General Topology

based on notes by Tom Leinster

2018–19

Contents

A Topological spaces **3**

Lecture 2 . . . 3
A1 Review of metric spaces . . . 3
Lecture 3 . . . 8
A2 The definition of topological space . . . 9
Lecture 4 . . . 13
A3 Metrics versus topologies . . . 13
Lecture 5 . . . 17
A4 Continuous maps . . . 17
Lecture 6 . . . 21
A5 When are two spaces homeomorphic? . . . 22
Lecture 7 . . . 25
A6 Topological properties . . . 26
A7 Bases . . . 28
Lecture 8 . . . 30
A8 Closure and interior . . . 31
Lecture 9 . . . 34
A9 Subspaces (new spaces from old, 1) . . . 35
Lecture 10 . . . 39
A10 Products (new spaces from old, 2) . . . 39
Lecture 11 . . . 44
A11 Quotients (new spaces from old, 3) . . . 44
Lecture 12 . . . 48
A12 Review of Chapter A . . . 48

B Compactness **52**

Lecture 13 . . . 52
B1 The definition of compactness . . . 52
Lecture 14 . . . 55
B2 Closed bounded intervals are compact . . . 56
B3 Compactness and subspaces . . . 56
B4 Compactness and products . . . 58
Lecture 15 . . . 58
B5 The compact subsets of $\mathbb{R}^n$ . . . 59
B6 Compactness and images and quotients . . . 60
Lecture 16 . . . 62
B7 Compact metric spaces . . . 62

C Connectedness **68**
Lecture 17 . . . 68
C1 The definition of connectedness . . . 68
Lecture 18 . . . 71
C2 Connected subsets of the real line . . . 71
Lecture 19 . . . 75
C3 Path-connectedness . . . 75
Lecture 20 . . . 79
C4 Connected-components and path-components . . . 79

Epilogue **85**

Chapter A

Topological spaces

A1 Review of metric spaces

For Lecture 2

Almost everything in this section should have been covered in Honours Analysis, with the possible exception of some of the examples. For that reason, this lecture is longer than usual.

Definition A1.1 Let X be a set. A **metric** on X is a function $d\colon X \times X \to [0, \infty)$ with the following three properties:

- $d(x, y) = 0 \iff x = y$, for $x, y \in X$;
- $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$ (triangle inequality);
- $d(x, y) = d(y, x)$ for all $x, y \in X$ (symmetry).

A **metric space** is a set together with a metric on it, or more formally, a pair (X, d) where X is a set and d is a metric on X.

Examples A1.2 i. The **Euclidean metric** d_2 on $\mathbb{R}^n$ is given by

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

for all $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$. So $(\mathbb{R}^n, d_2)$ is a metric space. The same formula defines a metric d_2 on X for any $X \subseteq \mathbb{R}^n$.

ii. This is not the only metric on $\mathbb{R}^n$. For example, there is a metric d_1 on $\mathbb{R}^n$ given by

$$d_1(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

(the 'taxicab' or 'Manhattan' metric), and another, d_∞, given by

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

(In fact, there is a metric d_p on $\mathbb{R}^n$ for each $p \geq 1$; perhaps you can guess what it is from the definitions of d_1 and d_2. The limit of $d_p(x, y)$ as $p \to \infty$ is $d_\infty(x, y)$, hence the name.)

iii. Let $a, b \in \mathbb{R}$ with $a \leq b$, and let $C[a, b]$ denote the set of continuous functions $[a, b] \to \mathbb{R}$. There are at least three interesting metrics on $C[a, b]$, which again are denoted by d_1, d_2 and d_∞. They are defined by

$$\begin{aligned} d_1(f, g) &= \int_a^b |f(t) - g(t)|\, dt, \\ d_2(f, g) &= \left(\int_a^b (f(t) - g(t))^2\, dt \right)^{1/2}, \\ d_\infty(f, g) &= \sup_{a \leq t \leq b} |f(t) - g(t)|. \end{aligned}$$

If you do the Linear Analysis or Fourier Analysis course, you'll get very used to the idea of spaces *whose elements are functions*. This is a tremendously important idea. Reasoning and writing about such spaces requires care, as they are sets whose elements are functions from one set to another set—and often, you need to talk about functions *between* such spaces.

iv. Let A be any set, which you might think of as an alphabet. Let $n \in \mathbb{N}$. The **Hamming metric** d on A^n is given by

$$d(x, y) = \big|\{ i \in \{1, \ldots, n\} : x_i \neq y_i \}\big|$$

for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in A^n. In other words, the Hamming distance between two strings or 'words' $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ is the number of coordinates in which they differ.

The Hamming metric is often used in the theory of information and communication. For instance, if I write the word 'needle' on the blackboard and you mistakenly copy it down as 'noodle', the Hamming distance between the words is 2, which is the number of errors of communication.

v. An informal example: consider any region of space X, such as the area within the King's Buildings accessible by foot. (This excludes the space occupied by trees, walls, etc.) We can certainly use the Euclidean metric on X, which is the distance as the crow flies. But in practical terms, we are often more interested in the 'shortest path metric', that is, the distance by foot. This is indeed a metric; you should be able to persuade yourself that the three axioms hold.

Strictly speaking, we should write metric spaces as pairs (X, d), where X is a set and d is a metric on X. But usually, I will just say 'a metric space X', using the letter d for the metric unless indicated otherwise.

Definition A1.3 Let X be a metric space, let $x \in X$, and let $\varepsilon > 0$. The **open ball around x of radius ε**, or more briefly the **open ε-ball around x**, is the subset

$$B(x, \varepsilon) = \{y \in X : d(x, y) < \varepsilon\}$$

of X. Similarly, the **closed ε-ball around x** is

$$\bar{B}(x, \varepsilon) = \{y \in X : d(x, y) \leq \varepsilon\}.$$

A good exercise is to go through Examples A1.2 and work out what the open and closed balls are in each of the examples given.

Now comes an extremely important definition.

Definition A1.4 Let X be a metric space.

i. A subset U of X is **open in** X (or an **open subset of** X) if for all $u \in U$, there exists $\varepsilon > 0$ such that $B(u, \varepsilon) \subseteq U$.

ii. A subset V of X is **closed in** X if $X \setminus V$ is open in X.

Thus, U is open if every point of U has some elbow room—it can move a little bit in each direction without leaving U.

Warning A1.5 Closed does not mean 'not open'! Subsets are not like doors. A subset of a metric space can be:

- neither open nor closed, such as $[0, 1)$ in $\mathbb{R}$
- both open and closed, such as $\mathbb{R}$ in $\mathbb{R}$
- open but not closed, such as $(0, 1)$ in $\mathbb{R}$
- closed but not open, such as $[0, 1]$ in $\mathbb{R}$.

Scumbag topologist 'opens' a store.

Warning A1.6 Another warning: properly, there's no such thing as an 'open set', only an open *sub*set. In other words, we should never say 'U is open'; we should always say 'U is open *in* X'. This can matter. For instance, $[0, 1)$ is not open in $\mathbb{R}$, but it is open in $[0, 2]$. (Why?)

In practice, it's often clear which space X we're operating inside, and then it's generally safe to speak of sets simply being 'open' without mentioning which space they're open in. (Wade's book *An Introduction to Analysis*, which was the main text for Honours Analysis, is often casual in this way.) Nevertheless, it's important to realize that this *is* a casual use of language, and can lead to errors if you're not careful.

Remark A1.7 Open balls are open and closed balls are closed. For a proof, see Remark 10.9 of Wade's book, or try it as an exercise.

Closed subsets of a metric space can be characterized in terms of convergent sequences, as follows.

Definition A1.8 Let X be a metric space, let $(x_n)_{n=1}^{\infty}$ be a sequence in X, and let $x \in X$. Then (x_n) **converges** to x if

$$d(x_n, x) \to 0 \text{ as } n \to \infty.$$

Explicitly, then, (x_n) converges to x if and only if: for all $\varepsilon > 0$, there exists $N \geq 1$ such that for all $n \geq N$, $d(x_n, x) < \varepsilon$. This generalizes the definition you're familiar with for $\mathbb{R}$.

Lemma A1.9 *Let X be a metric space and $V \subseteq X$. Then V is closed in X if and only if:*

> *for all sequences (x_n) in V and all x in X, if (x_n) converges to x then $x \in V$.*

A proof very similar to the following can also be found in Wade (Theorem 10.16).

Proof Suppose that V is closed, and let (x_n) be a sequence in V converging to some point $x \in X$. We must show that $x \in V$. Suppose for a contradiction that $x \in X \setminus V$. Since $X \setminus V$ is open in X, there is some $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq X \setminus V$. Now (x_n) converges to x, so there exists N such that $d(x_n, x) < \varepsilon$ for all $n \geq N$. In particular, $d(x_N, x) < \varepsilon$, that is, $x_N \in B(x, \varepsilon)$; hence $x_N \in X \setminus V$. This contradicts the hypothesis that (x_n) is a sequence in V.

Now suppose that V is not closed. We must show that the given condition does not hold, in other words, that there exists a sequence (x_n) in V converging to a point of X not in V. Since V is not closed, $X \setminus V$ is not open. Hence there is some point $x \in X \setminus V$ with the property that for all $\varepsilon > 0$, the ball $B(x, \varepsilon)$ has nonempty intersection with V. For each $n \geq 1$, choose an element $x_n \in B(x, 1/n) \cap V$. Then (x_n) is a sequence in V converging to $x \in X \setminus V$, as required. $\square$

Next we state some fundamental properties of open and closed subsets. In order to do this, we'll need to recall some basic set theory.

Remark A1.10 Let X be a set. A **family** $(A_i)_{i \in I}$ of subsets of X is a set I together with a subset $A_i \subseteq X$ for each $i \in I$. **De Morgan's laws** state that

$$X \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X \setminus A_i), \qquad X \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \setminus A_i).$$

The family $(A_i)_{i \in I}$ is said to be **finite** if I is finite. (It has nothing to do with whether the subsets A_i are finite.)

Lemma A1.11 *Let X be a metric space.*

i. *Let $(U_i)_{i \in I}$ be any family (finite or not) of open subsets of X. Then $\bigcup_{i \in I} U_i$ is also open in X.*

ii. *Let U_1 and U_2 be open subsets of X. Then $U_1 \cap U_2$ is also open in X.*

iii. *$\emptyset$ and X are open in X.*

Part (i) can be phrased less formally as 'a union of open sets is open'. Similarly, part (ii) (plus an easy induction) says 'a finite intersection of open sets is open'.

Proof For (i), let $x \in \bigcup_{i \in I} U_i$. Choose $j \in I$ such that $x \in U_j$. Since U_j is open in X, we can then choose $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U_j$. It follows that $B(x, \varepsilon) \subseteq \bigcup_{i \in I} U_i$.

For (ii), let $x \in U_1 \cap U_2$. For $i = 1, 2$, we can choose $\varepsilon_i > 0$ such that $B(x, \varepsilon_i) \subseteq U_i$. Put $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} > 0$. Then $B(x, \varepsilon) \subseteq U_1 \cap U_2$.

For (iii): any statement beginning 'for all $x \in \emptyset \ldots$' is trivially true, so $\emptyset$ is open. To see that X is open in X, let $x \in X$. Then $B(x, 1) \subseteq X$, simply because all balls in X are by definition subsets of X. □

It is *not* true that an arbitrary intersection of open subsets is open. For example, $(-\infty, 1/n)$ is an open subset of $\mathbb{R}$ for each $n \geq 1$, but $\bigcap_{n \geq 1} (-\infty, 1/n) = (-\infty, 0]$ is not open in $\mathbb{R}$.

Lemma A1.12 *Let X be a metric space.*

i. *Let $(V_i)_{i \in I}$ be any family (finite or not) of closed subsets of X. Then $\bigcap_{i \in I} V_i$ is also closed in X.*

ii. *Let V_1 and V_2 be closed subsets of X. Then $V_1 \cup V_2$ is also closed in X.*

iii. *$\emptyset$ and X are closed in X.*

Proof This follows from Lemma A1.11 by de Morgan's laws. □

Again, an *arbitrary* union of closed sets need not be closed. For example, $[1/n, \infty)$ is closed in $\mathbb{R}$ for each $n \geq 1$, but $\bigcup_{n \geq 1} [1/n, \infty) = (0, \infty)$ is not closed in $\mathbb{R}$.

Lemma A1.11 will be the key to making the leap from metric to topological spaces. We will see this in the next lecture.

Metric spaces do not live in isolation. We can also talk about functions (also called maps or mappings) between them. Typically, we are only interested in the *continuous* functions.

Definition A1.13 Let X and Y be metric spaces. A function $f : X \to Y$ is **continuous** if for all $x \in X$, for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$x' \in B(x, \delta) \Longrightarrow f(x') \in B(f(x), \varepsilon).$$

This generalizes the familiar definition for $X = Y = \mathbb{R}$.

The definition of continuity appears to make essential use of the metrics on X and Y. However, the following lemma reveals that this is not really so. In order to decide which functions are continuous, all we actually need is knowledge of the open (or closed) subsets.

Lemma A1.14 *Let X and Y be metric spaces and let $f : X \to Y$ be a function. The following are equivalent:*

i. f is continuous;

ii. for all open $U \subseteq Y$, the preimage $f^{-1}U \subseteq X$ is open;

iii. for all closed $V \subseteq Y$, the preimage $f^{-1}V \subseteq X$ is closed.

Recall that the **preimage** or **inverse image** $f^{-1}U$ is the subset $\{x \in X : f(x) \in U\}$ of X. It is defined whether or not f is invertible. Part (ii) says, informally: 'the preimage of an open set is open'.

Proof For (i)$\Longrightarrow$(ii), suppose that f is continuous and let U be an open subset of Y. We must show that $f^{-1}U$ is an open subset of X. Let $x \in f^{-1}U$. Then $f(x) \in U$, so we can choose $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subseteq U$. By continuity, we can then choose $\delta > 0$ such that

$$x' \in B(x, \delta) \Longrightarrow f(x') \in B(f(x), \varepsilon).$$

But then

$$x' \in B(x, \delta) \Longrightarrow f(x') \in U \iff x' \in f^{-1}U,$$

so $B(x, \delta) \subseteq f^{-1}U$, as required.

For (ii)$\Longrightarrow$(i), suppose that the preimage of every open set is open. We must prove that f is continuous. Let $x \in X$ and $\varepsilon > 0$. By Remark A1.7, the open ball $B(f(x), \varepsilon)$ is open, so $f^{-1}B(f(x), \varepsilon)$ is also open. Evidently it contains the point x, so there is some $\delta > 0$ such that

$$B(x, \delta) \subseteq f^{-1}B(f(x), \varepsilon).$$

But this says exactly that

$$x' \in B(x, \delta) \Longrightarrow f(x') \in B(f(x), \varepsilon),$$

as required.

Finally, (ii) $\iff$ (iii) follows from the fact that

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

for any $W \subseteq Y$. For instance, if (ii) holds then for any closed V in Y, the set $Y \setminus V$ is open in Y, so $f^{-1}(Y \setminus V)$ is open in X. But $f^{-1}(Y \setminus V) = X \setminus f^{-1}V$, so $f^{-1}V$ is closed in X, proving (iii). □

What next? We've just seen that continuity can be phrased in terms of open sets alone. We've also seen what properties the open sets in a metric space always have (Lemma A1.11).

Abstracting, we'll *define* a topological space to be a set X equipped with a collection of subsets (called 'open') satisfying the three properties in Lemma A1.11. We'll *define* a function between topological spaces to be continuous if the preimage of an open set is open. In that way, we'll have succeeded in generalizing the notion of continuity to a context where distance isn't even mentioned.

(We could equally well do this with closed sets instead of open sets. But it's usually the open sets that are given the upper hand.)

A2 The definition of topological space

For Lecture 3

We saw in Lemma A1.11 that the collection $\mathcal{T}$ of open subsets of a metric space has certain properties. Following the strategy explained at the end of the last section, we now turn those properties into a definition.

Definition A2.1 Let X be a set. A **topology** on X is a set $\mathcal{T}$ of subsets of X with the following properties.

- Whenever $(U_i)_{i \in I}$ is a family (finite or not) of subsets of X such that $U_i \in \mathcal{T}$ for all $i \in I$, then $\bigcup_{i \in I} U_i \in \mathcal{T}$.
- Whenever $U_1, U_2 \in \mathcal{T}$, then $U_1 \cap U_2 \in \mathcal{T}$.
- $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$.

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

Remarks A2.2 i. 'Topology' is both the name of the subject and the word for one of the central definitions of the subject! (If you do lots of algebra, you'll eventually learn that there is such a thing as 'an algebra' too.)

ii. We often write $(X, \mathcal{T})$ as just X, in situations where there is no ambiguity about which topology $\mathcal{T}$ we could mean. A single set X *can* carry many different topologies (as we shall see), but the context often makes clear which one is intended.

iii. We call the elements of $\mathcal{T}$ the **open** subsets of X. Thus, '$U \in \mathcal{T}$' and 'U is open in X' mean the same thing. Again, this is safe terminology when it is clear from the context which topology on X we are talking about.

To be clear: 'open' is just a word that we choose to use. By definition, an open subset of a topological space $(X, \mathcal{T})$ is simply an element of $\mathcal{T}$. It's like the way that in an abstract group, we refer to the group operation as 'multiplication', whether or not it's multiplication in any ordinary sense. Of course, mathematicians chose to use the word 'open' rather than some other word because it fits well with how it's used for metric spaces. We'll see this in a moment (Example A2.3(i)).

iv. The first axiom says that an arbitrary union of open subsets is open. The second implies (by induction) that a *finite* intersection $U_1 \cap \cdots \cap U_n$ of open subsets U_i is open, for any $n \geq 1$.

v. Strictly speaking, it is unnecessary to state the condition '$\emptyset \in \mathcal{T}$' in the third axiom. The first axiom already implies it. To see this, take $(U_i)_{i \in I}$ to be the empty family of subsets of X, that is, the unique family with $I = \emptyset$. Then $\bigcup_{i \in \emptyset} U_i$ is the set of all points x such that $x \in U_i$ for some $i \in \emptyset$; but there are *no* such points x, so $\bigcup_{i \in \emptyset} U_i = \emptyset$.

Similarly, the intersection of the empty family of subsets of X is the set of all $x \in X$ such that $x \in U_i$ for all $i \in \emptyset$; but *every* x has this property, so $\bigcap_{i \in \emptyset} U_i = X$. The definition of topology would therefore be unaffected

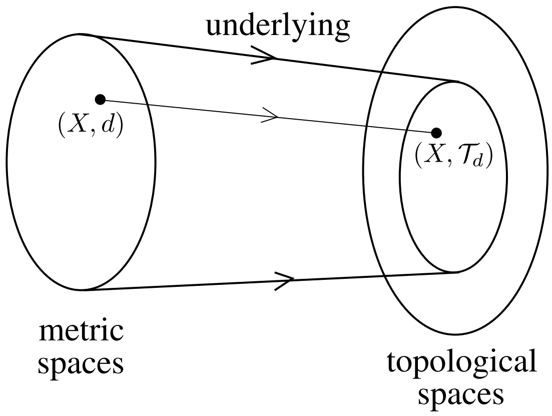


Figure A.1: From metric to topological spaces.

if we changed the second axiom to 'whenever $(U_i)_{i\in I}$ is a finite family of subsets of X such that $U_i \in \mathcal{T}$ for all $i \in I$, then $\bigcap_{i\in I} U_i \in \mathcal{T}$' and dropped the third axiom.

In summary: as long as you treat trivial cases with care, a topology on X can be defined as a collection of subsets of X that is closed under arbitrary unions and finite intersections. (This is a different mathematical usage of the word 'closed' from the topological one!)

Examples A2.3 i. Let (X, d) be a metric space. Put

$$\mathcal{T}_d = \{\text{open subsets of } (X, d)\}.$$

Lemma A1.11 says exactly that $\mathcal{T}_d$ is a topology on X. Thus, $(X, \mathcal{T}_d)$ is a topological space. We call $\mathcal{T}_d$ the topology **induced** by the metric d. We also call $(X, \mathcal{T}_d)$ the **underlying topological space** of the metric space (X, d). See Figure A.1.

It's worth taking a moment to think about the usage of the word 'open'. If (X, d) is a metric space, to say that $U \subseteq X$ is open means that for all $x \in U$, there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. If $(X, \mathcal{T})$ is a topological space, to say that $U \subseteq X$ is open means simply that $U \in \mathcal{T}$. These two usages are compatible, in the sense that U is open in the metric space (X, d) if and only if U is open in the topological space $(X, \mathcal{T}_d)$.

ii. The **standard topology** on $\mathbb{R}^n$ is the topology induced by the Euclidean metric d_2. We will see that this is the same as the topology induced by the metric d_1, and also the same as the topology induced by the metric d_∞. (In fact, the metrics d_p mentioned in Example A1.2(ii) all induce the same topology, for $1 \leq p \leq \infty$.)

iii. Let X be any set. The **discrete topology** on X is the collection of *all* subsets of X. It is induced by a metric, the so-called **discrete** metric d on X, which is defined by

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

(Exercise: check this is true!) It is the largest possible topology on X.

iv. Let X be any set. The **indiscrete topology** on X is the topology $\{\emptyset, X\}$; that is, it's the topology in which only $\emptyset$ and X are open. It is the smallest possible topology on X.

(Note: discreet means able to keep a secret, and indiscreet means gossipy. The topologies are discr*ete* and indiscr*ete*—there is no gossipy topology.)

When one topology $\mathcal{T}$ on a set X contains another topology $\mathcal{T}'$ on X (that is, every member of $\mathcal{T}'$ is a member of $\mathcal{T}$), we *could* say that $\mathcal{T}$ is 'larger' than $\mathcal{T}'$, as we just did. But in fact, it's more common to say that $\mathcal{T}$ is a **stronger** or **finer** topology than $\mathcal{T}'$, or that $\mathcal{T}'$ is a **weaker** or **coarser** topology than $\mathcal{T}$. So, the discrete topology is the strongest or finest topology on X, and the indiscrete topology is the weakest or coarsest.

Examples A2.3(iii) and (iv) show that the same set can have different topologies on it.

Is the process illustrated in Figure A.1 injective? In other words, if you have two different metrics on the same set, do they always give rise to different topologies?

Is the process illustrated in Figure A.1 surjective? In other words, is every topology on a set induced by some metric?

The following examples (Figure A.2) show that the answer to both questions is no.

Examples A2.4 i. Consider $\mathbb{Z}$ with its usual metric, $d(m, n) = |m - n|$. In this metric, every subset is open. (Proof: consider balls of radius $1/2$, say.) So the topology on $\mathbb{Z}$ induced by d is the discrete topology. But this is also the topology induced by the discrete metric d_{disc} on $\mathbb{Z}$. So $(\mathbb{Z}, d)$ and $(\mathbb{Z}, d_{\text{disc}})$ induce the same topology on $\mathbb{Z}$.

ii. Let $X = \{1, 2\}$ and let $\mathcal{T}$ be the indiscrete topology on X. We will show that $\mathcal{T}$ is not induced by any metric on X. (The same is true for an arbitrary set X with two or more elements.) Indeed, let d be a metric on X. Put $r = d(1, 2) > 0$. Then $B(1, r) = \{1\}$, so in the topology induced by d, the set $\{1\}$ is open in X. But $\{1\}$ is not open in the indiscrete topology.

Definition A2.5 Let $X = (X, \mathcal{T})$ be a topological space. A subset $V \subseteq X$ is **closed** (for $\mathcal{T}$) if $X \setminus V \in \mathcal{T}$.

Warning A1.5 applies here, too: closed does *not* mean 'not open'.

Examples A2.6 i. Let (X, d) be a metric space. Let $V \subseteq X$. Then

$$\begin{aligned}
& V \text{ is closed for } \mathcal{T}_d \\
\iff & X \setminus V \text{ is open for } \mathcal{T}_d \\
\iff & X \setminus V \text{ is open in the metric space } (X, d) \\
\iff & V \text{ is closed in the metric space } (X, d).
\end{aligned}$$

Conclusion: in the underlying topological space of a metric space, 'open' and 'closed' mean exactly the same as in the metric space itself.

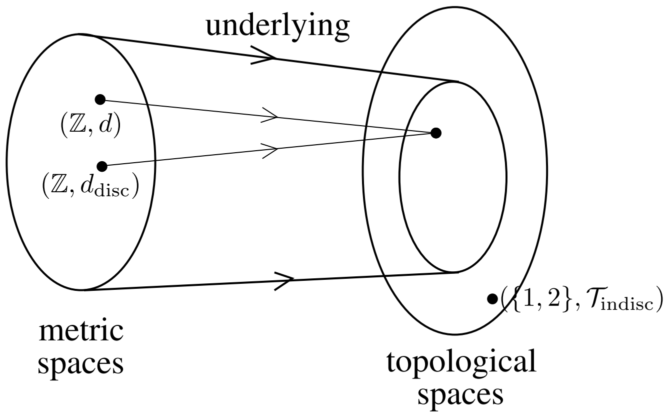


Figure A.2: The passage from metric to topological spaces is neither injective nor surjective.

ii. In particular, this applies to $\mathbb{R}^n$. So in the standard topology on $\mathbb{R}^n$, closed has its usual meaning.

iii. In the discrete topology, all subsets are closed (as well as open).

iv. In the indiscrete topology on a set X, only $\emptyset$ and X are closed.

Here are some basic facts about closed subsets of a topological space, generalizing Lemma A1.12 for metric spaces.

Lemma A2.7 *Let $X = (X, \mathcal{T})$ be a topological space.*

i. Whenever $(V_i)_{i \in I}$ is a family (finite or not) of closed subsets of X, then $\bigcap_{i \in I} V_i$ is closed in X.

ii. Whenever V_1 and V_2 are closed subsets of X, then $V_1 \cup V_2$ is also closed in X.

iii. $\emptyset$ and X are closed subsets of X.

Proof This follows from the definition of topological space by de Morgan's laws (Remark A1.10). □

In a general topological space, we cannot speak of balls around a point, because there is no notion of distance. However, we might still want to speak of 'small' regions around a point. The following terminology helps us do that.

Definition A2.8 Let X be a topological space and $x \in X$. An **open neighbourhood** of x is an open subset of X containing x. A **neighbourhood** of x is a subset of X containing an open neighbourhood of x.

For example, the subsets $[-\varepsilon, \varepsilon]$, $[-\varepsilon, \varepsilon)$, and $(-\varepsilon, \varepsilon)$ of $\mathbb{R}$ are all neighbourhoods of 0 (for any $\varepsilon > 0$), but only the last is an *open* neighbourhood of 0.

You should check that a subset of X is an open neighbourhood of x if and only if it is open (in X) and a neighbourhood of x. (In other words, check that the expression 'open neighbourhood' is unambiguous.)

The following lemma can be useful when you're trying to show that a subset is open.

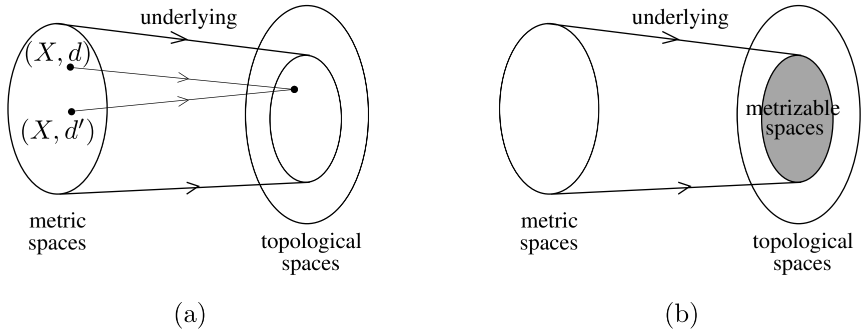


Figure A.3: (a) Topologically equivalent metrics d and d' on a set X; (b) the metrizable spaces.

Lemma A2.9 *Let X be a topological space and $U \subseteq X$. Then U is open in X if and only if for all $x \in U$, there is a neighbourhood of x contained in U.*

The same is true if 'topological' is replaced by 'metric' and 'neighbourhood of x' by 'open ball around x'. That's just the definition of open subset of a metric space.

Proof If U is open in X then for all $x \in U$, the set U itself is a neighbourhood of x contained in U.

Conversely, suppose that for each $x \in U$, there is a neighbourhood N_x of x contained in U. Then for each x, there is an open neighbourhood $U_x \subseteq N_x$ of x. Since $x \in U_x \subseteq U$ for each $x \in U$, we have $\bigcup_{x \in U} U_x = U$. But the union of open subsets of X is open, so U is open. $\square$

A3 Metrics versus topologies

For Lecture 4

We saw in Section A2 that the passage from metrics to topologies is neither injective nor surjective. That is, (i) different metrics on a set can induce the same topology, and (ii) some topologies are not induced by any metric at all. In this section, we take a closer look at these two phenomena.

First we consider the fact that different metrics can induce the same topology. Some terminology is useful (Figure A.3).

Definition A3.1 Let X be a set, and let d and d' be metrics on X. We say that d and d' are **topologically equivalent** if they induce the same topology on X.

It is immediate that topological equivalence is an equivalence relation (as the name suggests!) on the set of all metrics on X.

Example A2.4(i) describes two different but topologically equivalent metrics on $\mathbb{Z}$.

What tools do we have for showing that two metrics are topologically equivalent? We could use the definition directly (as in the example just mentioned), or we could try to verify the following useful condition.

Definition A3.2 Let X be a set, and let d and d' be metrics on X. We say that d and d' are **Lipschitz equivalent** if there exist real numbers $c, C > 0$ such that for all $x, y \in X$,

$$cd(x,y) \leq d'(x,y) \leq Cd(x,y).$$

Again, Lipschitz equivalence is an equivalence relation on the set of all metrics on X. (Check!)

Lemma A3.3 *Lipschitz equivalent metrics are topologically equivalent.*

Proof Let d and d' be Lipschitz equivalent metrics on a set X, and choose constants c and C as in Definition A3.2. Write B_d and $B_{d'}$ for balls with respect to d and d', respectively.

First note that for all $a \in X$ and $r > 0$,

$$B_d(a,r) \supseteq B_{d'}(a,cr),$$

since whenever $x \in B_{d'}(a,cr)$, we have

$$d(a,x) \leq \frac{1}{c} \cdot d'(a,x) < \frac{1}{c} \cdot cr = r.$$

Now let U be a subset of X that is open with respect to d. Let $a \in U$. We may choose $r > 0$ such that $B_d(a,r) \subseteq U$. But $B_{d'}(a,cr) \subseteq B_d(a,r)$, so $B_{d'}(a,cr) \subseteq U$, with $cr > 0$. Hence U is open with respect to d'.

We have now shown that any subset of X open with respect to d is open with respect to d', using the inequality $cd \leq d'$. The converse is proved similarly, using the inequality $\frac{1}{C}d' \leq d$. $\square$

Examples A3.4 i. Let (X,d) be a metric space and $t > 0$. Then there is a metric td on X defined by $(td)(x,y) = t \cdot d(x,y)$, for $x, y \in X$. This is Lipschitz equivalent to d, and therefore topologically equivalent too.

ii. The metrics d_1, d_2 and d_∞ on $\mathbb{R}^n$ are all Lipschitz equivalent, since for all $x, y \in \mathbb{R}^n$,

$$d_\infty(x,y) \leq d_1(x,y) \leq nd_\infty(x,y), \qquad d_\infty(x,y) \leq d_2(x,y) \leq \sqrt{n}d_\infty(x,y).$$

iii. On the other hand, the metrics d_1, d_2 and d_∞ on $C[0,1]$ are all topologically *in*equivalent. We prove that d_1 is not topologically equivalent to d_∞ later in this section, and the other cases can be proved by similar means.

iv. The standard metric d on $\mathbb{Z}$ is topologically equivalent to the discrete metric d_{disc} (Example A2.4(i)). However, it is not Lipschitz equivalent, since for distinct $m, n \in \mathbb{Z}$, the ratio $d(m,n)/d_{\text{disc}}(m,n)$ can be arbitrarily large. So the converse of Lemma A3.3 is false; in other words, Lipschitz equivalence is strictly stronger than topological equivalence.

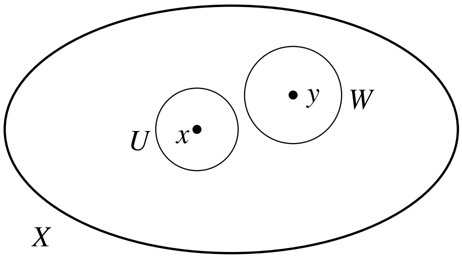


Figure A.4: The Hausdorff condition.

In the second (and longer) part of this section, we consider the fact that not every topology is induced by a metric. Again, some terminology is useful.

Definition A3.5 A topological space $(X, \mathcal{T})$ is **metrizable** if $\mathcal{T}$ is induced by some metric on X.

So, for instance, the two-point indiscrete topological space is not metrizable (Example A2.4(ii)).

Here are some special properties of metrizable spaces.

Definition A3.6 i. A topological space X is said to be T_1 if every one-element subset of X is closed.

ii. A topological space X is **Hausdorff** (or T_2) if for every $x, y \in X$ with $x \neq y$, there exist disjoint open subsets U, W of X such that $x \in U$ and $y \in W$.

As the odd names 'T_1' and 'T_2' hint, these definitions are members of a whole sequence of so-called 'separation conditions'.

The Hausdorff condition is illustrated in Figure A.4. Recall that for U and W to be disjoint means that $U \cap W = \emptyset$.

Lemma A3.7 *i. Every metrizable space is Hausdorff.*

ii. Every Hausdorff topological space is T_1.

Proof For (i), let (X, d) be a metric space and let x, y be distinct points of X. Put $r = d(x, y)/2 > 0$. Then $B(x, r)$ is an open subset of X containing x, and similarly $B(y, r)$ is an open subset of X containing y, so it suffices to show that $B(x, r)$ and $B(y, r)$ are disjoint. This follows from the triangle inequality, as if there exists a point $z \in B(x, r) \cap B(y, r)$ then

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

a contradiction.

For (ii), let X be a Hausdorff topological space and let $x \in X$. For each $y \in X$ with $y \neq x$, we can choose disjoint open neighbourhoods U_y of x and W_y of y. In particular, every point of $X \setminus \{x\}$ has a neighbourhood contained in $X \setminus \{x\}$. So by Lemma A2.9, $X \setminus \{x\}$ is open in X, or equivalently, $\{x\}$ is closed. $\square$

The Hausdorff condition is so useful that is was often included in early formulations of the definition of topological space (around 1910–20). Many geometrically interesting spaces, not only metrizable ones, are Hausdorff. Indeed, many present-day mathematicians assume silently that 'space' means 'Hausdorff space', regarding non-Hausdorff spaces as somehow unhealthy.

However, there *are* useful and important non-Hausdorff spaces, usually of the hard-to-visualize type. (Certainly they are not metrizable.) The Zariski topology is an example, central to algebraic geometry. More trivially, an indiscrete topological space with two or more points is also non-Hausdorff.

After the following short digression, we will see an example of how the Hausdorff condition can be useful.

Definition A3.8 Let X be a topological space, let $(x_n)_{n=1}^{\infty}$ be a sequence in X, and let $x \in X$. Then (x_n) **converges** to x if for all open sets U containing x, there exists N such that for all $n \geq N$, $x_n \in U$.

This is compatible with the familiar definition of convergence in a *metric* space:

Lemma A3.9 *Let (x_n) be a sequence in a metric space X, and let $x \in X$. The following are equivalent:*

i. (x_n) converges to x in the topological sense (Definition A3.8);

ii. (x_n) converges to x in the metric sense (that is, for all $\varepsilon > 0$, there exists N such that for all $n \geq N$, $d(x_n, x) < \varepsilon$).

Proof Exercise. □

So, convergence in a metric space can be expressed in terms of the topology alone.

Examples A3.10 i. When X has the discrete topology, a sequence (x_n) converges to x if and only if it is of the form

$$x_1, \ldots, x_{N-1}, x, x, x, \ldots$$

(in other words, there exists N such that $x_n = x$ for all $n \geq N$). To see this, use the fact that $\{x\}$ is an open set containing x.

ii. When X has the indiscrete topology, every sequence in X converges to every point in X. So, a sequence can converge to multiple points simultaneously.

The possibility of a sequence having multiple limits is avoided by assuming that our space is Hausdorff:

Lemma A3.11 *Let X be a Hausdorff topological space. Then each sequence in X converges to at most one point.*

Proof Let (x_n) be a sequence in X converging to distinct points x and y of X. Since X is Hausdorff, we can choose disjoint open neighbourhoods U of x and W of y. Since (x_n) converges to x, we can choose N such that $x_n \in U$ for all $n \geq N$. Since (x_n) converges to y, we can choose M such that $x_n \in W$ for all $n \geq M$. But then $x_{\max\{N,M\}} \in U \cap W = \emptyset$, a contradiction. □

The notion of convergence of a sequence in a topological space can also be used to prove that two metrics are not topologically equivalent:

Example A3.12 The metrics d_1 and d_∞ on $C[0,1]$ are not topologically equivalent (hence certainly not Lipschitz equivalent). To prove this, define $f_n \in C[0,1]$ by $f_n(x) = x^n \quad (n \geq 1,\ x \in [0,1])$. Let $g \in C[0,1]$ be the constant function 0. Then (f_n) converges to g with respect to d_1, since

$$d_1(f_n, g) = \int_0^1 |x^n|\, dx = \frac{1}{n+1} \to 0$$

as $n \to \infty$. However, (f_n) does not converge to g with respect to d_∞, since

$$d_\infty(f_n, g) = \sup_{x \in [0,1]} |x^n| = 1$$

for all n. Since convergence in a metric space can be expressed in terms of the induced topology alone (Lemma A3.9), the topologies induced by these two metrics are different.

For the record, here are two further 'separation conditions' on topological spaces, less important than Hausdorffness.

Definition A3.13 i. A topological space X is **regular** if for all closed sets $V \subseteq X$ and all $x \in X$ with $x \notin V$, there exist disjoint open sets $U, W \subseteq X$ such that $V \subseteq U$ and $x \in W$.

ii. A topological space X is **normal** if for all disjoint closed sets $V, Z \subseteq X$, there exist disjoint open sets $U, W \subseteq X$ such that $V \subseteq U$ and $Z \subseteq W$.

A normal T_1 space is regular (immediately from the definitions). Every metric space is normal. There are lots of complicated questions that can be asked about T_1, Hausdorff, regular and normal spaces, but we will mostly avoid them.

A4 Continuous maps

For Lecture 5

So far, we've been considering *individual* metric and topological spaces. But if we wish to be able to talk about deforming one space into another (such as a coffee cup into a doughnut), we need to start considering the *relationships between* spaces.

We will do this using the notion of continuous map. We already know what a continuous map between metric spaces is. To generalize the definition to topological spaces, we use the plan described at the end of Section A1.

Definition A4.1 Let X and Y be topological spaces. A function $f\colon X \to Y$ is **continuous** if for every open subset U of Y, the preimage $f^{-1}U$ is open in X.

Thus, continuity means that the preimage of an open set is open.

Remarks A4.2 i. The words 'function', 'map' and 'mapping' usually all mean the same thing, but in practice, people tend to speak of 'continuous maps' between topological spaces, rather than 'continuous functions'. In topology we are seldom interested in discontinuous maps, so in these notes, the word **map** can usually be taken to mean 'continuous map'. I will reserve the word 'function' for not-necessarily-continuous maps.

ii. The definition involves *preimages* (inverse images), not images. There is no easy way to rephrase the definition in terms of images. For instance, continuity is *not* equivalent to the condition that if $U \subseteq X$ is open then so is $fU \subseteq Y$. (Such functions f are called **open**. This condition is occasionally useful, but far less important than continuity.) Nor is it equivalent to the condition that for $U \subseteq X$, if $fU \subseteq Y$ is open then so is U.

Examples A4.3 i. Let (X, d) and (Y, d') be metric spaces. By Lemma A1.14, a function $f \colon X \to Y$ is continuous with respect to the metrics d and d' if and only if it is continuous with respect to the induced topologies $\mathcal{T}_d$ and $\mathcal{T}_{d'}$. In other words, Definition A4.1 for topological spaces is compatible with the familiar definition of continuity for metric spaces.

ii. Let X and Y be topological spaces. If X has the discrete topology then every function $X \to Y$ is continuous. Similarly, if Y has the indiscrete topology then every function $X \to Y$ is continuous.

iii. Let $\mathcal{T}$ and $\mathcal{T}'$ be two different topologies on the same set X. The 'identity' map $i : (X, \mathcal{T}) \to (X, \mathcal{T}')$ (defined by $i(x) = x$) is continuous if and only if every member of $\mathcal{T}'$ is also a member of $\mathcal{T}$, in other words, if $\mathcal{T}$ is finer than $\mathcal{T}'$. For example, the identity map

$$(\mathbb{R}, \text{discrete topology}) \to (\mathbb{R}, \text{standard topology})$$

is continuous.

Lemma A4.4 *Let X and Y be topological spaces. A function $f \colon X \to Y$ is continuous if and only if for every closed subset V of Y, the preimage $f^{-1}V$ is closed in X.*

Proof This is the same as the proof of (ii) $\iff$ (iii) in Lemma A1.14. □

Some topologies are most naturally defined by specifying the closed sets, then declaring the open sets to be their complements. This is the case for the cofinite topology and the Zariski topology. In such cases, Lemma A4.4 provides a useful way of showing that a function is continuous.

Example A4.5 Let k be a field and $n \geq 0$. Let $f \in k[X_1, \ldots, X_n]$ be a polynomial in n variables. Then f defines a function $k^n \to k$, which I claim is continuous with respect to the Zariski topologies on k^n and k.

By Lemma A4.4, it suffices to show that the preimage under f of any closed set is closed. Let V be a Zariski closed subset of k. By definition, there is some $S \subseteq k[X]$ such that $V = Z(S)$; that is,

$$V = \{x \in k : p(x) = 0 \text{ for all } p \in S\}.$$

Hence

$$f^{-1}V = \{(x_1, \ldots, x_n) \in k^n : p(f(x_1, \ldots, x_n)) = 0 \text{ for all } p \in S\}.$$

Put

$$R = \{p(f(X_1, \ldots, X_n)) : p \in S\} \subseteq k[X_1, \ldots, X_n].$$

Then $f^{-1}V = Z(R)$, which is a closed subset of k^n, as required.

Here are some basic properties of continuous maps.

Lemma A4.6 *Continuous maps preserve convergence of sequences. That is, let $f \colon X \to Y$ be a continuous map, and let (x_n) be a sequence in X converging to $x \in X$; then the sequence $(f(x_n))$ in Y converges to $f(x) \in Y$.*

Proof Let U be an open subset of Y containing $f(x)$. Then $f^{-1}U$ is an open subset of X containing x, so there exists N such that $x_n \in f^{-1}U$ for all $n \geq N$. But then $f(x_n) \in U$ for all $n \geq N$, as required. □

Warning A4.7 You may have encountered the fact that for *metric* spaces, a map is continuous *if and only if* it preserves convergence of sequences. For topological spaces, 'only if' is still true, as we have just proved. But 'if' is false: it is possible to construct examples of discontinuous maps of topological spaces that, nevertheless, preserve convergence of sequences.

Lemma A4.8 *i. The identity map on any topological space is continuous.*

ii. The composite of continuous maps is continuous.

Proof For (i), let X be a topological space, and write $\mathrm{id}_X \colon X \to X$ for the identity map on X. For any open $U \subseteq X$, the preimage $\mathrm{id}_X^{-1}U$ is just U, and is therefore also open in X.

For (ii), let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be continuous maps between topological spaces. We have to prove that $X \xrightarrow{g \circ f} Z$ is continuous. Let U be an open subset of Z. Then

$$(g \circ f)^{-1}U = f^{-1}g^{-1}U$$

(check!). But g is continuous, so $g^{-1}U$ is open in Y, and then f is continuous, so $f^{-1}g^{-1}U$ is open in X, as required. □

You can compare this lemma with the fact that the composite of two homomorphisms of groups or rings is again a homomorphism, or the fact that the composite of two linear maps is linear. However, there is a surprise. The inverse of a bijective homomorphism of groups or rings is again a homomorphism, and the inverse of a bijective linear map is again linear. In contrast:

The inverse of a continuous bijection need not be continuous.

It's easy to get caught out by this! Keep the following examples in mind.

Examples A4.9 i. We already saw in Example A4.3(iii) that the identity map

$$(\mathbb{R}, \text{discrete topology}) \to (\mathbb{R}, \text{standard topology})$$

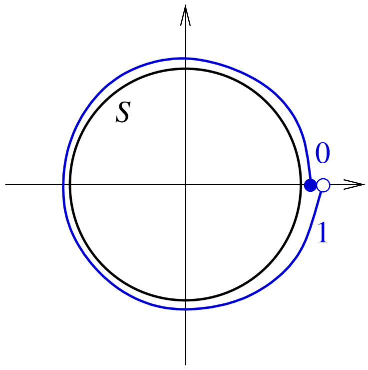


Figure A.5: A continuous bijection whose inverse is not continuous.

is continuous. However, its inverse is the identity map

$$(\mathbb{R}, \text{standard topology}) \to (\mathbb{R}, \text{discrete topology}),$$

which is not continuous: for instance, the subset $[0, 1)$ of $\mathbb{R}$ is open in the discrete topology but not in the standard topology.

ii. Here is a more geometrically intuitive example. Write

$$S = \{z \in \mathbb{C} : |z| = 1\}$$

(the unit circle). Give S the usual metric inherited from $\mathbb{C}$, that is, $d(w, z) = |w - z|$. Define $f\colon [0, 1) \to S$ by $f(t) = e^{2\pi it}$ (Figure A.5, which is not to be taken too literally; the lines are drawn slightly apart for visual clarity). Then f is a continuous bijection.

However, the inverse map $f^{-1}\colon S \to [0, 1)$ is not continuous. Intuitively, this is because it tears the circle at the point 1. Rigorously, if f^{-1} were continuous, then by Lemma A4.6, it would preserve convergence of sequences. But the sequence $(f(1 - 1/n))_{n=1}^{\infty}$ converges to 1 (since

$$f(1 - 1/n) = e^{2\pi i(1-1/n)} \to e^{2\pi i} = 1$$

as $n \to \infty$), whereas the sequence $(1 - 1/n)_{n=1}^{\infty}$ does not converge to $f^{-1}(1) = 0$.

Two spaces X and Y are supposed to be 'topologically the same' if we can deform X into Y and back again without tearing. How can we make this precise? In view of the last example, asking that there is a continuous bijection from X to Y is not the right thing to do, since although the bijection itself does not cause tearing (being continuous), its inverse might. What we should do is demand that not only the bijection, *but also its inverse*, is continuous.

Definition A4.10 Let X and Y be topological spaces.

i. A **homeomorphism** from X to Y is a continuous bijection whose inverse is also continuous.

ii. The spaces X and Y are **homeomorphic**, written $X \cong Y$, if there exists a homeomorphism from X to Y.

Note the 'e' in the words!
Here are some simple (non-)examples.

Examples A4.11 i. The maps in Examples A4.9 are continuous bijections but not homeomorphisms.

ii. Let $a, b \in \mathbb{R}$ with $a < b$. (In this course, subsets of $\mathbb{R}^n$ are always intended to be given the Euclidean metric unless otherwise indicated.) Define

$$f\colon [0,1] \to [a,b]$$

by

$$f(t) = (1-t)a + tb$$

($t \in [0,1]$). Then f is a continuous bijection. Its inverse is given by

$$f^{-1}(u) = \frac{u-a}{b-a}$$

($u \in [a,b]$), which is also continuous. Hence f is a homeomorphism, and the spaces $[0,1]$ and $[a,b]$ are homeomorphic.

The terminology 'X and Y are homeomorphic' introduced in Definition A4.10 would be highly misleading if it were not symmetric in X and Y (in other words, if it were possible that X and Y were homeomorphic but Y and X were not). Similarly, the notation $\cong$ would be misleading if $\cong$ were not an equivalence relation. We show that this terminology and notation do in fact behave sensibly.

Lemma A4.12 *i. Let X be a topological space. Then the identity map on X is a homeomorphism.*

ii. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be homeomorphisms. Then $g \circ f\colon X \to Z$ is a homeomorphism.

iii. Let $f\colon X \to Y$ be a homeomorphism. Then $f^{-1}\colon Y \to X$ is a homeomorphism.

Proof This follows from Lemma A4.8, using the facts that $\mathrm{id}_X^{-1} = \mathrm{id}_X$, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, and $(f^{-1})^{-1} = f$, for bijections f and g. □

Lemma A4.13 *Being homeomorphic is an equivalence relation on the class of all topological spaces.*

Proof Follows from Lemma A4.12. □

Example A4.14 In Example A4.11(ii), we showed that $[0,1]$ and $[a,b]$ are homeomorphic for all real $a < b$. Since $\cong$ is an equivalence relation, it follows that $[a,b] \cong [c,d]$ whenever $a < b$ and $c < d$.

In the next section, we will see some more substantial examples of homeomorphic spaces, and discuss the concept of a 'topological property'.

A5 When are two spaces homeomorphic?

For Lecture 6

We have defined what it means for two spaces to be homeomorphic. To a topologist, homeomorphic spaces are simply the same. The intuition is *roughly* that two spaces are homeomorphic if one can be deformed into the other by bending and reshaping, but without tearing or gluing. We'll see, though, that this description has to be taken with a pinch of salt.

Given two spaces, how can we decide whether they are homeomorphic? In general, there is no easy way. But we can begin to get a feel for it by working through some examples. This section will be about two different things:

- how to show that two spaces *are* homeomorphic;
- how to show that two spaces *are not* homeomorphic.

To show that two spaces *are* homeomorphic, we can simply write down a homeomorphism between them. We can also take advantage of the fact that being homeomorphic is an equivalence relation (Lemma A4.13).

Examples A5.1 Here we continue our analysis of homeomorphisms between intervals in $\mathbb{R}$, begun in Examples A4.11(ii) and A4.14.

i. We showed previously that $[a, b] \cong [c, d]$ whenever $a < b$ and $c < d$. Similarly,

$$\begin{aligned}(a, b) &\cong (c, d),\\ (a, \infty) &\cong (b, \infty) \cong (-\infty, b) \cong (-\infty, a),\\ [a, b) &\cong [c, d) \cong (c, d] \cong (a, b],\\ [a, \infty) &\cong [b, \infty) \cong (-\infty, b] \cong (-\infty, a],\end{aligned}$$

all via simple homeomorphisms of the form $x \mapsto rx + s$ for some $r, s \in \mathbb{R}$. (For instance, $[0, 1) \cong (0, 1]$ via the homeomorphism $x \mapsto 1 - x$.)

ii. There are some further, not so obvious, homeomorphisms between intervals. I claim that $(a, b) \cong \mathbb{R}$ whenever $a < b$. Since $(a, b) \cong (-1, 1)$, it is enough to prove that $(-1, 1) \cong \mathbb{R}$. Indeed, define $f\colon (-1, 1) \to \mathbb{R}$ by

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

($x \in (-1, 1)$). Then f is bijective, with inverse given by

$$f^{-1}(y) = \frac{y}{1 + |y|}$$

($y \in \mathbb{R}$); see Figure A.6. Both f and f^{-1} are continuous, so f is a homeomorphism. Hence $(-1, 1) \cong \mathbb{R}$.

The same formulas provide a homeomorphism between $[0, 1)$ and $[0, \infty)$; hence $[a, b) \cong [c, \infty)$ for all a, b, c with $a < b$. They also provide a homeomorphism between $(0, 1)$ and $(0, \infty)$; hence $(a, b) \cong (c, \infty)$ too.

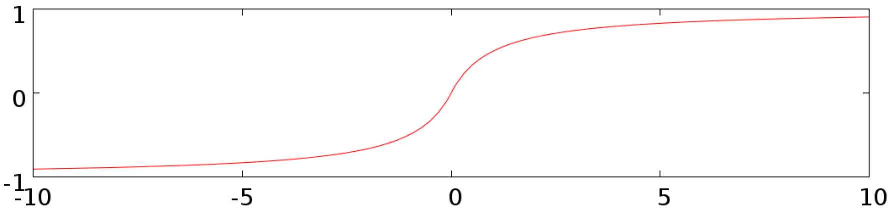
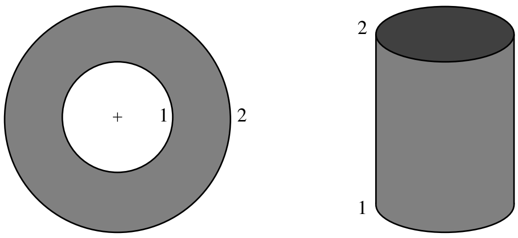


Figure A.6: The homeomorphism $f^{-1}\colon \mathbb{R} \to (-1, 1)$ of Example A5.1(ii). Axes are on different scales.

Bringing all this together: in the four lines of homeomorphisms in (i), all the intervals in the first two lines are homeomorphic to each other (and also homeomorphic to $(-\infty, \infty) = \mathbb{R}$), and all the intervals in the last two lines are homeomorphic to each other.

Example A5.2 Here is a more ambitious example. Consider an annulus and a cylinder:



The annulus includes its boundaries, and the cylinder is a hollow tube, with no lid at either end. Intuitively, we can deform the annulus into the cylinder by keeping the inner circle fixed and pulling the outer circle up towards us. So, they should be homeomorphic. Alternatively, imagine pressing your eye up to one end of the cylinder and looking through it. You would see an annulus. The projection of the cylinder onto your retina provides a homeomorphism.

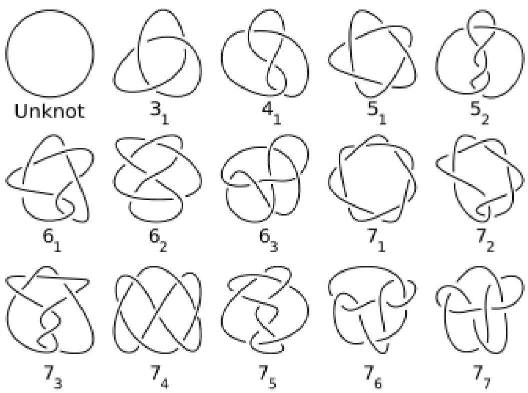
To make this precise, let us assume that the annulus has inner radius 1, outer radius 2, and is centred at the origin. Let us also assume that the cylinder has unit radius, that its bottom and top have z-coordinates 1 and 2, and that the central axis of the cylinder is the z-axis. (These choices make the calculations easier; different choices would give homeomorphic results.) Then there is a homeomorphism from the annulus to the cylinder given by

$$(r \cos\theta, r \sin\theta) \mapsto (\cos\theta, \sin\theta, r)$$

$(1 \leq r \leq 2,\ 0 \leq \theta < 2\pi)$.

Example A5.3 Similarly, the closed disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is homeomorphic to the closed square $[-1, 1] \times [-1, 1]$. Intuitively, the disk can be turned into the square by stretching. One can write down an exact formula for a homeomorphism, but it is not very illuminating.

Example A5.4 Any knot (in the mathematical sense) is homeomorphic to the circle. For example, all these knots are homeomorphic:



To see why, choose two knots, A and B, and let us suppose that they are both made of one metre of string. Imagine an ant crawling along A, starting at a point a_0, and a beetle crawling along B, starting at a point b_0. Assuming that the two insects crawl at the same speed, they will end up back where they started at the same time. There is a homeomorphism $f\colon A \to B$ defined by taking $f(a)$ to be the location of the beetle at the time when the ant is at $a \in A$. So, for instance, f matches up the point 10cm anticlockwise of a_0 on A with the point 10cm anticlockwise of b_0 on B.

The moral is that we have to be careful when thinking of homeomorphism as 'one space can be deformed into the other'. You might have thought that a (nontrivial) knot wouldn't be homeomorphic to the circle, since it cannot be unknotted. But when deciding what is homeomorphic to what, *we are not confined to* $\mathbb{R}^3$. If we can construct a homeomorphism then the spaces are homeomorphic, and that's that!

Example A5.5 Let S^n denote the n**-dimensional sphere**, defined by

$$S^n = \left\{ x \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} x_i^2 = 1 \right\}.$$

(You might have thought that this should be called $(n+1)$-dimensional, because it lives inside $\mathbb{R}^{n+1}$. But, for instance, the surface S^2 of the earth is best thought of as 2-dimensional, because any point on it can be specified by 2 coordinates, longitude and latitude.)

Let x be any point of S^n. Then $S^n \setminus \{x\} \cong \mathbb{R}^n$. For instance, the circle with a point removed is homeomorphic to $\mathbb{R}$ (or equivalently $(-1, 1)$), and the surface of the earth with the north pole removed is homeomorphic to $\mathbb{R}^2$. The most famous proof of this involves so-called stereographic projection.

Now we turn to the challenge of proving that two given spaces are *not* homeomorphic.

Example A5.6 Let us consider real intervals again. I claim that $[a, b]$ is not homeomorphic to $\mathbb{R}$ for any $a < b$. Suppose for a contradiction that $[a, b] \cong \mathbb{R}$. Then we may choose a homeomorphism $f\colon [a, b] \to \mathbb{R}$. By a basic theorem of analysis, f is bounded; that is, there exist $m, M \in \mathbb{R}$ such that $f[a,b] \subseteq [m,M]$. But f is a homeomorphism, and in particular a surjection, so $f[a,b] = \mathbb{R}$. This is a contradiction.

A similar argument shows that $[a,b]$ is not homeomorphic to $[c,d)$, for any $a < b$ and $c < d$.

Example A5.7 The interval $X = [0,1]$ is not homeomorphic to the union of intervals $Y = [3,5] \cup [10,13]$. (Here Y has the usual metric inherited from $\mathbb{R}$.) For suppose we have a homeomorphism $f\colon X \to Y$. Since f is surjective, there is some $c \in X$ such that $f(c) = 4$, and there is some $d \in X$ such that $f(d) = 11$. So f is a real-valued continuous function on $[0,1]$ taking the values 4 and 11. By the intermediate value theorem, f must somewhere take the value 6, a contradiction since f is a map into Y and $6 \notin Y$.

Less formally, the idea here is that X is in one piece and Y is in two. Homeomorphic spaces always have the same number of pieces, so X and Y are not homeomorphic.

Example A5.8 The letters T and U are not homeomorphic. Although we do not have the language to make this completely precise yet, the argument is as follows. The space T has a point with the property that when it is removed, what remains falls into three pieces. (This is the point where the vertical meets the horizontal.) If T and U were homeomorphic then U would have a point with this property too. But it does not: for when we remove either of the endpoints of U, what remains is in one piece, and when we remove any other point, what remains is in two pieces.

Example A5.9 The closed unit disk $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is not homeomorphic to the open unit disk $\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$. One argument for this would be to observe that the closed disk is compact and the open disk is not; we will come to compactness later. Another argument is that when we remove any point from the open disk, what remains has a hole in it; but there are certain points of the closed disk (the boundary points) which, when removed, leave a remainder that has no holes in it. If you take the Algebraic Topology course, you will learn that 'has a hole in it' can be made precise by 'has nontrivial fundamental group'.

Example A5.10 Consider the spaces $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \ldots$. Is it conceivable that $\mathbb{R}^m$ could be homeomorphic to $\mathbb{R}^n$ for some $m \neq n$?

It seems impossible that, for instance, $\mathbb{R}^2$ could be deformed into $\mathbb{R}^3$. But we should not dismiss the possibility too quickly. For a start, Cantor showed that for any $m, n \geq 1$, there is a *bijection* between $\mathbb{R}^m$ and $\mathbb{R}^n$. The bijection he constructed is not continuous—but still, this should make us pause for thought.

Furthermore, it is possible to construct a *continuous surjection* $\mathbb{R} \to \mathbb{R}^2$. (We will return to this when we do compactness, but if you're interested now, look up 'space-filling curves'.) Once you know this, it's easy to build a continuous surjection $\mathbb{R}^m \to \mathbb{R}^n$ for any $m, n \geq 1$.

But actually, our instinct about the original question is right: if $\mathbb{R}^m \cong \mathbb{R}^n$ then $m = n$. This is surprisingly hard to prove, so much so that we will not prove it in this course. Algebraic topology provides tools that make it easy.

A6 Topological properties

For Lecture 7; part one of two

These are not the positive integers:

1, 2, 3, 4, ...

Nor are these:

1, 2, 3, 4, ...

And nor are these:

I, II, III, IV, ...

The first are the Arabic numerals, the second are the Arabic numerals in a different typeface, and the third are the Roman numerals.

Behind this apparently pedantic distinction is an important mathematical point: *isomorphism is just renaming of elements.* All three number systems are isomorphic, which means that they are really the same, just with different names for their elements.

In many fields of mathematics there is a notion of isomorphism. For sets, an isomorphism is called a bijection. For groups and rings and vector spaces, an isomorphism is called an isomorphism. For metric spaces, an isomorphism is called an isometry. For topological spaces, an isomorphism is called a homeomorphism. In all these cases, an isomorphism is a bijection that respects all the structure: the multiplication in the case of groups, the distance in the case of metric spaces, and the topology (open sets) in the case of topological spaces. And in all cases, isomorphism can be viewed as simply renaming of elements.

We now focus on topological spaces. Let P be a true/false property that is defined for all topological spaces. We say that P is a **topological property** if whenever X and Y are homeomorphic spaces, X has property P if and only if Y has property P.

All sensible properties of topological spaces are topological properties. In other words, no sensible property of topological spaces depends on what the elements happen to be called.

Examples A6.1 Being T_1 is a topological property, as are being Hausdorff, or discrete, or indiscrete, or metrizable. Compactness and connectedness are topological properties too, which we will study later.

A more complicated topological property: let us say that a topological space X is 'purple' if there exists some $x \in X$ such that $X \setminus \{x\}$ is connected. Then purpleness is a topological property.

Example A6.2 Being a subset of $\mathbb{R}$ is *not* a topological property, as it depends on what the elements happen to be called. For instance, $\mathbb{R}$ is homeomorphic to the subset $L = \{(x, y) \in \mathbb{R}^2 : y = 3\}$ of $\mathbb{R}^2$, and $\mathbb{R}$ is a subset of $\mathbb{R}$ but L is not.

Example A6.3 Being bounded is *not* a topological property, for two reasons.

First, it is not a property of topological spaces, since it does not *make sense* to ask whether a topological space $(X, \mathcal{T})$ is bounded: you need a metric on X.

Second, it is not even true that if X and Y are homeomorphic metric spaces then X is bounded if and only if Y is bounded. For instance, the real interval $(0,1)$ is a bounded metric space homeomorphic to the unbounded metric space $\mathbb{R}$, by Example A5.1(ii).

I have argued that being Hausdorff, T_1, etc., *should* be topological properties. But we haven't actually *proved* it. Here is the proof for Hausdorffness. The proofs for the other properties are similar.

We will use the fact that for a homeomorphism $f\colon X \to Y$, a subset $U \subseteq X$ is open if and only if $fU \subseteq Y$ is open. (This is false for an arbitrary continuous map, as we saw in Remark A4.2(ii).) You should check this!

Lemma A6.4 *Hausdorffness is a topological property.*

Proof Let X and Y be homeomorphic topological spaces with X Hausdorff; we must prove that Y is Hausdorff. Choose a homeomorphism $f\colon X \to Y$. Let y and y' be distinct points of Y. Then $f^{-1}(y)$ and $f^{-1}(y')$ are distinct points of X. Since X is Hausdorff, we can choose disjoint open neighbourhoods U of $f^{-1}(y)$ and U' of $f^{-1}(y')$. Then fU and fU' are disjoint open neighbourhoods of y and y' respectively, since f is a homeomorphism. $\square$

Suppose we have two topological spaces X and Y in front of us, and want to show that they are *not* homeomorphic. It may be possible to argue directly that there is no homeomorphism from X to Y, as we did in Example A5.6 to show that $[0,1] \not\cong \mathbb{R}$. But more often, we do it by finding some topological property satisfied by X but not Y or vice versa.

For instance, the closed unit disk in $\mathbb{R}^2$ is not homeomorphic to the open unit disk, because the first is compact but the second is not (Example A5.9). Another example: the space $\mathbb{R}^2$ is 'purple' in the sense of Examples A6.1, but the space $\mathbb{R}$ is not; hence $\mathbb{R}^2 \not\cong \mathbb{R}$. This strategy also shows that $\mathbb{R}^n \not\cong \mathbb{R}$ for all $n > 1$. But it does not show, for instance, that $\mathbb{R}^3 \not\cong \mathbb{R}^2$, since both are purple. To prove that $\mathbb{R}^3 \not\cong \mathbb{R}^2$, a different strategy is needed.

More general than the idea of a topological property is the idea of a topological invariant. A **topological invariant** is a function assigning to each topological space X a mathematical object $I(X)$ of some kind, in such a way that that if X and Y are homeomorphic then $I(X)$ and $I(Y)$ are 'the same' (in whatever sense is appropriate for the mathematical objects in question). This is not quite a precise mathematical definition, but some examples should convey the idea:

Examples A6.5

i. We can assign to any topological space X the set $K(X)$ of connected-components of X (defined later). Intuitively, $K(X)$ is the set of 'pieces' of X. For instance, if $X = [3,5] \cup [10,13]$ then $K(X)$ is a two-element set.

   If there is a homeomorphism between two spaces X and Y then there is a bijection between their sets $K(X)$ and $K(Y)$ of connected-components. Since the existence of a bijection is the appropriate notion of 'sameness' of sets, this means that K is a topological invariant.

ii. Rather trivially, the set of points of a topological space is a topological invariant. In other words, if there is a homeomorphism between topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ then there is a bijection between the sets X and Y. This is immediate from the definition of homeomorphism.

iii. You will meet more interesting examples of topological invariants if you take Algebraic Topology. For instance, one can assign to every space X a certain group $\pi_1(X)$, called the fundamental group or first homotopy group of X. It measures the number of holes in X. If the spaces X and Y are homeomorphic then the groups $\pi_1(X)$ and $\pi_1(Y)$ are isomorphic, so π_1 is a topological invariant. (I am brushing under the carpet a detail concerning basepoints.)

iv. A topological property can be seen as a topological invariant taking values in the two-element set $\{\texttt{true}, \texttt{false}\}$. For instance, define $H(X) = \texttt{true}$ if X is Hausdorff and $H(X) = \texttt{false}$ if not. Then H defines a function from the class of all spaces to the set $\{\texttt{true}, \texttt{false}\}$, and it has the property that if X and Y are homeomorphic then $H(X)$ and $H(Y)$ are equal. In this way, the topological property of Hausdorffness can be seen as a topological invariant H.

Like topological properties, topological invariants are very useful for telling spaces apart. For example, the spaces $X = [3,5] \cup [10,13]$ and $Y = [1,2] \cup [3,4] \cup [5,6]$ are not homeomorphic, because if they were then there would be a bijection between the sets $K(X)$ and $K(Y)$, which is impossible because $K(X)$ has 2 elements and $K(Y)$ has 3. Briefly put: X and Y are not homeomorphic because they have different numbers of connected-components.

A7 Bases

For Lecture 7; part two of two

When we reason about metric spaces, it is often natural and convenient to use the open balls rather than arbitrary open sets. In a general topological space, we do not have balls available to us, but there may sometimes be a special collection of open sets with similar properties to those possessed by the open balls in a metric space. Such a collection is called a 'basis' (plural: 'bases').

Definition A7.1 Let X be a topological space. A **basis** for X is a set $\mathcal{B}$ of open subsets of X, such that every open subset of X is a union of sets in $\mathcal{B}$.

That is, a set $\mathcal{B}$ of open sets is a basis if for an arbitrary open $U \subseteq X$, we can find some family $(B_i)_{i \in I}$ with $B_i \in \mathcal{B}$ for all $i \in I$ and $\bigcup_{i \in I} B_i = U$.

Examples A7.2 i. Let X be a metric space. Let

$$\mathcal{B} = \{B(x,r) : x \in X, r > 0\}.$$

Then $\mathcal{B}$ is a basis for the induced topology on X. Indeed, let U be an open subset of X. For each $x \in U$, we can choose $r_x > 0$ such that $B(x, r_x) \subseteq U$. Then $\bigcup_{x \in U} B(x, r_x) = U$.

ii. The set

$$\mathcal{B} = \{(a,b) \times (c,d) : a,b,c,d \in \mathbb{R},\ a < b,\ c < d\}$$

is a basis for the standard topology on $\mathbb{R}^2$. Why? First, every element of $\mathcal{B}$ is certainly open. Second, among the elements of $\mathcal{B}$ are the open balls for the metric d_∞ on $\mathbb{R}^2$, which induces the standard topology (Example A3.4(ii)). Hence every subset of $\mathbb{R}^2$ that is open in the standard topology is a union of elements of $\mathcal{B}$.

This example shows that a topological space can have several different bases, since as well as this basis $\mathcal{B}$ for the standard topology on $\mathbb{R}^2$, we have the collection of Euclidean open disks (i.e. open balls with respect to d_2), which by (i) is also a basis for the standard topology.

The moral: you should never speak of 'the' basis of a topological space, any more than you should speak of 'the' basis of a vector space. (Actually, topological bases are more like spanning sets of vector spaces than bases of vector spaces, since if you take a topological basis and add some more open sets to it, it is still a basis. But one should not speak of 'the' spanning set of a vector space either.)

iii. In a discrete metric space, the collection of one-element subsets is a basis, since first, they are all open, and second, an arbitrary open subset is the union of its one-element subsets.

Here is a useful fact about bases.

Lemma A7.3 *Let $f\colon X \to Y$ be a function, not necessarily continuous, between topological spaces. Let $\mathcal{B}$ be a basis for Y, and suppose that $f^{-1}B$ is open in X for all $B \in \mathcal{B}$. Then f is continuous.*

Proof Let U be an open subset of Y. Then $U = \bigcup_{i\in I} B_i$ for some family $(B_i)_{i\in I}$ of elements of $\mathcal{B}$. Now

$$f^{-1}U = f^{-1}\left(\bigcup_{i\in I} B_i\right) = \bigcup_{i\in I} f^{-1}B_i,$$

which is a union of open subsets and therefore open. $\square$

So, one strategy for proving that a map is continuous is to find a convenient basis for the topology on the codomain, then prove that the preimage of any *basic* open set is open. (Compare (i)$\Longrightarrow$(ii) of Lemma A1.14, where we were implicitly using the open balls of Y as a basis for its topology.)

There is another way that bases are used, and that is to *specify* a topology. For instance, we might want to say something like 'define a topology on $\mathbb{R}^2$ by declaring every subset $(a, b) \times (c, d)$ to be open, then throwing in all the other open sets you need in order for the axioms for a topology to be satisfied'. (This should give the standard topology.) Note that in the following definition, X is only a *set*, not a topological space.

Definition A7.4 Let X be a set. A **synthetic basis** on X is a collection $\mathcal{B}$ of subsets of X, such that:

- the union of all the sets in $\mathcal{B}$ is X;
- whenever $B, B' \in \mathcal{B}$, then $B \cap B'$ is a union of sets in $\mathcal{B}$.

Warning A7.5 No one actually says 'synthetic basis'. That's just terminology we'll use for the next few paragraphs. In real life, everyone just says 'basis'. We now show how this fits with the meaning of 'basis' given in Definition A7.1.

Let X be a set and $\mathcal{B}$ a synthetic basis on X. The **topology generated by** $\mathcal{B}$ is the set of subsets U of X such that $U = \bigcup_{i \in I} B_i$ for some family $(B_i)_{i \in I}$ of elements of $\mathcal{B}$. Of course, simply calling it a 'topology' does not make it one! We have to check this:

Lemma A7.6 *Let X be a set and $\mathcal{B}$ a synthetic basis on X. Then the topology generated by $\mathcal{B}$ is indeed a topology. Moreover, it is the unique topology $\mathcal{T}$ on X such that $\mathcal{B}$ is a basis for $\mathcal{T}$.*

Proof Write $\mathcal{T}_{\mathcal{B}}$ for the topology generated by $\mathcal{B}$. We first prove that $\mathcal{T}_{\mathcal{B}}$ is a topology.

- Certainly $\emptyset \in \mathcal{T}_{\mathcal{B}}$, as it is the union of the empty family of elements of $\mathcal{B}$. Also, $X \in \mathcal{T}_{\mathcal{B}}$ by the first axiom for synthetic bases.
- Now let $(U_i)_{i \in I}$ be a family of elements of $\mathcal{T}_{\mathcal{B}}$. For each $i \in I$, the set U_i can be expressed as a union of elements of $\mathcal{B}$, so their union $\bigcup_{i \in I} U_i$ is also a union of elements of $\mathcal{B}$.
- Let $U, W \in \mathcal{T}_{\mathcal{B}}$. Then $U = \bigcup_{i \in I} B_i$ and $W = \bigcup_{j \in J} C_j$ for some families $(B_i)_{i \in I}$ and $(C_j)_{j \in J}$ of elements of $\mathcal{B}$. But then

  $$U \cap W = \bigcup_{i \in I, j \in J} B_i \cap C_j,$$

  and each of the sets $B_i \cap C_j$ is a union of elements of $\mathcal{B}$ (by the second axiom on synthetic bases), so $U \cap W$ is a union of elements of $\mathcal{B}$.

Finally, we must show that $\mathcal{T}_{\mathcal{B}}$ is the unique topology $\mathcal{T}$ on X such that $\mathcal{B}$ is a basis for $\mathcal{T}$. There are two parts to this statement: (i) that $\mathcal{B}$ *is* a basis for the topology $\mathcal{T}_{\mathcal{B}}$, and (ii) that $\mathcal{T}_{\mathcal{B}}$ is the *only* topology on X with this property. Part (i) is trivial, since by definition, every element of $\mathcal{T}_{\mathcal{B}}$ is a union of elements of $\mathcal{B}$. For (ii), let $\mathcal{T}$ be a topology on X such that $\mathcal{B}$ is a basis for $\mathcal{T}$. Then every element of $\mathcal{B}$ belongs to $\mathcal{T}$, so (by definition of topology) every union of elements of $\mathcal{B}$ belongs to $\mathcal{T}$, or equivalently $\mathcal{T}_{\mathcal{B}} \subseteq \mathcal{T}$. On the other hand, every element of $\mathcal{T}$ is a union of elements of $\mathcal{B}$, or equivalently $\mathcal{T} \subseteq \mathcal{T}_{\mathcal{B}}$. So $\mathcal{T} = \mathcal{T}_{\mathcal{B}}$, as required. $\square$

Because of Lemma A7.6, it is safe to say 'basis' instead of 'synthetic basis'. We will always do this from now on.

The word 'synthetic' is inspired by chemistry. In synthetic chemistry, you start with simple molecules and put them together to build more complex ones. When you define a topology from a synthetic basis, you start with simple (basic) open sets and put them together to build more complex open sets.

The bases defined in Definition A7.1 might also have been called 'analytic bases'. In analytical chemistry, you are given a complex molecule and try to decompose it into simpler molecules. Similarly, given a topology, we can try to decompose its open sets into simpler (basic) open sets—much as in a metric space, it is useful to see open sets as unions of open balls.

A8 Closure and interior

For Lecture 8

If I handed you an interval and told you 'Make it closed!', you'd know what to do. You'd turn (a, b) into $[a, b]$, and $[a, b)$ into $[a, b]$, and (a, ∞) into $[a, \infty)$. In short, you'd add endpoints wherever they're absent. And in the same way, if I told you to 'make an interval open', you'd remove endpoints wherever they were present.

One dimension up, if I handed you an open disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ and said 'Make it closed!', you'd turn it into the closed disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$; and similarly if I told you to make the closed disk open.

The point of this section is that the process of 'making a subset closed' (or open) is meaningful not only for familiar subsets of $\mathbb{R}^n$, but for completely arbitrary subsets of a completely arbitrary topological space. So, given a topological space X, we will define for each subset $A \subseteq X$ an associated closed set $\mathrm{Cl}(A)$, and also an associated open set $\mathrm{Int}(A)$.

We begin with the process of making a subset closed.

Definition A8.1 Let X be a topological space and $A \subseteq X$. The **closure** $\mathrm{Cl}(A)$ of A is the intersection of all the closed subsets of X that contain A.

In some texts, the closure of A is written as $\overline{A}$.

Lemma A8.2 *Let X be a topological space and $A \subseteq X$. Then:*

i. $\mathrm{Cl}(A)$ *is a closed subset of X containing A;*

ii. $\mathrm{Cl}(A) \subseteq V$ *for any closed subset V of X containing A.*

Proof Since $\mathrm{Cl}(A)$ is an intersection of closed sets in X, it is itself closed in X, and since it is an intersection of sets containing A, it itself contains A. On the other hand, let V be any closed subset of X containing A; then V is one of the sets in the intersection defining $\mathrm{Cl}(A)$, so $\mathrm{Cl}(A) \subseteq V$. $\square$

Remark A8.3 Lemma A8.2 is often phrased as '$\mathrm{Cl}(A)$ is the smallest closed set containing A'. This property of $\mathrm{Cl}(A)$ characterizes $\mathrm{Cl}(A)$ uniquely. In other words, if K is a subset of X such that (i) K is a closed subset of X containing A, and (ii) $K \subseteq V$ for any closed subset V of X containing A, then $K = \mathrm{Cl}(A)$. To see this, note that condition (i) for $\mathrm{Cl}(A)$ and condition (ii) for K together imply that $K \subseteq \mathrm{Cl}(A)$, and similarly the other way round.

Here are some further basic properties of closure.

Lemma A8.4 *Let X be a topological space and $A, B \subseteq X$. Then:*

i. $\mathrm{Cl}(A) = A \iff A$ *is closed;*

ii. $\mathrm{Cl}(\mathrm{Cl}(A)) = \mathrm{Cl}(A)$*;*

iii. if $A \subseteq B$ *then* $\mathrm{Cl}(A) \subseteq \mathrm{Cl}(B)$.

Proof For (i), $\Longrightarrow$ holds because $\mathrm{Cl}(A)$ is closed; $\Longleftarrow$ holds by definition of $\mathrm{Cl}(A)$, because if A is closed then A is a closed subset of X containing A.

Part (ii) follows from part (i), because $\mathrm{Cl}(A)$ is closed.

For (iii), suppose that $A \subseteq B$. Then $\mathrm{Cl}(B)$ is a closed subset of X containing A, so $\mathrm{Cl}(B) \supseteq \mathrm{Cl}(A)$ by definition of closure. $\square$

The definition of $\mathrm{Cl}(A)$ is in a sense unhelpful: given a specific point $x \in X$, it does not help us decide whether or not $x \in \mathrm{Cl}(A)$. The following lemma is much more useful in that respect.

Lemma A8.5 *Let X be a topological space and $A \subseteq X$. Then*

$$\mathrm{Cl}(A) = \{x \in X : \textit{every neighbourhood of } x \textit{ meets } A\}.$$

(Given two subsets A and B of a set X, we say that A **meets** B if $A \cap B \neq \emptyset$.)

Proof Let $K = \{x \in X :$ every neighbourhood of x meets $A\}$. In order to show that $K = \mathrm{Cl}(A)$, it is enough (by Remark A8.3) to prove that K is the smallest closed subset of X containing A: that is, (i) K is a closed subset of X containing A, and (ii) $K \subseteq V$ for any closed subset V of X containing A.

For (i), certainly K contains A. To show that K is closed in X, we show that $X \setminus K$ is open using Lemma A2.9. Let $y \in X \setminus K$. By definition of K, there exists an open neighbourhood W of y that does not meet A. No element w of W is in K, since the set W is a neighbourhood of w not meeting A. So W is an open neighbourhood of y contained in $X \setminus K$. Lemma A2.9 then implies that $X \setminus K$ is open in X.

For (ii), let V be any closed subset of X containing A; we show that $K \subseteq V$. Let $x \in K$. If $x \notin V$ then $X \setminus V$ is a neighbourhood of x not meeting A, so $x \notin K$, a contradiction. Hence $K \subseteq V$, as required. $\square$

To better understand the meaning of this lemma, it is useful to make another definition.

Definition A8.6 Let X be a topological space and $A \subseteq X$. A **limit point** of A is a point $x \in X$ such that every neighbourhood of x contains some point of A not equal to x.

Lemma A8.7 *Let X be a topological space and $A \subseteq X$. Then*

$$\mathrm{Cl}(A) = A \cup \{\textit{limit points of } A\}.$$

Proof Follows from Lemma A8.5. $\square$

Examples A8.8

i. In a *metric* space X, a point $x \in X$ is a limit point of $A \subseteq X$ if and only if for all $\varepsilon > 0$, the punctured ball $B(x, \varepsilon) \setminus \{x\}$ meets A. Equivalently, x is a limit point of A if and only if x can be expressed as a limit of some sequence in $A \setminus \{x\}$. The closure of A consists of all elements of X that can be expressed as a limit of some sequence in A.

ii. For instance, let $X = \mathbb{R}$. Then $\mathrm{Cl}((0,1)) = \mathrm{Cl}([0,1)) = \mathrm{Cl}([0,1]) = [0,1]$.

iii. Again in $\mathbb{R}$, the set of limit points of $\{0\} \cup [1,2)$ is $[1,2]$. This tells us that *a point of A may or may not be a limit point of A, and a limit point of A may or may not be a point of A.*

   The closure of $A = \{0\} \cup [1,2)$ in $\mathbb{R}$ consists of the points of $\mathbb{R}$ that are *either* in A *or* limit points of A; thus, $\mathrm{Cl}(A) = \{0\} \cup [1,2]$.

iv. In $\mathbb{R}$ once again, take a nonempty subset A that is bounded above, so that $\sup A$ exists. Then $\sup A \in \mathrm{Cl}(A)$ by (i), since $\sup A$ can be expressed as a limit of a sequence in A.

Before we go any further, we need to recall a little set theory.

Lemma A8.9 *Let X and Y be sets and let $f\colon X \to Y$ be a function.*

i. For $A \subseteq X$ and $B \subseteq Y$, $fA \subseteq B \iff A \subseteq f^{-1}B$.

ii. $f^{-1}fA \supseteq A$ for all $A \subseteq X$, and $ff^{-1}B \subseteq B$ for all $B \subseteq Y$.

Proof Part (i) is just the observation that for $a \in A$, we have $fA \subseteq B$ if and only if $f(a) \in B$ for all $a \in A$, if and only if $A \subseteq f^{-1}B$. Part (ii) follows, first by putting $B = fA$ and then by putting $A = f^{-1}B$. $\square$

We can now rephrase continuity in terms of closure (and *direct* images, not *inverse* images!) The rough idea is this. The closure of a set $A \subseteq X$ consists of A together with the points just outside A. For a function $f\colon X \to Y$ to be continuous should mean that f maps points only just outside A to points only just outside fA. And indeed:

Proposition A8.10 *Let $f\colon X \to Y$ be a function between topological spaces. Then f is continuous $\iff$ $f(\mathrm{Cl}(A)) \subseteq \mathrm{Cl}(fA)$ for all $A \subseteq X$.*

Proof Suppose that f is continuous, and let $A \subseteq X$. Then $\mathrm{Cl}(fA)$ is closed in Y, so by Lemma A4.4, $f^{-1}\,\mathrm{Cl}(fA)$ is closed in X. Now $f^{-1}\,\mathrm{Cl}(fA)$ contains $f^{-1}fA$ (by Lemma A8.2), which in turn contains A (by Lemma A8.9(ii)), so $f^{-1}\,\mathrm{Cl}(fA)$ is a closed subset of X containing A. Hence $\mathrm{Cl}(A) \subseteq f^{-1}\,\mathrm{Cl}(fA)$. Lemma A8.9(i) then gives $f\,\mathrm{Cl}(A) \subseteq \mathrm{Cl}(fA)$.

Conversely, suppose that $f(\mathrm{Cl}(A)) \subseteq \mathrm{Cl}(fA)$ for all $A \subseteq X$. By Lemma A4.4, it is enough to show that the preimage under f of a closed set is closed; so let $V \subseteq Y$ be closed. Then

$$f\,\mathrm{Cl}(f^{-1}V) \subseteq \mathrm{Cl}(ff^{-1}V) \subseteq \mathrm{Cl}(V) = V,$$

using our hypothesis in the first step (taking 'A' to be $f^{-1}V$) and Lemmas A8.9(ii) and A8.4(iii) in the second. Hence $\mathrm{Cl}(f^{-1}V) \subseteq f^{-1}V$, by Lemma A8.9(i). But also $\mathrm{Cl}(f^{-1}V)$ contains $f^{-1}V$, so they are equal. It then follows from Lemma A8.4(i) that $f^{-1}V$ is closed. $\square$

Some subsets A of a space X are so big that every point of X is either in A or a limit point of A:

Definition A8.11 Let X be a topological space. A subset $A \subseteq X$ is **dense** in X if $\mathrm{Cl}(A) = X$.

For example, $\mathbb{Q}$ is dense in $\mathbb{R}$ (with the usual topology).

Lemma A8.12 *Let $f, g\colon X \to Y$ be continuous maps of topological spaces, with Y Hausdorff. Then $\{x \in X : f(x) = g(x)\}$ is closed in X.*

Proof Exercise. $\square$

Corollary A8.13 *Let $f, g\colon X \to Y$ be continuous maps of topological spaces, with Y Hausdorff. Suppose there exists a dense subset $A \subseteq X$ such that $f(a) = g(a)$ for all $a \in A$. Then $f = g$.*

Proof The set $E = \{x \in X : f(x) = g(x)\}$ is closed in X and contains A, so $E \supseteq \mathrm{Cl}(A)$. But $\mathrm{Cl}(A) = X$, so $E = X$. $\square$

For example, two continuous functions $\mathbb{R} \to \mathbb{R}$ that agree on all rational numbers must in fact be equal.

Finally, we consider the mirror image of the closure operator: the interior operator. What we have proved about closures enables us to establish the basic properties of interiors very quickly.

Definition A8.14 Let X be a topological space and $A \subseteq X$. The **interior** $\mathrm{Int}(A)$ of A is the union of all the open subsets of X contained in A.

In some texts, the interior of A is written as A°.

Lemma A8.15 *Let X be a topological space and $A \subseteq X$. Then*

$$\mathrm{Cl}(X \setminus A) = X \setminus \mathrm{Int}(A), \qquad \mathrm{Int}(X \setminus A) = X \setminus \mathrm{Cl}(A).$$

Proof We have $\mathrm{Int}(A) = \bigcup_{\text{open } U \subseteq X \,:\, U \subseteq A} U$, so

$$X \setminus \mathrm{Int}(A) = \bigcap_{\text{open } U \subseteq X \,:\, U \subseteq A} X \setminus U = \bigcap_{\text{closed } V \subseteq X \,:\, V \supseteq X \setminus A} V = \mathrm{Cl}(X \setminus A).$$

The second identity follows by changing A to $X \setminus A$ in the first. $\square$

Lemma A8.16 *Let X be a topological space and $A \subseteq X$. Then $\mathrm{Int}(A)$ is an open subset of X contained in A. Moreoever, if U is any open subset of X contained in A, then $U \subseteq \mathrm{Int}(A)$.*

Less formally, $\mathrm{Int}(A)$ is the largest open subset of X contained in A.

Proof This is proved just as for Lemma A8.2. $\square$

Lemma A8.17 *Let X be a topological space and $A \subseteq X$. Then*

$$\mathrm{Int}(A) = \{x \in X : \textit{ some neighbourhood of } x \textit{ is contained in } A\}.$$

Proof This follows from Lemma A8.5, using Lemma A8.15. Explicitly,

$$\begin{aligned}\mathrm{Int}(A) &= X \setminus \mathrm{Cl}(X \setminus A)\\ &= \{x \in X : \text{not every neighbourhood of } x \text{ meets } X \setminus A\}\\ &= \{x \in X : \text{some neighbourhood of } x \text{ is contained in } A\}.\end{aligned}$$ $\square$

Warning A8.18 Closure and interior are not opposite processes. For example, $\mathrm{Cl}(\mathrm{Int}(A))$ is not in general equal to A, even if A is closed. Consider $X = \mathbb{R}$ and $A = \{0\}$; then $\mathrm{Int}(A) = \emptyset$, so $\mathrm{Cl}(\mathrm{Int}(A)) = \emptyset \neq A$.

A9 Subspaces (new spaces from old, 1)

For Lecture 9

The next three lectures are about three ways of constructing new topological spaces. We begin with subspaces.

Given a topological space X and a subset A of X, is there a sensible way of putting a topology on A? Questions like this are hard to answer in the abstract. It helps to start by examining a familiar related situation: that of metric spaces.

Given a metric space X, every subset A of X can be viewed as a metric space in its own right. For example, once we've defined a metric on $\mathbb{R}^n$, we automatically get a metric on any subset $A \subseteq \mathbb{R}^n$, simply by defining the distance between two points of A to be the same as the distance between them in $\mathbb{R}^n$. This process is so obvious and trivial that we hardly think about it. Nevertheless, it is useful to give it a name.

Definition A9.1 Let (X, d) be a metric space. Let $A \subseteq X$. The **subspace metric** on A is the function $d_A \colon A \times A \to [0, \infty)$ defined by $d_A(a, b) = d(a, b)$ for all $a, b \in A$. The metric space (A, d_A) is a **subspace** of the metric space (X, d).

It is very easy to check that the subspace metric is indeed a metric!

Trivial as this definition may seem, it has some consequences that may not be entirely obvious. To explain them, it will help to use a refined notation for balls. For a metric space $X = (X, d)$, a point $x \in X$, and $r > 0$, let us write

$$B_X(x, r) = \{y \in X : d(x, y) < r\},$$

which we would normally write as just $B(x, r)$.

Lemma A9.2 *Let X be a metric space and let A be a subspace of X. Then for all $a \in A$ and $r > 0$,*

$$B_A(a, r) = B_X(a, r) \cap A.$$

Proof Write the metric on X as d and the subspace metric on A as d_A. Then

$$\begin{aligned} B_A(a, r) &= \{b \in A : d_A(a, b) < r\} \\ &= \{b \in A : d(a, b) < r\} \\ &= B_X(a, r) \cap A. \end{aligned}$$ □

Now consider, for instance, the metric space $X = \mathbb{R}$. Let $A = [0, \infty) \subseteq \mathbb{R}$, and give A the subspace metric. Then

$$B_A(0, 1) = B_{\mathbb{R}}(0, 1) \cap A = (-1, 1) \cap [0, \infty) = [0, 1).$$

So $[0, 1)$ is an open ball in the metric space $[0, \infty)$. In particular, it is an open subset of the metric space $[0, \infty)$, even though it is not an open subset of $\mathbb{R}$.

This shows that when A is a subspace of a metric space X,

> *the open subsets of the metric space A are* not *simply the subsets of A that are open in X.*

In fact, the situation is as follows.

Lemma A9.3 *Let X be a metric space and $A \subseteq X$, and give A the subspace metric. Let $U \subseteq A$. Then U is open in the metric space A if and only if $U = W \cap A$ for some open subset W of X.*

Proof Suppose that U is open in the metric space A. For each $u \in U$, we may choose $r_u > 0$ such that $B_A(u, r_u) \subseteq U$. Put $W = \bigcup_{u \in U} B_X(u, r_u)$. Then W is open in X, and

$$W \cap A = \left(\bigcup_{u \in U} B_X(u, r_u) \right) \cap A = \bigcup_{u \in U} \big(B_X(u, r_u) \cap A \big) = \bigcup_{u \in U} B_A(u, r_u) = U,$$

using Lemma A9.2.

Conversely, suppose that $U = W \cap A$ for some open $W \subseteq X$. Let $u \in U$. Since $u \in W$ and W is open in X, we may choose $r > 0$ such that $B_X(u, r) \subseteq W$. But then

$$B_A(u, r) = B_X(u, r) \cap A \subseteq W \cap A = U$$

(using Lemma A9.2 again). Hence U is open in A. $\square$

Now imagine we are given a *topological* space X and a subset A. Lemma A9.3 strongly suggests how we should define a topology on A, namely:

Definition A9.4 Let $X = (X, \mathcal{T})$ be a topological space and $A \subseteq X$. The **subspace topology** $\mathcal{T}_A$ on A is defined by

$$\mathcal{T}_A = \{U \subseteq A : U = W \cap A \text{ for some } W \in \mathcal{T}\}.$$

(In words: a subset U of A is defined to be open in the subspace topology on A if and only if $U = W \cap A$ for some open subset W of X.)

The subspace topology really is a topology. (Check!) In the situation of the definition, we call $(A, \mathcal{T}_A)$ a **subspace** of $(X, \mathcal{T})$. So when we say 'let A be a subspace of X', this means 'let A be a subset of X, and give A the subspace topology'.

Remark A9.5 *Unless indicated otherwise, subsets of a topological space will always be given the subspace topology.* For instance, given a space X and $U \subseteq A \subseteq X$, we say that U is **open in** A to mean that U is open in the subspace topology on A. Similarly, if we have a space X and $A \subseteq X$, and we refer to 'continuous maps from A to $\mathbb{R}$', then it is implicitly intended that the continuity is with respect to the subspace topology on A.

This silent convention would be dangerous were it not for two facts. First, the subspace topology on a space X regarded as a subset of itself is the same as the original topology. Second, for a space X and $B \subseteq A \subseteq X$, we can either give A the subspace topology from X and then give B the subspace topology from A, or we can give B the subspace topology directly from X; and the fact is that these two topologies are the same. The proofs of these statements are left as exercises.

As usual, we have been thinking mostly about open rather than closed sets. But the situation for closed sets is similar:

Lemma A9.6 *Let X be a topological space and let A be a subspace of X. Let $V \subseteq A$. Then V is closed in A if and only if $V = S \cap A$ for some closed subset S of X.*

Proof

$$\begin{aligned}
V \text{ is closed in } A &\iff A \setminus V \text{ is open in } A \\
&\iff A \setminus V = W \cap A \text{ for some open subset } W \text{ of } X \\
&\qquad \text{(by definition of the subspace topology)} \\
&\iff A \setminus V = (X \setminus S) \cap A \text{ for some closed subset } S \text{ of } X \\
&\iff A \setminus V = A \setminus S \text{ for some closed subset } S \text{ of } X \\
&\iff V = A \setminus (A \setminus S) \text{ for some closed subset } S \text{ of } X \\
&\iff V = S \cap A \text{ for some closed subset } S \text{ of } X. \qquad \square
\end{aligned}$$

Lemma A9.7 *Let X be a topological space. Given an* open *subset A of X, a subset of A is open in the subspace topology on A if and only if it is open in X. Similarly, given a* closed *subset A of X, a subset of A is closed in the subspace topology on A if and only if it is closed in X.*

Proof The first statement follows from the definition of subspace topology, using the fact that the intersection of two open sets is open. The second follows similarly from Lemma A9.6. □

Warning A9.8 This result would fail without the hypothesis that A is open in X (in the first statement) or closed in X (in the second). For an arbitrary subspace A of a topological space, it is *not* true that if U is open in A then U is open in X. To see this, consider the case $U = A$ or the example $[0, 1) \subseteq [0, \infty) \subseteq \mathbb{R}$ mentioned above.

We justified the definition of topological subspace by considering metric spaces. But there is another kind of justification. Roughly speaking, it is that the subspace topology behaves as we would wish with respect to the notion of continuity.

Remark A9.9 Let us recall a little more set theory. Given any set X and subset $A \subseteq X$, there is an **inclusion function** $i\colon A \to X$ defined by $i(a) = a$ for all $a \in A$. For $W \subseteq X$, we have

$$i^{-1}W = \{a \in A : i(a) \in W\} = W \cap A.$$

Lemma A9.10 *Let X be a topological space and let A be a subspace of X. Then the inclusion function $i\colon A \to X$ is continuous.*

Proof Let W be an open subset of X. Then $i^{-1}W = W \cap A$ is an open subset of A, by definition of subspace topology. □

The subspace topology is defined in such a way that all the preimages $i^{-1}W$ of open sets are open, but nothing else is open. In other words, the subspace topology is the smallest (coarsest) topology on A such that i is continuous.

But that is not the only good continuity property of the subspace topology.

Consider, for instance, the map $f\colon \mathbb{R} \to [-1,1]$ defined by $f(x) = \sin x$. Is it continuous? (Here we have implicitly given $\mathbb{R}$ its standard topology and $[-1,1]$ the subspace topology from $\mathbb{R}$.) It would be nice if we could say 'yes, f is continuous, because $\sin\colon \mathbb{R} \to \mathbb{R}$ is continuous'.

The function $\sin\colon \mathbb{R} \to \mathbb{R}$ is the composite $i \circ f$ of the functions

$$\mathbb{R} \xrightarrow{f} [-1,1] \xrightarrow{i} \mathbb{R}.$$

So, the principle we would like to use is that if $i \circ f$ is continuous then so is f itself. That is the main content of the following theorem.

Theorem A9.11 *Let X be a topological space and let A be a subspace of X. Then for any topological space Z and any function $f\colon Z \to A$,*

$$f\colon Z \to A \text{ is continuous} \iff i \circ f\colon Z \to X \text{ is continuous.}$$

The maps involved can be illustrated in a triangle:

$$\begin{tikzcd} Z \arrow[r, "f"] \arrow[dr, "i\circ f"'] & A \arrow[d, "i"] \\ & X \end{tikzcd}$$

Proof $\Longrightarrow$ follows from Lemma A9.10 and the fact that the composite of continuous functions is continuous (Lemma A4.8(ii)). For $\Longleftarrow$, suppose that $i \circ f$ is continuous. Let U be an open subset of A. Then $U = W \cap A$ for some open subset W of X. Now

$$f^{-1}U = f^{-1}(W \cap A) = f^{-1}i^{-1}W = (i \circ f)^{-1}W,$$

which is open as $i \circ f$ is continuous. $\square$

Informally, Theorem A9.11 says that whether or not a map is continuous is unaffected by shrinking or enlarging the codomain, as long as you use the subspace topology. For instance, a map into $[-1,1]$ is continuous if and only if the corresponding map into $\mathbb{R}$ is continuous.

In fact, it can be shown that the subspace topology is the *only* topology on A for which Theorem A9.11 holds. This is a good reason for it to be defined the way it is.

We finish by considering the restrictions of continuous maps. Our experience of metric spaces suggests that the restriction of a continuous map should be continuous, and this is indeed the case:

Lemma A9.12 *Let $f\colon X \to Y$ be a continuous map between topological spaces. Let A be a subspace of X. Then the restricted map $f|_A\colon A \to Y$ is continuous.*

Proof Write $i\colon A \to X$ for the inclusion map. Then $f|_A$ is exactly the composite $f \circ i\colon A \to Y$. But f and i are both continuous (by Lemma A9.10), so $f \circ i$ is continuous. $\square$

Continuity is a local property, meaning that if we have a space X covered by 'small' open patches, and a function $f\colon X \to Y$ that is continuous on each patch, then f itself is continuous. This is the first part of our final result. The second part says that the same is true for closed sets as long as we stick to *finite* unions.

Lemma A9.13 *Let X and Y be topological spaces and let $f\colon X \to Y$ be a function.*

i. Let $(U_i)_{i \in I}$ be a family of open subsets of X such that $X = \bigcup_{i \in I} U_i$. Then f is continuous if and only if $f|_{U_i}$ is continuous for each $i \in I$.

ii. Let $(V_i)_{i \in I}$ be a finite *family of closed subsets of X such that $X = \bigcup_{i \in I} V_i$. Then f is continuous if and only if $f|_{V_i}$ is continuous for each $i \in I$.*

Proof Exxercise. □

A10 Products (new spaces from old, 2)

For Lecture 10

This section is longer than usual. Section A12 is correspondingly shorter.

What does it mean to say that the function

$$\begin{array}{ccc} [0,1] & \to & \mathbb{R}^2 \\ t & \mapsto & (t\cos t, t\sin t) \end{array}$$

is continuous?

You probably know at least two ways to interpret that statement. Perhaps you were first taught that a function

$$\begin{array}{ccc} [0,1] & \xrightarrow{f} & \mathbb{R}^2 \\ t & \mapsto & (f_1(t), f_2(t)) \end{array}$$

is defined to be continuous if the maps $f_1, f_2\colon [0,1] \to \mathbb{R}$ are both continuous. (In this case, $f_1(t) = t\cos t$ and $f_2(t) = t \sin t$, so f_1 and f_2 are indeed both continuous.) But later, you learned the definition of continuity for maps between metric spaces. Since $[0,1]$ and $\mathbb{R}^2$ are both metric spaces, this gives another definition of continuity for maps $[0,1] \to \mathbb{R}^2$.

Although you might not have noticed it before, there is a potential conflict here. What if it were possible to construct a function $[0,1] \to \mathbb{R}^2$ that was continuous in one sense but not the other? For instance, could it be that $f_1, f_2\colon [0,1] \to \mathbb{R}$ are continuous but $f\colon [0,1] \to \mathbb{R}^2$, as a map between metric spaces, is not? Do these notions of continuity ever disagree?

Fortunately, they don't. As results in this section will imply, a function $f\colon [0,1] \to \mathbb{R}^2$ is continuous as a map of metric spaces if and only if both its components $f_1, f_2\colon [0,1] \to \mathbb{R}$ are continuous. That's a relief!

Now let's consider the situation for *topological* spaces. Take two topological spaces, X_1 and X_2. We can form the cartesian product $X_1 \times X_2$ of the two sets. (Recall what this means: an element of $X_1 \times X_2$ is a pair (x_1, x_2) with $x_1 \in X_1$ and $x_2 \in X_2$.) Is there a sensible topology to put on $X_1 \times X_2$? Given what we just said about metric spaces, 'sensible' should mean that whenever Z is a topological space and

$$\begin{array}{ccc} Z & \xrightarrow{f} & X_1 \times X_2 \\ t & \mapsto & (f_1(t), f_2(t)) \end{array}$$

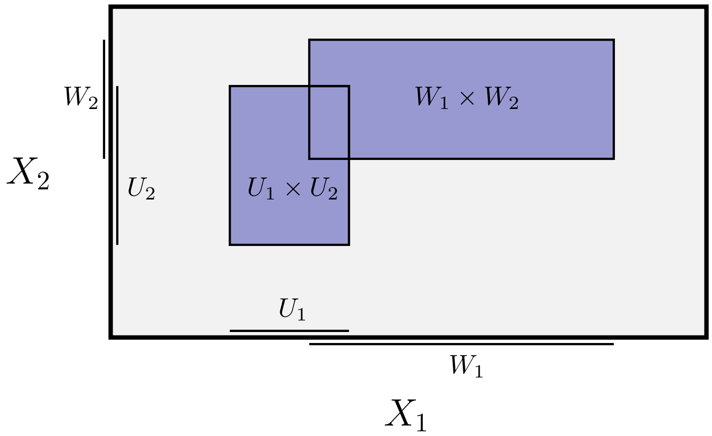


Figure A.7: The union of two rectangles is not usually a rectangle

is a function, f should be continuous if and only if $f_1 \colon Z \to X_1$ and $f_2 \colon Z \to X_2$ are continuous.

It turns out that there's exactly one topology on $X_1 \times X_2$ with this property. We will discover what it is.

First, let's guess. Perhaps we should define a subset of $X_1 \times X_2$ to be 'open' if it is of the form $U_1 \times U_2$ where U_1 is an open subset of X_1 and U_2 is an open subset of X_2. But there is an immediate problem: this doesn't define a topology, because *the union of two sets of this form is not necessarily of this form* (Figure A.7).

So, let's refine our guess. Put

$$\mathcal{B} = \{U_1 \times U_2 : U_1 \text{ is an open subset of } X_1 \text{ and } U_2 \text{ is an open subset of } X_2\}.$$

Thus, $\mathcal{B}$ is a set of subsets of $X_1 \times X_2$. We have just seen that $\mathcal{B}$ is not necessarily a topology on $X_1 \times X_2$. But it is a (synthetic) *basis* for a topology. The first axiom of Definition A7.4 holds because $X_1 \times X_2$ itself belongs to $\mathcal{B}$. The second holds because if U_1 and W_1 are open in X_1 and U_2 and W_2 are open in X_2 then

$$(U_1 \times U_2) \cap (W_1 \times W_2) = (U_1 \cap W_1) \times (U_2 \cap W_2)$$

is itself an element of $\mathcal{B}$, and is therefore (trivially!) a union of elements of $\mathcal{B}$.

Definition A10.1 Let X_1 and X_2 be topological spaces. The **product topology** on the set $X_1 \times X_2$ is the topology generated by the basis

$$\{U_1 \times U_2 : U_1 \text{ is an open subset of } X_1 \text{ and } U_2 \text{ is an open subset of } X_2\}.$$

The set $X_1 \times X_2$ equipped with the product topology is called the **product** of the spaces X_1 and X_2 (and the set $X_1 \times X_2$ is always assumed to be given the product topology unless indicated otherwise).

Before we go any further, let us record the following useful lemma (which really should have been in the section on bases):

Lemma A10.2 *Let $\mathcal{B}$ be a basis for a space X, and let $U \subseteq X$. Then U is open in X if and only if for all $x \in U$, there exists $W \in \mathcal{B}$ such that $x \in W \subseteq U$.*

Proof 'If' follows from Lemma A2.9. For 'only if', suppose that U is open and let $x \in U$. By definition of basis, U is a union of elements of $\mathcal{B}$, and at least one of them must contain x. $\square$

We can now rephrase the definition of the product topology:

Lemma A10.3 *Let X_1 and X_2 be topological spaces, and let $U \subseteq X_1 \times X_2$. Then U is open in the product topology if and only if for all $(x_1, x_2) \in U$, there exist open neighbourhoods U_1 of x_1 and U_2 of x_2 such that $U_1 \times U_2 \subseteq U$.*

Proof This is immediate from the definition of product topology and Lemma A10.2. $\square$

Warning A10.4 A common mistake is to think that in the product topology on $X_1 \times X_2$, the *only* open subsets are those of the form $U_1 \times U_2$ with U_i open in X_i. These are indeed open, *but they're not the only ones.*

This is clear if you think about the case $X_1 = X_2 = \mathbb{R}$: e.g. an open disk in $\mathbb{R}^2$ is not equal to $U_1 \times U_2$ for any $U_1, U_2 \subseteq \mathbb{R}$.

Our mission is to find a topology on $X_1 \times X_2$ such that a map $f: Z \to X_1 \times X_2$ is continuous if and only if its components $f_1: Z \to X_1$ and $f_2: Z \to X_2$ are both continuous. We'll show that the product topology has this property... but not just yet. First, we produce evidence of a different sort that the product topology is something sensible.

We defined the product $X_1 \times X_2$ of any two topological spaces. Exactly the same can be done for any finite collection of topological spaces, resulting in a product space $X_1 \times \cdots \times X_n$. (The product of *infinitely* many spaces can also be defined, but requires a little more care.)

Proposition A10.5 *Let $n \geq 0$. The product topology on $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ is the same as the standard topology.*

Proof Let $U \subseteq \mathbb{R}^n$. We will show that U is open in the product topology if and only if it is open in the standard topology, using the fact that the standard topology is induced by the metric d_∞ (Examples A2.3(ii) and A3.4(ii)).

First suppose that U is open in the product topology. Let $x = (x_1, \ldots, x_n) \in U$. By definition of the product topology, there are open subsets $U_1, \ldots, U_n$ of $\mathbb{R}$ such that

$$x \in U_1 \times \cdots \times U_n \subseteq U.$$

For each i, we have $x_i \in U_i$ and U_i is open in $\mathbb{R}$, so there exists $r_i > 0$ such that $(x_i - r_i, x_i + r_i) \subseteq U_i$. Put $r = \min\{r_1, \ldots, r_n\} > 0$; then $(x_i - r, x_i + r) \subseteq U_i$ for each i. Hence

$$(x_1 - r, x_1 + r) \times \cdots \times (x_n - r, x_n + r) \subseteq U, \tag{A:1}$$

that is, $B_{d_\infty}(x, r) \subseteq U$. So U is open in the topology induced by d_∞, which is the standard topology.

Conversely, suppose that U is open in the standard topology. Let $x \in U$. We may choose $r > 0$ such that $B_{d_\infty}(x, r) \subseteq U$, or equivalently, such that (A:1) holds. Put $U_i = (x_i - r, x_i + r)$: then U_i is an open subset of $\mathbb{R}$ and

$$x \in U_1 \times \cdots \times U_n \subseteq U.$$

So by Lemma A10.3, U is open in the product topology. $\square$

Remark A10.6 We have seen that there are many interesting metrics on $\mathbb{R}^n$, including d_1, d_2 and d_∞ (and more generally, d_p whenever $1 \leq p \leq \infty$). So there isn't a single 'product metric'. But there is a single product *topology*, and it's the topology induced by all the metrics just named.

Remark A10.7 We will need a little set theory. Let X_1 and X_2 be sets. There are **projection maps**

$$\begin{array}{rrcl} \mathrm{pr}_1 \colon & X_1 \times X_2 & \to & X_1, \\ & (x_1, x_2) & \mapsto & x_1 \end{array} \qquad \begin{array}{rrcl} \mathrm{pr}_2 \colon & X_1 \times X_2 & \to & X_2, \\ & (x_1, x_2) & \mapsto & x_2. \end{array}$$

A function $f\colon Z \to X_1 \times X_2$ can be written as $z \mapsto (f_1(z), f_2(z))$. (For instance, a curve $f\colon [0,1] \to \mathbb{R}^2$ can be written as $t \mapsto (f_1(t), f_2(t))$.) So a function $f\colon Z \to X_1 \times X_2$ amounts to a function $f_1\colon Z \to X_1$ together with a function $f_2\colon Z \to X_2$. The 'components' f_1 and f_2 of f can be expressed as follows:

$$f_1 = \mathrm{pr}_1 \circ f, \qquad f_2 = \mathrm{pr}_2 \circ f,$$

since, for instance, $(\mathrm{pr}_1 \circ f)(z) = \mathrm{pr}_1(f(z)) = \mathrm{pr}_1\big((f_1(z), f_2(z))\big) = f_1(z)$.

Lemma A10.8 *Let X_1 and X_2 be topological spaces. Then the projection maps*

$$X_1 \xleftarrow{\mathrm{pr}_1} X_1 \times X_2 \xrightarrow{\mathrm{pr}_2} X_2$$

are continuous.

Proof By symmetry, it is enough to prove this for pr_1. Let U_1 be an open subset of X_1. Then

$$\mathrm{pr}_1^{-1} U_1 = \{(x_1, x_2) \in X_1 \times X_2 : x_1 \in U_1\} = U_1 \times X_2,$$

which is open in $X_1 \times X_2$ since U_1 is open in X_1 and X_2 is open in X_2. □

The product topology is the smallest (coarsest) topology on $X_1 \times X_2$ such that the projections are continuous. Compare the remark on the subspace topology after Lemma A9.10.

Remark A10.9 Given a set A and functions $f_i\colon A \to X_i$ ($i \in I$) into topological spaces X_i, we can give A the smallest (or coarsest or weakest) topology such that every f_i is continuous. This is called the **weak topology** generated by the family of maps $(f_i)_{i \in I}$. For instance, the subspace topology on $A \subseteq X$ is the weak topology generated by the inclusion $A \to X$, and the product topology on $X_1 \times X_2$ is the weak topology generated by the two projections.

We now fulfil the mission described above: to show that a map into a product space is continuous if and only if its components are continuous.

Theorem A10.10 *Let X_1 and X_2 be topological spaces. Let Z also be a topological space, and let*

$$\begin{array}{rrcl} f\colon & Z & \to & X_1 \times X_2 \\ & z & \mapsto & (f_1(z), f_2(z)) \end{array}$$

be a function. Then

$$f \text{ is continuous} \iff f_1 \text{ and } f_2 \text{ are continuous.}$$

The maps involved are shown here:

$$\begin{tikzcd}
& Z \arrow[d, "f"] \arrow[ddl, bend right, "f_1"'] \arrow[ddr, bend left, "f_2"] & \\
& X_1 \times X_2 \arrow[dl, "\mathrm{pr}_1"] \arrow[dr, "\mathrm{pr}_2"'] & \\
X_1 & & X_2
\end{tikzcd}$$

Proof For $\Longrightarrow$, suppose that f is continuous. We have $f_1 = \mathrm{pr}_1 \circ f$ (Remark A10.7). But pr_1 is continuous (Lemma A10.8) and a composite of continuous functions is continuous (Lemma A4.8(ii)), so f_1 is continuous. The same goes for f_2.

For $\Longleftarrow$, suppose that f_1 and f_2 are continuous. We want to show that $f\colon Z \to X_1 \times X_2$ is continuous. Since the product topology was defined in terms of a basis, Lemma A7.3 implies that we need only show that $f^{-1}(U_1 \times U_2)$ is open for each open $U_1 \subseteq X_1$ and $U_2 \subseteq X_2$. So, let $U_1 \subseteq X_1$ and $U_2 \subseteq X_2$ be open sets. Then

$$\begin{aligned}
f^{-1}(U_1 \times U_2) &= \{z \in Z : (f_1(z), f_2(z)) \in U_1 \times U_2\} \\
&= \{z \in Z : f_1(z) \in U_1 \text{ and } f_2(z) \in U_2\} \\
&= f_1^{-1}U_1 \cap f_2^{-1}U_2,
\end{aligned}$$

which is open in Z since f_1 and f_2 are continuous. $\square$

Example A10.11 Consider a function

$$\begin{array}{cccc}
f\colon & [0,1] & \to & \mathbb{R}^2 \\
& t & \mapsto & (f_1(t), f_2(t))
\end{array}$$

where $\mathbb{R}^2$ has the standard topology. Is it continuous? By Theorem A10.10, f is continuous with respect to the product topology on $\mathbb{R}^2$ if and only if $f_1, f_2\colon [0,1] \to \mathbb{R}$ are both continuous. By Proposition A10.5, the product topology on $\mathbb{R}^2$ is the same as the standard topology. So, f is continuous with respect to the standard topology if and only if both f_1 and f_2 are continuous. Mission accomplished!

Remark A10.12 We made the convention that when A is a subset of a topological space X, we give A the subspace topology unless indicated otherwise (Remark A9.5). We also made the convention that when X_1 and X_2 are topological spaces, we give the set $X_1 \times X_2$ the product topology unless indicated otherwise (Definition A10.1).

Now suppose that we have topological spaces X_1 and X_2 and subsets $A_1 \subseteq X_1$ and $A_2 \subseteq X_2$. According to our conventions, what topology should we put on $A_1 \times A_2$? There are two possible answers. One is that we give A_1 the subspace topology from X_1 and A_2 the subspace topology from X_2, then put the product topology on $A_1 \times A_2$. The other is that we put the product topology on $X_1 \times X_2$, then the subspace topology on the subset $A_1 \times A_2 \subseteq X_1 \times X_2$. If these two topologies on $A_1 \times A_2$ were different, then the conventions we have made would be incompatible in a rather subtle and dangerous way. However, you can check that the two topologies on $A_1 \times A_2$ are, in fact, the same.

Other intuitively plausible properties of product spaces also hold:

$$(X_1 \times X_2) \times X_3 \cong X_1 \times (X_2 \times X_3), \qquad X \times 1 \cong X, \qquad X_1 \times X_2 \cong X_2 \times X_1$$

where 1 denotes the one-point space. These are straightforward to check.

A11 Quotients (new spaces from old, 3)

For Lecture 11

Someone hands you a pack of cards. You decide to ignore the suits and only pay attention to the numbers on the cards.

Someone hands you a straight line segment. You decide to glue the ends together to make a circle.

Someone hands you the integers. You decide that you only care what an integer is mod 10.

Someone hands you a group G. You pick a normal subgroup N and form the quotient group G/N.

Someone hands you a rectangle of paper. You twist it and glue the ends together to make a Möbius band.

These are all examples of 'quotient objects'. The common feature is that in each case, we take the object that we are handed and glue parts of it to itself, thus forming a new object. For instance, when we work mod 10, we are 'gluing together' all the integers with the same last digit. When we form a quotient group G/N, we glue g to g' whenever $g^{-1}g' \in N$.

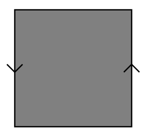
In this section, we will learn how to form quotients of topological spaces. To do this, we first need to remember how set-theoretic quotients work.

Remark A11.1 Recall the notion of an *equivalence relation* on a set. For example, every function $f\colon X \to Y$ between sets induces an equivalence relation $\sim_f$ on the domain X, where $x \sim_f x' \iff f(x) = f(x')$ $(x, x' \in X)$.

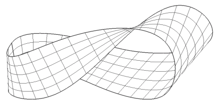
In fact, every equivalence relation arises in this way. (Better still, every equivalence relation is of the form $\sim_f$ for some *surjection* f.) For given an equivalence relation $\sim$ on a set X, we can form the set $X/{\sim}$ of equivalence classes, and there is a natural surjection $p\colon X \to X/{\sim}$ defined by taking $p(x)$ to be the equivalence class of x. Then $p(x) = p(x') \iff x \sim x'$, so $\sim\, = \,\sim_p$.

Examples A11.2 i. Think again about gluing together the two ends of a line segment $[0, 1]$ to make a circle. The gluing instruction amounts to an equivalence relation $\sim$ on $[0, 1]$; here $x \sim x'$ iff $\{x, x'\} = \{0, 1\}$ or $x = x'$. Then $X/{\sim}$ is the circle, at least as a set.

ii. The instruction to glue together the ends of a twisted rectangle of paper to make a Möbius band amounts to an equivalence relation on the rectangle.



The arrows on the diagram mean: glue together the edges marked by arrows in such a way that the arrows point in the same direction. Formally, this instruction amounts to the equivalence relation $\sim$ on $[0,1] \times [0,1]$ defined by $(x,y) \sim (x',y')$ iff either (a) $\{x,x'\} = \{0,1\}$ and $y+y' = 1$, or (b) $(x,y) = (x',y')$. Then we can reasonably define the Möbius band to be the set $X/\sim$:



But this only defines it as a *set*, not as a topological space.

In both these examples, what we are missing is a topology on $X/\sim$. We define one now.

Definition A11.3 Let $X = (X, \mathcal{T})$ be a topological space and let $\sim$ be an equivalence relation on the set X. Write p for the natural surjection $X \to X/\sim$. The **quotient topology** $\mathcal{T}'$ on $X/\sim$ is defined by

$$\mathcal{T}' = \{W \subseteq X/\sim \, : p^{-1}W \in \mathcal{T}\}.$$

The topological space $(X/\sim, \mathcal{T}')$ is called the **quotient space** of X by $\sim$ (and the set $X/\sim$ is always assumed to be given the quotient topology, unless otherwise mentioned).

Of course, one should check that the quotient topology really is a topology. This is left to you!

Like the definition of continuity, the definition of the quotient topology involves *preimages*, not images. There is no easy way to rephrase it in terms of images. For instance, it is *not* equivalent to say that the open subsets of $X/\sim$ are those of the form pU where U is an open subset of X. (If you think it is, try to prove it.)

Examples A11.4 i. Take the equivalence relation $\sim$ on $[0,1]$ defined in Example A11.2(i). Then $[0,1]/\sim$ is homeomorphic to the circle S^1 (with its subspace topology from $\mathbb{R}^2$). It is not obvious that the quotient topology is the same as the standard topology, but it is; we will come back to this.

ii. Take the equivalence relation $\sim$ on $\mathbb{R}$ defined by $x \sim x' \iff x - x' \in \mathbb{Z}$. Then $\mathbb{R}/\sim$ is again homeomorphic to S^1. (Again, we omit the proof that the quotient topology is the same as the standard topology.)

iii. We have an intuitive sense of what it means for two lines through the origin in three-dimensional space to be 'close' to one another, or of what it would mean for a line through the origin to move continuously through time. So, there ought to be a topological space of such lines. Here's how we define it.

First, what *is* a line through the origin in $\mathbb{R}^3$? Any such line is determined by choosing some point x on it other than the origin (since then, that line is the unique line passing through the origin and x). Two points $x, x' \neq 0$

determine the same line if and only if $x = \lambda x'$ for some scalar λ. So the set of all lines through the origin in $\mathbb{R}^3$ is

$$(\mathbb{R}^3 \setminus \{0\})/\sim$$

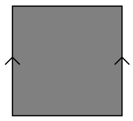
where $\sim$ is the equivalence relation on $\mathbb{R}^3 \setminus \{0\}$ defined by

$$x \sim x' \iff x = \lambda x' \text{ for some } \lambda \in \mathbb{R}$$

($x, x' \in \mathbb{R}^3 \setminus \{0\}$). But $\mathbb{R}^3 \setminus \{0\}$ carries a topology (the subspace topology from $\mathbb{R}^3$). So we can give $(\mathbb{R}^3 \setminus \{0\})/\sim$ the quotient topology. This quotient space is called the **real projective plane**.

The next few examples all involve gluing together edges of the square $D = [0, 1] \times [0, 1]$, using the arrow notation explained above.

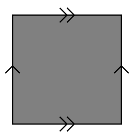
Examples A11.5 i. The diagram



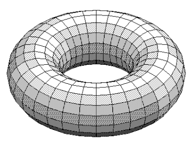
indicates the equivalence relation $\sim$ on D defined by $(x, y) \sim (x', y')$ if and only if *either* $\{x, x'\} = \{0, 1\}$ and $y = y'$ *or* $(x, y) = (x', y')$. The quotient space $D/\sim$ is a cylinder.

ii. Define $\sim$ as in Example A11.2(ii). The **Möbius band** is by definition the quotient space $D/\sim$.

iii. The diagram

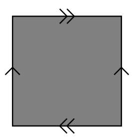


shows a more complicated equivalence relation on D: glue together the left- and right-hand edges, and glue together the top and bottom edges. The resulting quotient space is the **torus**:

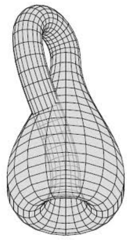


This can also be described as the product space $S^1 \times S^1$. We will prove this later in the course, once we have the right tools.

iv. Suppose that we reverse one of the arrows in this diagram:



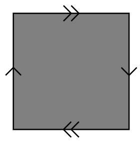
The resulting quotient space $D/\sim$ is called the **Klein bottle**. It cannot be embedded into $\mathbb{R}^3$; the best effort to do so looks something like this:



(The problem is where the neck crosses through the wall of the bottle. Really, the neck and the wall should both continue unobstructed, but there is no way to make this happen within the confines of $\mathbb{R}^3$.)

The Klein bottle is a 'non-orientable surface'. You'll learn more about surfaces if you take the Algebraic Topology course. Non-orientability means that it has no inside or outside. If you choose any spot on what you might think of as the 'outside' of the bottle, start painting, and keep painting the areas next to those you've already painted, then you'll eventually discover that you're painting the opposite side of where you started.

v. Finally, suppose that we reverse *another* arrow:



The resulting quotient space is, in fact, the real projective plane. (Again, this is not obvious; you can treat it as a non-examinable exercise.) It is also a non-orientable surface.

So far, we have not seen any justification for defining the quotient topology the way we did. In the examples above, we omitted all checks that actually used the definition. But now we give some justification of a theoretical rather than examples-based type. The next two results should remind you of results we proved for subspaces and results we proved for product spaces.

Lemma A11.6 *Let X be a topological space and $\sim$ an equivalence relation on X. Then the natural surjection $X \to X/\sim$ is continuous.*

Proof This is immediate from the definition of the quotient topology. $\square$

In fact, the quotient topology is the largest (or strongest, or finest) topology on $X/\sim$ that makes Lemma A11.6 true. This is immediate from the definition. Compare and contrast the remarks after Lemmas A9.10 (for subspaces) and A10.8 (for products). Also compare and contrast our next result with Theorems A9.11 (for subspaces) and A10.10 (for products).

Theorem A11.7 *Let X be a topological space and $\sim$ an equivalence relation; write p for the natural surjection $X \to X/\sim$. Let Z also be a topological space, and let $f\colon X/\sim \to Z$ be a function. Then*

$$f\colon X/\sim \to Z \text{ is continuous} \iff f \circ p\colon X \to Z \text{ is continuous.}$$

The maps involved are shown here:

$$\begin{tikzcd} X \arrow[r, "p"] \arrow[dr, "f\circ p"'] & X/\sim \arrow[d, "f"] \\ & Z \end{tikzcd}$$

Proof By definition of the quotient topology,

$$\begin{aligned} f \text{ is continuous} &\iff f^{-1}U \text{ is open in } X/\sim \text{ for all open } U \subseteq Z \\ &\iff p^{-1}f^{-1}U \text{ is open in } X \text{ for all open } U \subseteq Z \\ &\iff (f\circ p)^{-1}U \text{ is open in } X \text{ for all open } U \subseteq Z \\ &\iff f\circ p \text{ is continuous.} \end{aligned}$$ □

Example A11.8 View the circle as the quotient space $\mathbb{R}/\sim$, where $\sim$ is defined as in Example A11.4(ii). If $x, x' \in \mathbb{R}$ with $x - x' \in \mathbb{Z}$ then $\cos(2\pi x) = \cos(2\pi x')$. Hence there is a function $f\colon \mathbb{R}/\sim \to \mathbb{C}$ defined by $f([x]) = \cos(2\pi x)$ for all $x \in \mathbb{R}$, where $[x]$ denotes the equivalence class of x. Then $(f \circ p)(x) = f([x]) = \cos(2\pi x)$ for all $x \in \mathbb{R}$, so $f \circ p$ is continuous, so f is continuous.

(The moral here is that continuous 1-periodic functions on $\mathbb{R}$ are essentially the same thing as continuous functions on S^1. This point of view is useful in the theory of Fourier series.)

We still have the problem that it is hard for us to recognize the quotient topology when we see it. For example, I claimed in Example A11.4(i) that when we glue together the ends of the interval $[0, 1]$, the resulting quotient topology on the circle is the same as its standard topology. We will find it easier to prove statements like this once we have developed the theory of compactness.

A12 Review of Chapter A

For Lecture 12

From metric to topological spaces We started this course knowing about metric but not topological spaces. Topological spaces are more general than metric spaces, in the following sense: every metric space gives rise to a topological space, but not every topological space comes from a metric space. But as well as topological spaces *generalizing* metric spaces, they also *blur detail*: different metrics on a set can give rise to the same topology (for instance, if they are Lipschitz equivalent).

Topologists are less discriminating people than metric geometers—they only care about continuity, not distances.

A more subtle difference between metric and topological spaces is the status of the adjective ‘open’. In a metric space, one defines a subset to be open if and only if it satisfies a certain condition involving the metric, and one proves the lemma that a finite intersection or arbitrary union of open sets is open (Lemma A1.11). But in order to even specify a topological space, one needs to declare which sets are to be called ‘open’, and the conditions in the lemma just mentioned become a test that the open sets must pass in order for them to qualify as a topology.

Different kinds of space In the introductory lecture for this course, I showed you a large variety of different kinds of space, many of which we’ve now looked at rigorously.

Among metrizable spaces, the most obvious are the subsets of $\mathbb{R}^n$ with the standard topology (which is induced by each of the metrics d_1, d_2 and d_∞, among others). But we have met other metrics and metrizable spaces too: the discrete metric, the Hamming metric, and, informally, shortest-path metrics. We have also met some non-metrizable topological spaces, such as indiscrete spaces, spaces with the cofinite topology, spaces with the Zariski topology, and topological spaces with only finitely many points (e.g. you know all the topologies on the two-point set).

The anatomy of an individual topological space Part of what we’ve done is to fix a single topological space and consider basic questions about its anatomy. For instance, are one-elements subsets always closed? In other words, is it T_1? Is it Hausdorff, or regular, or normal? Aside from these separation conditions, one can also ask: is it metrizable? And given any subset A of our space, what’s the minimal way to expand A to a closed set or shrink it to an open set? These are the notions of closure and interior.

The point of topology: continuous maps between spaces The whole point of defining ‘topological space’ was to be able to define ‘continuous map’. Topological spaces are a good general context for studying continuity. The definition of continuous map between topological spaces generalizes the definition for metric spaces, but it also works well for spaces that aren’t metrizable. And the notion of continuous map leads seamlessly into the notion of ...

Homeomorphism Homeomorphism is the right notion of ‘sameness’ for topological spaces. (Remember the introduction to Section A6.) Two topological spaces are homeomorphic if they are the same in every way that we should care about. Homeomorphism preserves every conceivable topological feature. The only thing it doesn’t preserve is the names of the elements.

Homeomorphism is the precise formulation of the intuitive (but sometimes misleading) idea of ‘deformability’, in the doughnut/coffee cup sense. There are some obvious pairs of homeomorphic spaces, such as $(0, 1) \cong (12, 15)$, and some less obvious ones, such as $(0, 1) \cong (0, \infty) \cong \mathbb{R}$ and $S^n \setminus \{x\} \cong \mathbb{R}^n$ for any $x \in S^n$. There is also the fact that every knot is homeomorphic to the circle (Example A5.4). This reveals that homeomorphism does not, in fact, quite capture the idea of being able to deform one space continuously into the other within $\mathbb{R}^3$. (Indeed, it’s clear from the definition that homeomorphism has nothing to do with $\mathbb{R}^3$.) So while the conception of topology as ‘rubber geometry’ is sometimes helpful, it has to be taken with a pinch of salt.

Sometimes it’s hard to decide whether two given spaces are homeomorphic. For instance, it’s really surprisingly difficult to prove that $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic unless $m = n$. Example A5.10 gives some clues as to why this is harder than it sounds.

Constructing topological spaces Here's a challenge. Look back at the slides from the opening lecture of this course. (They're on the website.) Choose a page at random, then choose one of the spaces shown on that page. How would you actually define it, in a rigorous mathematical way?

For some of those spaces (such as the space of possible strains of a virus), that's not a question with a clear answer, although it's still worth your while to contemplate the possibilities. But for others, we now have exactly the tools necessary in order to give a definition that's both rigorous and intuitive.

Consider, for instance, the Möbius band. You *could* describe it as a subset of $\mathbb{R}^3$, via some complicated equations and inequalities describing how it twists in space. But you wouldn't really want to, and anyway, the complicated answer you'd end up with would conceal rather than reveal the way you actually *think about* a Möbius band: take a rectangle of paper, twist it, and stick its ends together.

Using the tools we've built, we can define the Möbius band, starting from scratch—not even knowing the standard topology on $\mathbb{R}$. Here goes.

- Give $\mathbb{R}$ the topology generated by the *basis* $\{(a, b) : a, b \in \mathbb{R}, a < b\}$.
- Give $[0, 1] \subseteq \mathbb{R}$ the *subspace topology*.
- Give $[0, 1] \times [0, 1]$ the *product topology*.
- Define the Möbius band to be $([0, 1] \times [0, 1])/\!\sim$ with the *quotient topology*, where $\sim$ is defined as in Example A11.2(ii).

The italicized terms show how many of our constructions we've called upon to make this definition. It's a demonstration of just how useful they are.

What really matters Our approach to defining topological spaces was based on the concept of open set: we recorded some of the properties of open sets in a metric space, and used those properties to make the definition of topological space. So you'd be forgiven for thinking that open sets are the most important aspect of topology.

They're not. Really, they're just a means to an end—namely, setting up a good general notion of continuous map.

In fact, there are several equivalent ways of defining 'topological space' without mentioning open sets. One (different, but not very different) is to use closed sets instead: a topological space could be defined as a set together with a collection $\mathcal{K}$ of subsets (the 'closed sets') such that every finite union or arbitrary intersection of sets in $\mathcal{K}$ is also in $\mathcal{K}$. More radically, a topological space can be defined as a set X together with a function $C: \mathscr{P}(X) \to \mathscr{P}(X)$ satisfying various properties (such as $C(C(A)) = C(A)$ for all $A \subseteq X$); here we think of C as the closure operator. There's a similar definition involving interiors. Or, you can axiomatize the neighbourhoods: a topological space is a set X together with a collection $\mathcal{N}_x$ of subsets of X for each $x \in X$, satisfying some axioms.

All the concepts mentioned in the previous paragraph (open sets, closed sets, closure, interior, neighbourhoods) are important. We phrased the definition of topological space in terms of open sets purely because it's technically convenient; the axioms are pretty simple. But you can find introductory topology books that use some of the other approaches. They're all equivalent, and in all of them, we can frame the definition of continuous map. That's what matters.

When we defined subspaces, product spaces and quotient spaces, we proved that each of these constructions behaves as we'd want it to with respect to continuous maps (Theorems A9.11, A10.10 and A11.7). This may have seemed somewhat abstract, but actually it gets right to the heart of the matter: fundamentally, continuity is what we care about, not open sets.

What's next? Most of the rest of the course is about two topological properties: compactness and connectedness. You've met both before in the context of metric spaces, but we'll generalize to topological spaces and dig deeper into their properties.

Compactness is to topological spaces as finiteness is to sets, or as finite-dimensionality is to vector spaces. It's probably the most important concept of this course, but it can take some effort to fully digest.

Connectedness is the intuitive idea of being all in one piece. If you do the Algebraic Topology course, you'll see that much of that subject can be seen as the study of 'higher-dimensional connectedness'. I'll leave that remark as a mystery, to whet your appetite for things to come.

Chapter B

Compactness

Compactness is one of the most important concepts in this course. The role played by compactness in the world of topological spaces is similar to the role played by finiteness in the world of sets. General topological spaces can be rather wild; compact topological spaces are much more tame (and compact Hausdorff ones even more so).

B1 The definition of compactness

For Lecture 13

You have already met the definition of compactness for metric spaces, and perhaps you can therefore guess the definition for topological spaces. But for the moment, forget what you know. We will work up to the definition by asking:

> What would we have to assume about a topological space X in order to prove that every continuous map $f\colon X \to \mathbb{R}$ is bounded?

Recall that a function $f\colon X \to \mathbb{R}$ is **bounded** if and only if its image fX is bounded, or equivalently, if there exists $M \geq 0$ such that for all $x \in X$, $|f(x)| \leq M$. For a general space X, continuous maps $X \to \mathbb{R}$ need not be bounded. For example, the map $f\colon (0,1) \to \mathbb{R}$ defined by $f(x) = 1/x$ is not bounded.

Observation 1 If X is finite then any function $f\colon X \to \mathbb{R}$ (continuous or not) is bounded, since we can take $M = \max_{x\in X} |f(x)|$. This M is a well-defined real number (not ∞) since X is finite.

Observation 2 More generally, suppose we know that X can be expressed as a finite union of subsets A_j ($j \in J$) such that f is bounded on each individual A_j. Then f itself is bounded. Why? Well, we can choose for each $j \in J$ some $M_j \geq 0$ such that for all $x \in A_j$, $|f(x)| \leq M_j$. Put $M = \max_{j\in J} M_j$ (which is a well-defined real number, not ∞, as J is finite). Then $|f(x)| \leq M$ for all $x \in X$.

Observation 3 A continuous function need not be bounded, but it is 'locally bounded' in the following sense. Let $x \in X$. If X is a *metric* space then there is some $\delta_x > 0$ such that $fB(x, \delta_x) \subseteq (f(x) - 1, f(x) + 1)$. For a general *topological* space, the set $U_x = f^{-1}(f(x) - 1, f(x) + 1)$ is an open neighbourhood of x, and satisfies $fU_x \subseteq (f(x) - 1, f(x) + 1)$. Hence, writing $M_x = |f(x)| + 1$, we have $|f(u)| \leq M_x$ for all $u \in U_x$. In particular, f is bounded on U_x (that is, fU_x is bounded). So each point of X has a neighbourhood on which f is bounded.

Observation 4 We could attempt to use this to prove that any continuous function $f: X \to \mathbb{R}$ on any topological space X is bounded. This is destined to fail, since we know that the statement we're trying to prove is false (e.g. by the $1/x$ example above). But let's try, and see what goes wrong.

For each $x \in X$, the function f is bounded on U_x, with $|f(u)| \leq M_x$ for all $u \in U_x$. Now put $M = \max_{x \in X} M_x$. The sets U_x $(x \in X)$ cover X, so $|f(x)| \leq M$ for all $x \in X$.

The trouble is that this maximum might not be a well-defined real number, since X is infinite. So this proves nothing.

Observation 5 On the other hand, suppose we somehow knew that it was possible to cover X with only *finitely many* of the sets U_x. In other words, suppose we knew that there was some finite $Z \subseteq X$ such that $\bigcup_{z \in Z} U_z = X$. Then we could legitimately put $M = \max_{z \in Z} M_z \in \mathbb{R}$. Since the sets U_z cover X, we have $|f(x)| \leq M$ for all $x \in X$, proving that f is bounded.

To make this work, we needed to assume that it was possible to cover X with only finitely many of the sets U_x. What would make that assumption valid? All we really know about the sets U_x is that they are open and cover X. We want a guarantee that under those circumstances, it is possible to select finitely many of them that still cover X. Compactness gives us exactly this guarantee.

Definition B1.1 Let X be a topological space. A **cover** of X is a family $(U_i)_{i \in I}$ of subsets of X such that $\bigcup_{i \in I} U_i = X$. It is **finite** if the indexing set I is finite, and **open** if U_i is open for each $i \in I$.

Given a cover $(U_i)_{i \in I}$ and $J \subseteq I$, we say that $(U_j)_{j \in J}$ is a **subcover** of $(U_i)_{i \in I}$ if it is itself a cover of X.

Definition B1.2 A topological space X is **compact** if every open cover of X has a finite subcover.

So, we have sketched a proof that every continuous map from a compact space to $\mathbb{R}$ is bounded. Later, we will redo this more formally.

Compactness is a topological property, in the sense of Section A6.

Warning B1.3 Compactness is not equivalent to any of the following:

i. there exists a finite open cover;

ii. every open cover is finite;

iii. for every open cover $(U_i)_{i \in I}$, every finite subfamily $(U_j)_{j \in J}$ is a subcover;

iv. every subcover of every open cover is finite;

v. some open cover has a finite subcover.

It is a good exercise to go through each of these conditions and find an example of a space that satisfies the condition but is not compact, or vice versa.

Examples B1.4 i. The sets $(n-1, n+1)$ $(n \in \mathbb{Z})$ form an open cover of $\mathbb{R}$. This cover has no finite subcover, since if J is a finite subset of $\mathbb{Z}$ then we may choose $N \in \mathbb{N}$ such that $|n| \leq N$ for all $n \in J$, and then $N + 1 \notin \bigcup_{n \in J}(n-1, n+1)$. So $\mathbb{R}$ is not compact.

ii. For $n = 0, 1, 2, \ldots$, define $U_n \subseteq [0, 1]$ by $U_0 = [0, 1/2)$ and $U_n = (2^{-n}, 1]$ $(n \geq 1)$. Then $(U_n)_{n \geq 0}$ is an open cover of $[0, 1]$. (Why open?) It has many finite subcovers, such as $(U_n)_{n \in J}$ where J is $\{0, 1, 2\}$ or $\{0, 2\}$ or $\{0, 10\}$. But this does not prove that $[0, 1]$ is compact, since compactness says that *every* open cover has a finite subcover, and we have only shown that *this particular* open cover has a finite subcover.

iii. In fact, the compact subspaces of $\mathbb{R}^n$ are precisely the closed bounded subsets. We will prove this later.

iv. Any indiscrete space is compact. For let $(U_i)_{i \in I}$ be an open cover of an indiscrete space X. Each U_i is either $\emptyset$ or X. Assuming that X is nonempty, we must have $U_j = X$ for some $j \in I$. Then the one-member family consisting of U_j alone is a finite subcover. (And if X is empty then the empty family $(U_i)_{i \in \emptyset}$ is a finite subcover.)

v. Any finite space is compact. For given an open cover $(U_i)_{i \in I}$ of a finite space X, we can choose for each $x \in X$ an element $i_x \in I$ such that $x \in U_{i_x}$. Put $J = \{i_x : x \in X\}$, a finite set. Then $\bigcup_{j \in J} U_j = \bigcup_{x \in X} U_{i_x} = X$, so $(U_j)_{j \in J}$ is a finite subcover.

vi. A discrete space is compact if and only if it is finite. We have just proved 'if'. For 'only if', let X be a compact discrete space. Then $(\{x\})_{x \in X}$ is an open cover, so has a finite subcover; that is, there exists a finite subset $Z \subseteq X$ such that $\bigcup_{z \in Z}\{z\} = X$. But the left-hand side is just Z, which is finite, so X is finite.

vii. This example may only make sense to those of you doing Linear Analysis; others can safely ignore it. In a normed vector space V, the closed unit ball is compact if and only if V is finite-dimensional. (The proof of 'only if' appears to need the Hahn–Banach theorem.)

Remark B1.5 The last two examples are evidence for the idea that compactness is a kind of finiteness condition. Here's a table of analogies:

sets	finite sets
vector spaces	finite-dimensional vector spaces
topological spaces	compact topological spaces.

If you ask a mathematician a hard question about topological spaces, they will probably be relieved if you tell them they're allowed to assume that the spaces are compact. It often makes things easier, as we will see.

Many of the topological spaces we are interested in, such as subspaces of $\mathbb{R}^n$, come presented as subspaces of some larger space. The definition of compactness

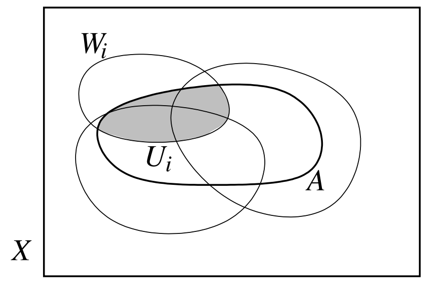


Figure B.1: Illustration of the proof of Lemma B1.7. The shaded region is U_i.

applied to a subspace (with the subspace topology, as always!) refers to the open sets of the subspace, which in turn are defined in terms of the open sets of the larger space. This is a little cumbersome. It is therefore convenient to reformulate the definition of compactness of a subspace directly in terms of the open sets of the larger space.

Definition B1.6 Let X be a set and $A \subseteq X$. A **cover of** A **by subsets of** X is a family $(W_i)_{i\in I}$ of subsets of X such that $A \subseteq \bigcup_{i\in I} W_i$. For $J \subseteq I$, we say $(W_j)_{j\in J}$ is a **subcover** of $(W_i)_{i\in I}$ if it is itself a cover of A by subsets of X.

In the special case $A = X$, a cover of X by subsets of X is simply a cover of X, and 'subcover' also has the same meaning as before.

Lemma B1.7 *Let X be a topological space and A a subspace. Then A is compact (as a topological space with the subspace topology) if and only if every cover of A by open subsets of X has a finite subcover.*

Proof Suppose that A is compact. Let $(W_i)_{i\in I}$ be a cover of A by open subsets of X (Figure B.1). For each $i \in I$, put $U_i = W_i \cap A$. Then U_i is an open subset of A by definition of the subspace topology (Definition A9.4). Also,

$$A = \left(\bigcup_{i\in I} W_i\right) \cap A = \bigcup_{i\in I} (W_i \cap A) = \bigcup_{i\in I} U_i.$$

Hence $(U_i)_{i\in I}$ is an open cover of A. Since A is compact, we can choose a finite $J \subseteq I$ such that $A = \bigcup_{j\in J} U_j$. But $U_j \subseteq W_j$ for all j, so $A \subseteq \bigcup_{j\in J} W_j$.

Conversely, suppose that every cover of A by open subsets of X has a finite subcover. Let $(U_i)_{i\in I}$ be an open cover of A (that is, a cover of A by open subsets of A). For each $i \in I$, the set U_i is open in the subspace A, so by definition of subspace topology, we may choose an open subset W_i of X such that $U_i = W_i \cap A$. Then $W_i \supseteq U_i$ for each $i \in I$, so

$$\bigcup_{i\in I} W_i \supseteq \bigcup_{i\in I} U_i = A.$$

By hypothesis, we can choose a finite subset $J \subseteq I$ such that $\bigcup_{j\in J} W_j \supseteq A$. Then

$$A = \left(\bigcup_{j\in J} W_j\right) \cap A = \bigcup_{j\in J} (W_j \cap A) = \bigcup_{j\in J} U_j.$$

Hence $(U_j)_{j\in J}$ is a finite subcover of $(U_i)_{i\in I}$, as required. $\square$

B2 Closed bounded intervals are compact

For Lecture 14; part one of three

We will soon prove that the compact subsets of $\mathbb{R}^n$ are exactly the closed bounded subsets. Some people call that result the Heine–Borel theorem; others use the same name for the following important special case.

Theorem B2.1 (Heine–Borel, weak version) *For any $a, b \in \mathbb{R}$ with $a < b$, the interval $[a, b]$ is compact.*

Proof Since $[a, b] \cong [0, 1]$, we might as well assume that $a = 0$ and $b = 1$.

We use Lemma B1.7. Let $(U_i)_{i\in I}$ be a cover of $[0, 1]$ by open subsets of $\mathbb{R}$. For the purposes of this proof, let us say that a point $c \in [0, 1]$ is *good* if $[0, c]$ is covered by $(U_j)_{j\in J}$ for some finite $J \subseteq I$. We must show that 1 is good.

Write $G = \{c \in [0, 1] : c \text{ is good}\}$. Certainly 0 is good, since we can choose $i \in I$ such that $0 \in U_i$, and then $[0, 0] = \{0\}$ is covered by the family consisting of U_i alone. So $G \neq \emptyset$. Also, $G \subseteq [0, 1]$, so G is bounded above. Hence $s = \sup G$ exists, with $s \in [0, 1]$.

(Pause for thought: we want to prove that $s = 1$. But on its own, that doesn't imply that 1 is good, since conceivably $\sup G \notin G$.)

Since $s \in [0, 1]$, we can choose $k \in I$ such that $s \in U_k$. Since U_k is open in $\mathbb{R}$, we can choose $\varepsilon > 0$ such that $(s - \varepsilon, s + \varepsilon) \subseteq U_k$. Since $s = \sup G$, we can choose a good $c \in (s - \varepsilon, s]$. And since c is good, we can choose a finite $J \subseteq I$ such that $[0, c] \subseteq \bigcup_{j\in J} U_j$. Now

$$\bigcup_{j\in J\cup\{k\}} U_j = \left(\bigcup_{j\in J} U_j\right) \cup U_k \supseteq [0, c] \cup (s - \varepsilon, s + \varepsilon) \supseteq [0, s + \varepsilon).$$

So every point $t \in [0, 1]$ satisfying $t < s + \varepsilon$ is good.

In particular, s is good. Suppose for a contradiction that $s < 1$. Then we can choose $t \in [0, 1]$ satisfying $s < t < s + \varepsilon$ (e.g. take $t = \min\{1, s + \varepsilon/2\}$). But then $t \in G$ with $\sup G = s < t$, a contradiction. So $s = 1$; but s is good, so 1 is good. □

We will use this result to prove the stronger theorem that every closed bounded subset of $\mathbb{R}^n$ is compact. In order to do this, we will need to know about compactness of subspaces and product spaces.

B3 Compactness and subspaces

For Lecture 14; part two of three

In Chapter A, we met three ways of constructing new topological spaces from old: subspaces, products and quotients. We can ask how these constructions behave with respect to compactness. For instance, is every subspace of a compact space compact? Is every product of compact spaces compact? Is every quotient of a compact space compact?

Here we look at subspaces.

Example B3.1 We have just seen that $[0,1]$ is compact. However, $(0,1)$ is not compact. (One way to see this: the cover $\big((\varepsilon,1)\big)_{\varepsilon>0}$ has no finite subcover. Another: we saw in Example A5.1(ii) that $(0,1)$ is homeomorphic to $\mathbb{R}$, and in Example B1.4(i) that $\mathbb{R}$ is not compact.) Hence a subspace of a compact space need not be compact.

This shows the limitations of the analogy between compact spaces, finite sets and finite-dimensional vector spaces (Remark B1.5), since every subset of a finite set is finite and every subspace of a finite-dimensional vector space is finite-dimensional. However, there is a result that comes close.

Lemma B3.2 *Every closed subspace of a compact space is compact.*

Proof Let X be a compact space and V a closed subset. We show that V is compact (with the subspace topology) using Lemma B1.7. Let $(U_i)_{i\in I}$ be a cover of V by open subsets of X. Then $(U_i)_{i\in I}$ together with $X \setminus V$ is an open cover of X. Since X is compact, it has some finite subcover; thus, there is some finite $J \subseteq I$ such that

$$\left(\bigcup_{j\in J} U_j\right) \cup (X \setminus V) = X.$$

Then $\bigcup_{j\in J} U_j \supseteq V$, as required. □

Is the converse true—that every compact subspace of a compact space is closed? For trivial reasons, no:

Example B3.3 Recall from Example B1.4(iv) that every indiscrete space is compact. Let X be any indiscrete space with two or more elements, and choose a nonempty proper subset Y (as we may). The subspace topology on Y is indiscrete, so Y is compact, as is X. But Y is not closed in X, since it is neither $\emptyset$ nor X.

On the other hand, it is only this sort of example that prevents the converse from being true:

Lemma B3.4 *Every compact subspace of a Hausdorff space is closed.*

(Note that the larger space need not be compact.)

Proof Let X be a Hausdorff space and A a compact subspace of X. To prove that A is closed, it is enough to prove that each point of X not in A has a neighbourhood that does not meet A (by Lemma A2.9). Let $x \in X \setminus A$.

Since X is Hausdorff and $x \notin A$, for each $a \in A$ we can choose disjoint open neighbourhoods U_a of a and W_a of x. Then $(U_a)_{a\in A}$ is a cover of A by open subsets of X. But the subspace A is compact, so by Lemma B1.7 (again!), there is a finite subset $B \subseteq A$ such that $(U_b)_{b\in B}$ covers A. Put $W = \bigcap_{b\in B} W_b$. Then W is a neighbourhood of x.

I claim that W does not meet A. Indeed, let $a \in A$; then $a \in U_b$ for some $b \in B$, and then $a \notin W_b$, so $a \notin W$, proving the claim. Hence W is a neighbourhood of x not meeting A, as required. □

Together, these two lemmas imply that when X is a compact Hausdorff space, a subspace of X is compact if and only if it is closed. For example, this is true when $X = [0,1]$ or (as we shall see) $X = [0,1]^n$.

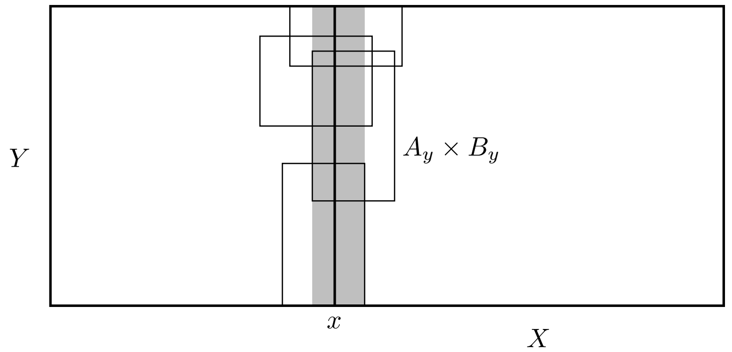


Figure B.2: Proof of Lemma B4.1. The shaded part indicates $N \times Y$.

B4 Compactness and products

For Lecture 14; part three of three

We show now that the product of two compact spaces is compact. This is one of the trickier proofs of the course so far. We begin with a lemma (Figure B.2).

Lemma B4.1 *Let X and Y be topological spaces with Y compact, and let $(U_i)_{i\in I}$ be an open cover of $X \times Y$. Then every point of X has an open neighbourhood N such that $N \times Y$ is covered by a finite subfamily of $(U_i)_{i\in I}$.*

Proof Fix $x \in X$. For each $y \in Y$, we can choose $i_y \in I$ such that $(x, y) \in U_{i_y}$, and then we can choose open neighbourhoods A_y of x and B_y of y such that $A_y \times B_y \subseteq U_{i_y}$ (by Lemma A10.3).

Now $(B_y)_{y\in Y}$ is an open cover of the compact space Y, so has a finite subcover $(B_y)_{y\in Y'}$. Put $N = \bigcap_{y\in Y'} A_y$, which is an open neighbourhood of x. Put $J = \{i_y : y \in Y'\}$, which is a finite subset of I. Then

$$\bigcup_{j\in J} U_j = \bigcup_{y\in Y'} U_{i_y} \supseteq \bigcup_{y\in Y'} (A_y \times B_y) \supseteq \bigcup_{y\in Y'} (N \times B_y) = N \times \bigcup_{y\in Y'} B_y = N \times Y,$$

as required. □

Theorem B4.2 *Let X and Y be compact topological spaces. Then $X \times Y$ is also compact.*

Proof Let $(U_i)_{i\in I}$ be an open cover of $X \times Y$. By Lemma B4.1, we can choose for each $x \in X$ an open neighbourhood N_x and a finite subset $J_x \subseteq I$ such that $N_x \times Y \subseteq \bigcup_{j\in J_x} U_j$. Then $(N_x)_{x\in X}$ is an open cover of the compact space X, so has a finite subcover $(N_x)_{x\in X'}$. Put $J = \bigcup_{x\in X'} J_x$, which is a finite subset of I (being a finite union of finite subsets). Then

$$\bigcup_{j\in J} U_j = \bigcup_{x\in X'} \bigcup_{j\in J_x} U_j \supseteq \bigcup_{x\in X'} (N_x \times Y) = \left(\bigcup_{x\in X'} N_x\right) \times Y = X \times Y,$$

as required. □

We now have everything we need to prove that closed bounded subsets of $\mathbb{R}^n$ are compact. It's just a matter of assembling results we've already proved. We'll do that next time, but a good exercise would be to try it yourself.

B5 The compact subsets of $\mathbb{R}^n$

For Lecture 15; part one of two

Our earlier work makes it easy to prove now that every closed bounded subset of $\mathbb{R}^n$ is compact. We can also prove the converse, with the aid of one further lemma:

Lemma B5.1 *Every compact metric space is bounded.*

Proof Let X be a compact metric space. If $X = \emptyset$ then certainly X is bounded. Otherwise, we can choose $a \in X$. There is an open cover $(B(a,r))_{r>0}$ of X, which has a finite subcover, say

$$B(a, r_1), \ldots, B(a, r_n).$$

Let $R = \max\{r_1, \ldots, r_n\}$. Then $X = \bigcup_{j=1}^{n} B(a, r_j) = B(a, R)$, so X is bounded. $\square$

Theorem B5.2 (Heine–Borel, strong version) *A subspace of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof Let $A \subseteq \mathbb{R}^n$. If A is compact then A is closed in $\mathbb{R}^n$ by Lemma B3.4, and bounded by Lemma B5.1. Conversely, suppose that A is closed and bounded. By boundedness, $A \subseteq [-M, M]^n$ for some $M > 0$. By Theorem B2.1, $[-M, M]$ is compact. By Theorem B4.2 and induction, $[-M, M]^n$ is compact. Since A is closed in $\mathbb{R}^n$, it is also closed in $[-M, M]^n$, so by Lemma B3.2, A is compact. $\square$

Warning B5.3 This result is specific to $\mathbb{R}^n$, and fails for many other metric spaces. For instance, let X be an infinite set with the discrete metric. Then X is closed in itself (of course!) and bounded, but by Example B1.4(vi), it is not compact.

Examples B5.4

i. Let $f_1, \ldots, f_k \colon \mathbb{R}^n \to \mathbb{R}$ be continuous functions. Then the set
$$S = \{x \in \mathbb{R}^n : f_1(x) = \cdots = f_k(x) = 0\}$$
is closed in $\mathbb{R}^n$, being the intersection of the closed sets $f_i^{-1}\{0\}$ $(1 \leq i \leq k)$. Hence S is compact if and only if S is bounded. Whether S *is* bounded depends on the equations. For instance, the plane
$$\{(x, y, z) \in \mathbb{R}^3 : x + 2y - 5z = 0\}$$
is unbounded and therefore not compact, but the ellipse
$$\{(x, y, z) \in \mathbb{R}^3 : x + 2y - 5z = 2x^2 + 3y^2 + 4z^2 - 1 = 0\}$$
is bounded and therefore compact.

ii. Similar statements can be made about subsets of $\mathbb{R}^n$ defined by non-strict inequalities. For example, the unit ball $\{x \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leq 1\}$ in $\mathbb{R}^n$ is compact.

iii. The **Cantor set** (or more properly **Cantor space**) C is the set of all $x \in [0,1]$ such that $x = \sum_{n=1}^{\infty} a_n 3^{-n}$ for some $a_1, a_2, \ldots \in \{0,2\}$. In other words, it consists of those numbers in $[0,1]$ that can be expressed in ternary without using the digit 1. (One has to be careful how one says this, as some numbers have more than one ternary expansion.)

As you may know from previous courses, C can also be described as the intersection of the sets

$$[0,1], \quad [0,1/3] \cup [2/3,1], \quad [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1], \quad \ldots$$

where at each stage we remove the open middle third from each of the intervals remaining. These sets are all closed in $\mathbb{R}$, so C is an intersection of closed sets, hence closed. It is also bounded. Hence the Cantor set is compact.

B6 Compactness and images and quotients

For Lecture 15; part two of two

We turn to the last of our three ways of building new spaces from old: quotient spaces.

Lemma B6.1 *Let $f \colon X \to Y$ be a continuous map of topological spaces, with X compact. Then fX is compact.*

Less formally: the continuous image of a compact space is compact.

Proof We use Lemma B1.7 (as we invariably do when dealing with compactness of a subspace). Let $(W_i)_{i \in I}$ be a cover of fX by open subsets of Y. Then

$$\bigcup_{i \in I} f^{-1} W_i = f^{-1} \bigcup_{i \in I} W_i \supseteq f^{-1} fX = X,$$

so $(f^{-1}W_i)_{i \in I}$ is an open cover of X. But X is compact, so there is some finite subcover $(f^{-1}W_j)_{j \in J}$. We show that $(W_j)_{j \in J}$ covers fX. For let $y \in fX$: then $y = f(x)$ for some $x \in X$, and $x \in f^{-1}W_j$ for some $j \in J$, so that $y = f(x) \in W_j$. $\square$

Corollary B6.2 *Let X be a compact space and $f \colon X \to \mathbb{R}$ a continuous function. Then f is bounded and (if X is nonempty) attains its bounds.*

Proof Since X is compact, Lemma B6.1 implies that fX is compact, and therefore closed and bounded by Theorem B5.2. Assume that X is nonempty. Then fX is nonempty and bounded above, so has a supremum. Now $\sup A \in \mathrm{Cl}(A)$ whenever $A \subseteq \mathbb{R}$ is nonempty and bounded above (by Lemma A8.5), and fX is closed, so $\sup fX \in fX$. Hence f attains its upper bound, and similarly its lower bound. $\square$

Corollary B6.2 fulfils the promise made after Definition B1.2: that we would prove that every real-valued function on a compact space is bounded. (Remember, I first introduced the definition of compactness by asking what conditions would suffice in order to prove that result.) By the weak version of the Heine–Borel theorem (Theorem B2.1), we can deduce a classic result of real analysis: every continuous function $[a,b] \to \mathbb{R}$ is bounded and attains its bounds.

Corollary B6.3 *Every quotient of a compact space is compact.*

Proof For any space X and equivalence relation $\sim$ on X, the natural surjection $X \to X/\sim$ is continuous (Lemma A11.6), so this follows from Lemma B6.1. $\square$

Examples B6.4 All the quotients of the square listed in Section A11 are compact: the cylinder, the Möbius band, the torus, the Klein bottle, and the projective plane. The compactness of some of these can also be established by other means. For example, the cylinder can alternatively be described as the product $S^1 \times [0, 1]$, and so is compact because it is the product of compact spaces (Theorem B4.2). The same goes for the torus $S^1 \times S^1$.

Remark B6.5 Secretly, compactness and Hausdorffness are 'mirror-image' or 'dual' or 'complementary' conditions, in a way that I won't be able to explain properly in this course. But we will catch a few glimpses of their complementary nature. For instance, we have seen that Hausdorffness guarantees that every sequence has *at most one* limit (Lemma A3.11), while compactness says—well, not quite that every sequence has *at least one* limit, but something similar involving subsequences (in metric spaces, anyway). Another glimpse: if you *add* more open sets to a Hausdorff topology, the result is still Hausdorff, while if you *remove* some open sets from a compact topology, the result is still compact.

The next few results give further hints of the interaction between compactness and Hausdorffness.

Definition B6.6 Let X and Y be topological spaces. A function $f : X \to Y$ is **closed** if for all closed $V \subseteq X$, the image $fV \subseteq Y$ is also closed.

Lemma B6.7 *Every continuous map from a compact space to a Hausdorff space is closed.*

Proof Let $f : X \to Y$ be a continuous map from a compact space X to a Hausdorff space Y. Let V be a closed subset of X. By Lemma B3.2, V is compact. So by Lemma B6.1, fV is compact. Then by Lemma B3.4, $fV \subseteq Y$ is closed. $\square$

I made a big deal of the fact that a continuous bijection need not be a homeomorphism. This is indeed important. But actually, there's a common situation where continuous bijections *are* automatically homeomorphisms:

Lemma B6.8 *A continuous bijection from a compact space to a Hausdorff space is a homeomorphism.*

Proof Let $f : X \to Y$ be such a bijection. For each closed $V \subseteq X$, the preimage $(f^{-1})^{-1}V$ of V under f^{-1} is fV, which is closed in Y by Lemma B6.7. So by Lemma A4.4, f^{-1} is continuous. $\square$

For example, any continuous bijection between compact Hausdorff spaces is a homeomorphism. Since compact Hausdorff spaces are quite widespread, this is a useful result.

Lemma B6.7 has a further useful consequence: it can help us to recognize quotient spaces.

Proposition B6.9 *Let $q\colon X \to Y$ be a continuous surjection from a compact space X to a Hausdorff space Y. Then Y is homeomorphic to the quotient space $X/\sim$, where $\sim$ is the equivalence relation on X induced by q.*

Corollary B6.3 also tells us that Y is compact.

Proof Write $p\colon X \to X/\sim$ for the natural surjection. For $x, x' \in X$, we have

$$p(x) = p(x') \iff x \sim x' \iff q(x) = q(x')$$

(by definition of p and $\sim$), so there is an injective function $f\colon X/\sim \to Y$ defined by $f(p(x)) = q(x)$ $(x \in X)$. Since q is surjective, so is f. Moreover, $f \circ p$ is the continuous map q, so f is continuous by Theorem A11.7.

We have constructed a continuous bijection $f\colon X/\sim \to Y$. But $X/\sim$ is compact by Corollary B6.3 and the hypothesis that X is compact, and Y is Hausdorff by hypothesis, so Lemma B6.8 guarantees that f is a homeomorphism. □

Remark B6.10 (Non-examinable.) Proposition B6.9 can be compared to the first isomorphism theorem in the theory of groups (or rings or vector spaces). Let $q\colon X \to Y$ be a surjective homomorphism of groups. The first isomorphism theorem states that Y is isomorphic to $X/\ker(q)$.

Notice that this is true for *all* groups X and Y, whereas in Proposition B6.9 we needed the hypotheses that X is compact and Y is Hausdorff. These extra hypotheses are used to upgrade $f\colon X/\sim \to Y$ from a continuous bijection to a homeomorphism, a distinction that has no analogue in group theory.

Examples B6.11

i. In Section A11, I claimed that the circle S^1 is the quotient of $[0, 1]$ by the equivalence relation $\sim$ defined by $x \sim y$ if and only if $\{x, y\} = \{0, 1\}$ or $x = y$, and I noted the difficulty of proving that the topology on S^1 that it gets as a subspace of $\mathbb{R}^2$ is the same as the topology that it gets as a quotient of $[0, 1]$.

   We can now prove it. Give S^1 the subspace topology from $\mathbb{R}^2 \cong \mathbb{C}$. Define $q\colon [0, 1] \to S^1$ by $q(x) = e^{2\pi i x}$ $(x \in [0, 1])$. Then q is a continuous surjection from a compact space to a Hausdorff space, and the equivalence relation $\sim$ on $[0, 1]$ that it induces is the one described in the previous paragraph. Hence by Proposition B6.9, S^1 is the quotient space $[0, 1]/\sim$.

ii. Similarly, the map
$$[0, 1] \times [0, 1] \xrightarrow{q \times q} S^1 \times S^1$$
   (where q is as in (i)) is a continuous surjection from a compact space to a Hausdorff space, so $S^1 \times S^1$ is homeomorphic to the quotient space $([0, 1] \times [0, 1])/\sim$, where $\sim$ is the equivalence relation induced by $q \times q$. That equivalence relation is exactly the one described in Example A11.5(iii). So our two descriptions of the torus, as a quotient of the square and as $S^1 \times S^1$, are equivalent.

B7 Compact metric spaces

For Lecture 16

You already know quite a few results about compactness for metric spaces. Here, we'll reprove most or all of the results you've seen before. But perhaps we'll prove some theorems that are new to you, and perhaps you'll meet here some new proofs of familiar theorems.

Definition B7.1 A topological space X is **sequentially compact** if every sequence in X has a convergent subsequence.

Warning B7.2 As we'll prove, sequential compactness is equivalent to compactness for *metric* spaces. However, there are examples of *topological* spaces that are compact but not sequentially compact, and also topological spaces that are sequentially compact but not compact. (Try a web search!) Sequences aren't very effective tools for probing an arbitrary topological space, as Example A3.10(ii) suggests.

Lemma B7.3 *Let X be a metric space, let $(x_n)_{n=1}^{\infty}$ be a sequence in X, and let $x \in X$. Then (x_n) has a subsequence converging to x $\iff$ for every neighbourhood W of x, there are infinitely many n such that $x_n \in W$.*

Proof $\implies$ is clear. For $\impliedby$, define integers $n_1 < n_2 < \cdots$ by putting $n_1 = 1$ and, inductively, choosing $n_{k+1} > n_k$ such that $d(x_{n_{k+1}}, x) < 1/(k+1)$ (which the hypothesis guarantees we can do). Then $(x_{n_k})_{k=1}^{\infty}$ converges to x. □

Warning B7.4 The condition 'there are infinitely many n such that $x_n \in W$' does *not* imply that W contains infinitely many points belonging to the sequence. For example, if $x \in W$ and $x_n = x$ for all n then there are infinitely many n such that $x_n \in W$, yet W contains only one point belonging to the sequence.

Proposition B7.5 *A compact metric space is sequentially compact.*

Proof Let X be a compact metric space and let $(x_n)_{n=1}^{\infty}$ be a sequence in X. Suppose for a contradiction that (x_n) has no convergent subsequence. Then by Lemma B7.3, every $x \in X$ has an open neighbourhood U_x such that $\{n \geq 1 : x_n \in U_x\}$ is finite. The open cover $(U_x)_{x \in X}$ of X has a finite subcover $(U_y)_{y \in Y}$. But then

$$\mathbb{Z}^+ = \{n \geq 1 : x_n \in X\} = \bigcup_{y \in Y} \{n \geq 1 : x_n \in U_y\},$$

which is a finite union of finite sets and therefore finite, a contradiction. □

Corollary B7.6 (Bolzano–Weierstrass) *Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

Proof Any bounded sequence in $\mathbb{R}^n$ is contained in $[-M, M]^n$ for some $M > 0$, which is compact by Theorem B5.2, and therefore sequentially compact. □

To prove that any sequentially compact metric space is compact, we introduce a further concept.

Definition B7.7 Let $(U_i)_{i \in I}$ be a cover of a metric space X. A **Lebesgue number** for $(U_i)_{i \in I}$ is a real number $\varepsilon > 0$ with the following property: for all $x \in X$, there exists $i \in I$ such that $B(x, \varepsilon) \subseteq U_i$.

What does this mean? Observe that given any $x \in X$, we can find $\varepsilon > 0$ such that for some $i \in I$, $B(x, \varepsilon) \subseteq U_i$. That's simply by definition of open set and cover. But in general, ε depends on x. A Lebesgue number for the cover is an ε that works *for all x at once.*

Example B7.8 Consider the cover $\big((n-1, n+1)\big)_{n \in \mathbb{Z}}$ of $\mathbb{R}$. This has 1/2 (or indeed any positive real smaller than 1/2) as a Lebesgue number, since for any real x, we can find some integer n such that $(x - 1/2, x + 1/2) \subseteq (n-1, n+1)$.

However, not every cover, or even every open cover, has any Lebesgue number at all. Can you think of an example?

Lemma B7.9 *Let X be a sequentially compact metric space. Then every open cover of X has a Lebesgue number.*

Proof Let $(U_i)_{i \in I}$ be an open cover of X. Suppose for a contradiction that it has no Lebesgue number.

For each $n \geq 1$, we can choose $x_n \in X$ such that $B(x_n, 1/n)$ is not a subset of any U_i. We can then choose a subsequence (x_{n_k}) of (x_n) convergent to x, say. Since (U_i) is an open cover, we can next choose $i \in I$ and $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U_i$. Then we can choose $K \geq 1$ such that $x_{n_k} \in B(x, \varepsilon/2)$ for all $k \geq K$, and finally we can choose $k \geq K$ such that $1/n_k < \varepsilon/2$.

It follows that $B(x_{n_k}, 1/n_k) \subseteq B(x, \varepsilon)$, since if $y \in B(x_{n_k}, 1/n_k)$ then

$$d(x, y) \leq d(x, x_{n_k}) + d(x_{n_k}, y) < d(x, x_{n_k}) + 1/n_k < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence $B(x_{n_k}, 1/n_k) \subseteq U_i$, contradicting the defining property of our sequence.□

To finish the proof that sequentially compact metric spaces are compact, we introduce another condition on metric spaces: total boundedness.

Definition B7.10 Let X be a metric space. For $\varepsilon > 0$, an **ε-net** on X is a subset $Z \subseteq X$ such that $(B(z, \varepsilon))_{z \in Z}$ covers X. We say that X is **totally bounded** if for all $\varepsilon > 0$, there exists a finite ε-net on X.

As an exercise, you should show that a totally bounded metric space is always bounded. But total boundedness is a *stronger* condition than boundedness:

Example B7.11 An infinite set with the discrete metric is bounded (because all distances are ≤ 1) but not totally bounded (e.g. there is no finite 1/2-net).

Proposition B7.12 *A sequentially compact metric space is totally bounded.*

Proof Let X be a sequentially compact metric space, and suppose for a contradiction that X is not totally bounded. Then we can choose $\varepsilon > 0$ such that X has no finite ε-net.

We construct a sequence $(x_n)_{n=1}^{\infty}$ recursively as follows. Let $n \geq 1$ and suppose that $x_1, \ldots, x_{n-1}$ have been defined. Since $\{x_1, \ldots, x_{n-1}\}$ is not an ε-net, we can choose $x_n \in X$ such that $d(x_i, x_n) \geq \varepsilon$ for all $i < n$.

Now $(x_n)_{n=1}^{\infty}$ is a sequence in X such that $d(x_m, x_n) \geq \varepsilon$ whenever $m \neq n$. By sequential compactness, it has a subsequence converging to x, say. By Lemma B7.3, there are infinitely many $n \geq 1$ such that $x_n \in B(x, \varepsilon/2)$. But then $d(x_m, x_n) < \varepsilon$ for infinitely many pairs (m, n) with $m \neq n$, a contradiction. □

Proposition B7.13 *A sequentially compact metric space is compact.*

Proof Let X be a sequentially compact metric space and $(U_i)_{i\in I}$ an open cover of X. By Lemma B7.9, this cover has a Lebesgue number ε. By Proposition B7.12, X has a finite ε-net Z. For each $z \in Z$, we can choose $i_z \in I$ such that $B(z, \varepsilon) \subseteq U_{i_z}$. Put $J = \{i_z : z \in Z\} \subseteq I$. Then

$$\bigcup_{j\in J} U_j = \bigcup_{z\in Z} U_{i_z} \supseteq \bigcup_{z\in Z} B(z, \varepsilon) = X$$

(the last by definition of ε-net), so $(U_j)_{j\in J}$ is a finite subcover of $(U_i)_{i\in I}$. $\square$

We now establish a third condition (or rather, pair of conditions) equivalent to compactness for metric spaces: completeness together with total boundedness. Recall the notion of completeness:

Definition B7.14 Let X be a metric space. A sequence (x_n) in X is **Cauchy** if for all $\varepsilon > 0$, there exists N such that for all $m, n \geq N$, $d(x_m, x_n) < \varepsilon$. The space X is **complete** if every Cauchy sequence in X converges.

(Conversely, a convergent sequence is always Cauchy, in *any* metric space.)

Lemma B7.15 *A Cauchy sequence that has a convergent subsequence is itself convergent.*

Proof This was probably in Honours Analysis. If not, it's a good exercise. $\square$

Proposition B7.16 *A sequentially compact metric space is complete.*

Proof Let (x_n) be a Cauchy sequence in a sequentially compact space. Then (x_n) has a convergent subsequence, so by Lemma B7.15, (x_n) is convergent. $\square$

We have now shown that

compact $\iff$ sequentially compact $\implies$ complete and totally bounded,

and we just have the final $\impliedby$ to prove. It amounts to the following lemma:

Lemma B7.17 *In a totally bounded metric space, every sequence has a Cauchy subsequence.*

Proof Let X be a totally bounded metric space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. By total boundedness, there is a finite cover of X by subsets of diameter ≤ 1 (for instance, balls of radius $1/2$), and the set $\mathbb{N}$ is infinite, so there must be some element of this cover that contains x_n for infinitely many values of $n \in \mathbb{N}$. Hence there is an infinite subset N_1 of $\mathbb{N}$ such that $d(x_n, x_m) \leq 1$ for all $n, m \in N_1$.

Similarly, there is a finite cover of X by subsets of diameter $\leq 1/2$, and the set N_1 is infinite, so there must be some infinite subset N_2 of N_1 such that $d(x_n, x_m) \leq 1/2$ for all $n, m \in N_2$.

Continuing like this, we construct a chain $\mathbb{N} \supseteq N_1 \supseteq N_2 \supseteq \cdots$ of infinite subsets of $\mathbb{N}$ such that $d(x_n, x_m) \leq 1/k$ whenever $n, m \in N_k$.

Choose $n_1 \in N_1$. Since N_2 is infinite, we can then choose $n_2 \in N_2$ with $n_2 > n_1$. Then we can choose $n_3 \in N_3$ with $n_3 > n_2$. Continuing like this, we construct $n_1 < n_2 < \cdots$ with $n_k \in N_k$ for all k. Then $d(x_{n_k}, x_{n_l}) \leq 1/k$ whenever $k \leq l$, so $(x_{n_k})_{k=1}^{\infty}$ is a Cauchy subsequence of our original sequence $(x_n)_{n=1}^{\infty}$. $\square$

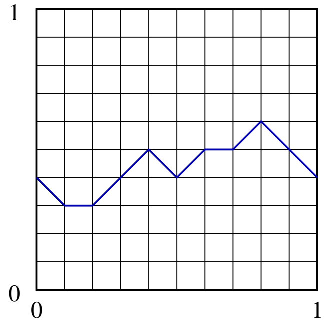


Figure B.3: One element of a (1/10)-net on the space X of Example B7.21.

Proposition B7.18 *A totally bounded complete metric space is sequentially compact.*

Proof This is immediate from Lemma B7.17 and the definitions. □

Assembling all five of the Propositions in this section so far, we arrive at:

Theorem B7.19 *The following are equivalent for a metric space X:*

i. *X is compact;*

ii. *X is sequentially compact;*

iii. *X is complete and totally bounded.* □

In particular, one way to prove that a metric space is compact is by showing that it is complete and totally bounded. The next two examples sketch applications of this strategy.

Example B7.20 Here, in outline, is an alternative proof of the Heine–Borel theorem. One can show directly that every bounded subset of $\mathbb{R}^n$ is totally bounded. On the other hand, $\mathbb{R}$ is complete, from which it can be shown that $\mathbb{R}^n$ is complete; moreover, any closed subspace of a complete space is complete. So, every closed subset of $\mathbb{R}^n$ is complete. Hence by Theorem B7.19, every closed bounded subset of $\mathbb{R}^n$ is compact.

Example B7.21 An important use of the 'complete and totally bounded' formulation of compactness is to prove that certain function spaces are compact. This is the subject of Chapter 10 of the first edition of Sutherland's book. (It's absent from the second edition.) Unfortunately, we don't have time to go into it in this course, but here is the flavour.

Let X be the set of functions $f\colon [0,1] \to [0,1]$ that are **distance-decreasing**: $|f(s)-f(t)| \leq |s-t|$ for all $s, t \in [0,1]$. Give X the d_∞ metric (also called the sup metric or uniform metric), as in Example A1.2(iii). Like many function spaces, X is complete. The distance-decreasing hypothesis guarantees that X is totally bounded; for instance, any $f \in X$ can be reasonably closely approximated by a function like the one in Figure B.3, of which there are only finitely many. It follows from Theorem B7.19 that X is compact.

We introduced the concept of Lebesgue number to help us prove that a sequentially compact metric space is compact. But it has other uses too:

Proposition B7.22 *Let $f\colon X \to Y$ be a continuous map of metric spaces. Suppose that X is compact. Then f is uniformly continuous.*

Proof Let $\varepsilon > 0$. For each $x \in X$, we can choose $\delta_x > 0$ such that for all $x' \in B(x, \delta_x)$, $d(f(x), f(x')) < \varepsilon/2$. Then $\big(B(x, \delta_x)\big)_{x \in X}$ is an open cover of the sequentially compact space X, and so has a Lebesgue number δ (by Lemma B7.9).

Let $x, x' \in X$ with $d(x, x') < \delta$. By definition of Lebesgue number, $B(x, \delta) \subseteq B(z, \delta_z)$ for some $z \in X$. Now $x, x' \in B(z, \delta_z)$, so

$$d(f(x), f(x')) \leq d(f(x), f(z)) + d(f(z), f(x')) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

as required. $\square$

This can also be proved without the use of Lebesgue numbers (as in Proposition 5.8.2 of the first edition of Sutherland), but the proof above is a bit shorter.

Chapter C

Connectedness

C1 The definition of connectedness

For Lecture 17

Intuitively, a topological space is said to be 'connected' if it is all in one piece. Of course, just about any topological space can be *torn* into two pieces; for instance, $\mathbb{R}$ can be torn into the two pieces $(-\infty, 0)$ and $[0, \infty)$. But tearing a space changes its topological nature. Connectedness means that the space does not *fall naturally* into two or more pieces.

Let us consider more closely what this means. The real line $\mathbb{R}$ is intuitively all in one piece. Why does the decomposition $(-\infty, 0) \cup [0, \infty)$ not contradict this? It is because the point 0 of the second piece is a limit point of the first piece. So, the two pieces are not entirely separated; the break between them is not clean.

More specifically, suppose we have a topological space X expressed as a disjoint union $X = U \cup V$. We should only think of this as separating X into two independent pieces if no point in U is a limit point of V, and vice versa. Equivalently, this means that $\mathrm{Cl}(U)$ is disjoint from V, and vice versa. Since $U \cup V = X$, that means that $\mathrm{Cl}(U) = U$ and $\mathrm{Cl}(V) = V$, or equivalently that U and V are both closed—or equivalently again, that U and V are both open.

Of course, we can write any space X as a disjoint union $U \cup V$ of open subsets by taking $U = \emptyset$ and $V = X$, or vice versa. We define a space to be 'connected' if there is no other way to do it.

Definition C1.1 A space X is **connected** if it is nonempty and if whenever $X = U \cup V$ for some disjoint open subsets U and V, then U or V is empty. A space is **disconnected** if it is not empty or connected.

Since connectedness is defined in terms of the topological structure alone, it is a topological property.

Remark C1.2 According to these definitions, the empty set is neither connected or disconnected. This convention is made for much the same reason that the number 1 is not counted as either prime or composite and the trivial group is not counted as simple. (Connected spaces are something like prime numbers or simple groups; they are the ones that cannot be broken down any further.) Many authors do not follow this convention, and count $\emptyset$ as connected.

Examples C1.3 i. $X = \mathbb{R} \setminus \{0\}$ is disconnected, since $(-\infty, 0)$ and $(0, \infty)$ are disjoint open subsets of X whose union is X.

ii. The space $\mathbb{Q}$ of rational numbers (topologized as a subspace of $\mathbb{R}$) is disconnected. Put $U = (-\infty, \sqrt{2}) \cap \mathbb{Q}$ and $V = (\sqrt{2}, \infty) \cap \mathbb{Q}$. Then U and V are open in $\mathbb{Q}$ (by definition of subspace topology), $U \cup V = \mathbb{Q}$ (since $\sqrt{2}$ is irrational), and $U \cap V = \emptyset$.

iii. Any discrete space X with two or more points is disconnected. For choose any $x \in X$; then $\{x\}$ and $X \setminus \{x\}$ are disjoint nonempty open subsets whose union is X.

iv. Any nonempty indiscrete space is connected.

v. As we will prove in the next section, all intervals in $\mathbb{R}$ are connected.

There are several equivalent ways of phrasing the definition of connectedness, and it is useful to have all of them available.

Lemma C1.4 *For a nonempty topological space X, the following are equivalent:*

i. X is connected;

ii. whenever $X = U \cup V$ for some disjoint closed subsets U and V, then U or V is empty;

iii. the only subsets of X that are both open and closed are $\emptyset$ and X;

iv. every continuous map from X to a discrete space is constant;

v. every continuous map from X to the two-point discrete space is constant.

Proof For (i)$\Rightarrow$(ii), suppose that X is the disjoint union of closed subsets U and V. Then X is also the disjoint union of open subsets $X \setminus U$ and $X \setminus V$. By (i), one of $X \setminus U$ and $X \setminus V$ is empty and the other is X; hence one of U and V is empty.

For (ii)$\Rightarrow$(iii), suppose that X satisfies (ii), and let U be an open and closed subset of X. Then X is the disjoint union of closed subsets U and $X \setminus U$, so one of them is empty.

For (iii)$\Rightarrow$(iv), suppose that X satisfies (iii), and let $f\colon X \to D$ be a continuous map to a discrete space D. Choose $x \in X$. Then $\{f(x)\}$ is an open and closed subset of D, so $f^{-1}\{f(x)\}$ is an open and closed subset of X. But $f^{-1}\{f(x)\}$ is nonempty (since it contains x), so $f^{-1}\{f(x)\} = X$. Hence $f(y) = f(x)$ for all $y \in X$.

Trivially, (iv)$\Rightarrow$(v).

Finally, for (v)$\Rightarrow$(i), suppose that X satisfies (v), and let U and V be disjoint open subsets of X with $U \cup V = X$. Let $\{0, 1\}$ denote the two-point discrete space, and define a function $f\colon X \to \{0, 1\}$ by

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

Then $f|_U$ and $f|_V$ are continuous, so f is continuous by Lemma A9.13. Hence f is constant, by (v). It follows that $U = \emptyset$ or $V = \emptyset$. $\square$

Sometimes, one of these equivalent conditions is more convenient than another. For instance, condition (iv) seems to be the most convenient one to use in the following proof.

Lemma C1.5 *Let X be a topological space. Let A and B be subspaces of X with $A \subseteq B \subseteq \mathrm{Cl}(A)$. If A is connected then so is B. In particular, the closure of a connected subspace is connected.*

Proof First note that A is dense in B. For if V is a closed subset of B containing A then $V = B \cap S$ for some closed $S \subseteq X$ (by definition of the subspace topology on B), and S is then a closed subset of X containing A, so $S \supseteq \mathrm{Cl}(A) \supseteq B$. Hence $V = B$, as required.

Since A is nonempty, so is B. Let f be a continuous map from B to a discrete space D. Then $f|_A : A \to D$ is also continuous, and therefore constant since A is connected; say $f(a) = d$ for all $a \in A$. Now f and the constant function d are continuous maps from B to the Hausdorff space D, and these two functions are equal on the dense subset A of B, so by Corollary A8.13, they are equal on all of B. Hence f is constant, as required. $\square$

Example C1.6 Let A be the open disk $\{z \in \mathbb{C} : |z| < R\}$, for some $R > 0$. Let B be the union of A with any set of points of modulus R. We will eventually show that A is connected; then Lemma C1.5 will imply that B is connected too. For instance, this implies that for any complex power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$, the set $\{z \in \mathbb{C} : f(z) \text{ converges}\}$ is connected.

In Chapter B, we explored how the concept of compactness related to constructions such as subspaces, products and quotients. We can do the same for connectedness. The results here are much easier.

Proposition C1.7 *Let $f : X \to Y$ be a continuous map of topological spaces. If X is connected then so is fX.*

That is, a continuous image of a connected space is connected. Recall the convention that unless mentioned otherwise, subsets of a topological space are always given the subspace topology (Remark A9.5). In particular, this applies to fX in the statement of this result.

Proof First we prove that for a continuous *surjection* $g : X \to Z$, if X is connected then so is Z. Assume that X is connected. First, Z is nonempty since X is. Now let U and V be disjoint open subsets of Z with $U \cup V = Z$. Then $g^{-1}U$ and $g^{-1}V$ are disjoint open subsets of X with $g^{-1}U \cup g^{-1}V = X$, so without loss of generality, $g^{-1}U = \emptyset$. But g is surjective, so $U = \emptyset$. Hence Z is connected.

Now consider an arbitrary continuous map $f : X \to Y$, with X connected. By Theorem A9.11, the map $g : X \to fX$ defined by $g(x) = f(x)$ is also continuous. It is also surjective. Hence by the first paragraph, fX is connected, as required. $\square$

Corollary C1.8 *Any quotient of a connected space is connected.* $\square$

Proposition C1.9 *The product of two connected spaces is connected.*

Proof Let X and Y be connected spaces. Both X and Y are nonempty, so $X \times Y$ is too. Let f be a continuous map from $X \times Y$ to a discrete space D. We prove that f is constant.

Let $(x, y), (x', y') \in X \times Y$. Then $\{x\} \times Y$ is homeomorphic to Y (as in Remark A10.12), hence connected. So $f|_{\{x\} \times Y}$ is constant. In particular, $f(x, y) = f(x, y')$. Similarly, since X is connected, $f(x', y') = f(x, y')$. Hence $f(x, y) = f(x', y')$, as required. □

Example C1.10 Once we have shown that real intervals $[a, b]$ are connected (which we will do in the next section), it will follow that all cuboids

$$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

in $\mathbb{R}^n$ are connected. But in fact, we will later prove the more general result that all convex subsets of $\mathbb{R}^n$ are connected.

We state one final basic result about connected spaces. Roughly, it says that if you glue together several connected spaces, and the spaces all overlap with each other, then the result is connected too.

Lemma C1.11 *Let X be a nonempty topological space and $(A_i)_{i \in I}$ a family of subspaces covering X. Suppose that A_i is connected for each $i \in I$ and that $A_i \cap A_j \neq \emptyset$ for each $i, j \in I$. Then X is connected.*

Proof Let f be a continuous map from X to a discrete space D. For each $i \in I$, we have the continuous map $f|_{A_i} \colon A_i \to D$, and A_i is connected, so $f|_{A_i}$ has constant value d_i, say. But for each $i, j \in I$, we have $A_i \cap A_j \neq \emptyset$, so $d_i = d_j$. Hence d_i is independent of $i \in I$, so that f is constant, as required. □

Example C1.12 The letter O is a quotient of $[0, 1]$, so once we have shown that real intervals are connected, it will follow from Corollary C1.8 that O is connected. Gluing O to another line segment gives P, which by Lemma C1.11 is connected too. Finally, gluing P to yet another line segment gives A, and one more application of Lemma C1.11 then tells us that A is also connected.

Clearly, our most pressing task is to show that real intervals are indeed connected. We do this next.

C2 Connected subsets of the real line

For Lecture 18

We now show that nonempty real intervals are connected. In order to do that, it will help to reflect on what an interval actually is. The following definition is the most convenient.

Definition C2.1 A subset $I \subseteq \mathbb{R}$ is an **interval** if for $x, y, z \in \mathbb{R}$,

$$(x \leq y \leq z \text{ and } x, z \in I) \Longrightarrow y \in I.$$

The next lemma says that this is equivalent to the much more explicit definition that you may be familiar with.

Lemma C2.2 *Let $I \subseteq \mathbb{R}$. Then I is an interval if and only if I is of one of the following eleven types:*

$$\begin{gathered}\{a\}, [a,b], [a,b), (a,b], (a,b), \\ [a,\infty), (a,\infty), (-\infty,a], (-\infty,a), \\ \emptyset, (-\infty,\infty), \end{gathered} \tag{C:1}$$

where $a, b \in \mathbb{R}$ with $a < b$.

Proof 'If' is clear. For 'only if', suppose that I is an interval.

First assume that I is nonempty and bounded. Put $a = \inf I$ and $b = \sup I$. Certainly $a \leq b$. If $a = b$ then $I = \{a\}$; suppose, then, that $a < b$. I claim that $(a, b) \subseteq I$. Indeed, let $y \in (a, b)$. By definition of infimum, there exists $x \in I$ with $x < y$, and similarly there exists $z \in I$ with $y < z$. So by definition of interval, $y \in I$, proving the claim.

On the other hand, $I \subseteq [a, b]$ by definition of infimum and supremum. So $(a, b) \subseteq I \subseteq [a, b]$, which means that I must be one of the sets in the first line of (C:1).

The other cases, where I is empty or not bounded above and/or below, are handled similarly. □

We begin our proof of the connectedness of intervals with a special case. Recall from Lemma C1.4(ii) that a nonempty space is connected if and only if it cannot be partitioned into two nonempty *closed* subsets.

Lemma C2.3 *Let $a, b \in \mathbb{R}$ with $a < b$. Then $[a, b]$ cannot be written as a disjoint union of closed subsets A and B such that $a \in A$ and $b \in B$.*

Proof Suppose that such A and B do exist. Since $[a, b]$ is closed in $\mathbb{R}$, Lemma A9.7 implies that A and B are also closed in $\mathbb{R}$.

The set $A \subseteq \mathbb{R}$ is nonempty and bounded above by b, so has a supremum $s \leq b$. The supremum of a subset of $\mathbb{R}$ always lies in its closure in $\mathbb{R}$ (Example A8.8(iv)), so $s \in A$. Since $b \in B$ and $A \cap B = \emptyset$, it follows that $s \neq b$, so $s < b$.

Now $(s, b]$ is a subset of $[a, b]$ disjoint from A, so $(s, b] \subseteq B$. But B is closed in $\mathbb{R}$, so B contains the closure of $(s, b]$ in $\mathbb{R}$, which is $[s, b]$. Hence $s \in A \cap B = \emptyset$, a contradiction. □

Remark C2.4 When we have a topological space X and subsets $A \subseteq I \subseteq X$, the notation $\mathrm{Cl}(A)$ does not specify whether the closure is intended to be taken in I or in X. For this reason, I have avoided this notation in the proof above (where $I = [a, b]$ and $X = \mathbb{R}$). If we wanted to distinguish the two, we could write $\mathrm{Cl}_I(A)$ and $\mathrm{Cl}_X(A)$. There can be a difference: e.g. if $A = I = (0, 1)$ and $X = \mathbb{R}$ then $\mathrm{Cl}_I(A) = (0, 1)$ and $\mathrm{Cl}_X(A) = [0, 1]$. But when I is closed in X, there is no difference, by Lemma A9.7.

Theorem C2.5 *A subset of $\mathbb{R}$ is connected if and only if it is a nonempty interval.*

Proof First let X be a connected subset of $\mathbb{R}$. Certainly X is nonempty. Assume for a contradiction that X is not an interval; then we can choose $x \leq y \leq z$ with $x, z \in X$ but $y \notin X$. Then $(-\infty, y) \cap X$ and $(y, \infty) \cap X$ are disjoint open subsets of X whose union is X, and are both nonempty since x belongs to the first and z to the second. This contradicts X being connected.

Conversely, let $I \subseteq \mathbb{R}$ be a nonempty interval. Let Y and Z be disjoint closed subsets of I with $Y \cup Z = I$. Suppose for a contradiction that Y and Z are nonempty; choose $a \in Y$ and $b \in Z$. We may assume without loss of generality that $a < b$ (otherwise swap the names of Y and Z). Put $A = Y \cap [a, b]$ and $B = Z \cap [a, b]$. Then $[a, b]$ is the disjoint union of closed subsets A and B with $a \in A$ and $b \in B$, contradicting Lemma C2.3. □

Corollary C2.6 (Intermediate value theorem) *Let $f\colon [a, b] \to \mathbb{R}$ be a continuous function, and let $y \in \mathbb{R}$ with $f(a) \leq y \leq f(b)$ or $f(b) \leq y \leq f(a)$. Then there exists $c \in [a, b]$ such that $f(c) = y$.*

Proof By the 'if' part of Theorem C2.5, $[a, b]$ is connected, so by continuity of f and Proposition C1.7, $f[a, b]$ is connected. Now by the 'only if' part of Theorem C2.5, $f[a, b]$ is an interval. The result now follows from the definition of interval. □

Remark C2.7 Here, we have derived the intermediate value theorem from the connectedness of intervals. But the chain of reasoning can also be reversed: the connectedness of intervals can be derived from the intermediate value theorem.

Indeed, assume the intermediate value theorem and let I be a nonempty interval. To prove that I is connected, take a continuous map f from I to the two-point discrete space $\{0, 1\}$. We can regard $\{0, 1\}$ as a subset of $\mathbb{R}$, and the subspace topology is discrete, so the inclusion $i\colon \{0, 1\} \to \mathbb{R}$ is continuous. Hence $i \circ f\colon I \to \mathbb{R}$ is a continuous map taking only the values 0 and 1. The intermediate value theorem then implies that $i \circ f$, hence f, is constant.

Which one to prove first is a matter of taste. My own feeling is that the connectedness of intervals is a more fundamental fact than the intermediate value theorem, and should therefore occupy the primary position.

If you take Algebraic Topology, you will meet the excellent Brouwer fixed point theorem. This says that every continuous map $\bar{B}^n \to \bar{B}^n$ has a fixed point, where

$$\bar{B}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leq 1\}$$

is the closed Euclidean n-ball. (A **fixed point** of a map $f\colon X \to X$ is a point $x \in X$ such that $f(x) = x$.) We do not have the tools to prove this, but we can at least prove the one-dimensional case.

Corollary C2.8 (One-dimensional Brouwer fixed-point theorem) *Every continuous map $[0, 1] \to [0, 1]$ has at least one fixed point.*

Proof Let $f\colon [0, 1] \to [0, 1]$ be a continuous map. Define $g\colon [0, 1] \to \mathbb{R}$ by $g(x) = f(x) - x$. Then $g(1) \leq 0 \leq g(0)$, so by the intermediate value theorem, $g(c) = 0$ for some $c \in [0, 1]$. Hence $f(c) = c$. □

Similarly, methods of algebraic topology can be used to show that $\mathbb{R}^n$ and $\mathbb{R}^m$ are never homeomorphic unless $n = m$. Indeed, something stronger is true: if there is a continuous injection $\mathbb{R}^n \to \mathbb{R}^m$ then $n \leq m$. (Why does this imply the result about homeomorphism?) But again, all we are capable of proving with our current technology is the one-dimensional case, Proposition C2.10 below.

Lemma C2.9 *There is no continuous injection $S^1 \to \mathbb{R}$.*

Proof Suppose for a contradiction that there is a continuous injection $f : S^1 \to \mathbb{R}$. Choose distinct points $a, b, c \in S^1$. Without loss of generality, $f(a) < f(b) < f(c)$. Let I be the arc between a and c in S^1 that does not contain b (but does contain its endpoints a and c). Then I is homeomorphic to a real interval, and its image under f contains $f(a)$ and $f(c)$, so by the intermediate value theorem, it also contains $f(b)$. That is, there exists $b' \in I$ such that $f(b') = f(b)$. But $b' \neq b$ (since $b \notin I$), so this contradicts f being injective. □

Proposition C2.10 *Let $n \geq 2$. Then there is no continuous injection $\mathbb{R}^n \to \mathbb{R}$.*

Proof There is a continuous injection $S^1 \to \mathbb{R}^n$, so if there is a continuous injection $\mathbb{R}^n \to \mathbb{R}$ then by composing, there is a continuous injection $S^1 \to \mathbb{R}$, contradicting Lemma C2.9. □

In case it seems that we have just proved the obvious, note that there *is* a continuous surjection $\mathbb{R} \to \mathbb{R}^n$ for every $n \geq 2$.

Connectedness can be a helpful tool in showing that two spaces are not homeomorphic. Certainly, if one of the spaces is connected and the other is not, then they are not homeomorphic. But even in situations where both spaces are connected, we can use connectedness to distinguish between them. This is best explained by some examples.

Examples C2.11 i. The intervals $[0, 1)$ and $(0, 1)$ are not homeomorphic. To prove this, let us temporarily call a space X 'good' if it has the following property:

there exists $x \in X$ such that $X \setminus \{x\}$ is connected.

Clearly goodness is a topological property. The space $[0, 1)$ is good, since $[0, 1) \setminus \{0\}$ is the connected space $(0, 1)$. On the other hand, $(0, 1)$ is not good, since for all $x \in X$, the subspace $(0, 1) \setminus \{x\} = (0, x) \cup (x, 1)$ of $\mathbb{R}$ is not an interval and therefore not connected (by Theorem C2.5). So $[0, 1)$ and $(0, 1)$ are not homeomorphic.

ii. The letters T and L are not homeomorphic. (View both as closed subspaces of $\mathbb{R}^2$.) One way to prove this would be to consider the following topological property of a space X:

there exists $x \in X$ such that $X \setminus \{x\}$ is in three pieces.

Although this seems to be satisfied by T but not L, the problem is that we do not yet have a precise way of saying 'in three pieces'. (We will very soon.) So let us consider this topological property of a space X instead:

there exist distinct $x_1, x_2, x_3 \in X$ such that $X \setminus \{x_i\}$ is connected for each $i = 1, 2, 3$.

We could call a point x of a space X 'removable' if $X \setminus \{x\}$ is connected; then the property is that X has at least three removable points. The space T has the property, since its three endpoints are removable. On the other hand, L does not, as if we take three distinct points of L then at least one is not an endpoint, and is therefore not removable. So T and L are not homeomorphic.

iii. We could attempt to prove that $\mathbb{R}^2$ and $\mathbb{R}^3$ are not homeomorphic by considering the following topological property of a space X:

> there exists a subspace Y of X such that $Y \cong \mathbb{R}$ and $X \setminus Y$ is disconnected.

Certainly $\mathbb{R}^2$ has this property (e.g. take Y to be the y-axis). It seems highly implausible that there could be a subspace Y of $\mathbb{R}^3$ such that $Y \cong \mathbb{R}$ and $\mathbb{R}^3 \setminus Y$ is disconnected, but it is not so easy to prove. Again, this kind of problem is best handled using the techniques of algebraic topology.

iv. A different strategy for proving that $\mathbb{R}^2$ and $\mathbb{R}^3$ are not homeomorphic is to consider this topological property of a space X:

> for every subspace L of X homeomorphic to S^1, the complement $X \setminus L$ is disconnected.

Certainly $\mathbb{R}^3$ does not have this property (that is, you can find a homeomorphic copy of the circle in $\mathbb{R}^3$ whose complement is connected). It seems plausible that $\mathbb{R}^2$ does have the property, because every closed curve in the plane should have an inside and an outside. But this is actually quite hard to prove. It follows from the Jordan curve theorem, a deep result that will be stated in the final section of the course.

C3 Path-connectedness

For Lecture 19

The physical space we exist in looks something like $\mathbb{R}^3$, the earth we walk on resembles $\mathbb{R}^2$ (at least over small distances), and we perceive time as moving along a line $\mathbb{R}$. For these and other, more subtle, reasons, Euclidean spaces $\mathbb{R}^n$ have a special place in mathematics as done by human beings.

For instance, a topological space is said to be an **n-dimensional manifold** if it is Hausdorff and has an open cover by subsets each homeomorphic to an open ball in $\mathbb{R}^n$. Typical examples of 2-dimensional manifolds (surfaces) are the sphere, the torus and the Klein bottle. Manifolds are enormously important in mathematics, especially in subjects such as mathematical physics and algebraic topology.

If we view topology through the lens of Euclidean space, it looks rather different. For instance, spaces with the cofinite or Zariski topology are so unlike Euclidean space that it becomes natural to simply ignore them (perhaps by restricting attention to Hausdorff spaces). And, as we are about to see, the Euclidean point of view suggests a variant on the notion of connectedness.

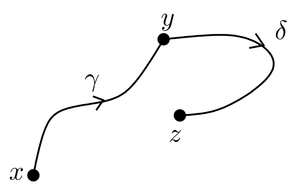


Figure C.1: Concatenating paths.

Definition C3.1 i. Let X be a topological space. A **path** in X is a continuous map $\gamma\colon [0,1] \to X$. If $\gamma(0) = x$ and $\gamma(1) = y$, we say that γ is a path **from** x **to** y.

ii. A space X is **path-connected** if it is nonempty and for all $x, y \in X$, there exists a path from x to y in X.

Lemma C3.2 *Every path-connected space is connected.*

Proof Let f be a continuous map from a path-connected space X to a discrete space D. By Lemma C1.4, it is enough to show that f is constant. Indeed, let $x, y \in X$. There is a path $\gamma\colon [0,1] \to X$ from x to y, and $f \circ \gamma$ is then a continuous map $[0,1] \to D$. But $[0,1]$ is connected, so $f \circ \gamma$ is constant, and in particular $f(x) = f(\gamma(0)) = f(\gamma(1)) = f(y)$, as required. □

It is sometimes quite hard to prove that a connected space really is connected. For example, how would you show that a disk in $\mathbb{R}^2$ cannot be expressed as a union of disjoint nonempty open subsets? The lemma we have just proved can be very useful in such situations:

Example C3.3 A subset X of $\mathbb{R}^n$ is **convex** if for all $x, y \in X$ and $t \in [0,1]$, we have $(1-t)x + ty \in X$. (For example, the convex subsets of $\mathbb{R}$ are precisely the intervals; see Definition C2.1.) Every nonempty convex subset of $\mathbb{R}^n$ is path-connected, since $t \mapsto (1-t)x + ty$ defines a path from x to y in X. Hence every nonempty convex subset of $\mathbb{R}^n$ is connected.

Given paths γ from x to y and δ from y to z in a space X, we can join them together (**concatenate** them) to form a new path $\gamma * \delta$ from x to z (Figure C.1). If we think of γ as a point moving from x to y over a period of one second, and similarly δ, then $\gamma * \delta$ performs γ at double speed in the first half-second, followed by δ at double speed in the second half-second. Formally, the **concatenation** $\gamma * \delta\colon [0,1] \to X$ is defined by

$$(\gamma * \delta)(t) = \begin{cases} \gamma(2t) & \text{if } t \in [0, 1/2] \\ \delta(2t-1) & \text{if } t \in [1/2, 1]. \end{cases}$$

Note that the two cases agree at $t = 1/2$, and that $\gamma * \delta$ is continuous (that is, a path) by Lemma A9.13.

Example C3.4 For $n \geq 2$, the space $X = \mathbb{R}^n \setminus \{0\}$ is path-connected. Indeed, let $x, y \in \mathbb{R}^n$. If x, y and 0 are not collinear, the straight line segment from x to y defines a path from x to y in X. If they are collinear, we may choose a point

$z \in X$ not on the straight line containing x, y and 0 (since $n \geq 2$). The straight line from x to z does not pass through 0, and therefore defines a path γ from x to z in X. Similarly, the straight line from z to y defines a path δ from z to y in X. So $\gamma * \delta$ is a path from x to y in X, as required.

Since $\mathbb{R} \setminus \{x\}$ is not connected (let alone path-connected) for any $x \in \mathbb{R}$, this provides another proof that $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}$ when $n \geq 2$. (Compare Proposition C2.10.)

We have shown that path-connectedness implies connectedness. The converse is false: there are topological spaces (even subsets of $\mathbb{R}^n$) that are connected but not path-connected. So path-connectedness is a stronger condition.

To prove this, we use a lemma that has nothing to do with connectedness.

Lemma C3.5 *Let $f\colon X \to Y$ be a continuous map between topological spaces. Then the subspace*

$$\{(x, y) \in X \times Y : f(x) = y\} \tag{C:2}$$

of $X \times Y$ is homeomorphic to X.

This subspace (C:2) is called the **graph** of f. (If you don't see why, draw a picture!)

Proof Write $\Gamma_f = \{(x, y) \in X \times Y : f(x) = y\}$. Define functions

$$\begin{array}{rccc} p\colon & \Gamma_f & \to & X \\ & (x, y) & \mapsto & x, \end{array} \qquad \begin{array}{rccc} q\colon & X & \to & \Gamma_f \\ & x & \mapsto & (x, f(x)). \end{array}$$

Evidently $p \circ q = \mathrm{id}_X$ and $q \circ p = \mathrm{id}_{\Gamma_f}$, so it remains to show that p and q are continuous.

Write $i\colon \Gamma_f \to X \times Y$ for the inclusion function (as defined in Remark A9.9). Then p is the composite of $i\colon \Gamma_f \to X \times Y$ with the first projection map $X \times Y \to X$. Both the inclusion and the projection are continuous, so p is continuous.

To show that q is continuous, it is enough to show that $i \circ q\colon X \to X \times Y$ is continuous (by Theorem A9.11). But $(i \circ q)(x) = (x, f(x)) = (\mathrm{id}_X(x), f(x))$ $(x \in X)$, and both id_X and f are continuous, so q is continuous by Theorem A10.10.□

Example C3.6 Define subspaces L, C and X of $\mathbb{R}^2$ by

$$\begin{aligned} L &= \{(0, y) \in \mathbb{R}^2 : -1 \leq y \leq 1\}, \\ C &= \{(x, \sin(1/x)) \in \mathbb{R}^2 : x > 0\}, \\ X &= L \cup C. \end{aligned}$$

The space X is called the **topologist's sine curve** (Figure C.2). We show that it is connected but not path-connected.

To show that is connected, first note that C is homeomorphic to $(0, \infty)$, by Lemma C3.5. Hence C is connected. Claim: the closure of C in $\mathbb{R}^2$ contains L. Proof: let $(0, y) \in L$ (so that $y \in [-1, 1]$) and $\varepsilon > 0$. There exists $z > 1/\varepsilon$ such that $\sin z = y$. Putting $x = 1/z$ gives $(x, y) \in C$ and $d_2((0, y), (x, y)) = |x| < \varepsilon$, proving the claim. Hence the closure of C in $\mathbb{R}^2$ contains X. (In fact, it *is* X, although we do not need to know this.) Since C is connected, Lemma C1.5 implies that X is connected too.

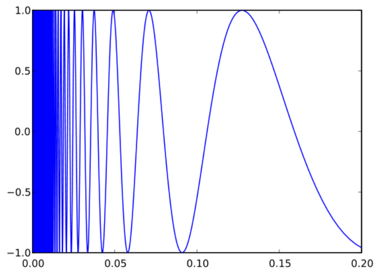


Figure C.2: The topologist's sine curve.

To prove that X is not path-connected, we show that there is no path in X from $(0,0)$ to $(1/\pi, 0)$. Suppose, for a contradiction, that γ is such a path, and write $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ $(t \in [0,1])$. By Theorem A10.10, the maps $\gamma_1, \gamma_2 \colon [0,1] \to \mathbb{R}$ are continuous.

The subset $\gamma_1^{-1}\{0\}$ of $[0,1]$ is nonempty and bounded above, so has a supremum $s \in [0,1]$. It is also closed, so $s \in \gamma_1^{-1}\{0\}$, which implies that $s < 1$.

(The idea now is as follows. Consider the movement of the point $\gamma(t)$ on X as t increases from 0 to 1. At the moment when t increases just beyond s, the point escapes from L onto C, never to return. However, the oscillation of C is so frantic that this is incompatible with continuity.)

Since γ_2 is continuous at s, there exists $\delta \in (0, 1-s)$ such that $|\gamma_2(t) - \gamma_2(s)| < 1$ for all $t \in [s, s+\delta)$.

First suppose that $\gamma_2(s) \leq 0$. We have $\gamma_1(s+\delta) > 0$ by definition of s, so there exists $x \in (0, \gamma_1(s+\delta))$ such that $\sin(1/x) = 1$. (Specifically, we can take $x = 1/((n+\frac{1}{2})\pi)$ for sufficiently large n.) Now $0 = \gamma_1(s)$ and γ_1 is continuous, so by the intermediate value theorem, there exists $t \in (s, s+\delta)$ such that $\gamma_1(t) = x$. Since $\gamma(t) \in X$, this implies that $\gamma_2(t) = \sin(1/x) = 1$. But $\gamma_2(s) \leq 0$, so $|\gamma_2(t) - \gamma_2(s)| \geq 1$ with $t \in [s, s+\delta)$, a contradiction.

A similar argument produces a contradiction if $\gamma_2(s) \geq 0$, using the fact that there are arbitrarily small positive solutions to $\sin(1/x) = -1$.

On the other hand, there are reasonable hypotheses under which connected spaces *are* automatically path-connected. We first give a necessary and sufficient condition for a general space to be path-connected, then give a sufficient condition for subsets of $\mathbb{R}^n$.

Proposition C3.7 *Let X be a topological space. Then X is path-connected if and only if X is connected and every point of X has at least one path-connected neighbourhood.*

Proof If X is path-connected then as we have already seen, X is connected, and X itself is a path-connected neighbourhood of every point.

Conversely, suppose that X is connected and that every point has a path-connected neighbourhood. Let $x \in X$, and write

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

We show that both U and $X \setminus U$ are open in X. Since $x \in U$, it will follow that $U = X$, and therefore that X is path-connected.

To show that U is open, let $y \in U$. Then we may choose a path γ from x to y. Also, by hypothesis, we may choose a path-connected neighbourhood W of y. For each $w \in W$, there is a path from y to w; concatenating it with γ gives a path from x to w. Hence $W \subseteq U$. So by Lemma A2.9, U is open in X.

The argument that $X \setminus U$ is open is similar. Let $y \in X \setminus U$. By hypothesis, we may choose a path-connected neighbourhood W of y. Then $W \subseteq X \setminus U$: for if $w \in U \cap W$ then there exist paths from x to w and from w to y, which when concatenated give a path from x to y, contradicting the fact that $y \notin U$. So by Lemma A2.9 again, $X \setminus U$ is open in X. □

Corollary C3.8 *Every connected open subset of* $\mathbb{R}^n$ *is path-connected.*

Proof Let U be a connected open subset of $\mathbb{R}^n$. For each $x \in U$, we can choose $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$; and $B(x, \varepsilon)$ is convex, hence path-connected. So by Proposition C3.7, U is path-connected. □

Path-connectedness has some of the same convenient properties as connectedness:

Lemma C3.9 *Let* $f\colon X \to Y$ *be a continuous map of topological spaces. If* X *is path-connected then so is* fX.

Proof Let $y, y' \in fX$. Then $y = f(x)$ and $y' = f(x')$ for some $x, x' \in X$. Since X is path-connected, there is a path $\gamma\colon [0,1] \to X$ from x to x' in X. Then $f \circ \gamma\colon [0,1] \to Y$ is a path from y to y' in Y. □

Proposition C3.10 *The product of two path-connected spaces is path-connected.*

Proof Exercise. □

C4 Connected-components and path-components

For Lecture 20

Suppose we have before us two spaces, X and Y. If X is connected and Y is not, then we know that they cannot be homeomorphic. However, what if both are disconnected, but they fall into different number of pieces? We would like to be able to conclude then that X is not homeomorphic to Y. For instance, Example C2.11(ii) was one such situation, but we did not have the formal language to make the argument precise. We develop it here.

Let X be a topological space. The **connectedness relation** on X is the relation $\sim$ on X defined by $x \sim y$ if and only if there exists a connected subspace $C \subseteq X$ such that $x, y \in C$.

Lemma C4.1 *The connectedness relation on a topological space is an equivalence relation.*

Proof Let X be a topological space. For reflexivity: given $x \in X$, the subspace $\{x\}$ of X is connected and contains x. Symmetry is immediate. For transitivity, let $x, y, z \in X$, and suppose we have connected subspaces C and D of X such that $x, y \in C$ and $y, z \in D$. By Lemma C1.11, $C \cup D$ is connected too, and $x, z \in C \cup D$. □

The equivalence classes of the connectedness relation on X are called the **connected-components** of X. So each point $x \in X$ is contained in precisely one connected-component, which is called the connected-component of x.

Warning C4.2 If x and y are in different connected-components of X, this does *not* guarantee that there are disjoint open sets U and V such that $x \in U$, $y \in V$ and $U \cup V = X$! It only means that there is no connected subspace of X containing both x and y.

If such U and V do exist then x and y *are* in different connected-components of X (since if $x, y \in C \subseteq X$ then $U \cap C$ and $V \cap C$ partition C into nonempty open subsets, so C is disconnected). But the converse fails (a hard exercise).

The next lemma expresses the idea that splitting a space into its connected-components amounts to dividing it into the 'biggest possible connected chunks'.

Lemma C4.3 *Let X be a topological space. Then:*

i. *Every connected-component of X is connected.*

ii. *Every connected-component C of X is a maximal connected subspace; that is, the only connected subspace of X containing C is C itself.*

iii. *Every maximal connected subspace of X is a connected-component.*

So, the connected-components of a nonempty space are exactly the maximal connected subspaces.

Proof Write $\sim$ for the connectedness relation on X.

For (i), let C be a connected-component of X. Since C is an equivalence class, it is nonempty, so we can choose some $x \in C$. For each $y \in C$, we have $x \sim y$, so there exists a connected subspace D_y of X such that $x, y \in D_y$. For each $y \in C$, we have $x \sim z$ for all $z \in D_y$; but C is the equivalence class of x, so $D_y \subseteq C$. Hence $C = \bigcup_{y \in C} D_y$. By Lemma C1.11, this union is connected, so C is connected.

For (ii), let C be a connected-component of X and D a connected subspace of X containing C. Again, since C is an equivalence class, it is nonempty, so we can choose some $x \in C$. Then $x \sim y$ for all $y \in D$, by definition of $\sim$. But C is the equivalence class of x, so $D \subseteq C$, so $D = C$.

For (iii), let C be a maximal connected subspace of X. Then C is nonempty (being connected), so we can choose a point $x \in C$. We have $x \sim y$ for every $y \in C$ (since C is a connected subspace containing x and y), so C is a subset of the connected-component $[x]$ of x. On the other hand, $[x]$ is connected, so by maximality, $C = [x]$. □

Examples C4.4 i. The connected-components of $\mathbb{R} \setminus \{0\}$ are $(-\infty, 0)$ and $(0, \infty)$. Indeed, $(-\infty, 0)$ is certainly a connected subspace, and it is a *maximal* connected subspace, since any subspace of $\mathbb{R} \setminus \{0\}$ strictly containing $(-\infty, 0)$ must have a positive element, and is therefore disconnected.

ii. Similarly, the connected-components of $\mathbb{R} \setminus \{0, 1\}$ are $(-\infty, 0)$, $(0, 1)$ and $(1, \infty)$.

A number is prime if and only if it has exactly one prime factor. Similarly:

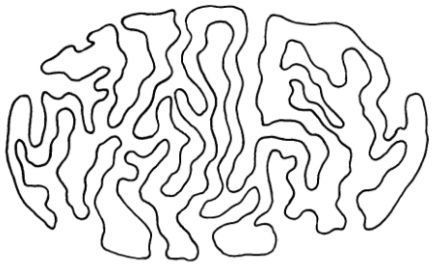


Figure C.3: Which points are inside and which are outside? (Image from Saunders Mac Lane, *Mathematics: Form and Function.*)

Lemma C4.5 *A topological space is connected if and only if it has exactly one connected-component.*

Proof Let X be a topological space, and write $\sim$ for the connectedness relation on X.

Suppose that X is connected. Then by definition, X is nonempty, so has at least one connected-component. On the other hand, any two points $x, y \in X$ are contained in the connected subspace X of X, so $x \sim y$. Hence X has exactly one connected-component.

Conversely, suppose that X has exactly one connected-component. This connected-component must be X itself, so by Lemma C4.3(i), X is connected. $\square$

Lemma C1.5 states that if you add some limit points to a connected subspace, it remains connected. We deduce:

Lemma C4.6 *Every connected-component of a topological space is closed.*

Proof Let C be a connected-component of a topological space X. By Lemma C1.5, $\mathrm{Cl}(C)$ is also connected, and of course $C \subseteq \mathrm{Cl}(C)$. But C is a maximal connected subspace of X, so $\mathrm{Cl}(C) = C$. Hence C is closed. $\square$

Here is a major theorem. It is the intuitively obvious statement that every closed curve (loop) in the plane has an inside and an outside (Figure C.3). Once again, this apparently obvious theorem is surprisingly hard to prove, and the proof is usually done using the tools of algebraic topology. We do not attempt it here.

Theorem C4.7 (Jordan curve theorem) *Let L be a subspace of $\mathbb{R}^2$ homeomorphic to the circle. Then $\mathbb{R}^2 \setminus L$ has exactly two connected-components, one bounded and one unbounded.*

Some spaces are highly fragmented:

Definition C4.8 A topological space is **totally disconnected** if every connected-component has only one element.

Equivalently, a space is totally disconnected if and only if every connected subspace has exactly one element. (Exercise: why is this equivalent?)

Examples C4.9 i. Every discrete space is totally disconnected, since every subspace of a discrete space is discrete, and the only connected discrete space is the one-element space (Example C1.3(iii)).

ii. $\mathbb{Q}$ is totally disconnected (but not discrete, since $\{0\}$ is not open). Indeed, let $x, y \in \mathbb{Q}$ with $x \neq y$. Then $x \not\sim y$: for we can choose an irrational number $u \in \mathbb{R}$ between x and y, and then whenever C is a subset of $\mathbb{Q}$ containing x and y, we have disjoint nonempty open subsets $(-\infty, u) \cap C$ and $(u, \infty) \cap C$ of C whose union is C, proving that C is disconnected.

Write $K(X)$ for the set of connected-components of X. (This is not standard notation; there is none, as far as I know.) It is clear in principle that the set of connected-components is a topological invariant; that is, if the spaces X and Y are homeomorphic then the sets $K(X)$ and $K(Y)$ are in bijection. This is because $K(X)$ is defined purely in terms of the topological structure of X.

So, for instance, if X has three connected-components and Y has two, then X and Y are not homeomorphic, since there is no bijection between the sets $K(X)$ and $K(Y)$. For example, $\mathbb{R} \setminus \{0, 1\}$ and $\mathbb{R} \setminus \{0\}$ are not homeomorphic.

In fact, we'll see that *any* continuous map $X \to Y$ (not necessarily a homeomorphism) induces a function $K(X) \to K(Y)$ (not necessarily a bijection).

Lemma C4.10 *Let $f\colon X \to Y$ be a continuous map and $x, x' \in X$. If x and x' are in the same connected-component of X then $f(x)$ and $f(x')$ are in the same connected-component of Y.*

Proof Let x and x' be points in the same connected-component of X. Then there is some connected subspace $C \subseteq X$ such that $x, x' \in C$. By Proposition C1.7, fC is a connected subspace of Y, and clearly $f(x), f(x') \in fC$. $\square$

(The image of a connected-component under a continuous map must be connected, but need not be a connected-*component*. Can you think of an example?)

Write $[x]$ for the connected-component of a point $x \in X$. Lemma C4.10 states that given a continuous map $f\colon X \to Y$, if $[x] = [x']$ then $[f(x)] = [f(x')]$. So we can legitimately define a function

$$K(f)\colon K(X) \to K(Y)$$

by

$$K(f)([x]) = [f(x)]$$

($x \in X$). Thus, a continuous map between topological spaces gives rise to a function between their sets of connected-components.

This process behaves well with respect to composition and identities:

Lemma C4.11 *i. Let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be continuous maps of topological spaces. Then $K(g \circ f) = K(g) \circ K(f)$.*

ii. Let X be a topological space. Then $K(\mathrm{id}_X) = \mathrm{id}_{K(X)}$.

Proof For (i), both $K(g \circ f)$ and $K(g) \circ K(f)$ are functions $K(X) \to K(Z)$. Let $x \in X$; we must show that

$$K(g \circ f)([x]) = (K(g) \circ K(f))([x]).$$

The left-hand side is $[(g \circ f)(x)] = [g(f(x))]$. The right-hand side is

$$K(g)\big(K(f)([x])\big) = K(g)([f(x)]) = [g(f(x))].$$

So the two sides are equal.

For (ii), both $K(\mathrm{id}_X)$ and $\mathrm{id}_{K(X)}$ are functions $K(X) \to K(X)$. Let $x \in X$. Then

$$K(\mathrm{id}_X)([x]) = [\mathrm{id}_X(x)] = [x] = \mathrm{id}_{K(X)}([x]),$$

as required. □

In the jargon, K is a **functor** from spaces to sets. Roughly, this means that K is a method for turning (i) spaces into sets and (ii) continuous maps between spaces into functions between sets, with the properties in the lemma just proved. If you take Algebraic Topology, you will meet some other important functors, mostly turning spaces into groups.

Corollary C4.12 *Let $f\colon X \to Y$ be a homeomorphism. Then $K(f)\colon K(X) \to K(Y)$ is a bijection.*

Proof By Lemma C4.11,

$$K(f) \circ K(f^{-1}) = K(f \circ f^{-1}) = K(\mathrm{id}_Y) = \mathrm{id}_{K(Y)},$$

and similarly $K(f^{-1}) \circ K(f) = \mathrm{id}_{K(X)}$. So $K(f^{-1})$ is a two-sided inverse to $K(f)$, from which it follows that $K(f)$ is a bijection. □

So we have now proved formally that if X and Y are homeomorphic then there is a bijection between $K(X)$ and $K(Y)$.

The notion of connected-component arises from the more basic notion of connectedness. This process can be imitated with *path*-connectedness in place of connectedness. Here is a quick sketch of how this works.

Let X be a topological space. The **path relation** on X is the relation $\approx$ on X defined by $x \approx y$ if and only if there exists a path from x to y in X.

Lemma C4.13 *The path relation on a topological space is an equivalence relation.*

Proof Let X be a topological space. Reflexivity follows from the fact that for every $x \in X$, the map $[0, 1] \to X$ with constant value x is a path from x to x. Symmetry follows from the fact that if γ is a path from x to y in X then γ' is a path from y to x in X, where we define $\gamma'(t) = \gamma(1-t)$ $(t \in [0, 1])$. Transitivity follows from the fact that if γ is a path from x to y and δ is a path from y to z then their concatenation $\gamma * \delta$ is a path from x to z. □

The equivalence classes of the path relation on X are called the **path-components** of X. Very much as in Lemma C4.3, they are the maximal path-connected subspaces of X. As in Lemma C4.5, a space is path-connected if and only if it has exactly one path-component. But in contrast to Lemma C4.6, the path-components of a space need *not* be closed: for instance, the path-components of the topologist's sine curve X (Example C3.6) are L and C, and C is not closed in X.

We write $\pi_0(X)$ for the set of path-components of a topological space X. (This *is* standard notation.) In algebraic topology, there are also sets called $\pi_n(X)$ for each $n \geq 1$. In fact, $\pi_n(X)$ is naturally a group for $n \geq 1$. These are the **homotopy groups** of X.

Just as for connected-components, we have:

Lemma C4.14 *Let $f\colon X \to Y$ be a continuous map and $x, x' \in X$. If x and x' are in the same path-component of X then $f(x)$ and $f(x')$ are in the same path-component of Y.*

Proof If x and x' are in the same path-component of X then we can find a path $\gamma\colon [0,1] \to X$ from x to x' in X. Then $f \circ \gamma\colon [0,1] \to Y$ is a path from $f(x)$ to $f(x')$ in Y. □

This enables us to define a function $\pi_0(f)\colon \pi_0(X) \to \pi_0(Y)$, sending the path-component of x in X to the path-component of $f(x)$ in Y. Just as for connected-components, this satisfies the equations

$$\pi_0(g \circ f) = \pi_0(g) \circ \pi_0(f), \qquad \pi_0(\mathrm{id}_X) = \mathrm{id}_{\pi_0(X)},$$

from which it follows that if $f\colon X \to Y$ is a homeomorphism then $\pi_0(f)\colon \pi_0(X) \to \pi_0(Y)$ is a bijection. In particular, the set of path-components is a topological invariant.

You will meet other, more powerful, topological invariants if you take Algebraic Topology next term.

* * *

Epilogue

This chapter will not be lectured, assessed or examined, and should be regarded as optional.

In some ways, we have reached our limit. We have been through all the basic concepts related to *general* topological spaces, covering the common core of topology that is used by many branches of mathematics. However, towards the end of the course, we increasingly often bumped up against statements about *particular* topological spaces that we were unable to prove with our existing techniques.

In this short chapter, we look at four large classes of topological spaces and note some of the things that we are unable to prove about them. We met representatives of all of these classes in the very first lecture (the slides for which are available as a separate file on Learn). Each of these classes points to one or more different branches of mathematics beyond the scope of this course, which you may be interested in pursuing.

'Nice' subsets of $\mathbb{R}^n$

By far the most commonly-studied spaces are those that arise as 'smooth' subspaces of Euclidean space $\mathbb{R}^n$: spheres, balls, toruses, planes, and so on. Because of the physical world we live in, these seem natural and appealing to us. Some mathematicians even use the word *pathological* (disease-related) to refer to spaces that are too much unlike these.

Algebraic topology provides extremely powerful tools for studying such spaces. Even a small amount of algebraic topology enables you to prove theorems that we have wanted to prove in this course but not been able to. For instance, it lets you prove the theorem that $\mathbb{R}^m \not\cong \mathbb{R}^n$ for $m \neq n$, and the Brouwer fixed point theorem, and the Jordan curve theorem. The fact that $\mathbb{R}^m \not\cong \mathbb{R}^n$ actually follows from another fundamental result, the *invariance of domain theorem*: any continuous injection from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^n$ is an open map. This is also provable using algebraic topology.

As the name of the subject suggests, its power comes from marrying topology with algebra. The standard pattern is to take a problem in topology, translate it into a problem about groups, solve that problem about groups, and translate that back into a solution to the original topological problem.

Function spaces

Much of our inspiration for topology comes from the space $\mathbb{R}^n$. A point of $\mathbb{R}^n$ can be regarded as a function from the set $\{1, \ldots, n\}$ into $\mathbb{R}$. Thus, $\mathbb{R}^n$ is the space of real-valued functions on $\{1, \ldots, n\}$.

One way of going beyond $\mathbb{R}^n$ is to think about real-valued functions of other kinds. For instance, back in Examples A1.2, we met the space $C[a, b]$ of continuous real-valued functions on the interval $[a, b]$, which comes up very naturally in analysis. We saw that it resembles $\mathbb{R}^n$ in the sense that one can define metrics d_1, d_2 and d_∞ on it (and, in fact, d_p whenever $1 \leq p \leq \infty$). But we also saw that it differs significantly from $\mathbb{R}^n$. For example, we saw that the metrics d_1, d_2 and d_∞ on $\mathbb{R}^n$ are all topologically equivalent (Example A3.4(ii)) but the metrics d_1, d_2 and d_∞ on $C[a, b]$ are not (Example A3.12).

Many other function spaces arise in analysis, such as the space $C^k(\mathbb{R}^n)$ of real-valued functions on $\mathbb{R}^n$ that are differentiable k times and have continuous kth derivative, or the space of holomorphic functions on any given open subset of $\mathbb{C}$. There are then more exotic relatives such as spaces of measures and distributions. Studying the spaces formed by functions (and measures etc.) turns out to be very fruitful; this is the subject known as functional analysis.

In most cases, the spaces concerned are *normed vector spaces* (vector spaces equipped with a norm). The norm gives rise to a metric, and thus a topology. However, some of these spaces carry more than one interesting topology.

For instance, it can be shown that in an infinite-dimensional normed vector space, the unit ball is never compact. At least, that's the case if you use the topology that comes from the norm. But there is another interesting and useful topology on certain normed vector spaces, called the weak* topology, and the Banach–Alaoglu theorem states that the unit ball *is* compact in the weak* topology even if the space is infinite-dimensional. Here 'weak' is used in much the same sense as in Remark A10.9 (and I'll leave the star as a mystery).

Hard-to-visualize spaces

Long ago in linear algebra, you learned that every square matrix A has a so-called adjugate matrix $\mathrm{adj}(A)$ (sometimes called the 'adjoint', confusingly). It has the property that if A is invertible then $A^{-1} = \frac{1}{\det(A)} \mathrm{adj}(A)$. Since $(AB)^{-1} = B^{-1}A^{-1}$ and $\det(AB) = \det(A)\det(B)$, it follows that

$$\mathrm{adj}(AB) = \mathrm{adj}(B)\,\mathrm{adj}(A)$$

when A and B are invertible. But the equation makes sense for arbitrary $n \times n$ matrices A and B, invertible or not. Is it true for all A and B?

I claim that it is. Here's a 'proof'. For a square matrix chosen at random, the determinant is a random scalar, so is probably not zero. So, a randomly-chosen square matrix is invertible. Hence (ahem) the set of invertible $n \times n$ matrices is dense in the space of all $n \times n$ matrices. Now adj can be viewed as a map from the space of $n \times n$ matrices to itself, and each entry of $\mathrm{adj}(A)$ is a polynomial in the entries of A, so adj is continuous. We know that $\mathrm{adj}(AB) = \mathrm{adj}(B)\,\mathrm{adj}(A)$ on the dense set of pairs (A, B) where A and B are invertible, and each side is continuous in A and B. Hence by Corollary A8.13 (or something like it), $\mathrm{adj}(AB) = \mathrm{adj}(B)\,\mathrm{adj}(A)$ for all A and B.

Obviously this 'proof' is full of enormous holes. Nevertheless, it has been understood since the 19th century that underlying it is a fundamentally sound principle. It has a wonderful old name: 'the principle of irrelevance of algebraic inequalities'. The 'proof' can be re-expressed like this: the equation we want to prove is a family of polynomial identities, which are known to hold as long as the algebraic 'inequality' $\det(A)\det(B) \neq 0$ is satisfied. The principle of irrelevance of algebraic inequalities tells us that we can simply ignore this condition (it's irrelevant!) and conclude that the equation *always* holds.

All of this can be made precise, and then becomes a statement about dense subsets in the Zariski topology. Closely related to the Zariski topology is the space $\mathrm{Spec}(R)$ that one can assign to any commutative ring R. (Its points are the prime ideals of R.) This is called the *spectrum* of R, a usage related to many other occurrences of the word in science, from 'spectrum' as the set of eigenvalues of a matrix to 'spectrum' as in the colours of the rainbow.

These spaces are hard to visualize but very important in algebraic geometry and number theory. For instance, the proof of Fermat's last theorem used a really enormous amount of machinery that started with this kind of topological space as a very, very basic ingredient.

The theory of ordered sets gives hard-to-visualize topological spaces of a completely different kind. These come up in certain branches of computer science, for instance. I'll leave you to look these up.

'Nasty' subsets of $\mathbb{R}^n$

For spaces that are not very like 'nice' subsets of $\mathbb{R}^n$, algebraic topology tends to be less effective. For example, non-Hausdorff spaces are mostly outside the scope of algebraic topology. However, present-day algebraic topology is also not much use when applied to subsets of $\mathbb{R}^n$ that are are highly irregular.

Such spaces arise naturally in the study of dynamical systems. Brownian motion is one example. Or consider, for instance, the map $f\colon \mathbb{C} \to \mathbb{C}$ defined by $f(z) = z^2 + i$. Given a point $z \in C$, apply f every time the clock ticks to get a sequence of points $z, f(z), f(f(z)), \ldots$. Depending on what z is, this sequence might converge to a point in the complex plane, or head to infinity, or settle into a cycle, or wander about chaotically. And a slight change to the starting point z might cause a qualitative change in its trajectory (e.g. changing it from converging to a point to wandering about chaotically). The set of starting points z that are unstable in this sense is called the *Julia set* of f, and it is a fractal: no matter how far you zoom in, it has an infinite amount of detail. This is what algebraic topologists would call a 'nasty' space.

That example of a dynamical system modelled time as a discrete sequence of steps, but continuous-time dynamical systems are also enormously important. These are usually presented as differential equations, and give rise in a natural way to other highly irregular subsets of $\mathbb{R}^n$ (often as 'attractors' of the system, such as the Lorenz attractor). With infinitely many loops and complications, these are again beyond the reach of algebraic topology.

The topological understanding of spaces like this is probably the least developed of the four classes of space that I have described here. Mathematics is never finished