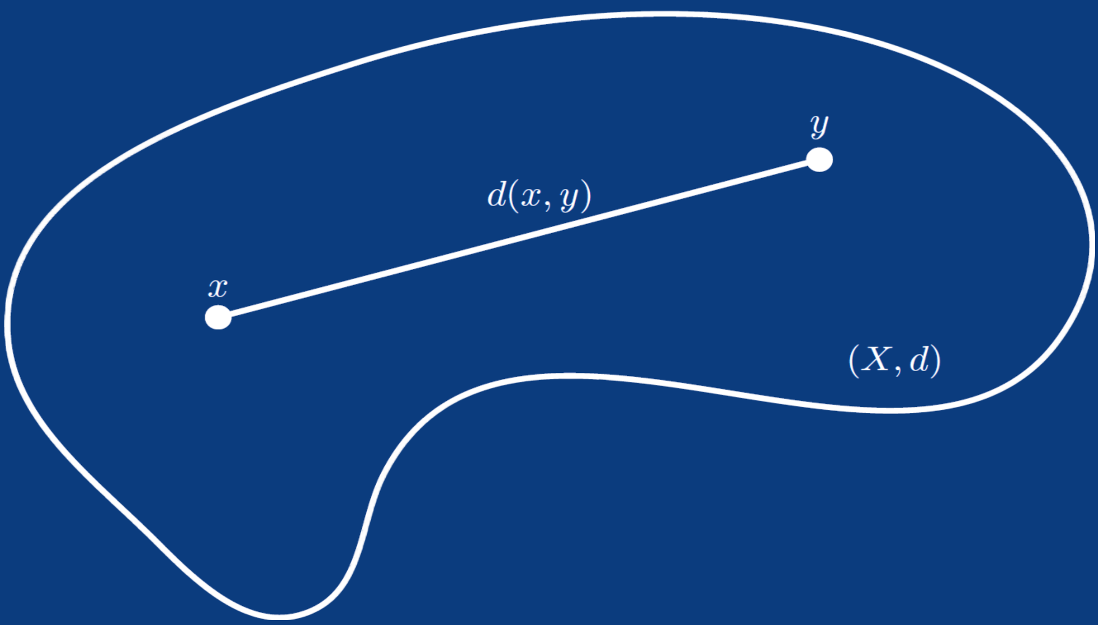
An Introduction to Metric Spaces

Samuel E. Andersson
Anders Björn
David Wiman



Nomenclature

Symbol	Explanation
$\mathbf{N}$	The set of all natural numbers, $\{0, 1, 2, ...\}$.
$\mathbf{Z}$	The set of all integers, $\{0, \pm 1, \pm 2, ...\}$.
$\mathbf{Q}$	The set of all rational numbers.
$\mathbf{R}$	The set of all real numbers.
$B_r(a)$	The open ball of radius r centered at a.
$B_r^X(a)$	The open ball of radius r centered at $a \in X$.
A^c	The complement of A, $X \setminus A$.
$f(A)$	The image of A under f, $\{f(x) \colon x \in A\}$.
∂A	The boundary of A.
$\subset$	Subset, not necessarily proper.
$X \times Y$	The set of points (x, y) such that $x \in X$ and $y \in Y$.
$\overline{A}$	The closure of A.
A°	The interior of A.

Preface

The study of metric spaces began in the early 20th century when Maurice Fréchet[1] published his PhD dissertation *Sur quelques points du calcul fonctionnel* in 1906. In it, he introduced the idea of a space where not much else but the notion of distance existed. However, Fréchet did not call it a metric space. The name is due to Felix Hausdorff[2], one of the founding fathers of modern topology.

In the beginning of the 20th century, functional analysis was still a novelty and an axiomatic foundation for mathematics had not yet been properly established. In the world of mathematics at that time, many people were studying spaces of functions and since many of them did not communicate with each other, they ended up with different notions of convergence in different spaces. The concept of convergence is strongly linked to the concept of distance, since we usually tie convergence to the notion of *approaching* certain points within a space. To make research simpler and more unified, Fréchet came up with the idea of axiomatizing the notion of distance so that a single definition of convergence could be used in most of the spaces at the time. Proofs based on these axioms would also hold in all of these spaces, which led to faster development as not every space had to be treated individually.

In this book, we will introduce you to metric spaces and some of the theory surrounding them. The purpose of this text is to be used as course material in TATA34 *Real analysis, honours course* at Linköping University. You will learn about how we define different distance functions, open sets, compactness and much more. At the end of the book, we will also quickly introduce you to topological spaces as to give you a hint of what further mathematics studies might hold.

[1] Maurice Fréchet (1878–1973), French mathematician.
[2] Felix Hausdorff (1868–1942), German mathematician.

The first edition of this book was written as a bachelor's thesis of mathematics at Linköping University by Samuel E. Andersson and David Wiman under the supervision of Anders Björn. Any potential later editions will feature changes by Anders Björn.

Contents

Chapter 1

Background

Before we introduce you to metric spaces, we want to remind you of a few definitions and theorems for $\mathbf{R}$ and $\mathbf{R}^n$ that are relevant in understanding this topic. Most of these should be recognizable from prior single variable and multivariable analysis courses, but some results from slightly more advanced analysis will also be presented.

In the theory of metric spaces, the convergence of sequences will play an important role and thus we remind ourselves of the definition of limits of sequences as well as the definition of Cauchy[1] sequences.

Definition 1.1. A sequence (x_n) in $\mathbf{R}$ *converges* to the *limit* L as $n \to \infty$, written

$$x_n \to L \quad \text{or} \quad \lim_{n\to\infty} x_n = L,$$

if for every $\varepsilon > 0$, there exists some $N \in \mathbf{N}$ such that $|x_n - L| < \varepsilon$ for all $n \geq N$. This can be expressed with quantifiers as

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbf{N} \colon n \geq N \implies |x_n - L| < \varepsilon.$$

Definition 1.2. A sequence (x_n) in $\mathbf{R}$ is a *Cauchy sequence* if for every $\varepsilon > 0$, there exists some $N \in \mathbf{N}$ such that $|x_m - x_n| < \varepsilon$ for all $m, n \geq N$. Expressed with quantifiers,

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbf{N} \colon m, n \geq N \implies |x_m - x_n| < \varepsilon.$$

Although this book is mainly concerned with sequences in other spaces than $\mathbf{R}$, it is still of great interest how these sequences specifically behave in $\mathbf{R}$.

[1] Louis Cauchy (1789–1857), French mathematician.

We continue with two theorems regarding convergence of sequences and subsequences. These theorems are not typically presented in a first course in single variable analysis, but they can be found on p. 67 and p. 64 respectively in [1].

Theorem 1.3. (Cauchy's convergence principle)
The sequence (a_n) *is convergent in* $\mathbf{R}$ *if and only if* (a_n) *is a Cauchy sequence in* $\mathbf{R}$.

Theorem 1.4. (The Bolzano[2]–Weierstrass[3] theorem)
If (a_n) *is bounded, then* (a_n) *has a convergent subsequence.*

Next we remind ourselves of the definitions of convergence and continuity of real functions.

Definition 1.5. A function $f\colon \mathbf{R} \to \mathbf{R}$ has the *limit* L as $x \to a$, written

$$f(x) \to L \quad \text{or} \quad \lim_{x \to a} f(x) = L,$$

if for every $\varepsilon > 0$, there exists some $\delta > 0$ such that $|f(x) - L| < \varepsilon$ whenever $0 < |x - a| < \delta$. Expressed with quantifiers,

$$\forall \varepsilon > 0 \;\; \exists \delta > 0\colon 0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon.$$

Definition 1.6. A function $f\colon \mathbf{R} \to \mathbf{R}$ is *continuous* at some point $a \in \mathbf{R}$ if

$$\lim_{x \to a} f(x) = f(a).$$

Remark 1.7. All the facts mentioned above also hold in $\mathbf{R}^n$, and Definitions 1.5 and 1.6 also hold for $f\colon \mathbf{R}^n \to \mathbf{R}^m$, where $n, m \geq 1$ if $|\cdot|$ is taken to be the *Euclidean*[4] *norm* (see Example 2.2).

In your multivariable analysis course, you hopefully came across the notion of compactness of sets. We remind ourselves of the definition in $\mathbf{R}^n$.

Definition 1.8. A set $K \subset \mathbf{R}^n$ is *compact* if it is closed and bounded.

You are now brought up to speed and we are ready to introduce you to metric spaces. We will look at limits and continuity in metric spaces in Chapter 3, while compactness is studied in Chapter 4.

---

[2] Bernard Bolzano (1781–1848), Czech mathematician.
[3] Karl Weierstrass (1815–1897), German mathematician.
[4] Euclid ($\sim$ 325–265 BC), Greek mathematician.

Chapter 2

Introduction to Metric Spaces

2.1 Definition of a Metric Space

Simply put, we want to establish a type of space with an inherent concept of distance between points within the space. The definition should be applicable to a wide variety of spaces, going beyond the standard geometric idea of distance. But there are some key properties we want to adhere to in order for this distance to make sense.

First, we would like two points to have zero distance between them only if they are one and the same. Second, the distance should be symmetric, meaning that the distance between two points should be the same regardless of which one you start from. Third, we want the triangle inequality to be true, i.e. the shortest distance between two points is the direct one without any "detours".

Definition 2.1. A *metric space* $X = (X, d)$ is a set X together with a *distance function*, or *metric*, $d\colon X \times X \to [0, \infty)$ such that for all $x, y, z \in X$,

(i) $d(x, y) = 0$ if and only if $x = y$,

(ii) $d(x, y) = d(y, x)$,

(iii) $d(x, y) \leq d(x, z) + d(z, y)$.

Notice that by setting $y = x$ in constraint (iii) we get $0 \leq d(x, y)$ since $d(x, x) = 0$ and $d(x, y) = d(y, x)$, so the three constraints would guarantee that the metric is a non-negative function even if we only required $d\colon X \times X \to \mathbf{R}$. This is of course very important in conveying the notion of distance.

From the small number of constraints of the definition, we can tell that metric spaces are quite general spaces. In fact, most spaces you have encountered are metric spaces.

Example 2.2. Below, we give some examples of metric spaces you have encountered before but perhaps not recognized as metric spaces.

- $X = (\mathbf{R}, d)$ with $d(x,y) = |x-y| = \begin{cases} x-y, & \text{if } x \geq y, \\ y-x, & \text{if } x < y. \end{cases}$
- $X = (\mathbf{R}^n, d)$ with the Euclidean distance

$$d(x,y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2},$$

  where $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.
- $X = (E, d)$, where $E \subset \mathbf{R}^n$, with $d(x,y)$ being the same as for $\mathbf{R}^n$.

Example 2.3. There are many other metric spaces that you might not have encountered yet which are important in other fields, both within and outside pure mathematics. The following are some examples of such metric spaces.

- The set of all binary numbers of length n with the *Hamming*[1] *distance*

$$d(x,y) = \sum_{k=1}^{n} |x_k - y_k|,$$

  where $x = (x_1 \dots x_n)$ and $y = (y_1 \dots y_n)$, with $x_k, y_k \in \{0,1\}$. This is intuitively understood as the smallest number of bit-flips required to change x into y. This distance function is also referred to as the ℓ^1-distance, see Example 2.4.
- The set of all compact $K \subset \mathbf{R}^n$ (see Definition 1.8) with the Hausdorff distance $d(K,L) = \inf\{\varepsilon > 0 \colon K \subset L_\varepsilon \text{ and } L \subset K_\varepsilon\}$, where

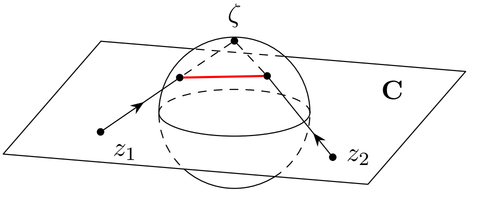
$$A_\varepsilon = \bigcup_{x \in A} \{y \in \mathbf{R}^n \colon |x-y| < \varepsilon\}.$$

  Intuitively, the Hausdorff distance between K and L is the largest possible distance between a point in K and the set L, or vice versa.

[1] Richard Hamming (1915–1998), American mathematician.

- The complex plane **C** with chordal distance on the Riemann[2] sphere, which you will encounter in a course in complex analysis if you have not already. The *chordal distance* between two points $z_1, z_2 \in \mathbf{C}$ is given by

$$d(z_1, z_2) = \frac{2|z_1 - z_2|}{\sqrt{(1+|z_1|^2)(1+|z_2|^2)}}.$$



Chordal distance on the Riemann sphere.

- The set $C([a,b])$ of all real-valued continuous functions on the interval $[a,b]$, with the L^1-norm,

$$d(f,g) = \int_a^b |f(x) - g(x)|\,dx.$$

- $C([a,b])$ with the supremum norm,

$$d(f,g) = \|f-g\|_\infty = \sup_{a \le x \le b} |f(x) - g(x)|.$$

If you have taken a course in Fourier[3] analysis you may recognize the L^1-norm and the supremum norm as these are norms on some of the most commonly studied spaces, $L^1([a,b])$ and $L^\infty([a,b])$ respectively. We will discuss these further in Chapter 7.

Of course there is not only one distance function per set that qualifies it as a metric space. In fact, for all sets with at least two points, there is an infinite number of possible distance functions.

[2] Bernhard Riemann (1826–1866), German mathematician.
[3] Joseph Fourier (1768–1830), French mathematician.

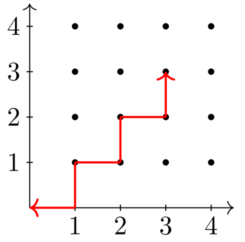
Example 2.4. The following are some examples of different distance functions for $\mathbf{R}^n$. Here $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

- ℓ^p-distance: $d(x,y) = \left(\sum_{k=1}^n |x_k - y_k|^p\right)^{1/p}$ for any $1 \le p < \infty$.
- ℓ^∞-distance: $d(x,y) = \max_k |x_k - y_k|$.
- Discrete distance: $d(x,y) = \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } x \neq y. \end{cases}$

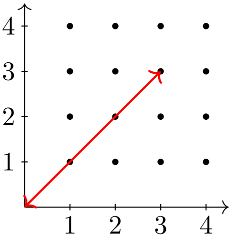
Remark 2.5. The ℓ^1-distance function is often called the *taxicab* distance. It is intuitive to see why when regarding the case of $\mathbf{R}^2$ if you are familiar with how taxicabs have to drive in New York City for example. Note that the shortest path is not unique.

The ℓ^2-distance is the Euclidean distance in Example 2.2. It is also the only ℓ^p-distance which has a corresponding inner product, being the standard dot product. Meaning that $d_{\ell^2}(0,x) = \sqrt{\langle x|x\rangle}$.

The discrete distance is the simplest distance function, and can be defined on any set X.



Taxicab distance, 6 units.



ℓ^2-distance, $\sqrt{18} \approx 4.24$ units.

2.2 Open Sets

Now that we have a concept of distance within spaces, we can introduce the concept of neighborhoods to points. We define the set of all points that are less than some distance $r > 0$ away from a given point, which can be referred to as an *r-neighborhood*. From now on, all sets X and Y will be metric spaces unless otherwise stated.

Definition 2.6. An *open ball* of radius r centered at the point a is defined by

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

When working with different metric spaces we sometimes write $B_r^X(a)$.

Notice that for the metric space $X = [0, \infty)$ with $d(x,y) = |x-y|$,

$$B_1(1) = (0,2), \quad \text{but } B_1(0) = [0,1).$$

In fact

$$B_{3/4}\left(\tfrac{1}{4}\right) = [0,1) = B_1(0),$$

so the center and radius of a ball are not unique. This may seem strange at first glance, but is in fact common in metric spaces.

Furthermore, we would like a way to categorize the points in a set based upon their surroundings. One can imagine a type of point in a set which is completely surrounded by other points within the set, i.e. has a neighborhood which is also contained by the set.

Definition 2.7. A point a is an *inner point*, or *interior point*, in $G \subset X$ if there exists an $\varepsilon > 0$ such that $B_\varepsilon(a) \subset G$.

Based on the characteristics of the points within a set, we can now begin to categorize sets. We introduce a category of sets which is exclusively made up of inner points.

Definition 2.8. $G \subset X$ is *open* if every point $a \in G$ is an inner point in G.

Note that open balls are indeed open sets, hence the name.

Example 2.9. If $X = \mathbf{R}$, then $\mathbf{R}$, $\varnothing$ and open intervals of the form (a,b) are open, just to name a few.

Now that we have defined open sets and seen some examples, we devise methods to determine whether or not sets are open, depending on how they are constructed.

Theorem 2.10. *If G_λ is open for every $\lambda \in \Lambda$, then $\bigcup_{\lambda \in \Lambda} G_\lambda$ is open.*

Notice that the theorem does not require any limitation on the index set Λ. This means that in addition to being valid for finite Λ, it is valid for infinite index sets, even uncountable ones.

Proof. If $x \in \bigcup_{\lambda \in \Lambda} G_\lambda$ then there is some $\lambda_0 \in \Lambda$ for which $x \in G_{\lambda_0}$. As G_{λ_0} is open there is some $\varepsilon > 0$ such that

$$B_\varepsilon(x) \subset G_{\lambda_0} \subset \bigcup_{\lambda \in \Lambda} G_\lambda. \qquad \square$$

Now that we have seen that we can construct open sets by taking a union of open sets, it is natural to consider intersections of open sets.

Theorem 2.11. *If $G_1, \dots, G_N$ are open, then $\bigcap_{j=1}^{N} G_j$ is open.*

Proof. If $x \in \bigcap_{j=1}^{N} G_j$, then $x \in G_j$ for every $j = 1, \dots, N$. Thus there are $\varepsilon_1, \dots, \varepsilon_N > 0$ such that $B_{\varepsilon_j}(x) \subset G_j$. If we let $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_N\} > 0$ then $B_\varepsilon(x) \subset \bigcap_{j=1}^{N} G_j$. Therefore the intersection is open. □

Example 2.12. The set

$$A = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$

is an intersection of open sets but is in fact not open. Note that this is because Theorem 2.11 is only valid for *finite* intersections.

Since not all sets are open and only consist of inner points, we might want to be able to single out and refer to the inner points of a set in a handy way.

Definition 2.13. The *interior* of A is the set of all inner points of A. It is denoted A°.

Theorem 2.14. *Let $A \subset X$ where X is a metric space. Then*

(a) *A° is open.*

(b) *A° is the largest open subset contained in A.*

Proof. (a) If $x \in A^\circ$ then there exists an $r_x > 0$ such that $B_{r_x}(x) \subset A$. Since $B_{r_x}(x)$ is open, it only contains inner points of A, thus $B_{r_x}(x) \subset A^\circ$. Hence A° is open per Definition 2.8.

(b) Fix a set A and let $G \subset A$ be an open set. Let $a \in G$ be arbitrary. Since G is open, a is an inner point and there exists some $\varepsilon > 0$ such that $B_\varepsilon(a) \subset G$. But since $G \subset A$, then $B_\varepsilon(a) \subset A$ and thus $a \in A^\circ$, showing that $G \subset A^\circ$. Since A° is always open, A° is the largest open subset contained in A. □

2.3 Closed Sets

A different way of categorizing points in a set based on their surroundings is to look at if there are other points belonging to the same set in every neighborhood of the point. We are also interested in categorizing points that have no other points from the same set that are arbitrarily close.

Definition 2.15. A point $x \in X$ is a *limit point*, or *accumulation point*, of A if $B_\varepsilon(x) \cap (A \setminus \{x\}) \neq \varnothing$ for every $\varepsilon > 0$. If $x \in A$ is not a limit point, it is called *isolated* in A.

Intuitively, limit points of a set are points which are arbitrarily close to other points of the set.

The definition of a limit point is quite similar to that of an inner point, and the statement that a certain point is a limit point is weaker than it being an inner point, as indeed all inner points that are not isolated are limit points.

Much like we derived open sets from the definition of inner points, we can categorize sets based upon their limit points. We introduce a category of sets which contain all of their limit points.

Definition 2.16. $F \subset X$ is *closed* if it contains all of its limit points.

Just like for Theorems 2.10 and 2.11 for open sets, we will show that intersections and finite unions of closed sets are indeed also closed. The proof however requires theory presented later in this book, so we will be coming back to this theorem later to prove it in Section 2.5.

Theorem 2.17.

(a) *If F_λ is closed for all $\lambda \in \Lambda$, then $\bigcap_{\lambda\in\Lambda} F_\lambda$ is closed.*

(b) *If $F_1, \dots, F_N$ are closed, then $\bigcup_{j=1}^{N} F_j$ is closed.*

Remark 2.18. We commonly use F for closed sets and G for open ones. This is a tradition that most likely came from French where the word for closed, *fermé*, begins with an f. G was simply the next letter in the alphabet. This book will use this convention where applicable. When we do not know if a set is open or closed, or it is not important for the statement, we will use some other letter.

Example 2.19. Let $A = [0, 1] \cup \{2\}$. Its limit points are all the points in $[0, 1]$ and its isolated point is 2. The set contains all of its limit points and is therefore closed.

Example 2.20. Let $A = [0, 1)$. Its limit points are all the points in $[0, 1]$ and it has no isolated points. Since the limit point 1 is not in the set, it is not closed.

Example 2.21. Let $A = \{\frac{1}{n} : n = 1, 2, \dots\}$. Then A consists only of isolated points. The only limit point to A is 0, which is not in the set, and therefore A is not closed. However, A is not open either since none of its points are inner points. This is fairly common and most sets do not fall into any of these two categories. Sets that fall into both categories are sometimes called *clopen*, more on them later on.

In Examples 2.19–2.21 we (implicitly) assumed them to be within the metric space $\mathbf{R}$ with the standard metric. Both openness and closedness depend on which metric space they are considered in. For instance $A = [0, 1)$ is clopen within itself.

Whether or not a set does not contain all of its limit points, we can add them to the set to create a new set.

Definition 2.22. $\overline{A} = A \cup L$, where L is the set of all limit points to A, is called the *closure* of A.

Note that $A \subset \overline{A}$. This is clear by the definition. From this point, it is not all that hard to realize that

$$A \subset B \quad \Longrightarrow \quad \overline{A} \subset \overline{B}. \tag{2.1}$$

The closure of a set A is always closed, and is the smallest closed set containing A. It is also easy to realize that A is closed if and only if it is equal to its closure, and thus the closure of an arbitrary closure $\overline{A}$ is equal to itself.

Theorem 2.23. *Let $A \subset X$ where X is a metric space. Then*

(a) *A is closed if and only if $A = \overline{A}$.*

(b) $\overline{A} = \overline{\overline{A}}$.

(c) *$\overline{A}$ is always closed.*

(d) *$\overline{A}$ is the smallest closed set containing A.*

Proof. (a) If $\overline{A} = A$ then A must contain all of its limit points, per definition, and is therefore closed. Conversely, if we know that A is closed and therefore contains all of its limit points, it must necessarily be equal to $\overline{A}$.

(b) By Definition 2.22, we conclude that $\overline{A} \subset \overline{\overline{A}}$ since

$$\overline{\overline{A}} = \overline{A} \cup \{\text{the limit points of } \overline{A}\} = \overline{A} \cup L_{\overline{A}}.$$

We now need to show that $L_{\overline{A}} \subset \overline{A}$. To this end, let $x \in L_{\overline{A}}$ and $x_j \in \overline{A}$ be such that $x_j \to x$. Then there are also $y_j \in A \cap B_{1/j}(x_j)$. It follows that also $y_j \to x$, and thus $x \in L_A$. Hence $L_{\overline{A}} \subset L_A \subset \overline{A}$ and therefore

$$\overline{\overline{A}} \subset \overline{A},$$

i.e. $\overline{\overline{A}} = \overline{A}$.

(c) $\overline{A}$ is closed if and only if $\overline{A} = \overline{\overline{A}}$ by (a) and since $\overline{A} = \overline{\overline{A}}$ is always true by (b), $\overline{A}$ is always closed.

(d) Take any closed set F such that $A \subset F$. Then $\overline{A} \subset \overline{F} = F$ by (2.1) and (a). Since the choice of F was arbitrary, all closed sets that contain A will also contain $\overline{A}$. As $\overline{A}$ is closed by (c), $\overline{A}$ is the smallest closed set containing A. □

2.4 Complement and De Morgan's Laws

Now that we have defined both open and closed sets, we are interested in their relationship. However, first we need to introduce some new tools.

We leave the world of analysis for a brief detour to the realm of discrete mathematics. We begin by defining the set of all points in a metric space which are not contained within a certain set.

Definition 2.24. The *complement* of a set $A \subset X$ is defined by

$$A^c = X \setminus A = \{b \in X : b \notin A\}.$$

Notice how applying the complement twice on a set will result in the original set, i.e.

$$(A^c)^c = A. \tag{2.2}$$

Note also that we can express the interior in terms of the closure of the complement by

$$A^\circ = X \setminus \overline{A^c}.$$

To be able to prove Theorem 2.17 we need to introduce you to De Morgan's[4] laws. You may have encountered them in a previous discrete mathematics course or perhaps in a course on logic gate networks.

Theorem 2.25. (De Morgan's laws)

(a) $\displaystyle\Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c = \bigcup_{\lambda\in\Lambda} A_\lambda^c.$

(b) $\displaystyle\Big(\bigcup_{\lambda\in\Lambda} A_\lambda\Big)^c = \bigcap_{\lambda\in\Lambda} A_\lambda^c.$

As before Λ is an index set that may be finite, countable or uncountable.

[4] Augustus De Morgan (1806–1871), British mathematician.

Proof. (a) Let

$$x \in \Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c.$$

This means that x is not contained in every set A_λ so there must be at least one $\lambda_0 \in \Lambda$ for which x is in $A^c_{\lambda_0}$. Therefore,

$$x \in A^c_{\lambda_0} \subset \bigcup_{\lambda\in\Lambda} A^c_\lambda.$$

This shows that

$$\Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c \subset \bigcup_{\lambda\in\Lambda} A^c_\lambda.$$

Conversely, let

$$y \in \bigcup_{\lambda\in\Lambda} A^c_\lambda.$$

This means that there is some λ_1 such that $y \in A^c_{\lambda_1}$. Therefore

$$y \notin A_{\lambda_1} \supset \bigcap_{\lambda\in\Lambda} A_\lambda.$$

So

$$y \in \Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c,$$

which in turns shows that

$$\bigcup_{\lambda\in\Lambda} A^c_\lambda \subset \Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c.$$

These two conclusions together show that

$$\Big(\bigcap_{\lambda\in\Lambda} A_\lambda\Big)^c = \bigcup_{\lambda\in\Lambda} A^c_\lambda.$$

(b) Using (2.2), and then (a) applied to A_λ, gives

$$\bigcap_{\lambda\in\Lambda} A^c_\lambda = \bigg(\Big(\bigcap_{\lambda\in\Lambda} A^c_\lambda\Big)^c\bigg)^c = \Big(\bigcup_{\lambda\in\Lambda} (A^c_\lambda)^c\Big)^c = \Big(\bigcup_{\lambda\in\Lambda} A_\lambda\Big)^c. \qquad \square$$

2.5 Connection Between Open and Closed Sets

We now have the tools necessary to examine the relationship between open and closed sets.

Theorem 2.26. *Let F and G be subsets of the metric space X. Then*

(a) *F is closed if and only if F^c is open,*

(b) *G is open if and only if G^c is closed.*

Remark 2.27. This gives an alternative way of defining closed sets assuming we have defined open sets. It is in fact very common to define closed sets in this manner.

Proof. (a) First we prove that F being closed implies that F^c is open: Let $x \in F^c$. Since F is closed, x is not a limit point to F. This in turn means there exists an $\varepsilon > 0$ such that $B_\varepsilon(x) \cap F = \varnothing$. Therefore $B_\varepsilon(x) \subset F^c$, and hence F^c is open.

Next we prove the converse; that F^c being open implies that F is closed: Let $x \in F^c$. Since F^c is open, there exists $B_\varepsilon(x) \subset F^c$ for some $\varepsilon > 0$. This means that x is not a limit point to F which in turn means that all limit points belong to F, i.e. F is closed.

(b) We prove that G is open if and only if G^c is closed by utilizing (a): We know that G^c is closed if and only if $(G^c)^c$ is open. But we also know $(G^c)^c = G$ by (2.2). Thus G is open if and only G^c is closed. □

We have now developed all the tools needed to prove Theorem 2.17.

Proof of Theorem 2.17. (a) By use of De Morgan's laws, we can write

$$\Big(\bigcap_{\lambda\in\Lambda} F_\lambda\Big)^c = \bigcup_{\lambda\in\Lambda} F_\lambda^c.$$

By Theorem 2.26 (a) each F_λ^c is open, and a union of open sets is open by Theorem 2.10. Therefore $(\bigcap_{\lambda\in\Lambda} F_\lambda)^c$ is also open. By applying Theorem 2.26 (a) again we conclude that $\bigcap_{\lambda\in\Lambda} F_\lambda$ is closed.

(b) Using De Morgan's laws again, we can write

$$\Big(\bigcup_{j=1}^{N} F_j\Big)^c = \bigcap_{j=1}^{N} F_j^c.$$

A finite intersection of open sets is open by Theorem 2.11 and therefore $(\bigcup_{j=1}^N F_j)^c$ is also open. By using Theorem 2.26 (a) we conclude that $\bigcup_{j=1}^N F_j$ is closed. □

2.6 Perfect, Dense, Bounded and Boundary

Now that we have laid some ground work for open and closed sets as well as looked at their relation, we shall dive a bit deeper and look at some finer details.

We are particularly interested in closed sets that lack isolated points. Such sets contain all of their limit points, and only limit points. We will see that this distinction from arbitrary closed sets is important for certain statements.

Definition 2.28. F is *perfect* if it is closed and has no isolated points.

As previously, you have encountered perfect sets before but perhaps not recognized them as such.

Example 2.29. The following are some examples of perfect sets.

- Any interval on the form $[a, b] \subset \mathbf{R}$, and any finite union of such intervals.
- The Cantor[5] set $\mathcal{C}$.

In certain contexts a set can be perfect while in other contexts it is not. For example, $\mathbf{Q}$ has no isolated points but is not closed in $\mathbf{R}$ and is therefore not perfect in $\mathbf{R}$. However, $\mathbf{Q}$ is closed in itself and therefore perfect in itself. In fact a metric space that lacks isolated points will always be perfect in relation to itself.

Another way we can categorize sets is by their relation to the metric space they are in. We introduce a category of sets where all points in the metric space are either in the set or limit points of the set.

Definition 2.30. $A \subset X$ is *dense* in X if $\overline{A} = X$.

Example 2.31. Let $x \in \mathbf{R}$. Then every $B_\varepsilon(x)$ contains $y \in \mathbf{Q} \setminus \{x\}$. Therefore x is a limit point of $\mathbf{Q}$. This means that $\overline{\mathbf{Q}} = \mathbf{R}$, i.e. $\mathbf{Q}$ is dense in $\mathbf{R}$.

We can also categorize points in a set based on their surroundings by looking at whether or not they make up the boundary of a set, i.e. if every neighborhood around the point is only partially contained in the set.

Definition 2.32. A point $x \in X$ is a *boundary point* of A if both $B_\varepsilon(x) \cap A$ and $B_\varepsilon(x) \setminus A$ are non-empty for every $\varepsilon > 0$. The set of all boundary points of A is called the *boundary* of A, written as ∂A.

We can conclude that a non-isolated point is a boundary point if and only if it is a limit point, but not an interior point. Thus, a set is closed if and only if it contains its boundary points. Moreover $\overline{A} = A \cup \partial A$.

We also define what it means for a set to be bounded. You probably have an intuitive understanding of this but here we give you the formal definition.

[5] Georg Cantor (1845–1918), German mathematician.

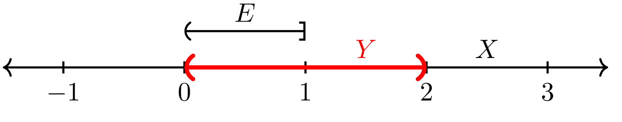
Definition 2.33. A set $A \subset X$ is *bounded* if there is some $x_0 \in X$ and $r_0 > 0$ such that $A \subset B_{r_0}(x_0)$.

Note that if $A \subset B_{r_0}(x_0)$, $x_1 \in X$ and $r_1 = r_0 + d(x_0, x_1)$, then $A \subset B_{r_1}(x_1)$. So Definition 2.33 is independent of the choice of $x_0 \in X$.

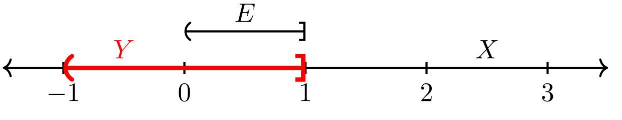
2.7 Relative or Induced Topology

Within a metric space we can also talk about the topology, i.e. which sets are open and which are closed, relative to subsets of the space. We call this the *relative* or *induced* topology. If (X, d) is a metric space and Y is a subset of X then d is automatically a distance function also on Y, and thus (Y, d) is also a metric space. If E is a subset of Y then we can talk about E being open or closed with respect to either X or Y. Within Y we often say relatively open/closed.

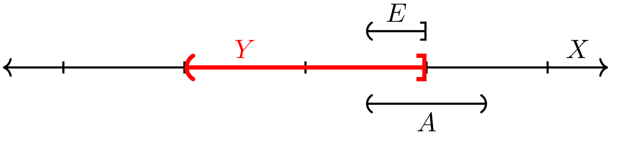
Example 2.34. Let $X = \mathbf{R}$. If $Y = (0, 2)$, then $E = (0, 1]$ is relatively closed with respect to Y but it is neither closed nor open with respect to X.



If we instead let $Y = (-1, 1]$, then E is relatively open in Y.



Theorem 2.35. *E is relatively open/closed in Y, where $Y \subset X$, if and only if there exists a set A that is open/closed in X such that $E = A \cap Y$, i.e. E is the part of A that is in Y.*



Proof. We only prove the theorem in the open case, but the proof for the closed case is very similar. First we prove that if there exists an open set $A \subset X$ such that $E = A \cap Y$, then E is open in Y:

A is open in X so for every $a \in A$, there exists an $\varepsilon > 0$ such that

$$B_\varepsilon^X(a) = \{x \in X : d(x,a) < \varepsilon\} \subset A.$$

As $Y \subset X$ it follows that

$$B_\varepsilon^Y(a) = \{x \in Y : d(x,a) < \varepsilon\} \subset A.$$

As also $B_\varepsilon^Y(a) \subset Y$, we see that

$$B_\varepsilon^Y(a) \subset E,$$

and thus E is open in Y.

We conclude the proof by proving the converse statement, i.e. that if E is open in Y, then there exists an open set $A \subset X$ such that $E = A \cap Y$.

Since E is open in Y we can for every point $x \in E$ choose an $r_x > 0$ such that $B_{r_x}^Y(x) \subset E$. Choose

$$A = \bigcup_{x \in E} B_{r_x}^X(x),$$

which is open in X, since the union of open balls is also open by Theorem 2.10. Furthermore, we realize that

$$A \cap Y = \bigcup_{x \in E} B_{r_x}^Y(x) \subset E,$$

and that since the set $\bigcup_{x \in E} B_{r_x}^Y(x)$ must contain at least all points in E,

$$E \subset \bigcup_{x \in E} B_{r_x}^Y(x) = A \cap Y.$$

Thus $E = A \cap Y$. $\square$

Chapter 3

Limits and Sequences

3.1 Sequences

We start by looking at sequences in metric spaces. A sequence, much like a set, is a collection of elements. But unlike a set, the order matters and the same element can appear more than once. Sequences can be finite or infinite, in which case we can talk about the limit of the sequence when we go arbitrarily far along the sequence. Here we will only consider infinite sequences (x_n) where n runs through $1, 2, 3, \dots$. Definitions 1.1 and 1.2 easily generalize to metric spaces, as we only need to replace $|x - y|$ with $d(x, y)$.

Definition 3.1. A sequence $(x_n) \subset X$ *converges* to the *limit* L as $n \to \infty$ if for every $\varepsilon > 0$ there exists some $N \in \mathbf{N}$ such that $d(x_n, L) < \varepsilon$ for all $n \geq N$. We write

$$x_n \to L \quad \text{or} \quad \lim_{n\to\infty} x_n = L.$$

Expressed with quantifiers, $x_n \to L$ as $n \to \infty$ if

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbf{N} \colon n \geq N \implies d(x_n, L) < \varepsilon.$$

The limit of a sequence in a metric space is unique, i.e. if $x_n \to L$ and $x_n \to M$ then $L = M$. We leave showing this as an exercise for the reader.

Now that we have defined convergence of sequences in metric spaces, we can present a new way of determining if a point is a limit point to a set or not.

Theorem 3.2. *x is a limit point to $A \subset X$ if and only if there exists a sequence $(x_n) \subset A \setminus \{x\}$ such that $x_n \to x$.*

Proof. We begin by proving that x being a limit point implies that there exists a sequence $(x_n) \subset A \setminus \{x\}$ such that $x_n \to x$: For all j, there is an x_j such that

$$x_j \in B_{1/j}(x) \cap (A \setminus \{x\}).$$

This means that $d(x_j, x) < 1/j$ and thus $x_j \to x$.

Next we want to prove the converse, i.e. that the existence of a sequence $(x_n) \subset A \setminus \{x\}$ such that $x_n \to x$ implies that x is a limit point: For every $\varepsilon > 0$, there is an n such that $d(x_n, x) < \varepsilon$, which means that

$$x_n \in B_\varepsilon(x) \cap (A \setminus \{x\})$$

and thus

$$B_\varepsilon(x) \cap (A \setminus \{x\}) \neq \varnothing. \qquad \square$$

A particularly interesting type of sequence that we can talk about when we have a notion of distance are so-called Cauchy sequences. Intuitively, a Cauchy sequence is a sequence where, arbitrarily far along the sequence, the distance between elements becomes arbitrarily small.

Definition 3.3. $(a_n) \subset X$ is a *Cauchy sequence* (or (a_n) is *Cauchy* for short) if for every $\varepsilon > 0$, there exists some $N \in \mathbf{N}$ such that $d(a_n, a_m) < \varepsilon$ for all $n, m \geq N$. Expressed with quantifiers, (a_n) is a Cauchy sequence if

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbf{N} \colon n, m \geq N \implies d(a_n, a_m) < \varepsilon.$$

This type of sequence is interesting because the property it has is closely tied to convergence, and entails other important properties of the sequence. First we look at the relation between Cauchy sequences and convergent sequences.

Theorem 3.4. *If (a_n) converges, then (a_n) is a Cauchy sequence.*

Proof. Let $\varepsilon > 0$. Suppose $a_n \to L$. Then there exists $N \in \mathbf{N}$ such that if $n \geq N$ then $d(x_n, L) < \frac{\varepsilon}{2}$. Then, for all $n, m \geq N$, we have $d(x_n, x_m) \leq d(x_n, L) + d(L, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. $\square$

Next we want to examine the relation between Cauchy sequences and bounded sequences. Recall that as a sequence is a set (if we ignore the ordering), Definition 2.33 of bounded sets applies here as well.

Theorem 3.5. *If (a_n) is a Cauchy sequence, then (a_n) is bounded.*

Proof. Since (a_n) is Cauchy, for every $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that $d(a_n, a_m) < \varepsilon$ for all $n, m \geq N$. Now let $\varepsilon = 1$, that is $d(a_n, a_m) < 1$ if $n, m \geq N$. Then

$$d(a_n, a_N) < 1$$

if $n > N$. Thus all a_n, $n \geq N$, are contained within the open ball $B_1(a_N)$, so $(a_n)_{n=N}^{\infty}$ is bounded. The sequence $(a_n)_{n=1}^{N}$ is finite and thus is bounded. Hence (a_n) is bounded. □

3.2 Completeness

From Theorem 3.4 we know that all convergent sequences in a metric space are Cauchy sequences. But what about the converse statement; are all Cauchy sequences convergent?

In the case of $\mathbf{R}$ with the standard distance function the answer is yes, by Cauchy's convergence principle (Theorem 1.3). However, using the same metric on $\mathbf{Q}$ we can easily see that the sequence $(x_n) = (3, 3.1, 3.14, 3.141, ...)$ is a Cauchy sequence, and that it approaches π, which is not in $\mathbf{Q}$. Hence the sequence is Cauchy but not convergent in $\mathbf{Q}$.

So the converse statement is not generally true. But the property of all Cauchy sequences being convergent is so important that metric spaces which have it are given their own category.

Definition 3.6. A metric space is *complete* if every Cauchy sequence is convergent.

As we saw above, when we use the standard distance function, $\mathbf{R}$ is complete but $\mathbf{Q}$ is not. Using different distance functions on the same set can change whether or not that metric space is complete. Looking back at Example 2.3 we can examine some of the different metric spaces to see if they are complete.

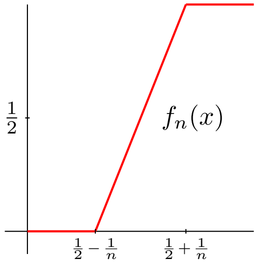
The L^1-norm on $C([0,1])$ does not give rise to a complete metric space. For example, consider the sequence of functions (f_n) where

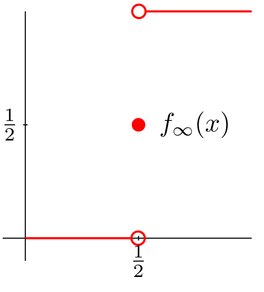
$$f_n(x) = \begin{cases} 0, & \text{if } 0 \leq x \leq \frac{1}{2} - \frac{1}{n}, \\ \frac{n}{2}\left(x - \left(\frac{1}{2} - \frac{1}{n}\right)\right), & \text{if } \frac{1}{2} - \frac{1}{n} \leq x \leq \frac{1}{2} + \frac{1}{n}, \\ 1, & \text{if } \frac{1}{2} + \frac{1}{n} \leq x \leq 1, \end{cases}$$

on $A = [0,1]$ with $1 \leq n < \infty$. All of these functions are continuous on the interval $[0,1]$ so they are in $C([0,1])$. They also get arbitrarily close to each other using the L^1-norm as $n \to \infty$ so (f_n) is Cauchy. However, the pointwise limit of this sequence is the function

$$f_\infty(x) = \begin{cases} 0, & \text{if } 0 \le x < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } x = \frac{1}{2}, \\ 1, & \text{if } \frac{1}{2} < x \le 1, \end{cases}$$

which is not continuous on $[0,1]$ and therefore not in $C([0,1])$. Moreover there is no continuous function f such that $f = f_\infty$ almost everywhere. Hence the sequence does not converge and $(X,d) = (C([0,1]), L^1)$ is not a complete space. However, using the supremum norm, $C([0,1])$ gives rise to a complete metric space.





Theorem 3.7. $C([a,b])$ *with the supremum norm is a complete metric space.*

We do not have the tools necessary to prove this yet so we will return to it later and prove it in Chapter 5.

Next we present a theorem relating completeness and perfect sets to cardinality.

Theorem 3.8. *If* X *is a complete metric space and* $F \subset X$ *is perfect, then* F *is uncountable.*

The proof for Theorem 3.8 is very tricky, even for simple special cases like $X = \mathbf{R}$, and is outside the scope of this book.

3.3 Continuous Functions

Close to the heart of analysis are limits of functions and continuity, which are essential for things such as derivatives and integrals, among many other things. Limits are most often very closely tied to notions of distance, as we talk about functions *approaching* a value for instance. You are likely familiar with the definition of limits for real functions (see Definition 1.5), so we start by generalizing that definition to metric spaces.

Definition 3.9. Let $f\colon X \to Y$ be a mapping, where X and Y are metric spaces with metrics d_X and d_Y respectively. Assume that a is not isolated in X. Then $f(x)$ *tends to* L as x tends to a if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_Y(f(x), L) < \varepsilon$ whenever $0 < d_X(x,a) < \delta$. We write

$$f(x) \to L \text{ as } x \to a \quad \text{or} \quad \lim_{x\to a} f(x) = L.$$

Using quantifiers to express the definition, we see that $f(x) \to L$ as $x \to a$ if

$$\forall \varepsilon > 0 \;\; \exists \delta > 0\colon 0 < d_X(x,a) < \delta \implies d_Y(f(x), L) < \varepsilon.$$

As with limits of sequences also here the limit L is unique if it exists. That is provided that a is not isolated. On the other hand, if a is isolated the condition $0 < d_X(x,a) < \delta$ is void if $\delta > 0$ is small enough and thus any L satisfies the definition. Thus we *do not* define limits as $x \to a$ if a is isolated.

Example 3.10. We begin with a familiar example in $\mathbf{R}$. We want to show that $\lim_{x\to a} \sqrt{x} = \sqrt{a}$ when $a > 0$. We use the standard distance function in $\mathbf{R}$ and we let $\varepsilon > 0$,

$$|\sqrt{x} - \sqrt{a}| = \frac{|x-a|}{\sqrt{x}+\sqrt{a}} \le \frac{|x-a|}{\sqrt{a}}.$$

If $|x - a| < \varepsilon\sqrt{a}$ then

$$\frac{|x-a|}{\sqrt{a}} < \varepsilon,$$

which is what we want. We set $\delta = \varepsilon\sqrt{a}$ so that the inequality holds.

Next we generalize the definition of continuity to functions on general metric spaces. This definition, much like the definition of limits we have just seen, will be very recognizable if you are familiar with the definition in the case of real-valued functions with the standard metric (see Definition 1.6).

Definition 3.11. $f\colon X \to Y$ is *continuous* at $a \in X$ if for every $\varepsilon > 0$, there exists some $\delta > 0$ such that $d_Y(f(x), f(a)) < \varepsilon$ whenever $d_X(x,a) < \delta$. Expressed with quantifiers, f is continuous at $a \in X$ if

$$\forall \varepsilon > 0 \;\; \exists \delta > 0\colon d_X(x,a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon.$$

If a function is continuous at every point of a set then the function is said to be *continuous* on that set.

Remark 3.12. An alternative, and completely equivalent, way of defining continuity is by using open balls. Expressed with quantifiers, f is continuous at $a \in X$ if

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \colon x \in B_\delta^X(a) \implies f(x) \in B_\varepsilon^Y(f(a)).$$

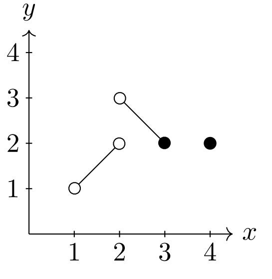
Here $B_\delta^X(a)$ is an open ball in X and $B_\varepsilon^Y(f(a))$ is an open ball in Y.

Note that if f is defined at $a \in X$ and a is isolated, then f is continuous at a. Thus every function $f \colon \mathbf{Z} \to \mathbf{Z}$ is automatically continuous. If a however is a limit point, then f is continuous at a if and only if $\lim_{x \to a} f(x) = f(a)$. Recall that limits are not defined at isolated points.

Example 3.13. The function

$$f(x) = \begin{cases} x, & \text{if } 1 < x < 2, \\ 5 - x, & \text{if } 2 < x \le 3, \\ 2, & \text{if } x = 4, \end{cases}$$

is continuous on $(1,2) \cup (2,3] \cup \{4\}$. We see this since all points between 1 and 3 are limit points and 4 is isolated but f is not defined at 1 and 2.



Theorem 3.14. *$f \colon X \to Y$ is continuous if and only if the preimage*

$$f^{-1}(G) = \{x \in X \colon f(x) \in G\}$$

is open for all open $G \subset Y$.

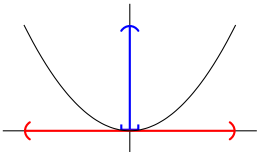
Proof. We prove that if f is continuous then $f^{-1}(G)$ is open for all open $G \subset Y$: Let $x \in f^{-1}(G)$. Then $f(x) \in G$. Since G is open, $B_\varepsilon(f(x)) \subset G$ for some $\varepsilon > 0$. Also, since f is continuous, there exists $\delta > 0$ so that $B_\delta(x) \subset f^{-1}(G)$, see Remark 3.12. Hence $f^{-1}(G)$ is open.

Next we prove the converse: Let $a \in X$ and $\varepsilon > 0$. The set $U = f^{-1}(B_\varepsilon(f(a)))$ is open and contains a, so there exists some $B_\delta(a) \subset U$. So if $d(x,a) < \delta$, then $x \in B_\delta(a) \subset U$. Therefore $f(x) \in f(U) = B_\varepsilon(f(a))$, i.e. $d(f(x), f(a)) < \varepsilon$. $\square$

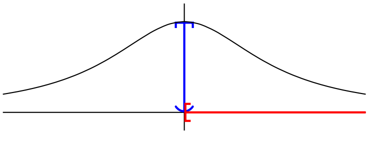
Remark 3.15. Theorem 3.14 gives an alternative to Definition 3.11 for defining continuous function. This is also the standard way of doing it in more general topological spaces (see Chapter 7). However a pointwise definition (as in Definition 3.11) is also essential. Note as well that we can substitute open sets with closed sets in Theorem 3.14, and the statement remains true.

The following two examples illustrate why one has to use preimages in Theorem 3.14.

Example 3.16. Let $f(x) = x^2$. Then $f((-1,1)) = [0,1)$. The set $[0,1)$ is not open while $(-1,1)$ obviously is. However, $f^{-1}((0,1)) = (-1,0) \cup (0,1)$. Here, both $(-1,0) \cup (0,1)$ and $(0,1)$ are open.



Example 3.17. Let $f(x) = 1/(1+x^2)$. Then the image of the closed set $[0,\infty)$ under f is $(0,1]$, which is not closed. However the preimage of the closed set $[\frac{1}{2},1]$ is $[-1,1]$, which is also closed.



Chapter 4

Compactness

In real analysis, you have likely noticed the importance of compact sets. For instance, functions defined on compact sets have more predictable properties than ones defined on other types of sets. The definition of compact sets in $\mathbf{R}^n$ (Definition 1.8) does not carry over to arbitrary metric spaces, and it is simply a special case of the "real" definition of compactness. For the general definition we first need a building block that utilizes open sets.

Definition 4.1. An *open cover* of A is a family of open sets G_λ such that

$$A \subset \bigcup_{\lambda} G_\lambda.$$

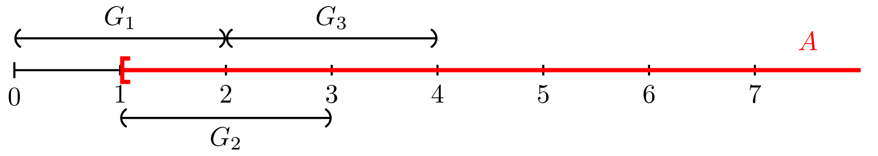
Definition 4.2. A set K is *compact* if every open cover $\{G_\lambda\}$ of K has a finite subcover $G_{\lambda_1} \cup ... \cup G_{\lambda_N}$ that also covers K.

Although this definition is seemingly very different from Definition 1.8, it is equivalent in the case of $\mathbf{R}^n$. Moreover, we can see that proving that a set is not compact is easier than proving that it is compact. All we need when proving non-compactness is a single example of an open cover that lacks a finite subcover, while to prove compactness, we need to show that *every* open cover has a finite subcover.

Example 4.3. The set $A = [1, \infty)$ has an open cover such that

$$A \subset \bigcup_{n=1}^{\infty} G_n, \quad \text{where } G_n = (n-1, n+1).$$

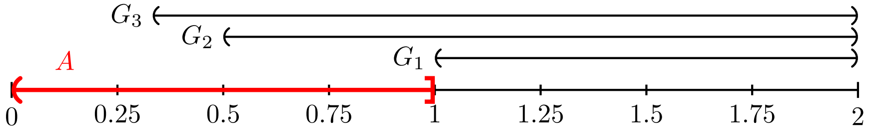
This open cover has no finite subcover that contains A, therefore A is not compact. Note that A is closed but not bounded.



Example 4.4. The set $A = (0, 1]$ has an open cover such that

$$A \subset \bigcup_{n=1}^{\infty} \left(\frac{1}{n}, 2\right).$$

This time, A is bounded but not closed. However, there is still no finite subcover. For any m, the points in the interval $(0, \frac{1}{m}]$ will not be covered by $\bigcup_{n=1}^{m} \left(\frac{1}{n}, 2\right)$. A is again not compact.



Next we define another type of compactness based on the convergence of subsequences, which we first need to define.

Definition 4.5. A *subsequence* of a sequence (x_n) is a sequence (y_k) such that $y_k = x_{n_k}$, where $n_1 < n_2 < ...$ is a sequence of indices.

Intuitively, we can think of a subsequence of (x_n) as a sequence constructed by removing some of the elements of (x_n) and keeping the relative order of the remaining elements.

Example 4.6. The sequence of the even positive integers and the sequence of prime numbers are subsequences of the natural numbers $\mathbf{N}$.

Definition 4.7. A set K is *sequentially compact* if every sequence in K has a subsequence that converges to a limit in K.

We will now examine how compactness and sequential compactness relate to each other.

Theorem 4.8. *Let X be a metric space. A set $K \subset X$ is compact if and only if it is sequentially compact.*

The proof for this theorem is very tedious and is outside the scope of this book. The interested reader can find the proof on pp. 149–150 in [3] under Theorem 9.14.

For the special case of $X = \mathbf{R}^n$ in the previous theorem, we get an additional equivalent property, as previously mentioned. This is formulated in a theorem that is very important in real analysis.

Theorem 4.9. (The Heine[1]–Borel[2]–Lebesgue[3] theorem)
For a set $K \subset \mathbf{R}^n$, the following are equivalent:

- *K is compact,*
- *K is sequentially compact,*
- *K is closed and bounded.*

The proof for this theorem will also be omitted but can be found in the same place as the proof for the previous theorem, pp. 149–150 in [3] under Theorem 9.14.

Remark 4.10. In [1], Abbott refers to *sequential compactness* simply as *compactness*. Since he is only concerned with $\mathbf{R}$, these are equivalent but it is still important to note that they are not the same. Both Definition 4.2 and 4.7 apply in more general spaces than metric spaces, namely topological spaces, in which they are not equivalent. More on this in Chapter 7.

Although a set being closed and bounded does not imply compactness in general metric spaces, compactness does imply closedness and boundedness.

Theorem 4.11. *Let X be a metric space. If $K \subset X$ is compact then K is closed.*

Proof. Let K be a compact set in a metric space and $a \in K^c$. For every $x \in K$, let $r_x = \frac{1}{2}d(x, a)$. The collection $\{B_{r_x}(x)\}_{x \in K}$ forms an open cover of K, and since K is compact there exists a finite subcover $\{B_{r_{x_j}}(x_j)\}_{j=1}^n$ that also covers K. Let $r_0 = \min_{1 \leq j \leq n} r_{x_j} > 0$. Then $B_{r_0}(a) \cap K = \varnothing$, and thus a is not a limit point to K. Hence K contains all of its limit points and is closed by Definition 2.16. $\square$

Theorem 4.12. *Let X be a metric space. If $K \subset X$ is compact then K is bounded.*

[1] Eduard Heine (1821–1881), German mathematician.
[2] Émile Borel (1871–1956), French mathematician.
[3] Henri Lebesgue (1875–1941), French mathematician.

Proof. Let $K \subset X$ be compact. Take the family of concentric open balls $\{B_n(a)\}$, where $a \in X$ is arbitrary. This is an open cover of the whole space, hence it is also an open cover of K. Since K is compact, this open cover has a finite subcover

$$\bigcup_{n=1}^{N} B_n(a) = B_N(a) \supset K.$$

Hence K is bounded by Definition 2.33.

□

Example 4.13. Here, we give an example of a closed and bounded subset of a metric space that is not compact nor sequentially compact. The space

$$\ell^2 = \left\{ x = (x_1, x_2, ...) \colon x_k \in \mathbf{R},\ \|x\|_2 = \left(\sum_{k=1}^{\infty} x_k^2 \right)^{1/2} < \infty \right\}$$

with $d(x, x') = \|x - x'\|$ is a complete metric space of infinite dimension. Take the set $E = \{e_1, e_2, ...\} \subset \ell^2$, where $e_1 = (1, 0, ...), e_2 = (0, 1, ...)$, etc. Note that $\|e_n\|_2 = 1$ for all n, so E is bounded. E has no limit points, neither in E nor outside of E, hence E is also closed. However, the sequence $(e_n)_{n=1}^{\infty} \subset E$ has no convergent subsequence so E is not sequentially compact. So by Theorem 4.8, E is not compact either. This can also be seen directly, take for instance the open cover

$$\bigcup_{n=1}^{\infty} B_1(e_n) \supset E,$$

which lacks a finite subcover.

Next we shall formulate several theorems relating to properties of compact sets. These will illustrate how important and useful the property of compactness is, both in general metric spaces and in the specific case of $\mathbf{R}$.

Theorem 4.14. (Cantor's encapsulation theorem for compact sets)
If $K_1 \supset K_2 \supset ...$ are non-empty compact subsets of a metric space, then

$$\bigcap_{n=1}^{\infty} K_n \neq \varnothing.$$

Proof. Choose $x_n \in K_n$ for each $n \in \mathbf{N}$. Because of the nesting, all x_n are in K_1 and thus there exists a subsequence (x_{n_k}) that converges to $x \in K_1$. Since $n_k > n_{k-1}$ we get that $n_k \geq k$ and thus $x_{n_k} \in K_m$ for all $k \geq m$. As K_m is closed, by Theorem 4.11, we get that $x \in K_m$ for all m. Therefore $x \in \bigcap_{m=1}^{\infty} K_m$, so it is non-empty. □

Theorem 4.15. *If $K \subset X$ is compact and $E \subset K$ is closed, then E is compact.*

Proof. Let $\{G_\lambda\}_{\lambda\in\Lambda}$ be an arbitrary open cover of E. Since E is closed, E^c is open by Theorem 2.26, and is, trivially, an open cover of itself. Thus, $\{E^c\} \cup \{G_\lambda : \lambda \in \Lambda\}$ is an open cover of K, which has a finite subcover $\{E^c\} \cup \{G_j\}_{j=1}^N$ since K is compact. Then $\{G_j\}_{j=1}^N$ is a finite subcover of $E \subset K$, and hence E is compact. □

Theorem 4.16. *If $f : A \to \mathbf{R}$ is continuous and $K \subset A$ is compact, then $f(K)$ is compact.*

Proof. Let $\{G_\lambda\}$ be an open cover of $f(K)$. Then $\{f^{-1}(G_\lambda)\}$ is an open cover of K by Theorem 3.14 as f is continuous. Since K is compact, there exists some finite subcover $\{f^{-1}(G_j)\}_{j=1}^N$ of K which means that $\{G_j\}_{j=1}^N$ is a finite subcover of $f(K)$. □

The next theorem will be very familiar from single variable analysis. However, the formulation of the theorem you have seen previously will most likely have only regarded functions of a real variable. Here we will give the theorem in its most general form.

Theorem 4.17. (Extreme value theorem)
If a function $f : K \to \mathbf{R}$ is continuous and K is compact, then there exist $x_1, x_2 \in K$ such that $f(x_1) \le f(x) \le f(x_2)$ for all $x \in K$.

Proof. Since f is continuous and K is compact, it follows from Theorem 4.16 that $f(K)$ is compact. As $f(K) \subset \mathbf{R}$ it is closed and bounded, by Theorem 4.9. Hence $\inf f(K) \in f(K)$ and $\sup f(K) \in f(K)$. □

Example 4.18. Let $f(x) = 1/x$ with $x \in (0, 1)$. The set $(0, 1)$ is not compact, and f lacks both maximum and minimum.

Example 4.19. Let

$$f(x) = \begin{cases} x, & \text{if } 0 \le x < 1, \\ 0, & \text{if } x = 1. \end{cases}$$

f is defined on a compact set but is not continuous, and lacks a maximum.

Theorem 4.20. *E is compact in $Y \subset X$ if and only if E is compact in X, where X and Y have the same metric.*

Proof. Sequential compactness is an inner property, i.e. it is not dependent on anything outside of E, so it is automatically the same for X and Y. By Theorem 4.8 the same is true for compactness. □

Chapter 5

Uniform Convergence

In Chapter 3 we discussed the convergence of sequences in metric spaces. We will now return to this subject, specifically looking at sequences in metric function spaces. New tools will be introduced which will help us study sequences of functions. We begin by defining two new types of convergence.

Definition 5.1. Let $f, f_n \colon X \to \mathbf{R}$. The sequence $f_n \to f$ *pointwise* if for every $x \in X$ and for every $\varepsilon > 0$, there exists some $N \in \mathbf{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $n \geq N$. Expressed with quantifiers,

$$\underline{\forall x \in X}\ \ \underline{\forall \varepsilon > 0\ \ \exists N \in \mathbf{N}} \colon n \geq N \implies |f_n(x) - f(x)| < \varepsilon. \tag{5.1}$$

Definition 5.2. Let $f, f_n \colon X \to \mathbf{R}$. The sequence $f_n \to f$ *uniformly* (sometimes written $f_n \rightrightarrows f$) if for every $\varepsilon > 0$ there exists some $N \in \mathbf{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in X$ and all $n \geq N$. Expressed with quantifiers,

$$\underline{\forall \varepsilon > 0\ \ \exists N \in \mathbf{N}}\ \ \underline{\forall x \in X} \colon n \geq N \implies |f_n(x) - f(x)| < \varepsilon. \tag{5.2}$$

Note that the difference between pointwise and uniform convergence is seen in the positioning of the quantifiers in (5.1) and (5.2).

The function sequence (f_n) converging to f uniformly means that, by (5.2), for every $\varepsilon > 0$ there exists N such that

$$\|f_n - f\|_\infty = \sup_{x \in X} |f_n(x) - f(x)| \leq \varepsilon \quad \text{if } n \geq N,$$

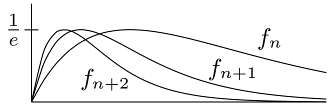
and thus $f_n \to f$ uniformly if and only if $\|f_n - f\|_\infty \to 0$. Uniform convergence is exactly convergence in the supremum norm.

Example 5.3. Let $f_n(x) = nxe^{-nx}$, defined on $A = [0, \infty)$. Since $x \geq 0$, $f_n(x) \to 0$ as $n \to \infty$ for every $x \in A$. Therefore $f_n \to 0$ pointwise. Notice that

$$f_n'(x) = ne^{-nx} - n^2 e^{-nx} = n(1 - nx)e^{-nx}.$$

So $\|f_n - 0\|_\infty = \sup_{x \geq 0} |f_n(x)| = \frac{1}{e} \not\to 0$ and thus $f_n \not\to 0$ uniformly.

x		$\frac{1}{n}$	
f_n'	$+$	0	$-$
f_n	$\nearrow$	$\frac{1}{e}$	$\searrow$



Example 5.4. The same function as in Example 5.3 but defined on $A_\delta = [\delta, \infty)$, $\delta > 0$. If $n \geq 1/\delta$, then f_n is decreasing on A_δ and thus

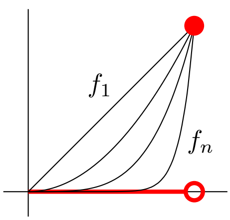
$$\sup_{x \in A_\delta} |f_n(x)| = f_n(\delta) \to 0.$$

So $f_n \to f$ uniformly on A_δ for all $\delta > 0$.

Example 5.5. Let $f_n(x) = x^n$, defined on $A = [0, 1]$. Then

$$f_n(x) \to f(x) := \begin{cases} 1, & \text{if } x = 1, \\ 0, & \text{if } 0 \leq x < 1, \end{cases}$$

but f is not continuous. We see that $\|f_n - f\|_\infty = 1$ for all n and therefore $f_n \not\to f$ uniformly.



As you know, in analysis we are particularly interested in continuous functions. Thus is it of great interest whether or not different types of convergence of function sequences preserve continuity.

Theorem 5.6. *If f_n is continuous at $a \in A$ and $f_n \to f$ uniformly then f is continuous at a.*

Proof. Let $\varepsilon > 0$. Then for every $x \in A$,

$$|f(x) - f(a)| \leq \underbrace{|f(x) - f_n(x)|}_{<\frac{\varepsilon}{3} \text{ if } n \geq N} + \underbrace{|f_n(x) - f_n(a)|}_{<\frac{\varepsilon}{3} \text{ if } d(x,a)<\delta} + \underbrace{|f_n(a) - f(a)|}_{<\frac{\varepsilon}{3} \text{ if } n \geq N}. \tag{5.3}$$

Since $f_n \to f$ uniformly, for every $\varepsilon > 0$ there exists some $N \in \mathbf{N}$ such that $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$ for all $n \geq N$ and all $x \in A$. Thus, the first and the third terms in (5.3) are bounded from above by $\frac{\varepsilon}{3}$ for all $n \geq N$.

Since f_n is continuous, for every $\varepsilon > 0$ there exists some $\delta > 0$ such that $|f_n(x) - f_n(a)| < \frac{\varepsilon}{3}$ whenever $d(x, a) < \delta$. Thus, the second term in (5.3) is also bounded from above by $\frac{\varepsilon}{3}$ when $d(x, a) < \delta$.

If we choose $n \geq N$ and then $\delta > 0$, we see that $|f(x) - f(a)| < \varepsilon$ whenever $d(x, a) < \delta$, and so f is continuous at a. □

Remark 5.7. If $f_n \to f$ pointwise and f and all f_n are continuous, then

$$\lim_{n\to\infty} \lim_{x\to a} f_n(x) = \lim_{x\to a} \lim_{n\to\infty} f_n(x)$$

since both sides evaluate to $f(a)$. But in Example 5.5,

$$\lim_{x\to 1^-} x^n = 1 \neq 0 = \lim_{x\to 1^-} \lim_{n\to\infty} x^n.$$

Recall 5.8. Convergence in the supremum norm is equivalent to uniform convergence.

We now return to the proof of Theorem 3.7 as we have developed the necessary tools.

Proof of Theorem 3.7. Let (f_n) be Cauchy in $C([a, b])$, i.e.

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbf{N} \colon n, m \geq N \implies |f_m(x) - f_n(x)| < \varepsilon \tag{5.4}$$

is true for all $x \in [a, b]$. Therefore $(f_n(x))$ is also Cauchy in $\mathbf{R}$ for every x, and hence it is also convergent since $\mathbf{R}$ is complete. Let

$$f(x) := \lim_{n\to\infty} f_n(x).$$

Now we need to show that $f_n \to f$ uniformly: Let $\varepsilon > 0$. Take N to be the same as in (5.4), then fix $n \geq N$ and let $m \to \infty$. This gives, if $n \geq N$, that

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{for all } x \text{ and all } m \geq N,$$

and hence

$$|f_n(x) - f(x)| \leq \varepsilon \quad \text{for all } x,$$

i.e. $f_n \to f$ uniformly. □

Since we did not use the fact that $f_n \in C([a,b])$ in the proof of Theorem 3.7, we can make an even stronger statement.

Theorem 5.9. (Cauchy's criterion)
The sequence (f_n) converges uniformly if and only if (f_n) is Cauchy with respect to the supremum norm.

Example 5.10. Let $f_n(x) = \sin(2^n \pi x)$. Then (f_n) is bounded in $C([0,1])$. Fix $m > n \geq 1$ and let $x = \frac{3}{2^{m+1}}$. Then

$$f_m(x) = \sin(2^m \pi x) = \sin\Big(\frac{3\pi}{2}\Big) = -1.$$

But $0 < 2^n \pi x \leq \frac{3\pi}{4}$ so $f_n(x) = \sin(2^n \pi x) > 0$. Hence

$$\|f_n - f_m\|_\infty \geq |f_n(x) - f_m(x)| > 1, \tag{5.5}$$

i.e. (f_n) is not Cauchy and therefore not convergent, by Theorem 3.4. Moreover, it also follows from (5.5) that (f_n) has no convergent subsequence.

Remark 5.11. Let $B = \{f \in C([0,1]) \colon \|f\|_\infty \leq 1\}$ be the closed unit ball in $C([0,1])$. This is a closed and bounded set which by Example 5.10 is not sequentially compact, hence it is not compact either.

Next, we present a famous theorem relating two properties of sequences of functions that guarantee uniform convergence of their subsequences. However, we first need to give the following definition.

Definition 5.12. A family of functions $f_n \colon X \to Y$ is *equicontinuous* at a if for every $\varepsilon > 0$ there exists some $\delta > 0$ such that for all n, $d_Y(f_n(a), f_n(x)) < \varepsilon$ whenever $d_X(a,x) < \delta$.

Theorem 5.13. (The Arzelà[1]–Ascoli[2] theorem)
If $f_n \colon [a,b] \to \mathbf{R}$ is an equicontinuous family of function and $|f_n(x)| \leq M$ for all n (uniformly bounded) and all $x \in [a,b]$, then (f_n) has a uniformly convergent subsequence.

Conversely, if every subsequence of (f_n), where $f_n \in C([a,b])$, has a uniformly convergent subsequence, then (f_n) is uniformly bounded and equicontinuous.

[1] Cesare Arzelà (1847–1912), Italian mathematician.
[2] Giulio Ascoli (1843–1896), Italian mathematician.

Chapter 6

Curves and Connectedness

6.1 Curves and Pathconnectedness

In your previous analysis courses, you have probably encountered curves. This might have been in your multivariable analysis or perhaps in a course in complex analysis. Now, we will revisit them as we have developed some tools that will help us analyse them. We begin by defining curves in general metric spaces.

Definition 6.1. A *curve* $\gamma\colon I \to X$, where X is a metric space, is a continuous function on an interval $I \subset \mathbf{R}$.

This allows us to depart from the usual geometric interpretation of curves. With the inherent distance function of our metric space, we can talk about how long a curve is.

Definition 6.2. Consider the curve $\gamma\colon [a,b] \to X$ where X is a metric space with the distance function d. The *length* of γ is,

$$\ell(\gamma) = \sup \sum_{j=1}^{n} d(\gamma(x_j), \gamma(x_{j-1}))$$

over all partitions $a = x_0 < x_1 < ... < x_n = b$ of $[a,b]$. We say that γ is *rectifiable* if $\ell(\gamma) < \infty$.

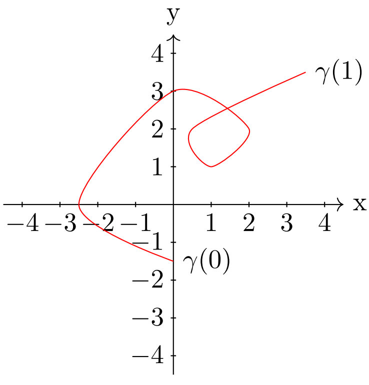
A rectifiable curve can be parameterized by arc length, denoted ds, so that

$$\ell(\gamma|_{[s,t]}) = t - s$$

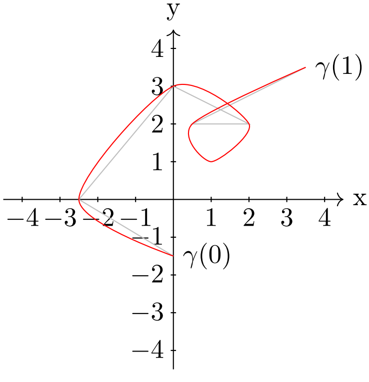
for all $a \leq s < t \leq b$. Then for all $s < t$,

$$d(\gamma(t), \gamma(s)) \leq \ell(\gamma|_{[s,t]}) = t - s.$$

Example 6.3. A curve $\gamma\colon [0,1] \to \mathbf{R}^2$.



We can get a lower bound on the length of the curve by drawing and measuring the length of straight lines between some of the points it crosses through. We will use the ℓ^2-distance for $\mathbf{R}^2$.



The cumulative length of the gray lines is approximately 13.92 units. So $\ell(\gamma) \geq 13.92$.

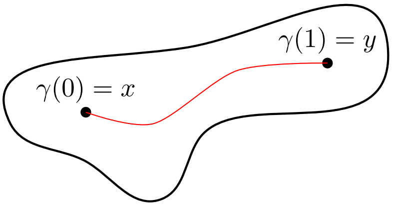
The image of a curve γ is a set in a metric space X. In mathematics, it is common that we want to integrate over such a set. We can do this using arc length parameterization.

Definition 6.4. A *curve integral*, or *path integral*, with respect to arc length on the arc length parameterized curve $\gamma\colon [a,b] \to X$ is defined by,

$$\int_\gamma f\, ds := \int_a^b f(\gamma(t))\, dt.$$

With curves, we can now express a notion of connectedness within sets in metric spaces.

Definition 6.5. X is *pathconnected*, or *arcwise connected*, if for every pair of points $x, y \in X$, there exists a curve $\gamma\colon [0,1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$.



If you have taken a course in complex analysis you may recognize this type of connectedness. Pathconnected sets are very important in complex analysis as curves and curve integrals are essential to the subject.

6.2 Connectedness

We have presented one way of describing connectedness in metric spaces. Now we present another way to express this notion of connectedness, this time using open sets as opposed to curves.

Definition 6.6. A metric space X is *disconnected* if there exists two sets, A and B, which are non-empty, open and disjoint (i.e. $A \cap B = \varnothing$), such that the union of A and B makes up the entire space, $X = A \cup B$. A metric space is *connected* if it is not disconnected.

Example 6.7. Let $X = \mathbf{Q}$, and let $A = \mathbf{Q} \cap (-\infty, \sqrt{2})$ and $B = \mathbf{Q} \cap (\sqrt{2}, \infty)$. Both A and B are non-empty and open. They are also disjoint since they share no points. Finally, A and B together make up the entire space, $\mathbf{Q} = A \cup B$. Hence $\mathbf{Q}$ is disconnected.

Remark 6.8. A set $A \subset X$ is *connected* in the metric space (X, d) if it is connected as a metric space (A, d) in its own right.

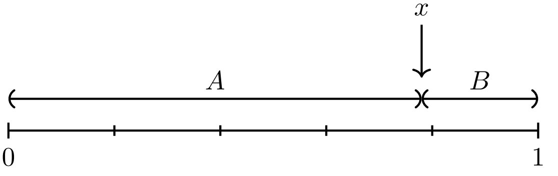
Next we regard the case of subsets of $\mathbf{R}$. The result presented in the next theorem will be very intuitive if we think about connectedness in a geometric sense.

Theorem 6.9. *A set $E \subset \mathbf{R}$ is connected if and only if it is either an interval, a single point set, $\varnothing$ or $\mathbf{R}$.*

The implication $P \Rightarrow Q$ is equivalent to its contrapositive statement $\neg Q \Rightarrow \neg P$, which will be used in the proof below.

Proof. First we show that a connected set $E \subset \mathbf{R}$ must be either an interval, a single point set, $\varnothing$ or $\mathbf{R}$. We prove the contrapositive statement: If a set E is not an interval, a single point set, $\varnothing$ or $\mathbf{R}$, then there exist $a < b < c$ such that $a, c \in E$ and $b \notin E$. Now, let $A = E \cap (-\infty, b)$ and $B = E \cap (b, \infty)$. Then A and B are non-empty, relatively open and disjoint, and they also make up the entire set, $E = A \cup B$. Hence E is disconnected.

Next we show that $E = [0, 1]$ is connected. The other sets are handled similarly. We give a proof by contradiction: Assume that E is disconnected. Then there exist A and B that are non-empty, relatively open and disjoint such that $E = A \cup B$. Now we may assume that $1 \in B$ (if this is false, just change names between A and B). Furthermore, let $x = \sup A$ (which exists by the supremum axiom since A is non-empty and bounded). Then x is a limit point of A and since A is relatively closed (as $B = E \setminus A$ is relatively open), we conclude that $x \in A$. This forces $x < 1$. However, since A is also relatively open, there exists some $\varepsilon > 0$ such that $x + \varepsilon \in A$. This directly contradicts the statement that $x = \sup A$. $\square$



Now we will examine some of the cases where connectedness is preserved. The next two theorems will explore preservation under two types of operations on sets.

Theorem 6.10. *If $E \subset \mathbf{R}^n$ is connected, then $\overline{E}$ is connected.*

We leave the proof of Theorem 6.10 as an exercise to the reader.

Theorem 6.11. *Let X and Y be metric spaces. If $f\colon X \to Y$ is continuous and X is connected, then $f(X)$ is connected.*

Proof. We prove the contrapositive statement: Assume $f(X)$ is disconnected. Then there exist non-empty, open and disjoint A and B such that $A \cup B = f(X)$. Let $C = f^{-1}(A)$ and $D = f^{-1}(B)$. Then C and D are non-empty, disjoint and $X = C \cup D$. As C and D are also open by Theorem 3.14, X is disconnected. □

A theorem you have most likely seen before in your first year single variable analysis course is the intermediate value theorem. We can now prove it using connectedness.

Theorem 6.12. (Intermediate value theorem)
If $f\colon [a,b] \to \mathbf{R}$ is continuous and $f(a) < m < f(b)$, then there exists some $x \in (a,b)$ such that $f(x) = m$.

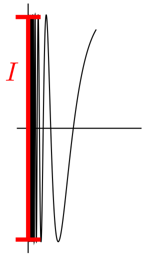
Proof. The set $f([a,b])$ is connected by Theorems 6.9 and 6.11, and so it is an interval I according to Theorem 6.9. As $f(a), f(b) \in I$, we have also $m \in I$ and thus there exists some $x \in [a,b]$ such that $f(x) = m$. □

Now that we have defined connectedness and present the necessary theory, we can return to pathconnectedness to formulate two theorems relating these two concepts. We look at their relation both in general metric spaces and in the case of $\mathbf{R}^n$.

Theorem 6.13. *If X is pathconnected, then X is also connected.*

Proof. We give a proof by contradiction: Assume X is disconnected. Then there exist A and B such that they are non-empty, open, disjoint and $X = A \cup B$. Now let $a \in A$ and $b \in B$. There exists a curve $\gamma\colon [0,1] \to X$ such that $\gamma(0) = a$ and $\gamma(1) = b$. Let $E = \gamma([0,1])$. Then E is connected by Theorems 6.9 and 6.11. However, the sets $A' = A \cap E$ and $B' = B \cap E$ are non-empty, relatively open, disjoint and $E = A' \cup B'$ which means that E is disconnected. □

Example 6.14. The following is an example of a set that is connected but not pathconnected. Let $E = \{(x, \sin\frac{1}{x}) \colon 0 < x \leq 1\}$. Since E is pathconnected (this is quite clear as it is a curve) it is also connected by Theorem 6.13. Thus $\overline{E} = E \cup I$ is connected by Theorem 6.10, where $I = \{0\} \times [-1,1]$. However, there are no curves between E and I and thus $\overline{E}$ is not pathconnected.



Theorem 6.15. *If $G \subset \mathbf{R}^n$ is open, then G is connected if and only if G is pathconnected.*

Proof. The implication that an open set $G \subset \mathbf{R}^n$ is connected if it is pathconnected is true by Theorem 6.13. Proving the converse statement, however, is very tricky and outside the scope of this book. $\square$

Chapter 7

Topological Spaces

7.1 Definition of a Topological Space

Much like when we defined metric spaces, we want to consider more general spaces then we previously had by means of fewer constraints. This time however, the structure of the space is not defined by the notion of *distance* but instead the notion of *openness*.

First, we want the entire space and the empty set to be open. Second, we would like arbitrary unions of open sets to remain open and lastly, finite intersections of open sets should also remain open.

Definition 7.1. A *topological space* $X = (X, \tau)$ is a set X together with a *topology* τ. A topology τ on X is a set of subsets to X such that

(i) $\varnothing$ and X are in τ,

(ii) a union of sets that are in τ, is in τ,

(iii) a finite intersection of sets that are in τ, is in τ.

Sets that are in τ are called *open*.

Almost all spaces you have encountered before are topological spaces, you might have just not recognized them as such. As an example, all metric spaces with $\tau = \{G \colon G \text{ is open}\}$ (using the definition of open sets from Definition 2.8), are topological spaces, by Theorems 2.10 and 2.11. If a topology can be defined using a distance function, we call that topology *metrizable*. We also say that the distance function *induces* the topology. Note that different distance functions can induce the same topology.

Example 7.2. Let $X = \{a, b, c\}$ and $\tau = \{\varnothing, X, \{a\}, \{b\}, \{b, c\}\}$. Is (X, τ) a topological space?

First we conclude that both $\varnothing$ and the whole space X, are open. Next, we can examine different intersections. Both $\{a\} \cap \{b\}$ and $\{a\} \cap \{b, c\}$ are equal to the empty set, which is open, and $\{b\} \cap \{b, c\} = \{b\}$ is also open. Lastly, we examine have the unions. The union $\{a\} \cup \{b\} = \{a, b\}$ is not open, and hence (X, τ) is not a topological space.

Remark 7.3. There are a few topologies that qualify a space as a topological space but which are not very interesting. The first one of these is the *discrete topology* which is the set of all subsets. Another one is the *trivial topology* which only contains $\varnothing$ and the set itself.

We continue by defining the convergence of sequences in arbitrary topological spaces. This definition is quite different from the one given for metric spaces in Definition 3.1 and thus gives rise to new properties which we will examine in more detail later on. First however, we need a new tool.

Definition 7.4. A *neighborhood* $V \subset$ of a point $a \in X$, where X is a topological space, is a subset of X that contains an open set G such that,

$$a \in G \subset V.$$

Definition 7.5. Let (X, τ) be a topological space. If (a_n) is a sequence in X and $L \in X$, then a_n *converges* to L, written $a_n \to L$, if for every open neighborhood G of L, there exists an $N \in \mathbf{N}$ such that $a_n \in G$ for all $n \geq N$.

Next, we define pointwise continuity of functions in arbitrary topological spaces.

Definition 7.6. Let (X, τ) and (X', τ') be topological spaces. A function $f\colon X \to X'$ is *continuous* at $a \in X$ if for every open neighborhood G' of $f(a)$ there exists some open neighborhood G of a such that $f(x) \in G'$ whenever $x \in G$. Expressed with quantifiers, f is continuous at $a \in X$ if

$$\forall\, G' \in \tau' \text{ such that } f(a) \in G' \;\; \exists\, G \in \tau \text{ such that } a \in G,$$

and

$$x \in G \implies f(x) \in G'.$$

Remark 7.7. This definition is very similar, and completely equivalent, to Remark 3.12 if X is a metric space, as every open neighborhood must contain an open ball around the point. We can take the set G' to be the open ball $B^Y_\varepsilon(f(a))$ as in the remark. Continuity implies the existence of the ball $B^X_\delta(a)$ as in Remark 3.12. This ball in turn qualifies as the open neighborhood G. Similarly, Definition 7.5 is equivalent to Definition 3.1 if X is a metric space.

We can also define continuity of a function on an entire space. This is similar to Theorem 3.14 just like we mentioned in Remark 3.15.

Definition 7.8. Let (X, τ) and (X', τ') be topological spaces. A function $f\colon X \to X'$ is *continuous* if $f^{-1}(G)$ is open in X for all open $G \subset X'$.

Just like we studied the interior and closure of sets in metric spaces, we can talk about them in topological spaces. However, instead of using open balls, we only use open and closed sets together with intersections and unions.

As previously stated, open sets in a topological space are the sets that are in the topology. Closed sets are defined by Theorem 2.26, i.e. a set is *closed* if its complement is open. This means that in all topological spaces, the whole space and the empty set are both open and closed at the same time, they are so-called *clopen* sets.

Definition 7.9. Let $Y \subset X$ where X is a topological space.

- The *closure* of Y, written $\overline{Y}$, is the intersection of all closed sets that contain Y.
- The *interior* of Y, written Y°, is the union of all open sets that are contained in Y.

Note that $Y^\circ \subset Y \subset \overline{Y}$. Definition 7.9 is completely equivalent to Definitions 2.13 and 2.22 in metric spaces, by Theorems 2.14 (b) and 2.23 (d), even though they look very different. We can however define interior just like we did for metric spaces if we first define inner points in topological spaces.

Definition 7.10. Let $A \subset X$, where X is a topological space. Then $a \in A$ is an *inner point* of A if $a \in G \subset A$ for some open set G.

Another category of points from earlier in this book, whose definition we would like to extend to topological spaces, is isolated points.

Definition 7.11. Let $A \subset X$, where X is a topological space. A point $a \in A$ is *isolated* in A if there exists a neighborhood of a which does not contain any other points of A.

We also revisit the definition of the boundary of a set but now in arbitrary topological spaces.

Definition 7.12. Let $A \subset X$, where X is a topological space. The *boundary* of A is the set of all points in the closure of A that are not in the interior of A, i.e. $\partial A = \overline{A} \setminus A^\circ$.

7.2 Unique and Non-unique Limits

In metric spaces, we know that the limit of a sequence is unique, i.e. a sequence can at most converge to one point. This is not generally true in topological spaces. We illustrate this with an example.

Example 7.13. Let $X = \{a, b\}$ be a topological space with $\tau = \{X, \varnothing, \{a\}\}$. Since $\{a\}$ and $X = \{a, b\}$ are all of the open sets which contain a, we can conclude that $a_n \to a$ and $b_n \not\to a$, where $a_n = a$ and $b_n = b$ for all n. $X = \{a, b\}$ is the only open set which contains b, so $a_n \to b$. Therefore (a_n) converges to both a and b which cannot be regarded as equivalent since $b_n \not\to a$.

If we let $K = \{a\}$ then K is compact (as all finite sets are compact), but since $K^c = \{b\}$ is not open we see that K is not closed. Note that in metric spaces, compact sets are closed and bounded (but these two properties are not enough to guarantee compactness), see also Theorem 4.9.

To keep sequences in topological spaces from converging to multiple points, we can put in place a constraint on the relationship between open sets around the points in question. We define a type of space with this constraint.

Definition 7.14. A topological space (X, τ) is a *Hausdorff space* if for every pair of points $a, b \in X$, where $a \neq b$, there exist open sets A and B such that $a \in A$, $b \in B$ and $A \cap B = \varnothing$.

Intuitively, we can see this as ensuring that all points are sufficiently far away from each other.

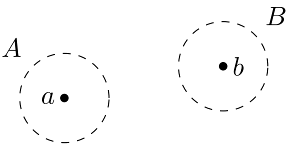
Theorem 7.15. *If $X = (X, \tau)$ be a Hausdorff space, then every convergent sequence in X has a unique limit.*

Proof. We give a proof by contradiction: Let $X = (X, \tau)$ be a Hausdorff space and (x_n) be a convergent sequence in X. Now assume that $\lim_{n\to\infty} x_n = a$ and $\lim_{n\to\infty} x_n = b$ where $a \neq b$. Since X is Hausdorff, there exist open disjoint sets A and B such that $a \in A$ and $b \in B$. From Definition 7.5, we know that there exist N_a and N_b such that $x_n \in A$ when $n \geq N_a$ and $x_n \in B$ when $n \geq N_b$. Now let $n \geq \max\{N_a, N_b\}$ so that x_n is in both A and B. But since $A \cap B = \varnothing$, this is impossible. $\square$

As previously stated, we know that sequences in metric spaces are unique but now we will prove it using Hausdorff spaces.

Theorem 7.16. *Every metric space is a Hausdorff space.*

Proof. Let (X, d) be a metric space and $a, b \in X$ where $a \neq b$. Then $A := B_r(a)$ and $B := B_r(b)$, with $r = \frac{1}{2}d(a, b)$, are open and disjoint. Hence the space is Hausdorff. $\square$



7.3 L^p-Spaces

For $1 \leq p < \infty$, consider

$$\tilde{L}^p([0,1]) = \left\{ f\colon \int_0^1 |f|^p \, dx < \infty \right\},$$

where f is *measurable* (see p. 127 in [2]) and defined on $[0, 1]$, together with the distance function

$$\tilde{d}_p(f, g) = \left(\int_0^1 |f - g|^p \, dx \right)^{1/p}, \tag{7.1}$$

where $f, g \in \tilde{L}^p([0, 1])$. This does not qualify as a metric space since there are many different functions that have the distance 0 between each other. As an example, consider the characteristic function at the point $1/2$,

$$\delta_{1/2} = \begin{cases} 1, & \text{if } x = 1/2, \\ 0, & \text{if } x \neq 1/2. \end{cases}$$

The distance between $\delta_{1/2}$ and the constant function $f(x) = 0$ will be zero even though they are not identical.

Instead we call this type of space a *semi-metric* space. It is, simply put, a metric space where the constraint (i) in Definition 2.1 is relaxed as we no longer require the distance between two distinct points to be non-zero. We still require that $d(x, x) = 0$ for every $x \in X$.

This semi-metric space does qualify as a topological space where the topology is induced by the semi-metric (7.1). A set $G \subset \tilde{L}^p([0, 1])$ is *open* if for every $f \in G$, there exists some $\varepsilon > 0$ such that $g \in G$ whenever $\tilde{d}_p(f, g) < \varepsilon$.

In order to create a metric space out of this set and distance function, we have to introduce equivalence classes.

Definition 7.17. A function f is *equivalent* to another function g, written $f \sim g$, if $\tilde{d}(f, g) = 0$.

Remark 7.18. Notice that this equivalence relation is reflexive, symmetrical and transitive, i.e.

(a) $f \sim f$,

(b) if $f \sim g$ then $g \sim f$,

(c) if $f \sim g$ and $g \sim h$, then $f \sim h$.

Definition 7.19. The *equivalence class* of f is a set such that

$$[f] = \{g \colon g \sim f\}.$$

Intuitively, we group together all functions that have zero $\tilde{d}$-distance to f, and by transitivity to each other.

Remark 7.20. If $\tilde{d}_p(f, g) = 0$ and $f_n \to f$ then $f_n \to g$. Thus we do not have unique limits in $\tilde{L}^p([0, 1])$.

With this equivalence class we can adjust $\tilde{L}^p([0, 1])$ slightly so that it becomes a proper metric space. The set

$$L^p([0, 1]) = \{[f] \colon f \in \tilde{L}^p([0, 1])\}$$

together with the L^p-norm

$$d([f], [g]) = \tilde{d}_p(f, g)$$

qualifies as a metric space since all functions that previously had zero distance between them are now one single element; an equivalence class. By Theorems 7.15 and 7.16, limits in $L^p([0, 1])$ are unique.

Remark 7.21. Note that in Example 7.13 we cannot create "natural" equivalence classes (so that $a_n \to a$ if and only if $[a_n] \to [a]$) to get unique limits.

7.4 Compactness

Since the compactness of a set $K \subset X$ is immediately dependent on which sets are open in X, we take particular interest in compactness when studying topological spaces. We begin with a definition which is identical to the corresponding Definition 4.2 for metric spaces.

Definition 7.22. A topological space (X, τ) is *compact* if every open cover $\{G_\lambda\}$ of X has a finite subcover $G_{\lambda_1} \cup \ldots \cup G_{\lambda_N}$ that also covers X.

Additionally, we can talk about sequential compactness. This is also very strongly related to which sets are open when we consider the definition of convergent sequences in topological spaces. Again, we give a definition which is identical to the corresponding Definition 4.7 for metric spaces.

Definition 7.23. A topological space (X, τ) is *sequentially compact* if every sequence in X has a convergent subsequence.

We previously stated in Theorem 4.8 that in all metric spaces, compactness and sequential compactness are equivalent. This is no longer true in arbitrary topological spaces. If you would like to see an example of this, we recommend Example 13.5 in [3] which shows a sequentially compact topological space that is not compact and Example 105 in [4] that shows a compact topological space that is not sequentially compact.

7.5 Further studies

What we have shown you so far contains some of the core elements for further studies in analysis. We will end with a few more examples of topological spaces. We begin with $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$ and $\overline{\mathbf{R}^n} = \mathbf{R}^n \cup \{\infty\}$.

Example 7.24. We can construct a "natural" topology on $\overline{\mathbf{R}}$ if we let G be *open* in $\overline{\mathbf{R}}$ if the following conditions hold:

(a) $G \cap \mathbf{R}$ is open in $\mathbf{R}$.

(b) There exists a point $a \in \mathbf{R}$ such that $(a, \infty] \subset G$ if $\infty \in G$,

(c) There exists a point $a \in \mathbf{R}$ such that $[-\infty, a) \subset G$ if $-\infty \in G$.

This topology is metrizable but there is no "natural" metric.

Example 7.25. Similarly, we can construct a topology on $\overline{\mathbf{R}^n}$, by letting G be *open* in $\overline{\mathbf{R}^n}$ if

(a) $G \cap \mathbf{R}^n$ is open in $\mathbf{R}^n$,

(b) there exists a compact set K such that $\overline{\mathbf{R}^n} \setminus K \subset G$ if $\infty \in G$.

This topology is also metrizable, but there is no natural metric.

If you have taken, or will take, a course in Fourier analysis you will most likely encounter distributions, a generalization of functions. Sequences of distributions can converge and will give rise to a non-metrizable topology. Thus we can talk about sets of distributions being compact or not. This is usually not very rigorously covered in first courses on Fourier analysis but nonetheless interesting to know.

Another example is weak and weak* convergence on a Banach space X which appears in functional analysis. They give rise to the weak and weak* topologies, which are non-metrizable if X is of infinite dimension. These types of convergence are also important in e.g. probability theory and when solving certain minimization problems.

Bibliography

[1] Stephen Abbott. *Understanding Analysis*. Springer, London, 2nd edition, 2015. 6, 31

[2] Frank Burk. *Lebesgue Measure and Integration: An Introduction*. Wiley, New York, 1st edition, 1997. 49

[3] Gerard Buskes and Arnoud van Rooij. *Topological Spaces*. Springer, New York, 1st edition, 1997. 31, 51

[4] Lynn Arthur Steen and J. Arthur Seebach. *Counterexamples in Topology*. Springer, New York, 2nd edition, 1978. 51

Appendix A

English–Swedish Glossary

English	**Swedish**
Ball	Klot
Boundary point	Randpunkt
Bounded	Begränsad
Closed	Sluten
Closure	Slutet hölje
Complete	Fullständig
Connected	Sammanhängande
Dense	Tät
Disjoint	Disjunkt
Interior	Inre
Limit point	Hopningspunkt
Open cover	Öppen övertäckning
Pathconnected	Bågvis sammanhängande
Sequence	Följd
Uniform convergence	Likformig konvergens

Appendix B

Swedish–English Glossary

Swedish	**English**
Begränsad	Bounded
Bågvis sammanhängande	Pathconnected
Disjunkt	Disjoint
Fullständig	Complete
Följd	Sequence
Hopningspunkt	Limit point
Inre	Interior
Klot	Ball
Likformig konvergens	Uniform convergence
Randpunkt	Boundary point
Sammanhängande	Connected
Sluten	Closed
Slutet hölje	Closure
Tät	Dense
Öppen övertäckning	Open cover