

Metric Spaces

Joseph Muscat 2003
(Last revised May 2009)

(A revised and expanded version of these notes are now published by
[Springer](#).)

1 Distance

A metric space can be thought of as a very basic space having a geometry, with only a few axioms. Metric spaces are generalizations of the real line, in which some of the theorems that hold for \mathbb{R} remain valid. Some of the main results in real analysis are

- (i) Cauchy sequences converge,
- (ii) for continuous functions $f(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} f(x_n)$,
- (iii) continuous real functions are bounded on intervals of type $[a, b]$ and satisfy the intermediate value theorem, etc.

At first sight, it is difficult to generalize these theorems, say to sequences and continuous functions of several variables such as $f(x, y)$. That is the aim of this abstract course: to show how these theorems apply in a much more general setting than \mathbb{R} . The fundamental ingredient that is needed is that of a distance or metric. This is not enough, however, to get the best results and we need to specify later that the distance be of a nice type, called a *complete* metric.

In what follows the metric space X will denote an abstract set, not necessarily \mathbb{R} or \mathbb{R}^n , although these are of the most immediate interest. When we refer to “points” we do not necessarily refer to geometrical points, although this is how most of us visualize them. They may in fact be sequences, functions, images, sounds, signals, etc.

Definition A **distance** or **metric** on a metric space X is a function

$$d : X^2 \mapsto \mathbb{R}^+ \\ (x, y) \mapsto d(x, y)$$

with the properties

- (i) $d(x, y) = 0 \Leftrightarrow x = y$
- (ii) $d(y, x) = d(x, y)$
- (iii) $d(x, y) \leq d(x, z) + d(y, z)$

for all $x, y, z \in X$.

1.0.1 Example

On \mathbb{N} , \mathbb{Q} , \mathbb{R} , \mathbb{C} , and \mathbb{R}^N , one can take the standard Euclidean distance $d(x, y) := |x - y|$. Check that the three axioms for a distance are satisfied (make use of the fact that $|a + b| \leq |a| + |b|$). Note that for \mathbb{R}^n , the Euclidean distance is $d(\mathbf{x}, \mathbf{y}) := \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$.

One can define distances on other more general spaces, e.g. the space of continuous functions $f(x)$ with $x \in [0, 1]$ has a distance defined by $d(f, g) := \max_{x \in [0, 1]} |f(x) - g(x)|$. The space of shapes (roughly speaking, subsets of \mathbb{R}^2 having an area) have a metric $d(A, B) := \text{area of } (A \cup B \setminus A \cap B)$. In all these cases we get an idea of which elements are close together by looking at their distance.

1.0.2 Exercises

1. Show that if x_1, \dots, x_n are n points, then

$$d(x_1, x_n) \leq d(x_1, x_2) + \dots + d(x_{n-1}, x_n).$$

2. Verify that the metric defined on \mathbb{R}^2 does indeed satisfy the metric axioms.
3. Show that if d is a metric, then so are the maps $D_1(x, y) := 2d(x, y)$ and $D_2(x, y) := \frac{d(x, y)}{1 + d(x, y)}$, but that $d(x, y)^2$ need not be a metric. Hence a metric space can have several metrics.
4. Show that if X, Y are metric spaces with distances d_X and d_Y , then $X \times Y$ is also a metric space with distance

$$D\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) := d_X(x_1, x_2) + d_Y(y_1, y_2).$$

(Note that for \mathbb{R}^2 , this metric is not the Euclidean one.)

5. Show further that

$$\tilde{D}\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) := \max(d_X(x_1, x_2), d_Y(y_1, y_2))$$

is also a metric for $X \times Y$.

6. Prove that the function $d(m, n) := |1/m - 1/n|$, on the natural numbers, is a metric.

7. Show that in fact the axioms (i) and (iii) imply both $d(x, y) \geq 0$ and axiom (ii).

8. Show that

$$d(x, y) \geq |d(x, z) - d(y, z)|.$$

9. Consider the set X of bytes, i.e., sequences of 0s and 1s of length 8. The Hamming distance between two bytes is defined as the number of places where their bits differ. Show that this is a distance on X .

1.1 Balls

Definition An (open) **ball** is the set

$$B_r(a) := \{x \in X : d(x, a) < r\}$$

where $a \in X$ is its *center* and $r > 0$ its *radius*.

Note: The ball is the “surroundings” of the point a . One can consider \emptyset and X to be balls with $r = 0$ and $r = \infty$ respectively, but this is unusual. Although we call it a ball, one must remove any preconceptions of it being round etc.

In \mathbb{R} , every ball $B_r(a) = \{x \in \mathbb{R} : |x - a| < r\} = (a - r, a + r)$ is just an open interval. In particular, $B_{1/2}(a) = (a - 1/2, a + 1/2)$. Conversely, any open interval of the type (a, b) is a ball in \mathbb{R} , with center $(a + b)/2$ and radius $(b - a)/2$.

In \mathbb{R}^2 , the ball $B_r(\mathbf{a})$ is the disk with center \mathbf{a} and radius r without the circular perimeter.

In \mathbb{N} , the ball $B_{1/2}(m) = \{m\}$, and $B_2(m) = \{m - 1, m, m + 1\}$.

Definition A point x of a set A is called an **interior** point of A when

$$\exists \epsilon > 0 \quad B_\epsilon(x) \subseteq A.$$

A point x (not in A) is an **exterior** point of A when

$$\exists \epsilon > 0 \quad B_\epsilon(x) \subseteq X \setminus A.$$

All other points of X are called **boundary** points. A point x of A is called an **isolated** point when there is a ball $B_\epsilon(x)$ which contains no points of A other than x itself.

Definition A set A is **open** in X when all its points are interior points.

Note: “An interior point of A can be surrounded completely by a ball inside A ”; “open sets do not contain their boundary”.

We also say that A is a *neighborhood* of a when a is an interior point of A .

The empty set is open by default, because it does not contain any points. The whole space X is also open because any ball about any point of X is a subset of X .

1.1.1 Example

Show that (a, b) is open in \mathbb{R} , but that $[a, b]$ and $\{a\}$ are not.

Theorem A

Balls are open sets in X .

PROOF Let $x \in B_r(a)$ be any point in the given ball. This means that $d(x, a) < r$. Let $\epsilon := r - d(x, a)$ which is positive. Then $B_\epsilon(x) \subseteq B_r(a)$ since for any $y \in B_\epsilon(x)$,

$$d(y, a) \leq d(y, x) + d(x, a) < \epsilon + d(x, a) = r.$$

Therefore $y \in B_r(a)$.

□

Open sets, however, need not be balls.

Note that we need to specify that A is open *in* X . For example we have seen that $\{m\}$ is open in \mathbb{N} (since it is a ball in \mathbb{N}) but not open in \mathbb{R} (see previous example).

Theorem B

A set is open \Leftrightarrow it is the union of balls.

PROOF Let A be an open set. Then

$$\forall x \in A \quad \exists \epsilon > 0 \quad x \in B_\epsilon(x) \subseteq A$$

which implies that

$$A = \bigcup_{x \in A} \{x\} \subseteq \bigcup_{x \in A} B_\epsilon(x) \subseteq A.$$

Hence $A = \bigcup_{x \in A} B_\epsilon(x)$, a union of balls.

Conversely, let $A = \bigcup_i B_{r_i}(a_i)$ be a union of balls, and let x be any point in A . Then x is in at least one of these balls, say, $B_r(a)$. But balls are open and hence $x \in B_\epsilon(x) \subseteq B_r(a) \subseteq A$. A is therefore open. \square

One can study open sets without reference to balls or metrics in the subject of *topology*. The basic properties of open sets are:

Theorem C

Any union of open sets is open.
Any finite intersection of open sets is open.

PROOF Consider $\bigcup_i A_i$ where A_i are all open. Given any $x \in \bigcup_i A_i$, it must lie in at least one of these sets A_i which is open. Therefore

$$x \in B_r(x) \subseteq A_i \subseteq \bigcup_i A_i$$

showing that the union is open.

For the second part it is enough to consider the intersection of two open sets $A \cap B$ (the rest can be done by induction). Let $x \in A \cap B$. Then $x \in A$ and $x \in B$, both sets being open. Therefore there are open balls $B_{r_1}(x) \subseteq A$ and $B_{r_2}(x) \subseteq B$. Pick the smaller of these two balls, say the one with radius r_1 . Then,

$$x \in B_{r_1}(x) \subseteq B_{r_1}(x) \cap B_{r_2}(x) \subseteq A \cap B.$$

\square

1.1.2 Example

In \mathbb{R} the intersection of two open intervals is either empty or else another open interval, both of which are open. Any collection of open intervals is open.

Note that an infinite intersection of open sets need not be open. For example, in \mathbb{R} , consider the open intervals $(-1/2^n, 1/2^n)$ which are nested one inside another. Their intersection is just the set $\{0\}$ (prove this!) which is not open in \mathbb{R} .

Theorem D

Distinct points in X can be separated by disjoint open balls.

$$x \neq y \Rightarrow \exists r > 0 \quad B_r(x) \cap B_r(y) = \emptyset$$

PROOF If $x \neq y$ then $d(x, y) > 0$. Let $r := d(x, y)/2$. Then $B_r(x) \cap B_r(y) = \emptyset$ since otherwise,

$$\begin{aligned} z \in B_r(x) \cap B_r(y) &\Rightarrow d(x, z) < r \text{ and } d(y, z) < r \\ &\Rightarrow d(x, y) \leq d(x, z) + d(y, z) < 2r = d(x, y) \end{aligned}$$

which is a contradiction. □

1.1.3 Exercises

1. If $r_1 < r_2$, show that $B_{r_1}(x) \subset B_{r_2}(x)$.
2. Show that in \mathbb{R} , the sets (a, ∞) and $(-\infty, a)$ are open sets.
3. Show that in \mathbb{R}^2 , (i) the half-plane $S = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ and (ii) the rectangles $R = \{(x, y) \in \mathbb{R}^2 : a < x < b, c < y < d\}$ are open sets.
4. Show that the interval (a, b) on the x -axis, is open in \mathbb{R} but not open in \mathbb{R}^2 .

5. Find an example in \mathbb{R}^2 in which the infinite intersection of open sets is not open.
6. Show that for any metric space X , the set $X \setminus \{x\}$ is open in X .
7. Prove properly by induction, that the finite intersection of open sets is open.
8. Show that if $\{x\}$ are open sets in X for all points $x \in X$, then all subsets of X are also open in X . In particular, all subsets of \mathbb{N} are open in \mathbb{N} .
9. Given any set X , let

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

Show that d is a metric (it is called the *discrete metric*), for which the open balls are just the sets $\{x\}$ and X . Deduce that all the subsets of X with this metric are open in X .

10. * Show that the set of interior points of A is the largest open set inside A , i.e., if A° denotes the set of interior points of A , and $V \subseteq A$ is an open set, then $V \subseteq A^\circ$.

2 Closed Sets and Convergence

2.1 Closed Sets

Definition A set F is **closed** in a space X when $X \setminus F$ is open in X .

2.1.1 Example

On \mathbb{R} the set $[a, b]$ is closed, since $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ is the union of two open sets, hence itself open. Similarly $[a, \infty)$ and $(-\infty, a]$ are closed in \mathbb{R} .

\mathbb{N} is closed in \mathbb{R} , but \mathbb{Q} is not.

On any metric space X , the sets X and \emptyset are closed since $X \setminus \emptyset = X$ and $X \setminus X = \emptyset$ are open in X .

Proposition 2.1

The sets $\{x\}$ are closed in X .

Proof. Exercise.

Theorem A

**The finite union of closed sets is closed.
Any intersection of closed sets is closed.**

PROOF

$$X \setminus (F \cup G) = (X \setminus F) \cap (X \setminus G)$$

which is the intersection of two open sets, hence open.

$$X \setminus \bigcap_i F_i = \bigcup_i (X \setminus F_i)$$

which is the union of open sets, hence open.

□

2.1.2 Exercises

1. Let U be an open set. Show that $X \setminus U$ is closed in X .
2. Show that the ‘closed ball’ $\{x \in X : d(x, a) \leq r\}$ is closed.
3. Let U be an open set and F a closed set in X . Show that $U \setminus F$ is open and $F \setminus U$ is closed.
4. Show that a finite collection of points $\{a_1, \dots, a_N\}$ form a closed set.
5. Find a set in \mathbb{R} which is (i) neither open nor closed; (ii) both open and closed.
6. Find an example in which the infinite union of closed sets (i) is not closed; (ii) is closed.
7. Determine which of the following sets in \mathbb{R} are closed, open, neither or both: (i) $(0, \infty)$; (ii) \mathbb{Z} ; (iii) $\bigcup_n [0, 2 - 1/n)$; (iv) \mathbb{R} ; (v) $[0, 1] \setminus \bigcup_{n=2}^{\infty} \{1/n\}$; (vi) $[0, 1] \cup \{5\}$.
8. * Let the Cantor set be defined as follows. Start with the closed interval $[0, 1]$; remove the middle interval $(1/3, 2/3)$ to end up with two closed intervals $[0, 1/3] \cup [2/3, 1]$. For each of these intervals remove the middle interval of each to end up with four closed intervals $[0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$. Show that if we continue this process indefinitely we end up with a closed set.

Definition A **limit point** of a set A is a point b (not necessarily in A) such that

$$\forall \epsilon > 0 \quad \exists a \neq b \quad a \in A \cap B_{\epsilon}(b)$$

In other words, a limit point is one that cannot be ‘separated’ from its set — it can be an interior point or part of the boundary. Every point of A is either a limit point or an isolated point.

Theorem B

A set is closed \Leftrightarrow it contains all its limit points.

PROOF Let F be a closed set, and let x be a limit point of it.

$$x \notin F \Rightarrow x \in X \setminus F \Rightarrow x \in B_\epsilon(x) \subseteq X \setminus F$$

which shows that $B_\epsilon(x) \cap F = \emptyset$, a contradiction since x is a limit point of F .

Conversely, suppose F contains all its limit points, and let $x \notin F$. Then x is not a limit point of F . Therefore,

$$\exists \epsilon > 0 \quad \forall y \neq x \quad y \notin F \cap B_\epsilon(x)$$

$$\text{i.e.,} \quad F \cap B_\epsilon(x) \subseteq \{x\} \subseteq X \setminus F$$

$$\therefore x \in B_\epsilon(x) \subseteq X \setminus F$$

which means that x can be surrounded by a ball inside $X \setminus F$. Therefore $X \setminus F$ is open, which by definition means that F is closed. □

Definition Any set A can be closed by adding its limit points to form the **closure** of A

$$\bar{A} = A \cup \{\text{limit points of } A\}$$

Note: the set of limit points of a set is sometimes called its *derived set*.

2.1.3 Exercises

1. Find the limit points and closure of \mathbb{Z} and \mathbb{Q} in \mathbb{R} .
2. Show that $\overline{(a, b)} = [a, b]$; $\overline{\{1/n\}} = \{1/n\} \cup \{0\}$.
3. Can a set *not* have limit points? Can an infinite set not have limit points?
4. A set A is said to be *dense* in X when $\bar{A} = X$. Show that any ϵ -ball of X must contain elements of A . Prove that \mathbb{Q} is dense in \mathbb{R} .
5. Let A be a bounded non-empty set in \mathbb{R} . Show that $\sup A$ is a limit point of both A and $\mathbb{R} \setminus A$.
6. Show that $x \in \bar{A} \Leftrightarrow \inf_{a \in A} d(x, a) = 0$.

7. Show that any point of A must be either a limit point of A or an isolated point of A .
8. * Show that $\overline{A \cup B} = \overline{A} \cup \overline{B}$; $A \subseteq B \Rightarrow \overline{A} \subseteq \overline{B}$; A is closed $\Leftrightarrow A = \overline{A}$.
9. * Show that \overline{A} is the smallest closed set containing A i.e., if $F \supseteq A$ is a closed set, then $\overline{A} \subseteq F$.
10. ** Show that the limit points of \overline{A} belong to \overline{A} so that $\overline{\overline{A}} = \overline{A}$.
11. ** Let the decimal expansion of x be $0.n_1n_2n_3\dots$. Show that

$$\left\{ x : \frac{n_1 + \dots + n_k}{k} \leq 5 \quad \forall k \right\}$$

is closed in \mathbb{R} .

2.2 Convergence

Definition A sequence (x_n) in the metric space X **converges** to x , written

$$x_n \rightarrow x \text{ as } n \rightarrow \infty, \quad \text{or} \quad \lim_{n \rightarrow \infty} x_n = x,$$

when

$$\forall \epsilon > 0 \quad \exists N \quad n \geq N \Rightarrow x_n \in B_\epsilon(x).$$

In other words, “any neighborhood of x contains all the sequence from N onwards.” This definition generalizes the definition of convergence for real sequences — the expression $|x_n - x| < \epsilon$ generalizes to $d(x_n, x) < \epsilon$ which is the same as $x_n \in B_\epsilon(x)$.

Proposition 2.2

A sequence can only converge to one point.

PROOF Suppose $x_n \rightarrow x$ and $x_n \rightarrow y$ as $n \rightarrow \infty$. Suppose that $x \neq y$ then they can be separated by two disjoint balls $B_r(x)$ and $B_r(y)$. But

$$\exists N_1 \quad n > N_1 \Rightarrow x_n \in B_r(x)$$

and

$$\exists N_2 \quad n > N_2 \Rightarrow x_n \in B_r(y).$$

Therefore for $n > N_1, N_2$ we have that $x_n \in B_r(x) \cap B_r(y) = \emptyset$ which is a contradiction.

□

Theorem C

If $x_n \in F$ converges to x , then $x \in \bar{F}$.

**If (x_n) is a convergent sequence in a closed set F ,
then $\lim_{n \rightarrow \infty} x_n \in F$.**

PROOF If x is in F then we're done. So suppose that the limit $x \notin F$. Take any ball $B_\epsilon(x)$ about x . Since x_n converges to x , we can find an integer N such that $n \geq N \Rightarrow x_n \in B_\epsilon(x)$. Hence x is a limit point of F and so is in \bar{F} .

The corollary follows immediately from the theorem.

□

2.2.1 Exercises

1. Show that if x is a limit point of the set A then there is a sequence (a_n) in A which converges to x . (Hint: take balls around x , with radii decreasing to 0)
2. Show that if $x_n \rightarrow x$ then $d(x_n, x_m) \rightarrow 0$ as $n, m \rightarrow \infty$.
3. Show that a sequence $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ in $X \times Y$ converges to $\begin{pmatrix} x \\ y \end{pmatrix}$ if, and only if, $x_n \rightarrow x$ and $y_n \rightarrow y$. (Recall the distance defined on $X \times Y$.)
4. Show that if $d_1(x, y) \leq \lambda d_2(x, y)$ ($\lambda > 0$) and the sequence (x_n) converges with respect to d_2 , then it converges with respect to d_1 .
5. Try to generalize the definition of $\lim_{x \rightarrow x_0} f(x)$ from the case of real functions to the case of functions $f : X \rightarrow Y$ on metric spaces.

2.3 Completeness

When does a sequence converge/diverge? Let us take some examples of divergent sequences:

(n) diverges in \mathbb{R} because the set $\{n\}$ seems to be too large;

$(1/n)$ diverges in the set $(0, 1)$ because the point it's trying to converge to, 0, is not in the set;

$(-1)^n$ diverges because it keeps alternating between two points;

(x_n) where $x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$ in \mathbb{Q} diverges because were it to converge then the limit would satisfy $x = x/2 + 1/x$ which implies that $x^2 = 2$ which we know cannot be satisfied by any rational number.

We seem to be able to make an intuitive distinction between sequences that do not converge at all (truly divergent), and those that seem to converge 'outside' the set — if we restrict ourselves to X however, it fails to converge because the metric space does not contain the 'limit'.

The same phenomenon can occur for other more general spaces e.g. a sequence of continuous functions can 'converge' to a discontinuous function.

It is possible to distinguish between the two ideas rigorously as follows:

Definition A **Cauchy sequence** is one such that

$$\forall \epsilon > 0 \quad \exists N \quad n, m > N \Rightarrow d(x_n, x_m) < \epsilon.$$

In other words, $\lim_{n, m \rightarrow \infty} d(x_n, x_m) = 0$. Intuitively speaking, a Cauchy sequence is one which is "trying" to converge. Convergent sequences are always Cauchy (see exercises above) but there are examples of Cauchy sequences which do not converge (e.g. the case of the rational sequence above, which tries to converge to $\sqrt{2}$).

Definition A metric space is **complete** when every Cauchy sequence converges.

It follows that the set of rational numbers \mathbb{Q} is not complete; but it will be shown that the set of real numbers \mathbb{R} is complete. This is the main reason why we use the real numbers rather than the rationals in most applications of calculus.

Note that every metric space X can be completed i.e., there is a complete metric space \tilde{X} which contains X ; for example $\tilde{\mathbb{Q}} = \mathbb{R}$ and $\widetilde{(0, 1)} = [0, 1]$.

Proposition 2.3

For a complete metric space X , any subset F is complete $\Leftrightarrow F$ is closed in X .

PROOF Let $F \subseteq X$ be complete i.e., any Cauchy sequence in F converges to a limit in F . Let x be a limit point of F . Then as shown in an exercise above, we can find a sequence (x_n) in F which converges to x , $x_n \rightarrow x$. But convergent sequences are Cauchy and F is complete. Hence (x_n) converges to a point in F i.e., $x \in F$. Thus F contains all its limit points and is therefore closed.

Conversely, let F be a closed set in X and let (x_n) be a Cauchy sequence in F . Then (x_n) is a Cauchy sequence in X , which is complete. Therefore $x_n \rightarrow x$ for some $x \in X$. But x must be a limit point of F and so $x \in F$. Thus any Cauchy sequence of F converges in F . □

Corollary

If X is a complete metric space, then a Cauchy sequence $x_n \in F$ converges to a limit in \bar{F} .

Theorem D

The set of real numbers \mathbb{R} is complete.

PROOF (From Real Analysis course) Let (x_n) be a Cauchy sequence, that is $|x_n - x_m| \rightarrow 0$ as $n, m \rightarrow \infty$. Set $\epsilon = 1$, then there must be an N_0 beyond which $|x_n - x_m| < 1$; let $y_0 := x_{N_0}$. Now set $\epsilon = 1/2$, then there must be an N_1 , with $N_1 > N_0$, beyond which $|x_n - x_m| < 1/2$; let $y_1 := x_{N_1}$. Continuing this way we get a sequence y_k for which $|y_{k+1} - y_k| < 1/2^k$.

Now consider the sum $y_0 - y_K = y_0 - y_1 + y_1 - y_2 + \dots - y_K$:

$$\left| \sum_{k=0}^K (y_k - y_{k+1}) \right| \leq \sum_{k=0}^K |y_k - y_{k+1}| \leq \sum_{k=0}^K 2^{-k} < 2.$$

The middle sum is increasing but bounded above, and so must converge. By comparison, the first sum must converge to, say, y ; set $x := y_0 - y$. Then $y_K \rightarrow y_0 - y = x$ as $K \rightarrow \infty$. That is, we have shown that the subsequence y_K converges to x ; it is left as an exercise to show that the full sequence x_n converges to x .

□

It follows that any closed subset of \mathbb{R} is complete. For example, any Cauchy sequence in $[0, 1]$ must converge to a limit in $[0, 1]$; and any Cauchy sequence in the Cantor set must converge there as well. However a Cauchy sequence of rational numbers need not converge to a rational number because \mathbb{Q} is not closed in \mathbb{R} .

Theorem E

**If X, Y are complete metric spaces
then so is $X \times Y$.**

PROOF Let $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ be a Cauchy sequence in $X \times Y$. This means that $d\left(\begin{pmatrix} x_n \\ y_n \end{pmatrix}, \begin{pmatrix} x_m \\ y_m \end{pmatrix}\right) = d_X(x_n, x_m) + d_Y(y_n, y_m) \rightarrow 0$ as $n, m \rightarrow \infty$. In particular $d(x_n, x_m) \rightarrow 0$, so that the sequence (x_n) is a Cauchy sequence in X , which is complete. We therefore get $x_n \rightarrow x \in X$, and similarly, $y_n \rightarrow y \in Y$. Now $d\left(\begin{pmatrix} x_n \\ y_n \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right) = d_X(x_n, x) + d_Y(y_n, y) \rightarrow 0$ as $n \rightarrow \infty$, so that the Cauchy sequence $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ converges in $X \times Y$.

□

Corollary

Hence $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$ and \mathbb{C} are complete.

2.3.1 Exercises

1. Show that the interval $(a, b]$ is incomplete, but $[a, b]$ is complete.
2. Show that if a Cauchy sequence (x_n) has a convergent *subsequence* $x_{n_i} \rightarrow x$ as $i \rightarrow \infty$, then (x_n) must converge to x .
3. A set A is called *bounded* when all the distances between points in A are bounded i.e., $d(x, y) \leq R$ for $x, y \in A$. Show that Cauchy sequences are bounded. Hence we get the following set inclusions: $\{\text{set of convergent sequences}\} \subseteq \{\text{set of Cauchy sequences}\} \subseteq \{\text{set of bounded sequences}\}$.
4. In the proof of the completeness of $X \times Y$ we used one particular distance $d_X + d_Y$. Show that the other distance $\max(d_X, d_Y)$ is also complete (if d_X and d_Y are complete).

3 Continuity

Definition A function $f : X \rightarrow Y$ is **continuous** when

$$\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad d_X(x, y) < \delta \Rightarrow d_Y(f(x), f(y)) < \epsilon$$

Since $d(a, b) < r \Leftrightarrow b \in B_r(a)$, this can be rewritten as

$$\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad y \in B_\delta(x) \Rightarrow f(y) \in B_\epsilon(f(x))$$

or even as

$$\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad fB_\delta(x) \subseteq B_\epsilon(f(x)).$$

Note also that a function is “discontinuous” at x if the opposite holds i.e.,

$$\exists \epsilon > 0 \quad \forall \delta > 0 \quad \exists y \in B_\delta(x) \quad f(y) \notin B_\epsilon(f(x)).$$

Proposition 3.1

**A function $f : X \rightarrow Y$ is continuous \Leftrightarrow
for every open set V in Y , $f^{-1}V$ is open in X .**

PROOF Let f be a continuous function and let V be an open set in Y . We want to show that $U := f^{-1}V$ is open in X . Let x be any point of U . Then $f(x) \in V$, which is open. Hence

$$f(x) \in B_\epsilon(f(x)) \subseteq V,$$

and so

$$\exists \delta > 0 \quad fB_\delta(x) \subseteq B_\epsilon(f(x)) \subseteq V.$$

In other words,

$$\exists \delta > 0 \quad B_\delta(x) \subseteq f^{-1}V = U.$$

Conversely, assume the second statement holds for all open sets V . In particular apply it to the open set $B_\epsilon(f(x))$. Then $f^{-1}B_\epsilon(f(x))$ is open in X , that is,

$$\exists \delta > 0 \quad x \in B_\delta(x) \subseteq f^{-1}B_\epsilon(f(x))$$

which implies

$$\exists \delta > 0 \quad fB_\delta(x) \subseteq B_\epsilon(f(x)),$$

as required. □

Note that if U is open in X , then fU need not be open in Y .

Proposition 3.2

If $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are continuous, then so is $g \circ f : X \rightarrow Z$.

PROOF Let W be any open set in Z . Then $g^{-1}W$ is an open set in Y , and so $f^{-1}g^{-1}W$ is an open set in X . But this set is precisely $(g \circ f)^{-1}W$. □

Note that f continuous need not imply that f^{-1} is continuous.

Theorem A

f is continuous \Leftrightarrow

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

PROOF Let f be a continuous function and let (x_n) be a sequence converging to x in the domain. We want to show that $f(x_n) \rightarrow f(x)$ in the codomain as $n \rightarrow \infty$. Consider the neighborhood $B_\epsilon(f(x))$ of $f(x)$. Then, since f is continuous,

$$\exists \delta > 0 \quad fB_\delta(x) \subseteq B_\epsilon(f(x)).$$

But $x_n \rightarrow x$ means

$$\exists N > 0 \quad n > N \Rightarrow x_n \in B_\delta(x)$$

which implies that

$$f(x_n) \in fB_\delta(x) \subseteq B_\epsilon(f(x)).$$

Conversely, suppose f is not continuous. Then there is a point x such that

$$\exists \epsilon > 0 \quad \forall \delta > 0 \quad fB_\delta(x) \not\subseteq B_\epsilon(f(x)).$$

In particular

$$\exists \epsilon > 0 \quad \forall n \quad fB_{1/n}(x) \not\subseteq B_\epsilon(f(x)).$$

Therefore we can find points $x_n \in B_{1/n}(x)$ for which $f(x_n) \notin B_\epsilon(f(x))$ i.e., $f(x_n) \not\rightarrow f(x)$ while $x_n \rightarrow x$. □

We thus see that continuous functions preserve *convergence*, a central concept in metric spaces; that is, if a sequence converges in X , and it is mapped continuously to Y , then it still converges. In a sense, they correspond to *homomorphisms* of groups and rings, which preserve the group and ring operations. The following concept is then analogous to an *isomorphism* between metric spaces.

Definition A **homeomorphism** between two metric spaces X and Y is a map $\phi : X \rightarrow Y$ such that

- ϕ is bijective (1-1 and onto)
- ϕ is continuous
- ϕ^{-1} is continuous

In other words, two spaces are homeomorphic when one can be obtained continuously and continuously invertibly from the other; in effect when convergence in one space is equivalent to convergence in the other. The most vivid example is of one space being continuously deformed into another homeomorphic to it; the classic ‘tongue-in-cheek’ example is that a ‘teacup’ is homeomorphic to a ‘doughnut’!

3.0.2 Exercises

1. Show that the constant functions $f(x) := y_0 \in Y$ are continuous.
2. Find a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ and an open set $U \subseteq \mathbb{R}$ such that fU is not open.

3. Let

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

When is χ_A continuous?

4. Show that if F is a closed set in Y and $f : X \rightarrow Y$ is a continuous function, then $f^{-1}F$ is closed in X .
5. Show that if $f, g : X \rightarrow \mathbb{R}$ are continuous functions, then so are the functions $f + g$ and λf .
6. Show that $\{x \in X : f(x) > 0\}$ is open in X when $f : X \rightarrow \mathbb{R}$ is a continuous function.
7. Reprove the proposition that $g \circ f$ is continuous using distances rather than open sets.
8. Show that if $x_n \rightarrow x$ and $y_n \rightarrow y$ then $d(x_n, y) \rightarrow d(x, y)$ and $d(x_n, y_n) \rightarrow d(x, y)$.
9. Show that continuous functions need not map Cauchy sequences to Cauchy sequences.
10. Find an example where f is continuous and bijective but f^{-1} is not.
11. Show that “ X is homeomorphic to Y ” is an equivalence relation on metric spaces, so that we can partition metric spaces into equivalence classes.
12. Show that $(0, 1)$ is homeomorphic to (a, b) .
13. Let $f : \mathbb{Q} \rightarrow \mathbb{Q}$ be a continuous function. Show that there is only one way of extending it to the reals as a continuous function $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$. Show further that the only function $f : \mathbb{Q} \rightarrow \mathbb{Q}$ satisfying $f(x + y) = f(x) + f(y)$ is $f(x) = \lambda x$. Deduce that there is only one continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ with this property.
14. * Let $f : A \rightarrow Y$ and $g : B \rightarrow Y$ be continuous functions with $A, B \subseteq X$. Suppose that the two functions agree on $A \cap B$, i.e., $x \in A \cap B \Rightarrow f(x) = g(x)$. Show that the function

$$h(x) = \begin{cases} f(x) & x \in A \\ g(x) & x \in B \end{cases}$$

is well-defined and continuous on $A \cup B$.

15. * A function $f : X \rightarrow Y$ is called *isometric* when f preserves distances i.e.,

$$d_Y(f(x), f(y)) = d_X(x, y) \quad \forall x, y \in X.$$

Show that onto isometric maps are continuous, invertible and have a continuous inverse (i.e., they are homeomorphisms).

3.1 Applications

Definition A function is called a **contraction** when there is a constant $0 \leq k < 1$ such that

$$\forall x, y \in X, d(f(x), f(y)) \leq k d(x, y).$$

It follows that f is continuous, because

$$d(x, y) < \delta := \epsilon/k \Rightarrow d(f(x), f(y)) < \epsilon.$$

Theorem B

The Banach contraction mapping theorem

Let X be a complete metric space, and suppose that $f : X \rightarrow X$ is a contraction map. Then f has a unique fixed point $x = f(x)$.

PROOF Consider the iteration $x_{n+1} = f(x_n)$ with $x_0 = a$ any point in X . Note that $d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq k d(x_n, x_{n-1})$; hence, by induction, $d(x_{n+1}, x_n) \leq k^n d(x_1, x_0)$.

(x_n) is a Cauchy sequence since

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + \dots + d(x_{m+1}, x_m) \\ &\leq (k^{n-1} + \dots + k^m) d(x_1, x_0) \leq \frac{k^m}{1-k} d(x_1, x_0) \end{aligned}$$

which converges to 0 as $m, n \rightarrow \infty$. Hence $x_n \rightarrow x$, and by continuity of f ,

$$x = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} f(x_n) = f(\lim_{n \rightarrow \infty} x_n) = f(x).$$

Moreover, the rate of convergence is given by $d(x_n, x) \leq \frac{k^n}{1-k} d(x, x_0)$.

Suppose there are two fixed points $x = f(x)$ and $y = f(y)$; then

$$d(x, y) = d(f(x), f(y)) \leq kd(x, y)$$

so that $d(x, y) = 0$ since $k < 1$.

□

4 Connectedness

Definition A set A is **disconnected** when it can be divided into (at least) two disjoint non-empty subsets $A = B \cup C$ and each subset can be covered exclusively by an open set (U, V respectively) i.e.,

$$B \subseteq U, \quad C \cap U = \emptyset,$$

$$C \subseteq V, \quad B \cap V = \emptyset.$$

A set is called **connected** otherwise.

4.0.1 Example

Two distinct points are disconnected by Theorem 1D.

Single points are always connected because they cannot be separated into two non-empty sets. Similarly the empty set is connected.

Any subset of the natural numbers is disconnected except the single points $\{n\}$ and the empty set.

The set of rational numbers \mathbb{Q} is disconnected e.g. $\mathbb{Q} \subseteq (-\infty, \sqrt{2}) \cup (\sqrt{2}, \infty)$.

Theorem A

The connected subsets of \mathbb{R} are precisely its intervals.

PROOF Let A be a connected subset of \mathbb{R} . Let $a, b \in A$ and let $a < x < b$. Then if $x \notin A$, it follows that $(-\infty, x) \cup (x, \infty)$ is a non-trivial open partition of A , which is a contradiction for a connected set. Hence $x \in A$, which means that A is an interval.

For the converse, first consider an interval $[a, b]$. Suppose it to be disconnected. Then it can be split up into disjoint subsets B and C , each covered exclusively by U and V respectively. Suppose b belongs to $B \subseteq U$ say. Consider the open set $V \cap [a, b]$ (open in $[a, b]$). Then this set is bounded above by b and is non-empty (because it contains the non-empty set C). Hence by the least upperbound property of the real numbers, $V \cap [a, b]$ has a least upperbound $\alpha \leq b$. Now α ought to belong to either U or V . If it belongs to U then we should find an ϵ -ball around it contained completely in U ; but this would make $\alpha - \epsilon$ a smaller upperbound. If instead it belongs to V , then

we can find an ϵ -ball around it contained completely in V ; but then there would be points of V larger than α . As both of these are contradictory, we conclude that the set $[a, b]$ is connected.

Now consider any other interval. The intervals $\emptyset = (a, a)$ and $\{a\} = [a, a]$ are always connected. So suppose the interval contains at least two points and is disconnected. Then we can find two points contained in separating open sets, $a \in U$ and $b \in V$. But this means that the set $[a, b]$ is disconnected using the same open sets U and V , which is false. Hence every interval is connected. □

Theorem B

If $f : X \rightarrow Y$ is continuous and $A \subseteq X$ is a connected set, then fA is a connected subset of Y .

PROOF Suppose fA is disconnected into the non-empty disjoint sets B and C covered exclusively by the open sets U and V . That is,

$$fA = B \cup C \subseteq U \cup V; \quad U \cap C = V \cap B = \emptyset.$$

Then,

$$A = f^{-1}B \cup f^{-1}C \subseteq f^{-1}U \cup f^{-1}V; \quad f^{-1}U \cap f^{-1}C = f^{-1}V \cap f^{-1}B = \emptyset.$$

Moreover $f^{-1}B$ and $f^{-1}C$ are disjoint and $f^{-1}U$ and $f^{-1}V$ are open sets. Hence fA disconnected $\Rightarrow A$ disconnected. □

This means that the property of connectedness is preserved by continuous maps. In particular a **path** which is a continuous function $\gamma : [0, 1] \rightarrow X$ has a connected image. For example, straight line segments, circles and parabolas in \mathbb{R}^2 are connected.

Corollary

The Intermediate Value Theorem.

If $f : [a, b] \rightarrow \mathbb{R}$ is continuous and $f(a) < c < f(b)$ then there exists an x such that $a \leq x \leq b$ and $f(x) = c$.

PROOF We have proved that $[a, b]$ is a connected set and hence $f[a, b]$ is also connected in \mathbb{R} . But connected subsets of \mathbb{R} are just the intervals. Now $f(a), f(b) \in f[a, b]$ and so $c \in f[a, b]$ i.e., $c = f(x)$ for some $x \in [a, b]$. \square

Similarly a continuous function $f : [0, 1]^2 \rightarrow \mathbb{R}$ has the same property: if $f(\mathbf{a}) < c < f(\mathbf{b})$ then there is an $\mathbf{x} \in [0, 1]^2$ such that $c = f(\mathbf{x})$. By the above theorem, this remains true if we substitute $[0, 1]^2$ by any connected subset of \mathbb{R}^2 .

Proposition 4.1

If A and B are connected sets and $A \cap B \neq \emptyset$ then $A \cup B$ is connected.

PROOF Suppose $A \cup B$ splits up into two parts covered exclusively by open sets U and V . Then A itself would split up into the two parts $A \cap U$ and $A \cap V$ were these non-empty. But A is known to be connected hence one of these must be empty, say $A \cap U = \emptyset$. Similarly for B , $B \cap V = \emptyset$. In other words, $A \subseteq V$ and $B \subseteq U$. Hence $A \cap B \subseteq A \cap U = \emptyset$. \square

4.0.2 Exercises

1. Show that the only connected subsets of \mathbb{Q} are the single points or the empty set.
2. Let y be a positive real number. Use the intermediate value theorem to show that $\sqrt[y]{y}$ exists.
3. Generalize the last proposition to the case when there are an infinite number of connected sets A_i with $\bigcap_i A_i \neq \emptyset$.
4. Show that \mathbb{R}^2 is connected by considering the radial lines and applying the previous exercise. More generally show that any vector space over the real numbers is connected.
5. Similarly show that $(a, \infty) \times \mathbb{R}$ is connected. Show further that $\mathbb{R}^2 \setminus \{\mathbf{x}\}$ is connected but that $\mathbb{R} \setminus \{x\}$ is not. Deduce that \mathbb{R} and \mathbb{R}^2 are not homeomorphic.
6. Use the same idea to show that $[a, b]$ is not homeomorphic to $[a, b)$.

7. Show that if a set contains a non-trivial proper subset which is both open and closed then it is disconnected. Equivalently, a connected set cannot contain a non-trivial open and closed proper subset.
8. Do balls have to be connected? Consider the space $X = (-\infty, -1) \cup (1, \infty)$. Find a ball in this space which is not connected.
9. Let A be a connected metric space and let $f : A \rightarrow \mathbb{N}$ be a continuous function. Show that f must be a constant function.
10. Suppose that there is an $x_0 \in X$ and an $r > 0$ such that $d(x_0, y) \neq r \forall y$. Moreover there is another point x_1 such that $d(x_0, x_1) > r$. Show that X must be disconnected.
11. * Show that \overline{A} is connected if A is.

5 Compactness

Definition A set B is **bounded** when

$$\exists R > 0 \quad \forall x, y \in B \quad d(x, y) < R.$$

That is, a set is bounded when the distances between any two points in the set have an upperbound. The least such upperbound is called the **diameter** of the set:

$$\text{diam}B = \sup\{d(x, y) : x, y \in B\}.$$

5.0.3 Example.

The set $[a, b)$ is bounded in \mathbb{R} because $d(x, y) \leq b - a$ for any two elements $x, y \in [a, b)$; in fact $\text{diam}[a, b) = b - a$. The set \mathbb{R} is unbounded because $d(0, x) = |x|$ can be made as large as needed.

Note: Boundedness is a property of how *large* a set is. However it is not a good “metric” property: if B is bounded and f is a continuous function, then fB need not be bounded (see Ex. 4). Thus a set may be bounded in one metric space X , but not remain so when X is deformed continuously to another.

Proposition 5.1

A set B is bounded when it can be covered by a ball,

$$\exists R > 0, a \in X, \quad B \subseteq B_R(a).$$

Proof. (Left as an exercise)

Hence, the set $[0, 1) \cup (2, 3)$ is bounded because it can be covered by the ball $B_4(0)$.

Definition A set B is **totally bounded** when

$$\forall \epsilon > 0 \quad \exists a_1 \dots a_N \quad B \subseteq \bigcup_{i=1}^N B_\epsilon(a_i).$$

That is, a set is totally bounded when it can be covered by a finite number of ϵ -balls, however small their radii ϵ .

5.0.4 Example.

The set $[0, 1]$ is totally bounded because it can be covered by the balls $B_\epsilon(n\epsilon)$ for $n = 0, \dots, N$ where $N > 1/\epsilon$.

Proposition 5.2

A totally bounded set is bounded.

PROOF Let B be a totally bounded set. We can therefore cover B with a finite number of ϵ -balls, in particular with 1-balls.

$$B \subseteq \bigcup_{i=1}^N B_1(a_i).$$

Since there are only a finite number of these balls, we can find the maximum distance between their centers.

$$D := \max\{d(a_i, a_j)\}.$$

Now, given any two points $x, y \in B$, they must be covered by two of these balls $B_1(a_I)$ and $B_1(a_J)$, say. Therefore, using the triangle inequality twice,

$$d(x, y) \leq d(x, a_I) + d(a_I, a_J) + d(a_J, y) \leq 1 + D + 1.$$

That is, $D + 2$ is an upperbound for the distances between points in B . □

In one of the exercises, you are asked to show that, in \mathbb{R} , bounded sets are totally bounded (i.e., boundedness and totally boundedness coincide for the real line). So a totally bounded set also need not be preserved by a continuous function.

Definition A set K is said to be **compactif**

$$K \subseteq \bigcup_i B_{\epsilon_i}(a_i) \quad \Rightarrow \quad \exists i = i_1, \dots, i_N : \quad K \subseteq \bigcup_{k=1}^N B_{\epsilon_{i_k}}(a_{i_k}).$$

In plain English, a set is compact when given any cover of balls of varying sizes, we can find a finite subcollection of them that still cover the set. This is a stronger property than totally-boundedness.

5.0.5 Example.

The set $[0, 1)$ is *not* compact. For example, the cover of balls $B_{1-1/n}(0)$ for $n = 2, \dots$ has no finite subcover. Similarly the sets \mathbb{R} and \mathbb{N} are not compact. On the other hand we will soon see that the sets $[a, b]$ are compact in \mathbb{R} .

By an *open cover* we mean a cover of open sets, and a *subcover* is a subcollection of a cover.

Theorem A

A set is compact \Leftrightarrow any open cover of it has a finite subcover.

PROOF Let K be the compact set and let the open sets A_i cover it i.e., $K \subseteq \bigcup_i A_i$. Now, by theorem 1B, the open sets A_i are composed of balls. Therefore it follows that K is inside a union of (a large number of) balls, i.e., it is covered by these balls. Therefore there is a finite number of these balls $B_{\epsilon_1}(a_1), \dots, B_{\epsilon_N}(a_N)$ which still cover the set K . Each of these balls is inside one of the open sets A_i . Therefore there is a finite number of these open sets (those which contain these balls) that still cover the compact set K .

Conversely, suppose K is such that any open cover of it has a finite subcover. In particular if we can cover it with (open) balls, then we can find a finite subcollection of them which still cover K .

□

We show next that “compactness” is a metric property: it is preserved by continuous functions.

Theorem B

If K is compact and f is a continuous function, then fK is compact.

PROOF Let A_i be an open cover for fK . We would like to show that a finite subcollection of them still covers fK .

From

$$fK \subseteq \bigcup_i A_i$$

we can deduce

$$K \subseteq f^{-1} \bigcup_i A_i = \bigcup_i f^{-1} A_i.$$

But $f^{-1}A_i$ are open sets since f is continuous (Prop. 3.0.1). Therefore the right-hand side is an open cover of K , which is compact. Therefore, by Theorem 5A, a finite number of these open sets will do to cover K .

$$K \subseteq \bigcup_{i=1}^N f^{-1} A_i.$$

It follows that

$$fK \subseteq \bigcup_{i=1}^N A_i.$$

That is, a finite number of the original open sets A_i will cover the set fK , which is therefore compact.

□

Theorem C

Compact sets are closed and bounded.

PROOF *Compact sets are bounded.* Let K be a compact set. For any fixed point $a \in X$, consider the balls $B_n(a)$ for $n = 1, 2, \dots$. These balls cover the set K (in fact they cover the whole space X). Therefore a finite number of these balls will cover K . But these balls are nested in each other, so that it follows that the largest of these finite number of balls, say $B_N(a)$, contains all the others. Hence $B_N(a)$ contains K , which is therefore bounded.

Compact sets are closed in X . Let $x \in X \setminus K$. We would like to surround it by a ball outside K . From Theorem 1D, x can be separated from $y \in K$ by open balls $B_{r_y}(x)$ and $B_{r_y}(y)$. Since $y \in B_{r_y}(y)$, these latter balls cover K . But K is compact, so that a finite subcollection of these balls still cover K ,

$$K \subseteq B_{r_1}(y_1) \cup \dots \cup B_{r_N}(y_N).$$

Now let $U := B_{r_1}(x) \cap \dots \cap B_{r_N}(x)$. Since this involves a *finite* intersection of open sets, then U is also an open set. Moreover $U \cap K = \emptyset$ since

$$\begin{aligned} z \in U &\Rightarrow z \in B_{r_i}(x) \text{ for } i = 1, \dots, N \\ &\Rightarrow z \notin B_{r_1}(y_1) \cup \dots \cup B_{r_N}(y_N) \supseteq K. \end{aligned}$$

Therefore,

$$x \in U \subseteq X \setminus K.$$

□

Theorem D

A closed subset of a compact set is itself compact.

PROOF Let F be a closed subset of a compact set K . Let the open sets A_i cover F . Then

$$K \subseteq F \cup (X \setminus F) \subseteq \bigcup_i A_i \cup (X \setminus F).$$

The right-hand side is the union of open sets since $X \setminus F$ is open when F is closed. But K is compact and therefore a finite number of these open sets will do to cover it i.e.,

$$K \subseteq \bigcup_{i=1}^N A_i \cup (X \setminus F).$$

Hence

$$F \subseteq \bigcup_{i=1}^N A_i,$$

since $X \setminus F$ is outside F and doesn't cover it.

□

5.0.6 Exercises

1. Show that B is bounded if and only if $\exists R > 0, a \in X, B \subseteq B_R(a)$.
2. Show that any subset of a bounded set is bounded; the union of two (or finite number of) bounded sets is bounded; repeat the exercise with totally bounded sets.

3. Show that a Cauchy sequence is bounded. Deduce that an unbounded sequence cannot converge.
4. Find an example of a bounded set $B \subseteq \mathbb{R}$ and a continuous function $f : B \rightarrow \mathbb{R}$ such that fB is not bounded.
5. Show that if $\text{diam}A \leq R$ and $A \cap B_R(a) \neq \emptyset$ then $A \subseteq B_{2R}(a)$.
6. Show that, in \mathbb{R} , a bounded set is totally bounded. (Hint: first show that a ball is totally bounded).
7. Theorem 5C shows that a compact set is bounded. Show further that a compact set is totally bounded. (Hint: consider ϵ -balls centred about all the points of the compact set).
8. Show that \mathbb{R} is not compact by finding an open cover that does not have a finite subcover.
9. Show that $[0, 1]$ is not compact in \mathbb{Q} by finding an infinite open cover that does not have a finite subcover. (Hint: you need to consider an irrational number in between 0 and 1).
10. Let X be a space with the discrete metric of Ex. 1.1.3.9. Show that a subset K is compact $\Leftrightarrow K$ consists of a finite number of points.
11. Show that compact sets behave like closed sets i.e., any intersection and any finite union of compact sets are compact. (Hint: use theorem 1D)
12. (Cantor) Let K_n be a decreasing sequence of nonempty compact sets. Show that if $\bigcap_n K_n = \emptyset$ then $X \setminus K_n$ form an open cover of K_1 . Deduce that $\bigcap_n K_n$ is compact and nonempty. Show moreover that if $\text{diam}K_n \rightarrow 0$ then $\bigcap_n K_n$ consists of a single point.
13. * Show that for a metric space with the distance function D_2 of Ex.1.0.2.3, any set is bounded. Show that not all sets are totally bounded and deduce that a bounded set need not be totally bounded.
14. * Let K precompact mean that \overline{K} is compact. Show that any subset of a precompact set is precompact, and that continuous images of precompact sets are precompact.

5.1 Compactness in \mathbb{R}

Theorem E

(Heine-Borel)

The closed interval $[a, b]$ is compact in \mathbb{R} .

PROOF Let $[a, b] \subseteq \bigcup_i A_i$ be an open cover of the set. Suppose there is no finite subcover, and we seek to obtain a contradiction. Divide the set $[a, b]$ in two $[a, (a+b)/2]$ and $[(a+b)/2, b]$. One of the two (and possibly both) does not admit a finite subcover: call it $[a_1, b_1]$. Repeat this process of dividing to get nested intervals $[a_n, b_n]$ of length $(b-a)/2^n$ which do not admit finite subcovers.

Now (a_n) is an increasing real sequence bounded above by b , while (b_n) is decreasing bounded below by a . They therefore both converge, say, $a_n \rightarrow x$ and $b_n \rightarrow y$ with $x \leq y$ since $a_n \leq b_n$. In fact $x = y$ since

$$a_n \leq x \leq y \leq b_n \text{ and } b_n - a_n \rightarrow 0.$$

This limit x is in the set $[a, b]$ and is therefore covered by some open set A_{i_0} . This implies that we can surround x with an ϵ -ball

$$x \in B_\epsilon(x) \subseteq A_{i_0}.$$

But $a_n \rightarrow x$ implies that there is an N such that $a_N \in B_\epsilon(x)$; similarly for b_n , so that $[a_N, b_N] \subseteq B_\epsilon(x) \subseteq A_{i_0}$. But $[a_N, b_N]$ was not supposed to be covered by a finite number of A_i 's. This implies that there must be a finite subcover to start with. □

Exercise: Generalize this to closed rectangles $[a, b] \times [c, d]$ in \mathbb{R}^2 , by repeatedly dividing it into four sub-rectangles and adapting the same argument. Can you extend this further to closed cuboids in \mathbb{R}^3 ?

Corollary

In \mathbb{R} , a set is compact \Leftrightarrow it is closed and bounded.

PROOF That a compact set is closed and bounded is shown in Theorem 5C. Conversely, let $F \subseteq \mathbb{R}$ be a closed and bounded set. By its boundedness $F \subseteq B_R(a) \subseteq [a, b]$. Hence F is a closed subset of a compact set and by 5C this implies that it is itself compact. \square

Exercise: Generalize this to show that any closed and bounded subset of \mathbb{R}^2 , \mathbb{R}^3 etc. is compact.

Corollary

Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function. Then its image $f[a, b] := \{f(x) : a \leq x \leq b\}$ is bounded, and the function attains its bounds, i.e.,

$$\exists c \in [a, b] \quad \forall x \in [a, b] \quad f(x) \leq f(c).$$

PROOF $[a, b]$ is a compact set. Its continuous image $f[a, b]$ is therefore also compact, and by 5C is bounded i.e., $f[a, b] \subseteq B_R(0)$ i.e., $|f(x)| \leq R$. Moreover compact sets are closed and so contain their limit points. In particular it contains $\sup f[a, b]$ i.e., $\sup f[a, b] = f(c)$ for some $c \in [a, b]$. \square

More generally, a continuous function $f : K \rightarrow \mathbb{R}$ on a compact set will be bounded and attain its bounds.

5.2 Compactness in a general Metric Space

We have seen that in the space of real numbers, a subset is compact if and only if it is closed and bounded. But this is not true for every metric space. Instead, we can prove the following:

Theorem F

A set is compact \Leftrightarrow it is complete and totally bounded.

PROOF The proof will be divided into several parts:

Every infinite subset of a compact set K has a limit point in K :

Let A be a subset of a compact set K , and suppose it has no limit points in K . This means that any $x \in K$ is not a limit point, and so we can find balls $B_\epsilon(x)$ which do not contain elements of A unless they are at the center. In particular, this is an open cover of K .

But K is compact and so must have a finite subcover of these,

$$K \subseteq B_{\epsilon_1}(x_1) \cup \dots \cup B_{\epsilon_N}(x_N).$$

$$\therefore A \cap K \subseteq A \cap \bigcup_{i=1}^N B_{\epsilon_i}(x_i) = \bigcup_{i=1}^N A \cap B_{\epsilon_i}(x_i) \subseteq \{x_1, \dots, x_N\}.$$

So A is a finite set. Equivalently, infinite subsets of K must have at least one limit point in K .

The property that every infinite subset of a set has a limit point is termed the *Bolzano-Weierstraß* property (BW-compact for short). So we have shown here that compact sets are BW-compact.

K BW-compact $\Rightarrow K$ complete:

Let (a_n) be a Cauchy sequence in a given BW-compact set K . If the set $\{a_n\}$ is infinite, then it has a limit point x in K . Now, from exercises 2.2.1.1 and 2.3.1.2, we can deduce that there is a subsequence of (a_n) which converges to x , and, since (a_n) is Cauchy, that the whole sequence (a_n) converges to x .

Otherwise, the set $\{a_n\}$ is finite, and the sequence (a_n) must repeat itself. The fact that the sequence is Cauchy implies that the sequence must eventually be constant $a_n = a_N$ for $n \geq N$, so that it converges to a_N in K .

K compact $\Rightarrow K$ totally bounded:

This is trivially true, since for any $\epsilon > 0$, each $x \in K$ belongs to $B_\epsilon(x)$. Hence, $K \subseteq \bigcup_{x \in K} B_\epsilon(x)$, which reduces to a finite subcover since K is compact i.e., $K \subseteq \bigcup_{i=1}^N B_\epsilon(x_i)$, making K totally bounded.

K complete and totally bounded $\Rightarrow K$ compact:

This part of the proof is really a generalization of the Heine-Borel theorem.

Let K be a complete and totally bounded set. Suppose K is covered by open sets V_i , and that no finite number of these open sets is enough to cover K . Since K is totally bounded, we can find a finite number of balls of radius 1 which cover K ,

$$K \subseteq \bigcup_{i=1}^N B_1(y_i).$$

If each of these balls had a finite subcover from the open cover $\{V_i\}$, then so would K . So at least one of these balls cannot be covered by a finite number of V_i 's; let us call this ball $B_1(x_1)$.

Now consider $B_1(x_1) \cap K$, also totally bounded. We can again cover it with a finite number of balls of radius $1/2$, one of which does not have a finite subcover, say $B_{1/2}(x_2)$. Repeat this process to get a nested sequence of balls $B_{1/2^n}(x_n)$, none of which have a finite subcover. The sequence (x_n) is Cauchy since $d(x_n, x_m) < 1/2^n$ (for $m > n$), and K is complete, hence $x_n \rightarrow x$ in K .

But x is covered by some open set V_i . Hence

$$x \in B_\epsilon(x) \subseteq V_i.$$

Moreover since $1/2^n \rightarrow 0$ and $x_n \rightarrow x$, we can find an N such that $1/2^N < \epsilon/2$ and $d(x_N, x) < \epsilon/2$, so that

$$B_{1/2^N}(x_N) \subseteq B_\epsilon(x) \subseteq V_i,$$

which contradicts the way that the balls $B_{1/2^n}(x_n)$ were chosen. □

Note that complete subsets are always closed, and totally bounded sets are always bounded. Hence the statement K compact $\Rightarrow K$ closed and bounded is really just a special case of this theorem. In general, however, closed and bounded sets are not complete and totally bounded, although we have seen that this is true in \mathbb{R}^N .

Theorem G

K is compact \Leftrightarrow every infinite subset of K has a limit point in K .

PROOF One implication has already been proved. Moreover, we have also shown that BW-compactness implies completeness, so we only need to prove that BW-compact sets are totally bounded to show the converse.

Let K be a (non-empty) BW-compact set. Let $\epsilon > 0$ and let a_1 be any point in K .

Either $B_\epsilon(a_1)$ covers K (and we're done i.e., K would be bounded) or if not, there is some point a_2 with $d(a_2, a_1) \geq \epsilon$. Repeating the argument

for a_2 , either $B_\epsilon(a_1) \cup B_\epsilon(a_2)$ cover K or else they omit some point, say a_3 . Continuing like this we get a sequence of points a_n each covered by $B_\epsilon(a_n)$ satisfying $d(a_n, a_m) \geq \epsilon$.

Suppose this process does not stop and we end up with an infinite set $\{a_n\}$. This has a limit point x since K is BW-compact, which implies that some of these a_n are close to x within $\epsilon/2$. But by the choice of the a_n , $d(a_n, a_m) \geq \epsilon$. This leads to a contradiction unless the process actually stops in a finite number of steps and the $B_\epsilon(a_n)$ cover the whole of K . □

5.2.1 Exercises

1. Show that a finite set of points is BW-compact.
2. Show directly that a BW-compact set is closed.
3. Show that $[a, b]$ is *not* compact in \mathbb{Q} by finding an infinite set of rational numbers in $[a, b]$ that do not have a limit point.
4. Let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a continuous curve. Show that there is a maximum and a minimum distance of points on the curve to the origin. Give an example to show that this is false if $[0, 1]$ is replaced by $(0, 1)$.
5. Show that the Bolzano-Weierstraß property is equivalent to the property that every sequence has a convergent subsequence.
6. Show directly that a continuous function maps a BW-compact set to a BW-compact set.
7. Show that if X, Y are compact metric spaces then so is $X \times Y$. (Hint: show that $X \times Y$ is BW-compact.)
8. Let X be a complete metric space. Show that a subset K is compact \Leftrightarrow it is closed and totally bounded. Deduce that for a complete metric space, a set is precompact (defined in 5.13.14) if and only if it is totally bounded.
9. * Let X be a totally bounded space. Consider finite covers of $\frac{1}{n}$ -balls and let A be the set of all the centers of all these balls. Show that A is countable and $\overline{A} = X$.

A Answers

Ex. 1.0.1

(7)

$$\begin{aligned} 0 &= d(x, x) \leq d(x, z) + d(x, z) = 2d(x, z) \\ d(x, y) &\leq d(x, x) + d(y, x) = d(y, x) \leq d(x, y) \end{aligned}$$

Ex. 1.1.3

(4) There are points in \mathbb{R}^2 , such as (x, ϵ) ($a < x < b$), which are arbitrarily close to points in the interval $(x, 0)$.

(5) For example, $\bigcap_n (-1/n, 1/n) = \{0\}$.

Ex. 2.1.2

(5) For example, (i) $[0, 1[$ and (ii) \mathbb{R} .

(6) For example, (i) $\bigcup_{n=1}^{\infty} [0, 1 - 1/n] = [0, 1)$, (ii) $\bigcup_{n=0}^{\infty} \{n\} = \mathbb{N}$.

(7) (i) open, (ii) closed, (iii) neither, (iv) both, (v) neither, (vi) closed.

Ex. 2.1.3

(1) $\overline{\mathbb{Z}} = \mathbb{Z}$, $\overline{\mathbb{Q}} = \mathbb{R}$.

(3) Yes to both, e.g. \mathbb{N} in \mathbb{R} .

Ex. 3.0.2

(2) For example, $f(x) := x^2$ with $U = (-1, 1)$, then $fU = [0, 1)$.

(3) For \mathbb{R} (or \mathbb{R}^N), never. But on “disconnected” spaces, such as \mathbb{Z} , χ_A may be continuous (e.g. $A = \{n\}$).

(10) For example, take $X := \mathbb{R} \setminus (0, 1]$ and $f(x) := \begin{cases} x & x \leq 0 \\ x - 1 & x > 1 \end{cases}$; then

$f : X \rightarrow \mathbb{R}$ is bijective and continuous but f^{-1} is discontinuous at 0.

Ex. 4.0.2

(8) For example, $B_4(2)$.

Ex. 5.0.6

(4) For example, $f(x) := 1/x$ maps $(0, 1) \rightarrow (1, \infty)$.

Ex. 5.2.1

(4) For example, any spiral curve, such as $\gamma(x) := e^{-1/x}(\cos(1/x), \sin(1/x))$, approaches but never reaches the origin, so has no minimum distance.

B Summary

B.1 Definitions

1. A **distance** (or **metric**) on a metric space X is a function

$$d : X^2 \mapsto \mathbb{R}^+$$

$$(x, y) \mapsto d(x, y)$$

with the properties

$$\begin{aligned} i \quad & d(x, y) = 0 \Leftrightarrow x = y \\ ii \quad & d(y, x) = d(x, y) \\ iii \quad & d(x, y) \leq d(x, z) + d(y, z) \end{aligned}$$

2. A **ball** is a set $B_r(a) = \{x \in X : d(x, a) < r\}$.
3. An **open** set, A , in X is a set such that

$$\forall x \in A \quad \exists \epsilon > 0 \quad B_\epsilon(x) \subseteq A.$$

Every point of A can be surrounded by a ball inside A .

4. A **closed** set, F , in X is a set such that $X - F$ is open in X .
Every point not in F can be surrounded by a ball outside F .
5. A **limit point** of a set A is a point b such that

$$\forall \epsilon > 0 \quad \exists a \neq b \quad a \in A \cap B_\epsilon(b)$$

Every ball around b contains points of A apart from b .

6. The **closure** of a set A is the set $\bar{A} = A \cup \{\text{limit points of } A\}$.
7. A sequence (x_n) **converges** to x , $x_n \rightarrow x$ means

$$\forall \epsilon > 0 \quad \exists N \quad n > N \Rightarrow x_n \in B_\epsilon(x).$$

The sequence of points eventually comes arbitrarily close to the limit.

8. A **Cauchy sequence** is one such that

$$\forall \epsilon > 0 \quad \exists N \quad n, m > N \Rightarrow d(x_n, x_m) < \epsilon.$$

Eventually the points get arbitrarily close to each other.

9. A **complete** metric space is one in which every Cauchy sequence converges.
10. A **continuous** function $f : X \rightarrow Y$ is one for which

$$\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad y \in B_\delta(x) \Rightarrow f(y) \in B_\epsilon(f(x))$$

Points sufficiently close to x are mapped to points arbitrarily close to $f(x)$.

11. Two metric spaces X and Y are **homeomorphic** when there is a map $\phi : X \rightarrow Y$ which is bijective (1-1 and onto), continuous, and its inverse is continuous.
12. A **connected** set is one such that

$$A \subseteq U \cup V \text{ with } U, V \text{ open sets} \Rightarrow A \cap U \cap V \neq \emptyset \text{ or } A \cap U = \emptyset \text{ or } A \cap V = \emptyset.$$

It cannot be separated by two non-trivial open sets.

13. A **bounded** set B is one for which $\exists R > 0 \quad \forall x, y \in B \quad d(x, y) < R$.
- It can be covered by a single ball.*
14. A **totally bounded** set B is one for which

$$\forall \epsilon > 0 \quad \exists a_1 \dots a_N \quad B \subseteq \bigcup_{i=1}^N B_\epsilon(a_i).$$

It can be covered by a finite number of ϵ -balls.

15. A **compact** set K is one for which

$$K \subseteq \bigcup_i A_i \quad \Rightarrow \quad \exists i = i_1, \dots, i_N : \quad K \subseteq \bigcup_{k=1}^N A_{i_k}.$$

Any cover of open sets (balls) has a finite subcover.

16. A **BW-compact** set K is one for which every infinite subset of K has a limit point in K .

B.2 Main Theorems

1. Balls are open sets.
2. Any union and any finite intersection of open sets is open.
3. A set is closed when it contains its limit points.
4. A Cauchy sequence in a set A converges to a limit in \overline{A} , for a complete metric space; closed sets of complete spaces are complete.
5. $f : X \rightarrow Y$ continuous and V open in $Y \Rightarrow f^{-1}V$ open in X .
6. A function f is continuous $\Leftrightarrow f(\lim_{n \rightarrow \infty} x_n) = \lim_{n \rightarrow \infty} f(x_n)$.
7. The connected sets of \mathbb{R} are the intervals.
8. A continuous function maps a connected set to a connected set.
9. The union of non-disjoint connected sets is connected.
10. Compact sets are closed and bounded.
11. A continuous function maps a compact set to a compact set.
12. (Heine-Borel theorem) The compact sets of \mathbb{R}^n are the closed and bounded sets. (In particular $[a, b]$ is compact.)
13. A set is compact \Leftrightarrow it is BW-compact \Leftrightarrow it is complete and totally bounded.

C Miscellaneous Exercises

Each question should take less than 3/4 hours to answer.

1. (a) Show that a continuous function maps a compact set to a compact set.
(b) Prove that the set $[0, 1]^2$ is compact.
(c) Deduce that if $f : [0, 1]^2 \rightarrow \mathbb{R}$ is continuous then there is a maximum point x :

$$\exists x \in [0, 1]^2, \quad \forall y \in [0, 1]^2 \quad f(y) \leq f(x).$$

2. (a) Define a *connected* set.
(b) Show that the circle S (subset of \mathbb{R}^2) is connected, but that its complement S' is not.
(c) Show that the only connected sets in \mathbb{N} are the single points $\{n\}$.
(d) Deduce that if $f : S \rightarrow \mathbb{N}$ is a continuous function then it must be constant.
3. (a) Let $K_n := [a_n, b_n]$ be closed and bounded intervals in \mathbb{R} satisfying $K_{n+1} \subseteq K_n$ and $b_n - a_n \rightarrow 0$. Show that

$$\bigcap_n K_n = \{x\}$$

and that x is a limit point of $\{a_n\}$, unless the sequence (a_n) is eventually constant.

- (b) Give an example to show that this intersection can be empty in \mathbb{Q} .