Week 5

Compactness in metric and Euclidean spaces. Connectedness

Version 2024/11/11 To accessible online version of this chapter

Metric spaces form a subclass of Hausdorff topological spaces. We can obtain further results about compact sets in this subclass. Recall the following from MATH21111 *Metric spaces*:

> **Definition: bounded set.**
>
> A subset A of a metric space (X, d) is **bounded** if $A \subseteq B_r(x)$ for some $x \in X$, $r > 0$.

We reproduce a result from MATH21111, but with a new proof which refers to the Hausdorff property:

> **Proposition 5.1.**
>
> In a metric space, a compact set is closed and bounded.

Proof. Let K be a compact subset of a metric space X. The metric topology on X is Hausdorff, Proposition 3.2, and compacts are closed in Hausdorff spaces, Proposition 4.4, so K is closed in X.

To show that K is bounded, fix any point x of X and consider the collection $\mathscr{C} = \{B_r(x)\}_{r\in\mathbb{R}_{>0}}$ of open balls. Clearly, $\bigcup\mathscr{C} = X$ and so $\mathscr{C}$ covers K. By Criterion of compactness 4.1, there exists a finite subcollection $\{B_{r_1}(x), \dots, B_{r_n}(x)\}$ of $\mathscr{C}$ which still covers K: that is,

$$K \subseteq B_{r_1}(x) \cup \dots \cup B_{r_n}(x) = B_R(x)$$

where $R = \max(r_1, \dots, r_n)$. We have shown that K is a subset of an open ball, that is, K is bounded. □

Remark: it is **not true** that every closed and bounded subset of a metric space is compact. An easy counterexample is given by X endowed with a **discrete metric,** $d(x,y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$ The metric topology defined by d on X is the **discrete topology** which is **not compact is X is infinite** (exercise).

A more conceptual example of a non-compact closed and bounded set in a metric space is the closed unit ball of an **infinite-dimensional Hilbert, or Banach, space.** This will be discussed in the second part of MATH31010, Functional Analysis.

The next result gives us a highly non-trivial example of a compact set in a Euclidean space.

> **Theorem 5.2: the Heine-Borel Lemma.**
>
> The closed bounded interval $[0, 1]$ is a compact subset of the Euclidean line $\mathbb{R}$.

Proof. There are several standard proofs of this result; we will present the proof by **bisection,** or **halving the interval.** Alternative proofs can be found in the literature.

Assume for contradiction that there exists a collection $\mathscr{C}$ of open subsets of $\mathbb{R}$ such that $[0, 1]$ is covered by $\mathscr{C}$, yet is not covered by any finite subcollection of $\mathscr{C}$. Then at least one of the two halves, the closed subintervals

$[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$

of $[0,1]$, has no finite subcover in $\mathscr{C}$: indeed, if $[0,\frac{1}{2}]$ were covered by finite $\mathscr{C}_1 \subseteq \mathscr{C}$, and $[\frac{1}{2},1]$, by finite $\mathscr{C}_2 \subseteq \mathscr{C}$, then the whole of $[0,1]$ would be covered by $\mathscr{C}_1 \cup \mathscr{C}_2$ which is a finite subcollection of $\mathscr{C}$. We therefore let

$$[a_1,b_1], \text{ where } 0 \le a_1 \le b_1 \le 1 \text{ and } b_1 - a_1 = \tfrac{1}{2},$$

denote one of the halves of $[0,1]$ which is not covered by any finite subcollection of $\mathscr{C}$. We can now apply the same argument to the closed bounded interval $[a_1,b_1]$ and obtain

$$[a_2,b_2], \text{ where } a_1 \le a_2 \le b_2 \le b_1 \text{ and } b_2 - a_2 = \tfrac{1}{2^2},$$

one of the halves of $[a_1,b_1]$ which is not covered by any finite subcollection of $\mathscr{C}$. Continuing this process, we will construct, for all $n \ge 1$, the interval

$$[a_n,b_n], \text{ where } a_{n-1} \le a_n \le b_n \le b_{n-1} \text{ and } b_n - a_n = \tfrac{1}{2^n},$$

which is not covered by any finite subcollection of $\mathscr{C}$.

Observe that the sequence $a_1 \le a_2 \le a_3 \le \ldots$ is increasing and bounded, as all its terms lie in $[0,1]$. By a result from Year 1 Foundations of Mathematics course, such a sequence converges to a limit ℓ, and moreover $\ell \in [0,1]$, $a_n \le \ell$ for all n. Note also that $\lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n + \frac{1}{2^n} = \ell$ and $\ell \le b_n$ for all n.

The point ℓ of $[0,1]$ must be covered by some set $U \in \mathscr{C}$. Since U is open, $U \supseteq (\ell-\varepsilon,\ell+\varepsilon)$ for some $\varepsilon > 0$. Take n such that $\frac{1}{2^n} < \varepsilon$. Then $\ell - \varepsilon < \ell - \frac{1}{2^n} \le a_n \le \ell$ and so $a_n \in (\ell-\varepsilon,\ell+\varepsilon)$. Similarly, $b_n \in (\ell-\varepsilon,\ell+\varepsilon)$, see Figure 5.1 for illustration.

Thus, the interval $[a_n,b_n]$ has a finite subcover in $\mathscr{C}$ — in fact, a cover by just one set, $U \in \mathscr{C}$ — contradicting the construction of $[a_n,b_n]$. This contradiction proves the Theorem. □

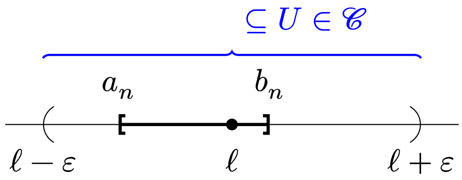

Figure 5.1: *the contradiction arrived at in the proof of the Heine-Borel Lemma*

Remark: it follows from the Heine-Borel Lemma that in every Euclidean space $\mathbb{R}^n$, a set is compact iff it is closed and bounded. This is because every such set is a subset of a cube of the form $[-M, M]^n \subseteq \mathbb{R}^n$. The closed bounded cube can be shown to be compact by adapting the subdivision argument in Theorem 5.2.

We will, however, not extend the subdivision argument to n dimensions: in a short while, compactness of $[-M, M]^n$ will follow from the **(baby) Tychonoff theorem** which will say that a direct product of n compact spaces is compact.

Connectedness

So far, we have considered two topological properties: the Hausdorff property and compactness. The third topological property, and the final one that we will study in this course, is connectedness.

Definition: connected.

A topological space X is **disconnected** if

$$\exists U, V \text{ open in } X\text{: } U \neq \emptyset,\ V \neq \emptyset,\ U \cap V = \emptyset,\ U \cup V = X.$$

That is, **a disconnected space is a disjoint union of two non-empty open sets.**

The space X is **connected** if it is not disconnected.

Let us construct two **disconnected** spaces. We will obtain them as subspaces of the Euclidean line $\mathbb{R}$:

Example.

Show that the subspace $X = (0,1) \cup (2,3)$ of $\mathbb{R}$ is disconnected.

Solution: let U be the open interval $(0,1)$ and V be the open interval $(2,3)$. Then U and V are disjoint non-empty sets open in X whose union is X. Hence, by definition, X is disconnected.

Example.

Show that the subspace $X = \{0,1\}$ of $\mathbb{R}$ is disconnected.

Solution. Consider the following non-empty disjoint subsets of X: $U = \{0\}$ and $V = \{1\}$. We note that both sets are open in X. For example, $\{0\} = (-\infty,1) \cap X$ where $(-\infty,1)$ is an open subset of $\mathbb{R}$, and so $\{0\}$ is open in X by definition of subspace topology. Similarly, $\{1\} = (0,+\infty) \cap X$.

We have written X as a disjoint union of two non-empty open subsets, so by definition, X is disconnected.

Remark. By showing $\{0\}$ and $\{1\}$ to be open in $\{0,1\}$, we have proved the following:

Claim.

The subspace $\{0,1\}$ of the Euclidean line $\mathbb{R}$ has discrete topology. $\square$

One can say that the discrete space $\{0,1\}$ is the "canonical" example of a disconnected space. This point of view is supported by the following result.

Proposition 5.3: conditions equivalent to connectedness.

The following are equivalent for a topological space X:

(i) X is connected.

(ii) For all continuous functions $f\colon X \to \mathbb{R}$, the image $f(X)$ is an interval.

(iii) Every continuous function $g\colon X \to \{0,1\}$ is constant.

Here $\mathbb{R}$ has Euclidean topology, and $\{0,1\}$ has discrete topology.

Before proving the Proposition, we formally define "interval".

Definition: interval.

An **interval** is a subset I of $\mathbb{R}$ such that

$$\forall x, y \in I, \quad \forall t, \quad x < t < y \implies t \in I.$$

Thus, **together with any two points, an interval contains all intermediate points.**

It is not difficult to establish the following classification of intervals:

Claim: classification of intervals in $\mathbb{R}$.

$I \subseteq \mathbb{R}$ is an interval $\iff$ I is a set of the form (a,b), $[a,b)$, $(a,b]$, $[a,b]$, $(a,+\infty)$, $[a,+\infty)$, $(-\infty,b)$ or $(-\infty,b]$ for some real a,b, or $I = \mathbb{R}$.

Remark: the empty subset $\emptyset$ of $\mathbb{R}$ can be written as (a,a) and is an interval. Any singleton $\{a\} \subseteq R$ can be written as $[a,a]$ and is an interval.

Sketch of proof of the Claim (not given in class). It is clear that every set (a,b), $[a,b)$, ... , $\mathbb{R}$, listed in the claim, is an interval. Conversely, assume that I is an interval in $\mathbb{R}$ and let $a = \inf I$, $b = \sup I$. If both a and b are real numbers and not $\pm\infty$ and so I is non-empty and bounded, one uses the definition of the "least upper bound" $(\sup)$ and the "greatest lower bound" $(\inf)$ to show that $(a,b) \subseteq I \subseteq [a,b]$, which leaves four possibilities for I. The remaining cases when one or both of a,b is infinite are handled similarly. Details are left to the student. □

Proof of Proposition 5.3. Recall for use in the proof that the preimage of intersection is the intersection of preimages; same for the union and the complement.

(i) ⇒ (ii): to prove the contrapositive, we must assume that there exists a continuous function $f\colon X \to \mathbb{R}$ such that $f(X)$ is not an interval. This means that there are $a, b \in f(X)$ and a real number t such that $a < t < b$ and $t \notin f(X)$. Consider

$$U = f^{-1}((-\infty,t)), \quad V = f^{-1}((t,+\infty)).$$

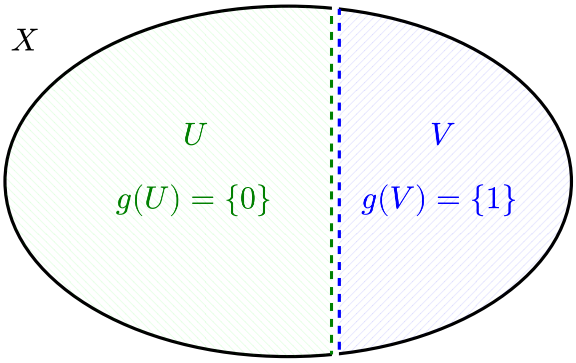

Figure 5.2: *defining $g\colon X \to \{0,1\}$ where X is disconnected*

The set $U \subseteq X$ is open in X because U is the preimage of the set $(-\infty, t)$, open in $\mathbb{R}$, under a continuous function f. Furthermore, U is not empty, as $f(U) \ni a$. Similarly, V is open in X and not empty; yet $U \cap V$ is

$$f^{-1}((-\infty,t)) \cap f^{-1}((t,+\infty)) = f^{-1}((-\infty,t) \cap (t,+\infty)) = f^{-1}(\emptyset),$$

so $U \cap V = \emptyset$ and U, V are disjoint. Finally,

$$U \cup V = f^{-1}(\mathbb{R} \setminus \{t\}) = X \setminus f^{-1}(\{t\}),$$

yet $t \notin f(X)$ by assumption, so $f^{-1}(t) = \emptyset$ and $U \cup V = X$. The construction of U, V shows that X is disconnected by definition. We have shown not(ii) $\Rightarrow$ not(i).

(ii) $\Rightarrow$ (iii): let $g\colon X \to \{0,1\}$ be continuous. As in an earlier example, we can view the discrete set $\{0,1\}$ as a subspace of the Euclidean line $\mathbb{R}$. The inclusion map $\mathrm{in}\colon \{0,1\} \to \mathbb{R}$ is continuous by Proposition 2.7, and so we have a continuous map $X \xrightarrow{g} \{0,1\} \xrightarrow{\mathrm{in}} \mathbb{R}$. By (ii), the image of this map must be an interval in $\mathbb{R}$; yet this image is a subset of $\{0,1\}$, and such an interval can be at most one point. Thus, g must be constant, proving (iii).

(iii) $\Rightarrow$ (i): to prove the contrapositive, assume that X is disconnected, so that $X = U \cup V$ where U, V are disjoint non-empty open sets. This allows us to define the following function (as illustrated in Figure 5.2):

$$g\colon X \to \{0,1\}, \quad g(x) = \begin{cases} 0, & x \in U, \\ 1, & x \in V. \end{cases}$$

We check that g is continuous by calculating the preimage of every open subset of $\{0, 1\}$. There are exactly 4 open subsets of the discrete space $\{0, 1\}$, and we have $g^{-1}(\emptyset) = \emptyset$, $g^{-1}(\{0\}) = U$, $g^{-1}(\{1\}) = V$ and $g^{-1}(\{0, 1\}) = X$. In each case, the preimage is an open subset of X, so g is continuous by definition. We have constructed a non-constant continuous function $X \to \{0, 1\}$, proving not(i) ⇒ not(iii). □

Here is an example where we use the Proposition 5.3 to establish that a space is connected.

Example: interval is connected.

Show that an interval $I \subseteq \mathbb{R}$ is connected.

Solution: we use condition (ii) from the Proposition. Let $f: I \to \mathbb{R}$ be a continuous function. To show that $f(I)$ is an interval, take two points $f(a)$ and $f(b)$ of $f(I)$, where $a, b \in I$. By the **Intermediate Value Theorem** from MATH11121 *Mathematical Foundations and Analysis,* all intermediate values between $f(a)$ and $f(b)$ belong to the image $f(I)$. Thus, by definition of "interval", $f(I)$ is an interval. Hence by Proposition 5.3, I is connected, as claimed.

References for the week 5 notes

The **bisection proof** of the **Heine-Borel Lemma** that we present in Theorem 5.2, is given in [Armstrong, Theorem (3.3)] and [Sutherland, Exercise 13.15]. An alternative "creeping-along" proof can be found in [Armstrong, Theorem (3.3)] and [Sutherland, Exercise 13.9].

Our **definition of "connected"** is identical to [Willard, Definition 26.1]. We thus deviate slightly from [Sutherland] which uses condition (iii) of our Proposition 5.3 as a definition [Sutherland, Definition 12.1], and turns our definition into a theorem, [Sutherland, Proposition 12.3]. The way [Armstrong] defines "connected", though equivalent to ours, requires the notion of closure of a set; but in our course closure comes after connectedness.

The definition of **interval** is elementary and is taken from [Smith, Definition 1.11]. The claim about classification of intervals is made in [Sutherland, Chapter 2: Notation and terminology].