2 Introduction to metric spaces
2.1 Metrics, pseudometrics and convergence of sequences

Notation: We denote the set $\{x \in \mathbb{R} : x \geq 0\}$ by $\mathbb{R}^+$.

So, in interval notation, $\mathbb{R}^+ = [0, \infty)$.

In particular, in this module $0 \in \mathbb{R}^+$.

Definition 2.1 Let X be a set. Then a **metric** on X is a function $d : X \times X \to \mathbb{R}^+$ which satisfies the following three conditions, which are known as the **metric space axioms**, for all x, y and z in X:
(1) $d(x, y) = 0$ if and only if $x = y$;
(2) $d(x, y) = d(y, x)$;
(3) $d(x, z) \leq d(x, y) + d(y, z)$.
A **pseudometric** on X is a function $d : X \times X \to \mathbb{R}^+$ such that, for all x, y and z in X, axioms (2) and (3) above are satisfied, and also
(1′) $d(x, x) = 0$.

Axiom (2) says that (pseudo)metrics are symmetric, and Axiom (3) is the **triangle inequality** or **triangle law** for (pseudo)metrics.

Definition 2.2 A **metric space** is a pair (X, d) where X is a set and d is a metric on X. Similarly, a **pseudometric space** is a pair (X, d) where X is a set and d is a pseudometric on X.

In this setting it is standard abuse of terminology to say ‘let X be a metric space’ or ‘let X be a pseudometric space’, in situations where the name of the (pseudo)metric on X is unimportant.

The easiest example of a pseudometric is the following.

Example 2.3 For any set X, there is a pseudometric on X given by $d(x, y) = 0$ for all x and y in X.

Example 2.4 (i) The **usual distance** in $\mathbb{R}$ is given by $d(x, y) = |x - y|$ for x and y in $\mathbb{R}$. This defines a metric on $\mathbb{R}$ which we call the **usual metric** on $\mathbb{R}$.

(ii) Similarly, for any subset A of $\mathbb{R}$ we define the **usual metric** on A to be the metric $d(x, y) = |x - y|$, for x and y in A.

(iii) The **usual metric** on $\mathbb{C}$ is given by $d(z, w) = |z - w|$ for z and w in $\mathbb{C}$. As for $\mathbb{R}$ above, the same formula is used to define the **usual metric** on any subset of $\mathbb{C}$.

(iv) For $n \in \mathbb{N}$, the **usual metric** on $\mathbb{R}^n$ is the **Euclidean distance**, denoted by d_2, which comes from Pythagoras's Theorem (or using the **Euclidean norm**): for $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, with $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$,

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^{n} |x_j - y_j|^2}.$$

Again, the same formula is used to define the **usual metric** on any subset of $\mathbb{R}^n$.

The same definition is used for convergence in pseudometric spaces, but there is a problem in that **limits in pseudometric spaces need not be unique**. This is not a problem in metric spaces: limits in metric spaces are unique when they exist.

Proposition 2.6 Let (X, d) be a metric space and let $(x_n) \subseteq X$. Then the sequence (x_n) has at most one limit in X with respect to d.

Definition 2.5 Let (X, d) be a metric space, let $x \in X$ and let $(x_n) \subseteq X$. We say that (x_n) **converges** to x in X, if with the usual notion of convergence in $\mathbb{R}$, $\lim_{n\to\infty} d(x_n, x) = 0$. If (x_n) converges to x, we write $x_n \to x$ **as** $n \to \infty$, or $\lim_{n\to\infty} x_n = x$. (If there is any ambiguity over the metric used, we say that the convergence is **with respect to** d.) We also say that the **limit as** $n \to \infty$ of the sequence (x_n) is x. If the sequence (x_n) does not converge, then we say that (x_n) **diverges**. In this case the notation $\lim_{n\to\infty} x_n$ does not mean anything. If we ever write $\lim_{n\to\infty} x_n = x$ we always mean that the sequence (x_n) converges and the limit of the sequence is x.

From now on, whenever we say that $\mathbf{x}$ is in $\mathbb{R}^n$ we will assume that $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and we will use corresponding notation for $\mathbf{y}$ or $\mathbf{z}$ in $\mathbb{R}^n$ (etc.).

Example 2.7 On $\mathbb{R}^n$, we may define the following metrics:

(i) $d_1(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} |x_j - y_j|$;

(ii) $d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_j - y_j| : 1 \le j \le n\}$;

(iii) (Generalization of d_1 and d_2) For any real number $p \in [1, \infty)$, we may define a metric d_p by

$$d_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{j=1}^{n} |x_j - y_j|^p \right)^{1/p}$$

Our next example of a metric is very different from the others we have met so far. It is often useful when you are looking for counterexamples, and it is valid on any set.

Example 2.8 Let X be any set. Then the **discrete metric** on X is the metric d defined on X by

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

We next consider the set of all continuous real-valued functions on $[0,1]$. **Note that this time the elements of our set are functions, and so we are discussing possible ways to define the distance between two functions.**

Notation: We denote the set of all continuous, real valued functions on $[0,1]$ by $C[0,1]$.

Example 2.9 Let $X = C[0,1]$, and let $p \in [1,\infty)$. Then we may define metrics d_∞, d_1, d_2 and (more generally) d_p on X as follows. For f and g in $C[0,1]$, we define

$$d_\infty(f,g) = \sup\{|f(t)-g(t)| : t \in [0,1]\},$$

$$d_1(f,g) = \int_0^1 |f(t)-g(t)|\,\mathrm{d}t,$$

$$d_2(f,g) = \sqrt{\int_0^1 |f(t)-g(t)|^2\,\mathrm{d}t}$$

and

$$d_p(f,g) = \left(\int_0^1 |f(t)-g(t)|^p\,\mathrm{d}t\right)^{1/p}.$$

In view of the boundedness theorem, we see that we may also write

$$d_\infty(f,g) = \max\{|f(t)-g(t)| : t \in [0,1]\}.$$

The metric d_1 is sometimes called the **area metric** while d_∞ is sometimes called the **uniform metric** or the **sup norm metric** (see the discussion of norms below). Convergence with respect to d_∞ is precisely the notion of uniform convergence for a sequence of functions, as discussed in G1BMAN.

We conclude this section with a rather different metric on $\mathbb{R}^2$.

Example 2.10 (The Jungle River Metric) The Jungle River Metric on $\mathbb{R}^2$, d_J, is defined as follows. Let $\mathbf{x} = (x_1,x_2)$ and $\mathbf{y} = (y_1,y_2)$ in $\mathbb{R}^2$. Then

$$d_J(\mathbf{x},\mathbf{y}) = \begin{cases} |y_2 - x_2|, & \text{if } x_1 = y_1; \\ |x_2| + |y_2| + |y_1 - x_1|, & \text{otherwise.} \end{cases}$$

2.2 Norms and normed spaces

Many of the metrics which we have met so far come from **norms** on vector spaces.

Definition 2.11 Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A **norm** on V is a function $\|\cdot\| : V \to \mathbb{R}^+$ satisfying the following three conditions **(normed space axioms)** for every scalar α and all $\mathbf{x}$ and $\mathbf{y}$ in V:
(1) $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$;
(2) $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$;
(3) $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$.

Note that the 'if' part of (1) is redundant: the fact that $\|\mathbf{0}\| = 0$ follows from (2) with $\alpha = 0$ and (for example) $\mathbf{x} = \mathbf{0}$.

Definition 2.12 A **normed space** is a pair $(V, \|\cdot\|)$ where V is a vector space over $\mathbb{R}$ or $\mathbb{C}$ and $\|\cdot\|$ is a norm on V.

As usual, where the name of the norm on V does not matter we will often say 'let V be a normed space'.

You have already met one standard example of a norm in the module G1BMAN: the Euclidean norm on $\mathbb{R}^n$.

Example 2.13 The **usual norm** on $\mathbb{R}^n$ is the **Euclidean norm** $\|\cdot\|_2$ defined, for $\mathbf{x} \in \mathbb{R}^n$, by

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2}.$$

We now see how norms induce metrics on vector spaces.

Proposition 2.14 Let $(V, \|\cdot\|)$ be a normed space. Then there is a metric d on V defined by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ (for $\mathbf{x}$ and $\mathbf{y}$ in V). This metric is called the **metric on V induced by $\|\cdot\|$.**

Example 2.15 Let $p \in [1, \infty)$.
(i) We may define norms $\|\cdot\|_\infty$, $\|\cdot\|_1$, and, more generally $\|\cdot\|_p$ on $\mathbb{R}^n$ as follows. For $\mathbf{x} \in \mathbb{R}^n$, we set

$$\|\mathbf{x}\|_\infty = \max_{1 \le j \le n} |x_j|, \quad \|\mathbf{x}\|_1 = \sum_{j=1}^{n} |x_j|$$

and

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}.$$

(ii) We may define norms $\|\cdot\|_\infty$, $\|\cdot\|_1$, and, more generally $\|\cdot\|_p$ on $C[0,1]$ as follows. For $f \in C[0,1]$, we set

$$\|f\|_\infty = \sup\{|f(t)| : t \in [0,1]\}$$
$$= \max\{|f(t)| : t \in [0,1]\},$$
$$\|f\|_1 = \int_0^1 |f(t)|\,\mathrm{d}t,$$
$$\|f\|_p = \left(\int_0^1 |f(t)|^p\,\mathrm{d}t\right)^{1/p}.$$

Every normed space V is thus also a metric space, using the metric induced by its norm. In this case, for $\mathbf{x} \in V$, the norm of $\mathbf{x}$ is equal to the distance from $\mathbf{x}$ to $\mathbf{0}$.

The next example gives the norms which correspond to metrics discussed earlier.

Normed spaces are a central theme in more advanced analysis, especially in Functional Analysis.

The norm $\|\cdot\|_\infty$ on $C[0,1]$ has some particularly useful properties.

It is often called the **sup norm** or the **uniform norm**, and (as mentioned above) the associated type of convergence for sequences of functions is **uniform convergence**.

2.3 Open balls in metric spaces

You have already met (in G1BMAN) open balls, open sets and closed sets in $\mathbb{R}^n$ (with respect to the Euclidean distance).

We shall see that it is easy to generalize these notions to metric spaces.

We begin by defining **open balls** in metric spaces.

Note that we have not yet defined open sets: this means that we will have to prove later that 'open balls' really are examples of open sets!

Definition 2.16 Let (X, d) be a metric space, let $x \in X$ and let $r > 0$. Then the **open ball in X centred on x and with radius r**, denoted by $B_X(x, r)$, is defined by

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

If there is no ambiguity over the metric space involved, we may write $B(x, r)$ instead.

We may simply replace the word 'metric' by 'pseudometric' in this definition in order to define open balls in pseudometric spaces.

We may also work with **closed balls** $\{y \in X : d(y, x) \leq r\}$ (as in G1BMAN), but we shall not introduce any special notation for these.

The open ball $B_X(x, r)$ depends on x, r, **the set X and the metric d used on X.**

The notation $B_{(X,d)}(x, r)$ would be even less ambiguous, but is perhaps too unwieldy to be practical.

Example 2.17 (i) As in G1BMAN, if you give $\mathbb{R}$ and $\mathbb{R}^2$ the usual metrics then open balls in $\mathbb{R}$ are open intervals and open balls in $\mathbb{R}^2$ are the usual open discs.

(ii) Similarly, if you give $\mathbb{C}$ the usual metric, then open balls are the usual open discs.

(iii) If X is $\mathbb{R}^2$ with the metric d_1, what do open balls look like? For example, what is $B_X(\mathbf{0}, 1)$?

(iv) Let A be a subset of $\mathbb{R}$ and give A the usual metric. What are the open balls in A? In particular, what are the open balls in $\mathbb{Q}$ when $\mathbb{Q}$ is given the usual metric?

(iv) Let X be $C[0, 1]$ with the metric d_∞, let $f \in C[0, 1]$ and let $r > 0$. How can we describe $B_X(f, r)$?

(v) If instead we take X to be $C[0, 1]$ with the metric d_1, the open balls are very strange: I do not know of any particularly illuminating way to describe them.

We may describe convergence of sequences in terms of open balls.

Informally, a sequence (x_n) converges to a point x if for every open ball B centred on x (no matter how small the radius) the sequence (x_n) eventually stays inside B.

More formally we have the following.

Proposition 2.18 Let (X, d) be a metric space, let $(x_n) \subseteq X$ and let $x \in X$. Then (x_n) converges to x in X if and only if the following holds: for all $r > 0$ there exists an $N \in \mathbb{N}$ such that whenever $n \geq N$ we have $x_n \in B_X(x, r)$.

2.4 Continuous functions
2.4.1 Standard terminology and notation for functions

Definition 2.19 Let X and Y be sets, let f be a function from X to Y, let $E \subseteq X$ and let $F \subseteq Y$. We define the **image** of E under f, denoted by $f(E)$, by

$$f(E) = \{f(x) : x \in E\}$$

$$= \{y \in Y : \text{ there exists } x \in E \text{ with } f(x) = y\}.$$

We define the **pre-image** of F, denoted by $f^{-1}(F)$, by

$$f^{-1}(F) = \{x \in X : f(x) \in F\}.$$

Note that we are **not** assuming that f has an inverse function here: $f^{-1}(F)$ is purely notational, and does not mean that we have a function f^{-1}.

Example 2.20 Let X and Y be sets and let $f : X \to Y$. What, if anything, can we say about $f(\emptyset)$? $f(X)$? $f^{-1}(\emptyset)$? $f^{-1}(Y)$? Now suppose that f is a constant function. What are the possibilities for $f^{-1}(F)$ where F is a subset of Y?

Definition 2.21 Let X and Y be sets and let f be a function from X to Y. Then f is **surjective** (or **onto**) if $f(X) = Y$. In this case we say that f is a **surjection**.
The function f is **injective** (or **one-to-one**, or 1-1) if the following condition holds: whenever x_1 and x_2 are in X with $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$ (i.e. no two different points of X are mapped by f to the same point of Y). In this case we say that f is an **injection**.
The function f is **bijective** if it is both injective and surjective. In this case f is called a **bijection** (or a **one-one correspondence**)between X and Y.

There are many equivalent ways of restating these definitions: some of these will be discussed in lectures.

Example 2.22 Let X be a set. The **identity function** on X, denoted by Id_X, is defined by $\mathrm{Id}_X(x) = x$ for all $x \in X$. It is obvious that $\mathrm{Id}_X : X \to X$ is a bijection.

We now discuss the connection between bijectivity and invertibility.

Definition 2.23 Let X and Y be sets and let $f : X \to Y$. We say that a function g is an **inverse function** for f if $g : Y \to X$ and, for all $x \in X$ and $y \in Y$, we have

$$f(g(y)) = y \quad \text{and} \quad g(f(x)) = x.$$

In other words, $f \circ g = \mathrm{Id}_Y$ and $g \circ f = \mathrm{Id}_X$.

The following result is standard.

Proposition 2.24 Let X and Y be sets and let $f : X \to Y$. Then f has an inverse function if and only if f is a bijection. In this case the inverse function is unique and is denoted by f^{-1}; f^{-1} is a bijection from Y to X.

This time f^{-1} is a genuine function: if $F \subseteq Y$, the two possible meanings of $f^{-1}(F)$ agree.

For general functions (not bijections), pre-image behaves better than image when it comes to intersections and complements.

Proposition 2.25 Let X and Y be sets and let $f : X \to Y$.

(a) Let $F \subseteq Y$. Then

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

In terms of complements (in Y and X respectively) we have $f^{-1}(F^c) = (f^{-1}(F))^c$.
The corresponding statement for images of sets under f is, in general, false.

(b) Let A and B be subsets of Y. Then

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B),$$

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$$

and

$$f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B).$$

Of these, only the first equality is generally true for images of sets under f.

We may also work with arbitrary intersections and unions.

Proposition 2.26 Let X and Y be sets and let $f : X \to Y$. Let F_i be subsets of Y $(i \in I)$. Then

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

and

$$f^{-1}\left(\bigcap_{i \in I} F_i\right) = \bigcap_{i \in I} f^{-1}(F_i)$$

Only the first of these equalities is generally true for images of sets under f.

2.4.2 Continuity, uniform continuity and Lipschitz continuity

We now discuss continuous and discontinuous functions, as well as the stronger notions of **uniform continuity** and **Lipschitz continuity**.

Definition 2.27 Let (X, d_X) and (Y, d_Y) be metric spaces, let $f : X \to Y$ and let $x \in X$. Then f is **continuous at** x if, for all $\varepsilon > 0$, there is a $\delta > 0$ such that

$$f(B_X(x, \delta)) \subseteq B_Y(f(x), \varepsilon)\,. \quad (*)$$

If f is not continuous at x then f is **discontinuous** at x.
We say that f is **continuous from X to Y** if f is continuous at all the points of X. Otherwise the function f is **discontinuous** (from X to Y).

Remarks. Let X, Y and f be as above.

(a) To say that f is discontinuous means that there exists **at least one point of** X at which f is discontinuous.

(b) In terms of pre-image, condition $(*)$ is equivalent to saying $B_X(x, \delta) \subseteq f^{-1}(B_Y(f(x), \varepsilon))$.

(c) Condition $(*)$ says that, for all $x' \in X$ with $d_X(x', x) < \delta$, we have $d_Y(f(x'), f(x)) < \varepsilon$.

Here x is fixed, and you should think of δ as a function of ε. Once x varies, δ will usually depend on both x and ε.

(d) In full, the function f is continuous from X to Y if and only if, for all $x \in X$ and for all $\varepsilon > 0$, there exists $\delta > 0$ such that $f(B_X(x, \delta)) \subseteq B_Y(f(x), \varepsilon)$.

Example 2.28 Let X, Y, f and x be as above. As an exercise in negation, write down the full definition of the statement 'f is discontinuous at x' in what you consider to be the most useful form.

Example 2.29 (a) Let X and Y both be $\mathbb{R}$ with the usual metric. Then the new definition of continuity for functions from X to Y agrees with the version met in earlier modules. The same is true if X and Y are any of $\mathbb{C}$ or $\mathbb{R}^n$ with the usual metrics.

(b) Let (X, d) be a metric space. Then the identity function Id_X is continuous from X to X.

Note that if we consider two different metrics d_1 and d_2 on X, then it is not necessarily the case that Id_X is continuous from (X, d_1) to (X, d_2). We will consider this situation in detail later.

(c) Let X and Y be metric spaces, and suppose that $y_0 \in Y$. We can then define a constant function $f : X \to Y$ by $f(x) = y_0$ for all $x \in X$. We see easily that all such constant functions are continuous.

Examples (b) and (c) are also examples of functions with the stronger properties (which we are about to define) that they are uniformly continuous, and indeed Lipschitz continuous.

Definition 2.30 Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$. Then f is **uniformly continuous** (from X to Y) if, for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in X$, $f(B_X(x, \delta)) \subseteq B_Y(f(x), \varepsilon)$.
We say that f is **Lipschitz continuous** from X to Y if there is a constant $A > 0$ such that, for all x and x' in X, we have

$$d_Y(f(x'), f(x)) \leq A d_X(x', x).$$

Remarks

(a) What is the difference between continuity and uniform continuity? Only the fact that the choice of δ depends only on ε and not on x. Given $\varepsilon > 0$, the same δ is required to work simultaneously for all $x \in X$.

(b) With X, Y and f as above, f is uniformly continuous from X to Y if and only if for all $\varepsilon > 0$ there is $\delta > 0$ such that whenever x and x' are in X and satisfy $d_X(x, x') < \delta$ then $d_Y(f(x), f(x')) < \varepsilon$.

(c) Every uniformly continuous function is continuous, and every Lipschitz continuous function is uniformly continuous. (Neither of the converses is true.)

You may find it useful to look at the document **Uniform continuity for functions between subsets of the real line** available from the module web page.

(d) Where we say that f is Lipschitz continuous, many authors would say that f **satisfies a Lipschitz condition of order** 1. (See books for information about Lipschitz conditions of other orders.)

(e) In the definition of Lipschitz continuous, A is a constant and must not depend on x and x'.

Example 2.31 (a) Let (X, d_X) and (Y, d_Y) be metric spaces. Then every constant function from X to Y is Lipschitz continuous and so is the identity map Id_X from X to X.

(b) Let I be a non-degenerate interval in $\mathbb{R}$ and let f be a differentiable function from I to $\mathbb{R}$. Then (with respect to the usual metrics) f is Lipschitz continuous if and only if f' is bounded on I.

(c) Let (X, d) be a metric space and let $x_0 \in X$. Define $f : X \to \mathbb{R}^+$ by $f(x) = d(x, x_0)$ $(x \in X)$. Then (using the usual metric on $\mathbb{R}^+$) f is Lipschitz continuous (and we may take $A = 1$).

Theorem 2.33 Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$.

(a) (Uniform continuity in terms of sequences) The following statements are equivalent.

  (i) The function f is uniformly continuous from X to Y.

  (ii) Whenever (x_n) and (x'_n) are sequences of elements of X such that $\lim_{n\to\infty} d_X(x_n, x'_n) = 0$ then we also have $\lim_{n\to\infty} d_Y(f(x_n), f(x'_n)) = 0$.

(b) (Failure of uniform continuity in terms of sequences) The following three statements are equivalent.

  (i) The function f is not uniformly continuous from X to Y.

  (ii) There is at least one pair of sequences (x_n) and (x'_n) of elements of X such that $\lim_{n\to\infty} d_X(x_n, x'_n) = 0$ but such that the sequence $(d_Y(f(x_n), f(x'_n)))$ does not converge to 0.

  (iii) There exist a positive number $\varepsilon > 0$ and a pair of sequences (x_n) and (x'_n) of elements of X such that $\lim_{n\to\infty} d_X(x_n, x'_n) = 0$ but, for all $n \in \mathbb{N}$, we have $d_Y(f(x_n), f(x'_n)) \geq \varepsilon$.

The following theorem gives the connection between our definition of continuity and continuity in terms of sequences. This is essentially the same as in earlier modules. (The details of the proof are on Question Sheet 2.)

Theorem 2.32 Let X and Y be metric spaces and let $f : X \to Y$.

(a) Let $x \in X$. Then the following statements are equivalent.

  (i) The function f is continuous at x.

  (ii) For every sequence $(x_n) \subseteq X$ which converges to x, we have $f(x_n) \to f(x)$ in Y as $n \to \infty$.

(b) The following two statements are equivalent.

  (i) The function f is continuous from X to Y.

  (ii) For every sequence $(x_n) \subseteq X$ which converges in X, the sequence $(f(x_n)) \subseteq Y$ is convergent in Y, and

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n).$$

Remark: In these results, the crucial difference between uniform continuity and continuity is that **it is possible to have**

$$\lim_{n\to\infty} d_X(x_n, x'_n) = 0$$

without either of the two sequences (x_n)**,** (x'_n) **converging.** A typical example of this phenomenon is when $X = \mathbb{R}$ and $x_n = n$, $x'_n = n + \frac{1}{n}$.

Note also that in (b)(ii), to say that the sequence $(d_Y(f(x_n), f(x'_n)))$ does not converge to 0 does not imply that it converges to some positive real number (it might not converge at all).

We shall see later that if the metric space X is (sequentially) compact, then continuity and uniform continuity coincide.

2.4.3 New continuous functions from old

Definition 2.34 Let X be a set, let f and g be real-valued functions on X and let $\alpha \in \mathbb{R}$. We define the functions αf, $f+g$, fg and (provided $g(x)$ does not take the value 0) f/g using the usual pointwise operations: for $x \in X$,

$$(\alpha f)(x) = \alpha \cdot f(x), \quad (f+g)(x) = f(x) + g(x),$$

$$(fg)(x) = f(x)g(x)$$

and, provided that $g(x) \neq 0$,

$$(f/g)(x) = f(x)/g(x).$$

With these operations, the set of all real-valued functions on X becomes a vector space over $\mathbb{R}$ and, indeed, a ring in which the invertible elements are those functions $f : X \to \mathbb{R}$ for which 0 is not an element of $f(X)$.

Notation. Let X be a metric space. Then we denote by $C(X)$ the set of all continuous functions from X to $\mathbb{R}$. (Note that we met $C([0,1])$ earlier, but we abbreviated it to $C[0,1]$.)

The content of the next proposition is, essentially, that for a metric space X, $C(X)$ is a vector subspace and also a subring of the vector space/ring of all real-valued functions on X.

Theorem 2.35 Let X be a metric space, let f and g be functions from X to $\mathbb{R}$ and let $\alpha \in \mathbb{R}$.
(a) Let $x \in X$. If f and g are both continuous at x, then so are αf, $f+g$, fg and (if $0 \notin g(X)$) f/g.
(b) If f and g are in $C(X)$ then so are αf, $f+g$, fg and (if $0 \notin g(X)$) f/g.

We conclude this subsection by noting that a composite of two continuous functions is again a continuous function.

Theorem 2.36 Let X, Y and Z be metric spaces, let $f : X \to Y$ and $g : Y \to Z$. Consider the composite function $g \circ f : X \to Z$ (recall that this is defined by $x \mapsto g(f(x))$).
(a) Let $x \in X$. If f is continuous at x and g is continuous at $f(x)$ then $g \circ f$ is continuous at x.
(b) If f is continuous from X to Y and g is continuous from Y to Z then $g \circ f$ is continuous from X to Z.

2.5 New metric spaces from old

Definition 2.37 Let (X, d) be a metric space and let $Y \subseteq X$. Then the **subspace metric on Y induced by the metric** d is the metric $\tilde{d}$ on Y obtained by restricting d to $Y \times Y$. Thus we have $\tilde{d} : Y \times Y \to \mathbb{R}^+$ and, for all y_1 and y_2 in Y,

$$\tilde{d}(y_1, y_2) = d(y_1, y_2).$$

Example 2.38 (a) As an easy exercise, you should check the following, with notation as above.

(i) The function $\tilde{d}$ really is a metric on Y.

(ii) Let $y \in Y$ and let $(y_n) \subseteq Y$. Then (y_n) converges to y in X with respect to d if and only if (y_n) converges to y in Y with respect to $\tilde{d}$.

In general, it is possible for a sequence in Y to converge in X (with respect to d) to a point in $X \setminus Y$. Such a sequence will be divergent with respect to $\tilde{d}$.

(b) Let X be any of the sets $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{C}$ and let d be the usual metric on X. Then, for $Y \subseteq X$, the subspace metric $\widetilde{d}$ on Y induced by d is also the usual metric on Y. For example, the usual metric on $[0,1]$ is also the subspace metric on $[0,1]$ induced by the usual metric on $\mathbb{R}$.

Example 2.39 Let X be a set, let d be a metric on X, let d' be a pseudometric on X and let α be a positive real number. Then the following functions of x and y are metrics on X:

(a) $d(x,y) + d'(x,y)$;

(b) $\alpha d(x,y)$;

(c) $\max\{d(x,y), d'(x,y)\}$;

(d) (i) $\min\{d(x,y), \alpha\}$, (ii) $\dfrac{d(x,y)}{1+d(x,y)}$ and (iii) $\sqrt{d(x,y)}$.

However $\min\{d(x,y), d'(x,y)\}$ often fails to be a metric on X.

The examples in (d) above can all be justified using the next result.

Theorem 2.40 Let X be a set and let d be a pseudometric on X. Let $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-decreasing function with $\phi(0) = 0$ and such that, for all a and b in $\mathbb{R}^+$,

$$\phi(a+b) \leq \phi(a) + \phi(b). \quad (*)$$

Then $\phi \circ d$ is also a pseudometric on X.

Remarks Given that $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ satisfies $\phi(0) = 0$, a sufficient condition for $(*)$ is that ϕ be **concave downwards**.

This means, essentially, that the slope of the curve decreases as you move from left to right.

If ϕ is continuous on $\mathbb{R}^+$ and twice-differentiable on $(0,\infty)$, then ϕ is concave downwards if and only if, for all $x \in (0,\infty)$, $\phi''(x) \leq 0$.

2.6 Open sets in metric spaces

All the definitions and results of this section are also valid for pseudometric spaces.

We will only state them for metric spaces here.

Note that 'openness' will depend on the metric space you are working in, i.e. both the set X which you regard as the whole space, and the metric d you use on X.

Definition 2.41 Let (X,d) be a metric space and let $U \subseteq X$. Then U is **open in** X (or **open with respect to** d or d-**open** to avoid ambiguity) if there is a function $r : U \to (0,\infty)$, such that, for all $x \in U$,

$$B_X(x, r(x)) \subseteq U\,.$$

In this setting we also say that U is an **open subset of** X, or simply U **is open** if there is no ambiguity over the metric space in question.

Remark Most authors do not explicitly state that r is a function of x.

With the above notation, the most common definition of 'U is open in X' is the following:

for all $x \in U$ there exists $r > 0$ such that $B_X(x,r) \subseteq U$.

Note that r depends on x here, so this is equivalent to our definition: you may use whichever you prefer!

Roughly speaking, you can take $r(x)$ to be 'the distance from x to the boundary of U, **or anything smaller**'.

We will make the concepts of boundary and distance to a set more formal later.

Example 2.42 (a) Let X be a metric space. Then X and $\emptyset$ are always open in X. For example, if $X = [0,1]$ with the usual metric, then X is open in X, even though X is not open in $\mathbb{R}$.

(b) Let X be a metric space. Then every open ball in X is an open subset of X, but, usually, **not every open subset of X is an open ball**.

(c) Let X be any of $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{C}$ and give X its usual metric: call this metric d. Then the d-open subsets of X are the same as the open sets met in earlier modules.

The next two theorems show that open sets behave well with respect to finite intersections and arbitrary unions.

Theorem 2.43 Let X be a metric space and let U and V be open subsets of X. Then $U \cap V$ is also open in X. More generally, **finite** intersections of open subsets of X are again open subsets of X.

This last result becomes false if you allow infinite intersections (see question sheets).

Theorem 2.44 Let X be a metric space and let U_i $(i \in I)$ be open subsets of X. Then $\bigcup_{i\in I} U_i$ is also open in X.

The following theorem summarizes what we know about open sets in metric spaces.

Theorem 2.45 Let X be a metric space.

(1) Both $\emptyset$ and X are open in X.

(2) Every finite intersection of open subsets of X is again an open subset of X.

(3) Arbitrary unions of open subsets of X are again open subsets of X.

Example 2.46 Let X be a set and give X the discrete metric d: recall that this is defined, for x and y in X, by

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

(a) What are the possible open balls in X?

(b) Which subsets of X are open?

What if, instead, d is the **indiscrete pseudometric** defined by $d(x,y) = 0$ for all x and y in X?

3 Introduction to topological Spaces
3.1 Topologies and topological spaces

With topological spaces, we do not need a metric: **we just decide which sets we want to be open!** We have a lot of freedom in our choice, but we are guided by the definition of **topology** below.

First we recall some notation and definitions concerning sets of sets.

Definition 3.1 Let X be a set. Then the **power set** of X is the set of all subsets of X: this is denoted by $\mathcal{P}(X)$ or 2^X. So if $\mathcal{C}$ is a set of subsets of X, then the elements of $\mathcal{C}$ are subsets of X, and $\mathcal{C} \subseteq \mathcal{P}(X)$. In this case, we often call $\mathcal{C}$ a **collection** of subsets of X.

Example 3.2 (a) How many subsets does $\emptyset$ have? Is there a difference between $\{\emptyset\}$ and $\emptyset$?

(b) How many subsets does $\{1, 2, 3\}$ have? More generally, if X has exactly n elements, how many subsets does X have? (This is the same as the number of **elements** of $\mathcal{P}(X)$.)

(c) (Optional) Show that $\mathcal{P}(\mathbb{N})$ is uncountable.

Definition 3.3 Let X be a set and let τ be a collection of subsets of X. Then τ is a **topology on** X if it satisfies the following three conditions (topological space axioms):

(1) $\emptyset$ and X are in τ;

(2) if U and V are in τ then $U \cap V \in \tau$;

(3) if U_i $(i \in I)$ is an indexed collection of elements of τ then $\bigcup_{i \in I} U_i$ is also in τ.

Given that τ is a topology on X, we say that (X, τ) is a **topological space**, and that the sets U in τ are **open in** X, or, in case of ambiguity, **open with respect to** τ or τ**-open**. For $U \in \tau$, we also say that U is an **open subset of** X, or simply U **is open** if the topological space under discussion is unambiguous.

Remark Conditions (2) and (3) tell us that, for topological spaces, finite intersections of open sets are open and arbitrary unions of open sets are open.

This is precisely what we proved earlier for open sets in metric spaces.

For topological spaces, it is part of the definition!

You can also say that τ is **closed under the operations of finite intersection and arbitrary union**, as long as you do not confuse this use of the word 'closed' with the closed sets which we are about to define.

Example 3.4 (a) **Every metric space is a topological space.**

Let (X, d) be a metric space.

We have defined what it means for a subset of X to be open with respect to d, or d-open. Set

$$\tau = \{U \subseteq X : U \text{ is } d\text{-open}\}.$$

Then τ is a topology on X, and is called the **topology on** X **induced by** d.

Whenever we have a metric space, we will assume by default that it has the topology induced by its metric.

(b) The definition of open set used for metrics applies equally well to a pseudometric d on a set X.

The collection of d-open subsets of X is again a topology on X: the topology **induced by the pseudometric** d.

(c) Given a set X, the **strongest** (or **finest**) topology possible on X is the topology $\tau = \mathcal{P}(X)$, which has the most possible open sets. This topology is called the **discrete topology** on X, and is induced by the discrete metric on X.

(d) The opposite of the discrete topology on X is the **indiscrete topology** on X: the topology with the fewest possible open sets, described as the **weakest** or **coarsest** possible topology on X. What is this topology? Can it be induced by a metric? Or a pseudometric?

(e) (This is also on question sheet 3). Find all the possible topologies on the set $X = \{1, 2\}$, and show that some of these topologies can not be induced by a pseudometric.

(f) (For the enthusiastic) Investigate the properties of the possible topologies on $\{1,2,3\}$.

From above we see that not all topologies can be induced by metrics, or even pseudometrics.

Definition 3.5 Let τ be a topology on a set X. Then τ is **pseudometrizable** if there exists a pseudometric d on X which induces the topology τ. If there is a metric d which induces τ then τ is **metrizable**.

Most commonly met topologies are metrizable, but we have mentioned above some standard examples which are not. Even if a topology is not metrizable, it may still satisfy the Hausdorff separation condition defined below.

Definition 3.6 Let (X, τ) be a topological space. Then we say that X is a **Hausdorff space**, or that X **is Hausdorff**, or that τ **is a Hausdorff topology**, if the following condition holds (the Hausdorff separation condition): for all x and x' in X with $x \neq x'$, there are open sets U and U' such that $x \in U$, $x' \in U'$ and $U \cap U' = \emptyset$.

Remarks

(a) The sets U and U' used will each depend on **both** x and x'.

(b) Every metric space is a Hausdorff topological space (see Question Sheet 3). A pseudometric d induces a Hausdorff topology if and only if the pseudometric d is actually a metric.

We now define closed sets in topological spaces (in terms of open sets). This definition applies also to metric spaces, and coincides with all previous definitions of closed sets from earlier modules.

Definition 3.7 Let (X, τ) be a topological space, and let $E \subseteq X$. If $X \setminus E$ is open in X then we say that E is **closed in** X, or, in case of ambiguity, **closed with respect to** τ or τ-**closed**. In this setting we also say that E is a **closed subset of** X, or simply E **is closed** if the topological space under discussion is unambiguous.

Remarks: Let (X, τ) be a topological space.

(a) Note! 'closed' does NOT mean 'not open'. Every combination of properties is possible: sets may be open, closed, or neither, or both. Sets which are both open and closed will be called **clopen** in this module.

(b) Let U and E be subsets of X. Working with complements in X, we have that U is open in X if and only if U^c is closed in X and E is closed in X if and only if E^c is open in X.

(c) The term 'half-open' is generally misleading, and should only be applied to very special kinds of sets such as intervals.

Example 3.8 (1) Let (X, τ) be any topological space. Since $\emptyset$ and X must be open in X, taking complements we see that they are also both closed in X. Thus $\emptyset$ and X are always clopen in X. (If X is **connected** then these are the only clopen subsets of X: see later.)

(2) Consider $X = \mathbb{R}$ with the usual topology. Set $A = \mathbb{Q}$, $B = [0, 1)$ and $C = [0, \infty)$. Then A is neither open nor closed in $\mathbb{R}$, and the same is true of B. It makes some sense to call B half-open, but this **makes no sense** for A. What can you say about C?

(3) Since the complement of every open set is closed, we may take complements in all our earlier examples of open sets to give examples of closed sets. In particular, for a metric space (X, d), given $x \in X$ and $r > 0$, the set $X \setminus B_X(x, r) = \{y \in X : d(y, x) \geq r\}$ is closed in X. The **closed ball** $\{y \in X : d(y, x) \leq r\}$ is also a closed subset of X (exercise).

(4) Let (X, τ) be a topological space, and let $x \in X$. Is it necessarily true that the single-point set $\{x\}$ is closed in X? What if X is a metric space? Or, more generally, a Hausdorff space? (See Question Sheet 3.)

Using the properties we know of open sets, an easy application of de Morgan's laws (again, see Question Sheet 3) proves the following for closed sets.

Theorem 3.9 Let (X, τ) be a topological space. Then finite unions of closed sets are again closed sets, and arbitrary intersections of closed sets are again closed sets.

There are many equivalent definitions of open sets and closed sets, especially for metric spaces. We shall meet some of these shortly.

3.2 Interior, closure and related concepts

3.2.1 Definitions for topological spaces

We now discuss the concepts interior, closure, boundary and denseness for subsets of topological spaces.

Informally speaking, if Y is a subset of a topological space X, then the interior of Y is the largest possible open subset of X which is contained in Y and the closure of Y is the smallest possible closed subset of X which contains Y.

The set difference between the closure and the interior is the boundary of Y.

Note that each subset Y of X does always have at least one subset which is open in X (namely $\emptyset$) and at least one superset which is closed in X (namely X).

Definition 3.10 Let (X, τ) be a topological space and let $Y \subseteq X$. Then the **interior** of Y (with respect to τ), denoted by $\operatorname{int} Y$, is the union of all the open subsets of X which are contained in Y. The **closure of** Y (with respect to τ), denoted by $\overline{Y}$ or $\operatorname{clos} Y$, is the intersection of all the closed subsets of X which contain Y. The **boundary** (or **frontier**) of Y (with respect to τ), denoted by ∂Y, is the difference between the closure and the interior of Y: $\partial Y = \overline{Y} \setminus \operatorname{int} Y$.

Since unions of open sets are still open, and intersections of closed sets are still closed, the following properties of interior and closure are almost immediate from their definitions.

We usually use these properties rather than the original definitions.

Proposition 3.11 Let X be a topological space and let $Y \subseteq X$.

(a) (i) For $x \in X$, we have $x \in \operatorname{int} Y$ if and only if there exists an open subset U of X with $x \in U \subseteq Y$.

(ii) The set $\operatorname{int} Y$ is open in X and $\operatorname{int} Y \subseteq Y$. Moreover, whenever U is an open subset of X with $U \subseteq Y$, then $U \subseteq \operatorname{int} Y$.

(iii) The set Y is open in X if and only if $Y = \operatorname{int} Y$.

(b) (i) The set $\overline{Y}$ is closed in X and $Y \subseteq \overline{Y}$. Moreover, whenever E is a closed subset of X with $Y \subseteq E$ then $\overline{Y} \subseteq E$.

(ii) The set Y is closed in X if and only if $Y = \operatorname{clos} Y$.

Example 3.12 (a) Let X be $\mathbb{R}$ with the usual topology. With respect to this topology, determine $\operatorname{int} A$, $\overline{A}$ and ∂A when A is: (i) $[0, 1)$ (ii) $\mathbb{Q}$.

(b) Let D be a disc in $\mathbb{R}^2$ or $\mathbb{C}$ (usual topology). Then the interior, closure and boundary of D are exactly what you expect them to be.

The following result relates interior to closure, via complements, and gives an alternative way to think about the boundary.

Theorem 3.13 Let X be a topological space and let $Y \subseteq X$. Then, working with complements in X,

$$\operatorname{clos}(Y^c) = (\operatorname{int} Y)^c, \quad \operatorname{int}(Y^c) = (\operatorname{clos} Y)^c$$

and

$$\partial Y = \operatorname{clos} Y \cap \operatorname{clos}(Y^c).$$

In the following, note that some authors insist that neighbourhoods must actually be open sets.

We do not require this.

Definition 3.14 Let X be a topological space and let $x \in X$. Then a **neighbourhood** of x is any subset N of X such that $x \in \operatorname{int} N$.

It is easy to see that a subset U of X is open if and only if U is a neighbourhood of every $x \in U$.

We conclude this subsection with several equivalent definitions of denseness for subsets of topological spaces.

We are already familiar with the idea that both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are 'dense' in $\mathbb{R}$.

What does this correspond to in the topological setting?

Definition 3.15 Let X be a topological space and let $Y \subseteq X$. Then Y is **dense** in X if, for every **non-empty** open subset U of X we have $Y \cap U \neq \emptyset$.

Remark We need to insist that U is non-empty, otherwise the condition can not be satisfied!

Another way to state the condition is to say that Y **meets every non-empty open subset of** X**.**

Example 3.16 (a) You may now easily check that, with the usual topology on $\mathbb{R}$, both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense in $\mathbb{R}$.

(b) As an exercise, you can prove that (with X and Y as above) the following statements are equivalent:

(i) Y is dense in X;

(ii) for every $x \in X$ and every neighbourhood N of x, we have $N \cap Y \neq \emptyset$;

(iii) $\overline{Y} = X$;

(iv) $\operatorname{int}(X \setminus Y) = \emptyset$.

3.2.2 Equivalent definitions for metric spaces

We now look again at the concepts introduced above from the point of view of metric spaces. First we give the metric space versions of interior, neighbourhood, closure and boundary.

Theorem 3.17 Let (X, d) be a metric space, let $x \in X$ and let $Y \subseteq X$.

(a) The following statements are equivalent:
  - (i) $x \in \operatorname{int} Y$;
  - (ii) Y is a neighbourhood of x;
  - (iii) there exists an $r > 0$ such that $B_X(x, r) \subseteq Y$.

(b) The following statements are equivalent:
  - (i) $x \in \overline{Y}$;
  - (ii) for all $r > 0$, $B_X(x, r) \cap Y \neq \emptyset$;
  - (iii) there exists a sequence $(y_n) \subseteq Y$ which converges to x with respect to d.

(c) The following statements are equivalent:
  - (i) $x \in \partial Y$;
  - (ii) for all $r > 0$, we have both $B_X(x, r) \cap Y \neq \emptyset$ and $B_X(x, r) \setminus Y \neq \emptyset$.

Corollary 3.18 Let (X, d) be a metric space and let $Y \subseteq X$. Then $\overline{Y}$ is equal to the set of points x in X which are limits (with respect to d) of sequences of elements of Y. In particular, Y is closed if and only if the following condition holds: whenever a sequence $(y_n) \subseteq Y$ has a limit x in X, we have $x \in Y$.

Corollary 3.19 Let (X, d) be a metric space and let $Y \subseteq X$. Then the following statements are equivalent:

(i) Y is dense in X;

(ii) for all $x \in X$ and all $r > 0$, $Y \cap B_X(x, r) \neq \emptyset$;

(iii) for all $x \in X$ there exists a sequence $(y_n) \subseteq Y$ which converges to x with respect to d.

You should check this corollary for the dense subsets $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ of $\mathbb{R}$.

3.3 Continuous functions

We have already discussed continuous functions between metric spaces.

Let us recall here some of the equivalent definitions available.

Theorem 3.20 Let X and Y be metric spaces and let $f : X \to Y$. Then the following statements are equivalent, and they all mean that f is continuous from X to Y.

(a) (Our standard definition of continuity for metric spaces) For all x in X and all $\varepsilon > 0$ there exists a $\delta > 0$ such that $f(B_X(x, \delta)) \subseteq B_Y(f(x), \varepsilon)$.

(b) For all x in X and all $\varepsilon > 0$ there exists a $\delta > 0$ such that $B_X(x, \delta) \subseteq f^{-1}(B_Y(f(x), \varepsilon))$.

(c) For every convergent sequence $(x_n) \subseteq X$, the sequence $(f(x_n))$ is convergent in Y and

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

How can we extend the notion of continuity to the purely topological setting?

We begin by reformulating continuity for metric spaces in terms of the topologies on the metric spaces involved.

Theorem 3.21 Let X and Y be metric spaces, and let $f : X \to Y$. Then the following statements are equivalent.
(a) The function f is continuous from X to Y.
(b) For every open subset U of Y, $f^{-1}(U)$ is open in X.
(c) For every closed subset E of Y, $f^{-1}(E)$ is closed in X.

We are now able to extend the definition of continuity to the topological space setting.

Definition 3.22 Let X and Y be topological spaces, and let $f : X \to Y$. Then f is **continuous** from X to Y if, for every open subset U of Y, $f^{-1}(U)$ is open in X.

Example 3.23 (a) By the theorem above, all earlier examples of continuous functions between metric spaces are also continuous in the topological space sense.

Also, constant functions between topological spaces and the identity map from a topological space to itself are easily seen to be continuous.

(b) Let X and Y be topological spaces and let $f : X \to Y$. Then (exercise) the following are equivalent.

(i) The function f is continuous from X to Y.

(ii) For every closed subset E of Y, $f^{-1}(E)$ is closed in X.

(iii) (Slightly harder) For every subset A of X, we have

$$f(\text{clos}\, A) \subseteq \text{clos}\,(f(A))\,.$$

(c) **Warning!** Suppose that f is continuous from X to Y, that V is an open subset of X and F is a closed subset of X. Then there is no reason to believe that $f(V)$ is an open subset of Y. Nor need $f(F)$ be closed in Y.

Easy exercise: find appropriate examples to verify these claims.

As before we may form new continuous functions from old. In particular, we can compose continuous functions.

Theorem 3.24 Let X, Y and Z be topological spaces. Let $f : X \to Y$ and $g : Y \to Z$ be continuous functions. Then the composite function $g \circ f$ is continuous from X to Z.

As usual, a function which is not continuous is said to be **discontinuous**.

We have not yet discussed any notion of continuity at a point for maps between topological spaces: this will not be an important notion in this module, but can be useful in more advanced study of analysis.

In the following definition, remember that we do **not** require neighbourhoods to be open.

Definition 3.25 Let X and Y be topological spaces, let $f : X \to Y$ and let $x \in X$. Then f is **continuous at** x if, for every neighbourhood N of $f(x)$ in Y, $f^{-1}(N)$ is a neighbourhood of x in X.

Exercise. With X, Y and f as above, prove that f is continuous from X to Y if and only if f is continuous at all the points of X.

3.4 Homeomorphisms and isometries

In this section we consider when two topological spaces may be regarded as essentially the same as each other (such topological spaces are **homeomorphic** to each other), and when two metric spaces may be regarded as essentially the same as each other (such metric spaces are **isometric** to each other).

Definition 3.26 Let X and Y be topological spaces. A **homeomorphism** from X to Y is a bijection $f : X \to Y$ such that both f and f^{-1} are continuous. If such a homeomorphism exists from X to Y, we say that X is **homeomorphic** to Y.

Note! The word **homeomorphism** should NOT be confused with the word **homomorphism**.

Example 3.27 (a) Using the continuous map $f(x) = \exp(x)$, we see that (with the usual topologies) $\mathbb{R}$ is homeomorphic to $(0, \infty)$. Note here that the inverse map is $f^{-1}(y) = \ln y$, which is continuous from $(0, \infty)$ to $\mathbb{R}$, as required.

(b) It is not enough to assume that f is a continuous bijection.

For example, with the usual topologies, there is an obvious continuous bijection from $X = [0, 2\pi)$ to the unit circle in $\mathbb{C}$, $Y = \{z \in \mathbb{C} : |z| = 1\}$.

(What is this bijection?)

However this bijection has a discontinuous inverse.

(In fact X is not homeomorphic to Y. This does not follow from what we have said above, but can be deduced easily from the fact that Y is compact and X is not: see later.)

(c) As an exercise you should check the following facts about homeomorphisms which should have a familiar flavour. Let X, Y and Z be topological spaces. Then

(i) X is homeomorphic to X (use the identity mapping Id_X);

(ii) if X is homeomorphic to Y then Y is homeomorphic to X (show that the inverse of a homeomorphism is a homeomorphism);

(iii) if X is homeomorphic to Y and Y is homeomorphic to Z then X is homeomorphic to Z (show that the composite of two homeomorphisms is a homeomorphism).

In view of (ii), rather than saying that X is homeomorphic to Y we may say that X **and** Y **are homeomorphic topological spaces**.

From the definition of continuity for topological spaces, we see that a homeomorphism induces a one-one correspondence between the collections of open sets, as follows.

Theorem 3.28 Let (X, τ_1) and (Y, τ_2) be homeomorphic topological spaces, and let f be a homeomorphism from X to Y. Then

$$\tau_2 = \{f(U) : U \in \tau_1\}$$

and

$$\tau_1 = \{f^{-1}(V) : V \in \tau_2\}.$$

From the point of view of metric spaces, homeomorphism is often not a strong enough condition, and we need the notion of **isometry**.

Definition 3.29 Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$. Then f is an **isometry** if, for all x and x' in X, we have

$$d_Y(f(x), f(x')) = d_X(x, x').$$

If there exists a **surjective** isometry from X to Y then we say that X **is isometric to** Y or X **and** Y **are isometric metric spaces**.

Remarks

(a) Note the subtle point that isometries need not be surjective, but that we require the existence of a surjective isometry if the metric spaces are to be isometric to each other.

(b) Isometries are Lipschitz continuous.

(c) Isometries are automatically injective. Thus surjective isometries are automatically bijections.

(d) The inverse of a surjective isometry is also a surjective isometry. In particular, every surjective isometry is also a homeomorphism.

Example 3.30 (a) Let X be $\mathbb{R}$ and let Y be $\mathbb{C}$, both with the usual metrics. Define $i : \mathbb{R} \to \mathbb{C}$ by $i(x) = x$ $(x \in \mathbb{R})$. Then i is an isometry, but it is not surjective. This does not prove that $\mathbb{R}$ is not isometric to $\mathbb{C}$, as you would need to prove that there does not exist a surjective isometry from $\mathbb{R}$ to $\mathbb{C}$. This is, in fact, true (exercise).

(b) Every metric space (X, d) is isometric to itself via the identity map Id_X.

(c) If we use two different metrics d_1 and d_2 on the same set X, the situation becomes more complex: the identity map Id_X need not be a homeomorphism from (X, d_1) to (X, d_2), and even if it is, it is only an isometry if $d_1 = d_2$.

Let X_1 be $\mathbb{R}^2$ with the metric d_1, and let X_2 be $\mathbb{R}^2$ with the metric d_∞. Then the identity map $\mathrm{Id}_{\mathbb{R}^2}$ is a homeomorphism from X_1 to X_2 but it is not an isometry.

However this still leaves open the possibility that X_1 might be isometric to X_2. What do you think?

(d) As a tricky exercise, try to prove that $(\mathbb{R}^2, d_\infty)$ is not isometric to $(\mathbb{R}^2, d_2)$.

In this module we will discuss many possible properties of spaces.

A **topological property** is one which is preserved by homeomorphisms: if X is homeomorphic to Y then any topological property that Y has is shared by X.

Such properties are often expressed in terms of open sets or closed sets: for example, Hausdorffness is a typical topological property.

Other properties have more of a metric space flavour: if a metric space X has such a property, so will every metric space which is isometric to X.

A typical example of such a property is **completeness** (see Chapter 7).

Note that, with the usual metrics, $\mathbb{R}$ is homeomorphic to $(0, \infty)$, but $\mathbb{R}$ is complete while $(0, \infty)$ is not.

3.5 Equivalence of metrics

In this section we look in more detail at the situation where we have different metrics on the same set.

Definition 3.31 Let X be a set and let d and d' be metrics on X. Then the two metrics d and d' are **equivalent** (or d **is equivalent to** d') if the following condition holds: for all $x \in X$ and all sequences $(x_n) \subseteq X$, $d(x_n, x) \to 0$ as $n \to \infty$ if and only if $d'(x_n, x) \to 0$ as $n \to \infty$.
The two metrics are **uniformly equivalent** to each other (or d **is uniformly equivalent to** d') if there are positive constants K_1 and K_2 such that, for all x and y in X,

$$d(x, y) \leq K_1 d'(x, y) \quad \text{and} \quad d'(x, y) \leq K_2 d(x, y)\,.$$

Remarks Let X, d and d' be as above.

(a) The metrics d and d' are equivalent if they give rise to the same convergent sequences and the same limits as each other.

(b) As an exercise you can prove that the following are equivalent:

- (i) d and d' are equivalent metrics;
- (ii) Id_X is a homeomorphism from (X, d) to (X, d');
- (iii) d and d' induce the same topology on X (i.e., a subset U of X is d-open if and only if it is d'-open).

(c) In the definition of uniform equivalence:

- (i) the constants K_1 and K_2 must not depend on x and y;
- (ii) by replacing K_1 and K_2 by $K = \max\{K_1, K_2\}$, it is easy to see that we may assume that $K_1 = K_2$ if we wish, without changing which metrics are or are not uniformly equivalent to each other.

(d) If d and d' are uniformly equivalent then they are equivalent (the converse is false). In fact, d and d' are uniformly equivalent if and only if both the map Id_X from (X, d) to (X, d') and its inverse, Id_X, from (X, d') to (X, d), are Lipschitz continuous,

Example 3.32 (1) Fix $n \in \mathbb{N}$ (so that in what follows, n is a constant). On $\mathbb{R}^n$, consider the metrics d_1, d_2 and d_∞. Then these three metrics are uniformly equivalent to each other. Therefore they are also equivalent to each other, and they give the same collection of open sets. They all induce what we call the **usual topology** on $\mathbb{R}^n$. Note that this is true even though their open balls are different.

(2) On $C[0,1]$, the two metrics d_1 and d_∞ are not equivalent to each other (hence they are not uniformly equivalent either).

(3) The usual metric on $\mathbb{N}$ (as a subset of $\mathbb{R}$) is equivalent to the discrete metric on $\mathbb{N}$, but these two metrics are not uniformly equivalent. (Why not?)

The relationship between equivalence and uniform equivalence is similar (but not identical) to that between homeomorphisms and surjective isometries.

The notion of equivalence is sufficient when discussing topological properties of metric spaces, while for properties of a more metric nature the notion of uniform equivalence is often more appropriate.

3.6 Normality and Urysohn's Lemma for metric spaces

In this section we discuss some separation properties for topological spaces, and in particular the existence of a good supply of continuous, real-valued functions on metric spaces.

Recall that a topological space X is **Hausdorff** if, whenever x and x' are in X with $x \neq x'$, then there are disjoint open sets U and V with $x \in U$ and $y \in V$.

To say that U and V are **disjoint** just means that $U \cap V = \emptyset$.

In the next definition, note that some authors insist that regular/normal topological spaces must be Hausdorff.

We do not require this.

Definition 3.33 Let X be a topological space. Then X is **regular** if, for every closed subset E of X and every point $x \in X \setminus E$ there exist disjoint open sets U and V with $x \in U$ and $E \subseteq V$. The topological space X is **normal** if for each pair of disjoint closed subsets E and F of X, there is a pair of disjoint open subsets of X, U and V with $E \subseteq U$ and $F \subseteq V$.

Note that, in general, single-point sets need not be closed.

For those topological spaces where single-point sets **are** closed, normality implies regularity and regularity implies Hausdorffness.

In particular, every normal Hausdorff space is regular.

From the question sheets, we know that all metric spaces are Hausdorff.

It is possible to prove directly that metric spaces are normal (see question sheet).

However we will prove this indirectly here via Urysohn's Lemma for metric spaces.

First we need to investigate the properties of the function given by taking the distance from a set to a point.

Definition 3.34 Let (X, d) be a metric space and let E be a non-empty subset of X. Then, for $x \in X$, we define the **distance from E to x** (or the **distance from x to E**), denoted by $\text{dist}(x, E)$ or $\text{dist}(E, x)$, by

$$\text{dist}(x, E) = \text{dist}(E, x) = \inf\{d(x, y) : y \in E\}.$$

Remarks: Note that this definition would not make sense for $E = \emptyset$. But provided that $E \neq \emptyset$, for each $x \in X$ the set $\{d(x, y) : y \in E\}$ is non-empty and bounded below by 0, so $\text{dist}(x, E) \in [0, \infty)$.

The inf need not be achieved, i.e. it may not be a minimum.

For example, with $X = \mathbb{R}$ with the usual metric $d(x, y) = |x - y|$, let $E = (0, 1)$ and $x = 3$.

We have $\text{dist}(x, E) = 2$, but there is no point y of E with $d(x, y) = 2$.

The function $x \mapsto \text{dist}(x, E)$ will be a very useful function for us. We begin by establishing that it is Lipschitz continuous.

Proposition 3.35 Let (X, d) be a metric space and suppose that E is a non-empty subset of X. Set $f(x) = \text{dist}(x, E)$ $(x \in X)$. Then the function $f : X \to \mathbb{R}$ is Lipschitz continuous, and we may take $A = 1$ as our Lipschitz constant.

Combining this with our earlier results gives us another characterization of closure in metric spaces.

Corollary 3.36 Let (X, d) be a metric space, and let E be a non-empty subset of X.

(a) In terms of distance from E, we have

$$\overline{E} = \{x \in X : \text{dist}(x, E) = 0\}.$$

(b) The following statements are equivalent:

(i) E is closed in X;

(ii) $E = \{x \in X : \text{dist}(x, E) = 0\}$;

(iii) whenever $x \in X$ satisfies $\text{dist}(x, E) = 0$ then $x \in E$;

(iv) for all $x \in X \setminus E$ we have $\text{dist}(x, E) > 0$.

Example 3.37 (a) Given two disjoint, closed intervals $[a, b]$ and $[c, d]$ in $\mathbb{R}$, it is not hard to show that there is a continuous function $f : \mathbb{R} \to \mathbb{R}$ which is constantly 0 on $[a, b]$ and constantly 1 on $[c, d]$.

Assuming that $a < b < c < d$, can you write down the definition of such a function?

(b) Given a topological space X and two non-empty, disjoint subsets E and F of X, it is not always the case that there is a continuous function $f : X \to \mathbb{R}$ which is constantly 0 on E and constantly 1 on F.

Can you think of an easy example where no such f exists?

What else would need to be true for there to be any chance of finding such a function f?

(b) The converse to Urysohn's lemma is fairly easy to prove: with our definition of normality (which does not require Hausdorffness), any topological space which satisfies the conclusion of Urysohn's Lemma must be a normal topological space (exercise).

We now prove Urysohn's Lemma for metric spaces.

This is an easy consequence of the facts we know about continuous functions in general and the distance from a set to a point in particular.

The proof of the following theorem is beyond the scope of this module, but you are expected to know the statement and be able to apply it.

Theorem 3.38 (Urysohn's Lemma) Let X be a normal topological space. Then, for each pair E and F of disjoint **closed** subsets of X, there exists a continuous function $f : X \to \mathbb{R}$ with the following properties.
(i) For all $x \in X$, $0 \le f(x) \le 1$.
(ii) For all $x \in E$, $f(x) = 0$.
(iii) For all $x \in F$, $f(x) = 1$.

Remarks

(a) Note that the function f is continuous on the whole of X, and is not just defined on E and F!

Theorem 3.39 (Urysohn's Lemma for metric spaces) Let (X, d) be a metric space and let E and F be disjoint **closed** subsets of X. Then there is a continuous function $f : X \to \mathbb{R}$ with the following properties.
(i) For all $x \in X$, $0 \le f(x) \le 1$.
(ii) For all $x \in E$, $f(x) = 0$.
(iii) For all $x \in F$, $f(x) = 1$.

Here the case where E or F is empty is trivial (a constant function will do).

When neither E nor F is empty, the following function has the desired properties:

$$f(x) = \frac{\mathrm{dist}(x, E)}{\mathrm{dist}(x, E) + \mathrm{dist}(x, F)}.$$

In view of the exercise above (the converse to Urysohn's Lemma), we obtain the following corollary (which may also be proved directly: see question sheets).

Corollary 3.40 Every metric space is normal.

Here (i) says that all the sets in $\mathcal{B}$ are τ-open, i.e. open in X, and (ii) says that for all $U \in \tau$, there are $V_i \in \mathcal{B}$ $(i \in I)$ with $U = \bigcup_{i \in I} V_i$. Here we allow the possibility that $I = \emptyset$, so that we always regard $\emptyset$ as being a union of a collection of sets taken from $\mathcal{B}$.

Our first example concerns the case of metric spaces.

Example 3.42 Let (X, d) be a metric space and let τ be the topology on X induced by d. Then each of the following three collections is a base for τ:

(i) $\mathcal{B}_1 = \{B_X(x, r) : x \in X, r > 0\}$;

(ii) $\mathcal{B}_2 = \{B_X(x, r) : x \in X, r \in \mathbb{Q} \cap (0, \infty)\}$;

(iii) $\mathcal{B}_3 = \{B_X(x, 1/n) : x \in X, n \in \mathbb{N}\}$.

3.7 Bases and sub-bases

Roughly speaking, bases and sub-bases for topologies are smaller collections of open sets which, by various means, generate all the remaining open sets in the topology.

For many purposes it is sufficient to work with a base for a topology rather than the whole topology, and this can also make certain properties easier to check.

Definition 3.41 Let (X, τ) be a topological space. Then a **base** (or **analytic base** or **basis**) for the topology τ is a collection $\mathcal{B}$ of subsets of X such that

(i) $\mathcal{B} \subseteq \tau$, and

(ii) for all $U \in \tau$, U is the union of some collection of sets taken from $\mathcal{B}$.

When checking to see whether we have found a base for our topology, the following result is useful.

Proposition 3.43 Let (X, τ) be a topological space and let $\mathcal{B}$ be a collection of subsets of X. Then the following statements (a) and (b) are equivalent.

(a) $\mathcal{B}$ is a base for τ.

(b) The following statements (i) and (ii) both hold:

  (i) $\mathcal{B} \subseteq \tau$, and

  (ii) for all $U \in \tau$ and all $x \in U$, there exists $V \in \mathcal{B}$ with $x \in V \subseteq U$.

The notion of base for a topology is rather different from the notion of a basis for a vector space.

Typically there is a lot of redundancy in a base for a topology, while there is never any redundancy in a basis for a vector space.

We see that a base for a topology generates the topology as the set of all possible unions of sets taken from the base.

If we are allowed to take finite intersections first, then we obtain the notion of a **sub-base**.

Definition 3.44 Let (X, τ) be a topological space and let $\mathcal{B}$ be a collection of subsets of X. Consider the collection $\mathcal{B}'$ defined by

$$\mathcal{B}' = \{X\} \cup \left\{ \bigcap_{k=1}^{n} V_k : n \in \mathbb{N}, V_1, \ldots, V_n \in \mathcal{B} \right\},$$

so that $\mathcal{B}'$ may be thought of as the set of all possible finite intersections of sets taken from $\mathcal{B}$ (including X as the 'empty intersection'). Then we say that $\mathcal{B}$ is a **sub-base** for τ if $\mathcal{B}'$ is a base for τ.

Note here that we always have both $\mathcal{B} \subseteq \mathcal{B}'$ and $X \in \mathcal{B}'$.

Interestingly, **every** collection $\mathcal{B}$ of subsets of X is a sub-base for a (unique) topology on X. This topology is the weakest possible topology on X such that every set in $\mathcal{B}$ is open. This fact follows from the theorem below, which is our main theorem concerning synthetic bases. First we note a lemma, which we essentially proved earlier during the proof of Proposition 3.43.

Lemma 3.47 Let X be a set, let $U \subseteq X$ and let $\mathcal{B}$ be a collection of subsets of X. Then the following statements are equivalent:

(i) U is a union of some collection of sets taken from $\mathcal{B}$;

(ii) $\forall x \in U$, $\exists V \in \mathcal{B}$ with $x \in V \subseteq U$. $(*)$

We now use this lemma to help us characterize synthetic bases.

Example 3.45 Take $X = \mathbb{R}$ with the usual topology, τ. Let $\mathcal{B}$ be the collection of all **unbounded** open intervals in $\mathbb{R}$. Then $\mathcal{B}$ is not a base for τ (why not?) but it is a sub-base. In fact it is easy to see here that $\mathcal{B}'$ (as in the definition above) is the set of all open intervals in $\mathbb{R}$ (empty, bounded or unbounded).

We now wish to look at things the other way round: given a collection $\mathcal{B}$ of subsets of X, is there a topology on X for which $\mathcal{B}$ is a base?

If so we say that $\mathcal{B}$ is a **synthetic base** for a topology on X.

Example 3.46 We may easily find an example of collection $\mathcal{B}$ of subsets of $\mathbb{R}$ which is not a synthetic base for a topology on $\mathbb{R}$. What is the most obvious way to do this? Can you think of a different way?

Theorem 3.48 Let X be a set and let $\mathcal{B}$ be a collection of subsets of X. Then the following statements are equivalent.

(a) The collection $\mathcal{B}$ is a synthetic base for a topology on X.

(b) There is a unique topology τ on X such that $\mathcal{B}$ is a base for τ.

(c) The collection $\mathcal{B}$ satisfies both of the following two conditions:

  (i) $\bigcup_{V \in \mathcal{B}} V = X$, and

  (ii) for all V and V' in $\mathcal{B}$ and $x \in V \cap V'$, there exists $W \in \mathcal{B}$ with $x \in W \subseteq V \cap V'$.

Condition (c)(ii) is equivalent to the following statement:

(ii′) for all V and V' in $\mathcal{B}$, $V \cap V'$ is a union of some collection of sets taken from $\mathcal{B}$.

Next we discuss the related notion of a **base of neighbourhoods** at a point.

Remember that, for us, neighbourhoods need not be open.

Definition 3.49 Let (X, τ) be a topological space and let $x \in X$, Then a **base of neighbourhoods** (or **neighbourhood base**) at x is a collection $\mathcal{N}$ of neighbourhoods of x with the property that for every neighbourhood V of x there is an $N \in \mathcal{N}$ with $N \subseteq V$. This property of a collection of neighbourhoods of x is equivalent to saying that for every open set U with $x \in U$ there is an $N \in \mathcal{N}$ with $N \subseteq U$.

We will meet later notions such as that of a **locally compact** topological space.

This means that every point of the space has a base of compact neighbourhoods.

If you replace compact by connected, you obtain the definition of **locally connected** topological space (etc.).

Example 3.50 Let (X, d) be a metric space and let $x_0 \in X$. As temporary notation, for $r > 0$ set $N_r = \{y \in X : d(y, x_0) \leq r\}$. Then each of the following collections is a neighbourhood base at x_0:

(i) $\mathcal{N}_1 = \{B_X(x_0, r) : r > 0\}$;

(ii) $\mathcal{N}_2 = \{B_X(x_0, r) : r \in \mathbb{Q} \cap (0, \infty)\}$;

(iii) $\mathcal{N}_3 = \{B_X(x_0, 1/n) : n \in \mathbb{N}\}$;

(iv) $\mathcal{N}_4 = \{N_r : r > 0\}$;

(v) $\mathcal{N}_5 = \{N_r : r \in \mathbb{Q} \cap (0, \infty)\}$;

(vi) $\mathcal{N}_6 = \{N_{1/n} : n \in \mathbb{N}\}$.

Remarks Let (X, τ_1) and (Y, τ_2) be topological spaces and let $f : X \to Y$.

(a) Let $x \in X$, let $\mathcal{N}$ be a base of neighbourhoods at $f(x)$ and let $\mathcal{N}'$ be a base of neighbourhoods at x. Then f is continuous at x if and only if the following condition holds: for all $N \in \mathcal{N}$ there exists $N' \in \mathcal{N}'$ with $f(N') \subseteq N$.

(b) Let $\mathcal{B}$ be a base for τ_2. Then f is continuous from X to Y if and only if, for all $U \in \mathcal{B}$, $f^{-1}(U) \in \tau_1$.

(c) Let τ_3 be another topology on Y. Then τ_2 is weaker than τ_3 (i.e., $\tau_2 \subseteq \tau_3$) if and only if $\mathcal{B} \subseteq \tau_3$.

4 Subspaces, quotients and products (new topological spaces from old)

4.1 The subspace topology

Definition 4.1 Let (X, τ) be a topological space, and let $Y \subseteq X$. Then the **subspace** (or **relative**) **topology on Y induced by the topology τ**, denoted by τ_Y, is defined by

$$\tau_Y = \{V \cap Y : V \in \tau\}.$$

Exercise: Check that τ_Y really is a topology on Y.

Example 4.2 (a) Let (X, d) be a metric space and let $Y \subseteq X$.

By Question 11(ii), Question Sheet 3, the subspace topology on Y induced by τ is the same as the topology on Y induced by the subspace metric on Y.

(b) In particular, for any subset Y of $\mathbb{R}^n$, the subspace topology on Y induced by the usual topology on $\mathbb{R}^n$ is the same as the usual topology on Y.

(c) (See Question Sheet 4) Let (X, τ) be a topological space, let $Y \subseteq X$ and let $Z \subseteq Y$.

There are two obvious ways to obtain a subspace topology on Z.

One is the subspace topology τ_Z induced by τ, regarding Z as a subset of X.

The other is to first use τ to induce the subspace topology τ_Y on Y and then to use this topology to induce the subspace topology $(\tau_Y)_Z$ on Z, regarding Z as a subset of the topological space (Y, τ_Y).

It is not hard to show that these two topologies on Z are the same.

(You could say that the subspace topology is the same as the subsubspace topology.)

4.2 The quotient topology

Using the quotient topology, we can make formal the fact that by (suitably) gluing together edges of the closed unit square $[0, 1] \times [0, 1]$, you can obtain (homeomorphic copies of) several standard surfaces.

These include the **cylinder**, the **Möbius band**, the **sphere** and the **torus**.

(See books for more details on these surfaces.)

Definition 4.3 Let (X, τ) be a topological space, let Y be a set and let $q : X \to Y$. Then the **quotient topology** (or **identification topology**) **on Y induced by q** is the topology τ' defined by

$$\tau' = \{U \subseteq Y : q^{-1}(U) \in \tau\}.$$

In other words, given $U \subseteq Y$, U is open in Y if and only if $q^{-1}(U)$ is open in X.

Remarks

(1) We need to check that the quotient topology τ' really **is** a topology on Y.

(2) Among various possible topologies you could put on Y in order to ensure that q is continuous from X to Y, the quotient topology τ' is the strongest possible. (Any other such topology is contained in τ'.)

(3) We do not require q to be surjective, though it usually is in applications.

(4) If q is a bijection, then (Y, τ') is homeomorphic to (X, τ), and q is a homeomorphism from X to Y.

(5) (See Question Sheet 4) Let (Z, τ_3) be another topological space, and let $g : Y \to Z$. Then the function g is continuous from (Y, τ') to Z if and only if $g \circ q$ is continuous from X to Z.

Example 4.4 Let (X, τ) be a topological space, and let $\sim$ be an equivalence relation on X.

For $x \in X$, set $[x] = \{x' \in X : x' \sim x\}$, and set $Y = \{[x] : x \in X\}$. Then Y is the **quotient space of** X **by** $\sim$, and may be denoted by $X/\!\sim$.

The **quotient map** q is then the surjective map from X to Y given by $x \mapsto [x]$.

We use the map q to define the quotient topology on Y.

Various examples of this will be discussed in lectures, including the fact that $\mathbb{R}/\mathbb{Z}$ is homeomorphic to a circle.

For more examples and standard properties of the quotient topology, see the question sheets and books on topology.

4.3 The product topology

Definition 4.5 Let (X, τ_1) and (Y, τ_2) be topological spaces. Then the **product topology** (or **weak topology**) on $X \times Y$ is the topology τ on $X \times Y$ which has base

$$\mathcal{B} = \{U \times V : U \in \tau_1, V \in \tau_2\}.$$

We define the **coordinate projections** $p_X : X \times Y \to X$, $(x, y) \mapsto x$ and $p_Y : X \times Y \to Y$, $(x, y) \mapsto y$.

Remarks In the following, the notation is as above.

(1) The sets in $\mathcal{B}$ are often called **basic** open subsets of $X \times Y$.

We need to check that $\mathcal{B}$ really is a (synthetic) base for a topology.

(2)(a) The coordinate projections are both continuous from $(X \times Y, \tau)$.

(b) If X or Y is empty, then $X \times Y$ is empty: this case is usually uninteresting. Otherwise the coordinate projections are both surjective.

(c) The product topology is the **weakest** topology on $X \times Y$ such that both p_X and p_Y are continuous: the collection

$$\{p_X^{-1}(U) : U \in \tau_1\} \cup \{p_Y^{-1}(V) : V \in \tau_2\}$$

is a sub-base for τ.

(3) With $X = Y = \mathbb{R}$ (usual topology), the product topology on $X \times Y$ is the usual topology on $\mathbb{R}^2$.

This follows from a more general result for metric spaces below.

(4) Given topological spaces X_1, X_2, ..., X_n, form the usual Cartesian product

$$\prod_{k=1}^{n} X_k = X_1 \times X_2 \times \cdots \times X_n .$$

We may define the product topology here inductively, using our definition for products of two spaces.

Or, in terms of coordinate projections, the product topology on the Cartesian product is the weakest topology such that all of the coordinate projections are continuous.

The product topology on $\mathbb{R}^n$ is again the same as the usual topology.

(5) In fact one can define products of infinitely many spaces (even uncountably many). This is, however, beyond the scope of this module.

(6) (See Question Sheet 4) Let Z be a topological space and let $f : Z \to X \times Y$. Then f is continuous from Z to $(X \times Y, \tau)$ if and only if both $p_X \circ f$ and $p_Y \circ f$ are continuous (from Z to X and from Z to Y, respectively).

We conclude this section by looking at products of metric spaces.

The main result is that the product topology on a product of two metric spaces is metrizable, and moreover there are several natural 'product metrics' you can define which induce it.

Theorem 4.7 Let (X, d_X) and (Y, d_Y) be metric spaces. Give each of X and Y the topology induced by its metric, and let τ be the resulting product topology on $X \times Y$. Then τ is metrizable, and is induced by each of the product metrics d_1, d_2 and d_∞ defined in the preceding lemma.

There is no problem in extending these results to products of finitely many metric spaces.

In particular, the usual topology on $\mathbb{R}^n$ is the same as the product topology (the product of n copies of $\mathbb{R}$).

In fact, the theorem above is true for a product of countably many metric spaces, though more care is needed in defining a suitable product metric.

The result fails for uncountable products.

Lemma 4.6 Let (X, d_X) and (Y, d_Y) be metric spaces. For $\mathbf{x} = (x, y)$ and $\mathbf{x}' = (x', y')$ in $X \times Y$, set

$$d_1(\mathbf{x}, \mathbf{x}') = d_X(x, x') + d_Y(y, y')\,;$$

$$d_2(\mathbf{x}, \mathbf{x}') = \sqrt{d_X(x, x')^2 + d_Y(y, y')^2}$$

and

$$d_\infty(\mathbf{x}, \mathbf{x}') = \max\{d_X(x, x'), d_Y(y, y')\}\,.$$

Then d_1, d_2 and d_∞ are all metrics on $X \times Y$. Moreover, these three metrics are all uniformly equivalent to each other, and convergence with respect to these metrics is **coordinatewise convergence**: given $(x, y) \in X \times Y$ and a sequence $((x_n, y_n)) \subseteq X \times Y$, then $(x_n, y_n) \to (x, y)$ as $n \to \infty$ if and only if (x_n) converges to x in X and (y_n) converges to y in Y.

Using these product metrics, we can prove the desired metrizability result for products of two metric spaces.

5 Compactness
5.1 Sequential compactness for metric spaces

The notion of sequential compactness which you met in G1BMAN (see also Chapter 1 of these notes) generalizes easily to metric spaces.

Definition 5.1 Let X be a metric space. Then X is **sequentially compact** if, for every sequence $(x_k) \subseteq X$, (x_k) has at least one subsequence which converges in X.
More generally, a subset A of X is **sequentially compact** if, for every sequence (a_k) of points of A, (a_k) has at least one subsequence which converges to a point of A. (This is the same as saying that A is a sequentially compact metric space when given the subspace metric.)

The main result here from G1BMAN is the characterization of sequentially compact subsets of $\mathbb{R}^n$.

Proposition 5.2 (Heine-Borel theorem, sequential compactness version) Let A be a subset of $\mathbb{R}^n$. Then A is sequentially compact if and only if it is both closed and bounded in $\mathbb{R}^n$.

One special case of this is the Bolzano-Weierstrass Theorem in $\mathbb{R}$, which tells you that the closed and bounded intervals $[a, b]$ are sequentially compact.

Note also that $\mathbb{R}$ is NOT sequentially compact.

We now look more closely at the metric space setting. We begin by proving the boundedness theorem for sequentially compact metric spaces.

Theorem 5.3 (Boundedness theorem, sequential compactness version) Let X be a non-empty, sequentially compact metric space, and let f be a continuous, real-valued function defined on X. Then Then f is bounded on X and moreover f attains both a minimum and a maximum value on X, i.e. there are b and b' in X such that, for all $x \in X$,

$$f(b) \leq f(x) \leq f(b').$$

Remark This result is equally valid for continuous functions defined on a sequentially compact subset of a metric space, since such subsets are themselves sequentially compact metric spaces when given the subspace metric.

We next prove a version of Tychonoff's theorem.

Theorem 5.4 (Tychonoff's theorem for two sequentially compact metric spaces) Let (X, d_X) and (Y, d_Y) be sequentially compact metric spaces. Give $X \times Y$ any of the standard product metrics. Then $X \times Y$ is also sequentially compact.

We conclude this section with a lemma on convergent subsequences.

Lemma 5.5 Let X be a metric space, let $(x_n) \subseteq X$ and let $x \in X$. For each $n \in \mathbb{N}$, set

$$A_n = \{x_n, x_{n+1}, x_{n+2}, \ldots\} = \{x_k : k \geq n\}.$$

Then the following statements are equivalent:

(a) (x_n) has a subsequence which converges to x;

(b) $x \in \bigcap_{n=1}^{\infty} \operatorname{clos} A_n$.

Thus (x_n) has a convergent subsequence if and only if

$$\bigcap_{n=1}^{\infty} \operatorname{clos} A_n \neq \emptyset.$$

5.2 Compactness for topological spaces

In this section, we discuss the standard topological space version of compactness, which is expressed in terms of **open covers**.

Definition 5.6 Let X be a set, let $E \subseteq X$ and let A_i $(i \in I)$ be subsets of X. We say that the indexed collection A_i $(i \in I)$ is a **cover** (or **covering**) of E if $E \subseteq \bigcup_{i \in I} A_i$. Such a cover of E **has a finite subcover** if there is a **finite** subset I' of I such that we also have $E \subseteq \bigcup_{i \in I'} A_i$.
In particular, a collection A_i $(i \in I)$ of subsets of X is a cover of X if and only if

$$X = \bigcup_{i \in I} A_i$$

and such a cover of X has a finite subcover if and only if there is a **finite** subset I' of I such that we also have

$$X = \bigcup_{i \in I'} A_i.$$

We are mainly interested in the case where the sets in the cover are open.

Definition 5.7 Let X be a topological space and let E be a subset of X. An **open cover** (or **open covering**) of E is a cover U_i $(i \in I)$ of E where all of the U_i are open subsets of X. In particular, an open cover of X is a collection U_i $(i \in I)$ of open subsets of X such that

$$X = \bigcup_{i \in I} U_i .$$

Example 5.8 (1) Every topological space X has a finite open cover, since X itself is open in X. In fact, every subset E of X has a one-set open cover using the set X. **This is obviously an uninteresting property!**

(2) (a) The sets $A_n = [-n, n]$ $(n \in \mathbb{N})$ form a cover of $\mathbb{R}$. This is not an open cover. This cover has no finite subcover.

(b) The sets $U_n = (-n, n)$ $(n \in \mathbb{N})$ form an **open** cover of $\mathbb{R}$ which has no finite subcover.

(c) There are **some** open covers of $\mathbb{R}$ which have finite subcovers, since (1) above easily shows that this is true for all topological spaces.

(3) Taking $X = \mathbb{R}$ and $E = [0, 2)$, then the collection $U_n = (-1, 2 - 1/n)$ $(n \in \mathbb{N})$ is an open cover of E which has no finite subcover.

We will see that this **never happens** for closed and bounded intervals such as $[0, 2]$: such intervals are **compact**

Definition 5.9 Let X be a topological space and let E be a subset of X. Then E is **compact** if **every** open cover of E has a finite subcover. In particular, X is compact if (and only if) every open cover of X has a finite subcover.

Remarks (1) It is not enough to find some finite covers of E (or X), or some open covers which have finite subcovers.

This would be true for every subset of every topological space.

For compactness we require that **every** open cover has a finite subcover.

(2) We noted above that **some** open covers of $\mathbb{R}$ do not have finite subcovers.

Thus $\mathbb{R}$ is **NOT** compact.

(3) Given a topological space X and a subset E, we can give E the subspace topology.

We then have two different possible notions of compactness for E: one as a subset of the topological space X, and another as a subset of itself, the topological space E with the subspace topology.

Easy exercise: check that these notions coincide.

As a result we may describe compact subsets of X as **compact subspaces** of X.

(4) Since compactness is defined in terms of open covers, it is easy to see that compactness is a topological property.

This also follows from the fact, to be proved later, that the continuous image of a compact set is compact.

(5) Let X be a metric space. In this case we will see that X is compact if and only if X is sequentially compact.

This is not an easy result: we will need to develop more theory in order to prove it.

(6) Every finite subset of a topological space is compact.

In particular the empty set is compact.

(7) Let X be a set and give X the indiscrete topology.

Then every subset of X is compact.

(8) Let X be a set, and give X the discrete topology.

Then the **only** compact subsets of X are the finite subsets.

5.3 Compactness in terms of closed sets

Suppose that U_i $(i \in I)$ is an open cover of a topological space X.

Working in X, set $F_i = U_i^c = X \setminus U_i$.

The sets F_i are closed, and de Morgan's rules tell us that

$$\bigcap_{i\in I} F_i = \bigcap_{i\in I} U_i^c$$

$$= \left(\bigcup_{i\in I} U_i\right)^c = X^c = \emptyset\,.$$

Moreover, there exists a finite subcover of U_i $(i \in I)$ if and only if there is a finite subset I' of I such that $\bigcap_{i\in I'} F_i = \emptyset$.

The next result is immediate from this correspondence between collections of open sets and of closed sets, and the definition of compact.

Proposition 5.10 Let (X, τ) be a topological space. Then the following are equivalent.
(a) The topological space X is compact.
(b) Whenever F_i $(i \in I)$ is a collection of closed subsets of X satisfying $\bigcap_{i\in I} F_i = \emptyset$, there exists a finite subset I' of I such that $\bigcap_{i\in I'} F_i = \emptyset$.

A more common way to express (b) involves taking the contrapositive, and using the notion of a collection of sets with the **finite intersection property**

Definition 5.11 Let X be any set and let A_i $(i \in I)$ be a collection of subsets of X. We say that this collection has **the Finite Intersection Property** (or **FIP**) if, for every **finite** subset I' of I we have $\bigcap_{i \in I'} A_i \neq \emptyset$.

Remarks and examples

(1) Any nested decreasing sequence of non-empty sets A_n $(n \in \mathbb{N})$ automatically has FIP: the intersection of any finite number of them will simply be the smallest one.

(2) There are many examples of nested decreasing sequences of non-empty sets A_n such that $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$. Can you think of one?

(3) In terms of FIP, condition (b) of Proposition 5.10 (equivalent to compactness of X) says the following:
Whenever F_i $(i \in I)$ is a collection of closed subsets of X satisfying $\bigcap_{i \in I} F_i = \emptyset$, then the collection F_i $(i \in I)$ does not have FIP.
We now take the contrapositive of this.

Theorem 5.12 Let X be a topological space. Then X is compact if and only if the following condition holds: whenever F_i $(i \in I)$ is a collection of closed subsets of X which has the finite intersection property, then

$$\bigcap_{i \in I} F_i \neq \emptyset \, .$$

Exercise. Prove that the following condition on a topological space is also equivalent to compactness: whenever A_i $(i \in I)$ is a collection of subsets of X which has FIP, then $\bigcap_{i \in I} \operatorname{clos} A_i \neq \emptyset$.

Theorem 5.13 Let X be a compact metric space. Then X is sequentially compact.

Proof Let (x_n) be a sequence in X. As in Lemma 5.5, set

$$A_n = \{x_n, x_{n+1}, x_{n+2}, \ldots\} \quad (n \in \mathbb{N}),$$

and set $E_n = \operatorname{clos} A_n$. Then the sets E_n are non-empty, closed subsets of X which are nested decreasing $(E_1 \supseteq E_2 \supseteq E_3 \ldots)$.

Thus the collection of closed sets E_n $(n \in \mathbb{N})$ has the finite intersection property.

Since X is compact, $\bigcap_{k=1}^{\infty} E_k \neq \emptyset$, i.e.,

$$\bigcap_{k=1}^{\infty} \operatorname{clos} A_k \neq \emptyset \, .$$

Thus, by Lemma 5.5, (x_n) has a convergent subsequence. $\square$

5.4 Equivalence of compactness and sequential compactness for metric spaces

In this section we finally prove that compactness and sequential compactness are equivalent for metric spaces.

Along the way we introduce the notions of total boundedness and the Lebesgue number for a cover.

Next we introduce the notion of total boundedness for a metric space.

Definition 5.14 Let X be a metric space. Then X is **totally bounded** if, for all $\varepsilon > 0$, there is a finite subset J of X such that

$$X = \bigcup_{x \in J} B_X(x, \varepsilon)\,.$$

In the case where $X \neq \emptyset$, this is equivalent to the following condition: for all $\varepsilon > 0$ there are $n \in \mathbb{N}$ and $x_1, \ldots, x_n$ in X with

$$X = \bigcup_{k=1}^{n} B_X(x_k, \varepsilon)\,.$$

Exercise Let X be a totally bounded metric space and let Y be any subset of X. Give Y the subspace metric. Prove that Y is also a totally bounded metric space.

Having chosen $x_1, \ldots, x_n$,

$$X \neq \bigcup_{k=1}^{n} B_X(x_k, \varepsilon)$$

(by assumption) so we may choose a point x_{n+1} in $X \setminus \bigcup_{k=1}^{n} B_X(x_k, \varepsilon)$.

Proceeding in this way we construct a sequence (x_n) with the property that

$$d(x_m, x_n) \geq \varepsilon$$

whenever $m, n \in \mathbb{N}$ with $m \neq n$.

Then no subsequence of (x_n) can converge in X (can you see why not?), and so we have the desired contradiction.

The result follows. $\square$

Lemma 5.15 Let X be a sequentially compact metric space. Then X is totally bounded.

Proof We may assume that $X \neq \emptyset$ (otherwise the result is trivial).

Suppose, for contradiction, that X is not totally bounded.

Then there exists an $\varepsilon > 0$ such that X is not equal to any finite union of open balls of radius ε.

Take such an ε (which will be fixed for the remainder of the proof).

We then choose a sequence (x_n) as follows: let x_1 be any point of X.

By assumption $X \neq B_X(x_1, \varepsilon)$.

The supremum of the numbers ε satisfying the conditions of the next Lemma is called the **Lebesgue number** for a given cover.

Lemma 5.16 Let X be a sequentially compact metric space. Let U_i $(i \in I)$ be an open cover of X. Then there is an $\varepsilon > 0$ such that, for all x in X there is at least one i in I satisfying $B_X(x, \varepsilon) \subseteq U_i$.

Proof Suppose for contradiction that $U_i (i \in I)$ is an open cover of X for which no such ε exists.

Then for all $\varepsilon > 0$ there exists an $x \in X$ such that for all i in I, $B_X(x, \varepsilon) \nsubseteq U_i$.

Applying this with $\varepsilon = 1, 1/2, 1/3, \ldots$, we may choose a sequence $(x_n) \subseteq X$ such that $B_X(x_n, 1/n)$ is not entirely contained in any one of the sets U_i.

Since X is sequentially compact, (x_n) has a convergent subsequence, say $x_{n_k} \to x$ in X as $k \to \infty$.

Then there is an $i \in I$ such that $x \in U_i$.

Since U_i is open, see easily that for large k,

$$B(x_{n_k}, 1/n_k) \subseteq U_i$$

(can you see why?), contradicting our choice of (x_n).

The result follows. $\square$

We are finally ready to prove the remaining part of the equivalence of compactness and sequential compactness.

Theorem 5.17 Let X be a sequentially compact metric space. Then X is compact.

Proof As usual, we may assume that $X \neq \emptyset$.

Let $(U_i)_{i \in I}$ be an open cover of X.

By Lemma 5.16, there is an $\varepsilon > 0$ such that for all x in X, there is an $i \in I$ with $B_X(x, \varepsilon) \subseteq U_i$.

Take such an ε.

By Lemma 5.15, there are $x_1, \ldots, x_n$ in X with

$$X = \bigcup_{k=1}^{n} B_X(x_k, \varepsilon).$$

Choose i_k in I such that $B_X(x_k, \varepsilon) \subseteq U_{i_k}$, $k = 1, 2, \ldots, n$.

Then

$$X = \bigcup_{k=1}^{n} U_{i_k}. \quad \square$$

Now that we know that compactness and sequential compactness coincide for metric spaces, we have a good supply of compact topological spaces to work with: all sequentially compact metric spaces are compact.

We also have the following corollaries (which can also be proved by more elementary means), simply because we know the corresponding results for sequentially compact spaces.

Corollary 5.18 Let a and b be in $\mathbb{R}$ with $a \leq b$. Then $[a, b]$ is compact.

Corollary 5.19 (Heine-Borel Theorem for $\mathbb{R}$) The compact subsets of $\mathbb{R}$ are precisely those sets which are both closed and bounded in $\mathbb{R}$.

Corollary 5.20 Let E be a non-empty, compact subset of $\mathbb{R}$. Then E has both a maximum and a minimum element.

Corollary 5.21 (Heine-Borel Theorem for $\mathbb{R}^n$) The compact subsets of $\mathbb{R}^n$ are precisely those sets which are both closed and bounded in $\mathbb{R}^n$.

Warning! Although it is true that compact subsets of general metric spaces must be closed and bounded, the converse is false. Can you think of an example?

Corollary 5.22 (Tychonoff's theorem for two compact metric spaces) Let X and Y be compact metric spaces. Then $X \times Y$ is compact with the product topology.

As usual this result is also valid for finite products.

In fact, Tychonoff's theorem is valid in far greater generality than this.

We do not need to work with metric spaces, and we can take a product of infinitely many (even uncountably many) compact topological spaces.

Every product of compact topological spaces is compact when given the product (weak) topology.

See books for details of the full version of Tychonoff's theorem.

We now show that closed subsets of compact spaces are compact.

Theorem 5.25 Let X be a compact topological space and let E be a closed subset of X. Then E is compact.

This result has a converse if we assume that the topological space is Hausdorff. In fact, every compact subset of a Hausdorff space is closed.

Theorem 5.26 Let X be a Hausdorff topological space and let E be a compact subset of X. Then E is closed in X. In particular, every compact subset of a metric space is closed.

5.5 New compact sets from old

The following result tells us that a continuous image of a compact set is compact. (This gives one easy way to see that compactness must be a topological property.)

Theorem 5.23 Let X and Y be topological spaces, let f be a continuous function from X to Y and let E be a compact subset of X. Then $f(E)$ is a compact subset of Y.

An immediate corollary of this is another version of the boundedness theorem.

Corollary 5.24 (Boundedness Theorem for compact topological spaces) Let X be a non-empty, compact topological space, and let f be a continuous function from X to $\mathbb{R}$. Then f is bounded on X and moreover f attains both a minimum and a maximum value on X.

These results lead to the following remarkable corollaries, which are extremely useful.

Corollary 5.27 Let X be a compact topological space and let Y be a Hausdorff space. Suppose that f is a bijection from X to Y, and that f is continuous. Then f^{-1} is also continuous, and so f is a homeomorphism.

Exercise. Using this corollary, prove carefully the homeomorphism claims made in the section on quotient spaces in Chapter 4.

Corollary 5.28 Let X be a set and let τ_1 and τ_2 be topologies on X, with $\tau_2 \subseteq \tau_1$ (i.e., τ_1 is stronger than τ_2). Suppose that (X, τ_1) is compact and (X, τ_2) is Hausdorff. Then $\tau_1 = \tau_2$: the two topologies coincide. In other words, no Hausdorff topology can be strictly weaker than a compact topology.

6 Connectedness
6.1 Connected and disconnected topological spaces and subsets

Consider the topological space $X = [0, 1] \cup [2, 3]$ (with the usual topology).

By observation, X splits up naturally into two pieces or 'components', namely $[0, 1]$ and $[2, 3]$.

If X is $\mathbb{Q}$, with the usual topology, it is less obvious how X splits up.

In order to understand what the appropriate definition of 'components' should be, we now develop the theory of connectedness, which has many important applications.

Some of the ideas may be familiar from earlier modules.

Definition 6.1 Let (X, τ) be a topological space. Let $E \subseteq X$. Then E is **clopen** in X (or **open-closed** in X) if E is both open and closed in X.

Example 6.2 The topological space $X = [0, 1] \cup [2, 3]$ mentioned above has exactly 4 clopen subsets. Can you see what they must be?

The proof that there are only four is probably not obvious: it will follow from the theory we develop later.

Recall that, for any topological space X, $\emptyset$ and X are **always** clopen in X.

Sometimes they are the **only** clopen subsets of X.

Definition 6.3 Let (X, τ) be a topological space. Then (X, τ) is **connected** if the **only** clopen subsets of X are $\emptyset$ and X. If X is not connected, then it is **disconnected**: thus X is disconnected if and only if there is a clopen subset E of X such that $E \neq \emptyset$ and $E \neq X$.
Let $Y \subseteq X$. Then we say that Y is **connected** or a **connected subset** (or **connected subspace**) of X if Y is a connected topological space when given the subspace topology. Again, if Y is not connected then it is **disconnected**.

There are many conditions equivalent to X being disconnected/connected.

We list some of the most standard of these in the next two propositions.

Proposition 6.4 Let (X, τ) be a topological space. The following conditions on X are all equivalent.

(i) (Our definition of disconnected) There exists a clopen subset E of X such that $E \neq \emptyset$ and $E \neq X$.

(ii) There exist two disjoint, non-empty open subsets U_1 and U_2 of X such that $X = U_1 \cup U_2$.

(iii) There exist two disjoint, open subsets U_1 and U_2 of X such that $X = U_1 \cup U_2$ and neither U_1 nor U_2 is equal to X.

(iv) There exist two disjoint, non-empty closed subsets F_1 and F_2 of X such that $X = F_1 \cup F_2$.

(v) There exist two disjoint, closed subsets F_1 and F_2 of X such that $X = F_1 \cup F_2$ and neither F_1 nor F_2 is equal to X.

In this setting, any such pair of open sets/closed sets is said to **disconnect** X.

Proposition 6.5 Let (X, τ) be a topological space. The following conditions on X are all equivalent.

(i) (Our definition of connected) The only clopen subsets of X are $\emptyset$ and X.

(ii) Whenever U_1 and U_2 are non-empty open subsets of X with $X = U_1 \cup U_2$, then $U_1 \cap U_2 \neq \emptyset$.

(iii) Whenever U_1 and U_2 are disjoint open subsets of X with $X = U_1 \cup U_2$ then $U_1 = \emptyset$ or $U_2 = \emptyset$.

(iv) Whenever U_1 and U_2 are disjoint open subsets of X with $X = U_1 \cup U_2$ then $U_1 = X$ or $U_2 = X$.

(v) Whenever F_1 and F_2 are non-empty closed subsets of X with $X = F_1 \cup F_2$, then $F_1 \cap F_2 \neq \emptyset$.

(vi) Whenever F_1 and F_2 are disjoint closed subsets of X with $X = F_1 \cup F_2$ then $F_1 = \emptyset$ or $F_2 = \emptyset$.

(vii) Whenever F_1 and F_2 are disjoint closed subsets of X with $X = F_1 \cup F_2$ then $F_1 = X$ or $F_2 = X$.

Example 6.6 (1) As we saw above, the topological space $X = [0,1] \cup [2,3]$ (usual topology) is disconnected, because $[0,1]$ is a clopen subset of X which is non-empty and not equal to X.

(2) We shall see later that $\mathbb{R}$ is connected (usual topology) and so is any **interval** in $\mathbb{R}$.

(3) There are many subsets of $\mathbb{R}$ which are not intervals. We shall see that these are disconnected.

(4) Although this fact is not obvious, **open** subsets of $\mathbb{C}$ are connected if and only if they are 'stepwise-connected', as defined in the module G1BCOF: every pair of points can be joined by a continuous path made up of line segments parallel to the axes.

This is a special result for **open** sets: it is easily seen to be false for **general** subsets of $\mathbb{C}$.

(5) We will see that, as with compactness, 'the continuous image of a connected set is also connected': in particular, connectedness is also a topological property.

Example 6.7 Consider the topological space $X = \mathbb{Q}$, with the usual topology.

Let us see that X is disconnected.

Note that $\sqrt{2} \in \mathbb{R} \backslash \mathbb{Q}$.

Set $U_1 = \{q \in \mathbb{Q} \mid q < \sqrt{2}\} = (-\infty, \sqrt{2}) \cap \mathbb{Q}$, and $U_2 = \{q \in \mathbb{Q} \mid q > \sqrt{2}\} = (\sqrt{2}, \infty) \cap \mathbb{Q}$.

Then U_1 and U_2 are open in $\mathbb{Q}$, disjoint, non-empty and not equal to $\mathbb{Q}$, and $U_1 \cup U_2 = \mathbb{Q}$.

Thus $\mathbb{Q}$ is disconnected: the open sets U_1 and U_2 disconnect $\mathbb{Q}$.

When investigating the connectedness of a subset Y of a topological space X (as defined above), it is often convenient to work with sets which are open or closed in X rather than working with the subspace topology on Y.

The following propositions show how the various formulations of connectedness and disconnectedness given above translate into this setting.

Each condition corresponds exactly to the condition with the same number in the appropriate earlier proposition, based on the definition of the subspace topology on Y.

From now on, all of these equivalent formulations will be regarded as standard: we are free to use whichever condition is the most convenient when investigating whether a set is connected or disconnected.

Proposition 6.8 Let (X, τ) be a topological space, let $Y \subseteq X$, and let τ_Y be the subspace topology on Y induced by τ. The following conditions are all equivalent.

(i) (Our definition of Y being disconnected) There exists a τ_Y-clopen subset E of Y such that $E \neq \emptyset$ and $E \neq Y$.

(ii) There exist two open subsets U_1 and U_2 of X such that $U_1 \cap Y \neq \emptyset$, $U_2 \cap Y \neq \emptyset$, $U_1 \cap U_2 \cap Y = \emptyset$ and $Y \subseteq U_1 \cup U_2$.

(iii) There exist two open subsets U_1 and U_2 of X with $U_1 \cap U_2 \cap Y = \emptyset$, and such that $Y \subseteq U_1 \cup U_2$ but neither U_1 nor U_2 contains Y.

(iv) There exist two closed subsets F_1 and F_2 of X such that $F_1 \cap Y \neq \emptyset$, $F_2 \cap Y \neq \emptyset$, $F_1 \cap F_2 \cap Y = \emptyset$ and $Y \subseteq F_1 \cup F_2$.

(v) There exist two closed subsets F_1 and F_2 of X with $F_1 \cap F_2 \cap Y = \emptyset$, and such that $Y \subseteq F_1 \cup F_2$ but neither F_1 nor F_2 contains Y.

As before, any such pair of open sets/closed sets is said to **disconnect** Y.

Proposition 6.9 Let (X, τ) be a topological space, let $Y \subseteq X$, and let τ_Y be the subspace topology on Y induced by τ. The following conditions are all equivalent.

(i) (Our definition of connectedness for Y) The only τ_Y-clopen subsets of Y are $\emptyset$ and Y.

(ii) Whenever U_1 and U_2 are open subsets of X such that $U_1 \cap Y \neq \emptyset$, $U_2 \cap Y \neq \emptyset$, and $Y \subseteq U_1 \cup U_2$, then $U_1 \cap U_2 \cap Y \neq \emptyset$.

(iii) Whenever U_1 and U_2 are open subsets of X with $U_1 \cap U_2 \cap Y = \emptyset$ and $Y \subseteq U_1 \cup U_2$ then $U_1 \cap Y = \emptyset$ or $U_2 \cap Y = \emptyset$.

(iv) Whenever U_1 and U_2 are open subsets of X with $U_1 \cap U_2 \cap Y = \emptyset$ and $Y \subseteq U_1 \cup U_2$ then $Y \subseteq U_1$ or $Y \subseteq U_2$.

(v) Whenever F_1 and F_2 are closed subsets of X such that $F_1 \cap Y \neq \emptyset$, $F_2 \cap Y \neq \emptyset$, and $Y \subseteq F_1 \cup F_2$, then $F_1 \cap F_2 \cap Y \neq \emptyset$.

(vi) Whenever F_1 and F_2 are closed subsets of X with $F_1 \cap F_2 \cap Y = \emptyset$ and $Y \subseteq F_1 \cup F_2$ then $F_1 \cap Y = \emptyset$ or $F_2 \cap Y = \emptyset$.

(vii) Whenever F_1 and F_2 are closed subsets of X with $F_1 \cap F_2 \cap Y = \emptyset$ and $Y \subseteq F_1 \cup F_2$ then $Y \subseteq F_1$ or $Y \subseteq F_2$.

6.2 Connected sets and continuous functions

We now show that a continuous image of a connected set is connected.

Theorem 6.10 Let X and Y be topological spaces. Let E be a connected subset of X, and let $f\colon X \to Y$ be continuous. Then $f(E)$ is a connected subset of Y.

It follows immediately from this theorem (or directly) that connectedness is a topological property.

Let X be a topological space.

It is easy to check that $\emptyset$ is a connected subset of X, and so is any subset with exactly one point.

Any disconnected subset of a topological space must have at least two points.

In the following we use the topological space $D = \{0, 1\}$ (which has exactly 2 points) with the usual topology as a subset of $\mathbb{R}$ (which in this case is the same as the discrete topology on D).

Note that D has exactly 4 subsets: $\emptyset$, $\{0\}$, $\{1\}$, $\{0, 1\}$.

These subsets are all open in D, and so they are also all clopen.

In particular, D is disconnected, as is shown by considering (for example) the clopen subset $E = \{0\}$.

Theorem 6.11 Let X be a topological space, and let D be as above. Then the following statements are equivalent.
(i) The topological space X is connected.
(ii) The only continuous functions from X into D are the constant functions.

Using this theorem, and the Intermediate Value Theorem, it is easy to show that, with the usual topology, every interval in $\mathbb{R}$ is connected. In particular, $\mathbb{R}$ itself is connected.

Theorem 6.12 Let I be an interval in $\mathbb{R}$. Then, with the usual topology, I is connected.

We now introduce the notion of a path.

Definition 6.13 A **path** in a topological space X is a continuous function γ from some non-degenerate closed and bounded interval $[a, b]$ into X.
Such a path is also called a **path from $\gamma(a)$ to $\gamma(b)$ in X**.

Note that paths need not be 1-1.

We conclude this section with a useful corollary.

Corollary 6.14 Let X be a topological space and let γ be a path in X. Then the image of γ in X is connected.

6.3 Classification of connected subsets

Recall that the intervals in $\mathbb{R}$ are $\emptyset$, $\mathbb{R}$, $\{a\}$, $(a \in \mathbb{R})$, $(-\infty, a)$, $(-\infty, a]$, (a, ∞), $[a, \infty)$, $(a \in \mathbb{R})$ and all the usual bounded intervals.

We know that every interval in $\mathbb{R}$ is connected. Are there any **other** connected subsets of $\mathbb{R}$? To see that the answer is no, we use the following lemma characterizing the intervals in $\mathbb{R}$.

Lemma 6.15 Let E be a subset of $\mathbb{R}$. Then E is an interval if and only if the following condition is satisfied: for all a and b in E with $a < b$, we have $[a, b] \subseteq E$.

This allows us to complete our classification of the connected subsets of $\mathbb{R}$.

Corollary 6.16 Let $E \subseteq \mathbb{R}$. Then E is connected if and only if E is an interval.

Example 6.17 Let X be $\mathbb{Q}$, with the usual topology. Which subsets of $\mathbb{Q}$ are connected? We already know that $\emptyset$ is connected, and that $\{q\}$ is connected $(q \in \mathbb{Q})$.

Are these the only ones? The answer is yes: no subset of $\mathbb{Q}$ with more than one point can be connected.

To see this, let E be a subset of $\mathbb{Q}$ which has more than one point. Then there must be q_1 and q_2 in E with $q_1 < q_2$.

Take any irrational number α with $q_1 < \alpha < q_2$. Set

$$U_1 = \{q \in \mathbb{Q} \mid q < \alpha\} \quad \text{and} \quad U_2 = \{q \in \mathbb{Q} \mid q > \alpha\}.$$

Then it is easy to see that these open subsets of $\mathbb{Q}$ disconnect E.

Alternatively

A more sophisticated argument is as follows.

If $E \subseteq \mathbb{Q}$ and E is connected then E must also be a connected subset of $\mathbb{R}$.

The only connected subsets of $\mathbb{R}$ are intervals in $\mathbb{R}$.

The only intervals in $\mathbb{R}$ which are subsets of $\mathbb{Q}$ are $\emptyset$ and single-point sets $\{q\}$ $(q \in \mathbb{Q})$.

6.4 Components, and new connected sets from old

It is clearly not true that a union of connected sets is connected. (Why not?)

We do, however, have the following useful result.

Lemma 6.18 Let X be a topological space, let E_i $(i \in I)$ be connected subsets of X and suppose that there is a point $x_0 \in \bigcap_{i \in I} E_i$. Then $\bigcup_{i \in I} E_i$ is connected.

From Question Sheet 5, we also have that the closure of a connected set is connected.

Lemma 6.19 Let X be a topological space and let E be a connected subset of X. Then $\operatorname{clos} E$ is also connected.

These results enable us to study connected components of a topological space.

Definition 6.20 Let X be a topological space, and let $x \in X$. We define K_x to be the union of all the connected subsets of X which have x as an element, and we call K_x the **component** (or **connected component**) of X containing the point x.

The following proposition summarizes the properties of the connected components K_x.

Proposition 6.21 Let X be a non-empty topological space.

(a) Let $x \in X$. Then $x \in K_x$, K_x is a connected subset of X, and no other connected subset of X strictly contains K_x. Also, K_x is closed in X, and whenever E is a connected subset of X with $E \cap K_x \neq \emptyset$, then $E \subseteq K_x$.

(b) Let x and y be in X. Then either $K_x \cap K_y = \emptyset$ or $K_x = K_y$. Thus the components partition X, and are the equivalence classes of an equivalence relation given by $x \sim y$ if $K_x = K_y$.

(c) The components of X are the maximal connected subsets of X.

(d) The topological space X is connected if and only if X has exactly one component.

Example 6.22 (1) Let $X = [0,1] \cup [2,3]$ (with the usual topology). Then the (connected) components of X are $[0,1]$ and $[2,3]$.

These two subsets of X are connected, and no connected subset of X **strictly** contains one of these intervals.

Note, however, that X has many connected subsets other than its components. (What are they?)

(2) Let X be $\mathbb{Q}$ (with the usual topology). We know that the only connected subsets here are $\emptyset$ and the single-point subsets $\{q\}$ ($q \in \mathbb{Q}$).

The components of $\mathbb{Q}$ are the single-point sets $\{q\}$ for $q \in \mathbb{Q}$ (and **not** $\emptyset$).

When investigating connected subsets of product spaces, it is useful to consider the coordinate projections (the continuous image of a connected set is connected), and also to use the following result.

Theorem 6.23 Let X and Y be connected topological spaces. Then $X \times Y$ is connected with the product topology.

Remarks

(1) It follows easily that finite products of connected topological spaces are still connected. In fact this is true for products of arbitrarily many connected topological spaces. (See books if interested.)

(2) Since intervals are connected in $\mathbb{R}$, it follows that rectangles are connected subsets of $\mathbb{R}^2$ (and this generalizes to higher dimensions).

There are, however, many connected subsets of $\mathbb{R}^2$ which are not rectangles.

In the next section we will see that we can say something useful about the connected **open** subsets of $\mathbb{R}^2$.

(3) A non-empty, topological space is said to be **totally disconnected** if all of its components are single-point sets. As we saw above, one example of this is $\mathbb{Q}$.

(What is strange about the case when a topological space has exactly one point?)
Exercise. Show that a product of two non-empty, totally disconnected topological spaces is also totally disconnected.

6.5 Path-connectedness and open subsets of $\mathbb{R}^n$

We discussed earlier the notion of a path from one point to another in a topological space.

Definition 6.24 A topological space X is **path-connected** if, for all x and y in X, there is a path from x to y in X. Similarly, a subset E of X is **path-connected** if for all x and y in E, there is a path from x to y whose image is contained in E. (This is equivalent to saying that E is path-connected when given the subspace topology.)

Path-connectedness is a stronger condition than connectedness.

Proposition 6.25 Let X be a topological space. Then every path-connected subset of X is connected.

The converse is false, in general (see Question Sheet 5). However one important case where the notions coincide is for **open** subsets of $\mathbb{R}^n$ (and hence also for open subsets of $\mathbb{C}$).

We conclude this chapter with the relevant theorem.

Theorem 6.26 Let U be an open subset of $\mathbb{R}^n$. Then the following conditions are equivalent.

(a) The open set U is connected.

(b) The open set U is path-connected.

(c) For all a and b in U, there is a path from a to b in U made up of finitely many line-segments each of which is parallel to one of the axes.

This last condition corresponds to the condition of 'stepwise connectedness' discussed in G1BCOF.

Remarks

(1) Every convergent sequence in a metric space is Cauchy, but in general Cauchy sequences need not be convergent (as the example above shows).

(2) If (x_n) is a Cauchy sequence in a metric space (X, d), then it is true, in particular, that $\lim_{n\to\infty} d(x_n, x_{n+1}) = 0$.

Warning! The converse is false: the sequence (x_n) defined by $x_n = \sum_{k=1}^{n} 1/k$ ($n \in \mathbb{N}$) is not Cauchy in $\mathbb{R}$, even though it is true that $d(x_n, x_{n+1}) \to 0$ as $n \to \infty$.

(3) Every non-empty metric space has at least one Cauchy sequence. This is not an interesting property of metric spaces.

7 Completeness

7.1 Cauchy sequences in metric spaces

Recall the following example. Let X be $(0, 1)$ with the usual metric, and set $x_n = \frac{1}{2n}$.

The sequence (x_n) **does not** converge in the metric space X.

Nevertheless, $|x_m - x_n| = \left|\frac{1}{2m} - \frac{1}{2n}\right|$ is very small when m and n are both large.

Definition 7.1 Let (X, d) be a metric space. A **Cauchy sequence** in X is a sequence $(x_n) \subseteq X$ satisfying the following condition: for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that, for all $m \geq N$ and $n \geq N$, we have $d(x_m, x_n) < \varepsilon$.
In this setting, we may also use the notation

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

as long as we remember the precise definition given above.

7.2 Complete metric spaces and incomplete metric spaces

We have seen above examples of metric spaces which have some Cauchy sequences which do not converge.

This never happens in a **complete** metric space.

Definition 7.2 A metric space (X, d) is **complete** if **every** Cauchy sequence in X converges in X. Otherwise X is **incomplete**: this means that there is at least one Cauchy sequence in X which does not converge in X.

Remarks Our earlier example shows that the metric space $X = (0, 1)$, with the usual metric, is incomplete: the Cauchy sequence (x_n) given by $x_n = 1/2n$ does not converge in the metric space X.

We will see that every (sequentially) compact metric space is complete, and that $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{C}$ are complete with any of their usual metrics.

Definition 7.3 Let (X, d) be a metric space and let (x_n) be a sequence in X. Then (x_n) is **bounded** if there are an $x \in X$ and an $R > 0$ such that $\{x_1, x_2, \ldots\} \subseteq B_X(x, R)$.

The proof of the following standard facts about Cauchy sequences is an **exercise.**

Proposition 7.4 Let (X, d) be a metric space, and let (x_n) be a Cauchy sequence in X. Then

(i) (x_n) is bounded in X,

(ii) if $x \in X$ is such that some subsequence of (x_n) converges to x, then (x_n) must itself converge to x .

The same argument (based on the sequential compactness of closed and bounded subsets) proves that $\mathbb{R}^n$ and $\mathbb{C}$ are complete with any of their usual metrics.

In fact, it turns out that every finite-dimensional normed space over $\mathbb{R}$ or $\mathbb{C}$ is complete. This is a standard result in the theory of functional analysis.

The situation is different for infinite-dimensional spaces. Recall that $C[0,1]$ is the set of continuous, real-valued functions on $[0,1]$. We have discussed two standard metrics on $C[0,1]$: d_1 and d_∞.

It is easy to give an example of a divergent Cauchy sequence in $(C[0,1], d_1)$.

Proposition 7.7 The metric space $(C[0,1], d_1)$ is incomplete.

Warning! (ii) is **NOT** true for general sequences: divergent sequences may have some convergent subsequences. But this can not happen for Cauchy sequences.

Next we show that every compact metric space is complete. Recall that, for metric spaces, compactness is equivalent to sequential compactness.

Theorem 7.5 Every (sequentially) compact metric space is complete.

We are now ready to prove that $\mathbb{R}$ is complete.

Theorem 7.6 With the usual metric, $\mathbb{R}$ is complete.

The metric d_∞ is better behaved.

The statement of the next theorem is examinable, and you will be expected to be able to apply this result.

However **the proof of this theorem (which will only be sketched) is not examinable.**

Theorem 7.8 The metric space $(C[0,1], d_\infty)$ is complete.

7.3 Characterization of complete subsets

In this section we investigate which subsets of a metric space can be complete when given the subspace metric.

We begin by showing that, with the subspace metric, every closed subset of a complete metric space is closed.

Theorem 7.9 Let (X, d) be a complete metric space, and let E be a closed subset of X. Let $\widetilde{d}$ be the subspace metric on E induced by d. Then $(E, \widetilde{d})$ is complete.

Next we show that, with the subspace metric, complete subsets of metric spaces must be closed.

Theorem 7.10 Let (X, d) be a metric space (not necessarily complete). Let $E \subseteq X$. If E is complete when given the subspace metric, $\widetilde{d}$, induced by d, then E is a closed subset of X.

Combining these two results gives us our characterization of which subsets of complete metric spaces are complete with the subspace metric: they are simply the closed subsets.

Corollary 7.11 Let (X, d) be a **complete** metric space, let $E \subseteq X$ and let $\widetilde{d}$ be the subspace metric on E induced by d. Then $(E, \widetilde{d})$ is complete if and only if E is closed in X.
In particular, if $E \subseteq \mathbb{R}$, then, with the usual metric, E is complete if and only if E is closed in $\mathbb{R}$.

The same characterization is also valid for subsets of $\mathbb{R}^n$ or $\mathbb{C}$ with any of their usual metrics, and also for subsets of the complete metric space $(C[0, 1], d_\infty)$.