = U$, because 1_X is the identity function. U, of course, is open with respect to $\mathcal{T}_1$, so 1_X is continuous. □

Homeomorphisms

The right notion of *equivalence* in topology might be more complicated than you expect if you're coming from algebra. After all, you remember the definitions; a homomorphism $f : G \to H$ is said to be an *isomorphism* if f is also a bijection. So you might guess that the right definition of 'equivalence' in topology is: 'A continuous bijection'.

This doesn't work. To point at the right definition, let me point out a special property that's true for groups.

Lemma 12 (Group isomorphisms are bijective homomorphisms)**.** *Suppose $f : G \to H$ is a group homomorphism, which is also a bijection. Then the inverse map $f^{-1} : H \to G$ is also a homomorphism.*

Proof. Pick $h_1, h_2 \in H$. Because f is a bijection, there are unique elements $g_1, g_2 \in G$ for which $f(g_i) = h_i$; these have $f(g_1 g_2) = f(g_1)f(g_2) = h_1 h_2$. Applying f^{-1} on both sides, we see that $f^{-1}(h_1 h_2) = g_1 g_2 = f^{-1}(h_1)f^{-1}(h_2)$. So f^{-1} is a group homomorphism as well. □

Example 11. Consider the interval $[0, 2\pi)$ equipped with the usual metric topology, and the unit circle $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ with the usual metric topology (inherited from $\mathbb{C}$).

The map $\exp : [0, 2\pi) \to S^1$, given by $\exp(\theta) = e^{i\theta}$, is continuous (because e^{iz} is continuous on all of $\mathbb{C}$, it's also continuous by restriction), and is a bijection (by your knowledge of trig, and the fact that $e^{i\theta} = \cos\theta + i \sin\theta$).

However, the inverse function $\exp^{-1} : S^1 \to [0, 2\pi)$ is *not* continuous. To see this, **CHECK** that $[0, \pi) \subset [0, 2\pi)$ is open in the metric topology; but $\exp^{-1}[0, \pi)$ is **not** open in the circle: any ball of radius $r > 0$ around $\exp^{-1}(0) = 1$ contains points in the circle with negative y-coordinate, but $\exp^{-1}[0, \pi)$ is contained entirely within the upper half-plane $y \geqslant 0$.

Topologically equivalent shapes should soooomehow look alike; you'll get more of an intuition for what that means as the course goes forward. But keep in mind what happened in the previous example. If topology is about nearness, abstractly, then the reason f^{-1} was not continuous above: points near 1 get sent by f^{-1} both to points near 0 and points near 2π. We 'broke' the circle in writing down the inverse; jumps like that are not continuous.

Inspired by our failure, we define the following.

Definition 10. *Let $(X, \mathcal{T}_1)$ and $(Y, \mathcal{T}_2)$ be topological spaces. A function $f : X \to Y$ is called a homeomorphism if f is a continuous bijection and the inverse map $f^{-1} : Y \to X$ is also continuous.*

We say that X and Y are homeomorphic if there exists a homeomorphism $f : X \to Y$.

As usual, continuity here means with respect to the given topologies.

At this point I am going to begin dropping the explicit mention of the topology. I will write X or Y for *topological spaces*, meaning sets equipped with a topology; when I write that something is a continuous map, it will mean with respect to the topology which was left implicit; if I write that a set is open, it will mean with respect to the topology. By abuse of notation, I will also write X for the underlying set.

Proposition 13 (Homeomorphism is an equivalence relation)**.** *The relation between topological spaces, $X \cong Y$ if X is homeomorphic to Y, forms an equivalence relation on topological spaces.*

Proof. First, we have reflexivity: $1_X : X \to X$ is a continuous bijection, whose inverse is $1_X^{-1} = 1_X$ — the same function, so also continuous.

Second, symmetry: if $X \cong Y$, then there's a homeomorphism $f : X \to Y$. Then because f is a homeomorphism, the inverse function $f^{-1} : Y \to X$ is a continuous bijection, and $(f^{-1})^{-1} = f$ is continuous; so $f^{-1} : Y \to X$ is a homeomorphism.

Third, transitivity: if $X \cong Y$ and $Y \cong Z$, write homeomorphisms $f : X \to Y$ and $g : Y \to Z$. Then their composite $gf : X \to Z$ is a homeomorphism: gf is continuous because composites of continuous functions are continuous, and $(gf)^{-1} = f^{-1}g^{-1}$ is continuous because f^{-1} and g^{-1} are continuous (f and g were both homeomorphisms) and composites are continuous. □

You should check all claims that are not immediately clear to you in the following examples; feel free to invoke facts from real analysis.

Example 12. Any two open intervals (a, b) in $\mathbb{R}$ (equipped with the metric topology), with $-\infty \leqslant a < b \leqslant \infty$, are homeomorphic.

First, all bounded intervals are homeomorphic: $f_{ab} : (a, b) \to (0, b - a)$, given by $f(x) = x - a$, is a homeomorphism. Then $g_{b-a} : (0, b - a) \to (0, 1)$, given by $g_{b-a}(x) = x/(b - a)$, is a homeomorphism. So

$$f_{cd}^{-1} g_{d-c}^{-1} g_{b-a} f_{ab} : (a, b) \to (c, d)$$

is a homeomorphism. (If this is spooky, you should write this out more explicitly by calculating a formula for the composite function. We stretch the real line and translate it so that $a \to c$ and $b \to d$.)

Next, all intervals of the form $(-\infty, a)$ are homeomorphic to each other (just like before, translate them so the right endpoint is zero), and these are all homeomorphic to any interval of the form $(-a, \infty)$, by the homeomorphism $f(x) = -x$.

Next, the exponential map $e^t : (-\infty, 0) \to (0, 1)$ is a homeomorphism with inverse $\log s$; so half-infinite open intervals are all homeomorphic to bounded intervals.

Lastly, $\mathbb{R}$ itself: the function $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ is a homeomorphism.

Similarly, one-side-closed intervals $[a, b)$ and $(a, b]$ in $\mathbb{R}$ are all homeomorphic, and two-side-closed intervals $[a, b]$ are all homeomorphic.

Example 13. The open ball of radius r in $\mathbb{R}^n$, given by

$$B_r(0) = \{(x_1, \cdots, x_n) \mid \sum x_i^2 < r^2\}$$

(with its metric topology) is homeomorphic to $\mathbb{R}^n$ itself (with the metric topology).

To see this, write x for a point in $\mathbb{R}^n$. The map $s : [0, r) \to [0, \infty)$ given by $s(\lambda) = \lambda/(r - \lambda)$, is a homeomorphism for which both

$$\lambda \mapsto s(\lambda)/\lambda : [0, r) \to [0, \infty)$$

and $\lambda \mapsto s^{-1}(\lambda)/\lambda : [0, \infty) \to [0, \infty)$, are also homeomorphisms.

Now check that $f : \mathbb{B}_r(0) \to \mathbb{R}^n$ given by

$$f(x) = \frac{s(|x|)x}{|x|},$$

is a continuous bijection, and that its inverse is

$$f^{-1}(x) = \frac{s^{-1}(|x|)x}{|x|},$$

which is continuous by the same logic.

Actually, every convex set in $\mathbb{R}^n$ is homeomorphic to the whole $\mathbb{R}^n$, though the proof is somewhat tricky. There is a more general notion of a 'star-shaped set': these too are homeomorphic to $\mathbb{R}^n$ but it is much harder to prove.

Example 14. The unit square $Sq = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| = 1 \text{ or } |x_2| = 1\}$ is homeomorphic to the unit circle $S^1 \subset \mathbb{R}^2$.

Proof. Notice that there's exactly one point of Sq and exactly one point of S^1 on each radial line through the origin. The idea is to write down how to scale from one to the other.

I will write $|(x_1, x_2)|_2 = \sqrt{x_1^2 + x_2^2}$ and $|(x_1, x_2)|_\infty = \max(|x_2|, |x_2|)$ in what follows. The map $f : Sq \to S^1$ given by $f(x) = x/|x|_2$ is a continuous bijection — it is continuous because $|x|_2$ is continuous on all of $\mathbb{R}^2$ and nonzero on Sq, and it is a bijection because there's exactly one point of each set on a given radial line through the origin.

Its inverse is $g(x) = x/|x|_\infty$; again, this is continuous because $|x|_\infty$ is continuous on all of $\mathbb{R}^2$, and nonzero on the unit circle. □

Example 15. The topological spaces underlying the following letters (considered as 1-dimensional subsets of $\mathbb{R}^2$) are homeomorphic:

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

Other clusters are $\mathsf{E, F, T, Y}$, and $\mathsf{A, R}$, and $\mathsf{D, O}$.

The famous example is, of course, that a coffee cup is homeomorphic to a donut. That will be not be easy to see right now rigorously, but following what you saw with intervals earlier, you might see the idea: move the handle down to underneath the coffee mug, lay out the vertical part of the "mug" portion flat (so it just looks like you've added a handle to a very flat ellipsoid), and then shrink the ellipsoid until it looks like the end of a donut. Or you'll see a rigorous proof at the end of the course (time permitting!)

Showing that two spaces are **not** homeomorphic is much harder! Following your homework, you will be able to see

Corollary 14 (Discreteness is a topological property)**.** *If $(X, \mathcal{T}_1)$ is homeomorphic to a discrete space, then $\mathcal{T}_1$ is discrete.*

In particular, topologies which are not discrete — in which there is some set which fails to be open — are not even *homeomorphic* to a discrete space.

We will see later in the class that no two of $(0, 1)$, $[0, 1)$, and $[0, 1]$ are homeomorphic.

Closed sets; open and closed maps

It is often useful to have a notion of *closed sets* to simplify our phrasing.

Definition 11. *Let X be a topological space. We say that $C \subset X$ is closed (with respect to the given topology) if its complement $X \backslash C$ is open (with respect to the given topology).*

Earlier, we made a fuss about the fact that for a metric topology $\mathcal{T}_d$, every set $X \backslash \{x\}$ is open. We can rephrase that as saying that points are closed sets in metric topologies.

Closed sets satisfy some properties directly from the axioms of a topology: the empty set and entire space are closed sets; **arbitrary** intersections of closed sets are closed; and **finite** unions of closed sets are closed. (It follows, for instance, that any finite subset of X is closed in a metric topology on X.)

Proposition 15 (Closed-continuous maps are continuous)**.** *Let X and Y be topological spaces. A map $f : X \to Y$ is continuous if and only if, for all closed subsets $C \subset Y$, the inverse image $f^{-1}(C) \subset X$ is closed.*

Proof. If f is continuous, and $C \subset Y$ is closed, write $U = Y \backslash C$ for its (open) complement. Then we have that $f^{-1}(U)$ is open, and by general properties of the inverse image, that $X \backslash f^{-1}(U) = f^{-1}(C)$; so $f^{-1}(C)$ is the complement of an open set, hence closed. This proves the forward direction.

The same argument proves the reverse direction: if U is open, its complement C is closed; we're given that $f^{-1}(C)$ is closed, so $X \backslash f^{-1}(C) = f^{-1}(U)$ is open, as desired. □

Warning. Unlike windows, most sets are neither open nor closed, and some sets (for instance, the empty set) are *both*. Be careful about this!

Open and closed maps

One might be inspired, after studying continuous maps (which are defined in terms of the inverse image) if there is a corresponding theory for the forward image. The answer is 'not much' — the kinds of functions I talk about below do not appear often in nature unless the functions are already continuous. These will show up again occasionally later, so best to define them now.

Definition 12. *Let X and Y be topological spaces. A function $f : X \to Y$ is open if, for every open subset $U \subset X$, the image $f(U)$ is open in Y. Similarly, f is closed if $f(C)$ is closed for every closed subset $C \subset X$.*

Many (most?) continuous maps you're comfortable with are not open!

Example 16. The map $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x^2$, is continuous but not open in the standard topology — note that $f(\mathbb{R}) = [0, \infty)$, which is not an open set.

And what's more, open maps don't need to be continuous:

Example 17. Consider $\mathbb{R}$ with the standard topology and the 3-element set $\{-, 0, +\}$ equipped with the discrete topology. The function $\mathsf{sign} : \mathbb{R} \to \{-1, 0, +\}$, defined by

$$\mathsf{sign}(x) = \begin{cases} + & x > 0 \\ 0 & x = 0 \\ - & x < 0 \end{cases}$$

is open but not continuous: it's not continuous because $\{0\} \subset \{-, 0, +\}$ is open in the discrete topology, but $\mathsf{sign}^{-1}(0) = \{0\} \subset \mathbb{R}$ is not open in the standard topology.

However, this map is open (and in fact any map to a discrete space is open) — since every subset is open in the discrete topology, so is $f(U)$, no matter what U is.

We mention one last property.

Proposition 16 (Homeomorphisms are open continuous bijections)**.** *Let X and Y be topological spaces, and let $f : X \to Y$ be a continuous bijection. The map f is a homeomorphism if and only if f is open.*

Proof. This is almost tautological: to say that f is a homeomorphism means that f^{-1} is a continuous map; to say that f^{-1} is a continuous map means that, for each open $U \subset X$, the set $(f^{-1})^{-1}(U)$ is open in Y; and $(f^{-1})^{-1}(U) = f(U)$. □

9/16: Closure, separability, and bases

Today, we'll cover *operations on subsets of topological spaces* (a basic tool which will recur for the rest of the course) and the idea of a *basis* for a topological space, which lets us more easily check whether maps are continuous.

Closure and interior

In analysis you learn about a number of operations on subsets of Euclidean space, such as the interior or closure of a set, and these turn out to be handly little gadgets. We can define them just as well in topology.

Definition 13. *Let X be a topological space.*[1] *If $A \subset X$ is* ***any*** *subset, we say that the* closure *of A — written $\overline{A}$ — is*

$$\overline{A} = \bigcap_{\substack{S \subset X \text{ closed} \\ A \subset S}} S.$$

Similarly we say the interior *of A, written A°, is*

$$A^\circ = \bigcup_{\substack{U \subset X \text{ open} \\ U \subset A}} U.$$

Proposition 17 ($\overline{A}$ is the smallest closed set containing A)**.** *The closure $\overline{A}$ is closed, and if $A \subset S$ where S is a closed set, then $\overline{A} \subset S$. It follows that $\overline{A}$ is the 'smallest' closed set containing A: it is a closed set containing A, and every other closed set containing A is contained in $\overline{A}$.*

Conversely, if $U \subset A$ and U is open, then $U \subset A^\circ$: the interior is the largest open set contained in A.

Proof. We will show this for the closure; a similar argument implies the corresponding result for the interior.

First, $\overline{A}$ is closed. It's given as the intersection of a family of closed sets. Just as the *union* of an arbitrary family of open sets is open, by taking complements we see that an arbitrary intersection of closed sets is closed. More precisely,

$$\overline{A}^c = \left(\bigcap_{\substack{S \subset X \text{ closed} \\ A \subset S}} S \right)^c = \bigcup_{\substack{S \subset X \text{ closed} \\ A \subset S}} S^c;$$

because each S is closed, this is a union of open sets, hence open. So $\overline{A}$ is closed.

Second, suppose S is a closed set with $A \subset S$. Then S is one of the sets we are taking the intersection of; it follows that

$$\overline{A} = \bigcap_{\substack{S \subset X \text{ closed} \\ A \subset S}} S \subset S.$$

□

It follows immediately that if A is closed, $\overline{A} = A$; if A is open, then $A^\circ = A$.

Example 18. In $\mathbb{R}_{\text{std}}$ we have

$$\overline{(a,b)} = [a,b],$$

and

$$\overline{\{1, 1/2, 1/3, \cdots\}} = \{1, 1/2, 1/3, \cdots, 0\} = \{1/n \mid n \in \mathbb{Z}_{\geqslant 1}\} \cup \{0\}.$$

In X_{disc} we have $\overline{S} = S$ and $S^\circ = S$ for any set S.

In X_{indisc}, if $S \subset X$ is a proper nonempty subset (so $S \neq \varnothing$ and $S \neq X$), we have $\overline{S} = X$ and $S^\circ = X$.

I suggest proving these claims, as practice, and drawing pictures of the real-line examples. Your intuition from real analysis should be that 'the closure adds every limit point of the original set.' (Does this help you understand the indiscrete space better?)

I will state the topological characterization of this claim later, but we won't use it as often as you might think.

[1]The topology $\mathcal{T}$ is suppressed from notation.

Closure plays well with subsets and unions:

Proposition 18 (Closure preserves subsets and unions)**.** *For any pair of nested subsets $A \subset B \subset X$, we have $\overline{A} \subset \overline{B}$.*

For any two subsets $A, B \subset X$ (not necessarily nested), we have $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

Proof. For the first claim, $A \subset B \subset \overline{B}$ — the latter inclusion by definition of the closure. Because $\overline{B}$ is a closed set containing A, and $\overline{A}$ is contained in any such set, it follows that $\overline{A} \subset \overline{B}$ as desired.

For the second claim, we will show that each set contains the other (and thus that they are equal).

Because $\overline{A}$ and $\overline{B}$ are closed sets, their union $\overline{A} \cup \overline{B}$ is closed (remember, unions of finitely many closed sets are closed, just as intersections of finitely many open sets are open). Because $A \subset \overline{A}$ and $B \subset \overline{B}$ it follows that $A \cup B \subset \overline{A} \cup \overline{B}$. Because this is a closed set containing $A \cup B$ — and the closure is the smallest such — we see that $\overline{A \cup B} \subset \overline{A} \cup \overline{B}$.

On the other hand, we know that $A \subset A \cup B \subset \overline{A \cup B}$, so we see by the first part of this proposition that $\overline{A} \subset \overline{A \cup B}$, and similaly for $\overline{B}$. Taking a union we see that $\overline{A} \cup \overline{B} \subset \overline{A \cup B}$ as desired. $\square$

Can you formulate and prove a corresponding statement for the interior? Your theorem will follow from the following proposition, but it may be good practice to prove your result first.

Proposition 19 (The complement of the closure is the interior of the complement)**.** *If $A \subset X$ is any set, then we have an equality*

$$X \backslash \overline{A} = (X \backslash A)^\circ.$$

Similarly, we have $\overline{X \backslash B} = X \backslash B^\circ$.

Proof. We will prove this using the characterization of the interior as the largest open set contained in a given set.

First, because $\overline{A}$ is closed, the complement $X \backslash \overline{A}$ is open (by definition of 'closed set') and contained in $X \backslash A$. It follows that $X \backslash \overline{A} \subset (X \backslash A)^\circ$ because the interior contains every open set contained in $X \backslash A$.

Now if we have an open set U with

$$X \backslash \overline{A} \subset U \subset X \backslash A,$$

then taking complements we see that $A \subset X \backslash U \subset \overline{A}$. Because U is open, the complement $X \backslash U$ is closed; because $\overline{A}$ is the smallest closed set containing A, but $X \backslash U$ is contained in it and contains A, we see that $X \backslash U = \overline{A}$. Thus $U = X \backslash \overline{A}$, and we have shown that $X \backslash \overline{A}$ is the largest open set contained in A.

The other relation can either be proved similarly or by applying the first relation to $A = X \backslash B$. $\square$

We will close out[2] our discussion of the closure with a topological version of the statement that 'the closure is obtained by adding all of the limit points'.

Theorem 20 (Limit-point characterization of closure)**.** *Let $A \subset X$ be an arbitrary subset. Then we have $x \in \overline{A}$ iff, for all open sets $x \in U \subset X$, the set $U \cap A$ is nonempty.*

The condition above could be read as 'x is in the closure if and only if every neighborhood of x meets A'.

Proof. It is easier to prove the contrapositive, taking advantage of the complement $X \backslash \overline{A} = (X \backslash A)^\circ$. (The statement is usually used in the form above, which is why I didn't just state the contrapositive.)

If $x \notin \overline{A}$, then $x \in (X \backslash A)^\circ$, by the formula we just proved. Taking $U = (X \backslash A)^\circ \subset X \backslash A$, this means there is an open set $x \in U$ so that $U \cap A$ is empty. This gives one implication of the contrapositive.

For the other direction, suppose that there exists an open set $x \in U \subset X$ so that $U \cap A$ is empty. The latter condition means $U \subset X \backslash A$ and hence $U \subset (X \backslash A)^\circ$, as this is the largest open set contained in $X \backslash A$. By the first condition, we see that

$$x \in U \subset (X \backslash A)^\circ = X \backslash \overline{A},$$

and hence $x \notin \overline{A}$, as desired. $\square$

---

[2] ☺

Separable spaces

Now that we have the notion of closure onhand, let me point out that a concept from analysis that can now be translated into topological terms (and hence applied to arbitrary topological spaces).

Definition 14. *Let X be a topological space. We say $S \subset X$ is* dense *if $\overline{S} = X$.*
We say a topological space is separable *if there is a* ***countable***[3] *set S with $\overline{S} = X$.*

Equivalently, a set $S \subset X$ is dense if and only if every nonempty open set U intersects S nontrivially (that is, for all nonempty U, we have $U \cap S \neq \varnothing$).

(Do you see why? Think about complements, or about the limit-point characterization of closure.)

Example 19. The set $(0,1)$ is dense in $[0,1]$ (with the standard topology induced by the metric $d(x,y) = |x-y|$). The set $\mathbb{Q}$ is dense in $\mathbb{R}_{std}$; the real line is separable. $(0,1) \cap \mathbb{Q}$ is dense in $[0,1]$, so $[0,1]$ is separable too. Any nonempty set is dense in the indiscrete space (so any indiscrete space is separable). The only dense set in a discrete space X_{disc} is X itself. It follows that X_{disc} is separable if and only if the underlying set X is countable.

Often one things of separable spaces as being 'not too large', while non-separable spaces are 'very spread out' — so spread out you need uncountably many points to get near every other point. You will see some more interesting non-separable examples in homework.

Here is some justification for the intuition above.

Proposition 21 (The image of a separable space is separable)**.** *If X is a separable topological space and $f : X \to Y$ is a surjective continuous function, then Y is also separable.*

Proof. Let $S \subset X$ be a countable dense set. I claim that $f(S) \subset Y$ is dense; it is certainly still countable. To show this, we will use the limit-point characterization of the closure. We want to show that for all $y \in Y$ and any open set $y \in U$, we have $U \cap f(S) \neq \varnothing$. Because $f^{-1}(\varnothing) = \varnothing$, equivalently, we want to show that

$$\varnothing \neq f^{-1}(U \cap f(S)) = f^{-1}(U) \cap f^{-1}(f(S)) \supset f^{-1}(U) \cap S.$$

Because f is surjective, the set $f^{-1}(U)$ is nonempty; pick some $x \in f^{-1}(U)$. Because $x \in f^{-1}(U)$ and S is dense, it follows that $S \cap f^{-1}(U) \neq \varnothing$ — which implies exactly what we want. □

Bases for topological spaces

In metric spaces, it was actually in some sense a *advantage* that we had the ϵ-δ definition of continuity onhand, despite how intense it looks. That definition actually vastly **reduces** the amount of computation we have to do, at the cost of making those computations perhaps less conceptual. For metric spaces, we saw that a map $f : X \to Y$ is continuous if, and only if, $f^{-1}(B_\epsilon y)$ is open for all $\epsilon > 0$ and all $y \in Y$; and we also saw (definitionally) that we can check whether a set is open by seeing whether it contains (sufficiently small) open balls around any point.

The idea of a basis for a general topological space is a useful gadget in a similar spirit, which (among other things!) allows us to reduce the amount of computation we do.

Definition 15 (Basis; topology generated by a basis)**.** *Let X be a set. A collection of subsets $\mathcal{B} \subset \mathcal{P}(X)$ is called a* basis *— and the sets $B \in \mathcal{B}$ called the* basic open sets *— if the following two conditions hold.*

a) Every point lies in some basic open set. That is, for each $x \in X$, there is some $B \in \mathcal{B}$ so that $x \in B$. We sometimes say that '$\mathcal{B}$ is a cover of X'.

b) For any finite collection of basic open sets $B_1, \cdots, B_n$, and any $x \in B_1 \cap \cdots \cap B_n$, there exists a basic open set B with

$$x \in B \subset B_1 \cap \cdots \cap B_n.$$

[3]In this course, 'countable' means 'finite or countably infinite'. A set is countable if it can be put in bijection with a subset of the naturals $\mathbb{N}$.

You should draw a picture of condition (b) above. Also, look back at the proof from Day 1 that open sets in a metric space formed a topology — you'll see that we proved condition (b) on the way to checking that a finite intersection of open sets is open.

Example 20. Let (X, d) be a metric space. Then $\mathcal{B}_d = \{B_r(x)\}_{\substack{x \in X \\ r > 0}}$ is a basis. Once you learn the definition you will be able to prove that $\mathcal{T}_d$ is the topology generated by the basis $\mathcal{B}_d$.

Definition 16. *Let $\mathcal{B}$ be a basis on the set X. The* topology generated by $\mathcal{B}$*, written $\mathcal{T}_{\mathcal{B}}$, is described as follows.*

We say a set $U \subset X$ is open (with respect to $\mathcal{B}$) if, for all $x \in U$, there exists a basis open set B_x with $x \in B_x \subset U$.

We say that $\mathcal{B}$ is a basis for the topology $\mathcal{T}$ if we have $\mathcal{T} = \mathcal{T}_{\mathcal{B}}$.

Compare this with the definition of open sets in a metric space. You should notice that they are not precisely the same, but they are very similar (and, as mentioned, open balls do form a basis for the topology on a metric space).

The collection of open sets $\mathcal{T}_{\mathcal{B}}$ forms a topology on X. You should confirm this yourself as practice with the definitions.

Warning. The bases you see in topology are not at all comparable to the bases you learned about in linear algebra. They are largely unrelated ideas with the same name. A given topology will have **many, many bases**, and we will see in an example later that in fact if $\mathcal{T}$ is a topology, then it is also a basis — and the topology generated by that basis is once again $\mathcal{T}$; and further, you may have bases of many different cardinalities. (By contrast, it's totally preposterous to say that V is a basis for a vector space V, and any two linear-algebra bases for V have the same cardinality.)

Example 21. We already know that $\mathcal{B} = \{(x - r, x + r)\}_{\substack{x \in \mathbb{R} \\ r > 0}}$ is a basis for $\mathbb{R}_{std}$. But if desired one may make this collection much smaller. For instance,

$$\mathcal{B}_Q = \{(x - r, x + r)\}_{\substack{x, r \in \mathbb{Q} \\ r > 0}}$$

is also a basis for the topology on $\mathbb{R}_{std}$ — but $\mathcal{B}_Q$ is a countable set. Much smaller!

Proposition 22 ($\mathcal{B}$-open sets are the unions of basic open sets)**.** *If $\mathcal{B}$ is a basis of open sets for the topology $\mathcal{T}$ on the set X, then a subset $U \subset X$ is open (that is, $U \in \mathcal{T}$) if and only if U is a union of basic open sets.*

Proof. First, if $V \in \mathcal{B}$, observe that V is open. By definition of open set in $\mathcal{T}_{\mathcal{B}}$, we need to show that for all $x \in V$, there is a basic open set B with $x \in B \subset V$. But this is tautological: take $B = V$.

Next, because arbitrary unions of open sets are open, it follows that every union of basic open sets is open. This gives one implication of our proposition.

On the other hand, suppose $U \subset X$ is open. This means that for each $x \in U$, there is some basic open set $x \in B_x \subset U$.

Now take the union of all of these basic open sets: set $V = \bigcup_{x \in U} B_x$. Because each $B_x \subset U$, we have $V \subset U$. On the other hand, for each $x \in U$, we have $x \in B_x \subset V$, so that $U \subset V$. Thus U is a union of open sets, hence open. $\square$

The trick we used in the last line above is very useful and in a moment we will name it. This in hand, let's demonstrate a rather tautological example.

Example 22. Let $(X, \mathcal{T})$ be a topological space. Then $\mathcal{T}$ is a basis, which generates its own topology.

That $\mathcal{T}$ is a basis is almost tautological (I will let you check this). That it generates the same topology follows from the previous proposition. In fact, this is so useful that we'll state it as a corollary.

Corollary 23 (Locally open implies open)**.** *Let X be a topological space. Suppose $S \subset X$ is a subset so that, for all $x \in S$, there exists an open set U_x with $x \in U_x \subset S$. Then S is open.*

Proposition 24 (Two bases give rise to the same topology iff they're open with respect to each other)**.** *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases on X, with $\mathcal{T}_i$ the topology they generate. If each $V \in \mathcal{B}_1$ is open with respect to $\mathcal{B}_2$, then $\mathcal{T}_1 \subset \mathcal{T}_2$. It follows that if the basic open sets of each $\mathcal{B}_1$ and $\mathcal{B}_2$ are open with respect to the other, they generate the same topology.*

Proof. If $U \in \mathcal{T}_{B_1}$, then U is a union of $\mathcal{B}_1$-basic open sets. Because each of these is $\mathcal{B}_2$-open, it follows that U is a union of $\mathcal{B}_2$-open sets, and hence $\mathcal{B}_2$-open itself, as desired. □

9/21: Operations on spaces

Subspaces

If you look at the statement of the Heine-Borel theorem in the Prerequisites document, you may spot something sly happening. We start with a statement about a *subset* of the Euclidean space $\mathbb{R}^n$ with the metric d_2, and then state a theorem about *that subset, as a metric space.*

To do that — to make $X \subset \mathbb{R}^n$ a metric space in its own right — we may *restrict the metric* to X, and the restricted function once again satisfies the axioms of a metric. (There is nothing particularly special about $\mathbb{R}^n$ here; if (X, d) is a metric space and $S \subset X$ is a subset, then $(S, d|_S)$ is again a metric space.

Our goal in this section is to explore an analagous construction for topological spaces. To do that, it is helpful to get some intuition from the metric context: how are the open sets in (X, d) and in $(S, d|_S)$ related?

To see this, think open balls. If $x \in S$, how are $B_r^X(x)$ and $B_r^S(x)$ related? Well,

$$B_r^S(x) = \{y \in X \mid d(x, y) < r \text{ and } y \in S\} = \{y \in X \mid d(x, y) < r\} \cap S = B_r^X(x) \cap S.$$

Now we know that open balls form a basis for the metric topology — if we set

$$\mathcal{B}_X = \{B_r^X(x) \mid x \in X, r > 0\}, \quad \mathcal{B}_S = \{B_r^S(x) \mid x \in S, r > 0\} = \{B \cap S \mid B \in \mathcal{B}_X\},$$

then we see that the basis for the topology on S is what you get when you intersect the basic open sets on X with S. More generally, you can show that an open set on S is of the form $U_S = U_X \cap S$ for some open set $U_X \subset X$.

This inspires the following definition, which works for arbitrary topological spaces.

Definition 17. *Let $(X, \mathcal{T})$ be a topological space. If $S \subset X$ is an arbitrary subset, we define the subspace topology on S to be*

$$\mathcal{T}_S = \{V \subset S \mid V = U_X \cap S \text{ for some } U_X \in \mathcal{T}\}.$$

That is, $V \subset S$ is open iff it can be written as the intersection $U \cap S$ for some open subset of X.

Check that this is a topology on S.

Let's start with something you might expect:

Lemma 25 (Inclusion of a subspace is continuous)**.** *Let X be a topological space and $S \subset X$ a subset equipped with the subspace topology*[4]*. Then the inclusion map $i : S \to X$ is continuous.*

Proof. If $U \subset X$ is any set, then $i^{-1}(U) = U \cap S$. If U is open in X, then (by definition of the subspace topology) $U \cap S$ is open in S. Therefore $i^{-1}(U)$ is open for any open set U, and thus i is continuous. ☐

While we're still early enough that we like to find ways to produce new topological spaces, the subspace topology is more than just a source of examples. It satisfies some very important properties. One of these properties, in fact, *determines the subspace topology completely*, and tells us what the 'point' of a subspace is. This is called a 'universal property'. These will be a theme today.

Suppose $f : X \to Y$ is a map of *sets*. If the image lands inside some subset $S \subset Y$, we may *restrict the codomain*: there is a function $g : X \to S$ defined by the same rule: $g(x) = f(x)$, and if $i : S \to X$ is the inclusion map, then $f = ig$.

(Remember that when I say '$f : X \to Y$ is a function', the domain and codomain are part of the *data* of that function: yes, g 'does the same thing', but it is a different function, because it has a different codomain.)

While this all probably sounds a little tautological when we're talking about set theory, it gets to be very handy when talking about spaces.

Theorem 26 (Universal property of the subspace topology)**.** *Suppose $f : X \to Y$ is a function between topological spaces, and $S \subset Y$ a subset equipped with the subspace topology. If $f(X) \subset S$ and $g : X \to S$ is the function obtained by restricting the codomain of f, then g is continuous if and only if f is continuous.*

[4]Soon, we'll just say 'S is a subspace of X'. You'll see that there's usually not much reason to consider any topology on a subset of X except the subspace topology.

Proof. If $g : X \to S$ is continuous, then we want to show that $f = ig : X \to S \to Y$ is as well. We showed i was continuous in the previous lemma, and we know that compositions of continuous functions are continuous, so it follows that f is continuous.

Now suppose f is continuous. Let's show g is, using the definition of continuity. If $U \subset S$ is an open set, we want to show that $g^{-1}(U)$ is open. By definition of the subspace topology, $U = V \cap S$ for $V \subset Y$ some open set.

Now

$$g^{-1}(U) = \{x \in X \mid g(x) \in U\} = \{x \in X \mid f(x) \in U\} = f^{-1}(U);$$

we want to show that this is open as a subset of X. We may rewrite this as

$$f^{-1}(U) = f^{-1}(V \cap S) = f^{-1}(V) \cap f^{-1}(S).$$

By the assumption that f is continuous, we know that $f^{-1}(V)$ is open. Because $f(X) \subset S$, we know that $f^{-1}(S) = X$. It follows that $g^{-1}(U) = f^{-1}(U) = f^{-1}(V)$, which is open as desired. □

Continuous maps to S are the same thing as continuous maps to X whose image lands in S. If you have a continuous map to X whose image lands in S, you may restrict the codomain to obtain a continuous map to S. Conversely, if you have a continuous map to S, you can compose with the inclusion into X to get a continuous map to X whose image lies inside of S.

This is extremely handy. For one specific trick, which we will use now and again: if you want a surjective continuous map for some argument you're making, but your map $f : X \to Y$ is very much not surjective — replace Y with $f(X)$ equipped with the subspace topology! For instance, do you remember the result that if $f : X \to Y$ is a continuous surjection, and X is separable, then Y is separable?

We have just proved, *for free*, that if $f : X \to Y$ is any continuous map and X is separable, then $f(X) \subset Y$ is separable when equipped with the subspace topology.

In fact, the property above uniquely determines the subspace topology. (This is something you can prove, but it is not obvious. If you want to think about it but get stuck, feel free to reach out to me.)

Products

Just like the previous section, we will start with a set-theoretic discussion and then move up to figure out what the right topological idea should be.

If Y and Z are sets, the Cartesian product $Y \times Z$ is the set of pairs $\{(y, z) \mid y \in Y, z \in Z\}$. An element of the Cartesian product is a choice of element of each of Y and Z.

If, furthermore, X is a set and $f : X \to Y \times Z$ is a function, it is uniquely determined by its components: $f(x) = (g(x), h(x))$, where $g : X \to Y$ and $h : X \to Z$ are functions.

That is: **A function to the Cartesian product of Y and Z is the same data as a function to each of Y and Z.**

Our goal now is to improve the above statement so that it holds true just as well for *spaces*, with *continuous maps* replacing the arbitrary functions above. To do that, we need a rule to produce a topology on $X \times Y$ given a topology on X and Y.

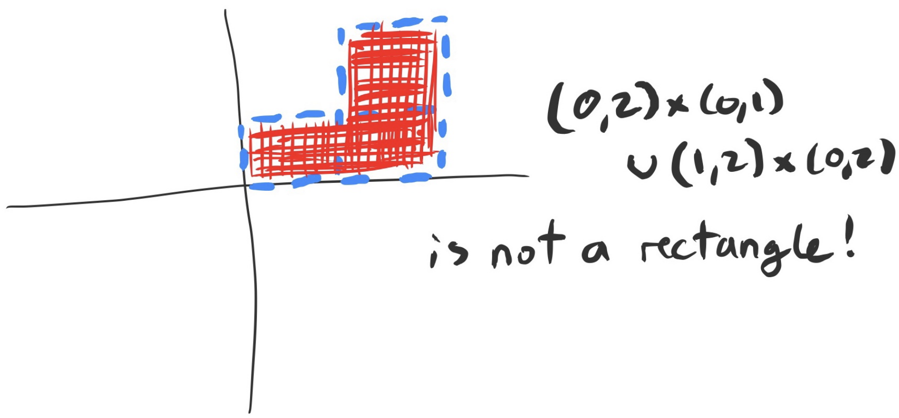
Before giving a formal definition, let me make a claim (which is hopefully intuitive, at least by comparison to $\mathbb{R}^2$). Whatever this 'product topology is', it should satisfy the following supposition[5]:

- If X and Y are topological spaces and $U \subset X, V \subset Y$ are open sets, then $U \times V \subset X \times Y$ should be open in the product topology.

A naive first guess is that the collection $\{U \times V \mid U \subset X, \quad V \subset Y \text{ open}\}$ forms a topology. This is not true: the union axiom fails.

But we can still take the smallest topology so that these sets are open.

[5]mathematicians like to say 'ansatz'



Definition 18. *Let X and Y be topological spaces. The* product basis *is*

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

The product topology *on $X \times Y$ is the topology $\mathcal{T}_{X\times Y}$ generated by the product basis.*

That is, a subset of $X \times Y$ is open in the product topology if and only if it may be written as a union of sets of the form $U \times V$, where both factors are open.

What we're going to do now is largely in the same spirit as what we did for subspaces. First, we'll check that a couple of standard maps are continuous; then we'll characterize continuous maps to $X \times Y$.

Lemma 27 (Projection is continuous in the product topology)**.** *If X and Y are topological spaces and $X \times Y$ is given the product topology, then $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are both continuous.*

Proof. If $U \subset X$ is open, then $\pi_1^{-1}(U) = U \times Y$, which is a product of open sets, hence open in the product topology. Thus π_1 is continuous. The same argument applied for π_2. □

Theorem 28 (Universal property of the product topology)**.** *Let X, Y, Z be topological spaces. If $g : X \to Y$ and $h : X \to Z$ are functions, then the function*

$$f : X \to Y \times Z \quad f(x) = (g(x), h(x))$$

is continuous if and only if f and g are.

Thus, **A continuous map to the product $Y \times Z$ is the same data as a pair of continuous maps: one to Y, one to Z.** It is difficult to understate how useful this will be: there is at least one much later homework exercise whose proof is applying this fact repeatedly to derive an interesting result. (I am being intentionally vague.) This also means that you can sometimes avoid thinking about what the open sets on a product space look like. If you want to check that a map to a product is continuous, you just need to check that its factors are.

Proof. The forward direction is the simpler one. If f is continuous, then $g = \pi_1 f$ and $h = \pi_2 f$ are continuous, as the π_i are continuous by the previous lemma and composites of continuous functions are continuous.

Conversely, suppose that g and h are continuous; we want to show that f is continuous. Let's first show that the inverse image of basic open sets are open. We have

$$f^{-1}(U \times V) = \{x \in X \mid f(x) \in U \times V\} = \{x \in X \mid g(x) \in U, \;\; h(x) \in V\} = g^{-1}(U) \cap h^{-1}(V).$$

Because g and h are continuous, both $g^{-1}(U)$ and $h^{-1}(V)$ are open. Because intersections of *finite* collections of open sets are again open, the intesection above is open, and hence $f^{-1}(U \times V)$ is open, as desired.

If $W \subset Y \times Z$ is an open set in general, we know that we may write

$$W = \bigcup_{i \in I} U_i \times V_i$$

for some index set I and some collection of open sets $U_i \subset Y$ and $V_i \subset Z$. Then

$$f^{-1}(W) = f^{-1}\left(\bigcup_{i\in I} U_i \times V_i\right) = \bigcup_{i\in I} f^{-1}(U_i \times V_i).$$

We have already established that each $f^{-1}(U_i \times V_i)$ is open, and unions of open sets are open, so it follows that $f^{-1}(W)$ is open, as desied; we have shown that f is continuous. ☐

Incidentally, in the proof above we implicitly showed a nice little fact.

Lemma 29 (Continuity may be checked on a basis)**.** *Suppose $\mathcal{B}$ is a basis generating the topology on a space Y. If $f : X \to Y$ is a function, then f is continuous if and only if $f^{-1}(B)$ is open for all $B \in \mathcal{B}$.*

The proof is contained in the argument above; just like 'locally open implies open', this often saves us a line or two.

Disjoint unions

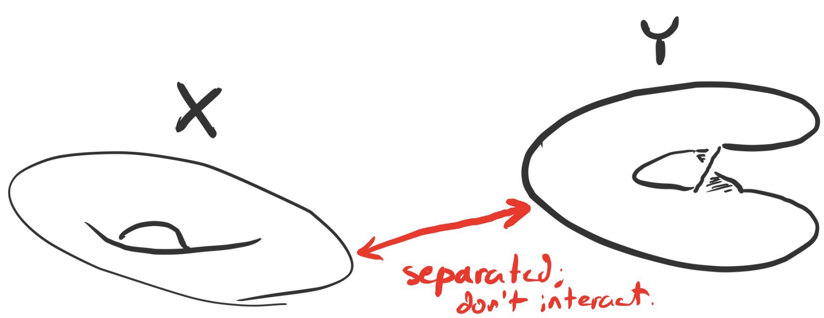
We characterized continuous maps *to* a product of topological spaces above: a continuous map $X \to Y \times Z$ is the same data as a continuous map to each of Y and Z.

Question. Is there a construction which 'goes the other way'? That is, given two spaces X and Y, is there a mystery space Q so that continuous maps *from* Q are the same data as a map from each of X and Y?

The category theorists call such spaces *coproducts*: the 'co' indicates that we flip the direction we're going around. (Instead of characterizing maps TO a product, we characterize maps FROM a coproduct.) We won't use that language in this course.

Unfortunately, constructing Q involves some set-theoretic pain (we have to make some 'arbitrary-looking' choices along the way to make this make sense). One of your homework exercises will show that these choices don't really matter, and that Q is determined by the property in the question above.

The idea is to take a copy of X and a copy of Y separated some distance from each other.



A continuous map from such a space should restrict to a continuous map on X and a continuous map on Y, and because these spaces are separated from one another, I expect that any continuous maps on X and Y give rise to a continuous map on this 'disjoint union'; the copy of X and the copy of Y shouldn't interact.

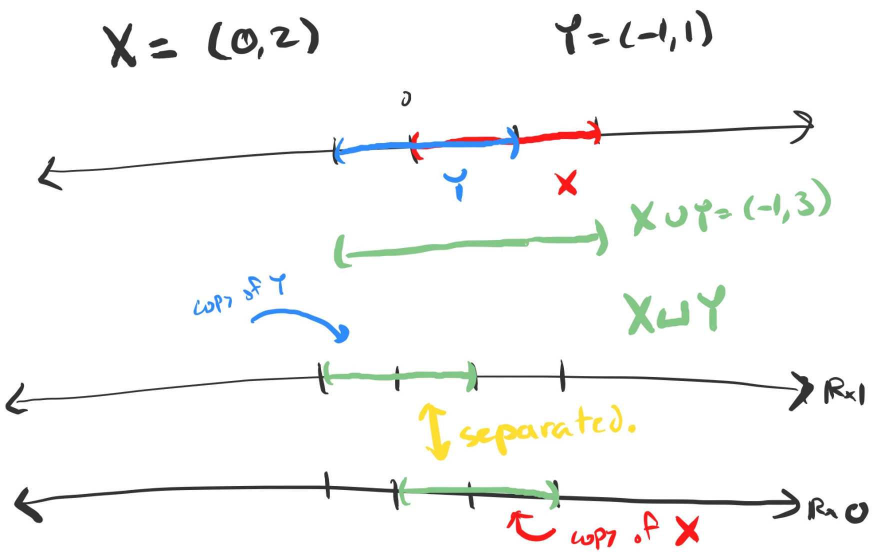
Problem. Given two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, the union $X \cup Y$ of the underlying sets makes sense, but may be nonempty. (For instance, what if $X = (-\infty, 1]$ and $Y = [-1, \infty)$? We want these to be separated from each other.

Definition 19. *Let X and Y be sets. We say their* disjoint union[6] *is*

$$X \sqcup Y = X \times \{0\} \cup Y \times \{1\} \subset (X \cup Y) \times \{0, 1\}.$$

[6] LaTeX code: \sqcup

If $S \subset X$ and $T \subset Y$ are subsets, we write $S \sqcup T \subset X \sqcup Y$ for the subset $S \times \{0\} \cup T \times \{1\}$. Convince yourself that every subset of $X \sqcup Y$ has this form.



This is kind of irritating. First off, the $\{0,1\}$ is not obviously important here; we could replace it with any other 2-point set (or any other many-point set and pick our two favorite points in it), and we get technically slightly different results depending on which set we use. But once we define the topology the resulting spaces are *canonically homeomorphic*, as you will see in your homework; so while this might give different results depending on the way we set it up, the resulting objects don't differ in any 'meaningful' way.

Definition 20. *Given topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, we say their* disjoint union *is the space $(X \sqcup Y, \mathcal{T}_\sqcup)$, where a set is open in $X \sqcup Y$ if it is of the form $U \sqcup V$ for open subsets of X and Y; that is,*

$$\mathcal{T}_\sqcup = \{U \sqcup V \mid U \in \mathcal{T}_X, \ \ V \in \mathcal{T}_Y\}.$$

You should check that this indeed forms a topology.

Now we will characterize the continuous maps *out* of the disjoint-union, following the same patterns as before.

Lemma 30 (Inclusion into disjoint-union is continuous)**.** *Let X and Y be topological spaces. Then the natural inclusion maps $i_1 : X \to X \sqcup Y$ and $i_2 : Y \to X \sqcup Y$ are both continuous.*

Proof. Let's check this for i_1. An open set in $X \sqcup Y$ is of the form $U \sqcup V$ for U, V open in X, Y, respectively. Then

$$i_1^{-1}(U \sqcup V) = U,$$

because i_1 identifies X with its image, and V lies inside Y (which is disjoint from X).

Because $U \subset X$ is open, we see that i_1 is continuous. The argument for i_2 is similar. □

For two sets X and Y, a function $f : X \sqcup Y \to Z$ is the same data as knowing the two functions $f|_X - fi_1 : X \to Z$ and $f|_Y = fi_2 : Y \to Z$. (Can you convince yourself of this? That this is true uses in an essential way that X does not intersect Y inside of $X \sqcup Y$.)

Theorem 31 (Universal property of the disjoint union)**.** *The map $f : X \sqcup Y \to Z$ is continuous with respect to the disjoint-union topology if and only if both $f|_X$ and $f|_Y$ are continuous.*

Proof. If f is continuous, then $f|_X = fi_1$ and $f|_Y = fi_2$ are, by the previous lemma and the fact that composites of continuous functions are continuous.

Conversely, suppose $f|_X$ and $f|_Y$ is continuous. We should show that f is. If $W \subset Z$ is an open subset, consider $f^{-1}(W) \subset X \sqcup Y$. This can be written as a disjoint union

$$f^{-1}(W) = S_X \sqcup S_Y,$$

where $S_X \subset X$ and $S_Y \subset Y$.

An open subset of $X \sqcup Y$ is of the form $U \sqcup V$ for some open subsets U, V of X and Y respectively. So we want to show that S_X and S_Y are open.

But

$$S_X = f^{-1}(W) \cap X = f|_X^{-1}(W), \quad S_Y = f^{-1}(W) \cap Y = f|_Y^{-1}(W),$$

and so S_X and S_Y are open by the assumption that $f|_X$ and $f|_Y$ are continuous. □

So we've characterized continuous maps out of the disjoint union: they're the same data as a continuous map from each of X and Y; if you want to check that a map from $X \sqcup Y$ is continuous, you just need to check it on each factor.

9/23: (Path)-connected spaces

Topological properties

I mentioned early on that to show two spaces are homeomorphic, you “just” need to construct a homeomorphism between them; but to show that they are *not* homeomorphic is much more difficult.

The next few weeks of the course are focused on what are called topological properties, which are precisely the tools we need to *distinguish* between non-homeomorphic spaces.

Definition 21. *A property P of topological spaces is called a topological property if:*

$$P(X) \text{ is true, and } X \text{ is homeomorphic to } Y \implies P(Y) \text{ is true.}$$

This is very vague, but we'll see what it means in practice by way of examples. (If you really want to make what a ‘property’ is more precise, imagine it's a way of assigning either True or False to each topological space X.)

Example 23. You showed on a homework exercise that ‘being discrete’ is a topological property. As another example, consider the property that *all points of X form closed sets*; equivalently, all sets of the form $X\backslash\{x\}$ are open. You can check that this is also a topological property.

Here is another example.

Definition 22. *We say that a topological space $(X, \mathcal{T})$ is metrizable if $\mathcal{T}$ is the metric topology induced by some metric d on X.*

Proposition 32. *Metrizability is a topological property.*

Proof. Suppose $(X, \mathcal{T}_d)$ is a metrizable topological space, with $\mathcal{T}_d$ induced by the metric d on X. Suppose $f : Y \to X$ is a homeomorphism. Consider the distance $d' : Y \times Y \to [0, \infty)$ given by $d'(a, b) = d(f(a), f(b))$. That this is a metric follows because f is a bijection and d is a metric.

Now given an open subset $U \subset Y$, we show that it is d'-open: because f is a homeomorphism, $f(U)$ is open in X, and hence for all $f(y) = x \in f(U)$, there is an open ball $B_r^d(x) \subset f(U)$. Now, by definition of the metric on Y, we have that $f^{-1}(B_r^d(x)) = B_r^{d'}(y)$; and because f^{-1} preserves containment, we have that

$$B_r^{d'}(y) \subset f^{-1}(f(U)) = U,$$

where the last equality follows because f is a bijection.

Conversely, if a subset $V \subset Y$ is d'-open, the same argument run in reverse shows that $f(V)$ is d-open in X, and so (because the topology on X is that induced by d) we have in fact that $f(V)$ is open. Because f is continuous, we have that $f^{-1}(f(V))$ is open; because f is a bijection, we have that $V = f^{-1}(f(V))$, and hence V is open.

So we have proven that the topology on Y is in fact the topology induced by d', and in particular, Y is a metrizable topological space. □

Remark 24. It is surprisingly difficult to write down a property of topological spaces which is *not* a topological property. Try to think of one; if you think you have one, tell me!

Next we will focus on a particularly important topological property.

Connectedness

Last time, given topological spaces X and Y, we defined a topological space $X \sqcup Y$ called the *disjoint union* of X and Y; we imagine it as being a single copy of each of X and Y, separated from each other and not at all interacting. One way to phrase the idea that these two pieces don't interact at all is that (so long as X and Y are both nonempty) $X \sqcup Y$ can be written as the union of two nonempty, non-intersecting open sets:

$$X \sqcup Y = (X \sqcup \varnothing) \cup (\varnothing \sqcup Y).$$

In fact, any space satisfying this property is a disjoint union:

Proposition 33 (Disconnected spaces are disjoint unions)**.** *Let X be a topological space, and suppose $X = U \cup V$, where U, V are non-intersecting open subsets ($U \cap V = \varnothing$). Equipping U and V with their subspace topologies, there is a canonical homeomorphism*

$$i = (i_U \sqcup i_V) : U \sqcup V \to X.$$

Proof. Write $i_U : U \hookrightarrow X$ and similarly i_V for the inclusions of these subspaces into X; the first thing we proved about the subspace topology is that these maps are continuous. By the Lemma above, $i = i_U \sqcup i_V$ is a continuous map.

Because $U \cap V = \varnothing$, the map i is injective; and because $U \cup V = X$, the map i is surjective.

Let's prove that i is a homeomorphism. Because i is a continuous bijection, it suffices to show that i is open (which, when i is a bijection, is equivalent to saying that i^{-1} is continuous).

Suppose we have an open subset of $U \sqcup V$; by definition of the disjoint union topology, such a subset may be written as $W_1 \sqcup W_2$, where $W_1 \subset U$ and $W_2 \subset V$ are open subsets. Further, by definition of the subspace topology, $W_1 = U \cap W$ for some open subset $W \subset X$; because U and W are both open, it follows that W_1 itself is open in X. Similarly, W_2 is open in X.

Then $i(W_1 \sqcup W_2) = W_1 \cup W_2$, which is the union of open sets in X, and hence open. So i is an open map. □

Warning. The above proof used in an essential way that U and V are open subsets. It is not true, for instance, that $\mathbb{R}$ (with the standard topology) is homeomorphic too $(-\infty, 0] \sqcup (0, \infty)$, even though these two sets don't intersect. This sometimes causes confusion, since we've written $\mathbb{R}$ as a *union* of *disjoint sets*; but that doesn't mean that $\mathbb{R}$ has the disjoint union topology with respect to these two sets.

Whatever connectedness means, $U \sqcup V$ should *not* be connected whenever U and V are nonempty: I imagine this as being two totally separate blobs. The previous proposition inspires the following definition.

Definition 23. *We say that a topological space X is disconnected if there are* ***nonempty*** *open subsets $U, V \subset X$ with $U \cup V = X$ and $U \cap V = \varnothing$.*

We say that X is connected if it is not disconnected. That is, for any nonempty open subsets $U, V \subset X$ with $U \cup V = X$, we have $U \cap V \neq \varnothing$.

Remark 25. With this definition, the empty topological space is connected. Some people don't like this. I don't really care either way.

By the previous proposition, a space is disconnected iff it is homeomorphic to a disjoint union of two **nonempty** spaces. Because homeomorphism is an equivalence relation, this is a topological property. (It is straightforward to see that connectedness/disconnectedness is a topological property without using that proposition, too.)

The following is a commonly stated and commonly used equivalent property.

Proposition 34 (Clopen characterization of connectedness)**.** *A space X is connected iff the only subsets of X which are simultaneously closed and open are $\varnothing$ and X itself.*

Proof. If X is connected, and $U \subset X$ is both closed and open, then $U \cup (X \backslash U)$ is a decomposition of X into non-intersecting open sets; it thus follows from connectedness that one of them is empty, and so either U is empty or $U = X$.

Conversely, suppose X is disconnected, so $X = U \cup V$ for U, V non-intersecting, nonempty open sets. Then $U = X \backslash V$ is both open and closed; but by assumption $U \neq \varnothing$ and $U \neq X$ (because V is nonempty), so X has nonempty proper subsets which are both open and closed. □

The following shows up again and again, and is one of the crucial properties of connectedness; it gives another way of seeing that connectedness is a topological property.

Proposition 35 (Images of connected spaces are connected)**.** *Let $f : X \to Y$ be a continuous map. If X is connected, then its image $f(X)$ (equipped with the subspace topology) is connected.*

Proof. It suffices to prove that if f is a continuous surjection, and X is connected, then Y is connected. We reduce to this case by restricting the codomain of f in the original statement to $f(X)$ equipped with the subspace topology; the universal property of the subspace topology implies that $f : X \to f(X)$ is still continuous.

So suppose f is a continuous surjection. To show that Y is connected, it suffices to show that for any nonempty open subsets $U, V \subset Y$ so that $U \cup V = Y$, we have $U \cap V \neq \varnothing$. Now, both $f^{-1}(U)$ and $f^{-1}(V)$ are open by continuity, and neither are nonempty because f is surjective. Because $U \cup V = Y$, and inverse image preserves unions, we have that $f^{-1}(U) \cup f^{-1}(V) = X$. Because X is connected, it follows that

$$\varnothing \neq f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V),$$

and hence that $U \cap V \neq \varnothing$, as desired. □

For free as a result, we get the following equivalent condition.

Corollary 36 (Functional characterization of connectedness)**.** *A topological space X is connected iff every continuous map to a discrete space is constant.*

Proof. If Y is discrete and $S \subset Y$ is any subset, then the subspace topology on S is again the discret topology; further, discrete spaces are connected if and only if they consist of a single point (otherwise you can take $\{x\} \cup (Y \backslash \{x\})$ to write Y as a union of disjoint open sets).

So if X is connected and $f : X \to Y_{\text{disc}}$ is continuous, then by the previous proposition $f(X)$ must be a single point, so f is constant.

Conversely, if $X = U \cup V$ where U and V are open and nonempty, with $U \cap V = \varnothing$, then define $f : X \to \{0, 1\}_{\text{disc}}$ by

$$f(x) = \begin{cases} 0 & x \in U \\ 1 & x \in V \end{cases}$$

Because $f^{-1}(0) = U$ is open and $f^{-1}(1) = V$ is open, it follows that f is continuous; because U and V are nonempty, it follows that f is nonconstant. So if X is disconnected, there exists a nonconstant continuous map to a discrete space. □

The following fundamental result is our first example of a familiar space which we can prove is connected. It will later inspire a new definition, which is easier to check than connectedness.

Somehow this is one of the first times so far we feel like we need to get into the nitty gritty, and that's because the only way to understand the topology of $\mathbb{R}$ and its subsets is explicitly in terms of the metric and the completeness of the reals.

Theorem 37. *The unit interval $[0, 1]$, with the standard topology, is a connected topological space.*

Proof. Suppose $[0, 1] = U \cup V$ where U, V are nonempty open sets. Without loss of generality, suppose $0 \in U$. In the standard topology on $[0, 1]$, this implies that $[0, r/2] \subset [0, r) \subset U$ for some $r > 0$.

Consider the set $B = \{r \in [0, 1] \mid [0, r] \subset U\}$. Because all elements of B are at most one (and hence there is an upper bound on B), by the least upper bound property of the real numbers (equivalent to completeness), there is a least upper bound $x \in \mathbb{R}$ for our set B. Noting that if $r \in B$ and $0 < r' < r$, then $r' \in B$ as well — because $[0, r'] \subset [0, r]$ — it follows that $[0, r] \subset B$ for all $r < x$, and hence that $[0, x) \subset B$.

We will complete the proof by casework, depending on whether or not $x \in V$ and whether or not $x = 1$.

If $x \in V$, then by the openness of V there is some $(x - s, x + s) \cap [0, 1] \subset V$ as well; in particular, $(x - s, x) \subset U \cap V$, and so $U \cap V$ is nonempty.

So suppose $x \in U$. If $x = 1$, then $U = [0, 1]$, and thus by nonemptiness of V we have $U \cap V = \varnothing$.

So suppose $x < 1$ (and still $x \in U$). Then because x is the least upper bound of B, there exists $y \notin B$ with $x < y$, and $y - x$ arbitrarily small; in particular, we may take $y < 1$, Note further that by definition of B, we see that there are points in $[x, y]$ which are *not in* U, and hence which are in V — so there are points in V which are arbitrarily close to x.

Because $x \in U$ and U is open, it follows that $(x - s, x + s) \subset U$ also contains points in V for any sufficiently small $s > 0$, and so $U \cap V$ is nonempty.

In all four cases, we've seen that $U \cap V$ is nonempty, as desired. □

Path-connected spaces

The fact that the interval $[0,1]$ is connected allows us to define a condition which is usually easier to check than, and implies, connectedness.

Definition 24. *Let X be a topological space. We say that X is path-connected if, for any two points $p, q \in X$, there is a continuous map $\gamma : [0,1] \to X$ with $\gamma(0) = p$ and $\gamma(1) = q$.*

This is perhaps what most people's immediate intuitive understanding is of what 'connectedness' should be: you can walk from anywhere on X to any other point on X. And indeed:

Proposition 38. *If X is path-connected, then X is connected.*

Proof. Suppose $f : X \to Y_{\text{disc}}$ is continuous; we aim to show that f is constant. For pick any two points $p, q \in X$, and choose a continuous map $\gamma : [0,1] \to X$ with $\gamma(0) = p$ and $\gamma(1) = q$. because γ is continuous, the composite $f\gamma : [0,1] \to Y_{\text{disc}}$ is continuous, too; and because $[0,1]$ is connected, $f\gamma$ is constant. In particular,

$$f(p) = f(\gamma(0)) = f(\gamma(1)) = f(q).$$

Because $p, q \in X$ were chosen arbitrarily, we see that f takes the same value at any two points; hence f is constant. □

Images of path-connected spaces are path-connected; it follows that path-connectedness is a topological property.

Proposition 39 (Image of path-connected space is path-connected)**.** *If X is path-connected, and $f : X \to Y$ is a continuous map, then the image $f(X) \subset Y$ (equipped with the subspace topology) is path-connected.*

Proof. By considering $f(X)$ with the subspace topology and restricting the codomain, it suffices to assume f is a continuous surjection.

If $p, q \in Y$, we need to show there is a path connecting them. Choose $a, b \in X$ with $f(a) = p$ and $f(b) = q$; because X is path-connected, we may choose a continuous path $\gamma : [0,1] \to X$ with $\gamma(0) = a$ and $\gamma(1) = b$. Then the compoiste $f\gamma : [0,1] \to Y$ is a continuous path with $f(\gamma(0)) = f(a) = p$ and $f(\gamma(1)) = f(b) = q$, as desired. □

And lastly, we may generate new path-connected spaces by taking products. (If you have not yet done the homework exercise about products over arbitrary index sets, you may read this as a proof about $X = X_1 \times \cdots \times X_n$.)

Proposition 40 (Product of path-connected spaces is path-connected)**.** *If X_i is a family of path-connected spaces, then the product $X = \bigtimes_{i \in I} X_i$ is path-connected.*

Proof. Pick any points $p = (p_i)_{i \in I}$ and $q = (q_i)_{i \in I}$ in X. For each $i \in I$, we know that there is a continuous map $\gamma_i : [0,1] \to X_i$ with $\gamma_i(0) = p_i$ and $\gamma_i(1) = q_i$. We also know that the map $\gamma = (\gamma_i)_{i \in I} : [0,1] \to X$, whose components are the maps γ_i, is continuous; this is the point of the product topology! Thus γ is a path connecting p and q. □

This provides us with a great many examples of path-connected spaces. For instance, any real interval $(a,b), [a,b)$, or $[a,b]$ is path-connected; if $p < q$ are points in the interval, then just take the path to be $\gamma(t) = (1-t)p + tq$. (This lies in $[p,q]$, hence stays in the given interval.) In particular, the real numbers $\mathbb{R}_{\text{std}}$ are connected, hence path-connected.

We conclude that $\mathbb{R}^n$ is path-connected in the standard topology for any n; as is the space $\mathbb{R}^\omega$ of arbitrary sequences, with the product topology, introduced in the homework.

Example 26. The space $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ is path-connected: if $p = (\cos\theta, \sin\theta)$ and $q = (\cos\psi, \sin\psi)$ are points on the circle (with, say, $\theta < \psi$), there is a path between them given by

$$\gamma(t) = (\cos(t\psi + (1-t)\theta), \sin(t\psi + (1-t)\theta)).$$

The n-fold product $(S^1)^n$ is called the *n-dimensional torus T^n*; this is also connected. (You know T^2 as the surface of a donut.)

Just like the circle is connected, spheres are connected.

Example 27. Consider $S^n = \{x \in \mathbb{R}^{n+1} \mid |x|_2 = 1\}$, the space of points of (Euclidean) distance 1 from the origin; this is the sphere of dimension n. This is equipped with the subspace topology as a subset of $\mathbb{R}^{n+1}$; equivalently, with the metric topology induced by the restriction of the Euclidean metric.

Write N for the north pole of the sphere, $N = (0, \cdots, 0, 1)$. First, we will prove that $S^n \backslash \{N\}$ (with the subspace topology) is homeomorphic to $\mathbb{R}^n$.

First, consider the map $f : S^n \backslash \{N\} \to \mathbb{R}^n$, given by

$$f(x_1, \cdots, x_{n+1}) = \left(\frac{x_1}{1 - x_{n+1}}, \cdots, \frac{x_n}{1 - x_{n+1}} \right).$$

Each component of the map f is continuous as a function on $\mathbb{R}^{n+1} \backslash \{x \mid x_{n+1} = 1\}$; more generally, if $f, g : X \to \mathbb{R}$ are continuous functions so that $g \neq 0$ on X, the map f/g is again a continuous function on X. (This follows because $x/y : \mathbb{R}^2 \backslash \{(x, y) \mid y = 0\} \to \mathbb{R}$ is continuous.) It follows that it each component is continuous when restricted to the sphere $S^n \backslash \{N\}$. Because the standard topology on $\mathbb{R}^n$ is the product topology (with respect to the standard topology on each factor), a map to $\mathbb{R}^n$ is continuous iff all of its components are; so we have seen that f is continuous.

The map f is called *stereographic projection*, and has a geometric interpretation: $f(x)$ is the unique point in the plane $x_{n+1} = 0$ on the line through both x and N.

One may check that f is a bijection, and that its inverse is $g : \mathbb{R}^n \to S^n \backslash \{N\}$, given explicitly by

$$g(x_1, \cdots, x_n) = \left(\frac{2x_1}{1 + x_1^2 + \cdots + x_n^2}, \cdots, \frac{2x_n}{1 + x_1^2 + \cdots + x_n^2}, \frac{-1 + x_1^2 + \cdots + x_n^2}{1 + x_1^2 + \cdots + x_n^2} \right).$$

A rather tedious explicit computation shows that g is indeed inverse to f, and in particular has image in $S^n \backslash \{N\}$. Again, the components of g are continuous, because they're quotients of continuous functions whose denominator is nonzero, so $g' : \mathbb{R}^n \to \mathbb{R}^{n+1}$ is continuous; because the image of g' lies in $S^n \backslash \{N\}$, and $S^n \backslash \{N\}$ is equipped with the subspace topology, it follows that g, which is obtained by restricting the codomain, is continuous.

Since f and g are both continuous and inverse to one another, we see that $S^n \backslash \{N\}$ is homeomorphic to $\mathbb{R}^n$. In particular, it's path-connected. A similar argument applies to $S^n \backslash \{S\}$, for $S = (0, \cdots, 0, -1)$ the south pole.

We have almost proved that S^n is path-connected: so long as one of p, q is neither N or S, we may connect the two by a path. So all that is left is to connect N and S by a path. This can be done explicitly: take $\gamma(t) = (\sin(\pi t), 0, \cdots, 0, \cos(\pi t))$; we have $\gamma(0) = N$ and $\gamma(1) = S$.

(In fact, you can check that if $X = S \cup T$ where S, T are path-connected subspaces, and $S \cap T$ is nonempty, then X is path-connected.)

9/28: Components and cut-points

Connected components

The next idea we explore is: to what degree can we decompose a space into connected pieces? This question inspires the following definition.

Definition 25. *Let X be a topological space. A connected component of X is a maximal connected subspace of X: a subset $S \subset X$ so that S is connected (when equipped with the subspace topology), and so that if $S \subset S'$ and S' is connected, we have $S = S'$.*

First we will need to see that such maximal connected subspaces even exist! To do that, we'll use the following lemma on the way.

Lemma 41 (Union of mutually intersecting connected sets is connected)**.** *Suppose X is the union of connected subspaces S_j, so that $S_j \cap S_k \neq \varnothing$ for every $j, k \in I$, where I is the index set for our collection of subspaces.*

Then X is connected.

Proof. We will show that every function $f : X \to Y_{\text{disc}}$ is constant; we showed earlier this is equivalent to connectedness.

First, note that the inclusion map $i_j : S_j \to X$ is continuous (with respect to the subspace topology on S_j) so the restriction $f|_{S_j} = fi_j : S_j \to Y_{\text{disc}}$ is continuous. Because S_j is connected, $f|_{S_j}$ is constant. Suppose f takes the value y_j on S_j.

Because $S_j \cap S_k \neq \varnothing$ for all j, k, we conclude that $y_j = y_k$ — for if $x \in S_j \cap S_k$, we see that

$$y_j = f|_{S_j}(x) = f(x) = f|_{S_k}(x) = y_k.$$

Since this holds for all j, k, we conclude that all of the y_j are equal to a single value, which may be called (say) y.

Because f is constant on each S_j, and every $\bigcup_{j \in I} S_j = X$, we conclude that f is constant on X, as desired. ☐

This in hand, we'll now show that not only do connected components exist, but they partition our space into connected pieces: that is, the connected components do not overlap.

Proposition 42. *Every point $x \in X$ lies in a unique connected component of X. In fact, every connected subspace of X lies in a unique connected component.*

Proof. Consider the set

$$C_x = \bigcup_{\substack{x \in S \subset X \\ S \text{ connected}}} S,$$

the union of all connected subspaces containing x.

First, note that C_x is connected by the previous lemma: we've written C_x as a union of connected subspaces any two of these subspaces S, S' intersect, because $x \in S \cap S'$.

Then note that C_x is maximal by its very definition. If $C_x \subset S$ and S is connected, then in particular $x \in S$; so S is one of the sets in the union that defines C_x, and in particular $C_x \subset S$ as well. Because we have both containments $C_x \subset S \subset C_x$, it follows that $C_x = S$; we have proved that C_x is maximal.

The same argument, without substantial change, shows that any connected subspace lies in some connected component.

Now what remains is to see that two distinct connected components C and C' do not intersect.

If C and C' are maximal connected components, and $C \cap C' \neq \varnothing$, it follows from the previous lemma that $C \cup C'$ is connected. Because C is a maximal connected space and we have $C \subset (C \cup C')$, it follows that we have $C = C \cup C'$ so $C' \subset C$. Similarly, because C' is maximal, we have that $C' = C \cup C'$ and thus $C \subset C'$. Therefore $C = C'$ as desired, and we've shown that any two distinct connected components are disjoint. ☐

Thus we may write X as a union of its connected components, which are pairwise non-intersecting. However, *this does not necessarily mean* that X is homeomorphic to the disjoint union of its components. (There is also an example in Example 6 below).

In proving that a set is a connected component, the following is a useful fact. I do not have a good name for this proposition or the previous one. If you have a recommendation, please tell me!

Proposition 43. *Suppose a space X is written as $X = U \cup V$ where U, V are open and non-intersecting. If $C \subset X$ is a connected subspace, then either $C \subset U$ or $C \subset V$.*

Proof. Let's phrse this in terms of maps to discrete spaces[7]. Consider the continuous map $f : X \to \{0,1\}_{\text{disc}}$ given by

$$f(x) = \begin{cases} 0 & x \in U \\ 1 & x \in V \end{cases}$$

Because $f(C)$ is connected, the map $f|_C$ is constant (and here we use that C is given the subspace topology to conclude that $f|_C = fi_C$ is continuous); so either $f(C) = 0$, in which case $C \subset U$, or $f(C) = 1$, in which case $C \subset V$, as desired. ☐

Example 28. The space $X = [0,1] \cup [3,4] \cup [5,6]$ has three connected components: $[0,1], [3,4]$, and $[5,6]$.

Note that these are all connected spaces. To see that $[0,1]$ is a maximal connected subspace, note that $X = ([0,1]) \cup ([3,4] \cup [5,6])$ decomposes X into non-intersecting open subspaces of X; because $[0,1]$ is not a subset of $[3,4] \cup [5,6]$, we conclude from the previous proposition that any connected set containing $[0,1]$ is in fact contained in $[0,1]$, so that $[0,1]$ is a maximal connected set — a connected component.

Similara arguments show that $[3,4]$ and $[5,6]$ are connected components.

Example 29. In the space $\mathbb{Q}$ of rational numbers, in the subspace topology, the connected components are all the singleton sets $\{x\}$.

To see this, suppose $C \subset \mathbb{Q}$ is connected and nonempty; say $x \in C$ is an element. Let $y < x$ be an **irrational** number; then $((-\infty, y) \cap \mathbb{Q}) \cup ((y, \infty) \cap \mathbb{Q}))$ is a decomposition of $\mathbb{Q}$ into disjoint open subsets. By the previous proposition, we see that because $x \in (y, \infty) \cap \mathbb{Q}$, we in fact have $C \subset (y, \infty) \cap \mathbb{Q}$. Because this argument holds for any $y < x$ irrational, we have

$$C \subset \bigcap_{\substack{y<x \\ y \notin \mathbb{Q}}} (y, \infty) \cap \mathbb{Q} = [x, \infty) \cap \mathbb{Q},$$

the last equality because we may choose y arbitrarily close to x.

A similar argument shows that C is contained in $(-\infty, y) \cap \mathbb{Q}$ for all $x < y$ and y irrational; it follows as before that $C \subset (-\infty, x]$. Putting these together, we conclude that

$$C \subset (-\infty, x] \cap [x, \infty) = \{x\};$$

because $x \in C$, we see that in fact $C = \{x\}$, and we've proved the desired statement.

The number of connected components is preserved under homeomorphisms, and so one can use this to distinguish between non-homeomorphic spaces (such as $(0,1)$ and $(0,1) \cup (1,2)$).

Proposition 44 (Number of components is a topological property)**.** *Write* Comp(X) *for the set of connected components of a topological space X.*

Then a continuous map $f : X \to Y$ induces a function

$$f_* : \text{Comp}(X) \to \text{Comp}(Y),$$

and if f is a homeomorphism, then f_* is a bijection.

In particular, homeomorphic spaces have the same number of connected components.

[7]This was the way of phrasing it that I was able to write most easily, but you can prove this using any of the equivalent definitions of connectedness

Proof. If $C \subset X$ is a connected component, its image $f(C)$ is connected, and so is contained in a unique connected component $C' \subset Y$. We define $f_*(C)$ to be the unique connected component containing C.

Note that if $f : X \to Y$ and $g : Y \to Z$ are continuous, then $(gf)_* = g_* f_*$; this is because if $f(C) \subset C'$, where C' is a connected component of Y, and $g(C') \subset C''$, where C'' is a connected component of Z.

It follows that $(gf)(C) \subset C''$; and in particular $(gf)_*(C) = C'' = g_*(f_*(C))$.

Now suppose f is a homeomorphism; so there is a continuous map

$$f^{-1} = g : Y \to X$$

with $1_X = gf$ and $1_Y = fg$.

Further, note that $(1_X)_*$ is the identity map on $\mathrm{Comp}(X)$, since after all $1_X(C) = C$.

It follows from everything said thus far that

$$f_* g_* = 1_{\mathrm{Comp}(Y)} \qquad g_* f_* = 1_{\mathrm{Comp}(X)}.$$

In particular, f_* and g_* are inverse to one another, so that f_* is a bijection. $\square$

In due time we will use this to distinguish between non-homeomorphic spaces. First, we will investigate the corresponding notion for path-connectedness.

Path components

Just like in the context of connectedness, there is a way to decompose a space into path-connected pieces.

Definition 26. *A subset P of X is called a path component of X if it is a maximal path-connected-subspace: P is path-connected, and if $P \subset P'$ for some path-connected P', we have $P = P'$.*

One can argue that a space is given as a union of its path components, all of which are disjoint, very similarly to how we did before. However, I'm interested in phrasing the argument a different way, because the phrasing that follows hides a useful trick that shows up quite a lot.

Lemma 45 (Path-connectedness is an equivalence relation)**.** *The relation between points on X, given by $p \sim q$ if there is a continuous path $\gamma : [0,1] \to X$ joining p and q, is an equivalence relation.*

Proof. Reflexivity follows because the path $\gamma_p : [0,1] \to X$ given by $\gamma_p(t) = p$ is continuous (all constant maps are).

Symmetry follows because the map $g : [0,1] \to [0,1]$ given by $g(t) = 1 - t$ is continuous; if $\gamma : [0,1] \to X$ is a continuous map with $\gamma(0) = p$ and $\gamma(1) = q$, so that $p \sim q$, then $(\gamma g) : [0,1] \to X$ has $(\gamma g)(0) = \gamma(1) = q$ and $(\gamma g)(1) = \gamma(0) = p$, so that $q \sim p$.

To prove transitivity, suppose $p \sim q$ and $q \sim r$; so there is a continuous map $\gamma : [0,1] \to X$ with $\gamma(0) = p$ and $\gamma(1) = q$, as well as a continuous map $\eta : [0,1] \to X$ with $\eta(0) = q$ and $\eta(1) = r$.

We define a map $(\gamma * \eta) : [0,1] \to X$, as follows, called the *concatenation* of γ and η:

$$(\gamma * \eta)(t) = \begin{cases} \gamma(2t) & 0 \leqslant t \leqslant 1/2; \\ \eta(2t-1) & 1/2 \leqslant t \leqslant 1 \end{cases}$$

We can prove that the concatenation is continuous quite explicitly. Suppose $U \subset X$ is open. Write $I = \gamma^{-1}(U)$ and $J = \eta^{-1}(U)$; these are open by continuity. We have that

$$(\gamma * \eta)^{-1}(U) = \{t \in [0,1/2] \mid 2t \in I\} \cup \{t \in [1/2,1] \mid 2t - 1 \in J\}.$$

We claim that this is open. For $t < 1/2$, because I is open, we know that

$$(2t - r, 2t + r) \cap [0,1) \subset I$$

for some $r > 0$; in particular

$$(t - r/2, t + r/2) \cap [0,1/2) \subset (\gamma * \eta)^{-1}(U).$$

Similarly, this follows for $t > 1/2$, because J is open in $[1/2, 1]$. It remains to check this for $t = 1/2$. Now becauue I is open and $1 \in I$, we have that $(1-r, 1] \subset I$ for some $r > 0$; and because J is open and $0 \in J$, we have that $[0, s) \subset J$ for some $s > 0$. Putting this together, we find that

$$(1/2 - r/2, 1/2 + s/2) \subset (\gamma * \eta)^{-1}(U)$$

for some $r, s > 0$. It follows that $(\gamma * \eta)^{-1}(U)$ is open, as desired.

So $\gamma * \eta$ is continuous, and hence there is a path connecting p and r, and $p \sim r$.

Thus we've proven that the existence of a path connecting p and q is an equivalence relation between points in X. ☐

Remark 30. Actually, the proof that $\gamma * \eta$ is continuous in fact can be modified to show something much more general. If you can cover X by closed sets so that $f : X \to Y$ is continuous on each of these closed sets, then f is continuous on all of X. We'll talk about this much later in the context of quotient spaces.

Now we can quickly prove the following.

Proposition 46 (Path-components are equivalence classes)**.** *Every path-connected subspace of X lies in a unique path-component; from this, it follows that the path-components are disjoint. Furthermore, the path component P_x containing $x \in X$ is the set $P_x = \{p \in X \mid p \sim x\}$ of points with a path connecting them to x — that is, it's the equivalence class containing x under the relation of path-connectedness.*

Proof. First, observe that P_x is a path-connected subspace: if $y, z \in P_x$, then $y \sim x$ and $z \sim x$; because $\sim$ is an equivalence relation, $x \sim z$, and so $y \sim x \sim z$ and by transitivity $y \sim z$. This means precisely there is a path connecting y and z. (Less formally: take the path connecting y to x, and the path connecting z to x; first follow the path from y to x, then concatenate it with the path going from z to x, run in reverse.)

Now if $P_x \subset P$, where P is path-connected, then in particular any $z \in P$ has $z \sim x$ (by definition of path-connectedness); so it follows that $z \in P_x$, and since z was an arbitrary element of P, that $P \subset P_x$. So $P = P_x$, as desired — the equivalence classes are path-components.

The rest follows from formal properties of equivalence relations: two distinct equivalence classes are disjoint, and every point is contained in an equivalence class. Further, if P is a path-connected subspace and $x \in P$, it follows that $P \subset P_x$, because every $z \in P$ has $z \sim x$ by assumption. (If you are not very familiar with equivalence relations, I suggest you check this.) ☐

Because path-connected spaces are connected, every path-component is contained in a unique connected component. It is not generally true that path-components and connected components coincide. (For instance, every path-component of $\mathbb{R}_{cc}$ is a singleton, but $\mathbb{R}_{cc}$ is connected. You can definitely prove this, but it's quite tricky: only try it if you want a challenge.)

Remark 31. Can you check confirm that the path-components of $\mathbb{Q}$ and $[0, 1] \cup [3, 4] \cup [5, 6]$ coincide with the connected components?

The following result is perhaps not terribly exciting, but the idea behind it shows up again and again; this idea is perhaps the key reason connectedness is so useful. The catchy name would be 'connected and locally path-connected implies path-connected'. However, 'locally path-connected' is a name already in use in topology for a stronger condition than the one below.

Proposition 47. *We say a topologicial space X* has path-connected neighborhoods *if for each point $x \in X$, there is an open set $x \in U \subset X$ with U path-connected.*

If X is connected and has path-connected neighborhoods, then X is path-connected.

Proof. I claim that if X has path-connected neighborhoods, its path-components are open sets. For if P is a path-component and $y \in P$, then P consists of all points which can reach y by a continiuous path. Since by assumption there is an open $U_y \subset X$ which is path-connected and contains y, it follows that $U_y \subset P$. It follows that P is "locally open", and we have proven that such sets are open. To remind you of the argument[8],

$$P = \bigcup_{y \in P} \{y\} \subset \bigcup_{y \in P} U_y \subset P,$$

[8]After this, I will just say 'locally open implies open'

because each $U_y \subset P$. Because P is a union of open sets, it is open.

I claim that each path-component is also *closed*. This follows because the path-components partition the space X; we have that

$$X \backslash P_x = \bigcup_{y \notin P_x} P_y.$$

(If this is not immediately clear, you should check it!) Because each P_y is open, we see that the complement is a union of opens, hence open; thus each path-component P_x is also closed.

Because X is connected, the only subsets which are both open and closed are the empty set and X itself. Because $x \in P_x$, it follows that $P_x = X$; every point can be connected to x by a path, and hence X itself is path-connected. $\square$

Remark 32. A more general result is true. There is a natural map

$$\sqcup_{X_C \in \mathrm{Comp}(X)} X_C \to X,$$

where the domain is the disjoint union of the path-components of X, and the map on each term $X_C \to X$ is the inclusion of that component.

This map is a homeomorphism if and only if X has path-connected neighborhoods. The above argument is vey close to giving a proof of this already; the hardest part is understanding how to clearly notate your work with infinite disjoint unions.

Cut points, and more non-homeomorphic spaces

The number of connected components distinguishes two spaces, but this is not a very subtle approach. For instance, all $\mathbb{R}^n$ are connected, so this certainly cannot distinguish between those. But a slightly more subtle idea helps a lot: what happens when we delete a point? How many connected components result from that?

We demonstrate this idea with an example.

Proposition 48. *$\mathbb{R}$ is not homeomorphic to $\mathbb{R}^n$, for any $n > 1$.*

Proof. Suppose, towards a contradiction, that $f : \mathbb{R} \to \mathbb{R}^n$ was a homeomorphism. Then, restricting the domain and codomain, $f' : \mathbb{R}\backslash\{0\} \to \mathbb{R}^n\backslash\{f(x)\}$ would also be a homeomorphism. (This follows from general properties of the subspace topology: do you see why?)

The domain of f' is disconnected: it has two connected components, $(-\infty, 0)$ and $(0, \infty)$. The codomain, however, is connected: translating by $-f(x)$ shows that $\mathbb{R}^n\backslash\{f(x)\}$ is homeomorphic to $\mathbb{R}^n\backslash\{0\}$; and this latter space is homeomorphic to $S^{n-1} \times (0, \infty)$, the inverse of the homeomorphism given by $(x, r) \mapsto rx$. The product of path-connected spaces is path-connected (and for $n > 1$, the sphere S^{n-1} is path-connected), so $\mathbb{R}^n\backslash\{0\}$ is path-connected. This could also be seen explicitly without using this homeomorphism by constructing by-hand all of the needed paths.

We have arrived at a contradiction: if f was a homeomorphism, then f' would be as well; but $\mathbb{R}\backslash\{0\}$ and $\mathbb{R}^n\backslash\{f(x)\}$ are not homeomorphic. It follows that there is no homeomorphism from $\mathbb{R}$ to $\mathbb{R}^n$. $\square$

We turn the idea behind this argument into a definition.

Definition 27. *We say that $x \in X$ is a cut point if $X\backslash\{x\}$ is disconnected. More precisely, for a cardinal number n, we say that $x \in X$ is an n-cut point if $X\backslash\{x\}$ has n connected components.*

Proposition 49 (Number of cut points is a topological property)**.** *If $f : X \to Y$ is a homeomorphism, then f restricts to a bijection $\mathrm{Cut}(X) \to \mathrm{Cut}(Y)$ between the sets of cut-points. What's more, for a cardinal number n, f restricts to a bijection $f_n : \mathrm{Cut}_n(X) \to \mathrm{Cut}_n(Y)$. (Here one should interpret $Cut_1(X)$ as the set of non-cut points.)*

Proof. If $f : X \to Y$ is a homeomorphism, and $x \in X$, then as in the previous proposition we see using general properties of the subspace topology that $f' : X\backslash\{x\} \to Y\backslash\{y\}$ is a homeomorphism. Thus $X\backslash\{x\}$ and $Y\backslash\{y\}$ have the same number of connected components; if x is an n-cut point, then y is too, and vice versa. Thus f sends n-cut points to n-cut points, and the map $f_n : \mathrm{Cut}_n(X) \to \mathrm{Cut}_n(Y)$ is surjective. Because f is an injection, f_n is as well, and thus f_n is a bijection. $\square$

It follows that the number of n-cut points is the same for any two homeomorphic spaces. This is what we used in the previous proposition: in $\mathbb{R}$ every point is a cut point (a 2-cut point), whereas in $\mathbb{R}^n$ no points are cut points.

Proposition 50. *No two of the spaces* $S^1, (0,1), [0,1)$, *and* $[0,1]$, *are homeomorphic, nor are they homeomorphic to the topological space underlying the letter* Y *(not including the points at the ends of the Y).*

Proof. S^1 has no cut points: deleting any point you get a space homeomorphic to the interval, which is connected.

In $(0,1)$, every point is a cut point (a 2-cut point). Thus $\mathrm{Cut}_2(0,1) = (0,1)$.

In $[0,1)$, the only point which is not a cut point (a 2-cut point) is 0. Thus $\mathrm{Cut}_1[0,1) = \{0\}$, and $\mathrm{Cut}_2[0,1) = (0,1)$.

In $[0,1]$, there are exactly two non-cut points: $\mathrm{Cut}_1[0,1] = \{0,1\}$, whereas $\mathrm{Cut}_2[0,1] = (0,1)$.

In Y, every point is a cut point, but not all of the same type! $\mathrm{Cut}_3(\mathsf{Y})$ consists of exactly one point (the center of the Y), whereas every other point is a 2-cut point.

Y is not homeomorphic to any of the other spaces because it's the only space with a 3-cut point. $[0,1]$ is not homeomorphic to any of the other spaces because it's the only space with exactly two non-cut points. $[0,1)$ is not homeomorphic to any of the other spaces because it's the only space with exactly one non-cut point; and S^1 is not homeomorphic to any of the other spaces because it's the only space with no cut points. $\square$

Remark 33. Can you come up with a definition of 'local cut point', and prove that the space given by the symbol

$$\otimes,$$

though of as a subspace of $\mathbb{R}^2$, is not homeomorphic to the circle, even though neither of them have any cut points?

9/30: Compact spaces

We start by naming a concept we've seen special cases of before.

Definition 28. *Let X be a topological space. A collection of open sets $\mathcal{U} = \{U_i\}_{i\in I}$ is called* an open cover *of X if $\bigcup_{i\in I} U_i = X$ — that is, if every point is in at least one of the open sets U_i.*

*A **subcover** is a subcollection $\mathcal{U}' \subset \mathcal{U}$ whose union is still X. Said another way, it is a subset $J \subset I$ of the indexing set so that the corresponding collection of open sets $\{U_j\}_{j\in J}$ stiill covers X — so every $x \in X$ is contained in at least one U_j, now with j in this smaller set J.*

Example 34. Every basis $\mathcal{B}$ for the topology on a space X gives an open cover for X; in fact, one of the axioms of a basis is that it gives a cover.

Example 35. Suppose that you have a metric space (X, d), and you know that some condition is true near each point (but 'near' might depend on the point). That is, for each $x \in X$, there is an $r(x) > 0$ so that your condition holds in $B_{r(x)}(x)$. Then $\mathcal{U} = \{B_{r(x)}(x)\}_{x\in X}$ is an open cover of X. (We will do constructions like this pretty often.)

The following definition is perhaps one of the most important definitions in topology. I'll try to give some intuition in a moment.

Definition 29. *A topological space X is* compact *if, for every open cover $\{U_i\}_{i\in I}$ of X, there is a* ***finite*** *subcover $U_1, \cdots, U_n$.*

Compactness is a kind of finiteness condition. If you know how to check something on a little open neighborhood of every point, compactness says you only need to check finitely many of those little open neighborhoods. The following statement might help with your intuition.

Proposition 51. *A discrete topological space X is compact if and only if it's finite.*

Proof. Suppose X is a finite topological space. Then it's certainly compact: enumerate the points as $x_1, \cdots, x_n$. If $\{U_i\}$ is an open cover, each x_j is contained in some U_j; so take $U_1, \cdots, U_n$ to be open sets in the cover, each containing the corresponding x_j. Then this is a finite subcover, because each x_j is in one U_j, and there are only finitely many open sets in the subcover.

This proves (more than) the forward direction: any finite space is compact.

For the reverse direction, suppose X is a compact discrete space. Now, because X is discrete, the collection

$$\mathcal{U} = \{\{x\}\}_{x\in X}$$

is an open cover of X. *No proper subcover of this collection covers* X, because if you remove the set $\{x\}$ from the cover, then x is no longer in any other sets in the cover!

The compactness condition thus says that this collection *itself* is the finite subcover: so X must be finite. □

The following characterization of compactness in metric spaces is very important for my intuition. This new notion of compactness is the one I'm used to from analysis.

Proposition 52 (Analysis characterizations of compactness)**.** *Let (X, d) be a metric space. A subset $S \subset X$ is* sequentially compact *if, for every sequence x_n in S, there is a* subsequence *x_{n_k} which converges to a point in S.*

Then S is sequentially compact if and only if S is compact.

The Heine-Borel theorem further states that if $X = (\mathbb{R}^d, d_2)$, then a subset is compact iff it is closed and bounded; or more generally, a metric space is compact iff it is complete and 'totally bounded'.

I will consider this a fact from analysis and take take it for granted as a way to generate examples of compact spaces. The ideas in this proof do not really recur; it's mainly useful to have this onhand to be able to observe that, say, Euclidean spheres are compact, Euclidean discs are compact, the Cantor set is compact, and so on.

The next proposition shows that compactness — a sort of finiteness condition — is preserved under taking images. The intuition that leads me to believe this is that the image of a finite set can only be at most as big as the original set was; so if compactness is a finiteness criterion, you'd expect the same.

Proposition 53 (Images of compact spaces are compact)**.** *If X is a compact space, and $f : X \to Y$ is a continuous map, then $f(X)$ is a compact space (when equipped with the subspace topology).*

Proof. By restricting the codomain if necessary (which results in a continuous map, because $f(X)$ is given the subspace topology), we may assume that f is surjective.

Suppose $\mathcal{U}$ is an open cover of Y — a collection of open sets whose union is Y. We may take the inverse image of this cover to obtain

$$f^{-1}\mathcal{U} = \{f^{-1}(U) \mid U \in \mathcal{U}\},$$

a cover of X. Because X is compact, there is a finite subcover of $f^{-1}\mathcal{U}$: a finite collection of open sets of the form $f^{-1}(U_j), 1 \leqslant j \leqslant n$, where each $U_j \in \mathcal{U}$, so that every $x \in X$ lies in one of these (finitely many) open sets.

Then I claim that U_j, for $1 \leqslant j \leqslant n$, is an open cover of Y. For if $y \in Y$, choose x with $f(x) = y$ (using surjectivity of f). We know that $x \in f^{-1}(U_j)$ for some j because these finitely many sets cover X; this means precisely that $y = f(x) \in U_j$, so that y is in one of our finitely many open sets.

We've shown that any open cover of Y has a finite subcover: that is, that Y is compact. □

A *closed* subspace (a closed subset, equipped with the subspace topology) of a compact space is again compact.

Proposition 54. *Let X be a compact space, and $C \subset X$ a closed subspace. Then C is also compact.*

Proof. The strategy is to start with an open cover of C; (sort of) turn it into an open cover of X by adding the complement of C, which is open; and then use that X is compact to reduce this to a finite cover, giving us a finite cover of C to begin with.

Let $\mathcal{U}$ be an open cover of C. Each $U \in \mathcal{U}$, because it is open in the subspace topology, is given as an intersection $U = C \cap V_U$ for some open set $V_U \subset X$. We use this to create an open cover of X: write

$$\mathcal{V} = \{V \subset X \mid V \text{ open, and } V \cap C \in \mathcal{U}\}.$$

Then $\mathcal{V} \cup \{X\backslash C\}$ is an open cover of X: every point not in C is in $X\backslash C$ (which is open, because C is closed). If $x \in C$, then because $\mathcal{U}$ is a cover of C, there is some $U \in \mathcal{U}$ containing x. By the definition of the subspace topology, there is some open $V \subset X$ with $V \cap C = U$. Using this, we see that $x \in V$ and (by definition of $\mathcal{V}$) that $x \in \cup_{V \in \mathcal{V}} V$. Thus every point of X lies in some element of the collection $\mathcal{V} \cup \{X\backslash C\}$: it is indeed an open cover.

Because X is compact, there is a finite subcover — so a finite set of $V_1, \cdots, V_n \in \mathcal{V}$ so that

$$V_1 \cup \cdots \cup V_n \cup (X\backslash C) = X.$$

Interesting with C, and using distributivity of intersection over union, we find that

$$(V_1 \cap C) \cup \cdots \cup (V_n \cap C) = C.$$

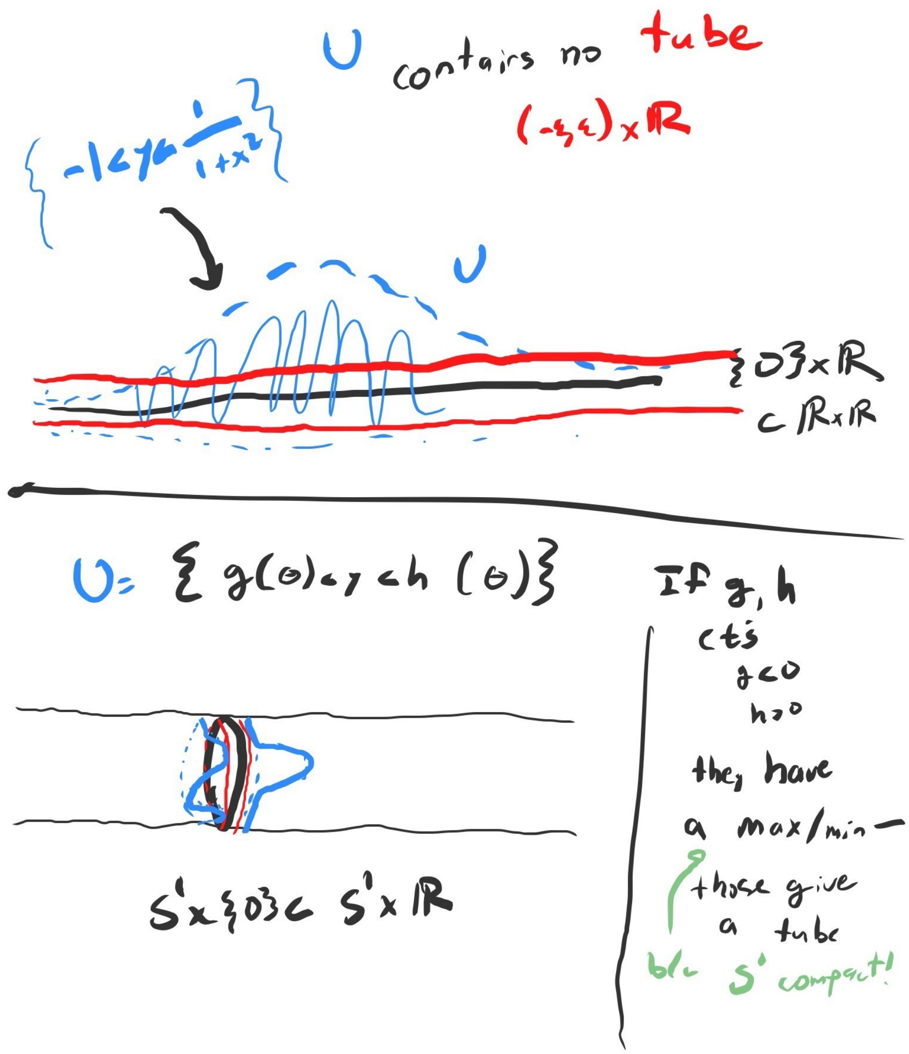
Since each $V_i \cap C$ is, by definition, an element of the collection $\mathcal{U}$, we've found a finite subcover of $\mathcal{U}$. □

To help build your intuition with compact spaces and the kind of way you'll use the property, and to give a result which shows up surprisingly often in (non-topology) mathematical practice, we have the following fact:

Lemma 55 (The tube lemma)**.** *Suppose X is a compact space and Y is any topological space.*

For any open set $U \subset X \times Y$ containing the subset $X \times \{y\}$ ('a slice'), there is an open set $W \subset Y$ containing y so that $X \times W \subset U$ ('a tube').

It follows that the projection $p_2 : X \times Y \to Y$, given by $p_2(x, y) = y$, is a closed map.



Proof. For each point $x \in X$, we have $(x, y) \in U$. By definition of the product topology, there is an open set $V_x \subset X$ and $W_x \subset Y$ with $(x, y) \in V_x \times W_x \subset U$ and in particular $x \in V_x$ and $y \in W_x$.

Write $\mathcal{V} = \{V_x\}_{x \in X}$, chosen as above. This is an open cover of X. Because X is compact, there's a finite subcover of V; let's say it's V_{x_i}, where $x_1, \cdots, x_n$ is a finite collection of points in X.

But now we look at the second factor; each of these V_{x_i}'s had a corresponding W_{x_i} so that $V_{x_i} \times W_{x_i} \subset U$. Write $W = W_{x_1} \cap \cdots \cap W_{x_n}$ — this is again an open subset of Y (it's a finite intersection of open sets), and this intersection is nonempty because $y \in W_{x_i}$ for all i (so $y \in W$ as well).

By definition, we have that

$$V_{x_i} \times W \subset V_{x_i} \times W_{x_i} \subset U;$$

taking a union, we find that

$$X \times W = \bigcup_{i=1}^{n} V_{x_i} \times W \subset U.$$

Since $y \in W$, we've constructed the claimed set W.

Now let's show that p_2 is closed. We start with a closed set $C \subset X \times Y$; we'll show that $p_2(C)^c$ is open, and we'll use one of our common strategies (locally open sets are open).

Pick $y \notin p_2(C)$ — so $(x, y) \notin C$, for all $x \in X$. Thus $X \times \{y\} \subset C^c$.

By the tube lemma, there is an open $y \in W \subset Y$ so that $X \times W \subset C^c$. It follows that $W \subset p_2(C)^c$.

Since we've shown that every $y \in p_2(C)^c$ has an open set $y \in W$ around it also contained in $p_2(C)^c$, it follows that $p_2(C)^c$ is a union of open sets, hence open; so $p_2(C)$ is closed, as desired. □

What's going on here is the general trick that compact spaces are good for. You have some condition, which you've encoded in a family of open sets. You know what to do when there's only finitely many of them (above, we knew that the intersections of the W_{x_i} were open — when we were only intersecting finitely many); so you use the compactness condition to reduce your family to a finite one.

As a simple topological application of the tube lemma, we have the following.

This argument is somewhat long and complicated, but can be made very visual. I recommend you attempt to 'draw' this argument as you go along to improve and check your understanding, and talk to me about your drawing so we can see if we agree about the picture.

Proposition 56. *If X and Y are compact spaces, the product $X \times Y$ is also a compact space.*

Proof. We start with an open cover $\mathcal{U}$ of $X \times Y$. We'll use compactness of X to cook up — for each $y \in Y$ — a *finite* open cover $\mathcal{U}_y$ of X. Then we'll use compactness of Y to reduce this to a finite collection of finite open covers — giving us a finite subcover of $X \times Y$.

To make that first step more precise: for each $y \in Y$, we look at the intersection $\mathcal{U}_y = \{U \cap (X \times \{y\}) \mid U \in \mathcal{U}\}$. This gives an open cover of $X = X \times \{y\}$. (I am sweeping under the rug here an argument that X is homeomorphic to $X \times \{y\}$ with the subspace topology, but you can check this if you want.)

We then apply compactness of X to get a finite sequence $U_{y,1}, \cdots, U_{y,n(y)} \in \mathcal{U}$ so that $X \times \{y\} \subset U_{y,1} \cup \cdots \cup U_{y,n(y)}$. (The number of these open sets depends on our choice of y, so I write the finite number of them — $n(y)$ — as a function of y.)

Now, applying the tube lemma, pick an open $y \in W_y \subset Y$ such that $X \times W_y \subset U_{y,1} \cup \cdots \cup U_{y,n(y)}$.

Now *this* gives us an open cover $\mathcal{W} = \{W_y\}_{y \in Y}$ — and because Y is compact, there's a finite collection $W_{y_1}, \cdots, W_{y_k}$ which covers Y — so that $Y = W_{y_1} \cup \cdots \cup W_{y_k}$.

Taking this a step back... this suggests we look at the finite collection of open sets

$$\mathcal{U}' = \{U_{y_i,j}\}_{\substack{1 \leqslant i \leqslant k \\ 1 \leqslant j \leqslant n(y_k)}}.$$

We already know that

$$X \times W_{y_i} \subset U_{y_i,1} \cup \cdots \cup U_{y_i,n(y_i)} \subset \bigcup_{U \in \mathcal{U}'} U;$$

taking the union, we see that

$$X \times Y = X \times \left(\bigcup_{i=1}^{k} W_{y_i} \right) \subset \bigcup_{U \in \mathcal{U}'} U.$$

So $\mathcal{U}'$ is actually an open cover of $X \times Y$. We've shown that an arbitrary open cover has a finite subcover — so $X \times Y$ is compact. □

It follows inductively that any finite product $X_1 \times \cdots \times X_n$ of compact spaces is compact.

It also happens to be true that *arbitrary* products of compact spaces are compact. This argument either requires (1) that we develop the theory of convergence of nets (which is too much set-theoretic work for my taste) or (2) that we write an extended argument whose ideas aren't used elsewhere.

If you want to read a proof, it's called Tychonoff's theorem; one gets, for instance, that $[0,1]^{[0,1]}$ is compact.

10/5: Separation axioms and Hausdorff spaces

There are a sequence of increasingly more demanding conditions that one can demand of a topological space. They are called *separation axioms*, and are about how effectively you can differentiate points by open sets around them.

Despite how many of them there are, very few of these appear in the actual study of a mathematician in the 21st century. I'll list the first few here for culture's sake, but we'll care mostly about a specific one.

Definition 30. *A topological space X is said to be T_0 if, for every two distinct points $x, y \in X$, there is an open set containing one of these points but not the other.*

Equivalently, the map $\in: X \to \mathcal{P}(\mathcal{T})$, sending a point $x \in X$ to the collection of all open sets containing x, is injective.

The idea is that in a space that is *not* T_0, you can't even tell every pair of points apart; some of them are just indistinguishable from the POV of the open sets. This is usuaally pretty undesirable; if you can't tell two points apart, why are we considering them to be different?

The indiscrete space is an example of a space which is *not* T_0; all those points are more or less completely indistinguishable. In fact, if you want, you can prove that a space X is *not* T_0 if and only if X has an indiscrete subspace: a subset which, when equipped with the subspace topology, is indiscrete.

Every space you actually see in real life is T_0, and usually much more. Indiscrete spaces are a bit silly and do not really show up except when you're trying to be a contrarian.

Next is a property we've already seen mentioned before:

Definition 31. *A topological space X is said to be T_1 if every singleton set $\{x\}$ is closed; that is, $X \backslash \{x\}$ is open.*

A T_1 space is T_0: if x, y are distinct points, then $X \backslash \{x\}$ is ann open set containing y but not x.

Not every space in mathematical practice is T_1, but for a topologist most are. The most famous space which is not T_1 is the *spectrum of a ring* equipped with the *Zariski topology*. Talking about this would be too much of a digression for us, but you'll learn about it in either a commutative algebra or algebraic geometry class, where it serves a very important book-keeping role.

Some mathematicians really like to think about finite topological spaces; they can be used to model interesting combinatorial questions about even normal-looking spaces, like a tetrahedron. This is another place where T_1 is too restrictive: every finite T_1 space is discrete. (Do you see why?)

However, as we've seen, every metric space is T_1.

The following condition is perhaps one of the single most important assumptions in topology; you'll see this assumption on most statements. Oftentimes topologists are so used to this assumption it's not even mentioned when speaking informally; the original definition of a topological space (due to Felix Hausdorff) included this as an axiom.

Definition 32. *A topological space X is said to be T_2 (or more commonly,* Hausdorff*) if, for every distinct points $x, y \in X$, there are open sets $U, V \subset X$ with $x \in U, y \in V$, and $U \cap V = \varnothing$.*

This is perhaps where it becomes clear why this sequence of axioms ought to be called separation axioms: the Hausdorff axiom quite literally says we can separate points by open sets.

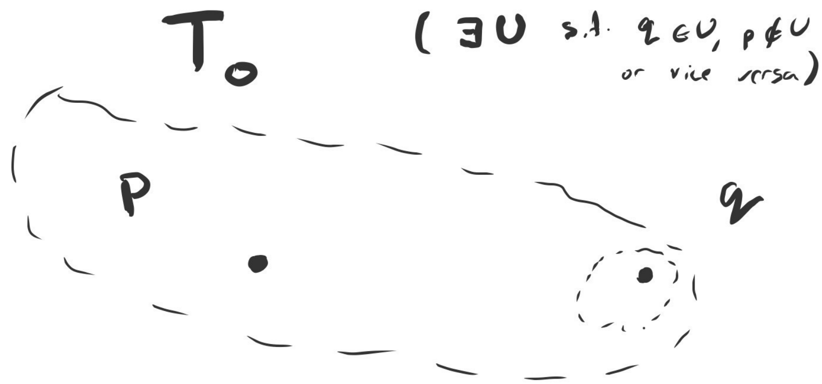
Being able to separate points by open sets is more than enough to show that points are closed sets:

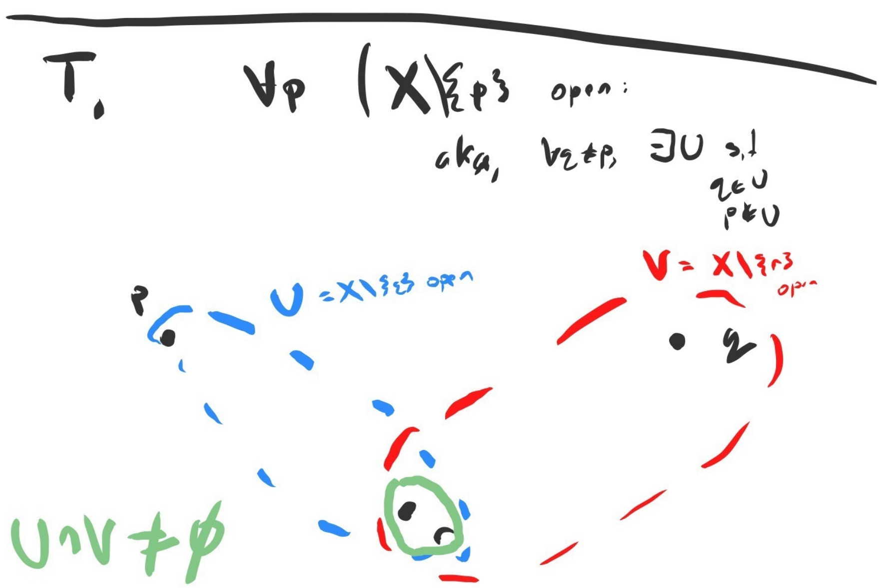
Proposition 57. *A Hausdorff space is T_1.*

Proof. Pick $x \in X$; we want to show that $X \backslash \{x\}$ is open. For any $y \in X \backslash \{x\}$, by assumption, we can find open sets $U, V_y \subset X$ with $x \in U$ and $y \in V_y$, with $U \cap V_y = \varnothing$. In particular, $x \notin V_y$, so that $V_y \subset X \backslash \{x\}$. We have shown that $X \backslash \{x\}$ is locally open, hence open. □

The spaces we usually have in our back pocket are always Hausdorff, too. Try drawing a picture and then coming up with the following proof on your own (if you have trouble, try reading the proof then coming up with the right picture).

Proposition 58. *Metric spaces are Hausdorff spaces.*





Proof. Pick $x, y \in X$, a metric space with distance function d. If x and y are distinct, we have $d(x, y) > 0$. The idea is to take as our open sets small balls around x and y, so small that they can't possibly overlap; we use the fact that the distance is positive to choose a sufficiently small radius. In fact, take as our open sets $U = B_r(x)$ and $V = B_r(y)$, where $0 < r$ and $r \leqslant d(x, y)/2$, then these balls are not large enough to intersect: if $z \in B_r(x)$, then

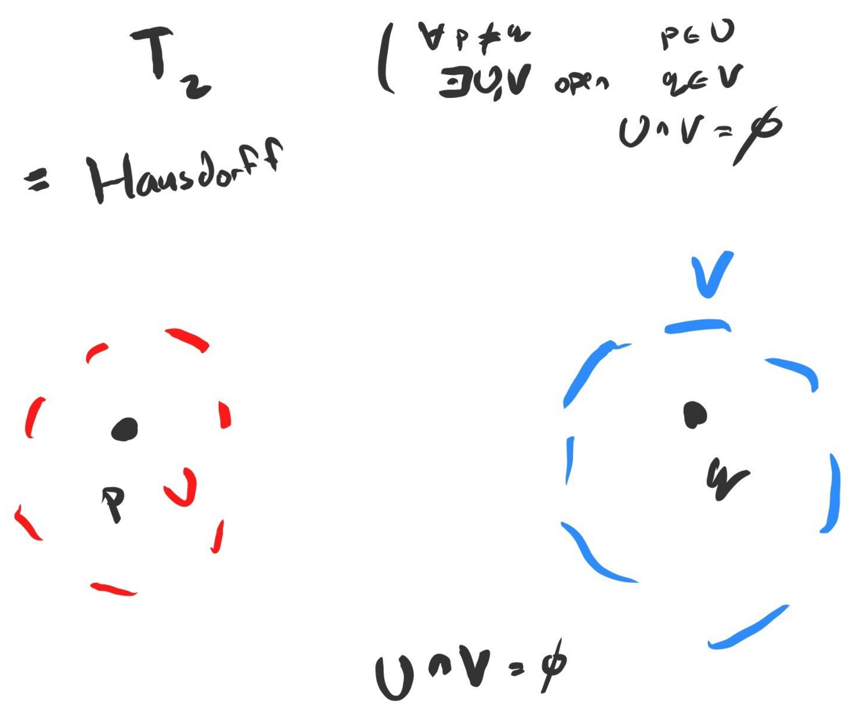
$$d(x, y) \leqslant d(x, z) + d(z, y) < r + d(z, y),$$

so that $d(z, y) > d(x, y) - r$. Since $d(x, y) \geqslant 2r$, we find that $d(z, y) > d(x, y) - r \geqslant r$, so that $d(z, y) \geqslant r$ — and in particular, $z \notin B_r(y)$. So $B_r(x) \cap B_r(y) = \varnothing$; these two sets do not intersect. We checked at the very beginning of the semester that open balls are indeed open sets, so we have proved that the Hausdorff condition holds. □

Hausdorff spaces are ubiquitous (though the spaces we mentioned earlier which are not T_1 are also necessarily not Hausdorff.) It is fortunate, then, that the main operations we have to produce new spaces also produce Hausdorff spaces.

Proposition 59 (Subspaces of Hausdorff spaces are Hausdorff)**.** *Let X be a Hausdorff space, and S a subspace. Then S is also Hausdorff.*

Proof. Pick $x, y \in S$ distinct points. By the assumption that X is Hausdorff, there are open sets $U, V \subset X$ with $x \in U, y \in V$, and $U \cap V = \varnothing$. by the definition of the subspace topology, the sets $U \cap S$ and $V \cap S$ are



open in S; we have $x \in U \cap S, y \in V \cap S$; and lastly

$$(U \cap S) \cap (V \cap S) = (U \cap V) \cap S = \varnothing.$$

So we've proved the Hausdorff property. □

Proposition 60 (Products and disjoint unions of Hausdorff spaces are Hausdorff)**.** *Let X_i be a family of Hausdorff spaces, indexed by $i \in I$. Then both the product*

$$\bigtimes_{i \in I} X_i$$

and the disjoint union

$$\bigsqcup_{i \in I} X_i$$

are Hausdorff spaces.

Proof. Pick $x, y \in \bigtimes X_i$ distinct points; that means that $x = (x_i)_{i \in I}$, an I-tuple of elements, one for each X_i. Similarly, $y = (y_i)_{i \in I}$.

That these points are distinct means that for *some* component of the I-tuple, they disagree; so $x_i \neq y_i$ for *some* $i \in I$. Now we use that X_i is Hausdorff: there are open sets $U_i, V_i \subset X_i$ with $x_i \in U_i, y_i \in V_i$, with $U_i \cap V_i = \varnothing$.

We set

$$U = \bigtimes_{j \in I} U_j,$$

where U_i is the set we constructed earlier but for $j \neq i$ we have $U_j = X_j$, and similarly with V — the only factor that's actually a proper subset is the i'th factor. Said another way, $U = \{x \in X \mid x_i \in U_i\}$. This is a basic open set inn the definition of the product topology, as is V. But

$$U \cap V = \{x \in X \mid x_i \in U_i \cap V_i\} = \varnothing,$$

because $U_i \cap V_i = \varnothing$. Thus we have found disjoint open sets containing x and y, respectively.

The argument is a little bit easier for the disjoint union, where the idea to begin with is that we just have all of the X_i's totally separate from each other, doing their own thing. So if we want to separate points, they're already separate if they're in different factors X_i; otherwise, we separate them inside of X_i.

More formally: if $x, y \in \bigsqcup_{i \in I} X_i$, then $x \in X_i$ for a unique i, and $y \in X_j$ for a unique j. If $i \neq j$, then $x \in X_i, y \in X_j$, and $X_i \cap X_j = \varnothing$, while both of these are (by definition) open sets in the disjoint-union topology.

If $i = j$, then $x, y \in X_i$ are distinct points; pick open sets $U, V \subset X_i$ with $U \cap V = \varnothing$ and $x \in U, y \in V$. Then

$$U, V \subset X_i \subset \sqcup_{i \in I} X_i$$

are disjoint open subsets of the disjoint union which contain x and y respectively. We've checked the Hausdorff condition in both cases, and thus the disjoint union of Hausdorff spaces is Hausdorff. □

It follows that most spaces we know are Hausdorff. There are some silly examples that aren't:

Example 36. Let X be a set with $|X| > 1$. Equipping it with the indiscrete topology, X_{indisc} is not Hausdorff (it's not even T_0).

If X is uncountable and X_{cc} equips it with the countable-complement topology, then X_{cc} is T_1 (points are singletons, hence countable, hence closed), but not Hausdorff: if $U, V \subset X$ are open, then their complements U^c and V^c are countable; so $(U \cap V)^c = U^c \cup V^c$ is countable, and in particular, not all of X. That is, any two nonempty open subsets of X_{cc} intersect non-trivially. So we certainly cannot separate two points by disjoint open sets: there aren't any two nonempty disjoint open sets!

I don't know how often you're expecting to see $\mathbb{R}_{cc}$ on the street, but: don't. It's a great example for understanding the complexity of topological spaces, but it doesn't appear in reality.

The following properties characterize Hausdorff spaces in terms of graphs being closed — a property you might be familiar with if you've studied continuous maps between metric spaces.

Theorem 61 (Graph characterization of Hausdorffness)**.** *Let X be a topological space. Then the following are equivalent.*

(a) X is Hausdorff,

(b) For any topological space Y and any continuous map $f : Y \to X$, the graph

$$\Delta_f = \{(y, f(y)) \mid y \in Y\} \subset Y \times X$$

is a closed subset in the product topology,

(c) The subset $\Delta_X = \{(x, x) \mid x \in X\} \subset X \times X$ is closed in the product topology.

Proof.

$$(a) \implies (b)$$

Suppose X is Hausdorff, and $f : Y \to X$ is continuous. Let's prove that the complement $\Delta_f^c = (Y \times X) \backslash \Delta_f$ is open.

Pick a point $(y,x) \in \Delta_f^c$; it suffices to show that there is an open set

$$(y,x) \in U_{y,x} \subset \Delta_f^c,$$

because locally open sets are open.

By definition of Δ_f^c, we have that $x \neq f(y)$. Because X is Hausdorff, we may choose open sets V_x and $V_{f(y)}$ with $x \in V_x$ and $f(y) \in V_{f(y)}$, but $V_x \cap V_{f(y)} = \varnothing$. In particular, $f(y) \notin V_x$, and so $y \notin f^{-1}(V_x)$.

Then take $U_{y,x} = f^{-1}(V_{f(y)}) \times V_x$. This is a basic open set for the product topology, hence open. We have immediately that $(y,x) \in U_{y,x}$; what we need to show is that $U_{y,x} \subset \Delta_f^c$. So pick $(z,w) \in U_{y,x}$ — this means that $f(z) \in V_{f(y)}$ and $w \in V_x$. Because $V_{f(y)} \cap V_x = \varnothing$, we certainly have that $f(z) \neq w$ — so $(z,w) \notin \Delta_f$. Since we've shown this for arbitrary $(z,w) \in U_{y,x}$, it follows that

$$U_{y,x} \subset \Delta_f^c$$

as desired.

$$(b) \implies (c)$$

The identity map $1_X : X \to X$ is continuous, and Δ_X is its graph.

$$(c) \implies (a)$$

Suppose Δ_X is closed in the product topology on $X \times X$; so its complement Δ_X^c is open. For distinct points $x, y \in X$, we have that $(x,y) \in \Delta_X^c$.

Because Δ_X^c is open in the product topology, there is a basic open set $U \times V$ containing (x,y) with $U \times V \subset \Delta_X^c$. To say that $(x,y) \in U \times V$ means that $x \in U$ and $y \in V$. To say that $U \times V \subset \Delta_X^c$ means that $U \times V \cap \Delta_X = \varnothing$ — that is to say,

$$\{(u,v) \mid u \in U, v \in V, u = v\} = \varnothing.$$

This is just a very fancy way to say that $U \cap V = \varnothing$.

So for distinct points $x, y \in X$, we have produced open sets $x \in U, y \in V$ which do not intersect: $U \cap V = \varnothing$. Therefore, X is a Hausdorff space. □

Our characterization of Hausdorff spaces buys us the following fact — which I would intuitively have said is true for *any* space, but really does require the Hausdorff condition.

Corollary 62 (Functions to a Hausdorff space agree on a closed set)**.** *If Y is a Hausdorff space, and $f, g : X \to Y$ is a pair of continuous maps, the set*

$$S = \{x \in X \mid f(x) = g(x)\},$$

consisting of points where f and g coincide, is a closed set.

Proof. If f and g are continuous, then $(f,g) : X \to Y \times Y$ is continuous with respect to the product topology on the codomain. The set S is equal to $(f,g)^{-1}\Delta_Y$. Because Y is Hausdorff, this is the inverse image of a closed set by a continuous map, whence S is closed. □

Convergent sequences

You may remember the following fact from a course in analysis.

Proposition 63 (Continuity is sequential continuity)**.** *Let $f : X \to Y$ be a function between metric spaces (with metrics d, d' respectively). Then f is continuous iff, for every convergent sequence $x_n \to x$, the sequence $f(x_n)$ is also a convergent sequence, with* $\lim f(x_n) = f(x)$.

We do not currently have the technology to study how this works in topological spaces, but we can at least set up the right notions.

Definition 33. *The* convergent sequence space *is the set* $\mathbb{N} \cup \{\infty\} = \{1, 2, \cdots, \infty\}$*, equipped with the topology*

$$\mathcal{T} = \{U \mid \textit{if } \infty \in U, \textit{ then for some } n, \ (n, \infty] \subset U\}.$$

That is, each $\{n\}$ *is open, for finite* n*; but any open set containing* ∞ *also contains every sufficiently large positive integer.*

If you like, the map $n \mapsto 1/n$ gives a homeomorphism from the convergent sequence space to $\{1, 1/2, 1/3, \cdots, 0\}$, which you can think of as being the canonical convergent sequence.

Definition 34. *A sequence* x_n *in a topological space (*$n \in \mathbb{N}$*) is* convergent *if, considered as a map* $x : \mathbb{N} \to X$ *given by* $n \mapsto x_n$*, the sequence extends to a continuous map* $x : \mathbb{N} \cup \{\infty\} \to X$ *from the convergent sequence space to* X*. In this case, we say the sequence* x_n *converges to the point* x_∞*.*

This fancy definition can be stated more explicitly as follows; the definition is really just an encoding of this way of thinking of convergent sequences.

Proposition 64 (Open set characterization of convergence)**.** *A sequence* x_n *converges to* $x \in X$ *if, for every open set* U *containing* x*, there is some large* $N \in \mathbb{N}$ *so that, for all* $n > N$*, we have* $x_n \in U$*. That is: every term in the sequence sufficiently far in is contained in* U*.*

Proof. Write $f : \mathbb{N} \cup \{\infty\} \to X$ for the function with $f(n) = x_n$ and $f(\infty) = x$. First, note that *any* subset of $\mathbb{N} \cup \{\infty\}$ *not* containing ∞ is open, by definition of the convergent sequence topology. So if $x \notin U$, then $f^{-1}(U) \subset \mathbb{N}$, and hence $f^{-1}(U)$ is open.

Now if $x \in U$, then $f^{-1}(U)$ contains ∞. Then $f^{-1}(U)$ being open is *equivalent* to $x_n \in f^{-1}(U)$ for all sufficiently large n (say, $n > N$); and this is the same as saying that $f(n) = x_n \in U$ for all sufficiently large n (say, $n > N$). Thus continuity of f is equivalent to the stated condition. □

This definition is kind of unsatisfactory in general. One major failure is that *sequences do not need to converge to a unique limit!* For instance, *any sequence in an indiscrete space converges to any point of that space.* The reason I mention this now is that Hausdorffness fixes this problem.

Proposition 65 (Sequences have at most one limit in a Hausdorff space)**.** *Let* X *be a Hausdorff space. Then a convergent sequence converges to exactly one point.*

Proof. We will prove the contrapositive: if $x_n \to x$ and $x_n \to y$, where $x \neq y$, then X is not a Hausdorff space.

Pick open sets U, V, with $x \in U$ and $y \in V$. I claim that $U \cap V$ is always nonempty, and thus X is not Hausdorff. To see this, note that by assumption of convergence, $x_n \in U$ for all sufficiently large n (say, $n > N_1$); and also $x_n \in V$ for all sufficiently large n (say, $n > N_2$). Then taking $n > \max(N_1, N_2)$, we find that $x_n \in U \cap V$, so that $U \cap V$ is nonempty. □

This proposition might give a sense of why Hausdorff spaces are so important and ubiquitous: sequential thinking only really works in Hausdorff spaces.

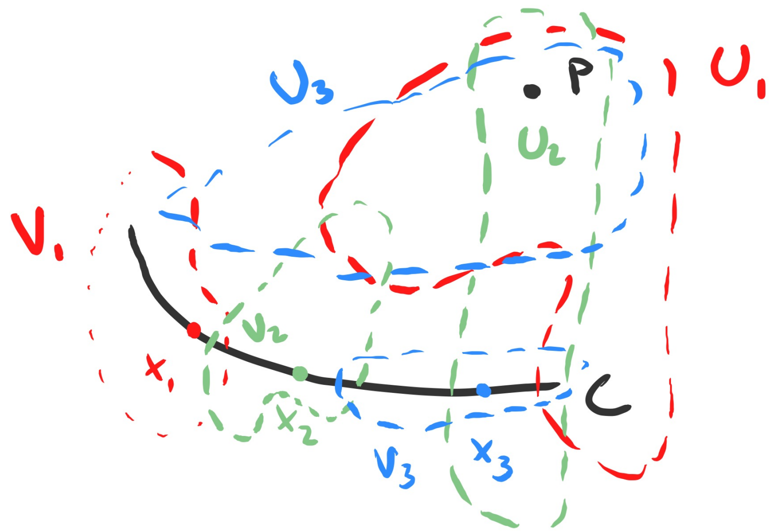
If you really, really like thinking of continuous functions in terms of convergent sequences, there is a replacement for topological spaces in general called *nets.* We won't talk about them because the set-theoretic setup is kind of confusing (at least to me), and they take us kind of far afield from the rest of what we want to do.

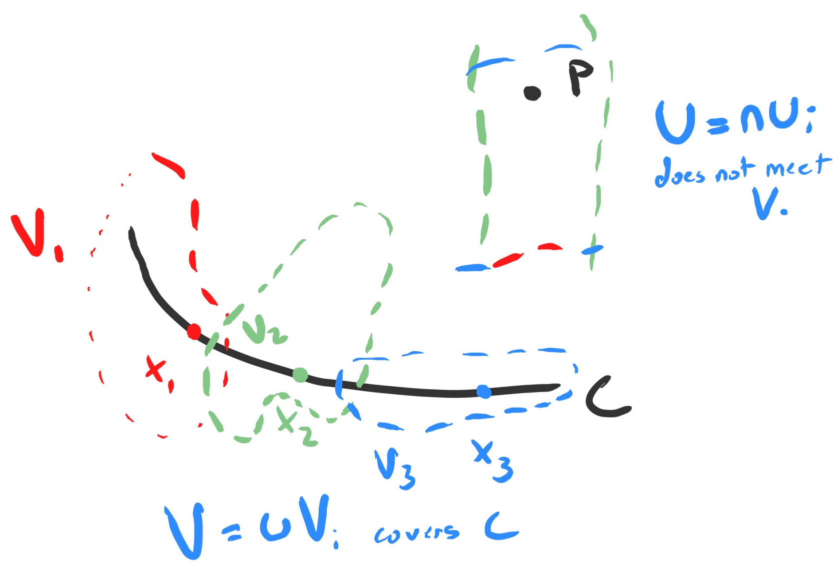
The point is, somehow, that topological spaces can be too "big", and sequences indexed by $\mathbb{N}$ don't cut it — so one needs to index by larger sets, and then understanding what that means (and what convergence should mean) gets confusing.

Compactness and Hausdorffness

The Hausdorffness condition can be used to *produce* open sets, and in particular, open covers; the compactness condition can be used to reduce these to finite covers. Put together, this is a formidable force. For instance, we have the following:

Proposition 66. *A compact subspace of a Hausdorff space is closed.*





Proof. The pictures above illustrate the proof below (in the special case that it only takes three of the V_x's to cover all of C).

Let X be Hausdorff, and $C \subset X$ be a compact space. Our goal is to show that C^c is open; let's aim to show that it's locally open.

To that end, pick a point $p \notin C$. For each $x \in C$, we know by the Hausdorff property that we can find $x \in V_x, p \in U_x$ so that $V_x \cap U_x = \varnothing$, and both sets are open. The collection $\{V_x \cap C\}_{x \in C}$ gives an open cover of C; appealing to compactness, there's a finite subcover $V_{x_1}, \cdots V_{x_n}$.

Write $U = U_{x_1} \cap \cdots \cap U_{x_n}$. Because $p \in U$, and we intersect only finitely many open sets, this is a nonempty open set. Further, I claim $U \cap C = \varnothing$ — for $U \cap V_{x_i} = \varnothing$, and so

$$U \cap C \subset U \cap \left(\bigcup_{1 \leqslant i \leqslant n} V_{x_i} \right) = \bigcup_{1 \leqslant i \leqslant n} (U \cap V_{x_i}) = \varnothing.$$

We have shown that C^c is locally open — so it is a union of open sets (these U we constructed for each p), whence open. □

This leads to an incredibly efficient tool for checking that certain maps are homeomorphisms: if you can verify some simple properties on the domain and codomain, continuous bijections are homeomorphisms for free.

Corollary 67. *Let X be a compact space and Y a Hausdorff space. A continuous bijection $f : X \to Y$ is necessarily a homeomorphism.*

Proof. I claim that f is a closed map, whence f^{-1} is closed-set-continuous, so continuous.

Let $C \subset X$ be closed. We have already seen that closed subspaces of X are compact; so C is compact. We have also already seen that images of compact sets are compact, so $f(C)$ is compact. And we just saw that *in a Hausdorff space,* compact subspaces are closed. So $f(C)$ is closed, as desired. ☐

I cannot stress how **wildly useful** this result is. It saves you half of the work! And sometimes more than that, when f^{-1} is harder to show is continuous than f is.

Example 37 (Non-examples). In the case of the map $\exp : [0, 2\pi) \to S^1$, which is a continuous bijection but not a homeomorphism, the domain is not compact.

Let X be a finite set with more than one element, and let $\mathcal{T}$ be a topology in which not every set is open. Consider the identity map $1_X : X_{\text{disc}} \to (X, \mathcal{T})$. This is continuous, because every map from a discrete space is continuous; it's a bijection because it's the identity map. It is not a homeomorphism: pick any set $S \subset X$ which is not open in $\mathcal{T}$; then $1_X(S) = S$ is not open, so 1_X is not an open map.

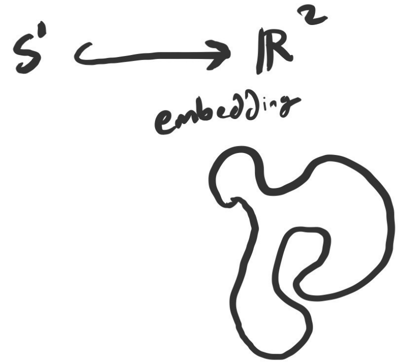
Note that $\mathcal{T}$ cannot be Hausdorff. If it was, points would be closed, which we saw earlier would force $\mathcal{T}$ to be the discrete topology; but we've assumed that there is at least one set which is not open in $\mathcal{T}$.

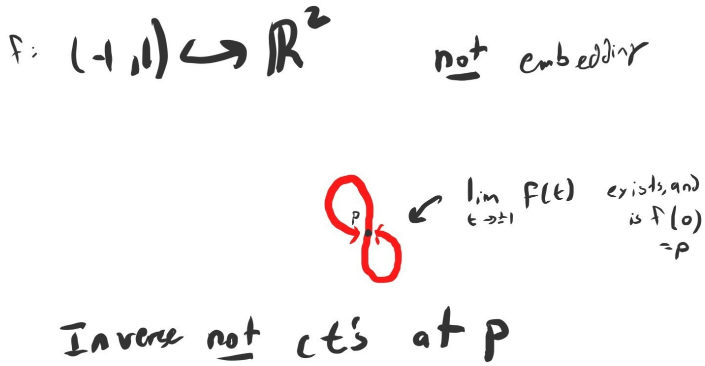
Example 38. This property (that continuous bijections are homeomorphisms) sometimes holds for non-compact spaces, but you shouldn't expect it to most of the time. It is in fact true that every continuous bijection $f : \mathbb{R} \to \mathbb{R}$ is a homeomorphism (prove it using the intermediate value theorem, if you like).

Even more generally, it is true that a continuous bijection $f : \mathbb{R}^n \to \mathbb{R}^m$ for $n \neq m$ is in fact a homeomorphism. But this is quite hard and requires tools we won't develop in this class.

Definition 35. *A continuous map $f : X \to Y$ is a* topological embedding *(or just an embedding) if the map $f : X \to f(X)$, given by restricting codomain to $f(X)$ with the subspace topology, is a homeomorphism.*

One says 'f is a homeomorphism onto its image'.



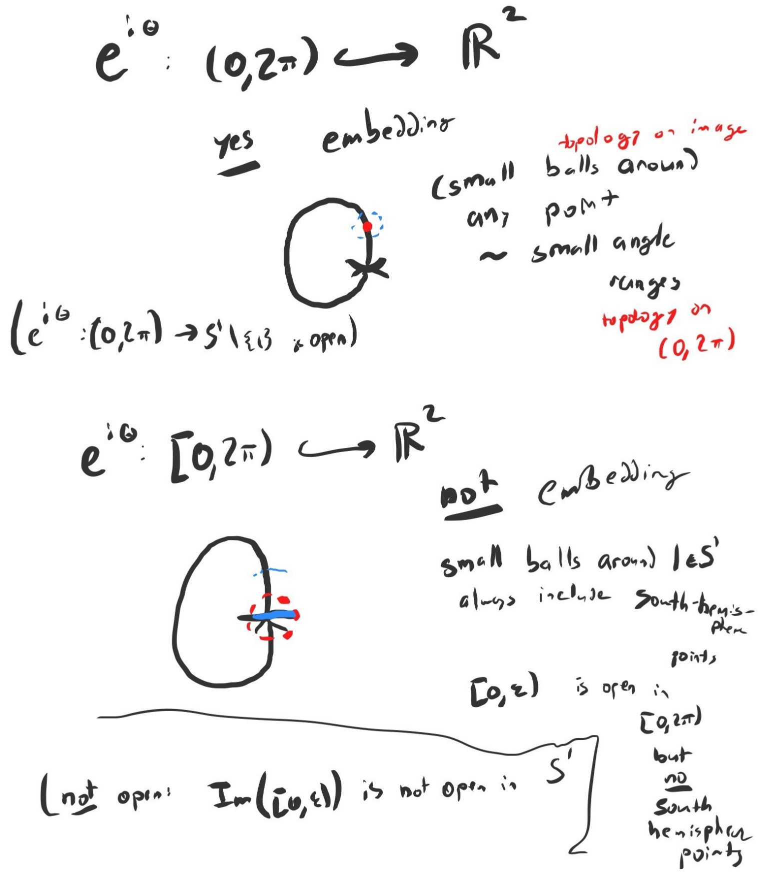


Corollary 68. *Let X be compact and Y be Hausdorff. A continuous injection $f : X \to Y$ is a topological embedding.*

Proof. Subspaces of Hausdorff spaces are Hausdorff, so the map $f' : X \to f(X)$ given by restricting the codomain is a continuous map from a compact space to a Hausdorff space — and because f surjects onto its image, but was assumed injective, the map f' is a bijection.

Because X is compact and $f(X)$ is Hausdorff, f' is a homeomorphism. □

Referring to the pictures above, the idea to me is that if X is compact, f cannot 'accumulate onto itself'.



10/7: One-point compactifications and proper maps

We'll spend today covering one more topic that shows up occasionally in topology: the one-point compactification and proper maps. A discussion of proper maps could be done in more generality than our approach permits, but we choose the current approach to connect these two topics together and because many spaces you encounter in modern topology will be *locally compact*, a term you will see below.

The one-point compactification

We've learned that compact spaces are great. In particular, we just learned one of my favorite and most useful theorems: a continuous bijection from a compact space to a Hausdorff space is a homeomorphism.

We would really love it if we could take a space X, turn it into a compact space X', and then apply nice theorems like the one above to get results like that more generally. This leads us to the notion of *compactification*, and then to a particularly comprehensible compactification: Alexandrov's *one-point compactification.*

Definition 36. *Let X be a topological space. A* compactification of X *is a pair (X', j) of a* ***compact*** *topological space X' and an embedding $j : X \to X'$ so that $j(X)$ is a dense subset of X'.*

The idea is that we've taken a space and 'made it compact' by adding some extra points 'at infinity'. Whatever it means to add points at infinity,

Definition 37. *Let X be a topological space. The* Alexandrov one-point compactification*, written X_∞, is the topological space whose underlying set is $X \sqcup \{\infty\}$. A set $U \subset X_\infty$ is open if either $\infty \notin U$ and $U \subset X$ is an open subset of X, or if $\infty \in U$ and $U^c \subset X$ is both closed and compact.*

Lemma 69. *The one-point compactification topology described above is, in fact, a topology. The natural inclusion $i : X \to X_\infty$ is a compactification if and only if X is noncompact.*

Proof. You actually gave much of this argument already on a previous homework!

First, observe that $\varnothing$ is open in X_∞ (it is an open subset of X), and X_∞ is open in X_∞ (it contains ∞ and its complement $\varnothing$ is a compact, closed subset of X).

Next, suppose U_i is an arbitrary collection of open sets in X_∞; we want to show their union is open. If $\infty \notin \bigcup_i U_i$ then it is not in any of the U_i's, and (by definition of the one-point compactification topology) each U_i is an open subset of X. Because unions of open sets on X are open, it follows that $\bigcup_i U_i$ is an open subset of X — and hence an open subset of X_∞, as desired.

If, on the other hand, $\infty \in \bigcup_i U_i$, then we need to check that $(\bigcup_i U_i)^c = \bigcap_i U_i^c$ is a compact and closed subset of X. Note first that for all i, $U_i^c \cap X$ is a closed subset of X (either it is the complement in X of an open set, or it is compact and closed by definition).

Therefore $\bigcap_i U_i^c = \bigcap_i (U_i^c \cap X)$ is a closed subset of X, and we are taking the intersection of closed sets, *at least one of which is compact.* A previous homework exercises shows the result is compact, as desired.

Lastly, suppose $U_1, \cdots, U_n$ is a finite collection of open sets in X_∞; we want to show the intersection is open. Again, we split into cases. Whether or not $\infty \in U_i$, we still have that $U_i \cap X$ is open in X, so $\bigcap_i U_i \cap X$ is open in X. If $\infty \notin U_i$ for some i, then that's all we needed to show: $\bigcap_i U_i$ is open.

If $\infty \in U_i$ for all i, then $\infty \in \bigcap_i U_i$.

To see the claim about X_∞ being a compactification of X, observe that $i : X \to X_\infty$ is always continuous, and $\overline{i(X)} = X_\infty$ if and only if $i(X)$ is not closed. Now $i(X)$ is closed if and only if $i(X)^c = \{\infty\}$ is an open set; and $\{\infty\}$ is an open set if and only if its complement — $i(X)$ — is compact. Therefore

$$\overline{i(X)} = X_\infty \iff X \text{ is noncompact},$$

as desired. □

Example 39. I previously said the 'convergent sequence space' is the set $\{0, 1, 2, \cdots, \infty\}$, where a set is open if either $\infty \notin U$ or $[n, \infty] \subset U$ for some sufficiently large n. This is precisely the one-point compactification $\mathbb{N}_\infty$ of the discrete space $\mathbb{N}$. (Note that a compact subset of $\mathbb{N}$ is finite.)

The following proposition will be helpful to produce more examples without doing any extra work.

Proposition 70. *Let X be a compact Hausdorff space. If $p \in X$ is any point, then $(X\backslash\{p\})_\infty \cong X$.*

It follows that if $X\backslash\{p\} \cong Y$, then $X \cong Y_\infty$.

Proof. Define the map $f : (X\backslash\{p\})_\infty \to X$ by the formula

$$f(x) = \begin{cases} x & x \neq \infty \\ p & x = \infty \end{cases}.$$

This is clearly a bijection. To see that it is continuous, pick $U \subset X$ open. If $p \notin U$ then $f^{-1}(U) = U$ which is open in X by hypothesis, hence open in X_∞. If $p \notin U$ then $f^{-1}(U) = (U\backslash\{p\}) \sqcup \{\infty\}$. This is open in the one-point-compactification if and only if

$$(X\backslash\{p\})\backslash(U\backslash\{p\}) = U^c$$

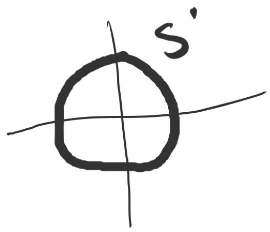
is both compact and a closed subspace of $X\backslash\{p\}$. But $U^c \subset X$ is closed in X because $U \subset X$ was open by hypothesis; because U^c is a closed subset of a compact space it is compact. Lastly, $U^c \cap (X\backslash\{p\}) = U^c$ is closed by definition of the subspace topology. Thus $f^{-1}(U)$ is open and we have shown that f is continuous.

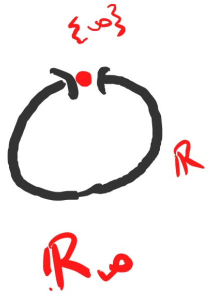
Now f is clearly a bijection; the domain is compact (the one-point compactification always is) and the codomain Hausdorff by hypothesis, so as a result this map is a homeomorphism.

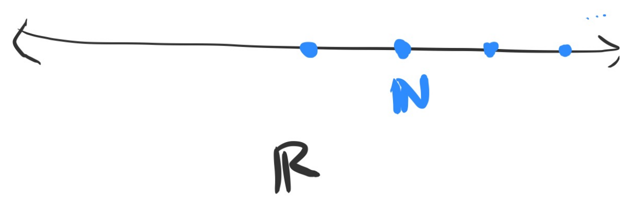
For the claim after that, prove that if $f : A \to B$ is a homeomorphism, it induces a homeomorphism on their one-point compactifications. □

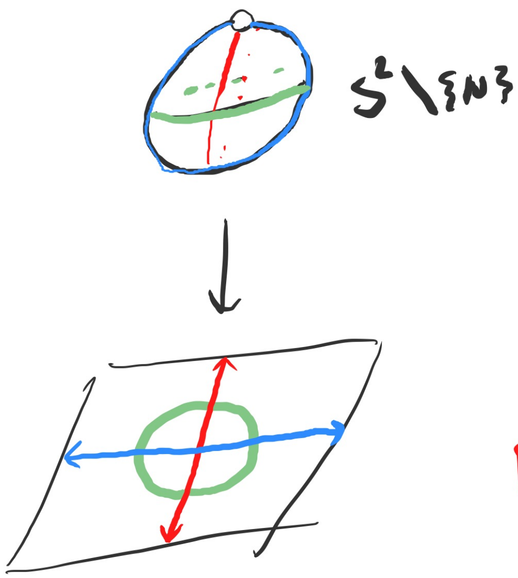
Example 40. We've previously shown that $S^n\backslash\{N\}$ — the n-dimensional sphere inside of $\mathbb{R}^{n+1}$, with a point deleted — is homeomorphic to $\mathbb{R}^n$. As an immediate corollary, we see that $(\mathbb{R}^n)_\infty \cong S^n$ — the one-point compactification of Euclidean space is a Euclidean sphere of the same dimension. For instance, $\mathbb{R}_\infty \cong S^1$.

Here I will draw some pictures of the cases $n = 1, 2, 3$ to hopefully help give some intuition.



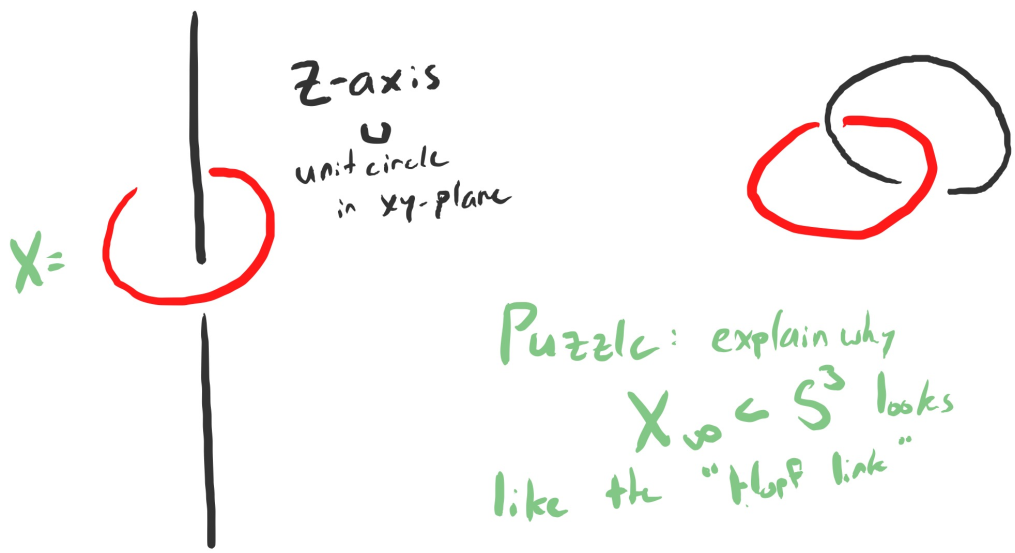






One-pt compactification of lower half-plane is "front hemisphere"

Puzzle: what does $\mathbb{N}_\infty^2 \subset S^2$ look like?



In order to get at our technical tool from yesterday — continuous bijections from a compact space to a Hausdorff space are homeomorphisms — we'll want some analysis of how this construction behaves with the property of being Hausdorff. We say that a space *has compact neighborhoods*[9] if for each $x \in X$, there is an open set $x \in U$ so that $\overline{U}$ is compact.

Most spaces you think of intuitively have compact neighborhoods: $\mathbb{R}^n$ has compact neighborhoods; any open subspace has compact neighborhoods; any closed subspace of $\mathbb{R}^n$ has compact neighborhoods. (Do you see why all of these are true?) However, $\mathbb{Q}$ does not have compact neighborhoods.

Proposition 71. *Let X be a topological space. The one-point compactification X_∞ is Hausdorff if and only if X is Hausdorff and has compact neigborhoods.*

The proof is left to you as homework.

Proper maps

We've seen that maps from compact spaces are very well-behaved. Oftentimes one can't actually work with compact domain spaces, but there is a next-best thing which does appear in real life. A proper map is a map which 'behaves like a map from a compact space would'.

Definition 38. *Let $f : X \to Y$ be a continuous map. We say f is* ***proper*** *if, for all compact sets $K \subset Y$, the inverse image $f^{-1}(K)$ is compact as well.*

For instance, any singleton set $\{y\}$ is compact, so each fiber $f^{-1}(y)$ is compact. People often think of properness as being like *compactness in families* — points in the codomain parameterize a family of compact sets in the domain (looking at preimages).

First we'll justify this intuition about what a proper map is. The idea in the following proof is a lot like the proof of the tube lemma.

Proposition 72 (Closed maps with compact fibers are proper)**.** *Let X and Y be spaces. Suppose $f : X \to Y$ is a continuous map which is closed and has compact fibers: that is, for all $y \in Y$, the inverse image $f^{-1}(y)$ is compact.*

Then f is proper.

[9] Some authors say 'X is locally compact', but there is another inequivalent property which also sometimes goes by that name, and probably deserves that name more.

Proof. Pick $K \subset Y$ compact; we want to show $f^{-1}(K)$ is compact.

It is going to save us a *lot* of notational heartache in what follows to change our conception of an open cover slightly. We'll refer to a collection $\mathcal{U}$ of open sets *in* X as an ambient open cover of $f^{-1}(K)$ if $f^{-1}(K) \subset \bigcup \mathcal{U}$ — so that every point $x \in f^{-1}(K)$ lies in some $U \in \mathcal{U}$, but so that $\mathcal{U}$ need not cover all of X. If we have an open cover $\mathcal{V}$ of $f^{-1}(K)$, then the definition of the subspace topology provides us with an ambient open cover

$$\mathcal{U} = \{U \subset X \mid U \text{ open, and } U \cap f^{-1}(K) \in \mathcal{V}\};$$

and intersecting an ambient open cover with $f^{-1}(K)$ provides an open cover in the usual sense. So these two notions are (roughly) equivalent. Certainly $\mathcal{U}$ has a finite subcollection that covers $f^{-1}(K)$ iff $\mathcal{V}$ does, so we may as well discuss compactness in terms of these ambient open covers.

Be warned that this notation is nonstandard.

Start with an ambient open cover $\mathcal{U}$ of $f^{-1}(K)$.

I want to start by making this collection *smaller*, even if it's not finite yet. For each $x \in K$, note that $\mathcal{U}$ also gives an ambient open cover of $f^{-1}(x) \subset f^{-1}(K)$. Since $f^{-1}(x)$ is compact, there are finitely many $U_{x,1}, \cdots, U_{x,n} \in \mathcal{U}$ whose union contains $f^{-1}(x)$.

Write

$$V_x = U_{x,1} \cup \cdots \cup U_{x,n}.$$

The set $\mathcal{V} = \{V_x\}_{x \in K}$ gives us an ambient open cover of $f^{-1}(K)$ — if we can show that it has a finite subcover, then we're finished, as each of the U_x's is a finite union of elements of $\mathcal{U}$.

The idea is to now transport these sets downstairs, to K, so that we can reduce it to only finitely many x — which would give us a finite subcover of $\mathcal{U}$. To make this precise, we have to be a little careful. We can't actually take $f(V_x)$, since f is not assumed to be an open map, but rather a *closed* map. So we should look at $f(V_x^c)$ instead. (This is where it's nice that the V_x's are open in X and not in $f^{-1}(K)$.)

Because V_x^c is closed, and f is a closed map, the image $f(V_x^c)$ is as well. Thus $f(V_x^c)^c$ is open.

Observation: $x \in f(V_x^c)^c$. This is equivalent to saying that $x \notin f(V_x^c)$, or that $f^{-1}(x) \cap V_x^c = \varnothing$. But remember that $f^{-1}(x) \subset U_x \subset V_x$, so that $f^{-1}(x) \cap V_x^c = \varnothing$, indeed. Therefore

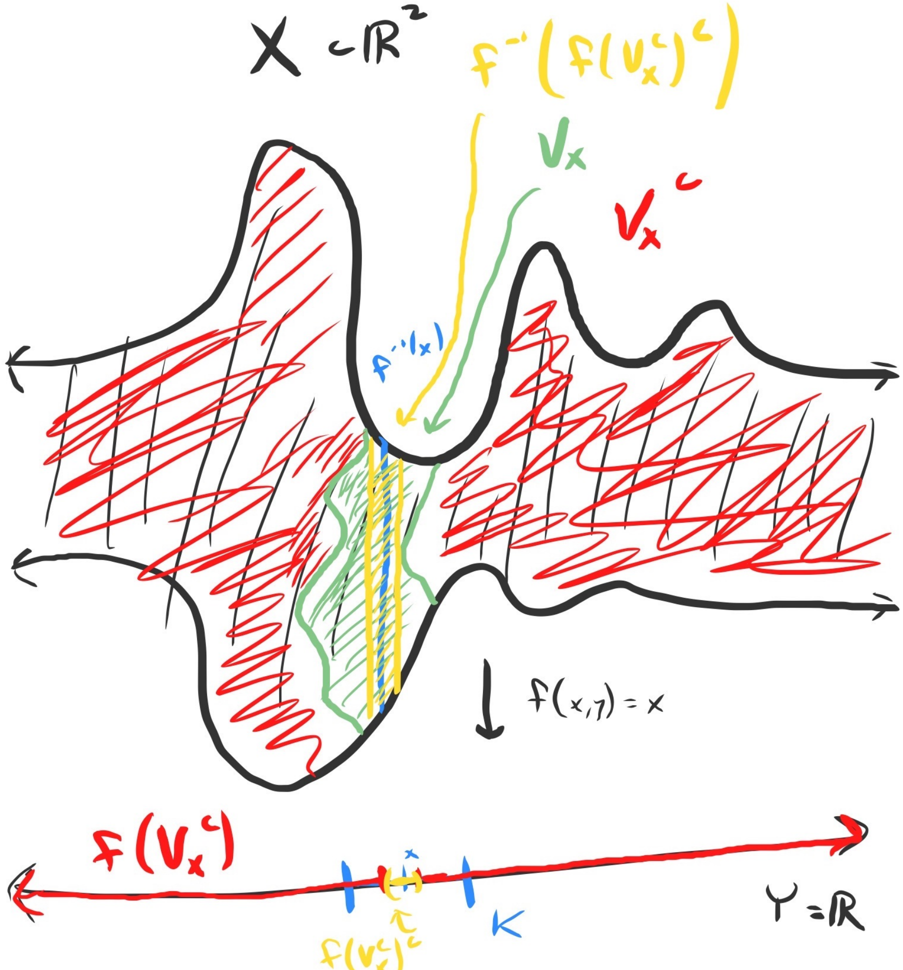
$$\mathcal{V}' = \{f(V_x^c)^c\}_{x \in K}$$

gives an ambient open cover of K in Y. By compactness of K, it therefore has a finite subcover.

Observation 2:

$$f^{-1}(f(V_x^c)^c) = f^{-1}(f(V_x^c))^c \subset (V_x^c)^c = V_x,$$

by applying the usual argument and taking complements. This very silly looking statement captures what the following picture intends to.



Upshot: if it only takes finitely many $f(V_x^c)^c$ to cover K (which it does!), then taking their preimages, we see that it only takes finitely many of the V_x's to cover $f^{-1}(K)$, and hence it only takes finitely many of the original $U_{x,i}$'s to cover $f^{-1}(K)$. So our cover has a finite subcover, as desired. $\square$

Theorem 73 (Proper maps induce maps on the one-point compactification)**.** *If $f : X \to Y$ is a proper map, then the map $f_\infty : X_\infty \to Y_\infty$ with*

$$f_\infty(x) = \begin{cases} x & x \neq \infty \\ \infty & x = \infty \end{cases}$$

is continuous if and only if f is.

Proof. Notice that the subspace topology on $X \subset X_\infty$ on X coincides with the topology it had originally. If f_∞ is continuous, then f is too: the composite $X \xrightarrow{i} X_\infty \xrightarrow{f_\infty} Y_\infty$ is continuous because i is and f_∞ is; its image lies in the subspace $Y \subset Y_\infty$; therefore restricting the codomain we see that $f : X \to Y$ is continuous.

Now for the harder direction. If f is continuous, we want to show that f_∞ is continuous. Pick $U \subset Y_\infty$ open. If $\infty \notin U$ then $U \subset Y$ is an open set, and $f_\infty^{-1}(U) = f^{-1}(U) \subset X \subset X_\infty$ is open in X (because f is continuous) and doesn't contain ∞, so defines an open subset of X_∞.

On the other hand, if $\infty \in U$, then by assumption $U^c \subset Y$ is both closed and compact. Now $f_\infty^{-1}(U)$ contains ∞, and

$$f_\infty^{-1}(U)^c = f_\infty^{-1}(U^c) = f^{-1}(U^c) \subset X \subset X_\infty.$$

Now f is proper, so $f^{-1}(U^c)$ is compact; and f is continuous, so $f^{-1}(U^c)$ is closed. Therefore $f_\infty^{-1}(U)^c$ is a closed and compact subset of X; because $f_\infty^{-1}(U)$ contains ∞, it follows that it is open, as desired. □

This leads to a great and totally immediate upshot, which (to me) justifies the discussion of proper maps.

Proposition 74 (The 'continuous-bijection' theorem for proper maps)**.** *Let $f : X \to Y$ be a proper map, where Y is a Hausdorff space with compact neighborhoods. If f is a bijection, then f is a homeomorphism. If f is injective, then f is a topological embedding (homeomorphism onto its image). Further, a topological embedding is proper if and only if $f(X)$ is closed in Y.*

Proof. By the assumption that f is proper, we get a continuous map $f_\infty : X_\infty \to Y_\infty$. The domain of f_∞ is compact and the codomain is Hausdorff, because Y is a Hausdorff space with compact neighborhoods.

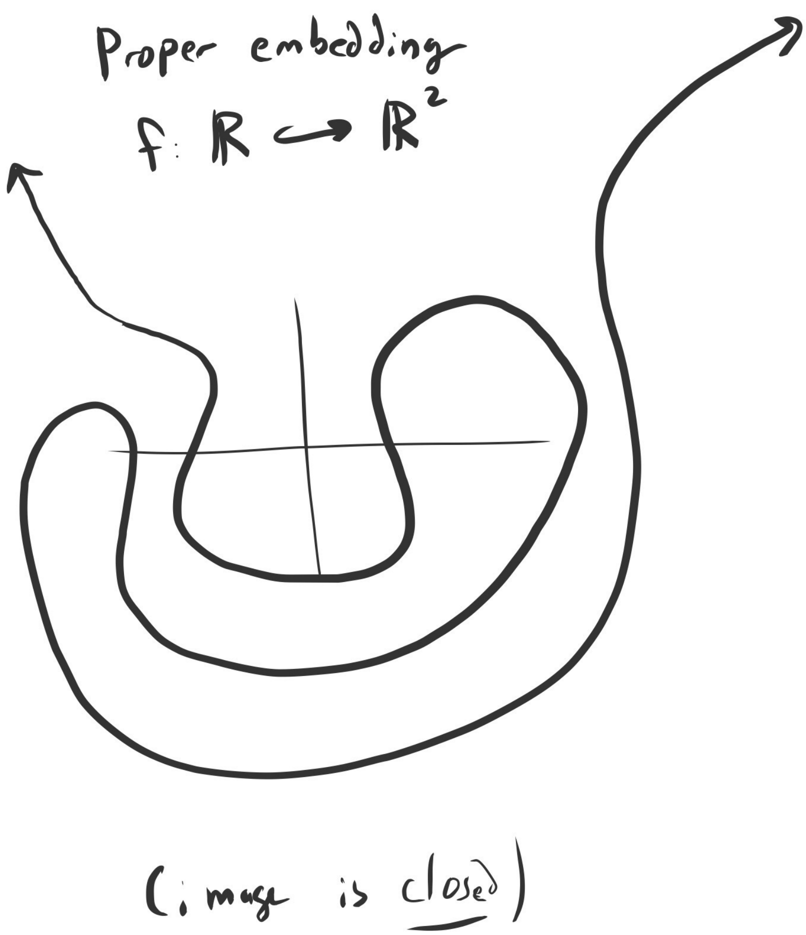
If f is a bijection, so is f_∞. Because the domain of f_∞ is compact and the codomain Hausdorff, it is a homeomorphism. Restricting the domain to X and the codomain to $f_\infty(X) = Y$, we see that f is a homeomorphism as well.

If f is injective, so is f_∞. The image of f_∞ is precisely $f(X)_\infty \subset Y_\infty$; restricting codomains and running the same argument we see that $f : X \to f(X)$ is a homeomorphism, and thus f is a topological embedding. Further, $f(X_\infty) = f(X)_\infty$ is a compact subset of a Hausdorff space, hence closed. Thus $f(X) = f(X)_\infty \cap Y$ is closed in the subspace topology (which is the original topology on Y), as desired.

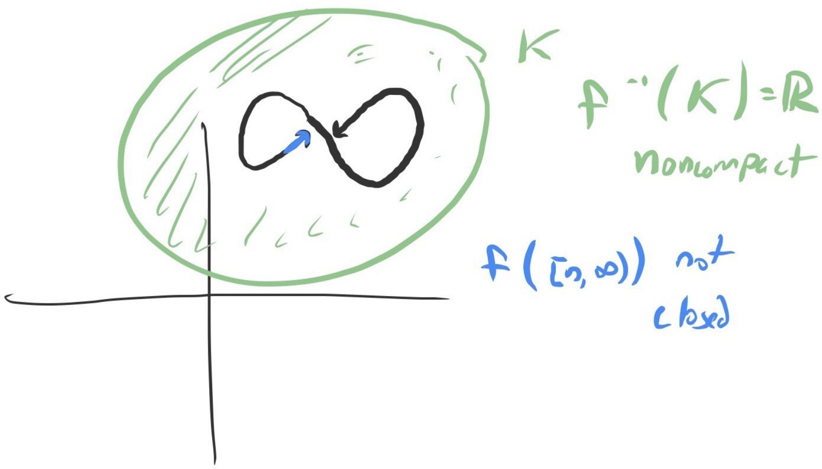
Now for the very last statement, suppose f is a topological embedding with $f(X)$ closed; we want to show f is proper. If $K \subset Y$ is any compact set, then $K \cap f(X)$ is a closed subset of the compact space K whence compact. If you write $g : f(X) \to X$ for the inverse of f, then

$$f^{-1}(K) = f^{-1}(K \cap f(X)) = \phi(K \cap f(X))$$

is the continuous image of a compact set, hence compact. So f is proper. □



Not an embedding
(__not__ proper)
($f(X)$ __is__ closed but f not a closed map)



Embedding but not proper
& $f(\mathbb{R})$ not closed

$f(\mathbb{R}) = S^1 \setminus \{i\}$

$\overline{f(\mathbb{R})} = S^1$

$f^{-1}(S^1) = \mathbb{R}$ __not__ compact

10/12: Quotient spaces

Quotient spaces are an incredibly useful and incredibly subtle topic. Most students who have taken a first course in topology leave without really understanding them — which I would blame on the fact that they're usually introduced as an aside and quickly moved on from!

We'll spend the next two weeks talking about quotient spaces — understanding the definition, seeing examples and building intuition, and seeing **weird** examples which destroy (or rather refine) that intuition.

Warning. The initial setup of quotient spaces can be confusing. I've thought long and hard about how avoidable this is, and I've come to the conclusion: it's kind of not. That's why we're going to be so grounded in examples; we'll see the definition, be baffled, and then start to get it.

Equivalence relations and quotients

For sets

You'll recall the notion of an equivalence relation from when we discussed connected components:

Definition 39. *Let X be a set. A relation $\sim$ on X is nothing more than a subset $G(\sim) \subset X \times X$. We write $a \sim b$ if $(a, b) \in G(\sim)$. We think of a relation as being a rule that says whether or not 'a is related to b', with no further conditions whatsoever, which we encode it by its 'graph' $G(\sim)$.*

An equivalence relation is a relation satisfying the three properties:

- *For all $x \in X$, we have $x \sim x$* (reflexivity)
- *If $x \sim y$, then we also have $y \sim x$* (symmetry)
- *If $x \sim y$ and $y \sim z$, then we also have $x \sim z$* (transitivity)

We will mainly appeal to the graph $G(\sim)$ for some technical statements below. Mostly, we just think of a relation as a rule that says whether or not $a \sim b$.

Definition 40. *If $\sim$ is an equivalence relation on a set X, an equivalence class is a subset*

$$C_x \subset X = \{z \in X \mid z \sim x\};$$

equivalently, an equivalence class is a subset for which all elements $y, z \in C_x$ are similar ($y \sim z$), and so that if $w \in X$ is similar to an element of C_x, then $w \in C_x$.

There is a set of equivalence classes

$$X/\sim = \{C_x \mid x \in X\};$$

if $x \sim y$, then the corresponding equivalence classes are equal: $C_x = C_y$.

Because X is partitioned into its equivalence classes (no point lies in two distinct equivalence classes), there is a canonical projection or quotient map $p : X \to X/\sim$, sending a point x to the equivalence class C_x.

Warning. Here we are abusing notation a little bit. C_x is **simultaneously** used to denote a *subset* $C_x \subset X$, as well as an *element* $C_x \in X/\sim$. This gives rise to the bizarre-looking relation for the quotient map p:

$$p^{-1}(C_x) = C_x.$$

Sometimes authors write $[x]$ for the equivalence class of x. These authors would write the above relation explicitly:

$$p^{-1}([x]) = \{y \in X \mid y \sim x\} = C_x \subset X.$$

I think this is obviously much less confusing. That in mind, I'm going to write $[x]$ for an element of $X/\sim$, and C_x for a subset of X, slightly contradicting my earlier notation.

This is analagous to how, for quotient groups, one might at the beginning write gH for an element of G/H, but later start to write $[g]$ instead, since why bother carrying around the baggage of H?

Remark 41. Why the word 'quotient'? Well, when we divide (divisible) integers, what does a/b mean? We have *divided* a into pieces (each of size b). Here we are dividing a space X into pieces C_x — unlike in the case of division of numbers, we do not demand they are all the same size!

The way I try to think of the set of equivalence classes is: it's what's left after I **identify** all points which are similar to each other; I imagine collapsing every equivalence class down to a point.

For spaces

This is a new way of constructing spaces, like the disjoint union, product, and subspace were. If you remember how those discussions went, you may be able to predict what comes next: if X is a topological space, and $\sim$ is an equivalence relation on X, there ought to be a topology on $X/\sim$ and $p : X \to X/\sim$ ought to be continuous!

Well, you're right.

Definition 41. *Let X be a topological space, and $\sim$ an equivalence relation on (the underlying set of) X.*
The quotient topology on $X/\sim$ has

$$U \subset X/\sim \text{ open} \iff p^{-1}(U) \subset X \text{ open.}$$

We can think about this upstairs: a subset S of X is called saturated *if $x \in S$ and $y \sim x$ implies that $y \in S$, too. (Equivalently, $x \in S \implies C_x \subset S$; equivalently, S is a union of equivalence classes.)*

If $U \subset X/\sim$ is open, the set $p^{-1}(U)$ is called a saturated open set *in X: it is an open subset of X which is a union of equivalence classes. Conversely, if $V \subset X$ is a saturated open set, then $p(V)$ is open (because $p^{-1}(p(V)) = V$).*

(The last sentence of this definition is really a proposition. This material is confusing: **it is a very useful exercise** to check your understanding by writing down a careful proof.)

Example 42. Consider $[0, 1]$, with the equivalence relation

$$t \sim s \iff t = s \quad \text{or} \quad t = 0, s = 1 \quad \text{or} \quad t = 1, s = 0.$$

(This is the simplest equivalence relation under which $0 \sim 1$; we are forced to have $1 \sim 0$ by symmetry, and we're forced to have $t \sim t$ by reflexivity.)

The points in $[0, 1]/\sim$ come in two types: for $0 < t < 1$, the point $[t] \in [0, 1]/\sim$ corresponds to the equivalence class $\{t\}$ upstairs. But for $t = 0$ or $t = 1$, the point $[t] \in [0, 1]/\sim$ corresponds to the equivalence class $\{0, 1\}$ upstairs.

The open intervals (a, b) for $0 < a < b < 1$ are saturated open sets, since $t \in (a, b)$ has equivalence class $C_t = \{t\}$; thus $t \in (a, b) \implies C_t \subset (a, b)$. These project to open sets in the quotient.

The interval $[0, b)$ for $b < 1$ (which is open in $[0, 1]$) does **not** project to an open set $p\big([0, b)\big)$ downstairs. It's not saturated: if it was, then we would have $1 \in [0, b)$. The smallest saturated set containing $[0, b)$ — which is $p^{-1}\Big(p\big([0, b)\big)\Big)$ — is $[0, b) \cup \{1\}$, which is not open. Thus $p\big([0, b)\big)$ is not open by definition of the quotient topology.

This little example is great, and we'll come back to it later to understand it even more explicitly.

First properties

The above is a *formal definition* of a quotient space. But it's rarely how anybody thinks about quotient spaces. The following proposition gives the basis for the way I think about quotients.

Proposition 75 (Universal property of quotient spaces)**.** *If X is a topological space, and $\sim$ is an equivalence relation on X, the projection map $p : X \to X/\sim$ is continuous.*

Suppose $f : X \to Y$ is a continuous map which is constant on equivalence classes: if $x \sim z$, then $f(x) = f(z)$. Then there is a continuous map

$$g : X/\sim \to Y,$$

defined by $g([x]) = f(x)$; we call g the map induced by f, and we have that $gp = f$.

Proof. If $U \subset X/\sim$ is open, then by definition $p^{-1}(U)$ is an open set! So p is continuous.

For the next part, we've already said how we should define g. But we need to check that the definition actually makes sense — remember that if $x \sim z$, we have $[x] = [z]$ (upstairs, we would say $C_x = C_z$). To see that this is *well-defined* (independent of the representative $y \in C_x$), we need to check that $g([x]) = g([z])$ when $[x] = [z]$. But this precisely means that $x \sim z$ — and we've assumed that $f(x) = f(z)$ whenever $x \sim z$. So g is well-defined. (This is all set-theory, and all of this works at the level of sets.)

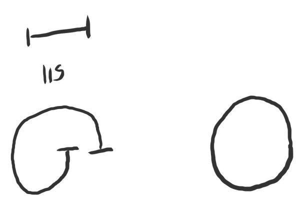
We also have $gp = f$, more or less tautologically —

$$(gp)(x) = g(p(x)) = g([x]) = f(x).$$

Now let's check that g is continuous. This is an application of the definition and the relation we just saw. Suppose $U \subset Y$ is open; we want to show that $g^{-1}(U)$ is open in $X/\sim$. By definition of the quotient topology, this is true iff $p^{-1}(g^{-1}(U))$ is open — but this double-inverse-image is precisely $(gp)^{-1}(U) = f^{-1}(U)$, and f is continuous. Thus $g^{-1}(U)$ is open in $X/\sim$, and so g is continuous, as desired. □

This makes some sense to me: the quotient space is what you get when you crush all equivalence classes down to points; and if f is constant on equivalence classes, you can certainly crush the equivalence classes down to points and still have a well-defined map. The proposition says that our topology was chosen *precisely* so that this map is still continuous. One says that f 'factors through' the quotient $X/\sim$.

This proposition is *almost always* how we actually reason about quotient spaces in practice: we try to avoid being explicit about open sets as much as possible and go back to the universal property. As an example:

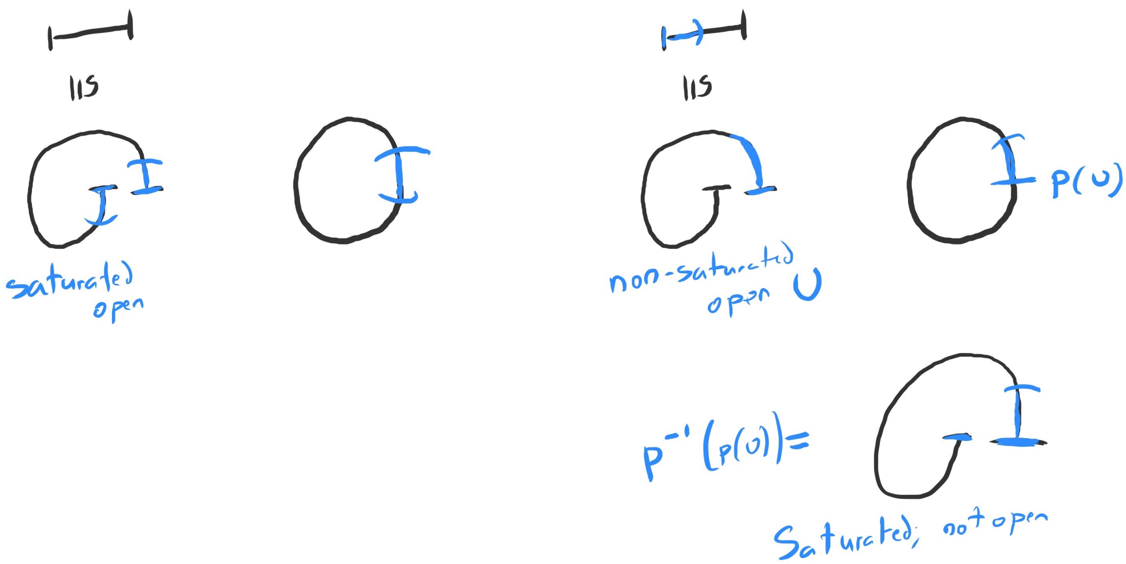


Proposition 76 (An interval with its endpoints glued together is a circle)**.** *The quotient from Example 42, written here as $[0,1]/\sim$, is homeomorphic to the circle.*

Proof. Consider the map $f : [0,1] \to S^1$, given by $f(t) = e^{2\pi i t}$. This is a continuous surjection, with $f(0) = f(1)$. Thus f is constant on the equivalence classes $\{0,1\}$ and $\{t\}$ for $0 < t < 1$ — it thus induces a continuous map $g : [0,1]/\sim \to S^1$.

g is a continuous bijection — note that f was surjective, and was injective on $(0,1)$; the only failure of injectivity is that $f^{-1}(1) = \{0,1\}$.

The domain is compact (it's the continuous image of the compact space $[0,1]$) and the codomain Hausdorff. Thus $g : [0,1]/\sim \to S^1$ is a homeomorphism. □



When we defined the relation $\sim$ on $[0,1]$, it was kind of cumbersome. It seemed a little silly, given that the main 'point' of the relation is that it had $0 \sim 1$ — everything else was given to us by the axioms of an equivalence relation. This inspires us to give the following.

Definition 42. *Let $\sim$ be a relation on X. The equivalence relation generated by $\sim$ is the equivalence relation $\hat{\sim}$ whose graph $G(\hat{\sim})$ contains $G(\sim)$ and is as small as possible. Said in another way, if $\sim$ is a relation, $\hat{\sim}$ is the relation with*

- $x \hat{\sim} x$,
- $x \hat{\sim} y$ *if there is a sequence $x = x_0, \cdots, x_n = y$ so that for each $1 \leqslant i \leqslant n$ either $x_{i-1} \sim x_i$ or $x_i \sim x_{i-1}$.*

In practice, we will abuse notation and write $\sim$ for the equivalence relation *generated by* the relation $\sim$.

The explicit formulation of $\hat{\sim}$ is not really important at all. Most of the time when I hand you a relation, it will already be transitive, and sometimes even symmetric — the point of this notion is that I don't want to write all that junk about what $\sim$ is when it really only does one important thing. For instance, in the above example, $\sim$ was the equivalence relation generated by the relation with $0 \sim 1$ and nothing else.

We write, for instance,

$$[0,1]/(0 \sim 1)$$

for the quotient space by the equivalence relation generated by the relation with $0 \sim 1$. This is much more compact, and says what the point *is*: we're pasting 0 and 1 together.

Our definition of 'quotient' is not necessarily the most common way people formulate these ideas. Often people talk about *quotient maps*, as in the following proposition.

Definition 43. *A continuous map $f : X \to Y$ is called a quotient map if f is surjective and a subset $U \subset Y$ is open iff $f^{-1}(U)$ is open.*

Proposition 77. *[The relationship between quotient maps and quotient spaces] If $f : X \to Y$ is a continuous map, consider the equivalence relation $\sim$ on X defined by*

$$x \sim z \iff f(x) = f(z).$$

Then the induced map $g : X/\sim \to Y$ is a homeomorphism if and only if f is a quotient map.

Proof. By the universal property of quotient spaces, there is indeed an induced continuous map $g : X/\sim \to Y$.

Let's first suppose f is a quotient map, and show that g is a homeomorphism.

First, g is a bijection: iff $g([x]) = g([z])$, then $f(x) = f(z)$ by definition. The definition of the above equivalence relation, this means that $x \sim z$, so that $[x] = [z]$. Thus g is injective. Further, we have assumed that f is surjective; so for all $y \in Y$, there is some $x \in X$ with $f(x) = y$. Correspondingly, we have $g([x]) = y$. So g is surjective.

To conclude we need to show that g is open. Suppose $U \subset X/\sim$ is open. This means precisely that $p^{-1}(U)$ is open in X. Because g is a bijection, we have that

$$p^{-1}(U) = p^{-1}(g^{-1}(g(U))) = (gp)^{-1}(g(U)) = f^{-1}(g(U)).$$

Thus $f^{-1}(g(U))$ is open. But because f is a quotient map, this means that $g(U)$ is open, as desired.

Next let's suppose g is a homeomorphism. To show that f is a quotient map, we need to show that it is surjective and that $U \subset Y$ is open if and only if $f^{-1}(U)$ is. Because g is surjective, the projection $p : X \to X/\sim$ is surjective, and $f = gp$, we see that f is surjective as well. Because g is a homeomorphism, it follows that $U \subset Y$ is open if and only if $g^{-1}(U)$ is open. By definition of the quotient topology, a subset $V \subset X/\sim$ is open if and only if $p^{-1}(V)$ is open. Therefore $U \subset Y$ is open if and only if

$$p^{-1}(g^{-1}(U)) = (gp)^{-1}(U) = f^{-1}(U)$$

is open, as desired. □

It follows that our idea of quotient in terms of the quotient space construction, and other authors' idea in terms of the notion of quotient map, are equivalent. There is no intrinsic value in one over the other. It will occasionally happen that one phrasing might be more immediately applicable, but we can always translate back and forth between them as desired.

I will only occasionally use the term 'quotient map'.

The following proposition is often useful in identifying quotient spaces. (We basically reasoned like this in the $[0,1]/\sim$ example.)

Proposition 78. *[Surjections from compact to Hausdorff are quotients] Let $f : X \to Y$ be a continuous **surjection** from a compact space to a Hausdorff space. Then if $\sim$ is the equivalence relation*

$$x \sim z \iff f(x) = f(z),$$

the induced map $g : X/\sim \to Y$ is a homeomorphism.

Other authors would conclude 'then f is a quotient map', but I think this statement shows more clearly what the point is!

Proof. Because $p : X \to X/\sim$ is continuous and surjective, and X is compact, so too is $p(X) = X/\sim$.

Because f is constant on equivalence classes (by definition of the equivalence relation!) it induces a continuous map $g : X/\sim \to Y$; in fact, a continuous bijection (as you saw in the proof of Proposition 77). Because the domain of g is compact and the codomain Hausdorff, it follows that g is a homeomorphism. □

10/14: Special cases and examples

There are two simple examples of quotient spaces which appear extremely often, and thus deserve special attention.

Definition 44. *Let X be a topological space, and $A \subset X$ a subspace. We write X/A for the quotient by the equivalence relation generated by $a \sim b$ if $a, b \in A$. Intuitively, we crush A down to a point, and leave everything else untouched.*

Example 43. The quotient

$$(S^{n-1} \times [0,1])/(S^{n-1} \times \{0\})$$

is homeomorphic to the closed unit disc D^n.

To see this, consider the map $f : S^{n-1} \times [0,1] \to D^n$ defined by $f(x,t) = tx$. This is a continuous map, and it's injective away from $S^{n-1} \times \{0\}$, while we have $f^{-1}(0) = S^{n-1} \times \{0\}$.

The equivalence relation of Proposition 78 is thus precisely the equivalence relation generated by $x \sim y$ if $x, y \in S^{n-1} \times \{0\}$. Because $S^{n-1} \times [0,1]$ is compact, and D^n is Hausdorff, the result follows from Proposition 78.

Example 44. Similarly, let $\sim$ be the relation on $S^{n-1} \times [-1,1]$ with $x \sim y$ for $x, y \in$ ‘$S^{n-1} \times \{-1\}$ and for $x, y \in S^{n-1} \times \{1\}$. (That is, we are crushing both the top sphere and the bottom sphere down to a point.) Then the quotient $S^{n-1} \times [-1,1]/\sim$ is homeomorphic to the unit sphere S^n

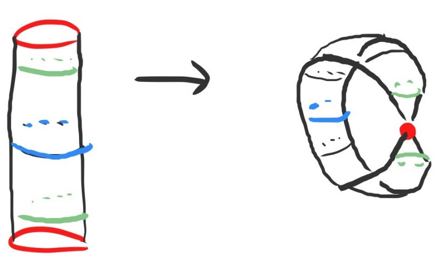
We'll do something along the lines of the previous example. Let $f : S^{n-1} \times [-1,1] \to \mathbb{R}^{n+1}$ be the map defined by

$$f(x,t) = (\sqrt{1-t^2}x, t).$$

This map is continuous because both of its components are, and its image is contained in S^n (as its Euclidean norm is $(1-t^2)\|x\|_2 + t^2 = (1-t^2) + t^2 = 1$). Thus we may restrict the codomain to obtain a continuous map to S^n. In fact, f surjects on to S^n (which you may check, if you like), and the preimages are all points, except for $S^{n-1} \times \{-1\}$ (the inverse image of the south pole) and $S^{n-1} \times \{1\}$ (the inverse image of the north pole). Because the domain is compact and the codomain Hausdorff, we may apply Proposition 4 to obtain the desired result.

Notice that the proof does what the picture tells us to! We're getting to the part of topology where we really just want to draw pictures all the time — they give us way more intuition for what's going on than the explicit functions do.

Note that collapsing the top sphere and bottom sphere to points (one for each) is different than collapsing both to a common point — the latter looks more like an earring, or a crescent moon with touching tips.



This allows us to generalize one of our first examples.

Example 45. Write D^n for the closed unit disc in Euclidean space and S^{n-1} for its boundary, the unit sphere. Then there is a homeomorphism $D^n/S^{n-1} \cong S^n$.

This follows immediately from the previous example and one of your homework problems, which proves that iterated quotients, like

$$D^n/S^{n-1} \cong \Big((S^{n-1} \times [0,1])/(S^{n-1} \times \{0\}) \Big)/(S^{n-1} \times \{1\}),$$

are quotients of the original space $S^{n-1} \times [0,1]$ by an explicit equivalence relation — which in this case collapses the top sphere and bottom sphere of $S^{n-1} \times [0,1]$ down to points.

Similarly, we have...

Example 46. If S^n is the n-sphere, and $H \subset S^n$ is the **closed lower hemisphere**, then S^n/H is homeomorphic to S^n itself.

Example 47. If $S^1 \subset S^2$ is the equator, the quotient S^2/S^1 is homeomorphic to two spheres which meet tangentially at a common point. (Can you visualize the continuous surjection which is injective except along the equator, which all gets sent to one point?)

I would write this formally if it didn't seem like a hassle to give the explicit formulas. ☺ However, if you see the picture, you can translate that into formulas and back again. (If you have trouble doing so, reach out to me! I'd love to help you understand this.)

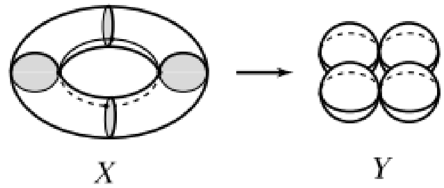
Example 48. Let $T^2 = S^1 \times S^1$ be the 2-dimensional torus, where we write points of $S^1 \subset \mathbb{C}$ in complex number notation. Consider four circles

$$S^1 \times \{1\}, \quad S^1 \times \{i\}, \quad S^1 \times \{-1\}, \quad S^1 \times \{-i\}.$$

If you collapse these circles in succession (equivalently, quotient by the equivalence relation in which all the points in any given circle are identified), what remains is homeomorphic to four spheres, touching tangentially.

Picture stolen from Hatcher's Algebraic Topology book:

These are all examples of *collapsing disjoint subspaces*. Another (perhaps even more common) example of a quotient space is a *pasting operation*.



Definition 45. *Let X and Y be topological spaces. Suppose $A \subset X$ is a subspace, and $f : A \to Y$ a topological embedding (homeomorphism onto $f(A)$). Then we write*

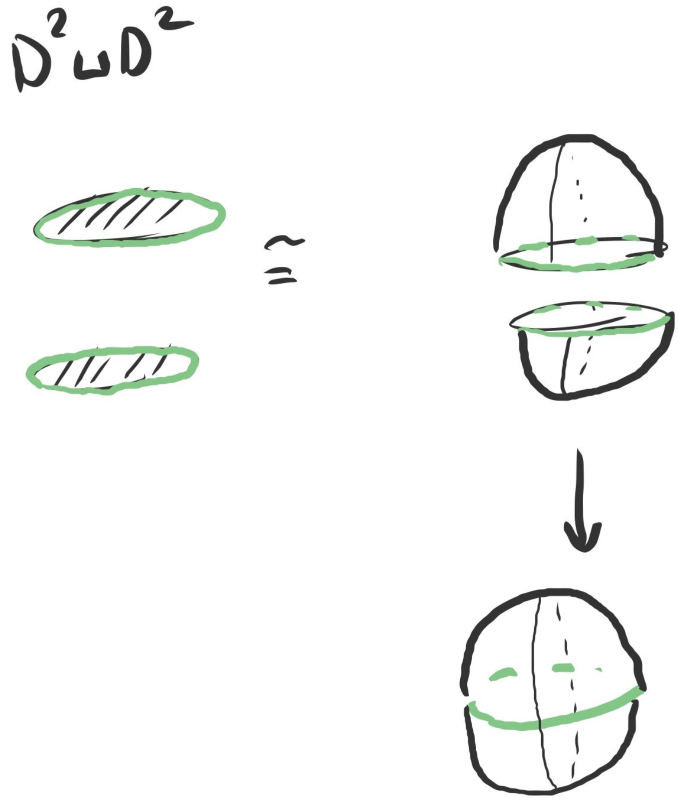
$$X \sqcup_f Y = X \sqcup Y/(a \sim f(a));$$

that is, we quotient by the equivalence relation generated by $a \sim f(a)$, whenever $a \in A$. This is sometimes called an adjunction space.

When A is also a subspace of Y in an obvious way, we usually do not write down the homeomorphism f, and just write $X \sqcup_A Y$.

Example 49. Pasting two discs along their boundary is homeomorphic to the sphere. That is,

$$D^n \sqcup_{S^{n-1}} D^n \cong S^n.$$



To see this, consider the two maps

$$f_\pm : D^n \to S^n \subset \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R},$$

given by

$$f_\pm(x) = (x, \pm\sqrt{1 - \|x\|_2}).$$

These combine to give a map $f = f_+ \sqcup f_- : D^n \sqcup D^n \to S^n$. This is a surjective map which is injective on D^n; the images only intersect on the equator, which is mapped to identically by the boundary of both copies of D^n.

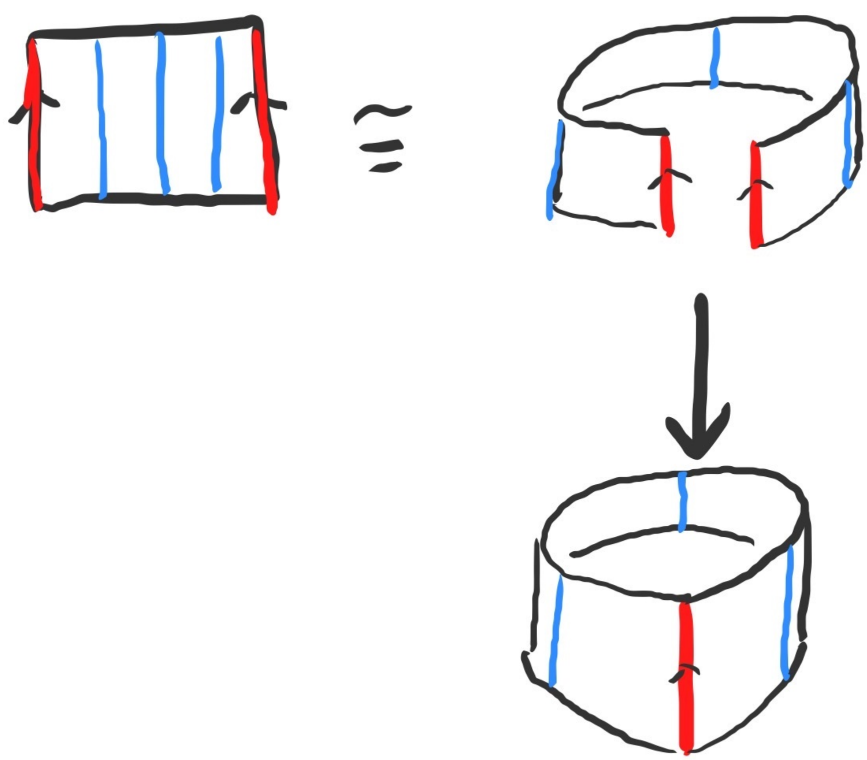
Writing $(x, 1)$ or $(x, 2)$ for the copies of $x \in D^n$ in each of the two factors, the preimages of points are thus either points themselves, or pairs $\{(x, 1), (x, 2)\}$ for $x \in S^{n-1}$. The quotient by the correpsonding equivalence relation is precisely $D^n \sqcup_{S^{n-1}} D^n$.

The result thus follows from Proposition 78.

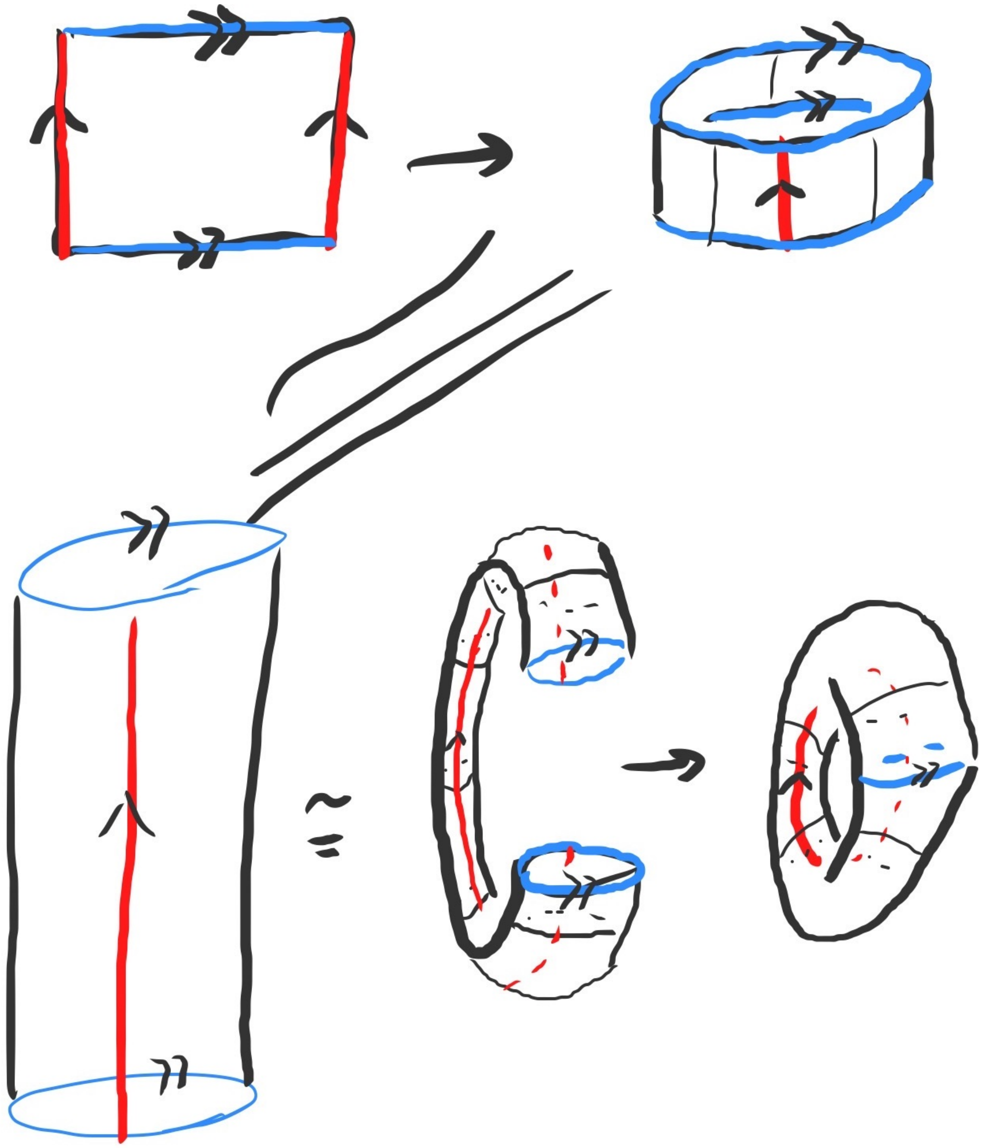
You can even paste parts of a space onto other parts of the same space. This is not usually notated in any special way, we just write down the relation we're quotienting by. I'll draw some examples below. For the first two, can you write down explicit maps, like above, proving that these are the quotients I claim they are?

Example 50. The cylinder and torus are both quotients of the unit square $[0, 1]^2$.

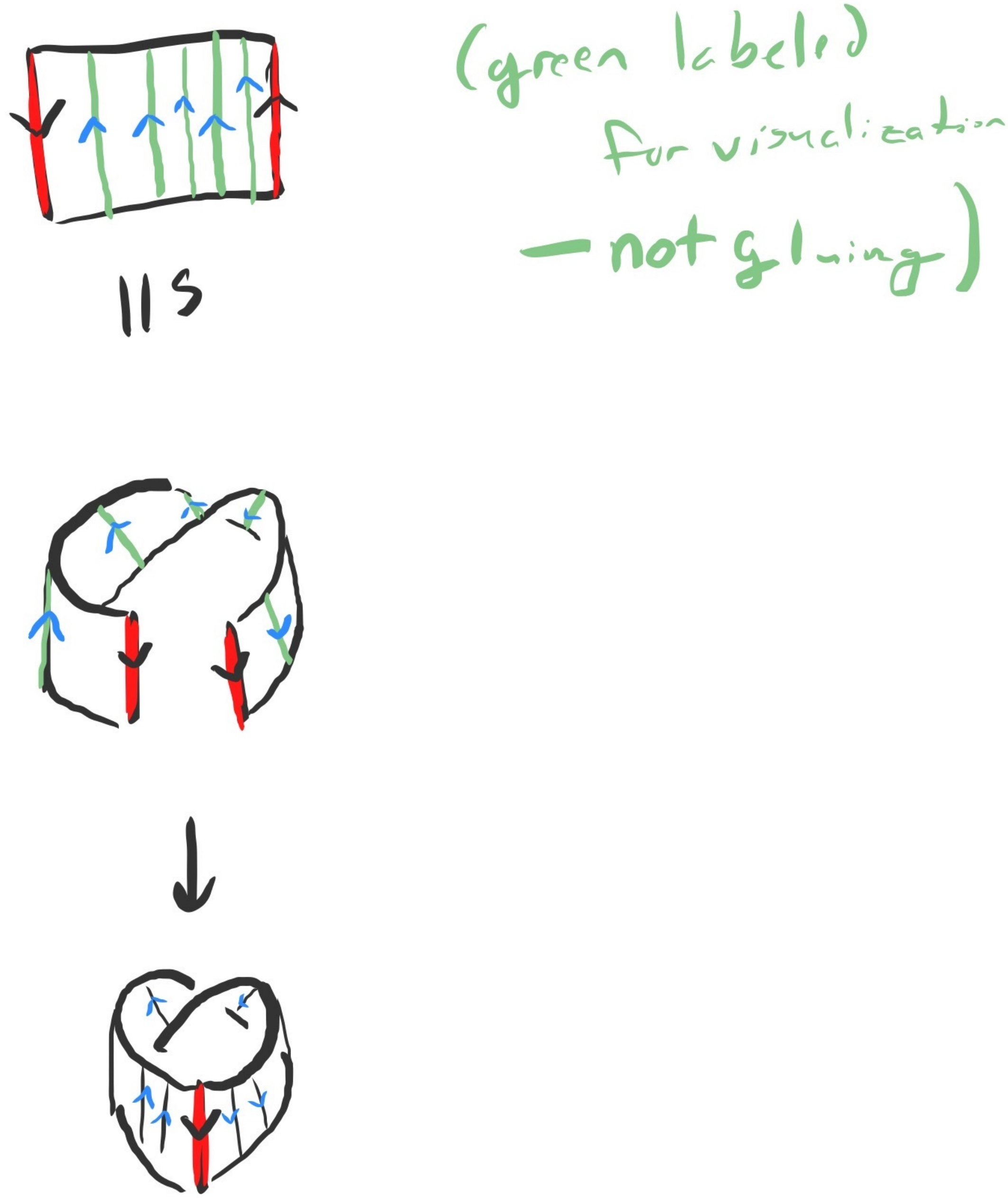
$$S^1 \times [0,1] \cong [0,1] \times [0,1] \,/\, (0,t) \sim (1,t)$$



$$[0,1]\times[0,1] \Big/ \begin{matrix}(0,t)\sim(1,t)\\ (s,0)\sim(s,1)\end{matrix} \cong T^2$$



Example 51. So is the Mobius band!



10/19: Group actions and their quotients

Definition 46. *Let X be a topological space. A group action on X is the data of a group G and a map $\alpha : G \times X \to X$ satisfying the following properties:*

- $\alpha(e, x) = x$*;*
- $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$*;*
- *For all $g \in G$, the map $\alpha_g : X \to X$ given by $\alpha_g(x) = \alpha(g, x)$, is continuous.*

Usually we do not write a group action as α; we often just record what α_g is for all g. Including 'α' in our notation is cumbersome. Just like products in groups $m(g, h)$ are usually just written gh, we usually write the action of g on an element $x \in X$ as

$$\alpha(g, x) =: g \cdot x;$$

and often we specify an action by specifying what $g \cdot x$ means for all $g \in G$ and $x \in X$.

Remark 52. Let X be a topological space. Because the inverse of a homeomorphism is a homeomorphism, and the composite of two homeomorphisms is a homeomorphism, and composition is associative, the set $\text{Homeo}(X)$ of all homeomorphisms of X forms a group.

You should check your understanding by showing that a *group action* in the sense above is uniquely determined by a group homomorphism $\alpha : G \to \text{Homeo}(X)$, and vice versa. (Hint: to show that α_g is a homeomorphism, first show that $\alpha_g^{-1} = \alpha_{g^{-1}}$.)

Group actions appear *all over the place* in math. We'll spend today going over a couple examples, the notion of quotients by group actions, and how to find simpler presentations of quotients when they're available.

Example 53. Probably one of the simplest examples is the action of $\mathbb{Z}$ on $\mathbb{R}$ by translation:

$$\alpha(n, t) = n + t, \quad \text{aka} \quad n \cdot t = n + t.$$

You can check that this is a group action without too much difficulty. In particular, the translation map $\alpha_n : \mathbb{R} \to \mathbb{R}$ is continuous because we know addition of real numbers is continuous.

Given a group action, the most common topological construction one can take is the *orbit space*, also called the

Definition 47. *Let $\alpha : G \curvearrowright X$ be a group action. An orbit of α is a set of points*

$$O_x = G \cdot x = \{g \cdot x \mid g \in G\},$$

the set of all points G can move a given point x to.

The group action α gives rise to an equivalence relation $\sim_\alpha$ on X: we say

$$x \sim_\alpha y \iff \exists_{g \in G} 7;\ gx = y.$$

That is, $x \sim_\alpha y$ if and only if x, y are in the same G-orbit (that is, $y \in G \cdot x$). The equivalence classes are precisely the G-orbits.

Then the quotient by the group action or the orbit space is the quotient space

$$X/G := X/{\sim_\alpha}.$$

Let's try our hand at using this definition with our first nontrivial example of group actions.

Proposition 79 (The circle is homeomorphic to $\mathbb{R}/\mathbb{Z}$)**.** *Let $\mathbb{Z}$ act on $\mathbb{R}$ by translations, as above. Then $\mathbb{R}/\mathbb{Z} \cong S^1$.*

Proof. Consider the map $f : \mathbb{R} \to S^1$ given by $f(t) = e^{2\pi i t}$. Because

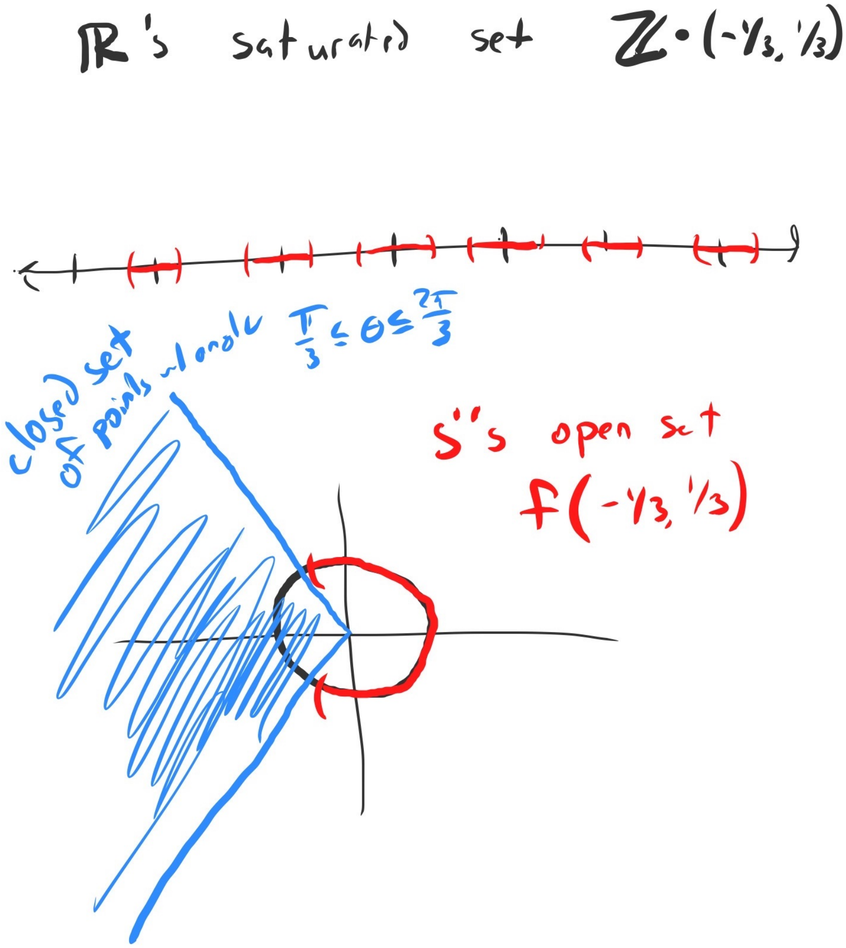
$$f(t+n) = e^{2\pi i(t+n)} = \cos(2\pi t + 2\pi n) + i\sin(2\pi t + 2\pi n) = \cos(2\pi t) + i\sin(2\pi t) = e^{2\pi i t} = f(t),$$

we see that f is constant on the orbits of the group action. Therefore, by the universal property of quotient spaces, there is a continuous induced map $g : \mathbb{R}/\mathbb{Z} \to S^1$.

This map is a continuous bijection — $(\cos t, \sin t) = (\cos s, \sin s)$ if and only if $t - s = 2\pi n$ for some n, essentially because these functions are 2π-periodic.

We will soon see a trick that makes it easier to conclude, but let's explicitly show that f is open, as that may help guide your intuition. Because forward-image commutes with unions, it suffices to show that $f(a, b)$ is open for all open intervals (a, b).

If $b - a > 2\pi$, then $f(a, b)$ is the entire unit circle, which is certainly open. So let's assume $b - a \leqslant 2\pi$. Then $f(a, b)$ is a circular arc between the endpoints $f(a)$ and $f(b)$ (which are distinct if $b - a < 2\pi$, but equal if $b - a = 2\pi$). Because the portion of the plane lying strictly between the two angles $f(a)$ and $f(b)$ (and excluding zero) is open, and its intersection with the circle is $f(a, b)$, it follows that $f(a, b)$ is open in the circle, as desired. ☐



You may at this point remember that we proved a similar-looking theorem with a similar proof: that

$$[0, 1]/(0 \sim 1) \cong S^1.$$

How do these relate?

We introduce two lemmas to help us work with orbit spaces. The first introduces the a notion that will return when you study algebraic topology, but we will not use it in this course. I hate the name. It's very confusing. But it's also well-established. **You may want to skip the proof on a first read; we will skip the proof in class. I include it because you're ready to learn it now, if you want, and you may want to understand it later.**

Lemma 80 (Properly discontinuous actions often give Hausdorff quotients)**.** *We say that $G \curvearrowright X$ is a properly discontinuous action if, for all compact subsets $K \subset X$, the set*

$$\{g \in G \mid g \cdot K \cap K \neq \varnothing\}$$

is finite.

If X has compact neighborhoods, is Hausdorff, and G is a properly discontinuous action, then the space X/G is Hausdorff.

(To understand the definition, think about the case of $\mathbb{Z}$ acting on $\mathbb{R}$. This is a properly discontinuous action.)

The idea will be: start by separating x from y in X by open sets. Then see if we can *shrink* these so that V is still disjoint from $G \cdot U$ (the union of all translates of U). Then we'll see that, for free, $G \cdot U \cap G \cdot V = \varnothing$.

The finiteness assumption will be used together with the locally compact assumption in the shrinking process.

Proof. Pick $[x], [y]$ distinct points in X/G. Because X is Hausdorff and locally compact, we can choose sets $x \in U_c, U_h$ and $y \in V_c, V_h$ so that $\overline{U_c}, \overline{V_c}$ are compact and $U_h \cap V_h = \varnothing$; the U_c, V_c are guaranteed by the local compactness condition, and the U_h, V_h are guaranteed to exist by the Hausdorff condition.

Taking the intersection, we see that $U = U_c \cap U_h$ and $V = V_c \cap V_h$ have both of these properties — $\overline{U}$ is closed in $\overline{U_c}$, so that $\overline{U}$ is compact (and similarly $\overline{V}$); and $U \cap V \subset U_h \cap V_h = \varnothing$.

Note that since U, V are open, and $U \subset V^c$, it follows also that

$$\overline{U} \subset \overline{V^c} = V^c,$$

as closure preserves containment, V^c is a closed set, and the closure of a closed set is the same set. In particular, $y \in V$ is not in $\overline{U}$, and x is not in $\overline{V}$.

Now we'll want to start our shrinking argument. Write $K = \overline{U} \cup \overline{V}$. By the assumption that the action is properly discontinuous, there are only finitely many elements $e = g_0, g_1, \cdots, g_n$ so that $g \cdot K \cap K \neq \varnothing$.

In particular, there are only finitely many $g_1, \cdots, g_n$ so that $\overline{U} \cap (g_i \cdot \overline{V})$ is nonempty.

Now we start our shrinking process. We know that we may choose open sets $U_1, \cdots, U_n$ around x and $V_1, \cdots, V_n$ around $g_i \cdot y$ so that $U_i \cap V_i \neq \varnothing$; note that because $[x] \neq [y]$ we have $x \neq g_i \cdot y$ for any g_i. Now replace U with

$$x \in W_x = U \cap U_1 \cap \cdots \cap U_n,$$

and V by

$$y \in W_y = V \cap \bigcap_{i=1}^{n} g_i^{-1} V_i.$$

Here we use that α_{g_i} is continuous to see that $g_i^{-1} \cdot V_i$ is still open.

Now notice that we have

$$W_x \cap g \cdot W_y \subset U \cap g \cdot V \subset \overline{U} \cap g \cdot \overline{V};$$

this can only be nonempty if $g = g_i$ for one of our g_i above. But in that case,

$$W_x \cap g_i \cdot W_y \subset U_i \cap V_i = \varnothing.$$

So we see that $W_x \cap g \cdot W_y$ is empty for all g. To put this another way,

$$G \cdot W_y = \bigcup_{g \in G} g \cdot W_y$$

is a saturated open set which does not intersect W_x.

I claim now that $g \cdot W_x \cap h \cdot W_y$ is also empty for all $g, h \in G$. To see this, simply note that

$$g^{-1} \cdot (g \cdot W_x \cap h \cdot W_y) = W_x \cap (g^{-1}h) \cdot W_y,$$

which we already know is empty. It thus follows that both $G \cdot W_x$ and $G \cdot W_y$ are saturated open sets containing x and y, respectively, which do not intersect.

Projecting these to X/G, their images are open (as these are saturated open sets in X) and disjoint (as the saturated open sets do not intersect). Therefore $[x]$ and $[y]$ may be separated by open sets in X/G, and so X/G is Hausdorff. □

We will use this to write down simpler quotients that give rise to the same quotient space. It will help us visualize group quotients and work with examples.

Lemma 81. *[If a compact domain meets every orbit, the orbit space is a quotient of that domain] Let X be a topological space, and $G \rightharpoondown X$ be a group action such that X/G is Hausdorff. If $D \subset X$ has*

$$\forall_{x \in X} \exists_{g \in G} g \cdot x \in D,$$

then write $\sim_\alpha$ for the relation on D given by

$$x \sim_\alpha y \iff \exists_{g \in G} g \cdot x = y;$$

the equivalence classes of $\sim_\alpha$ on D are the sets $G \cdot x \cap D$, the part of a G-orbit contained in D.

If D is compact, then the natural map $D/{\sim_\alpha} \to X/G$ is a homeomorphism.

Proof. The given assumptions imply that the composite

$$D \xrightarrow{i} X \xrightarrow{p} X/G$$

is surjective; call that composite $pi = q$.

Immediately from the definition, we see that $x, y \in D$, we have $x \sim_\alpha y$ if and only if $q(x) = q(y)$. It follows from this that there is an induced map $\overline{q} : D/\sim_\alpha \to X/G$ which is also a bijection. Now observe that the domain is compact (as a quotient of a compact space) and the codomain Hausdorff, so that this continuous bijection is in fact a homeomorphism. □

These out of the way, we can explore examples in detail.

Example 54. Earlier we studied the orbit space $\mathbb{R}/\mathbb{Z}$. The action of $\mathbb{Z}$ on $\mathbb{R}$ is properly discontinuous: because every K is bounded, it suffices to check for $K = [-r, r]$, but then $n \cdot K \cap K \neq \varnothing$ is only possible if $|n| \leqslant 2r$, and so there are only finitely many g with $g \cdot K \cap K \neq \varnothing$.

Notice that $D = [0, 1]$ satisfies the conditions in the lemma: it is compact, and if $x \in \mathbb{R}$, then writing $\lfloor x \rfloor$ for the greatest integer less than x, we have

$$(-\lfloor x \rfloor) \cdot x \in [0, 1) \subset [0, 1].$$

Notice further that the relation $\sim_\alpha$ on $[0, 1]$ is almost trivial — the only points $x, y \in [0, 1]$ with $n \cdot x = y$ for some $n \neq 0$ are $\{x, y\} = \{0, 1\}$ (every number strictly between 0 and 1 is moved outside the interval by translation).

Thus $\sim_\alpha$ on $[0, 1]$ is the relation generated by $0 \sim 1$. By the previous lemma, we can conlude that

$$\mathbb{R}/\mathbb{Z} \cong [0, 1]/(0 \sim 1) \cong S^1.$$

Often, a D chosen as in the above example (where the only gluing that happens is on its boundary) is called a 'fundamental domain'.

Example 55. The action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ by

$$(n, m) \cdot (x, y) = (x + n, y + m)$$

is also properly discontinuous (you should check this). The map

$$f : \mathbb{R}^2 \to T^2,$$

given by $f(t, s) = (e^{2\pi it}, e^{2\pi is})$, is constant on orbits — $f(t + n, s + m) = f(t, s)$. It therefore gives rise to a quotient map $g : \mathbb{R}^2/\mathbb{Z}^2 \to T^2$, which is a continuous bijection (check that it is a bijection).

T^2 is Hausdorff, so it would suffice to show that the domain is compact. But if $p : \mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2$ is the quotient map, then note that $p([0, 1]^2) = \mathbb{R}^2/\mathbb{Z}^2$ — and $[0, 1]^2$ is compact, so the domain of g is compact as well, hence a homeomorphism.

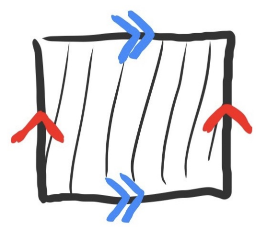
But this actually inspired us to look more seriously at $[0, 1]^2$. That satisfies the hypotheses of Lemma 81, so we have $[0, 1]^2/\sim_\alpha \cong T^2$. What is this relation?

Note that no two translates of $(0, 1)^2$ intersect each other, so the relation $\sim_\alpha$ only identifies points on the boundary. Explicitly, we have $(t, 0) \sim_\alpha (t, 1)$, and $(0, s) \sim_\alpha (1, s)$, and every other relation (like $(0, 0) \sim (0, 1) \sim (1, 0) \sim (1, 1)$) is generated by these.

Thus we may write T^2 as

$$[0, 1]^2/\big((t, 0) \sim (t, 1), \quad (0, s) \sim (1, s)\big).$$

This is usually drawn as in the following picture:



Example 56. Let $\mathbb{Z}$ act on $\mathbb{R}^2$ by

$$n \cdot (x, y) = (x + n, (-1)^n y);$$

the action of $1 \in \mathbb{Z}$ is to move the plane one unit to the right then flip it upside down. The quotient $\mathbb{R}^2/\mathbb{Z}$ by this action usually called the *open Mobius strip*. (It's important to realize this is *not* the action by translation — the quotient by that action is homeomorphic to an open cylinder.) The open Mobius strip is homeomorphic to the quotient of $\mathbb{R} \times (-\epsilon, \epsilon)$ by the same action — can you prove it?

The previous lemmas are better suited to dealing with the *closed Mobius strip*, given as

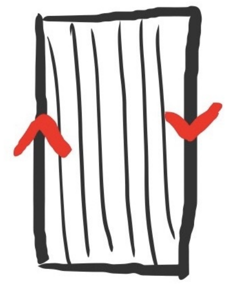
$$(\mathbb{R} \times [-1, 1])/\mathbb{Z},$$

with the action defined by the same formula as above. This has a fundamental domain given by $[0, 1] \times [-1, 1]$. The relation $\sim_\alpha$ in this case is given as $(0, t) \sim (1, -t)$, and is usually drawn as in the following picture:

10/21: The real projective plane: A case study

Today I want to focus on an object which is rather difficult to visualize — and not just because of its dimension.

Intuitively, the *real projective space* $\mathbb{RP}^n$ is the *space of lines through the origin in* $\mathbb{R}^{n+1}$. It is not at all obvious how this ought to be formally defined. Sure, I know the underlying set — each point in $\mathbb{RP}^n$ is



a line through the origin, and two points are distinct if they correspond to different lines. But what is the topology?

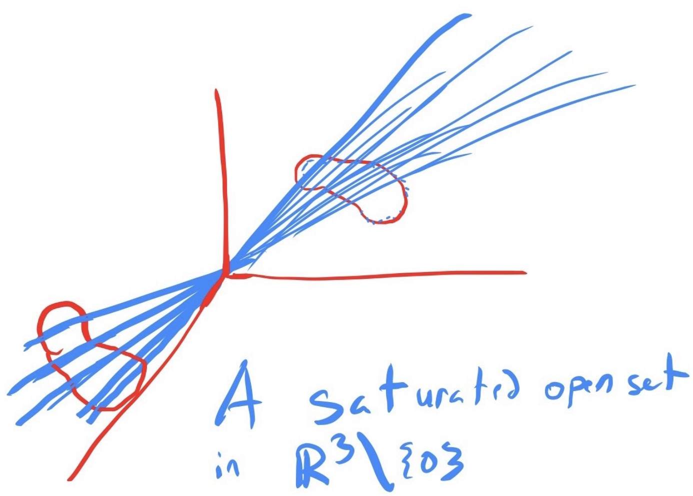
For the following definition to make sense, note that two **nonzero** points $x, y \in \mathbb{R}^{n+1}$ are in the same line if, and only if, $x = \lambda y$ for some $\lambda \in \mathbb{R}^\times := \mathbb{R}\backslash\{0\}$. And of course, every line through the origin contains some nonzero point! So we may as well focus in on the nonzero points in $\mathbb{R}^{n+1}$.

Definition 48. *Consider $\mathbb{R}^\times$, the set of nonzero real numbers, equipped with the group structure given by multiplication of real numbers.*

There is an action of $\mathbb{R}^\times$ on $\mathbb{R}^{n+1}\backslash\{0\}$, given by scaling: $\lambda \cdot x = \lambda x$.

*The n-*dimensional real projectivespace*, written $\mathbb{RP}^n$, is defined as the quotient space*

$$\mathbb{R}P^n = \left(\mathbb{R}^{n+1}\backslash\{0\}\right)/\mathbb{R}^\times.$$



Our first visual aids will be two facts that let me view $\mathbb{RP}^n$ as a less complicated quotient than the above.

Proposition 82 (Projective space is a quotient of the sphere)**.** *Consider the n-sphere $S^n \subset \mathbb{R}^{n+1}$. The group $\{\pm 1\}$ acts on S^n by scaling, too:*

$$(-1) \cdot x = -x.$$

(This map, sending a point to its negative / opposite, is called the antipodal map.) There is a natural homeomorphism

$$S^n / \pm 1 \cong \mathbb{RP}^n.$$

This group quotient is sometimes written

$$S^n/(x \sim -x).$$

Proof. We will want to invoke Lemma 81. Note that all we really need there is that X/G is Hausdorff, and that's true in this case; I won't give a careful proof, but can you draw a picture justifying why? (If you have two lines through the origin in 3-space, do you see how to get small 'bundles of lines' around each, which do not intersect except at the origin?)

Note that every $\mathbb{R}^\times$-orbit intersects the sphere, so it is a suitable choice of D — precisely, if $x \in \mathbb{R}^{n+1}\backslash 0$, then

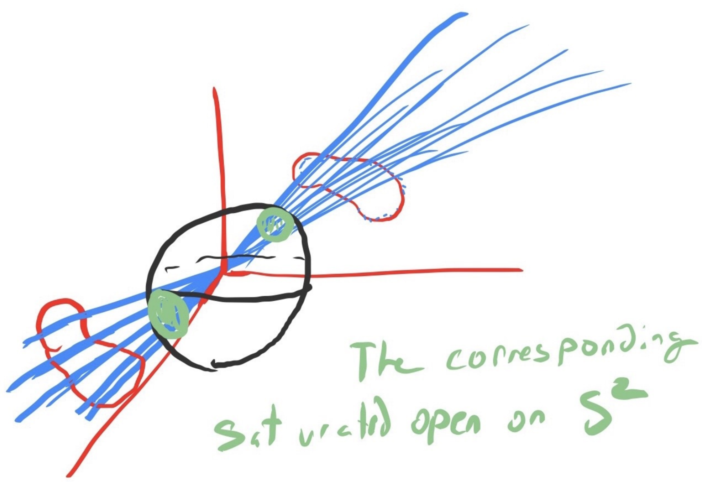
$$\frac{1}{\|x\|} \cdot x \in S^n.$$

So it suffices to figure out what the induced relation $\sim_\alpha$ is on the sphere S^n. Remember that $x \sim_\alpha y$ iff there is some $\lambda \in \mathbb{R}^\times$ with $\lambda x = y$. Note that if this is the case, then

$$|\lambda| = |\lambda|\|x\| = \|\lambda x\| = \|y\| = 1,$$

so that $\lambda \in \{\pm 1\}$; so that if $x \sim_\alpha y$, then either $x = y$ or $x = -y$. This is exactly the relation given by the action of ± 1 on S^n — if that action is called β, then $x \sim_\beta y$ iff $x = y$ or $x = -y$.

Thus the result follows as we obtain a continuous bijection $S^n/\pm 1 \to \mathbb{RP}^n$ from a compact space to a Hausdorff space. ☐



This is how one often thinks of projective space — as *the sphere with antipodal points identified.* This is not a bad definition, it works pretty well for geometry. But it's still not quite suitable (to me) for visual intuition.

Proposition 83. *Let $\sim$ be the relation on D^n, given by $x \sim y$ if either $x = y$ or $x, y \in S^{n-1} \subset D^n$ and $y = -x$. Then there is a homeomorphism $D^n/\sim \to \mathbb{RP}^n$.*

Proof. This will follow once more from Lemma 81. Note that D^n is homeomorphic to the closed upper-hemisphere of the sphere, by the map $f(x) = (x, \sqrt{1 - \|x\|^2})$.

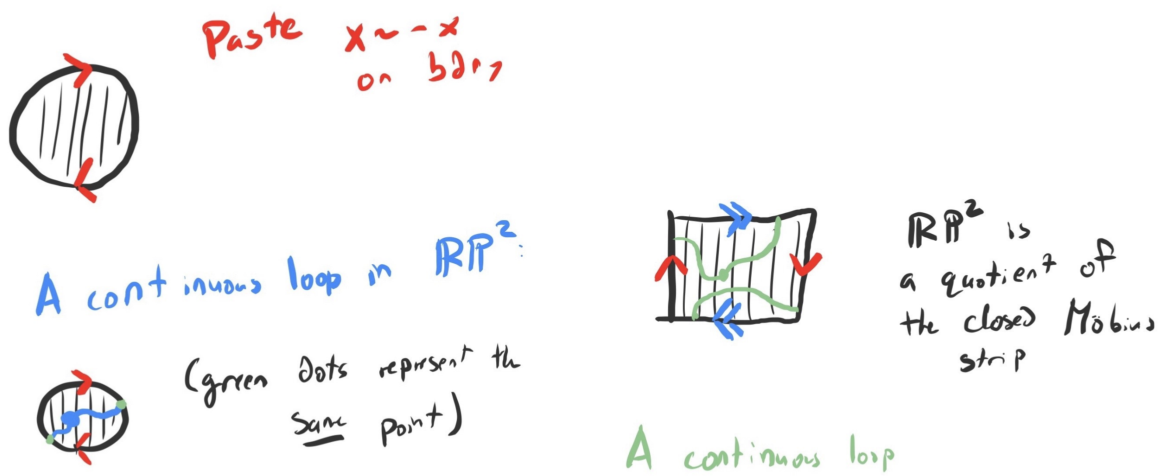
Second, note that the upper hemisphere is a fundamental domain for the action of ± 1 on S^n — if x is in the open lower hemisphere, then $-x$ is in the open upper hemisphere. So $D^n/\sim_\alpha \cong \mathbb{RP}^n$ by Lemma 81, so it is just up to us to determine what $\sim_\alpha$ is. But by this same observation, the only points $x \neq y \in D^n$

with $x = -y$ are those on the equator (corresponding to S^{n-1}, the boundary of D^n). So $\sim_\alpha$ is the relation generated by $x \sim -x$ for $x \in S^{n-1}$, as claimed, and

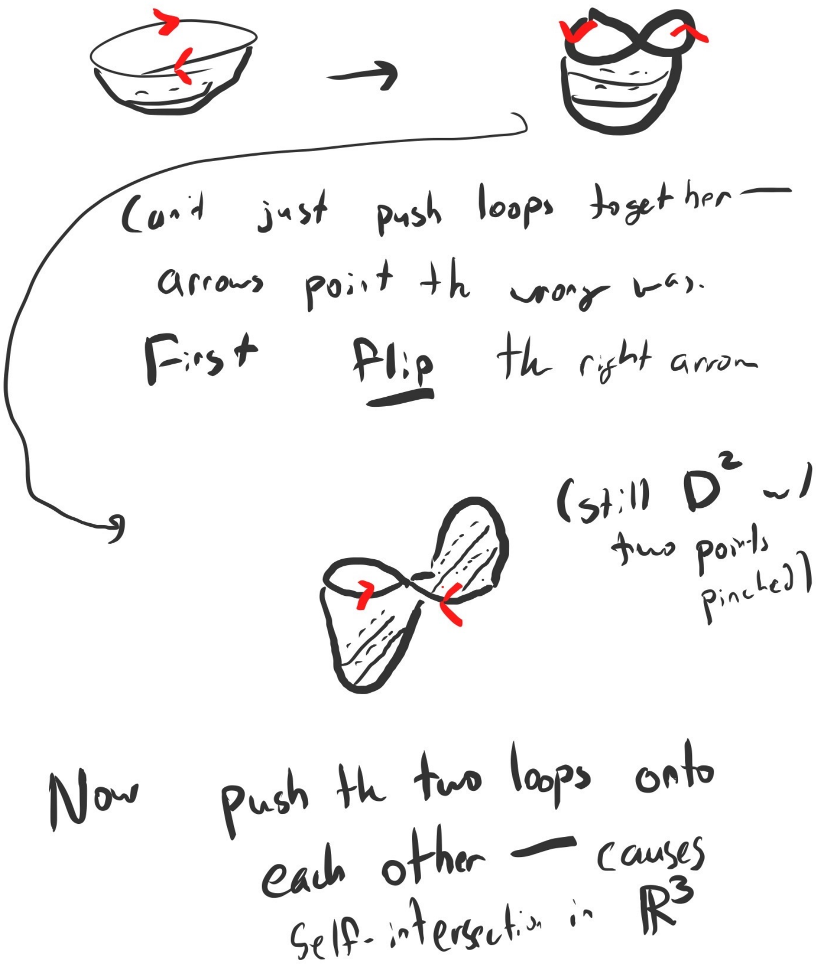
$$D^n / \sim \cong \mathbb{RP}^n,$$

as desired. $\square$

This gives us more room to actually visualize what's going on here for $\mathbb{RP}^2$. Here are two different pictures of the same identifications giving rise to $\mathbb{RP}^2$ as a quotient of the 2-dimensional disc. I like to think of this picture as meaning that if I'm a little ant walking around on the disc, if I walk through one side, I magically reappear at the opposite side, marching ever forward.



The following sequence of pictures is an attempt to show what's going on here. Note the self-intersections that arise while trying to describe this in $\mathbb{R}^3$; in fact, there is no topological embedding $\mathbb{RP}^2 \hookrightarrow \mathbb{R}^3$ (though this is hard to prove).



10/26: Countability axioms

Earlier in the course we discussed *Hausdorff spaces*, and mentioned that there are a zoo of related 'separation axioms': $T_0, T_1, T_2, T_3, \cdots$ of various strength levels; we only focused on Hausdorff ($= T_2$), as this condition appears far more often in practice than the others.

There are also a long list of conditions about the *size* of a topological space, measured in many different ways.

We've previously seen the notion of separability: a space is separable if there is a countable dense subset. One may generalize this notion and say that if X is a space, its *density* is the least possible cardinality of a dense subset; a separable space is a space with $d(X) \leqslant |\mathbb{N}|$.

Density measures how many points you need to be 'near' everything in the space. Having small density means everything is packed together, while having large density means everything is spread out. The indiscrete topology has $d(X) = 1$ and the discrete topology has $d(X) = |X|$.

Today's notions, on the other hand, measure how locally complex the topology is by measuring how many open sets you need to generate it, and the most extreme examples are neither discrete nor indiscrete, as neither of these topologies is terribly complicated to describe.

Second countability as 'size condition'

Confusingly, 'second countability' is the condition one sees more often in practice. Because it is more important, and its definition is easier, I will start with that first.

Definition 49. *A second-countable space is a topological space X so that there* ***exists*** *a basis $\mathcal{B}$ of countably many open sets generating the topology on X.*

More generally, we say the ***weight*** *of a topological space $w(X)$ is the least possible cardinality of a basis for X; then X is second countable iff $w(X) \leqslant |\mathbb{N}|$.*

Let me start by giving some examples and counterexamples. First I'll give some examples of a more set-theoretic nature, and then I'll talk about examples from analysis.

Example 57. Every indiscrete space is second-countable because one may take $\mathcal{B} = \{X\}$: we're generated by a basis of ONE open set! Thus $w(X_{indisc}) = 1$.

On the other hand, a discrete space X is second-countable if and only if X consists of at most countably many points. More precisely, $w(X_{disc}) = |X|$.

To see this, notice that $\mathcal{B}_{singleton} = \big\{\{x\}\big\}_{x \in X}$ is a basis for X, which has the same cardinality as X. This implies that $w(X_{disc}) \leqslant |X|$; there exists a basis of cardinality $|X|$.

I claim that **every** basis $\mathcal{B}$ generating the discrete topology contains $\mathcal{B}_{singleton}$. This would imply that $|\mathcal{B}| \geqslant |X|$ for all bases $\mathcal{B}$, and thus that $w(X_{disc}) \geqslant |X|$, proving the desired equality.

To see this, observe that $\{x\}$ is open in the discrete topology. We may thus write it as a union of basic open sets in $\mathcal{B}$; but the only nonempty subset contained in $\{x\}$ is $\{x\}$ itself. So $\{x\}$ must be in $\mathcal{B}$, and this logic applies for arbitrary $x \in X$. Thus if $\mathcal{B}$ is a basis for the discrete topology on X, we have $\mathcal{B}_{singleton} \subset \mathcal{B}$.

Proposition 84. *Let X be a metric space (with the induced topology). Then X is second-countable if and only if X is separable.*

Proof. Suppose X is separable, and suppose $\{x_1, x_2, \cdots\}$ is the countable dense subset. I claim

$$\mathcal{B} = \{B_{1/n}(x_m) \mid m, n \in \mathbb{N}\}$$

is a basis for the topology on X; certainly this is a countable collection of open sets.

To see that this forms a basis, let U be an arbitrary open subset: we will show that it is $\mathcal{B}$-open. (The converse, that $\mathcal{B}$-open sets are open in the topology on X, is immediate from the fact that all sets in the basis $\mathcal{B}$ are open in the given topology on X.)

Pick $x \in U$. By the definition of the metric topology, there exists some $r > 0$ so that $B_r(x) \subset U$. Further, because the set above is dense, it must intersect the nonempty open set $B_t(x)$ non-trivially for any $t > 0$. In particular, there exists some x_m in our countable dense set so that $d(x_m, x) < r/2$.

Now pick n so that $1/n < r/2$. Then we have

$$B_{1/n}(x_m) \subset B_{r/2}(x_m) \subset B_r(x) \subset U,$$

the middle inclusion coming from the triangle inequality: because $d(x_m, y) < r/2$ and $d(x, x_m) < r/2$ implies $d(x, y) < r$. Therefore U is $\mathcal{B}$-open, and this indeed is a countable basis for the topology on X.

On the other hand, if X is second-countable with countable basis $\{U_1, U_2, \cdots\}$, pick an element $x_i \in U_i$ for all i (at least, for all i where that basic open set is nonempty). I claim that $\{x_i\}$ is a countable dense subset. To show that it is dense, it is equivalent to show that for any nonempty open set U, there exists some $x_i \in U$. (Do you see why? Refer back to the open-set characterization of closure.)

To see why my claim holds, observe that because $\{U_i\}$ is a basis and U is a nonempty open set (with say $x \in U$), there exists some i so that $x \in U_i \subset U$. Because $x_i \in U_i$, we see that $x_i \in U$, as desired; the set $\{x_i\}$ is indeed a countable dense set. □

Second-countability and separability are not at all equivalent in general; see the next subsubsection for a counterexample.

Most of our standard constructions preserve second-countability.

Proposition 85. *If X and Y are second-countable, then $X \times Y$ is second-countable, as is $X \sqcup Y$. If X is second-countable and $S \subset X$ is a subspace, then S is second-countable.*

Countable products of second-countable spaces are usually not second-countable; countable disjoint unions always are. Second-countability does not behave well with respect to arbitrary quotient space constructions.

Proof of Proposition 85. If X has countable basis $\mathcal{B}_X$ and Y has countable basis $\mathcal{B}_Y$, then $X \times Y$ has countable basis $\{U \times V \mid U \in \mathcal{B}_X, \ Y \in \mathcal{B}_Y\}$, while $X \sqcup Y$ has countable basis

$$\mathcal{B} = \{U \sqcup \varnothing \mid U \in \mathcal{B}_X\} \cup \{\varnothing \sqcup V \mid V \in \mathcal{B}_Y\}.$$

Lastly, S has countable basis $\mathcal{B}_S = \{U \cap S \mid U \in \mathcal{B}_X\}$. □

In practice, second-countability is used as a tameness assumption, to guarantee that the spaces we study are well-behaved. The most common usage is in the following deep result, sometimes called 'the first theorem of point-set topology'; Curio 2 guides you through a proof, and points out corollaries.

Theorem 86 (Urysohn's theorem)**.** *Let X be a compact Hausdorff space. Then X is metrizable if and only if X is second-countable.*

In fact, Urysohn's theorem holds in somewhat more generality than this (we may assume that X is a 'regular space', meaning that singleton sets are closed and any two closed sets may be separated by open sets; this is true when X is compact Hausdorff). The second-countability condition is mainly used to demand that the spaces we construct are 'well-behaved' or 'small'. For instance, manifold theorists like to know that their manifolds are metrizable, even if they aren't interested in working with a particular distance function. It just lets them know that all of their metric-space intuition will work without change.

A peculiar non-example

Below I will give an example of a non-second countable space and a proof that it is not second countable. In fact, this space is not even first-countable. The language below uses some ideas from the theory of (partially) ordered sets, but I think it's useful to parse through what it's saying; some of this logic can be applied in more concrete settings. Allow me to give that language first so it doesn't appear in the proof itself.

Definition 50. *A partially ordered set $(S, \leqslant)$ is a set euipped with a relation which is reflexive ($x \leqslant x$), transitive ($x \leqslant y$ and $y \leqslant z$ implies $x \leqslant z$), and 'antisymmetric' (if $x \leqslant y$ and $y \leqslant x$ then in fact $x = y$).*

We do ***not*** *assume that for any two elements $x, y \in S$, either $x \leqslant y$ or $y \leqslant x$; this would be called a totally ordered set.*

For instance, $\mathbb{N}^2$ is a partially ordered set, with the relation $(m, n) \leqslant (m', n')$ if $m \leqslant m'$ and $n \leqslant n'$. This means that $(0, 0) < (1, 1)$, but that $(0, 1)$ and $(1, 0)$ are incomparable: neither is bigger than the other.

Definition 51. *Let $(S, \leqslant)$ be a partially ordered set. A* ***cofinal subset*** *is a subset $C \subset S$ so that, for all $x \in S$, there is* some *$c \in C$ with $x < c$.*

That is, 'everything is less than something in C'; you can think of this as saying that C 'goes off to infinity in S'. A good example is $C = \mathbb{N} \subset \mathbb{R}$. For every real number, there is some natural number bigger than it.

In the example of $\mathbb{N}^2$ with the order given above, the set $\Delta_{\mathbb{N}} = \{(n, n) \mid n \in \mathbb{N}\}$ is cofinal. If $(m, n) \in \mathbb{N}^2$, then this is less than the element $(\max(m, n), \max(m, n))$ on the diagonal. The same argument works to show that the 'diagonal' in $\mathbb{N}^n$, with the product order

$$(x_1, \cdots, x_n) \leqslant (y_1, \cdots, y_n) \iff \forall_i x_i \leqslant y_i$$

is a cofinal subset.

For a more complicated version of the same example, consider the set of functions $\mathbb{N} \to \mathbb{N}$; we write this set as $\mathbb{N}^{\mathbb{N}}$. This is a partially ordered set, too, with $f \leqslant g$ if and only if $f(n) \leqslant g(n)$ for all natural numbers

n. However, unlike the previous examples, **there is no countable cofinal subset of** $\mathbb{N}^{\mathbb{N}}$. To see this, suppose we have a countable set $\{f_1, \cdots,\}$ of functions $f_i : \mathbb{N} \to \mathbb{N}$. Set

$$g(n) = f_n(n) + 1$$

for all n. Then g is not less than f_n, for any n, so this is not a cofinal set. (This is a variation on Cantor's diagonalization argument.)

We are now ready to see the example of a non-second countable space.

Example 58. Despite what your intuition might say, even if the space X is countable (that is, it has countably many points), there need not be a countable basis for the topology. The simplest counterexample I can think of is the following topological space.

Set $X = \mathbb{N}^2 \cup \{\infty\}$. We say $U \subset X$ is open if either

- $\infty \notin U$ (any subset not containing ∞ is open), or
- $\infty \in U$ and, for each $y \in \mathbb{N}$, there exists an $n(y) \in \mathbb{N}$ so that $(x, y) \in U$ whenever $x \geqslant n(y)$.

Try visualizing the second condition on an infinite grid: it says that on each horizontal line, we contain everything sufficiently far to the right (where 'sufficiently far' depends on the line). This space X should be thought of as the having countably many convergent sequences (the horizontal lines $\mathbb{N} \times \{y\}$) all converging to the same point $\{\infty\}$, but not converging uniformly.

In fact, if we write $\mathbb{N}_+$ for the one-point compactification of $\mathbb{N}$ (which we think of as being a space corresponding to one convergent sequence), the above space is given by taking the disjoint union of countably many copies of $\mathbb{N}_+$, and gluing together all of the limit-points; this is called the wedge product of the countable family of spaces $\mathbb{N}_+$. It is **not** homeomorphic to any subspace of $\mathbb{R}^n$, as such subspaces are second-countable.

I claim there is no countable collection of open sets which generates this topology. The idea is that each open set contains 'too much information': it should know the function $\mathbb{N} \to \mathbb{N}$ in the second bullet point, which we called $n(y)$. The set of functions $\mathbb{N}^{\mathbb{N}}$ has the cardinality of the reals; it seems far too big.

Precisely, we saw before this example that $\mathbb{N}^{\mathbb{N}}$ is a partially ordered set which contains no countable cofinal subset.

How do we use this?

Observe that there is a function $E : \mathcal{T}_\infty \to \mathbb{N}^{\mathbb{N}}$, described as follows.

Here the domain is the collection of open subsets of X which contain ∞, and the codomain is the set of functions $\mathbb{N} \to \mathbb{N}$. The E stands for 'end', as in, 'the end of each sequence lies in U, and we record when that end starts'). We write

$$E(U)(y) = \min\{n \mid \forall x \geqslant n,\ (x, y) \in U\}.$$

That is, $E(U)(y)$ records the point after which the horizontal strip $\mathbb{N} \times \{y\}$ lies entirely in U. Notice that this is order-preserving: if $U \subset V$ then $E(U) \geqslant E(V)$. (The smaller U is, the 'later' the ends of the strips start to lie in U.) It follows that for all i we have $E(\bigcup_{i \in I} U_i) \leqslant E(U_i)$.

Notice that E is surjective; if $f : \mathbb{N} \to \mathbb{N}$ is a function and $U = \{(x, y) \mid x \geqslant f(y)\} \cup \{\infty\}$, then U is an open set with $E(U) = f$.

Together, this implies that if $\mathcal{B}$ is a basis, then

$$\{E(U_i) \mid U_i \in \mathcal{B}\}$$

is a cofinal subset of $\mathbb{N}^{\mathbb{N}}$. To see this, observe that every open set may be written as a union of the various $U_i \in \mathcal{B}$; pick an arbitrary $f \in \mathbb{N}^{\mathbb{N}}$, and construct a U with $E(U) = f$. Then $U = \bigcup_{i \in J \subset I} U_i$, and $f = E(U) \leqslant E(U_i)$ for $i \in J$. In particular, f is less than at least one of the terms in the above set; so the above set is cofinal.

It follows that there is no countable basis for the topology on X. If there was, then $E(U_1), E(U_2), \cdots$ would be a countable cofinal subset of $\mathbb{N}^{\mathbb{N}}$, and the lemma below shows that no such countable collection exists.

While the proof above will seem non-intuitive at first, I encourage you to compare some of the ideas to what goes on in the metric space setting up above, where the idea of cofinality is secretly used: to get our countable collection of small open neighborhoods of a given point, we use that $1/n$ is cofinal in the ordered set $(\mathbb{R}_{>0}, <)$, and that this is a countable cofinal set.

First-countabillity and sequential continuity

The notion of first-countable is slightly more subtle. First, I will give the definition, as well as some examples and non-examples. Then I'll say why this is an interesting condition. It's related to the idea of sequential convergence, so you may need to review the brief section on sequences from the 10/5 lecture notes.

Definition 52. *Let X be a topological space and $x \in X$ a point. We say* ***a neighborhood basis of*** *x is a collection $\mathcal{B}_x = \{U_i\}_{i \in I}$ of open subsets of X with $x \in U_i$ for all i, so that for any open subset $x \in U \subset X$ containing the given point, there exists some $U_i \in \mathcal{B}_x$ with $x \in U_i \subset U$.*

We say that a topological space X is first-countable if, for all $x \in X$, there exists a countable neighborhood basis of x.

This is more concrete than it sounds. The following proposition shows that you've seen this idea before.

Proposition 87. *If X is a metric space, then X is first-countable.*

Proof. Pick $x \in X$. Set the neighborhood basis to be $\mathcal{B}_x = \{B_{1/n}(x) \mid n \in \mathbb{N}\}$. This is clearly a countable collection of open sets containing x. To see that this is a neighborhood basis at x, observe that if $x \in U$ is open, then $x \in B_r(x) \subset U$ by definition of the metric topology, for some $r > 0$. Pick $n \in \mathbb{N}$ with $1/n < r$. (There's cofinality again.) Then

$$x \in B_{1/n}(x) \subset B_r(x) \subset U,$$

as desired. □

Finite products of first-countable spaces are first-countable; arbitrary disjoint-unions of first-countable spaces are first-countable; subspaces of first-countable spaces are first-countable. Quotients need not be.

You should convince yourself that the argument in the previous subsubsection actually shows that the space constructed there fails to be first-countable.

Whereas second-countability is a sort of tameness condition, the weaker notion of first-countability asks that you be 'tame near each point' but not necessarily overall. It also is very closely related to the notion of sequential continuity. First, recall the notion of convergent sequence $x_n \to x$ in a topological space X: this sequence converges to x if, for every open set U, we have $x_n \in U$ for all sufficiently large n.

To analyze why first-countability is related to this notion, we need to prove a small lemma.

Lemma 88 (Nested neighborhood bases recover the metric-space notion of convergence)**.** *Let X be a first-countable space and let $(U_1, U_2, \cdots)$ be a neighborhood basis of x as in the previous lemma.*

Then a sequence $x_n \in X$ converges to $x \in X$ if and only if, for each $m \in \mathbb{N}$, there exists an $N(m)$ so that $x_n \in U_m$ for all $n > N(m)$.

Notice that if you replace U_m with $B_{1/m}(x)$, this is exactly the definition of convergence in a metric space.

Proof. Suppose $x_n \to x$. Because each U_m is open, the definition of convergence implies that $x_n \in U_m$ for all sufficiently large n. It is worth being precise: there exists some $N = N(m)$ depending on m so that $x_n \in U_m$ for all $n > N(m)$. So this proves one direction of the claim.

For the other direction, suppose that $x_n \in U_m$ for all m and all $n > N(m)$. Let's show $x_n \to x$. To see this, pick an arbitrary open set $x \in U \subset X$; we aim to show that $x_n \in U$ for all sufficiently large n.

Because U_m is a neighborhood basis of x, we know that there exists some m with $x \in U_m \subset U$. Further, we know that $x_n \in U_m$ for all $n > N(m)$. It follows that $x_n \in U$ for all $n > N(m)$, as desired. □

For technical reasons in a few minutes, it will be useful to slightly restrict the kind of neighborhood bases we work with, so they behave more like $B_{1/n}(x)$.

Lemma 89 (Nested neighborhood bases exist in first-countable spaaces)**.** *If X is a first-countable space and $x \in X$, there is a neighborhood basis of x given by open sets $U_1, U_2, \cdots$ with $U_{n+1} \subset U_n$ for all n. This is called a nested neighborhood basis.*

Proof. Let $U_1, U_2, \cdots$ be an arbitrary neighborhood basis of x, and replace it with $V_1 = U_1, V_2 = U_1 \cap U_2$, and in general $V_n = V_{n-1} \cap U_n = U_1 \cap \cdots \cap U_n$. Notice that these are still open sets (as they are finite intersections of open sets), and they still contain x (as $x \in U_i$ for all i). To see that this is indeed a neighborhood basis, let $x \in U \subset X$ be arbitrary; we know by hypothesis that $x \in U_n \subset U$ for some n. Because $V_n \subset U_n$, we see that in fact $x \in V_n \subset U$, as desired. ☐

Our goal will be to use this to show that sequences know everything there is to know about the topology of a first-countable space. We make this precise with the following definition.

Definition 53. *Let X be a topological space. We say $U \subset X$ is sequentially open if, for every convergent sequence $x_n \to x$ where $x \in U$, we have $x_n \in U$ for all sufficiently large n.*

If Y is another topological space we say $f : X \to Y$ is sequentially continuous if, for every convergent sequence $x_n \to x$, the sequence $f(x_n)$ converges to $f(x)$.

We are going to use the previous lemma to show that for first-countable spaces, the 'sequential' concepts coincide with the usual ones. The arguments are very much like arguments you may have seen in analysis. Instead of using the neighborhood basis $B_{1/n}(x)$, we use the neighborhood basis U_n from the previous lemma.

Theorem 90 (For first-countable spaces, continuity is sequential continuity)**.** *If X is a first-countable space, then $U \subset X$ is open if and only if it is sequentially open. Further, for X first-countable and Y any space, $f : X \to Y$ is continuous if and only if f is sequentially continuous.*

Proof. If U is open, it is sequentially open (immediately from the definition of sequential convergence).

The hard part is the other direction. Let U be a sequentially open set. We will show that U is locally open, by showing that for all $x \in U$, there exists an open set V with $x \in V \subset U$.

Pick a countable nested neighborhood basis $U_1 \supset U_2 \supset \cdots$ of $x \in X$. I claim that for n sufficiently large, we have $U_n \subset U$. **Towards a contradiction, assume this is not true.** That means that for each n, we may choose $x_n \in U_n$ with $x_n \notin U$.

Because our neighborhood basis is *nested*, this implies that $x_n \in U_m$ for all $m \geqslant n$. By our lemma on sequential convergence, we see that x_n converges to x.

Because $x \in U$ but $x_n \notin U$ for any sufficiently large n, this contradicts the fact that U is sequentially open. We have thus proved that $U_n \subset U$ for sufficiently large n. Because $x \in U$ was arbitrary, this implies U is sequentially open.

Now let's analyze sequential continuity. **Verify** that a map $f : X \to Y$ is sequentially continuous precisely when, for all $U \subset Y$ open, the inverse image $f^{-1}(U)$ is **sequentially open**. (This is just a rephrasing of the definition.)

If X is first-countable, we have shown that sequentially open sets are open, and vice versa. It thus follows that a map out of a first-countable space is sequentially continuous if and only if it is continuous. ☐

WHERE DID I USE THAT THIS IS A NESTED BASIS? This tells us that for a wide class of spaces, we can still use the sequential thinking we got used to when working with metric spaces. If you found this unsatisfying and you'd like a suitable notion of sequential convergence for **all** spaces (as the functionalanalysts do), you have to learn about nets. This is out of the scope of this class but could make a good project!

11/4: Algebraic topology: Motivation

A basic question in any field of study is: *how do I show that two (blahs) are equivalent?* In algebra, you often try to study whether two groups are isomorphic or not; you may have done an exercise enumerating the isomorphism types of groups of cardinality pq, where p and q are prime. In some sense, the study of finite simple groups was launched to study this question, as you'd like to enumerate isomorphism types of finite groupsby studying how these simple pieces patch together.

So this leaves us with a basic question in topology: *if I have two spaces X and Y, how do I tell whether or not they're homeomorphic?*

If I want to prove that X and Y *are* homeomorphic, I know what I have to do: I have to construct a homeomorphism. An example where the process to get to the homeomorphism was long and abstract was given in Curio 1, but in the end *we still construct a map, and checked that it was a homeomorphism.* I don't know a single instance of two spaces known to be homeomorphic where the proof, ultimately, does not construct a homeomorphism (compare: the Ax-Grothendieck theorem).

If we want to show that two spaces are *not* homeomorphic, the only real approach we have is to find some property that is preserved under homeomorphism (a so-called topological property), which one space has and the other doesn't. So for instance...

- $D^n \not\cong \mathbb{R}^n$, because the former is compact and the latter is not;
- The Cantor set C is not homeomorphic to the unit interval, because the former is disconnected and the latter is connected;
- The topologist's sine curve S is not homeomorphic to the unit interval, because the latter is path-connected but the former is not;
- $\mathbb{R}$ with the standard topology is not homeomorphic to $\mathbb{R}_{cc}$, itself equipped with the countable-complement topology, because the former is Hausdorff and the latter is not;
- $\mathbb{R}$ is not homeomorphic to $\mathbb{R}^n$ for $n > 1$, because $\mathbb{R}\backslash\{0\}$ is not homeomorphic to $\mathbb{R}^n\backslash\{p\}$ for any p, and that follows because the former is disconnected while the latter is connected.
- The unit interval is not homeomorphic to the circle, because the former has cut-points while the latter does not.

But eventually we run into walls. There are only so many of these properties we can come up with! Already I don't know how to, say, show that S^2 is not homeomorphic to $S^1 \times S^1$; these are compact path-connected Hausdorff spaces, both have a countable basis, neither of them have cut-points...

We also don't yet have the tools to distinguish $\mathbb{R}^2$ from $\mathbb{R}^n$ for $n > 2$ (though there actually is a point-set topology way to do this, under the name *dimension theory*). However, what happens when I try to mimic the above argument?

I look at $\mathbb{R}^2\backslash\{0\}$. In some sense, what we did with $\mathbb{R}\backslash\{0\}$ is show that I cannot 'deform' a point in $(-\infty, 0)$ to a point in $(0, \infty)$, thinking of a path as tracing out how that point moves over time.

When I look at $\mathbb{R}^2\backslash\{0\}$, I see that there seems to be a 'loop' wrapping around the origin I cannot seem to deform into a loop which does not wrap around the origin. This justifies the investigation of this idea of *deformation.* We will obtain a new, more flexible tool, which we can use to distinguish topological spaces.

Homotopies and homotopy equivalences

Definition 54. *Let X and Y be topological spaces. We say two continuous maps $f, g : X \to Y$ are homotopic, written $f \sim g$, if there is a continuous map $F : X \times [0, 1] \to Y$ so that $f(x) = F(x, 0)$ and $g(x) = F(x, 1)$. Oftentimes, we write F_t for the map $X \to X$ given by $F_t(x) = F(t, x)$; in that language, the assumption is that $F_0 = f$ and $F_1 = g$.*

The idea to me is that the F_t is a continuously-varying family of functions from X to Y (the idea that this is continuously-varying is encoded by the assumption that $F : X \times [0,1] \to Y$ is continuous as a function on $X \times [0,1]$, and **not** just that each F_t is separately continuous); watching what happens as t goes from 0 to 1 is like watching a movie of the map f deforming, slowly, into the map g.

The relation of being homotopic is, in fact, an equivalence relation, as you will prove on your homework.

Example 59. Let $\mathrm{rot}_\theta : S^1 \to S^1$ be the map that rotates a circle by a given angle $\theta \in \mathbb{R}$. Then rot_θ is homotopic to rot_ψ for any other angle ψ; the homotopy is given by $F_t = \mathrm{rot}_{(1-t)\theta+t\psi}$.

To check that this is continuous, I'll change the language a little bit. The map rot_θ is given by *complex multiplication*: $\mathrm{rot}_\theta(z) = e^{i\theta}z$. Then our homotopy is

$$F(t,z) = e^{i[(1-t)\theta+t\psi]}z.$$

Because addition, scalar multiplication, exponentiation, and complex multiplication are all continuous, F is continuous as a map $[0,1] \times S^1 \to S^1$.

The idea is that we just keep rotating until we've rotated by the extra $\psi - \theta$.

Example 60. Any two maps $f, g : X \to \mathbb{R}^n$ are homotopic. To see this, consider the map $F_t(x) = (1-t)f(x) + tg(x)$; this is called the *straight-line homotopy*, since I go from $f(x)$ to $g(x)$ by following the straight line between them.

Why is this continuous as a map from $[0,1] \times X$? F is the composite of the continuous maps

$$X \times [0,1] \xrightarrow{(f,g,1_{[0,1]})} \mathbb{R}^n \times \mathbb{R}^n \times [0,1] \xrightarrow{(x,y,t)\mapsto tx+(1-t)y} \mathbb{R}^n;$$

the first map is continuous because all of its components are continuous (pedantically, the first component is projection onto X and then applying f; the second is projection onto X then applying g; the last is projection onto $[0,1]$), and the second map is continuous because addition and scalar multiplication both are (you can write it as a composite of even more transparently continuous functions, if you like).

Example 61. Write $S^0 = \{1, -1\} \subset \mathbb{R}$ for the 0-sphere; this is two points with the discrete topology. I claim that the identity $1 : S^0 \to S^0$ and negation $m : S^0 \to S^0$ with $m(x) = -x$ are not homotopic. You will argue this on your homework.

What is more important, though, is to understand why the straight-line homotopy does not work. Why does $F_t(x) = (1-2t)x$ not define a homotopy between 1 and m? I **strongly** recommend understanding this as soon as possible, to minimize later confusion.

Two maps being *homotopic* intuitively means I can deform one of those maps into the other. I can use this to cook up a notion of *homotopy equivalence* of topological spaces, intuitively capturing the idea that I can deform one *space* into another.

Definition 55. *Let X and Y be topological spaces. We say that a continuous map $f : X \to Y$ is a homotopy equivalence if there is a continuous map $g : Y \to X$ (called a homotopy inverse) so that $fg : X \to X$ is homotopic to the identity 1_X, and $gf : Y \to Y$ is homotopic to the identity 1_Y.*

If there exists a homotopy equivalence from X to Y, we say that X and Y are homotopy equivalent, written $X \simeq Y$.

Why this definition? Well, recall this useful alternate definition of homeomorphism:

A continuous map $f : X \to Y$ is a homeomorphism if there exists a continuous map $g : Y \to X$ so that $gf = 1_X$ and $fg = 1_Y$. This is the same as demanding that f is a bijection with continuous inverse: the fact that g is both a left- and right- inverse implies that f is bijective and $g = f^{-1}$. 'You can go from X to Y and back continuously and end up where I started, and vice versa.'

Then a homotopy equivalence is like this, where we relax the demand that fg and gf are actually the identity. We say, ok, well I don't mind if I have to do a little deforming/wiggling to make these into the identity. *I can go from X to Y continuously and back again continuously, so that if I wiggle a bit, I'm back where I started.*

Notice that there may be a great many homotopy inverses in general. It follows from a result on your homework this week that if f is a homotopy equivalence with homotopy inverse g, then in fact any map homotopic to g is also a homotopy inverse to f.

I think this notion will be made clearer with examples.

Example 62. The space $S^1 \times D^2$ (the *solid torus*) is homotopy equivalent to S^1. The homotopy equivalence is $p : S^1 \times D^2 \to S^1$, given by $p(x, y) = x$; a homotopy inverse is given by $i(x) = (x, 0)$.

To prove these are homotopy inverse, we need to prove that $pi \sim 1_{S^1}$ and that $ip \sim 1_{S^1 \times D^2}$. The former condition holds for free, as $pi = 1_{S^1}$; no need to wiggle.

The second condition is not automatic. The map $ip : S^1 \times D^2 \to S^1 \times D^2$ is given by $(ip)(x, y) = (x, 0)$.

For the homotopy, I will take

$$F_t(x, y) = (x, ty).$$

Notice that $F_0 = ip$ and $F_1 = 1_{S^1 \times D^2}$.

The idea is that I am slowly shrinking the disc to 0, which shrinks the whole space $S^1 \times D^2$ to zero. This is continuous as a function

$$S^1 \times D^2 \times [0, 1] \to S^1 \times D^2$$

because each of its components are continuous.

There was something special about the above example — the latter space (S^1) lived as a subspace in the first, and we never moved it in the homotopy; it sat where it was. We will investigate this special case more soon. First, I'd like to give an example justifying how this is special.

Example 63. Let Q denote what it looks like: the union of a circle and a line segment which crosses both into and out of the circle. Let $\mathsf{6}$ denote what it looks like: the union of a circle with a curved line segment leaving the circle (but not going inside).

Then these two spaces are homotopy equivalent. The map $f : \mathsf{Q} \to \mathsf{6}$ collapses the line segment to a point, and identifies the circle on the first space with the circle in the second space. The map $g : \mathsf{6} \to \mathsf{Q}$ collapses the curved segment to a point, and identifies the circle on $\mathsf{6}$ with the circle on Q.

The composite $gf : \mathsf{Q} \to \mathsf{Q}$ is the identity on the circle, but sends the line interval to the point it intersects the circle. This is homotopic to the identity; take $F_t : \mathsf{Q} \to \mathsf{Q}$ to be a map which collapes the line interval to t times its original length, so that $F_0 = gf$ and $F_1 = 1_{\mathsf{Q}}$.

The discussion for $fg : \mathsf{6} \to \mathsf{6}$ is similar.

11/9: Deformation retractions and homotopies rel A

I want to pin down the idea of what happened in Example 62. Let me first give an auxiliary condition, which explains something about the map p.

Definition 56. *Let X be a topological space and $A \subset X$ a subspace. A **retraction** of X onto A is a continuous map $r : X \to X$ with $r(a) = a$ for all $a \in A$, and $r(X) \subset A$. It follows that $Im(r) = A$, and that $rr = r$; in fact, r being a retraction onto A is precisely these two conditions.*

If A is the image of a retraction $r : X \to X$, then we say A is a retract of X.

The map $p : S^1 \times D^2 \to S^1 \times D^2$ given by $p(x, y) = (x, 0)$ is a retraction onto $S^1 \times \{0\}$. One tends to intuit that a retract of X is in some sense 'simpler' than X. We cannot prove this yet, but for instance, the unit circle S^1 is not a retract of the plane $\mathbb{R}^2$.

In Example 62, the time-1 map of our homotopy was a retraction onto A, with a handful of special properties. We record the properties we saw in that example as a definition.

Definition 57. *Let X be a topological space with $A \subset X$ a subspace. We say that a map $F : X \times [0, 1] \to X$ is a **deformation retraction** of X onto A if:*

- $F_0 = 1_X$ *(that is, $F(0, x) = x$),*
- $F_t|_A = 1_A$ *(that is, if $a \in A$, then $F(t, a) = a$ for all t),*
- $F_1(X) \subset A$ *(that is, $F_1(x) \in A$ for all $x \in X$).*

If there exists a deformation retraction onto A, we say that A is a deformation retract of X.

The idea here is that we are slowly collapsing X onto its subspace A. Note that the map F_1 is a retraction: $F_1F_1 = F_1$, because $F_1(x) \in A$, and we know that $F_1|_A = 1_A$, so that $F_1(F_1(x)) = F_1(x)$, which is why this is called a *deformation retraction*: we have a homotopy from the identity to a retraction onto A, which never destroys the property that A is unmoved by F_t.

We saw earlier that $S^1 \times \{0\}$ is a deformation retract of $S^1 \times D^2$. Here are two more examples.

Example 64. Let $X = \mathbb{R}^3 \backslash \{(0,0,z) \mid z \in \mathbb{R}\}$; that is, the complement of the z-axis. Let $A = S^1 \times \{0\} \subset \mathbb{R}^3$. Then X deformation retracts onto A.

For a formula, I will write $\mathbb{R}^3$ as $\mathbb{C} \times \mathbb{R}$ and its points as (w, z). Take

$$F_t(w,z) = \left(\frac{w}{t|w| + 1 - t}, (1-t)z\right).$$

The idea is that we are rescaling the complex coordinate to lie on the circle, while crushing the real coordinate until it is zero.

Now note that $F_0(w,z) = (w,z)$, so that $F_0 = 1_X$; note that if $w \in S^1$, we have $F_t(w,0) = (x,y,0)$, so that $F_t|_A$ is the identity on A; and we have

$$F_1(w,z) = (w/|w|, 0) \in S^1 \times \{0\}.$$

Thus $S^1 \times \{0\}$ is a deformation retract of $\mathbb{R}^3$ minus the z-axis.

Example 65. Let $X = \mathbb{R}^3 \backslash \left(S^1 \times \{0\} \cup \{(0,0)\} \times \mathbb{R}\right)$, the complement of both the unit circle in the xy-plane and the z-axis. I claim that X deformation retracts onto

$$\{(x,y,z) \mid (1 - \sqrt{x^2+y^2})^2 + z^2 = 1/4\} \cong S^1 \times S^1,$$

a 2-torus in 3D space. I won't write down the formula, because it's awful, but can you see what's going on?

In fact, this is a great place to apply our homeomorphism $(\mathbb{R}^3)_\infty \cong S^3$ from the one-point compactification of $\mathbb{R}^3$ to the 3-sphere S^3. This homeomorphism sends the subset we've taken the complement of to a pair of linked circles in the 3-sphere, called the *Hopf link* —

$$S^1 \times \{0\} \cup \{0\} \times S^1 \subset S^3 \subset \mathbb{C}^2.$$

The deformation retraction of this onto the 2-torus

$$\{(z,w) \mid |z| = |w| = 1/\sqrt{2}\} \subset S^3,$$

which is called a *Clifford torus*. Here the deformation retraction of of the complement of the linked circles onto the 2-torus is easier to write down (though perhaps harder to see!)

The reason we're talking about deformation retractions is because of the following result.

Proposition 91. *If $A \subset X$ is a deformation retract of X, then the inclusion $i : A \to X$ is a homotopy equivalence.*

Proof. If $F : X \times [0,1] \to X$ is the deformation retraction, I claim the necessary homotopy inverse is $F_1 : X \to A$ (thought of as a map with codomain A with the subspace topology).

By the definition of deformation retraction, we have $F_1 i = 1_A$. What we need to show is that $iF_1 : X \to X$ is homotopic to the identity map. But F itself provides such a homotopy.

Notice that in this argument we never used the property that $F_t|_A = 1_A$ for all t; only that $F_1(A) \subset A$, that $F_1|_A = 1_A$, and that $F_0 = 1_X$. The result is true under weaker conditions; the idea of a deformation retraction is a nice geometric picture, so we tend to talk about those instead of these weaker conditions. □

Note that both Q and 6 deformation retract onto the circle. You will prove in your homework that homotopy equivalence is an equivalence relation, and thus this gives a proof that Q and 6 are homotopy equivalent. This is basically the proof written above in different language.

Contractible spaces

In Example 2, I pointed out that *any two maps* to $\mathbb{R}^n$ are homotopic. I did this by using a sort of *canonical* homotopy on $\mathbb{R}^n$, the straight-line homotopy. Another approach would be to show that every map to $\mathbb{R}^n$ is homotopic to the constant map c_0; then because homotopy is an equivalence relation, it follows that any two maps are homotopic.

This idea informs the following two definitions.

Definition 58. *We say that a map $f : X \to Y$ is null-homotopic if, for some $y \in Y$, f is homotopic to the constant map $c_y : X \to Y$ (with $c_y(x) = y$ for all x).*

We think of constant maps as being particularly uninteresting maps, which don't involve the topology on X at all; if we can deform a map to a constant map, then (from the perspective of maps up-to-homotopy) it wasn't super interesting to begin with, whence the name.

(Remark: two constant maps are homotopic if and only if they lie in the same path-component. Can you prove this?)

Definition 59. *We say that a topological space is contractible if the identity map $1_X : X \to X$ is null-homotopic.*

This captures the idea of the straight-line homotopy, to some degree: we can collapse the space down to a point, so *any maps to our space are homotopic.* In fact:

Proposition 92. *Let X be a topological space. If X is contractible, then any map $f : X \to Y$ is null-homotopic, as is any map $g : Z \to X$.*

Conversely, if for all spaces Y and all continuous maps $f : X \to Y$, f is null-homotopic, it follows that X is contractible. (Similarly for maps to X.)

Proof. For the forward direction, set $H : [0,1] \times X \to X$ to be the contracting homotopy (a continuous map with $H_0 = 1_X$ and $H_1 = c_{x_0}$ the constant map $c_{x_0}(x) = x_0$ for some fixed point $x_0 \in X$). Then $fH : [0,1] \times X \to Y$ is a continuous map with

$$(fH)_0 = f1_X = f$$

and

$$(fH)_1 = fc_{x_0} = c_{f(x_0)},$$

so f is null-homotopic; similarly if $(1 \times g) : [0,1] \times Z \to [0,1] \times X$ is given by $(1 \times g)(t, z) = (t, g(z))$, then $H(1 \times g)$ gives a homotopy from g to c_{x_0}.

For the reverse direction, take $Y = X$ and $f = 1_X$ to see that 1_X is null-homotopic. □

In particular, it follows by taking $Z = \{*\}$, the one-point space, that...

Proposition 93. *If X is contractible, then X is path-connected.*

This is a very weak implication! Path-connected spaces are common; contractible spaces are rare.

Proposition 94. *A topological space X is contractible if and only if it is homotopy equivalent to the one-point space $\{*\}$.*

Proof. Suppose X is contractible; so there is a homotopy F from 1_X to c_{x_0}, for some $x_0 \in X$.

I claim that $i : \{*\} \to X$, given by $i(*) = x_0$, is a homotopy equivalence, with homotopy inverse the unique map $p : X \to \{*\}$. For $pi = 1_*$ rather tautologically (the one-point space only has one map to itself, the identity map!), and $ip = c_{x_0}$ is homotopic to the identity by assumption.

Conversely, suppose X is homotopy equivalent to the 1-point space. The homotopy equivalence must be that same map $p : X \to \{*\}$, because there is only one map to the one-point space. The inclusion $i : \{*\} \to X$ picks out a point $x_0 = i(*)$. The assumption that these two maps are homotopy inverse means that $pi \sim 1_*$ (automatic; they are equal) and $c_{x_0} = ip \sim 1_X$, thus furnishing a homotopy from ip to the identity. □

Remark 66. At this point we can prove that not all retractions $f : X \to X$ are the time-1 map F_1 of a deformation retraction of X onto $f(X)$. Write S for the topologist's sine curve; the map $p : S \to \{0\} \times [-1, 1]$, given by $p(x, y) = y$, is a retraction, but S does not deformation retract onto the interval. If it did, S would be homotopy equivalent to the interval, hence contractible; but S is not path-connected.

This is all well and good, but how about some examples?

Example 67. $\mathbb{R}^n$ is contractible: the map $F_t(x) = (1-t)x$ is a homotopy from the identity to the constant map at 0. So is the disc D^n. More generally, a *star-shaped set* is a set $A \subset \mathbb{R}^n$ with a chosen point $x_0 \in X$ so that, for $x \in X$, the entire straight line from x_0 to x is also contained in X — that is,

$$x \in X \implies \{(1-t)x + tx_0 \mid t \in [0, 1]\} \subset X.$$

Then any star-shaped set is contractible: the map $F_t(x) = (1-t)x + tx_0$ is a null-homotopy of the identity map.

You should draw a picture of a star-shaped set which is not convex to see why they deserve this name.

Not every contractible subset of $\mathbb{R}^n$ is star-shaped: for instance, the letter N is contractible (first collapse the left and right vertical lines, then collapse the diagonal to a point), but it is clearly not star-shaped.

Remark 68. The straight-line homotopy does **not** imply that any subspace $S \subset \mathbb{R}^n$ is contractible. **Why not?**

We do not currently have the technology to *prove* that any given space is non-contractible; it will take us some time to build the necessary tools.

Let me conclude by pointing out that the relationship between *homotopy equivalence* and *homeomorphism* is often very subtle, even when you make some assumptions that a space is 'well-behaved'.

Conjecture 1 (Poincare conjecture)**.** *Let X be a compact Hausdorff space so that for each point $p \in X$, there is an open set $p \in U \subset X$ which is homeomorphic to $\mathbb{R}^3$. (X is said to be a compact 3-manifold; manifolds are the objects most topologists in the 21st century study.)*

If X is homotopy equivalent to S^3, then X is homeomorphic to S^3.

This question remained a conjecture for about a century, until it was proven by Perelman in the early 2000s. His proof uses techniques from PDE (!) which are something like studying heat flow.

Homotopy relative to a subspace

Our first goal is to take the idea of *deformation retraction* and put it into an appropriate context. This uses the notion of homotopy relative to a subspace, which we will then proceed to use to better understand paths and loops in a given space.

Definition 60. *Let $f, g : X \to Y$ be continuous maps, and let $A \subset X$ be a subspace. Suppose that $f|_A = g|_A$.*

We say that f and g are homotopic relative to A*, or simply* homotopic rel A*, if there is a homotopy $F : X \times [0, 1] \to Y$ so that $F_0 = f, F_1 = g$, and $(F_t)|_A = f|_A$ — that is, $F_t|_A$ is the same for all time t.*

Here is a result from smooth manifold theory / analysis that makes use of this notion, to give you a sense of how one might want to appeal to it in practice.

Proposition 95 (Relative smooth approximation)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous map. Then f is homotopic to a smooth map by an arbitrarily small homotopy. If f is already smooth on a subspace $A \subset \mathbb{R}^n$, then f is homotopic rel A to a smooth map by an arbitrarily small homotopy.*

You don't need to understand anything in the statement here. The point is simple: A homotopy allows me to wiggle a map a bit ('small homotopy' means I'm only wiggling a tiny bit). A homotopy *relative to* A is one that does not adjust my map on a subspace I'm already happy with it on. In this example, it's already smooth on a subspace — why bother wiggling there?

We have already seen an example of homotopies relative to A.

Proposition 96. *Let $A \subset X$ be a subspace. A deformation retraction $F_t : X \times [0, 1] \to X$ onto A is the same as a homotopy* **rel** *A from the identity to a map $F_1 : X \to X$ with $F_1(X) \subset A$.*

Proof. Recall the definition of deformation retraction. It had three conditions.

The first was that $F_0 = 1_X$; in our phrasing here, this is the assumption that F is a homotopy from the identity to 1_X.

The second was the statement that $(F_t)|_A$ is the identity on A for all t — $F(a,t) = a$ for all $a \in A$ and $t \in [0,1]$. Because $F_0 = 1_X$, certainly $(F_0)|_A$ is the identity on A; then the assumption that $(F_t)|_A$ is the identity on A for all t is equivalent to demanding that $(F_t)|_A$ is independent of t.

The last condition was that $F_1(X) \subset A$, which is reproduced here. ☐

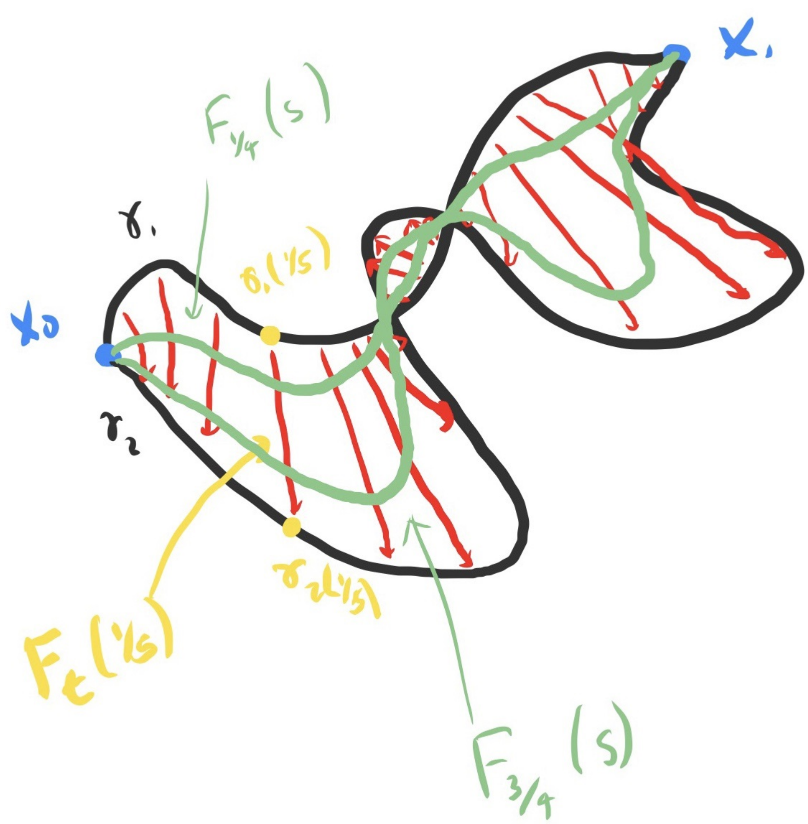
The idea is the same: we're happy with what our map (the identity) does **on** A; we want to homotope the rest of the space into A while leaving A itself alone.

Path homotopy, loops, and based homotopy

We concluded last week by defining the notion of *path homotopy.*

Definition 61. *Let X be a space with two chosen points $x_0, x_1 \in X$ (not necessarily distinct). We say two paths $\gamma_i : [0,1] \to X$ with $\gamma_i(0) = x_0$ and $\gamma_i(1) = x_1$ are path-homotopic if they are homotopic rel $\{0,1\}$; spelled out, they are path-homotopic if there is a homotopy $F : [0,1] \times [0,1] \to X$ satisfying*

1. $F(x,0) = \gamma_1(x)$
2. $F(x,1) = \gamma_2(x)$
3. $F(0,t) = x_0$
4. $F(1,t) = x_1$.



We will be primarily interested in a special case: paths *which start and end at the same point* — or *loops.* These can be interpreted in terms of maps from the circle.

Proposition 97 (Loops rel basepoint are paths rel boundary)**.** *The data of a continuous map $f : [0,1] \to X$ with $f(0) = f(1) = x_0$ is the same data as that of a continuous map $\bar{f} : S^1 \to X$ with $\gamma(1) = x_0$.*

The data of a path-homotopy from f to g is the same data as that of a basepoint-preserving homotopy from $\bar{f}$ to $\bar{g}$.

It will sometimes be useful to us to go back and forth between these two perspectives: loops as maps from the interval and loops as maps from the circle.

Proof. A map $f : [0,1] \to X$ with $f(0) = f(1)$ induces a continuous map $\bar{f} : ([0,1])/(0 \sim 1) \to X$, and vice versa. Since $[0,1]/(0 \sim 1)$ is homeomorphic to S^1 by a homeomorphism sending $[0]$ to 1, we get the first statement in the proposition.

The next part is similar: a path-homotopy is a map $F : [0,1] \times [0,1] \to X$ satisfying the conditions outlined above. The assumption that f and g have the same endpoints means, then, that $F(0,t) = F(1,t)$ for all t. There is thus an induced map

$$\bar{F} : [0,1] \times [0,1]/\big((0,t) \sim (1,t)\big) \to X.$$

The domain is homeomorphic to $([0,1]/(0 \sim 1)) \times [0,1] \cong S^1 \times [0,1]$ by a homeomorphism sending $[(0,t)]$ to $(1,t) \in S^1 \times [0,1]$. Thus the fact that $F(0,t) = x_0$ for all t gives that $\bar{F}(1,t) = x_0$ for all t; similarly, $\bar{F}_0 = \bar{f}$ and $\bar{F}_1 = \bar{g}$.

So from a path-homotopy, we can construct a basepoint-preserving homotopy between the two paths thought of as loops; the converse direction is similarly straightforward (start with a homotopy $S^1 \times [0,1] \to X$ and precompose with our quotient map $[0,1]^2 \to S^1 \times [0,1]$). □

This suggests a new collection of notions, which we will have to keep handy for some time to come.

Definition 62. *A pair of a space X and a chosen basepoint $x_0 \in X$, written (X, x_0), is called a pointed space or a based space.*

A based map between based spaces, written $f : (X, x_0) \to (Y, y_0)$, is a continuous map $f : X \to Y$ with $f(x_0) = y_0$.

A based homotopy between based maps is a homotopy $F : X \times [0,1] \to Y$ from f to g so that $F_t(x_0) = y_0$ for all $t \in [0,1]$.

Remark 69. A based space is based homotopy equivalent to a point if and only if the identity map $1_X : X \to X$ is based-homotopic to the constant map c_{x_0} if and only if X deformation retracts to the basepoint x_0.

Finally we assemble all of what we've talked about so far into a new notion, which we will spend quite some time investigating.

Definition 63. *Let (X, x_0) be a based topological space. We write $\pi_1(X, x_0)$ for the set of* based homotopy classes of loops in X based at x_0. *That is,*

$$\pi_1(X, x_0) = \big\{\gamma : S^1 \to X \mid \gamma(1) = x_0\big\}\Big/\big(\gamma_1 \sim \gamma_2 \text{ if } \gamma_1, \gamma_2 \text{ are based-homotopic}\big).$$

11/11: The fundamental group

Here is an eminently reasonable question: *why are we spending so much time talking about basepoints?* I tried to convince you a while back that we were going to get a lot of mileage out of loops, sure, but I never said anything about this irritating basepoint stuff.

It turns out it's for a very good reason — which will be fleshed out both in this lecture and in your homework.

Theorem 98. *There is a canonical group operation on $\pi_1(X, x_0)$.*

Proof. For this result, the easiest way to represent loops is to think of them as paths $\gamma : [0,1] \to X$ with $\gamma(0) = \gamma(1)$. We know how to *concatenate* continuous paths (assuming the first one ends where the second one starts); that is, we know that if $\gamma_1(1) = \gamma_2(0)$, then

$$(\gamma_1 * \gamma_2)(t) = \begin{cases} \gamma_1(2t) & 0 \leqslant t \leqslant 1/2 \\ \gamma_2(2t-1) & 1/2 \leqslant t \leqslant 1 \end{cases}$$

is a continuous function; it first runs over γ_1 at double speed and then over γ_2 at double speed.

Here we're working with *loops*, so the starting point and ending point are always the same — x_0. So if γ_1, γ_2 are loops based at x_0, then $\gamma_1 * \gamma_2$ is also a loop based at x_0.

I claim that $[\gamma_1] \cdot [\gamma_2] = [\gamma_1 * \gamma_2]$ is a well-defined group operation on $\pi_1(X, x_0)$. This requires we check a number of things.

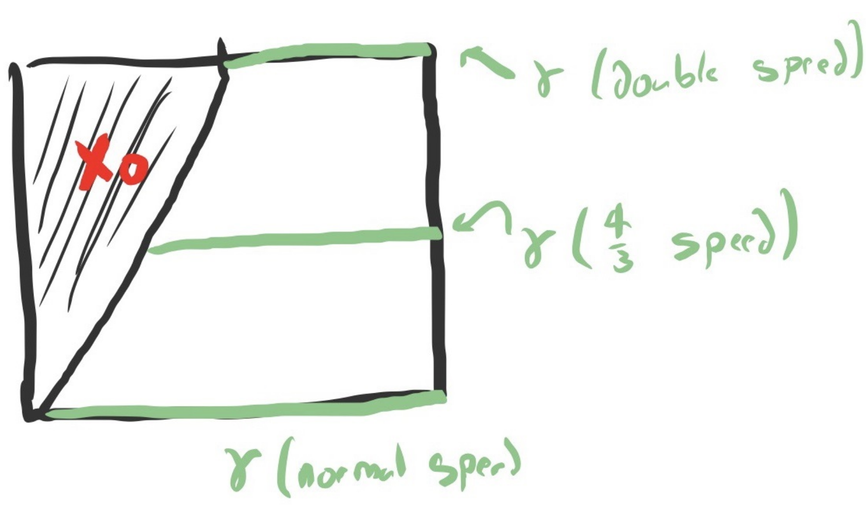
1. Well-defined. If γ_1 is based-homotopic to η_1, and γ_2 is based-homotopic to η_2, we need to show that $\gamma_1 * \gamma_2$ is based-homotopic to $\eta_1 * \eta_2$. If F gives a homotopy from γ_1 to γ_2, and G is a homotopy from η_1 to η_2, then I claim

$$(F * G)(s,t) = \begin{cases} F(2s,t) & 0 \leqslant s \leqslant 1/2 \\ G(2s-1,t) & 1/2 \leqslant s \leqslant 1 \end{cases}.$$

   Then one directly checks that $(F * G)_0 = \gamma_1 * \eta_1$, and $(F * G)_2 = \gamma_2 * \eta_2$. $F * G$ is continuous because it's continuous on each of the closed subsets $[0,1/2] \times [0,1]$ and $[1/2,1] \times [0,1]$, and the union of these subsets is all of $[0,1]^2$.

   It follows that $[\gamma_1 * \eta_1] = [\gamma_2 * \eta_2]$, so that $[\gamma] \cdot [\eta] = [\gamma * \eta]$ is independent of the representatives chosen for the homotopy-classes $[\gamma]$ and $[\eta]$.

2. The first group axiom: the existence of an identity. Write $c : [0,1] \to X$ for the constant map $c(t) = x_0$. I claim that $[c]$ is an identity element for the group operation of composition. The problem is that $c * \gamma$ is not *literally the same* as γ. First we dawdle at x_0 until time $1/2$, and then we do γ at double speed. We'd like to perform a homotopy that 'undoes that': we dawdle for less and less time, and do γ at the appropriate speed. This is encoded in the following picture:



   The left triangle is the part of the homotopy F that just stays constant at x_0, and the part on the right displays the part of the interval $[0,1] \times \{t\}$ that we are running γ along. Formally, the homotopy is given by

$$F_t(s) = \begin{cases} x_0 & 0 \leqslant s \leqslant 1/2 - t/2 \\ \gamma\left(\frac{2(s-1/2+t/2)}{1+t}\right) & 1/2 - t/2 \leqslant s \leqslant 1 \end{cases}$$

This is a piecewise defined map on $[0,1]^2$, with both pieces being closed; since this map is continuous on both of the closed subsets, and agrees on their intersection, it defines a continuous map on all of $[0,1]^2$. Note that $F_t(0) = F_t(1) = x_0$ for all t, so this is a path-homotopy.

F is thus a path-homotopy from $c * \gamma$ to γ, so that $[c] \cdot [\gamma] = [\gamma]$. A similar argument shows that $[\gamma] \cdot [c] = [\gamma]$.

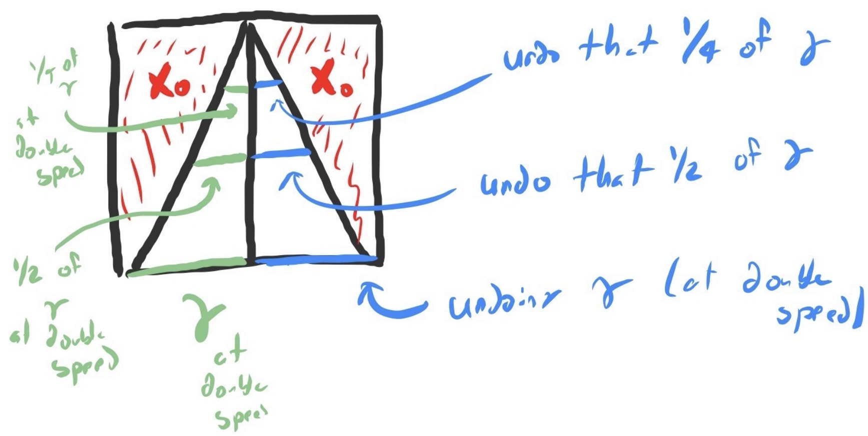
3. The second group axiom: existence of inverses. I claim that if $\gamma : [0,1] \to X$ is a loop based at x_0, then the loop $\overline{\gamma} : [0,1] \to X$ defined by

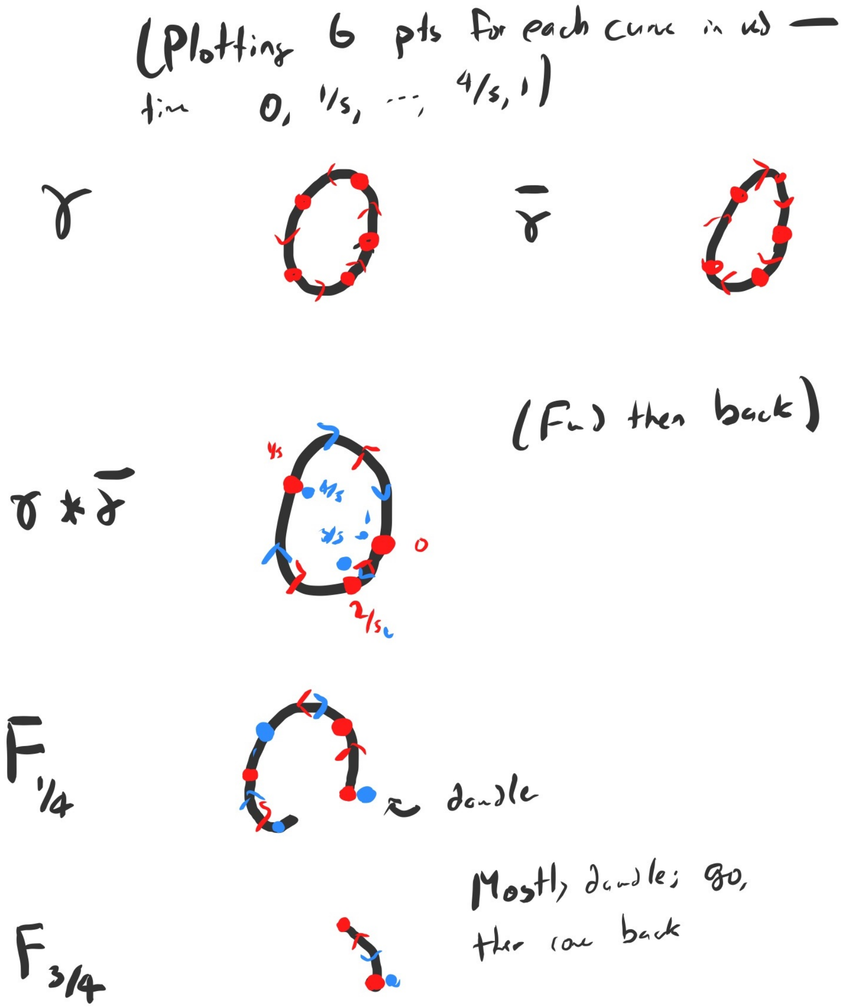
$$\overline{\gamma}(t) = \gamma(1-t)$$

has

$$[\overline{\gamma}] \cdot [\gamma] = [\gamma] \cdot [\overline{\gamma}] = [c].$$

The idea is that we're moving along γ and then walking backwards. We want to just sit at x_0 instead of moving at all. Now imagine we dawdle for a bit at x_0, before following γ *part of the way but not all the way*, and then going back the way we started, and dawdling at x_0 some more.





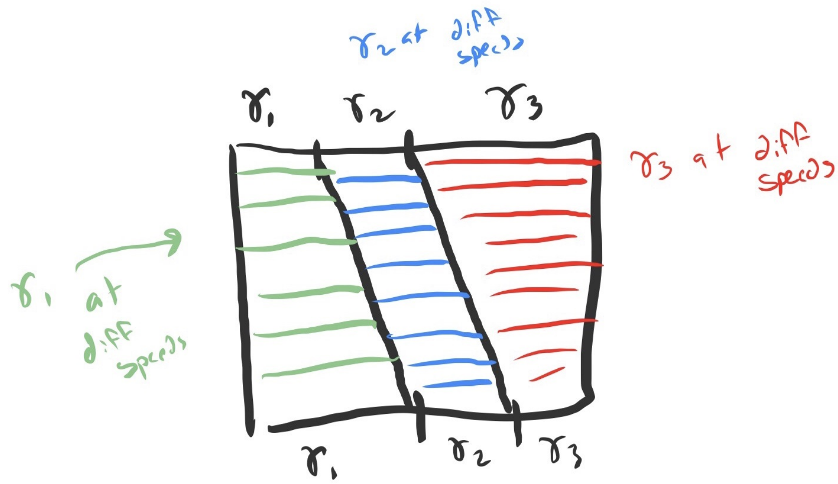
We encode this in a homotopy as

$$F_t(s) = \begin{cases} x_0 & 0 \leqslant s \leqslant t/2 \\ \gamma(2(s - t/2)) & t/2 \leqslant s \leqslant 1/2 \\ \gamma(2(1 - t/2 - s)) & 1/2 \leqslant s \leqslant 1 - t/2 \\ x_0 & 1 - t/2 \leqslant s \leqslant 1 \end{cases}.$$

Again one checks that $F_0 = \gamma * \overline{\gamma}$ and $F_1 = c$; again one checks that $F_t(0) = F_t(1) = x_0$ for all t; again one checks that F is piecewise continuous on $[0,1]^2$ and the definitions match up on the boundary, so that F defines a continuous map on $[0,1]^2$.

Because $\gamma * \overline{\gamma}$ is path-homotopic to c, it follows that $[\gamma] \cdot [\overline{\gamma}] = [c]$. Since we could run the same argument on $\overline{\gamma}$ as well, we also see that $[\overline{\gamma}] \cdot [\gamma] = [c]$. Thus $[\overline{\gamma}] = [\gamma]^{-1}$, and $\pi_1(X, x_0)$ does have inverses.

4. The last group axiom: associativity. The problem is that $\gamma_1 * (\gamma_2 * \gamma_3)$ and $(\gamma_1 * \gamma_2) * \gamma_3$ are not literally the same path. In the first, we traverse γ_1 at double-speed, but γ_2 and γ_3 at *quadruple-speed*; in the second, we traiverse γ_1 and γ_2 at quadruple-speed, but γ_3 at double-speed. So we need to use another homotopy like the above, modulating these different speeds.



This is encoded in formula as

$$F_t(s) = \begin{cases} \gamma_1\left(\frac{4s}{2-t}\right) & 0 \leqslant s \leqslant 1/2 - t/4 \\ \gamma_2(4(s - 1/2 + t/4)) & 1/2 - t/4 \leqslant s \leqslant 3/4 - t/4 \\ \gamma_3\left(\frac{4(s-3/4+t/4)}{1+t}\right) & 3/4 - t/4 \leqslant s \leqslant 1 \end{cases}$$

Once more you check that F is piecewise continuous, is a path-homotopy, and that $F_0 = \gamma_1 * (\gamma_2 * \gamma_3)$ while $F_1 = (\gamma_1 * \gamma_2) * \gamma_3$. Therefore these two elements are path-homotopic, so that the group operation on $\pi_1(X, x_0)$ is indeed associative.

□

Remark 70. You can check by bashing out definitions that if $f : X \to Y$ is a homeomorphism, there is an isomorphism of groups $\pi_1(X, x_0) \cong \pi_1(Y, f(x_0))$. Instead of writing this out, we'll explore a broader approach soon (in lecture, and in Curio).

11/16: Categories, functors, and induced maps

When studying most mathematical objects you have found it useful not just to study the individual objects, but also the *relationships* between them.

- When studying sets, you were interested in sets and the functions between them (and various properties of these functions, such as injectivity and surjectivity).
- When studying linear algebra (over a fixed ground field k — most likely $\mathbb{R}$), you were interested in vector spaces and the *linear maps* between them.
- When studying groups, you were interested in groups and the *homomorphisms* between them.
- When studying topology, you were interested in topological spaces and continuous functions between them.

A category formalizes the notion of studying some collection of objects and a certain class of *morphisms* saying how you can go from one object to another. As in all of the cases above, one should always have an identity map, and one ought to be able to compose morphisms.

Definition 64. *A **category** C consists of the following data.*

1. *A collection*[10] $\mathsf{Ob}(\mathsf{C})$ *of* objects *in the category. We write generic elements as things like X, Y, Z.*
2. *For each $X, Y \in \mathsf{Ob}(\mathsf{C})$, a **set***
$$\mathrm{Hom}_{\mathsf{C}}(X, Y)$$
*called the set of **morphisms** from X to Y. We usually write a generic element as $f : X \to Y$, and use similar letters like g, h, etc.*
3. *For each X, we specify an* identity morphism $1_X \in \mathrm{Hom}_{\mathsf{C}}(X, X)$.
4. *For each X, Y, Z, we have a composition map*
$$\mathrm{Hom}_{\mathsf{C}}(Y, Z) \times \mathrm{Hom}_{\mathsf{C}}(X, Y) \to \mathrm{Hom}_{\mathsf{C}}(X, Z),$$
written $(f, g) \mapsto fg$.

This data is required to satisfy the following properties.

a) *If $f : X \to Y$ is a morphism, then $f1_X = 1_Y f = f$; that is, composing with the identity morphism leaves a morphism unchanged.*

b) *The composition map is associative. If $f : Z \to W$, $g : Y \to Z$, $h : X \to Y$ are composable morphisms, then*
$$(fg)h = f(gh).$$

.

It is often convenient to abbreviate $\mathrm{Hom}_{\mathsf{C}}(X, Y)$ to $\mathsf{C}(X, Y)$. I will mostly do so below.

Example 71. The category Set is the one we are most used to. Its objects are sets. The set of morphisms $\mathsf{Set}(X, Y)$ is precisely the set of functions $f : X \to Y$. The identity morphism is the identity function $1_X(x) = x$. Composition is given by composition of functions.

Example 72. The category Vect_k has objects V, a vector space over the field k, and morphisms $\mathsf{Vect}_k(V, W)$ the set of k-linear maps $L : V \to W$. (This is often denoted $L(V, W)$ in linear algebra courses.) The identity morphism is the identity linear function, and composition is aas you expect.

[10] If you like, you could say there is a set of objects, but this would be something of a lie: we want to say Set is a category, and there is no set of all sets. One normally formalizes this by talking about proper classes, but the set-theoretic stuff is not the most interesting part here.

Example 73. Let me give two more examples here. First is Top, the category whose objects are topological spaces, and whose morphisms $\mathsf{Top}(X, Y)$ are continuous maps $X \to Y$. We have been working with this category for the whole class so far.

The second we have just learned about: the category of based spaces. Write Top_* for the category whose objects are pairs (X, x_0) of a topological space X and a chosen basepoint x_0. The morphisms $f : (X, x_0) \to (Y, y_0)$ in this category are continuous maps $f : X \to Y$ so that $f(x_0) = y_0$.

Let me briefly mention that categories can be of a more unfamiliar flavor as well; not all categories have objects which are sets with some extra structure, and not all morphisms take the form of literal functions between sets (satisfying some properties).

Example 74. Let $\mathsf{C}_{\to}$ be the category with two objects i (input) and o (output), and for which $\mathsf{C}_{\to}(i, i) = \{1_i\}$ and $\mathsf{C}_{\to}(o, o) = \{1_o\}$ consist of only the identity arrows, while $\mathsf{C}_{\to}(i, o) = \{f\}$ consists of a single morphism $f : i \to o$, and $\mathsf{C}_{\to}(o, i)$ is empty.

We can depict this category as the following diagram:

$$\circlearrowleft \bullet \longrightarrow \bullet \circlearrowright$$

The dots represent the objects, the arrows represent the morphisms.

Example 75. The category Rel of sets and relations has as its objects sets, but has more interesting morphisms. A morphism $R : X \to Y$ is no more than a subset $\Gamma_R \subset X \times Y$; this is called a relation, and one usually writes xRy if $(x, y) \in \Gamma_R$. (Think $\sim = R$ when R is an equivalence relation.) If $R : X \to Y$ and $S : Y \to Z$ are relations, their composite is the relation

$$RS = \{(x, z) \mid \exists_{y \in Y} xRy \text{ and } ySz\}.$$

That is, $xRSz$ if and only if there is some y in between with $xRySz$.

The identity relation is $1_X = \Delta_X$ given by the diagonal; if you want, you can verify that $1_Y R = R 1_X$ for any relation $R : X \to Y$.

In the above example, the morphisms cannot at all be interpreted as functions between sets, like in the previous examples.

Example 76. The following categories are the categories of interest in algebraic topology.

Let hTop be the category whose objects are topological spaces X, but whose morphisms

$$\mathsf{hTop}(X, Y) = [X, Y] = \mathsf{Top}(X, Y)/\sim$$

are *homotopy classes* of continuous maps $X \to Y$. That is, two continuous maps are considered equal in this category.

To verify that this is a category, one needs to know that the composition $[f][g]$ of homotopy classes of maps is well-defined up to homotopy, so descends to give a composition map

$$\mathsf{hTop}(Y, Z) \times \mathsf{hTop}(X, Y) \to \mathsf{hTop}(X, Z).$$

You verified this on your homework.

Similarly, one may define hTop_* to be the category whose objects are *based spces*, and whose morphisms are *based homotopy classes* of based maps $f : (X, x_0) \to (Y, y_0)$.

Notice that *in no way* is the set $\mathsf{hTop}(X, Y)$ a set of functions from X to Y satisfying some conditions. Rather, it is a set of *equivalence classes* of functions.

Before moving on, let me point out that this gives us a unified way of talking about equivalences between mathematical objects.

Definition 65. *Let $f : X \to Y$ be a morphism in the category C. We say that f is an isomorphism if there exists a morphism $g : Y \to X$ so that $fg = 1_Y$ and $gf = 1_X$. (If such a g exists, it is necessarily unique. Can you prove it?)*

We say that X, Y are isomorphich if there exists some isomorphism $f : X \to Y$.

This captures the notion of bijection between sets; isomorphism of vector spaces; isomorphism of groups; homeomorphism of topological spaces; and in $\mathsf{hTop}(X)$, homotopy equivalence between topological spaces.

Functors

Very frequently in mathematics we find tools to pass between domains of discourse. In the language of category theory, these are usually encoded by *functors*.

Definition 66. *Let* C, D *be categories. A* ***functor*** $F : \mathsf{C} \to \mathsf{D}$ *is the following data.*

1. *For each* $X \in \mathsf{Ob}(\mathsf{C})$*, an object* $F(X) \in \mathsf{Ob}(\mathsf{D})$*;*
2. *For each* $f \in \mathsf{C}(X, Y)$*, a morphism* $F(f) : F(X) \to F(Y)$ *in* D*.*

These are required to satisfy the following conditions, amounting to saying that F *preserves the structure in the relevant categories.*

a) *The identity morphisms are sent to identity morphisms;* $F(1_X) = 1_{F(X)}$*.*

b) *Composite morphisms are sent to composite morphisms; more precisely,* $F(fg) = F(f)F(g)$*.*

The simplest examples of functors are *forgetful functors*. For instance, there is a functor $U : \mathsf{Top} \to \mathsf{Set}$ which sends a space to its underlying set; if we write a space as (X, T) where X is a set and $\mathcal{T}$ is a topology on X, this functor has $U(X, \mathcal{T}_X) = X$.

On morphisms, it sends a continuous function $f : (X, \mathcal{T}_X) \to (Y, \mathsf{T}_Y)$ to the underlying function $f : X \to Y$, forgetting the topologies involved.

Another good example of a forgetful functor is $U : \mathsf{Top}_* \to \mathsf{Top}$, the functor $U(X, x_0) = X$ which sends a based space to its underlying space; that is, it forgets the basepoint.

Lastly, there is a funcntor $\mathsf{Top} \to \mathsf{hTop}$ which is *the identity on objects*, but sends a map $f : X \to Y$ to its homotopy class.

Another common class of functors are *free functors*.

Example 77. There are two rather silly functors $D, I : \mathsf{Set} \to \mathsf{Top}$; the first is called the *discrete space functor* and the latter the *indiscrete spaace functor*.

If X is a set, $D(X)$ is X equipped with the discrete topology. If $f : X \to Y$ is a function, $D(f)$ is the same function, considered as a map between the topological spaces $D(X) \to D(Y)$. Notice that this is always continuous, because any function from a discrete space is continuous.

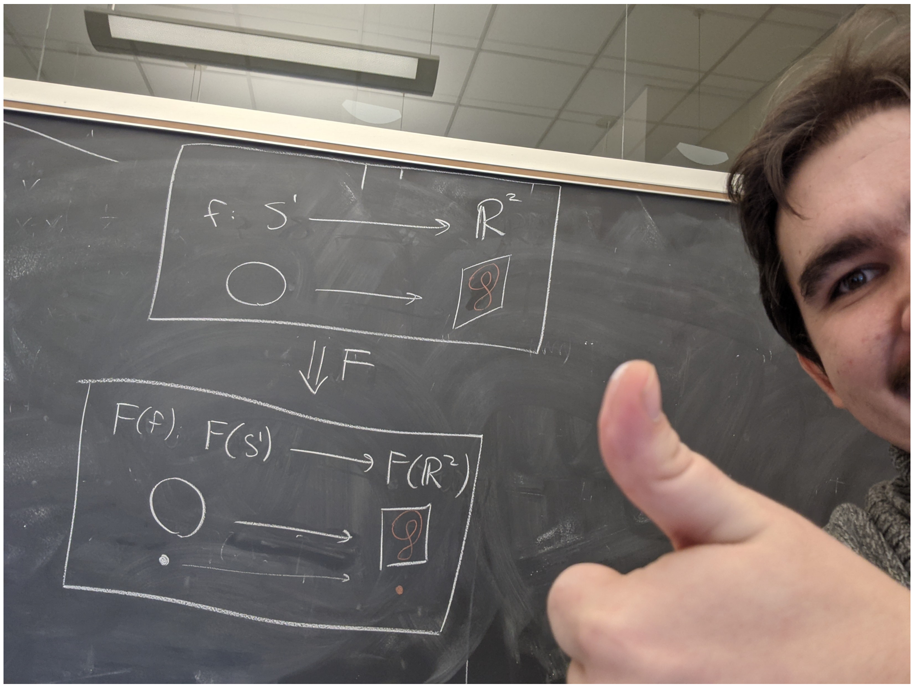
I is similar; $I(X)$ is X equipped with the indiscrete topology, and $I(f)$ is again the same underlying function, continuous because maps *to* indiscrete spaces are continuous.

Example 78. There is a free functor $F : \mathsf{Top} \to \mathsf{Top}_*$, called the *disjoint basepoint functor*, which sends

$$F(X) = (X \sqcup \{*\}, *).$$

A space X is sent to the disjoint union with a singleton space, and that new point is taken to be the basepoint. (Some authors write $F(X)$ as X_+, but this clashes with our notation for the 1-point compactification.) If $f : X \to Y$ is a continuous map, the new map $F(f) : X \sqcup \{*\} \to Y \sqcup \{*\}$ is $f : X \to Y$ on the first component and the identity $* \to *$ on the second component.

A picture of this example is drawn on the blackboard below.



Example 79. One I like is the *free abelian group functor* $F : \mathsf{Set} \to \mathsf{AbGp}$, where the latter is the category whose objects are abelian groups and whose morphisms are group homomorphisms. This functor sends a set X to the abelian group $\mathbb{Z}[X]$, where

$$\mathbb{Z}[X] = \Big\{ \sum_{x \in X} n_x x \mid n_x = 0 \text{ for all but finitely many } x \Big\}.$$

This is a formal sum — one cannot evaluate it more than it has already been written. If we were working over a field instead of $\mathbb{Z}$, you could think of this as being a vector space whose basis vectors correspond to points in X. I think of an element of $\mathbb{Z}[X]$ as a finite sum of delta masses above different points in X.

If $f : X \to Y$ is a set function, the induced map $F(f) : \mathbb{Z}[X] \to \mathbb{Z}[Y]$ sends

$$F(\sum n_x x) = \sum_{x \in X} n_x f(x) = \sum_{y \in Y} \sum_{\substack{x \in X \\ f(x) = y}} n_x y.$$

This is always a finite sum, because $n_x = 0$ for all but finitely many x.

Functors preserve any property that can be stated purely categorically. We only have one of these so far, so I'll state that as a proposition.

Proposition 99 (Functors preserve isomorphism)**.** *If X and Y are isomorphic objects of C, and $F : \mathsf{C} \to \mathsf{D}$ is a functor, then $F(X) \cong F(Y)$ in D, as well.*

Proof. Suppose $f : X \to Y$ and $g : Y \to X$ are morphisms in C so that $fg = 1_Y$ and $gf = 1_X$. Then I claim $F(f) : F(X) \to F(Y)$ and $F(g) : F(Y) \to F(X)$ are inverse isomorphisms as well. To see this, observe that

$$F(f)F(g) = F(fg) = F(1_Y) = 1_{F(Y)};$$

the outer two identities are by the definition of a functor, and the inner by the assumption that $fg = 1_Y$. Similarly, we have

$$F(g)F(f) = F(gf) = F(1_X) = 1_{F(X)}.$$

Therefore $F(X)$ and $F(Y)$ are isomorphic as well. □

Functors in topology

You have already seen a small handful of interesting functors in this course which are not of the above types.

Example 80. Let Comp be the category whose objects are compact topological spaces, and whose morphisms are the continuous maps between them. Let $\mathsf{PropTop}$ be the category whose objects are topological spaces and whose morphisms $f : X \to Y$ are the **proper** continuous maps from X to Y.

The *one-point compactification* gives a functor

$$C : \mathsf{PropTop} \to \mathsf{Comp},$$

given by $C(X) = X_\infty$ (the Alexandroff one-point compactification) on objects, and $C(f) : X_\infty \to Y_\infty$ given by

$$C(f)(x) = \begin{cases} f(x) & x \neq \infty \\ \infty & x = \infty \end{cases}.$$

We showed in class that when f is a proper map, $C(f)$ is continuous. It is also clear that $C(1_X) = 1_{C(X)}$ and $C(fg) = C(f)C(g)$.

Example 81. In your homework, you showed that the *path components* define a functor

$$\pi_0 : \mathsf{Top} \to \mathsf{Set},$$

given by sending a space X to its set of path-components $\pi_0 X$, and if $f : X \to Y$ is a continuous map, recording which path components it sends each into; that is,

$$\pi_0 f : \pi_0 X \to \pi_0 Y$$

is given by sending $(\pi_0 f)[x] = [f(x)]$, where $[x]$ is the path-component of $x \in X$.

You checked that this was well-defined, and that $\pi_0(1_X) = 1_{\pi_0 X}$, and $\pi_0(fg) = \pi_0(f)\pi_0(g)$. So you showed that this defines a functor.

You actually showed slightly more: you verified that when $f : X \to Y$ is homotopic to $g : X \to Y$, the induced maps $\pi_0(f) = \pi_0(g)$ are the same. This proves that this actually defines a functor

$$\mathsf{hTop} \to \mathsf{Set}$$

on the *homotopy category of spaces.*

It follows that if X and Y are homotopy equivalent, they have the same number of path-components.

Induced maps on the fundamental group

What we'll show next is that π_1 fits into this framework as well.

Theorem 100 (Existence of induced maps)**.** *Let $f : (X, x_0) \to (Y, y_0)$ be a based map. There is an induced homomorphism*

$$f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0),$$

satisfying

1. *If $g : (X, x_0) \to (Y, y_0)$ is a based map, and $f : (Y, y_0) \to (Z, z_0)$ is a based map, then $(fg)_* = f_* g_*$;*
2. *If $f, g : (X, x_0) \to (Y, y_0)$ are **based** homotopic, then $f_* = g_*$.*
3. *If $1_X : X \to X$ is the identity map, the induced map $(1_X)_* : \pi_1(X, x_0) \to \pi_1(X, x_0)$ the identity homomorphism.*

This theorem, phrased in the language above, says that π_1 defines a functor $\mathsf{hTop}_* \to \mathsf{Gp}$; to each space-with-basepoint we assign a group $\pi_1(X, x_0)$, and to each continuous-map-preserving-basepoints $f : (X, x_0) \to (Y, y_0)$ we associate a map $f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$, and this map is the same for any two f, g in the same based-homotopy class. Furthermore, these send the identity to the identity and send composites to comoposites. This is exactly what a functor is.

Proof. The map we have in mind is $f_*[\gamma] = [f\gamma]$ — that is, if $\gamma : [0,1] \to X$ is a loop based at x_0 (so $\gamma(0) = \gamma(1) = x_0$), then the composite $f\gamma : [0,1] \to Y$ is also because

$$f(\gamma(0)) = f(x_0) = y_0,$$

and similarly with $f(\gamma(1))$. So we want to send f_* to send the *based homotopy class* of γ in $\pi_1(X, x_0)$ to this homotopy class in $\pi_1(Y, y_0)$.

We first need to justify two things. First, that this map is well-defined, and second, that this map is a group homomorphism.

To see that this is well-defined, notice that if γ_1 is based-homotopic to γ_2 via the based homotopy $H_s : [0,1] \times [0,1] \to X$, then $fH_s : [0,1] \times [0,1] \to Y$ also defines a based homotopy from $f\gamma_1$ to $f\gamma_2$. (Spell this out, as practice.)

Thus $[f\gamma_1] = [f\gamma_2]$, and $f_*[\gamma]$ is well-defined (the output is independent of choice of representative of $[\gamma]$).

To see that f_* is a *homomorphism*, we need to know that $f(\gamma_1 * \gamma_2) = f\gamma_1 * f\gamma_2$. But explicitly,

$$(\gamma_1 * \gamma_2)(t) = \begin{cases} \gamma_1(2t) & 0 \leqslant t \leqslant 1/2 \\ \gamma_2(2t-1) & 1/2 \leqslant t \leqslant 1 \end{cases},$$

while

$$(f\gamma_1 * f\gamma_2)(t) = \begin{cases} (f\gamma_1)(2t) & 0 \leqslant t \leqslant 1/2 \\ (f\gamma_2)(2t-1) & 1/2 \leqslant t \leqslant 1 \end{cases};$$

we explicitly see from this that $f(\gamma_1 * \gamma_2) = f\gamma_1 * f\gamma_2$, as desired.

The next claim is that $(gf)_* = g_* f_*$. But this is also a definition chase:

$$g_* f_*[\gamma] = g_*([f\gamma]) = [g(f\gamma)] = [(gf)\gamma] = (gf)_*[\gamma].$$

The penultimate claim is that if f, g are based homotopic, then $f_* = g_*$. But if H_s is a based homotopy from f to g, and if $\gamma : [0,1] \to (X, x_0)$ is a loop, then $H_s\gamma$ is a based homotopy from $f\gamma$ to $g\gamma$ (again, check this explicitly). Hence

$$f_*[\gamma] = [f\gamma] = [g\gamma] = g_*[\gamma].$$

The last claim is that $(1_X)_* = 1_{\pi_1(X)}$. This again follows from the definition: $(1_X)_*[\gamma] = [1_X\gamma] = [\gamma]$; the induced map is the identity. ☐

These homomorphisms are what make the fundamental group tick; they are what make it so useful.

Corollary 101. *If (X, x_0) and (Y, y_0) are* ***based*** *homotopy equivalent, then $\pi_1(X, x_0) \cong \pi_1(Y, y_0)$.*

Proof. Suppose $f : (X, x_0) \to (Y, y_0)$ is a (based) homotopy equivalence, with homotopy inverse $g : (Y, y_0) \to (X, x_0)$, so that $gf \sim 1_X$ by a based homotopy, and $fg \sim 1_Y$ by a based homotopy.

Then we have homomorphisms $f_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$ and $g_* : \pi_1(Y, y_0) \to \pi_1(X, x_0)$.

Because $gf \sim 1_X$, we see that

$$g_* f_* = (gf)_* = (1_X)_* = 1_{\pi_1(X)};$$

similarly $f_* g_* = 1_{\pi_1(Y)}$. Because g_* is both a right and left inverse for f_*, it follows that f_* is a bijection; a bijective homomorphism is an isomorphism, so $\pi_1(X, x_0) \cong \pi_1(Y, y_0)$. ☐

So far we still haven't proved that there is a single space for which $\pi_1(X, x_0)$ is nontrivial. Next week we will show that $\pi_1(S^1, 1)$ is nontrivial, which implies the following corollary.

Corollary 102 (No-retract theorem)**.** *There is no continuous retraction $D^2 \to S^1$.*

Proof. We have a continuous map $i : S^1 \to D^2$ including the boundary circle. Suppose, towards a contradiction, that there is a continuous map $r : D^2 \to S^1$ with $ri = 1_{S^1}$; that is, a retraction.

Consider the induced maps

$$\pi_1(S^1, 1) \xrightarrow{i_*} \pi_1(D^2, 1) \xrightarrow{r_*} \pi_1(S^1, 1).$$

The middle term is the trivial group, and so

$$r_* i_* [\gamma] = r_* [c] = [c];$$

so $r_* i_*$ is the trivial homomorphism. On the other hand, $r_* i_* = (ri)_* = (1_{S^1})_*$ is the identity map on $\pi_1(S^1, 1)$. As we will see tomorrow, this group is nontrivial; this is a contradiction, and no such map r can exist. □

In the next section we get our hands dirty with some examples.

11/18: Calculations and covering spaces

The following is our first truly non-trivial result on fundamental groups. It requires new ideas. Pursuing these ideas leads to *covering space theory*, which is often a few weeks in an introductory algebraic topology course; I just want to give you a taste.

Theorem 103. *There is an isomorphism* $\pi_1(S^1, 1) \cong \mathbb{Z}$.

Proof. It will be convenient, notationally, to think of S^1 as $\mathbb{R}/\mathbb{Z}$, the orbit space of $\mathbb{R}$ under the action of the integers by translation. These are homeomorphic, hence we certainly have $\pi_1(\mathbb{R}/\mathbb{Z}, [0]) \cong \pi_1(S^1, 1)$. We will need the following two facts.

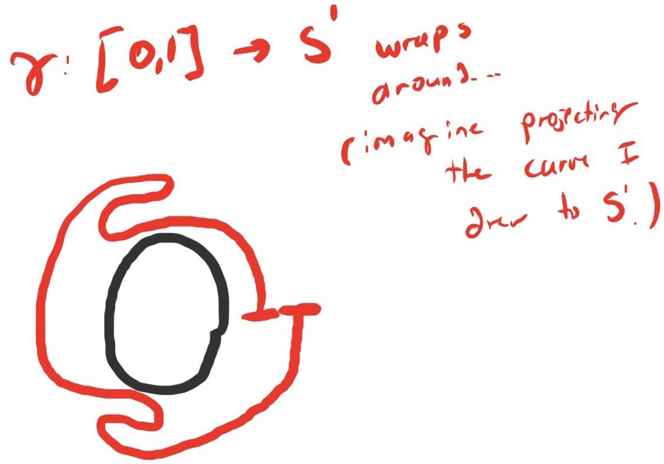
Fact 1. If $\gamma : [0,1] \to \mathbb{R}/\mathbb{Z}$ is a continuous loop with $\gamma(0) = \gamma(1) = [0]$, there is a unique continuous lift to a path $\widetilde{\gamma} : [0,1] \to \mathbb{R}$ so that $p\widetilde{\gamma} = \gamma$ and $\widetilde{\gamma}(0) = 0$. The path $\widetilde{\gamma}$ need not be a loop.

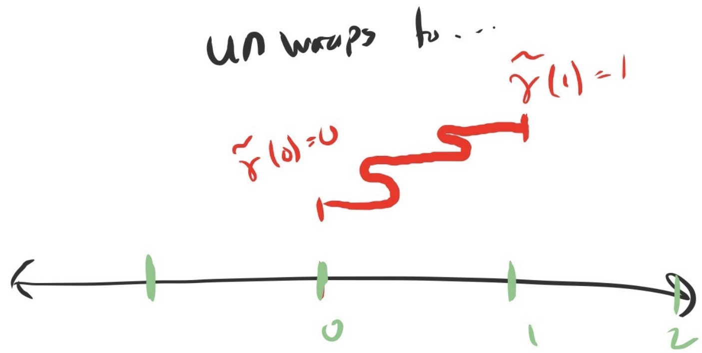
Fact 2. If $F : [0,1] \times [0,1] \to \mathbb{R}/\mathbb{Z}$ is a path-homotopy between two loops (so $F(0,t) = F(1,t) = [0]$ for all t, while $F_0 = \gamma_0$ and $F_1 = \gamma_1$, then there is a unique continuous lift $\widetilde{F} : [0,1] \times [0,1] \to \mathbb{R}$ to a path-homotopy between the lifted paths $\widetilde{\gamma}_0$ and $\widetilde{\gamma}_1$.

Putting these together... I claim that there is a homomorphism $\mathrm{wn} : \pi_1(\mathbb{R}/\mathbb{Z}, [0]) \to \mathbb{Z}$, called the *winding number*, defined as follows. If $\gamma : [0,1] \to \mathbb{R}/\mathbb{Z}$ is a loop based at $[0]$, write $\widetilde{\gamma}$ for the unique lift promised by Fact 1. Then since

$$p\widetilde{\gamma}(1) = \gamma(1) = [0],$$

it follows that $\widetilde{\gamma}(1) \in \mathbb{Z}$; we write $\mathrm{wn}(\gamma) = \widetilde{\gamma}(1)$. This is well-defined because the lift $\widetilde{\gamma}$ is *unique.*





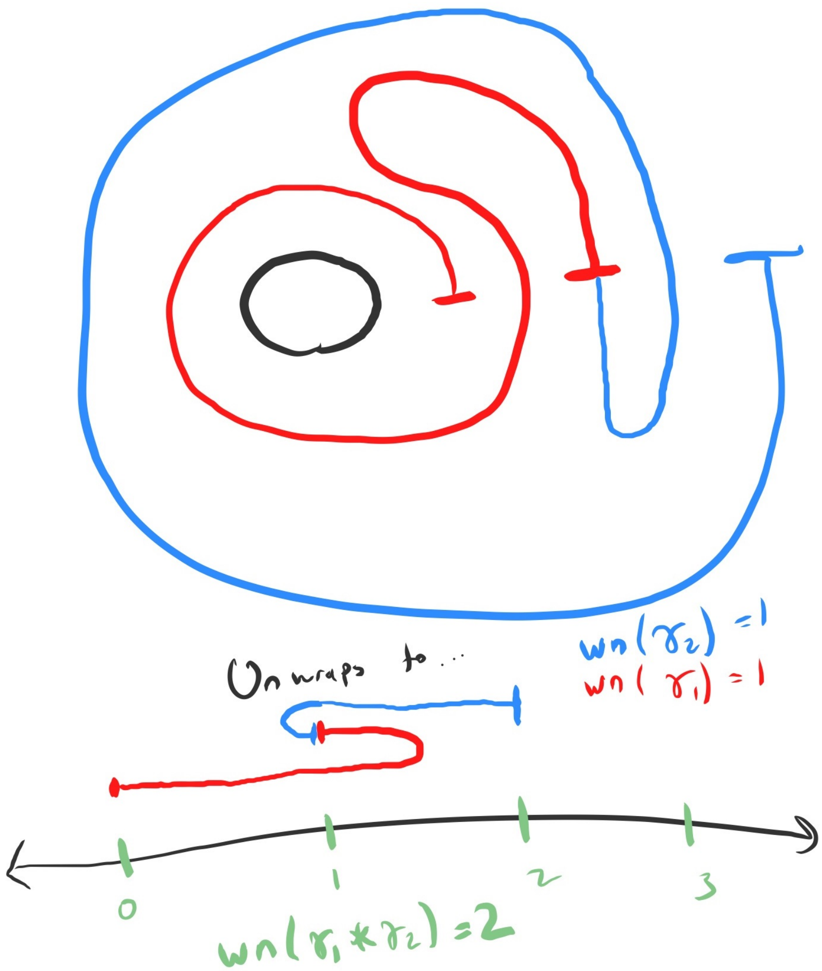
To see that wn defines a map on π_1, we need to see that based-homotopic loops have the same winding number. This follows from Fact 2 above: if $\gamma_0 \sim \gamma_1$, then we may lift the based homotopy F to a path-homotopy $\widetilde{F}$; because $\widetilde{F}$ is a path-homotopy, we have in particular that $\widetilde{\gamma}_0(1) = \widetilde{\gamma}_1(1)$.

To see that this is a homomorphism, consider $\gamma_2 * \gamma_1$. To concatenate the lifts, we should take $\widetilde{\gamma}_2 + \mathrm{wn}(\gamma_1)$, which indeed starts at $\widetilde{\gamma}_1(0)$. Then

$$(\widetilde{\gamma}_2 + n) * \widetilde{\gamma}_1$$

is a continuous lift of $\gamma_2 * \gamma_1$, hence the unique continuous lift; so

$$\mathrm{wn}(\gamma_2 * \gamma_1) = \widetilde{\gamma}_2(1) + n = \mathrm{wn}(\gamma_2) + \mathrm{wn}(\gamma_1).$$



Therefore $\mathrm{wn} : \pi_1(\mathbb{R}/\mathbb{Z}, [0]) \to \mathbb{Z}$ is a well-defined homomorphism. We need to show that it is a bijection.

To see surjectivity, observe that for any integer n, the path $\gamma_n : [0,1] \to \mathbb{R}/\mathbb{Z}$, given by $\gamma_n(t) = [nt]$, is a loop ($\gamma_n(1) = [n] = [0]$) and has as lift $\widetilde{\gamma}(t) = nt$. Because $\widetilde{\gamma}(1) = n$, this path has $\mathrm{wn}(\gamma_n) = n$. So wn is surjective.

To see injectivity, suppose $\mathrm{wn}(\gamma) = 0$; we want to show that γ there is a path-homotopy from γ to the constant map.

By definition, $\mathrm{wn}(\gamma) = 0$ means that the lift $\widetilde{\gamma} : [0,1] \to \mathbb{R}$ is a *loop*: it starts and ends at 0. Consider the path-homotopy given by $\widetilde{F} : [0,1] \times [0,1] \to \mathbb{R}$ given by $\widetilde{F}(x,t) = (1-t)\widetilde{\gamma}(x)$. Then $p\widetilde{F}$ is a perfectly good path-homotopy from $p\widetilde{\gamma} = \gamma$ to the constant map at $[0]$. Therefore, if $\mathrm{wn}(\gamma) = 0$, then $[\gamma] = 0$.

Putting all of this together, we have shown that

$$\mathrm{wn} : \pi_1(\mathbb{R}/\mathbb{Z}, [0]) \to \mathbb{Z}$$

is an isomorphism. □

Remark 82. There is no corresponding version of our 'wrapping around' map $\mathbb{R} \to S^1$ for the sphere. In fact, we will see in a couple weeks (depending on when you're reading this) that $\pi_1(S^n)$ is the trivial group for all $n > 1$. If you want to learn a proof early, Hatcher has an argument in Chapter 1.1 of his book.

Intuitively, any loop that is *not surjective* has to be null-homotopic, because $S^n \backslash \{p\}$ is homeomorphic to $\mathbb{R}^n$. Unfortunately, surjective loops exist (think space-filling curves). So one first homotopes the path so that it is not so hideous, and in particular not space-filling.

Now let's prove those technical facts at the beginning of this proof.

Lemma 104 (The lifting lemma)**.** *Given any continuous map $\gamma : [0,1] \to \mathbb{R}/\mathbb{Z}$ and any choice of $x_0 \in \mathbb{R}$ with $\gamma(0) = [x_0]$, there is a unique continuous lift $\tilde{\gamma} : [0,1] \to \mathbb{R}$ with $\tilde{\gamma}(0) = x_0$.*

Proof. Consider the projection map $p : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$, sending $p(x) = [x]$. We will use the following repeatedly:

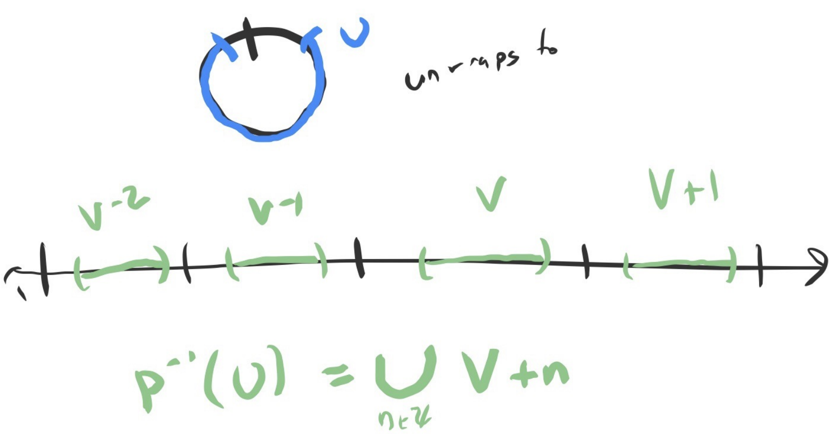
Fact 1. *When restricted to any interval (a,b) with $b < a + 1$, the map $p|_{(a,b)} : (a,b) \to \mathbb{R}/\mathbb{Z}$, is a homeomorphism onto its image.*

This follows because $p|_{(a,b)}$ is an open injection (hence an open bijection onto its image, hence a homeomorphism onto its image). It is clearly an injection, because if $p(x) = p(y)$, then $y = x + n$ for n an integer; but $|y - x| < |b - a| \leqslant 1$, so if their difference is an integer they must be equal.

To see why $p|_{(a,b)}$ is open, note that if $U \subset (a,b)$ is open, then $p(U)$ is open iff $p^{-1}(p(U))$ is open (by definition of the quotient topology). But

$$p^{-1}(p(U)) = \bigcup_{n \in \mathbb{Z}} U + n = \bigcup_{n \in \mathbb{Z}} \{x \mid x - n \in U\}.$$

Because 'translation by n' is a homeomorphism, this is a union of open sets, hence is itself open. Therefore we get the desired claim.



Fact 2. *If $\gamma : [a,b] \to \mathbb{R}/\mathbb{Z}$ is any continuous path whose image misses a point $[t] \in \mathbb{R}/\mathbb{Z}$, and we choose a representative $x_0 \in \mathbb{R}$ for the class $\gamma(a) \in \mathbb{R}/\mathbb{Z}$, then there is a unique continuous 'lift' $\tilde{\gamma} : [a,b] \to \mathbb{R}$ with $\tilde{\gamma}(a) = x_0$ with $p\tilde{\gamma} = \gamma$.*

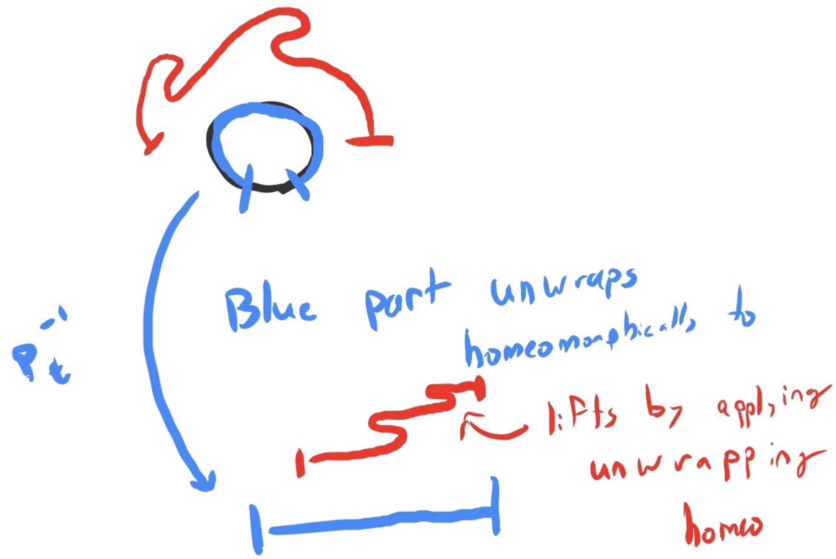
This means that we can find a continuous map $\tilde{\gamma}$, so that $\tilde{\gamma}(t)$ picks out an element of the equivalence class $\gamma(t) \in \mathbb{R}/\mathbb{Z}$. If we think of $\mathbb{R}$ as being an 'unfolding' of $\mathbb{R}/\mathbb{Z}$, this is which 'layer' of the unfolding we're on.

To see this, note that $\gamma([a,b])$ is not all of $\mathbb{R}/\mathbb{Z}$; it misses the class $[t]$. Write t for the unique representative of $[t]$ with $t < x_0 < t + 1$. Then Fact 1 tells us that

$$p|_{(t,t+1)} : (t, t+1) \to (\mathbb{R}/\mathbb{Z}) \backslash \{[t]\}$$

is a homeomorphism; we write it as p_t for convenience. If we set $\tilde{\gamma} = p_t^{-1}\gamma$, then we find $p_t^{-1}\gamma(a) = x_0$, and by definition, pf_t^{-1} is the identity, so $p\tilde{\gamma} = \gamma$.

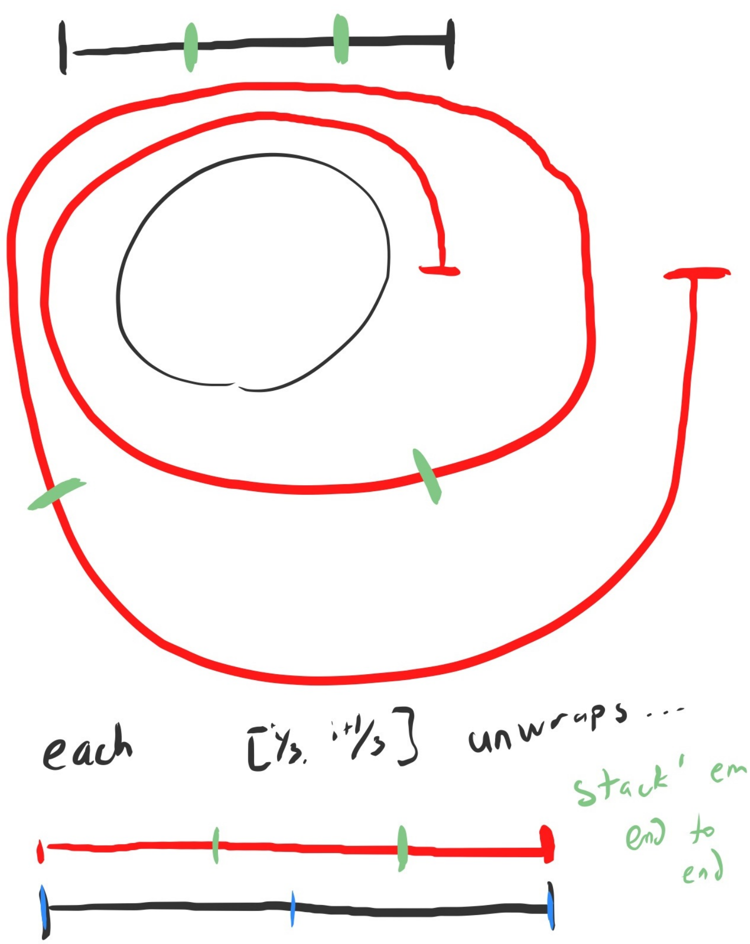
That $\widetilde{\gamma}$ is unique follows from the identity $p_t\widetilde{\gamma} = \gamma$, as one may apply p_t^{-1} on both sides to obtain $\widetilde{\gamma} = p_t^{-1}\gamma$.



Fact 3. For any continuous path $\gamma : [0, 1] \to \mathbb{R}/\mathbb{Z}$, there is an integer $n > 0$ so that all of $\gamma([i/n, (i + 1)/n]) \subset \mathbb{R}/\mathbb{Z}$ are proper subsets of $\mathbb{R}/\mathbb{Z}$; that is, so long as we take these little intervals small enough, γ cannot wrap around the entire circle.

To see this we apply the Lebesgue number lemma. Write $U_t = \{[t]\}^c \subset \mathbb{R}/\mathbb{Z}$; we are trying to show that γ sends every sufficiently small interval into one of these U_t's.

Consider the open cover $\mathcal{U} = \{\gamma^{-1}(U_t)\}$. Because $[0, 1]$ is compact, the Lebesgue number lemma tells us that there is a uniform $\epsilon > 0$ so that every interval $(x - \epsilon, x + \epsilon)$ is contained in some $\gamma^{-1}(U_t)$. Taking $1/n < \epsilon$, this implies the stated claim.



Finishing the proof.

To see the claim in the Lemma, we know by Fact 3 that for n sufficiently small, each of $\gamma([i/n,(i+1)/n])$ misses a point $[t] \in \mathbb{R}/\mathbb{Z}$. Let's inductively show that we can uniquely lift our map along $[0, i/n]$. For the first interval $[0, 1/n]$, we know that $\gamma([0, 1/n])$ misses some point $[t]$, so by Fact 2 there is a unique lift $\widetilde{\gamma} : [0, 1/n] \to \mathbb{R}$ so that $\widetilde{\gamma}(0) = x_0$ with $p\widetilde{\gamma} = \gamma$.

Now by inductive hypothesis we know we can find a unique lift $\widetilde{\gamma} : [0, i/n] \to \mathbb{R}$ so that $\widetilde{\gamma}(0) = x_0$ and $p\widetilde{\gamma} = \gamma$. Write $x_i = \widetilde{\gamma}(i/n)$. As in the previouos paragraph, we can combine Fact 3 and Fact 2 to see that there is a lift $\widetilde{\gamma} : [i/n,(i+1)/n] \to \mathbb{R}$ with $\widetilde{\gamma}(i/n) = x_i$, and $p\widetilde{\gamma} = \gamma$.

We thus can define $\widetilde{\gamma}$ as a continuous function on both $[0, i/n]$ and $[i/n,(i+1)/n]$, so that the values at the two copies of i/n agree; because $\widetilde{\gamma}$ is continuous on both of these closed sets and agrees on their intersection, it follows that $\widetilde{\gamma}$ defines a continuous function on $[0,(i+1)/n]$ with $p\widetilde{\gamma} = \gamma$ and $\widetilde{\gamma}(0) = x_0$. By inductive hypothesis there is a unique way to lift on $[0, i/n]$, and by Fact 2 there is a unique way to extend this lift to $[i/n,(i+1)/n]$, so this lift on $[0,(i+1)/n]$ is unique.

It follows that we can construct a unique lift on the entirety of $[0, 1]$ once we specify $\widetilde{\gamma}(0) = x_0$.

□

I would somehow summarize this as: *we can locally lift paths uniquely, so after chopping our path into a bunch of pieces, we can globally lift paths uniquely, too; we just need to know where to start from. If they* *started as loops, they might not be after lifting.*

Lemma 105 (Homotopy lifting lemma). *If $\gamma_1, \gamma_2 : [0,1] \to \mathbb{R}/\mathbb{Z}$ are path-homotopic (and in particular have the same endpoints) via the homotopy F, and if $\widetilde{\gamma}_1, \widetilde{\gamma}_2 : [0,1] \to \mathbb{R}$ are lifts with the same starting-point ($\widetilde{\gamma}_1(0) = \widetilde{\gamma}_2(0)$), then there is a unique lift of F to a path-homotopy $\widetilde{F}$ from $\widetilde{\gamma}_1$ to $\widetilde{\gamma}_2$. In particular, $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ have the same end-point as well.*

Proof. The proof of this fact is largely the same as the previous one; I will only outline the proof.

First, we chop $[0,1]^2$ into the sub-squares $[i/n, (i+1)/n] \times [j/n, (j+1)/n]$. For $1/n^2$ sufficiently small, the Lebesgue number lemma and the argument in the previous proof guarantees that we can lift all of these squares uniquely once we specify the value of $\widetilde{F}(i/n, j/n)$. One first argues inductively that we can paste these together to get a continuous lift on $[0,1] \times [j, (j+1)/n]$ once we specify $\widetilde{F}(0, j/n)$, and then pastes these rectangles together to argue that there is a unique lift $\widetilde{F}$ once you specify $\widetilde{F}(0,0)$.

It remains to justify that $\widetilde{F}$ is a path-homotopy from $\widetilde{\gamma}_1$ to $\widetilde{\gamma}_2$.

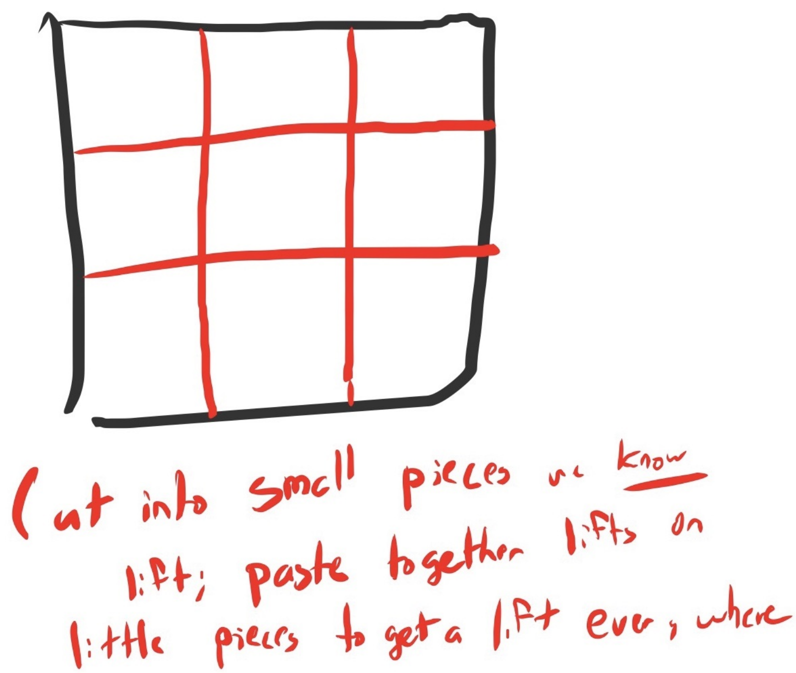
This justifies the existence of a continuous lift $\widetilde{F}$. Note that $\widetilde{F}_0$ is a lift of $F_0 = \gamma_1$; choosing the lift with $\widetilde{F}_0(0,0) = \widetilde{\gamma}_1(0)$, it then follows that $\widetilde{F}_0 = \widetilde{\gamma}_1$.

Because $F(0,t) = \widetilde{\gamma}_1(0)$ for all t, it follows that $\widetilde{F}(0,t) \in \mathbb{Z}$ for all t; because $\widetilde{F}(0,t)$ is continuous in t, it follows that $\widetilde{F}(0,t)$ is a continuous map from a connected space to a discrete space, hence is constant. So $\widetilde{F}(0,t) = \widetilde{\gamma}_1(0)$ for all t. The same argument applies to see that $\widetilde{F}(1,t) = \widetilde{\gamma}_1(t)$ for all t.

Lastly, because $\widetilde{F}(1,0) = \widetilde{\gamma}_1(0)$, which is also $\widetilde{\gamma}_2(0)$, and $\widetilde{F}_1$ is a lift of $\widetilde{\gamma}_2$, it follows that $\widetilde{F}_1 = \widetilde{\gamma}_2$. Thus $\widetilde{F}$ is indeed a path-homotopy from $\widetilde{\gamma}_1$ to $\widetilde{\gamma}_2$. Because $\widetilde{F}(1,t)$ is constant in t, it follows that

$$\widetilde{\gamma}_1(1) = \widetilde{\gamma}_2(1).$$

□



I find this proof remarkable, pleasant, and insightful. I will try to summarize the ideas as well as I can before we see how it generalizes.

- There is a way to 'unfold' the circle: the circle appears to be the real line 'folded up on itself'. It is fruitful to compare paths 'downstairs' to paths 'upstairs'.

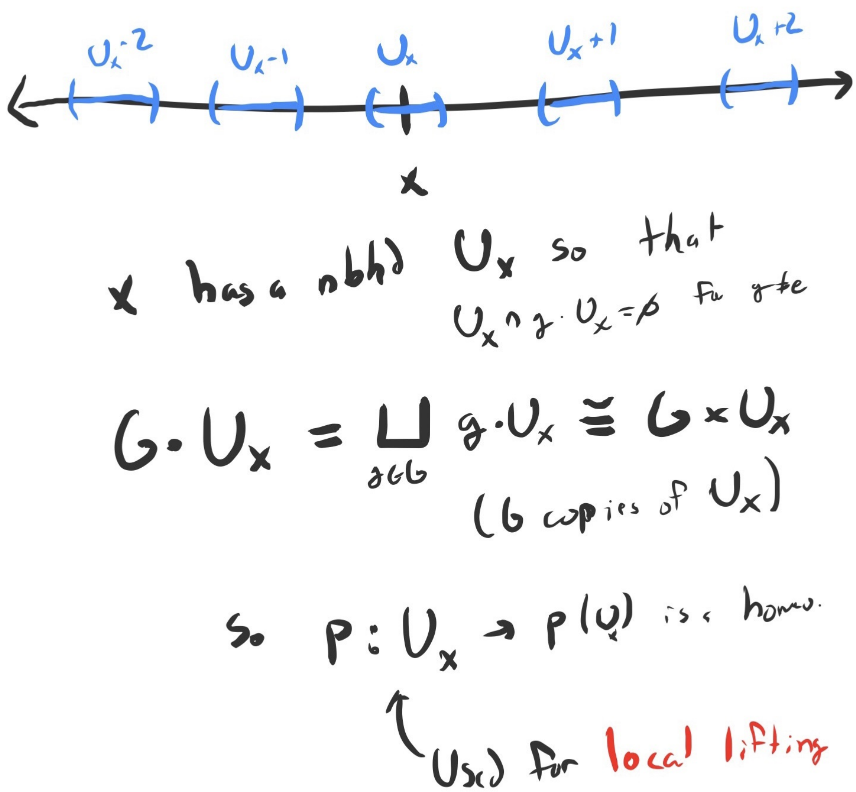
- Formally, the way to go from $\mathbb{R}$ to the circle is by taking the orbit space $\mathbb{R}/\mathbb{Z}$ for a particularly well-behaved group action on $\mathbb{R}$. We needed to be able to lift paths from downstairs that only moved 'a little bit'.

- We needed to know that the 'upstairs space' was simply connected to show that our 'winding number' map is injective.

We formalize the first two of these as a definition and the last of these as a theorem.

Definition 67. *Let X be a topological space. A covering space action of a group G on X is a group action $G \curvearrowright X$ so that, for all $x \in X$, there is an open set $x \in U \subset X$ so that $g \cdot U \cap U = \varnothing$ for all $g \neq e$.*

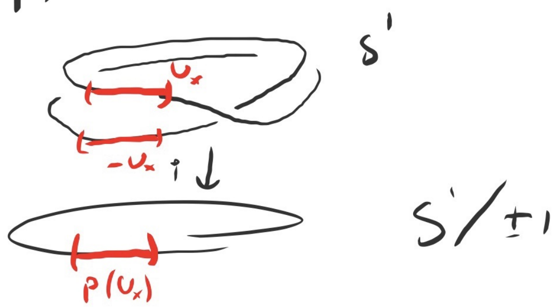
In the case of $X = \mathbb{R}$, such open sets were easily obtained — take $U = (x - 1/2, x + 1/2)$. This is basically what we did in the proof above, in 'Fact 1' of the technical lemma. Under the covering space assumption, each point in X/G has a neighborhood $[x] \in U$ so that there is a neighborhood $V \subset X/G$ upstairs with $x \in V$, $V \cap gV = \varnothing$ for g not the identity, and $p : gV \to U$ a homeomorphism for all $g \in G$, so that $p^{-1}(U) = \bigcup_{g \in G} gV \cong G \times U$, endowing G with the discrete topology. This is exactly the kind of thing going on in that first lemma above — as long as my image lands in U, I can lift to $p^{-1}(U)$, since the projection from each chunk of $p^{-1}(U)$ is a homeomorphism.



Covering quotients are like X is a "stack of records" above X

eg $S^1/\pm 1 \cong S^1$ is a covering quotient.

Picture



Remark 83. If G is finite, covering space actions are the same as *free actions*: actions such that $gx = x$ implies $g = e$. (No non-trivial element of g fixes a point.) Can you prove the equivalence?

Theorem 106. *Suppose X is a simply-connected topological space. If $G \curvearrowright X$ is a covering space action, then there is an isomorphism $\pi_1(X/G) \cong G$.*

I will not provide a proof here, though we will use this result freely. The idea is *very similar* to what we did for the circle above; in fact, the proof is almost unchanged.

Let me point out a quick corollary.

Proposition 107. *For $n > 1$, we have $\pi_1(\mathbb{RP}^n) \cong \mathbb{Z}/2$.*

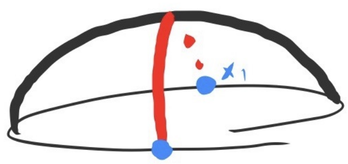
Proof. From a remark earlier (still unproved), we know (believe) that $\pi_1(S^n)$ is trivial — so S^n is simply connected.

We also have $\mathbb{RP}^n = S^n/(\pm 1)$, with the action being $(-1) \cdot x = -x$. I claim that this is a covering space action. In fact, if $x \in S^n$, then x lies in an open hemisphere U_x (top or bottom unless x is on the equator; in that case, front or back unless x is the east or west poles; in that case, left or right).

Then $-U_x$ is the opposite open hemisphere, and in particular $-U_x \cap U_x = \varnothing$. Since -1 is the only non-identity element of ± 1, it follows that this is a covering space action.

By the theorem above, $\pi_1(\mathbb{RP}^n) \cong \pm 1$, and this group is isomorphic to $\mathbb{Z}/2$; there is only one two-element group up to isomorphism. □

$$\pi_1(\mathbb{RP}^2) = \mathbb{Z}/2 \,?$$



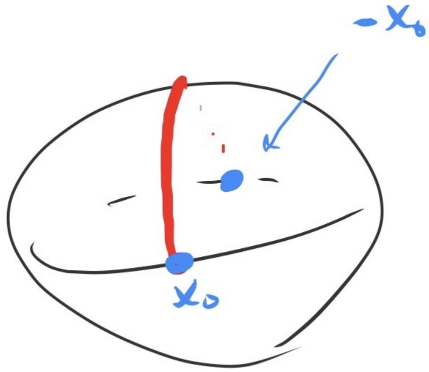
In $\mathbb{RP}^2 = D^2/(x \sim -x \text{ on } \partial D^2)$

$x_0 = x_1$

γ is a loop based at x_0 —

it represents $[1] \in \mathbb{Z}/2 = \pi_1(\mathbb{RP}^2)$

It lifts to a path not a loop in S^2



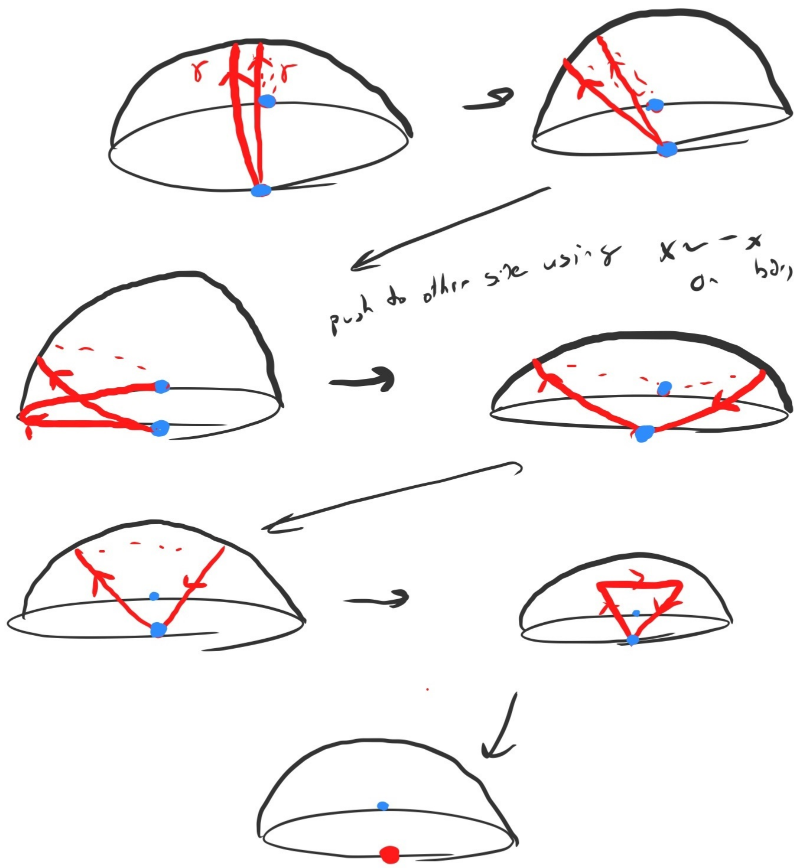
$\tilde{\gamma}(1) = [1] \cdot x_0 \quad (= -x_0)$

so "$wn(\gamma) = [1]$"

↖ measures $g \in G$ s.t.

$\tilde{\gamma}(1) = g \cdot x_0$

$\gamma * \gamma$ is null-htpic!



11/30: Topology of surfaces, and the boundary of a surface

What we have done thusfar in this course is develop a general notion of 'space where we can talk about continuous maps' and their properties. One of the basic questions we had about these spaces was: *can we determine whether or not two spaces are homeomorphic?*

This question is completely intractable for arbitrary spaces. But, in practice, nobody wants to know about a *completely arbitrary* space. Modern topology restricts its attention to special classes of spaces which are both geometrically/visually interesting, broad enough to provide interesting questions, and restrictive enough to enable the possibility of *classification.*

More precisely, 'geometric topology' is largely about the study of *manifolds*: spaces that locally look like Euclidean space.

Definition 68. *A naive manifold of dimension n is a topological space X so that for every point $x \in X$, there exists some open set $x \in U_x \subset X$ for which there is a homeomorphism $\varphi_x : U_x \to \mathbb{R}^n$.*

A naive manifold with boundary of dimension n is a topological space so that every point has a neighborhood U_x homeomorphic to $\mathbb{R}^n$ as above, or homeomorphic to the half-space

$$H^n = [0, \infty) \times \mathbb{R}^{n-1}.$$

You will notice the use of the word 'naive' above. The problem is that we want the notion of 'manifold' in low dimensions to coincide with our visual intuition: there are curves and surfaces in $\mathbb{R}^3$. However, the given definition allows both for *very large manifolds* (the 'long line', for instance) which are not even metrizable — $\mathbb{R}$ is not large enough to capture a good notion of distance! — and for non-Hausdorff objects like the line with two origins. (A naive manifold is still T_1, though. Do you see why?)

We correct this as follows.

Definition 69 (Topological manifolds)**.** *A topological manifold (resp. manifold with boundary) of dimension n is a naive manifold (resp. naive manifold with boundary) X which is furthermore second-countable (there exists a countable basis of open sets for the topology on X) and Hausdorff.*

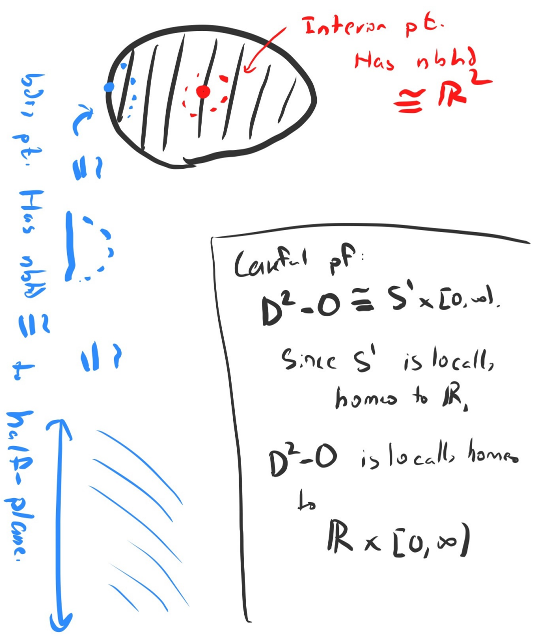
Remark 84. The second-countable condition is harmless in real practice. Every separable metric-space is second-countable; if $\{x_k\}$ is the countable dense subset, then $\{B_{1/n}(x_k)\}$ is a countable basis for the topology. Subspaces of second-countable spaces are second-countable. Quotients do not need to be in general, but if $p : X \to X/\sim$ is an *open map* and X is second-countable, then $X/\sim$ is as well. This is enough to deal with most examples we care about.

We say that a 2-dimensional manifold (possibly with boundary) is a *surface.*

Example 85. The real numbers $\mathbb{R}^n$ are the basic n-manifold. The sphere S^n is an n-manifold for all n; any hemisphere is homeomorphic to $\mathbb{R}^n$, and S^n is a subspace of the Hausdorff and second-countable space $\mathbb{R}^{n+1}$, so is itself Hausdorff and second-countable.

Quotients of manifolds by *covering space actions* are again manifolds (so long as the quotient is Hausdorff), as the quotients are locally homeomorphic to the original space and the projection map is open (so that the quotient is second-countable). In particular, $\mathbb{RP}^n$ is an n-manifold for all n. The Klein bottle K is a surface. The Mobius band is a surface with boundary.

The disc D^n is an example of an n-manifold with boundary.



Products of manifolds are again manifolds, so the torus T^2 and in general T^n are manifolds.

Our goal for the end of this class will be to understand as much as we can in low dimensions. Let me point out a fact; your final homework has extra credit for proving a special case:

Theorem 108. *Every 1-dimensional manifold (possibly with boundary) is homeomorphic to exactly one of* $\mathbb{R}, [0, \infty), [0, 1]$, *or* S^1.

This is the first local-to-global theorem you have seen in topology. A statement about *local* properties about these spaces — that they locally look like either $\mathbb{R}$ or $[0, \infty)$ (with some simple 'global' requirements about Hausdorffness and the existence of a countable basis) — is enough for us to conclude something about their *global* structure.

The statement for surfaces is harder and more exciting, to my eye.

Theorem 109. *There is a classification of compact surfaces — an explicit list of countably many compact surfaces so that every compact surface is homeomorphic to exactly one on the list.*

On Thursday we will have enough technology to write down the list.

Topology of surfaces

The first thing we want to do is focus on connected spaces. The following shows that the disconnected case is reducible to the connected case:

Proposition 110. *If S is a topological surface, it has at most countably many path-components, and S is homeomorphic to the disjoint union of its path-components.*

Proof. First, observe that each point in S has an open set around it homeomorphic to either $\mathbb{R}^2$ or H^2. It follows that the path-components in S are open sets. We have seen on a previous homework that a space whose path-components are open is the disjoint union of its path-components, so we are reduced to showing that S has at most countably many.

Now we remark that if a space X has a basis $\mathcal{B}$, and an open cover of pairwise disjoint nonempty open sets $\{U_i\}_{i \in I}$, we may find a subject $B' \subset \mathcal{B}$ and a surjection $B' \to I$, so that $|I| \leqslant |\mathcal{B}|$. In particular, when

X is second-countable, there can be no more than countably many nonempty pairwise disjoint open sets.

To see this, let $B' \subset \mathcal{B}$ be the subset of nonempty basis elements $V \in \mathcal{B}$ so that $V \subset U_i$ for some i. I claim that the map $f : B' \to I$, defined by

$$f(V) = i \text{ such that} V \subset U_i\},$$

is well-defined and surjective.

To see that f is well-defined, observe that since $U_i \cap U_j = \varnothing$ for all i, j, it follows that V cannot simultaneously be contained in both of them unless V is empty. Because elements of B' are by assumption nonempty and contained in some U_i, it follows that $V \subset U_i$ for a *unique* i, as desired.

Surjectivity is clear: each U_i is open, so by definition for each $x \in U_i$ is some basis element $x \in V \subset U_i$. By definition, $V \in B'$, and $f(V) = i$. Since i was arbitrary, it follows that f is surjective.

It follows that there are at most as many path-components as there are basis elements. By assumption, S is second-countable, so it has at most countably many path-components. □

From now on we essentially always work with connected surfaces.

Intuitively, a surface should have some number of boundary curves. But this is **not obvious at all!**. It requires proof.

Definition 70. *Let M be an n-dimensional topological manifold. We say that $x \in M$ is a boundary point if there exists an open set $x \in U_x$ and a homeomorphism $\varphi : U_x \to [0, \infty) \times \mathbb{R}^{n-1}$ so that $\varphi(x) \in \{0\} \times \mathbb{R}^{n-1}$.*

We say that $x \in M$ is an interior point if there exists an open set $x \in U_x$ and a homeomorphism $\varphi : U_x \to \mathbb{R}^n$.

We write ∂M for the set of boundary points and M° for the set of interior points.

Note that these notions of boundary and interior do not coincide with the usual notions in general topology. If M is a topological space, then the interior of M in the sense of general topology is M itself; its boundary is empty. **I hope this will not cause confusion; we will not use the general topology notions here.** If you feel a need to refer to the notions from general topology, I suggest writing them in text, as $\text{int}(M)$ and $\text{bd}(M)$ instead of the symbols above, reserved for the manifold notions.

A priori, it could be the case that $\mathbb{R}^n$ is homeomorphic to $[0, \infty) \times \mathbb{R}^{n-1}$ — this would mean that the boundary and interior *coincide*. We do not have the technology to deal with this possibility. However, we can deal with this for 1-manifolds and surfaces. In these notes we will investigate these two sets ∂S and S° for a surface S.

Proposition 111. *If S is a topological surface (possibly with boundary), then*

$$\partial S \cup S^\circ = S, \qquad \partial S \cap S^\circ = \varnothing.$$

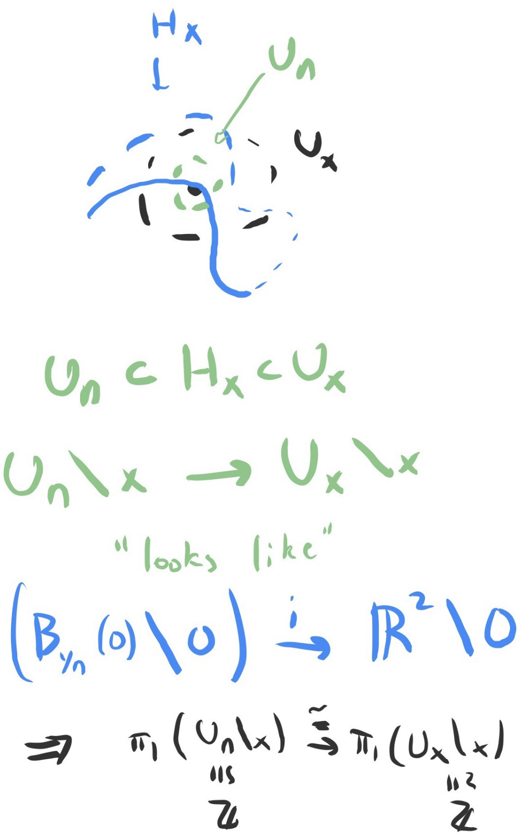
The set ∂S is a closed subset of S, and the set S° is an open subset. The boundary ∂S is a 1-dimensional topological manifold.

It is clear from the definition of topological surface that $S = \partial S \cup S^\circ$. The non-obvious thing is that a boundary point can be distinguished from an interior point.

Proof. Suppose, towards a contradiction, that $x \in \partial S$ and $x \in S^\circ$. Then there are open sets H_x, U_x with homeomorphisms $\varphi : U_x \to \mathbb{R}^n$ and $\psi : H_x \to H^2$.

I claim that $U_n = \varphi^{-1}(B_{1/n}(0))$ is contained in H_x for some sufficiently large n. Observe that $H_x \cap U_x \subset U_x$ is an open subset of U_x containing x, and φ is a homeomorphism, so that $\varphi(H_x \cap U_x) \subset \mathbb{R}^n$ is an open subset containing 0; it follows that $B_{1/n}(0) \subset \varphi(H_x \cap U_x)$ for some $n > 0$, and thus that we can arrange for $U_n \subset H_x$.

This gets for us a sequence of inclusions $x \in U_n \hookrightarrow H_x \hookrightarrow U_x$, where the first and third terms are homeomorphic to the half-plane, while the second and last terms are homeomorphic to the plane.



Consider the composite

$$\pi_1(H_m\backslash\{x\}) \xrightarrow{i_1} \pi_1(U_n\backslash\{x\}) \xrightarrow{i_2} \pi_1(H_x\backslash\{x\}) \xrightarrow{i_3} \pi_1(U_x\backslash\{x\}).$$

It is straightforward to see that the induced map

$$\pi_1(\mathbb{B}_{1/n}(0)\backslash\{0\}) \to \pi_1(\mathbb{R}^2\backslash\{0\})$$

is an isomorphism $\mathbb{Z} \to \mathbb{Z}$ (both deformation retract to the circle of radius $1/2n$).

This leads us to a contradiction. On the one hand, the map $i_3 i_2$ above should be an isomorphism. On the other hand, the maps i_2 and i_3 are zero, as their codomain (resp. domain) are the trivial group, which would imply $i_3 i_2 = 0$. This contradicts our original assumption that $S^\circ \cap \partial S \neq \varnothing$; so these two sets are disjoint.

The rest follows straightforwardly. Every point is either an interior point or a boundary point (but not both), so that $S^\circ = (\partial S)^c$. The interior is open because, if $x \in S^\circ$, there is an open set U_x and a homeomorphism $\varphi : U_x \to \mathbb{R}^2$, so that $U_x \subset S^\circ$ as well; so the interior is locally open, hence open. It follows that the boundary is closed.

To see that the boundary is a 1-dimensional manifold, note that it is Hausdorff as it is a subspace of a Hausdorff space, and second-countable because if $\mathcal{B}$ is a countable basis of S, then $\mathcal{B}' = \{U \cap \partial S \mid U \in \mathcal{B}\}$ is a countable basis of ∂S.

Then note that if $x \in \partial S$, there is an open set $x \in U_x \subset S$ homeomorphism $\varphi : U_x \to H^2$ with $\varphi(x) \in \{0\} \times \mathbb{R}$. Then $y \in U_x$ is also a boundary point if and only if $\varphi(y) \in \{0\} \times \mathbb{R}$ (the other points are interior points, as they have neighborhoods homeomorphic to $\mathbb{R}^2$). Therefore $\varphi^{-1}(\{0\} \times \mathbb{R}) = U_x \cap \partial S$ is an open subset of ∂S consisting entirely of boundary points, which is homeomorphic to $\mathbb{R}$. Since we can do this for an arbtirary $x \in \partial S$, it follows that ∂S is locally Euclidean, and thus a topological 1-manifold. Notice that ∂S has no boundary of its own. ☐

Corollary 112. *If S is a compact surface (possibly with boundary), then ∂S is a disjoint union of circles. If $f : S \to S'$ is a homeomorphism, then it restricts to a homeomorphism $\partial S \to \partial S'$. It follows that the number of boundary components — $|\pi_0 \partial S|$ — is topological property of compact surfaces.*

Proof. The property of being a boundary point is a topological property; if $f : S \to S'$ is a homeomorphism, then $f(\partial S) \subset \partial S'$, as if $x \in \partial S'$ and $U_{f^{-1}(x)} \xrightarrow{\varphi} [0, \infty) \times \mathbb{R}$ is a homeomorphism, then

$$f(U_{f^{-1}(x)}) \xrightarrow{\varphi f^{-1}} [0, \infty) \times \mathbb{R}$$

is a homeomorphism, so that if $f^{-1}(x)$ is a boundary point, so is x, justifying the stated inclusion.

It follows that f restricts to a continuous bijection $f : \partial S \to \partial S'$. The same argument applies to f^{-1}, and thus f and f^{-1} also give inverse homeomorphisms between ∂S and $\partial S'$, as desired. We have not yet used compactness.

Because S is assumed compact, ∂S is a compact 1-manifold without boundary. By the classification of 1-manifolds, it follows that each connected component of ∂S is a circle, and hence (because ∂S is the disjoint union of its path components) that ∂S is homeomorphic to the disjoint union of circles. f Finally, because ∂S is compact and its path-components open, it follows that it has at most finitely many path-components (otherwise this would give an infinite cover of ∂S without a finite subcover), so the boundary is the disjoint union of finitely many circles. ☐

12/2: Operations on surfaces

The goal today is two-fold:

- We want to be able to relate surfaces with boundary to surfaces without boundary. The latter seems simpler, so this will reduce the work we have to do for the general classification theorem (maybe!)

- We want to be able to build more complicated surfaces out of simple ones.

To do these, we develop three operations on surfaces: *puncturing*, *capping*, and *connected-sum*.

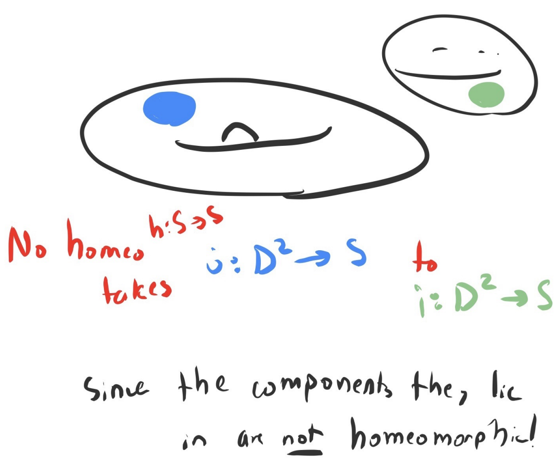
The first operation — *puncturing* — adds boundary components.

Definition 71. *Let $f : D^2 \to S$ be a topological embedding of the closed disc into S so that $f(D^2) \subset S^\circ$. Write D° as an abbreviateion for $(D^2)^\circ = \{(x, y) \mid x^2 + y^2 < 1\}$.*
We define $P_f(S) = S \backslash (D^2)^\circ$.

This is not really an operation on S itself — it requires extra data. However (nontrivial!), so long as S is connected, the resulting topological space is independent of the choice of f (up to homeomorphism).

Theorem 113 (The disc-moving theorem)**.** *Suppose S is a connected surface, and $f_1, f_2 : D^2 \to S$ are topological embeddings of the closed 2-dimensional disc with $f_i(D^2) \subset S^\circ$.*
Then there is a homeomorphism $h : S \to S$ so that $(hf_1)(D^2) = f_2(D^2)$. (Note that we do not conclude that $hf_1 = f_2$, just that the images are the same.)

Why is this not true for disconnected surfaces?



Corollary 114. *If S is a connected surface, the topological space $P_f(S)$ is well-defined up to homeomorphism; we therefore drop f from the notation and write $P(S)$. It is again a compact surface, which has exactly one more boundary component than S.*
We write $P_n(S) = P(P_{n-1}(S))$ recursively; $P_n(S)$ has exactly n more boundary components than S does.

This is our first basic operation: *puncturing*. It increases the number of boundary components.

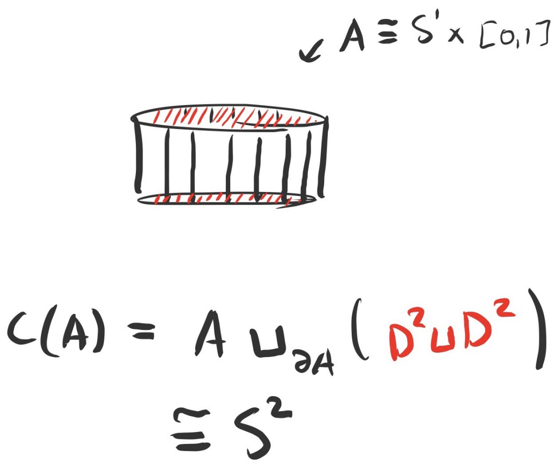
Example 86. S2. Puncturing once, twice, three times. Puncturing the torus and RP2 once.

Definition 72. *Let S be a compact surface with boundary with n boundary components. Choose a homeomorphism $g : \sqcup_{i=1}^n S^1 \to \partial S$. Then we write*

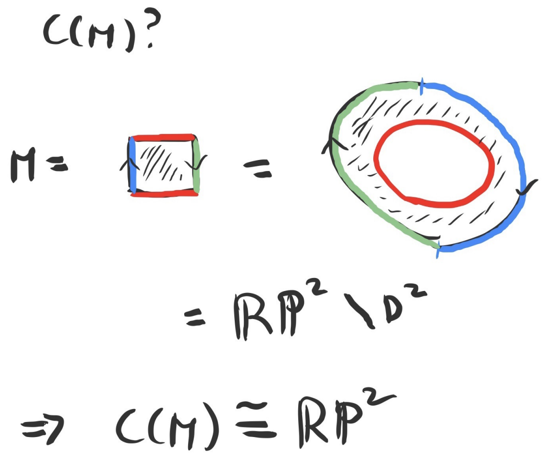
$$C_g(S) = S \sqcup_{i=1}^n D^2 / \big(x \in \sqcup_{i=1}^n S^1 \sim g(x) \in \partial S\big).$$

Example 87. We have that $C(D^2) \cong S^2$ (we are adding the upper hemisphere of the sphere back — this is what inspired the term 'cap').

Similarly, we have that $C(A) \cong S^2$, where A is the annulus.

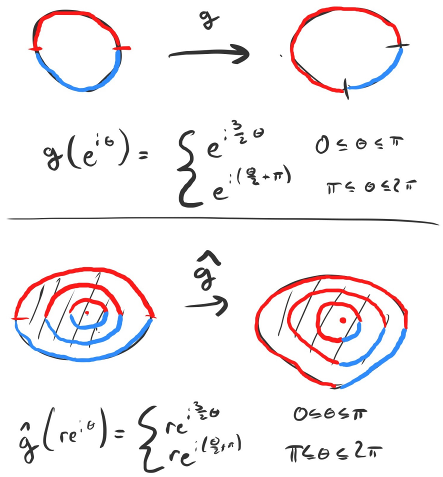


For a non-orientable example, if M is the Mobius band, then $C(M) \cong \mathbb{RP}^2$. It is easiest to see this by identifying M with the complement of a disc in $\mathbb{RP}^2$ to begin with.



Proposition 115 (Alexander trick, first version)**.** *Any homeomorphism $g : S^1 \to S^1$ extends to a homeomorphism $\hat{g} : D^2 \to D^2$.*

Proof. Define $\hat{g}(re^{i\theta}) = rg(e^{i\theta})$. One can either check continuity by hand (using the epsilon-delta definition) or observe that this is the quotient of a homeomorphism $S^1 \times [0,1] \to S^1 \times [0,1]$, given by $(e^{i\theta}, r) \mapsto (g(e^{i\theta}), r)$. It is clearly a bijection from a compact Hausdorff space to itself, so (because it is continuous) necessarily a homeomorphism.



□

Corollary 116. *The surface $C_g(S)$ is well-defined up to homeomorphism: for any two choices g, g' of parameterizations of the boundary circles, the resulting surfaces $C_g(S)$ and $C_{g'}(S)$ are homeomorphic. The result is a compact surface without boundary.*

These two operations allow us to relatively freely pass between *compact surfaces with boundary* and *compact surfaces without boundary.*

Corollary 117. *Write $\mathcal{S}_n$ for the set of compact connected surfaces with exactly n boundary components, considered up to homeomorphism.*

Then $P_n : \mathcal{S}_0 \to \mathcal{S}_n$ and $C : \mathcal{S}_n \to \mathcal{S}_0$ are well-defined functions which are inverse to each other. In particular, these two sets are in bijection; two compact surfaces S, S' are homeomorphic iff they have the same number of boundary components and $C(S) \cong C(S')$.

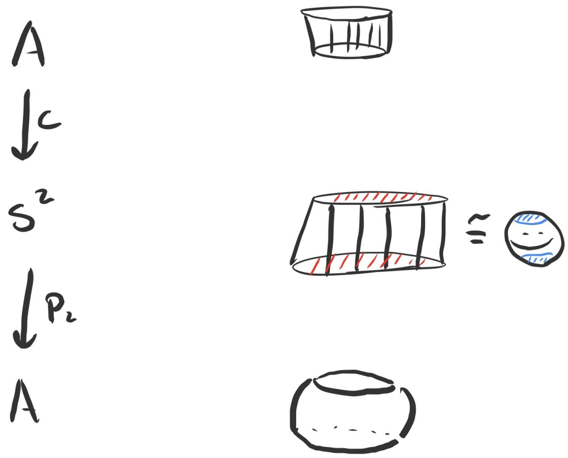
Proof. Let us show that $(CP_n)(S) \cong S$; the other direction is similar.

Choose an embedding $i : \sqcup_{j=1}^n D^2 \to S^\circ$, so that $P_n(S) \cong S \backslash i(\sqcup (D^2)^\circ)$ — that is, $P_n(S)$ is obtained by deleting the interiors of these discs. The restriction of i to the boundary circles gives us a parameterization $i : \sqcup_{j=1}^n S^1 \to \partial P_n(S)$ along which we can define

$$C(P_n(S)) = P_n(S) \sqcup_{j=1}^n D^2 / \big(x \in \sqcup_{j=1}^n S^1 \sim i(x)\big).$$

There is a natural map $f : C(P_n(S)) \to S$ — on the $P_n(S)$ factor, this map is just the identity, and on the disc-factor, $f(x) = i(x)$. To descend to the quotient $C(P_n(S))$ this should respect the equivalence relation $x \in \sqcup_i S^1 \sim i(x) \in P_n(S)$ — but note that if $i(x) \in P_n(S)$, we have $f(i(x)) = i(x)$, while if x is the corresponding point on $\sqcup_i S^1$, then $f(i(x)) = i(x)$ by definition — so this does respect the equivalence relation.

The map f is a continuous bijection from a compact space to a Hausdorff space, so a homeomorphism.

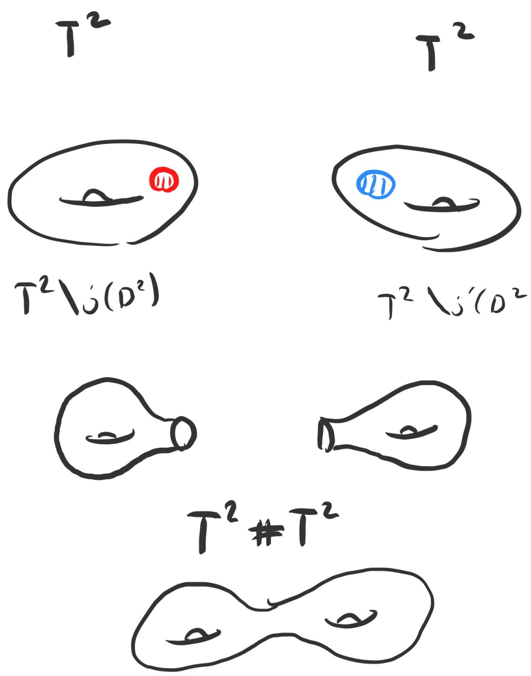


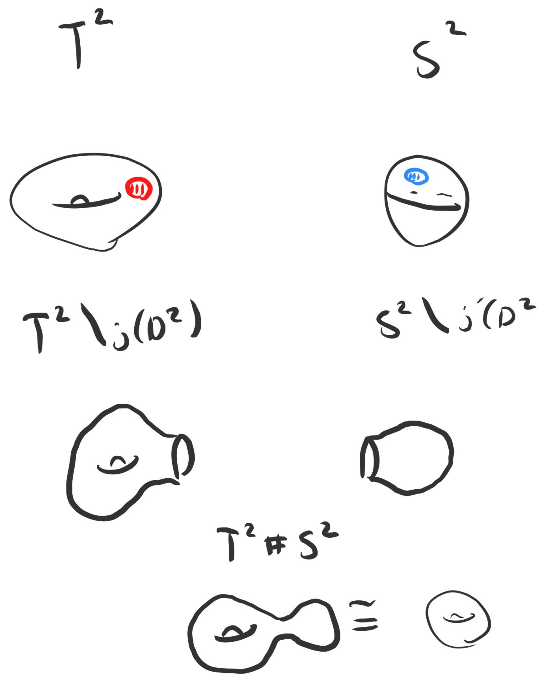
□

The one remaining tool we need is a notion of building more complicated surfaces out of simpler ones.

Definition 73. *Let S and S' be connected surfaces. We say that the connected sum is obtained by choosing a pair of embeddings $j : D^2 \to S^\circ$ and $j' : D^2 \to (S')^\circ$, deleting the interiors, and pasting along the boundary circles:*

$$S\#S' = (S\backslash j((D^2)^\circ)) \sqcup (S\backslash j'((D^2)^\circ))/\big(j(x) \sim j'(x) \mid x \in S^1\big).$$





The idea is relatively simple. Pasting two surfaces together along part of their boundary should again get us a surface; to combine two surfaces, we forcibly add boundary, which we then get rid of by pasting.

Example 88. The connected sum of any surface S with the sphere S^2 is just homeomorphic to S. First, you delete a disc in both S and S^2, and paste the resulting surfaces along their boundary. But S^2 with a disc deleted is just a disc again, so this can be restated as: 'delete a disc from S, then reglue a disc in along its boundary'. In the end, nothing has changed, so $S\#S^2 \cong S$.

The annulus is the connected sum $D^2\#D^2$. Notice that deleting a disc gives us a pair of annuli which we want to paste along one boundary curve, and the result of that is again an annulus; after all,

$$\{(x,y) \mid 1/4 \leqslant x^2+y^2 \leqslant 4\} = \{(x,y) \mid 1 \leqslant x^2+y^2 \leqslant 4\} \sqcup_{S^1} \{(x,y) \mid 1/4 \leqslant x^2+y^2 \leqslant 1\}.$$

In fact, in general, $S\#D^2 \cong P(S)$ — connect-summing with a disc is the same as *deleting* a disc. This follows from the fact that $S\#S^2 \cong S$; deleting a disc from the second factor gives us $S\#D^2$, but deleting a disc from S gives us $P(S)$, so that $S\#D^2 \cong P(S)$.

We will black-box the following. One needs to be somewhat more careful in the definition past the 2-dimensional setting; it is always well-defined but one has to pay close attention to orientations in general.

The proof is similar to the facts used above — first you use that there is a homeomorphism taking any disc to another to show that this was independent of the choice of disc, then you show that it's independent of the parameterization of the boundary discs.

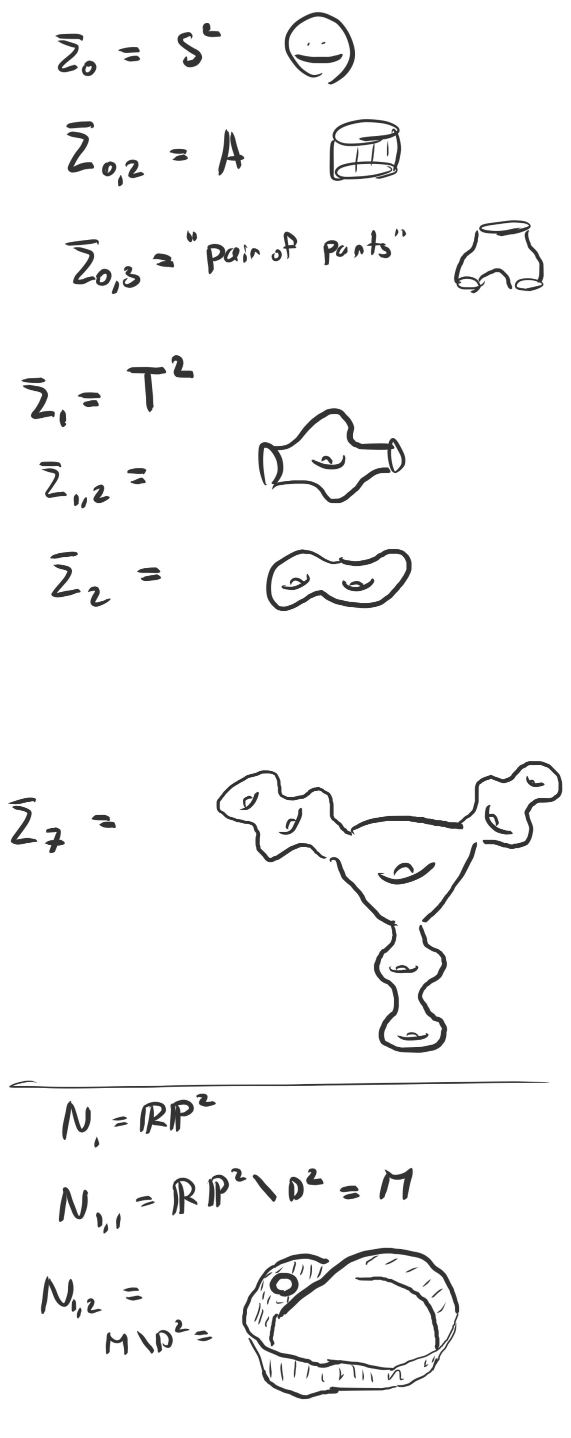
Proposition 118. *The connected sum is well-defined: if S and S' are connected surfaces, then the result $S\#S'$ of the above procedure are homeomorphic, and hence the homeomorphism type is independent of the choices made above.*

Definition 74. *We write Σ_g recursively as $\Sigma_g = \Sigma_{g-1}\#T^2$; that is, $\Sigma_g = \#^g T^2$. This is called the surface of genus g.*

Similarly, we write $N_h = N_{h-1}\#\mathbb{RP}^2$, or $\#^h\mathbb{RP}^2$. There is not a uniformly used name for these, but a common one is the non-orientable surface with unoriented genus h.

We write $\Sigma_{g,n} = P_n(\Sigma_g)$ for the surface of genus g with n boundary components.

Similarly, we write $N_{h,m} = P_m(N_h)$ for the non-orientable surface of unoriented genus h and m boundary components.



We can finally state the 'big theorem'.

Theorem 119 (Classification of compact surfaces)**.** *Let S be a compact, connected surface. Then S is* *homeomorphic to exactly one of the surfaces*

$$\{\Sigma_{g,n}\}_{g\geqslant 0, n\geqslant 0} \cup \{N_{h,m}\}_{h\geqslant 1, m\geqslant 0}.$$

That is, the surfaces listed in the previous definition are pairwise non-homeomorphic, and every compact surface is homeomorphic to one of them.

The disconnected case follows from this by Proposition 3 above — a surface is homeomorphic to the disjoint union of its path-components, so every compact surface is homeomorphic to the (finite) disjoint union of some on the list above.

Example 89. The Klein bottle, K, is homeomorphic to $\mathbb{RP}^2\#\mathbb{RP}^2 = N_2$.