§17 - Connected Spaces

1 Motivation

Getting from one place to another is important. Connectedness is the notion that a space "is in one piece". It is slightly surprising that connectedness is a property best described in topological language, but it is. Unsurprisingly it is easier to describe what it means for a space to be "in two pieces". We make this precise, then we refine it a bit. We will see that connectedness is related to some theorems of calculus and some very deep theorems in mathematics.

In common language, you would say that your kitchen is "connected to" your living room if you can walk from your kitchen to your living room without leaving your house. This gives rise to the notion of "path-connectedness" which is a very natural form of connectedness. We will investigate some examples of spaces that challenge our intuitive ideas of connectedness.

Finally, we will look at local properties of spaces. Previously, we have looked at separability and second countability which are global properties (because they say something about the whole space), but now we will look at properties which are only true in small open sets around points. (Of course we will make this precise!)

2 The Definition

Definition. *A topological space $(X, \mathcal{T})$ is said to be* ***disconnected*** *if there are non-empty $A, B \subseteq X$ such that $A \cup B = X$, $A \cap \overline{B} = \emptyset$ and $\overline{A} \cap B = \emptyset$. Such a pair (A, B) is called a* ***separation of*** *X. A space that is not disconnected is called* ***connected****.*

Notation: Recall that $X = A \sqcup B$ means that $X = A \cup B$ *and* $A \cap B = \emptyset$. This will save us some words later on.

Obviously, we see that $X_{\text{indiscrete}}$ is always connected, but X_{discrete} is disconnected (provided that X has at least two points). Before we see more examples, let us restate the definition of connectedness in ways that will be useful at various times.

Theorem. *Let $(X, \mathcal{T})$ be a topological space. TFAE:*

1. *X is disconnected;*
2. *X has a non-trivial clopen subset (i.e. other than $\emptyset$ and X);*
3. *X can be written as the disjoint union of two non-empty open sets;*

4. X can be written as the disjoint union of two non-empty closed sets.

Proof. These are all just unwinding definitions, but let us prove that [1] implies all the others.

[$1 \Rightarrow 2, 3, 4$] Suppose that $X = A \sqcup B$ such that A, B are non-empty, $A \cap \overline{B} = \emptyset$ and $\overline{A} \cap B = \emptyset$. This tells us that $\overline{B} \subseteq B$ and $\overline{A} \subseteq A$, which means that both A and B are closed. Since $X \setminus A = B$ we see that B is open, and since $X \setminus B = A$ we see that A is open. This tells us that X has a non-trivial clopen subset (A or B) and X can be written as the union of two disjoint, non-empty open (and closed!) sets A and B.

□

Another Boring Exercise: Complete the above proof by checking that the other conditions above are really equivalent to being disconnected.

Some Examples:

- $\mathbb{R}_{\text{usual}}$ is connected. We saw this on Assignment 5, A.1.
- $(0,1), [0,1)$ and $(0,1]$, each with the usual topology, are connected also by Assignment 5, A.1. (If we are being careful we should check that these intervals really are Dedekind complete, but this is straightforward if we assume that $\mathbb{R}$ is Dedekind complete.)
- S^1, the circle; T^1, the torus; $B_\epsilon(0) \subseteq \mathbb{R}^n$ are all connected, but these will require arguments.
- $\mathbb{R}_{\text{co-finite}}, \mathbb{R}_{\text{co-countable}}$ and $\mathbb{R}_{\text{indiscrete}}$ are all very connected; they don't have any disjoint non-empty open sets!

Some Non-Examples:

- $\mathbb{R} \setminus \{0\}$ with the usual topology is disconnected as $\mathbb{R} \setminus \{0\} = (-\infty, 0) \sqcup (0, +\infty)$, which is the disjoint union of two open sets.
- $\mathbb{R}_{\text{Sorgenfrey}}$ is disconnected, and is in fact very disconnected. Every non-empty basic open set in the Sorgenfrey line is actually a non-trivial clopen subset.
- $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$, with the usual subspace topology, is disconnected. For example $\{\frac{1}{2}\}$ is a non-trivial clopen subset of the space. For basically the same reason, $\omega + 1$ and ω_1 and $\omega_1 + 1$ are not connected.

3 Tools and Proofs

Now we will simultaneously give a tool for showing that certain spaces are connected and that connectedness is a topological invariant.

Proposition. *If $f : X \longrightarrow Y$ is a continuous surjection and X is connected, then Y is connected.*

Proof. We prove the contrapositive. Suppose that Y is not connected. Let $C \subseteq Y$ be a non-trivial clopen subset of Y, and let $A := f^{-1}[C]$ which is a clopen subset of X since f is continuous. Moreover, we see that since $Y \neq C$ we have a $y \in Y \setminus C$, so $f^{-1}(y) \notin f^{-1}[C] = A$. Thus A is a *non-trivial* clopen subset of X and so X is not connected. □

Corollary. *Connectedness is a topological invariant.*

The thing to notice here isn't just that this corollary says something about connectedness and homeomorphisms, but the proposition gives us a useful way for showing that certain spaces are connected. To show that a space Y is connected it is enough to find a continuous surjection $f : X \longrightarrow Y$ where X is connected. Similarly, if we want to show that X is disconnected then it is enough to find such a continuous surjection with Y disconnected.

Proposition 1. *S^1 is connected.*

Proof. We have already established that $[0, 1]$ is connected, so it is enough to point out that the map $f : [0, 1] \longrightarrow S^1$ given by $f(x) = (\cos(2\pi x), \sin(2\pi x))$ is a continuous surjection.

In some ways, the more natural thing to do is to think about $S^1 \subseteq \mathbb{C}$, the complex plane, and to take the (continuous) map that sends $\theta \in [0, 1]$ to $e^{2\pi i\theta}$. □

Let us look at a similar argument that shows that a space is disconnected.

Proposition 2. *$\mathbb{R} \setminus \{0\}$ with the usual subspace topology is disconnected.*

Proof. Yes we have already seen this, but let us look at a different reason. Consider the map $f : \mathbb{R} \setminus \{0\} \longrightarrow \{-1, 1\}$ given by $f(x) := \frac{x}{|x|}$, where $\{-1, 1\}$ is given the discrete topology and so is not connected.

Now it is easy to see that f is continuous and a surjection, hence $\mathbb{R} \setminus \{0\}$ cannot be connected. □

Punctured Exercise: Use the proof of the previous argument to show that $\mathbb{R}^2 \setminus \{(0, 0)\}$, the punctured plane, is a connected space.

4 Subspaces and Products

As always we should discuss how connectedness interacts with subspaces and products. It should already be clear that connectedness is not a hereditary property.

Proposition 3. *Connectedness is not a hereditary property. For example,* $[0,1]$ *is connected, but* $\{0,1\}$ *is not.*

In fact, on Assignment 4, A.2 we investigated the notion of cut-points, which shows how connectedness is definitely not hereditary.

Fill-in Exercise: Go through Assignment 4, A.2 again and verify all of the assumptions we made (e.g. $\mathbb{R}^2$ is connected).

What about products?

Proposition. *If* X, Y *are connected spaces then* $X \times Y$ *is connected.*

Proof. This is left as an exercise. The proof is a direct proof (i.e. assume X and Y are both connected and show that $X \times Y$ is connected). The proof uses the fact that $X \times \{y\} \cong X$ for any point $y \in Y$. $\square$

Arbitrary Exercise: Does your argument of the previous proposition extend to arbitrary products?

Taking stock of what we've done we now additionally know that the following spaces are connected:

1. $T^1 = S^1 \times S^1$, the torus, is connected as it is the product of two connected spaces.
2. $B_\epsilon(0) \subseteq \mathbb{R}^n$ is connected since it is homeomorphic to $(0,1)^n$ which is connected.

5 Cool Applications

Connectedness is instrumental in many applications in mathematics, and here are two such examples:

Theorem (The Intermediate Value Theorem)**.** *Let* $I \subseteq \mathbb{R}$ *be a connected space and let* $f : I \longrightarrow \mathbb{R}$ *be continuous. If* $a, b \in X$ *are such that* $f(a) < f(b)$ *then for every* $c \in (f(a), f(b))$ *there is a* $x_0 \in I$ *such that* $f(x) = c$*.*

Proof. The proof is by contrapositive. Let $c \in (f(a), f(b))$ be such that $c \notin f[I]$. Then we see that

$$f[I] \subseteq (-\infty, c) \sqcup (c, +\infty)$$

So in the image, $(-\infty, c)$ and $(c, +\infty)$ are both clopen (and non-empty since they contain $f(a)$ and $f(b)$ respectively).

Since f is continuous we have $a \in f^{-1}(\infty, c)$ a clopen subset of I, and we have $b \in f^{-1}(c, +\infty)$ another non-empty clopen subset of I. Moreover we see that

$$I = f^{-1}(\infty, c) \sqcup f^{-1}(c, +\infty)$$

Hence I is not connected. $\square$

Go Further Exercise: We can see that this is a slight generalization of the IVT that you saw in first year calculus, but can you do even better? Can this theorem be generalized to arbitrary connected spaces? Does that even make sense? What is the right amount of structure you need the domain to have so that a version of the IVT is still true?

Let us also state the first year calculus version of the intermediate value theorem. This one will be true since we have already established that all intervals in $\mathbb{R}$ are connected.

Theorem (IVT, first year calculus style)**.** *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \longrightarrow \mathbb{R}$ be continuous. If $a, b \in X$ are such that $f(a) < f(b)$ then for every $c \in (f(a), f(b))$ there is a $x_0 \in I$ such that $f(x) = c$.*

Now we give a theorem that has many brothers and sisters in mathematics. Fixed point theorems are quite interesting.

Theorem. *Let $f : [0,1] \longrightarrow [0,1]$ be a continuous map. There is a $c \in [0,1]$ such that $f(c) = c$, which is called a fixed point of f.*

Proof. Consider the function $g : [0,1] \longrightarrow [-1,1]$ which is defined by $g(x) = f(x) - x$, which is continuous since f is continuous and the difference of two continuous functions is again continuous, (although you should convince yourself that the range of g is appropriate). Now if $g(c) = 0$ for some $c \in [0,1]$ then we see that $0 = g(c) = f(c) - c$ hence $c = f(c)$ as desired.

Assume that $g(x) \neq 0$ for all $x = 0, 1$ (because otherwise we are finished). Thus $g(0) = f(0) > 0$ and $g(1) = f(1) - 1 < 0$. Since g is continuous, and $[0,1]$ is connected, by the Intermediate Value Theorem there is a $c \in [0,1]$ such that $g(c) = 0$. $\square$

The study of fixed points is quite interesting and is a theme of many branches of topology. Let's leave with another fixed point theorem that is called Brower's Fixed Point Theorem and has some neat corollaries!

Theorem (Brower's Fixed Point Theorem)**.** *Any continuous map from the closed unit disc in $\mathbb{R}^2$ to the closed unit disc in $\mathbb{R}^2$ has a fixed point.*

Here is a taste of the consequences of Brower's Fixed Point Theorem and its siblings. You can look up their statements on Wikipedia if you would like more information.

1. The fundamental theorem of algebra;

2. The existence of the Nash equilibrium;

3. Invariance of Domain (and as a corollary, $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}^m$ for any $n \neq m$).

Now, Connectedness can also be used in a rather clever way to show that some property is "true of the whole space". For example, if we have a connected space X and some property ϕ that is true of elements of X and then we manage to show that the collection $B \subseteq X$ of points with property ϕ is (1) closed and open, and (2) non-empty, then we can conclude that $B = X$. Let's use that idea to show a fact from Complex Analysis. Here it is the proof technique that we are interested in, not the concepts from Complex Analysis.

Definition. *Let $D \subseteq \mathcal{C}$ be connected with non-empty interior. A function $f : D \longrightarrow \mathbb{R}$ is said to be* ***analytic*** *if for every point $p \in D$ there is an open ball $B_\epsilon(p) \subseteq D$ such that f has a power series representation (with real coefficients) that is valid on $B_\epsilon(p)$.*

Proposition. *Let $f : D \longrightarrow \mathbb{R}$ be an analytic function that is zero on a set $A \subseteq D$ that has a limit point in A, then f is the constant zero function.*

Proof. This proof is adapted from Gamelin's "Complex Analysis", p.156. Again, let me stress that the technique involved is the most interesting part for us; focus on the form and the first two parts of the proof. Let

$$B := \{z \in D : f^{(m)}(z) = 0, \forall m \geq 0\}$$

and we will show that $B = D$ by showing that B is a non-empty clopen subset of D. From there we conclude that the power series representation of f at any point has only 0 coeffiecients, hence f is the constant zero function.

[Open] Let $p \in B$. Since f is analytic, f has a power series representation $f(z) = \sum_{n\geq 0} a_n(z-p)^n$ which is valid in some $B_\epsilon(p) \subseteq Z$. Since $p \in B$, we see that the coefficients

$$a_n = \frac{f^{(n)}(p)}{n!} = 0, \forall n \in \mathbb{N}$$

Thus $B_\epsilon(p) \subseteq B$ so B is open.

[Closed] Let $p \in D \setminus B$. So then $f^{(m)}(p) \neq 0$ for some m. Since f is analytic, f has a power series representation $f(z) = \sum_{n \geq 0} a_n (z-p)^n$ which is valid in some $B_\epsilon(p) \subseteq Z$. Thus we see that

$$a_m = \frac{f^{(m)}(p)}{m!} \neq 0$$

so $B_\epsilon(p) \subseteq D \setminus B$. Hence $D \setminus B$ is open and B is closed.

[Non-empty] This part really belongs to analysis, so don't worry too much about it, but here it is for completeness. Let $\langle a_n \rangle$ be a (non-trivial) sequence in A that converges to $a \in A$ such that $f(a_n) = 0 = f(a)$ for all $n \in \mathbb{N}$. Since f is analytic, f has a power series representation

$$f(z) = \sum_{n \geq 0} a_n (z-a)^n$$

which is valid in some $B_\epsilon(a) \subseteq Z$. Suppose for the sake of contradiction that $f^{(m)}(a) \neq 0$ for some m. Then observe that

$$f(z) = (z-a)^m h(z)$$

for some $h(z)$ analytic and $h(a) \neq 0$. (This is simply writing out the power series of f!) But now, for any $a_n \in B_\epsilon(a)$ we see that

$$0 = f(a_n) = (a_n - a)^m h(a_n),$$

so $h(a_n) = 0$. By continuity of h we get that $h(a) = 0$ a contradiction. □

I know that this isn't a Complex Analysis course, but I can't resist stating this great corrolary, which is proved by applying the previous proposition to the function $h(z) := f(z) - g(z)$.

Corollary. *If two analytic functions f, g agree on a non-discrete set, then they are identical.*

6 Path Connectedness

Connectedness lines up with our intuition that a space "comes in one piece", but we also have the intuition that a connected space should be one where "you can get from any point to any other point by taking a path". This notion turns out to be stronger than connectedness, and is quite useful.

Definition. *A path $p : [0,1] \longrightarrow X$ is a continuous map into a topological space X.*

We remark that the image of a path is always a conected subspace of X, so this lines up with our notion of "connecting points in X by a path". That leads to the following definition.

Definition. *A topological space X is said to be* ***path-connected*** *provided that for any $a, b \in X$ there is a path $p : [0,1] \longrightarrow X$ such that $p(0) = a$ and $p(1) = b$.*

Before we get into the thick of things, let us look at some examples:

Example 1: $\mathbb{R}^n$ is path-connected.

Proof. Let $a, b \in \mathbb{R}^n$. The idea is to first take a straight line from a to the origin, then take a straight line from the origin to b. Makes sense right? Well here's the horrible mathematically precise version of this easy idea:

$$p(x) = \begin{cases} (1-2t) \cdot a & : x \in [0, \frac{1}{2}] \\ (2t-1) \cdot b & : x \in [\frac{1}{2}, 1] \end{cases}$$

□

The only thing you might be concerned about is the continuity of this map, so let us just recall the following easy lemma, called the pasting lemma:

Lemma (The Pasting Lemma)**.** *Let $X = C \cup D$ where C and D are closed, and let $f : C \longrightarrow Y$ be a continuous function and let $g : D \longrightarrow Y$ be continuous, and suppose that $f(x) = g(x)$ for all $x \in C \cap D$. Then the function $G : X \longrightarrow Y$ defined by:*

$$G(x) = \begin{cases} f(x) & : x \in C \\ f(x) & : x \in C \cap D \\ g(x) & : x \in D \end{cases}$$

is continuous.

The proof of this is ommited, but it just invloves showing that the preimage of a closed set is closed, which is easy here.

Example 2: Let $\mathbb{M}$ be the collection of 3×3 matrices with real coefficiants who satisfy $A^3 = [0]$, the zero matrix. Then this space is path connected.

Proof. At first this problem seems intractible, but it has the same idea as the previous example, first take a straight line from A to $[0]$ and then take a straight line to B. The thing that isn't so clear here is that the path stays inside $\mathbb{M}$.

Let $A, B \in \mathbb{M}$ and let $[0]$ be the zero matrix. Define a path $p : [0,1] \longrightarrow \mathbb{M}$ by:

$$p(x) = \begin{cases} (1-2t)\cdot A & : x \in [0, \frac{1}{2}] \\ (2t-1)\cdot B & : x \in [\frac{1}{2}, 1] \end{cases}$$

This is continuous since multiplication and addition is continuous and by the Pasting Lemma. Clearly $p(0) = A$ and $p(1) = B$. Now observe that if $C \in \mathbb{M}$ and $t \geq 0$, then

$$(t \cdot C)^3 = t^3 \cdot C^3 = t^3 \cdot [0] = [0]$$

So p is a map into $\mathbb{M}$. □

Now that we have some examples, let's observe the following nice fact (which you may have guessed) which relates connectedness and path-connectedness.

Proposition. *Every path-connected space is connected.*

Proof. Suppose that X is path-connected, but not connected. Let A, B be two disjoint (non-empty) clopen subsets of X. Let $a \in A$ and $b \in B$, since they are non-empty. Let $p : [0,1] \longrightarrow X$ be a path from a to b. Then we see by continuity of p that $p^{-1}[A]$ is a non-empty clopen subset of I that isn't all of $[0,1]$ since $p^{-1}(b) \notin p^{-1}[A]$. This contradicts connectedness of $[0,1]$. □

7 Topologist's Things

Here we introduce two interesting topological spaces: The Topologist's Sine Curve, and the Topologist's Comb. These examples serve to constrast the notions of connectedness and path-connectedness. Specifically we will see two examples of spaces that are connected but not path connected.

Example. ***The Topologist's Sine Curve***

Discussion. Most of the intuition of this example comes from the picture, so be sure to draw one. This space will be a subspace of $\mathbb{R}^2$.

Let

$$TS := \{(0,y) : -1 \leq y \leq 1\} \cup \{(x, f(x)) : 0 < x \leq 1, f(x) := \sin(\frac{1}{x})\}$$

which looks like a very wiggly sine curve to the right of the y-axis, and a small subinterval on the y-axis itself.

Observe that this space is a closed bounded subset of $\mathbb{R}^2$ so it is compact.

This space hace the property that it is connected but is not path connected (there is no path connecting the sine portion to the y-axis). Connectedness is straightforward as the sine portion and the y-axis portion are each clearly path-connected, and any open set that contains the y-axis portion must also intersect the sine portion. Path-connectedness is annoying to prove, but seems clear enough.

Let the point $a = (1, \sin(1))$ and let $b = (0, 0)$. For the sake of contradiction, assume that $p : [0, 1] \longrightarrow TS$ is a path that sends 0 to a and 1 to b. Since p is continuous, the idea is that p must "travel along the sine curve" and cannot "jump off of the sine curve to get to the y-axis". We will omit the proof since it is really just an exercise in playing with ϵs. $\square$

On its own this example isn't so interesting, but it can be used to construct interesting subsets of $\mathbb{R}^2$. Playing around with the definition can lead to some interesting results. Try drawing pictures of the following:

1. $TS_2 := \{(0, 0)\} \cup \{(x, f(x)) : 0 < x \leq 1, f(x) := x \sin(\frac{1}{x})\}$, which is a shrinking sine curve. This one is actually path connected.
2. This one is given in polar coordinates. $TS_3 := \{(0, y) : 1 \leq y \leq 3\} \cup \{(\theta, f(\theta)) : 0 < \theta \leq 2\pi, f(x) := 2 + \sin(\frac{1}{\theta})\}$. This is the topologist's Sine Curve on a circle.
3. $TS_4 := \{(0, y) : -1 \leq y \leq 1\} \cup \{(x, f(x)) : 0 < x \leq 1, f(x) := e^{-x} \sin(\frac{1}{x})\}$. This version of the Topologist's Sine Curve always has the sine portion bounded by $e^{-x} < 1$.
4. $TS_5 := \{(0, -1), (0, 1)\} \cup \{(x, f(x)) : 0 < x \leq 1, f(x) := \sin(\frac{1}{x})\}$ is a version that is no longer closed, so it is not compact.

It is often helpful to be able to adapt the clever examples by making slight changes. We should think of the topologist's sine curve as a building block for interesting examples.

Example. ***The Topologist's Combs***

Discussion. There are two main examples here, one is called the Topologist's Comb and the other is the Deleted Topologist's Comb (which will be like TS_5). Define the **Topologist's Comb** as

$$TC := \{(x, 0) : 0 \leq x \leq 1\} \cup \{(0, y) : 0 \leq y \leq 1\} \cup \bigcup_{n \in \mathbb{N}} \{(\frac{1}{n}, y) : 0 \leq y \leq 1\}$$

This looks like a comb with infinitely many teeth converging to the y-axis.

This space is clearly path-connected as each point clearly has a path between it and $(0,0)$: you just travel down to the x-axis (if needed) and then travel across to $(0,0)$ (if needed). Again we also see that this is a compact space.

Similarly, we define the **Deleted Topologist's Comb** by

$$TC := \{(x,0) : 0 \leq x \leq 1\} \cup \{(0,1)\} \cup \bigcup_{n \in \mathbb{N}} \{(\frac{1}{n}, y) : 0 \leq y \leq 1\}$$

where this is the Topologist's Comb except that we have removed the set $\{(0,y) : 0 < y < 1\}$. This is no longer a path-connected space, but it remains connected; any open set around $(0,1)$ must intersect the rest of the comb.

Both combs have a very strange property: Any small open set $B_\epsilon((0,1))$ around $(1,0)$, with $\epsilon < 1$ is not path-connected (or even connected!). No matter what $\epsilon < 1$ you take there will be an $n \in \mathbb{N}$ such that $(\frac{1}{n}, 1) \in B_\epsilon((0,1))$ but there is no path from $(0,1)$ to $(\frac{1}{n}, 1)$ in $B_\epsilon((0,1))$, because this open set does not contain any part of the x-axis.

This property is so useful that we give it a name: local connectedness and local path-connectedness. □

Definition. *A topological space $(X, \mathcal{T})$ is said to be* ***locally connected at a point*** $p \in X$ *provided that whenever U is an open set containing p then there is an open set $V \subseteq U$ containing p such that V is connected. If X is locally connected at p for each $p \in X$ then we say that X is* ***locally connected****.*

Definition. *A topological space $(X, \mathcal{T})$ is said to be* ***locally path-connected at a point*** $p \in X$ *provided that whenever U is an open set containing p then there is an open set $V \subseteq U$ containing p such that V is path-connected. If X is locally path-connected at p for each $p \in X$ then we say that X is* ***locally path-connected****.*

We often think about objects in terms of their local properties. For example, the Earth is locally flat, and we often think of it as being a flat object; the fact that the Earth is spherical does not usually come up in the construction of buildings (for example). Often in mathematics we will need both an assumption about the space locally (like local connectedness) together with some assumption about the space globally (like separability) to get a particular result. Local connectedness and its siblings are extremely useful in the study of algebraic topology; The property of being locally Euclidean is the start of the study of manifolds; local compactness is a key property is analysis. We will investigate local compactness in the next section.

Local Exercise: Try to define local compactness. Make sure that your definition doesn't make every space locally compact.