Optimization, WS 2017/18

Lecturer – Dan Rust, Tutor – Mima Stanojkovski

Part 1. Convergence in Metric Spaces

(about 7 Lectures)

Supporting Literature: ***Angel de la Fuente*, "Mathematical Methods and Models for Economists",** Chapter 2

Contents

1 Convergence in Metric Spaces

Acknowledgements These Lecture notes were kindly provided by Dr. Tanja Pasurek. They were edited by Dr. Dan Rust and Chrizaldy Neil Manibo.

1.1 Definition of Metric and Normed Spaces

Of fundamental importance in *mathematical analysis* is the notion of *limit* and *convergence* (e.g., for real numbers, complex numbers, vectors in $\mathbb{R}^n$, functions etc.). This constitutes a basis for defining two fundamental operations: *differentiation* and *integration* of functions.

The notion of limit involves a possibility to measure a *"distance"* between the objects. Extending the well-known notion of the Euclidean distance

$d(x, y) := |x - y|$ between two reals $x, y \in \mathbb{R}$, or

$d(x, y) := \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$ between two points $x := (x_1, x_2, x_3)$, $y := (y_1, y_2, y_3) \in \mathbb{R}^3$,

we naturally come to the notion of *metric space*, which is the basic one in modern mathematics. Roughly speaking, a metric space is a set in which we have defined a distance, i.e., a *metric*. Introducing the metric, we can study the properties of limits independently on the concrete nature of objects under consideration.

History: The notion of a metric space was first introduced by the French mathematician *M. Fréchet* (1906). This area of mathematics is now seen as a part of *Functional Analysis (FA).*

Functional Analysis (as a branch of *Mathematical Analysis*) studies vector space (in general, infinite-dimensional) endowed with some kind of *limit-related structure* (e.g. *inner product*, *norm*, *metric*, *topology*, etc.) and the functions (so-called *functionals* or *operators*) acting upon these spaces and respecting these structures in a suitable sense. *(see Wiki).*

FA started out at the beginning of the last century in works of the French mathematicians Hadamard, Fréchet, Lévy, Riesz, ..., and by the group of Polish mathematicians around Stefan Banach (1892-1945). The PhD of Stephan Banach (1922) included the basic ideas of functional analysis, which was soon to become an entirely new branch of mathematics. Banach's most influential work was Théorie des opérations linéaires (Theory of Linear Operations, 1932), in which he formulated the concept now known as *"Banach spaces"* (i.e., complete normed vector space) and proved many basic theorems of *FA*. An important example of such spaces is a *Hilbert space* (named after David Hilbert, 1862-1943), where the norm arises from an inner product. These spaces are of fundamental importance in many areas.

Definition 1.1.1. *Let X be a nonempty set (i.e, a collection of objects we call elements).*

*A **metric** (or **distance** function) on X is a mapping*

$$d\colon X \times X \ni (x, y) \to d(x, y) \in \mathbb{R}_+ \quad (\geq 0)$$

with the following properties:

(i) $d(x, y) = 0 \Longleftrightarrow x = y$ *(positive definite);*

(ii) $d(x,y) = d(y,x)$, $\forall x, y \in X$ *(symmetry);*

(iii) $d(x,z) \leq d(x,y) + d(y,z)$, $\forall x, y, z \in X$ *(triangle inequality).*

*The **pair** (X,d) is called a metric space. The elements x of a metric space X are also called **points**.*

On the same set one can define different metrics.

Example 1.1.2. *The **trivial** (or **discrete**) metric on any set X:*

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

Exercise 1.1.3. *Prove that this d is a metric.*

Exercise 1.1.4. *Prove that (i)–(iii) imply the **inverse** triangle inequality*

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

Given a metric space (X,d) and a (nonempty) subset $Y \subset X$, it is clear that (Y,d) is also a metric space; it is called a metric **subspace** of X.

Definition 1.1.5. *Let V be a vector (i.e., linear) space. A **norm** is a (nonnegative) mapping*

$$\|\cdot\| : V \to \mathbb{R}_+ \quad (\geq 0)$$

obeying the following properties:

(i) $\|x\| = 0 \iff x = 0$ *(identity axiom);*

(ii) $\|x+y\| \leq \|x\| + \|y\|$, $\forall x, y \in V$ *(triangle inequality);*

(iii) $\|\alpha x\| = |\alpha| \cdot \|x\|$, $\forall$ *scalar* $\alpha \in \mathbb{R}$, $\forall x \in X$ *(homogeneity).*

*The **pair** $(V, \|\cdot\|)$ is called a **normed space**.*

Remark 1.1.6 (For Supporting material, see A. de la Fuente, p. 28.)**.** *A **vector** (or **linear**) space is a set V of elements called vectors, together with a binary operation $V \times V \to V$ called vector **addition** (and denoted by "+") and an operation $\mathbb{R} \times V \to V$ called **scalar multiplication**. These operations have the following properties:*

for all $x, y, z \in V$ and $\alpha, \beta \in \mathbb{R}$
(1) $x + y = y + x$ (commutative property);
(2) $x + (y + z) = (x + y) + z$ (associative property);
(3) $\exists!\ 0 \in V : x + 0 = 0 + x = x$ (existence of the zero element);

(4) $\forall\, x \in V$, $\exists!\ (-x) \in V : x + (-x) = 0$ *(existence of inverse elements);*
(5) $\alpha(x + y) = \alpha x + \alpha y$ *and* $(\alpha + \beta)x = \alpha x + \beta x$ *(distributive property);*
(6) $\alpha(\beta x) = (\alpha\beta)x$ *(associative law for scalars);*
(7) $1 \cdot x = x$ *(multiplicative identity).*
Then one can define the ***difference*** *operation*

$$x - y := x + (-y) \text{ where } -y := (-1)y,$$

with the property $x - y = 0 \iff x = y$.

Lemma 1.1.7. *Let* $(V, \|\cdot\|)$ *be a normed space. Then*

$$d(x, y) := \|x - y\|,\ x, y \in V,$$

is a metric on V.

Proof. By construction $d(x, y) \geq 0$. Let us check (i)–(iii) in the definition of a metric.

(i) $d(x, y) = \|x - y\| = 0 \iff x - y = 0 \iff x = y$ by Definition 1.1.5(i).

(ii) $d(y, x) = \|y - x\| = \|(-1)\cdot(x - y)\| = |-1|\cdot\|x - y\| = d(y, x)$ by Definition 1.1.5(iii).

(iii) $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$ by Definition 1.1.5(ii).

□

Examples of normed and metric spaces.

(1) The set of real numbers $\mathbb{R}$ with $\|x\| := |x|$ (absolute value).

(2) Euclidean space $\mathbb{R}^n$ with $n \geq 1$;

vectors $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $x_i \in \mathbb{R}$, $1 \leq i \leq n$,

$$\textbf{norm} \quad \|x\| := \sqrt{\sum_{i=1}^{n} x_i^2},$$

$$\textbf{Euclidean distance} \quad d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

$\|x + y\| \leq \|x\| + \|y\|$? The triangle inequality is not trivial. To check it one needs the **Cauchy-Schwarz** inequality (Theorem 1.1.10 below).

(3) Another norm on $\mathbb{R}^n$, the **maximum** norm

$$\|x\| := \max_{1\leq i\leq n} |x_i|.$$

(4) $\boxed{!}$ *Counterexample*: **Not every** metric defined on a vector space corresponds to some norm.

Exercise 1.1.8. *Show that the trivial metric on* $\mathbb{R}$ *does not come from any norm.*

(5) The space of continuous functions $C([0,1])$ with the **uniform** (or **maximum**) norm.

$$f\colon [0,1] \to \mathbb{R}, \text{ continuous},$$

$$\|f\| := \sup_{t\in[0,1]} |f(t)| = \max_{t\in[0,1]} |f(t)|.$$

Exercise 1.1.9. *Prove that the examples* (1), (2), (3) *and* (5) *are norms. You will need:*

Theorem 1.1.10 (Cauchy–Schwarz Inequality)**.** *For* $x, y \in \mathbb{R}^n$ *we define the scalar product* $x \cdot y := \sum_{i=1}^n x_i y_i \in \mathbb{R}$. *Then*

$$|x \cdot y| = \left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2} = \|x\| \cdot \|y\|.$$

Proof. Note that for any $k \in \mathbb{R}$

$$\begin{aligned} 0 \ &\leq \ \textstyle\sum_{i=1}^n (x_i - k y_i)^2 = k^2 \underbrace{\sum_{i=1}^n y_i^2}_{a} - 2k \underbrace{\sum_{i=1}^n x_i y_i}_{b} + \underbrace{\sum_{i=1}^n x_i^2}_{c} \\ &= \ k^2 a - 2kb + c. \end{aligned}$$

Quadratic polynomial $P(k) := k^2 a - 2kb + c \geq 0$ for any $k \in \mathbb{R}$ iff its discriminant $\Delta := b^2 - ac \leq 0$, i.e.,

$$\begin{aligned} \Delta \ &= \ \left(\sum_{i=1}^n x_i y_i\right)^2 - \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right) \leq 0. \\ &\Leftrightarrow \ \left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}. \end{aligned}$$

□

Definition 1.1.11. *The **open ball** with center at point $x \in X$ and radius $\epsilon > 0$ is*

$$B_\epsilon(x) := \{y \in X \mid d(x,y) < \epsilon\}.$$

*The **closed ball** with center at point $x \in X$ and radius $\epsilon > 0$ is*

$$\overline{B_\epsilon(x)} := \{y \in X \mid d(x,y) \leq \epsilon\}.$$

*The subset $U \subseteq X$ is **open** if for **each** $x \in U$ there exists an open ball $B_\epsilon(x) \subset U$ (with radius $\epsilon = \epsilon(x) > 0$ depending on x).*
*$U \subseteq X$ is **closed** if its complement $U^c := X \backslash U$ is open.*
*$U \subseteq X$ is called **bounded** if there exists some ball $B_R(x)$ (or $\overline{B_R(x)}$) such that $U \subseteq B_R(x)$ (resp. $U \subseteq \overline{B_R(x)}$).*

Lemma 1.1.12. *Open balls are open sets.*

Proof. Let $y \in B_R(x)$. We have to find $\epsilon > 0$ such that $B_\epsilon(y) \subseteq B_R(x)$.

$$\text{Set} \quad \epsilon := R - d(x,y) > 0 \quad (\text{since } y \in B_R(x)).$$

For any $z \in B_\epsilon(y)$, by the triangle inequality we have

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

which means $z \in B_\epsilon(y)$. Thus, $B_\epsilon(y) \subset B_R(x)$. □

Example 1.1.13. *The intervals (a,b), (b,∞), $(-\infty,b)$ are open sets in $\mathbb{R}$. Their unions are also open. The inverse statement is also true:*

Proposition 1.1.14. *Any open set Y in $\mathbb{R}$ can be represented as a **finite** or **countable** union of **non-intersecting** open intervals:*

$$Y = \bigcup_{i=1}^{N \leq \infty} (a_i, b_i), \quad (a_i,b_i) \bigcap (a_j,b_j) = \emptyset, \quad i \neq j.$$

Exercise 1.1.15. *What do open balls in $\mathbb{R}^2$ look like with:*

(i) the Euclidean norm;

(ii) the maximum norm;

(iii) the trivial metric?

Exercise 1.1.16. *Let (X,d) be a metric space. Show that:*

(i) $\emptyset$ and X are open sets in X;

(ii) the intersection of two (or a finite number of) open sets is open;

(iii) the union of arbitrarily many open sets is open.

What do points (i)-(iii) imply for closed sets?

Exercise 1.1.17. *Let $X = [0, 1]$ with the metric $d(x, y) = |y - x|$. Characterise all open balls in X. Is $[0, 1]$ open? Is $[0, 1/2]$ open?*

Exercise 1.1.18. *Show that the following are equivalent*

(i) $Y \subseteq X$ is bounded;

(ii) there is $x \in X$ and $C > 0$ such that $d(x, y) \leq C$ for all $y \in Y$;

(iii) $\operatorname{diam} Y := \sup_{x,y \in Y} d(x, y) < \infty$.

Definition 1.1.19. *If $A, B \subseteq X$, the **Hausdorff distance** between the sets A and B is given by*

$$d_H(A, B) := \max\{\sup_{x\in A} \inf_{y\in B} d(x, y), \sup_{y\in B} \inf_{x\in A} d(x, y)\}.$$

In particular, the distance from x to B is given by

$$d_H(x, B) := \inf_{y\in B} d(x, y).$$

Exercise 1.1.20. [*] *Show that, in general, the Hausdorff distance between two sets is **not** a metric. [**Hint:** Show that $d_H(A, B) = 0$ if and only if $\bar{A} = \bar{B}$. Hence, construct a counterexample—find two subsets $A \neq B \subset \mathbb{R}$ such that $\bar{A} = \bar{B}$.]*

What *is* true is that the Hausdorff distance on $\mathbb{R}^n$ defines a metric on the set of all non-empty **closed and bounded** subsets of $\mathbb{R}^n$ (we'll see what this means soon).

Basic Examples

Now we will introduce some of the more important metrics, as well as some metrics which are a bit 'non-standard' (just to illustrate that metrics can be unintuitive sometimes):

(i) **Euclidean space** $(\mathbb{R}^n, |\cdot|)$ of vectors $x = (x_i)_{i=1}^n$ with the metric

$$d(x, y) = |x - y| := \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

(ii) **Manhattan** (or taxicab, or city-block, or l^1-) metric (norm) on $\mathbb{R}^n$ (in particular, $n = 2$):

$$d(x, y) = \|x - y\|_{l^1} := \sum_{i=1}^{n} |x_i - y_i|.$$

(Taxicabs cannot drive through buildings. They have to drive either North-South or East-West).

(iii) **British rail** (or post) metric on $\mathbb{R}^n$ (with centre 0:=London):

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

(This metric is not generated by any norm $\|\cdot\|$. Indeed, $\|x - y\| = d(x, y) = |x| + |y|$ (for $x \neq y$) implies $\|x\| = d(x, 0) = |x|$ (for $x \neq 0$ and $y = 0$). But then we must have $|x - y| = \|x - y\| = d(x, y) := |x| + |y|$, which is surely wrong for arbitrary $x \neq y$.) **Exercise:** Let $n = 2$ Show that the set $\{(0, 1)\}$ is open in the British rail metric. Find all points $x \in \mathbb{R}^2$ such that the set $\{x\}$ is **not** open in the British rail metric. Describe all the open sets in this metric.

(iii) **French Metro** metric on $\mathbb{R}^n$ (with centre 0:=Paris):

$$d(x, y) = \begin{cases} |x - y|, & x = cy, \ c \in \mathbb{R} \\ |x| + |y|, & \text{otherwise.} \end{cases}$$

This metric is similar to the British rail metric, but now passengers are allowed to take a shorter journey if their destination is on the same rail line coming from Paris (both points lie on the same ray emanating from the origin). **Exercise:** Can you find an open ball in the British rail metric which is not open in the French Metro metric?

(iv) $C([a, b])$ – Space of all **continuous** functions on a **bounded** interval $[a, b]$ with the maximum norm

$$\|f - g\|_\infty := \max_{t \in [a,b]} |f(t) - g(t)|.$$

(v) $\mathbb{R}^\infty$ – space of **all real sequences** $x = (x_i)_{i \geq 1}$ with $x_i \in \mathbb{R}$. The metric (which cannot be generated by any norm) is given by

$$d(x, y) := \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}.$$

(vi) l_∞ – the space of all **bounded sequences**

$$l_\infty := \left\{ x = (x_i)_{i\geq 1} \in \mathbb{R}^\infty \;\middle|\; \sup_{i\geq 1} |x_i| < \infty \right\}$$
$$\text{with } d(x,y) := \|x-y\|_\infty := \sup_{i\geq 1} |x_i - y_i|.$$

(vii) The space of p-**summable sequences** l_p, $1 \leq p < \infty$,

$$l_p := \left\{ x = (x_i)_{i\geq 1} \in \mathbb{R}^\infty \;\middle|\; \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}$$
$$\text{with the norm } \|x\|_p := \sqrt[p]{\sum_{i=1}^{\infty} |x_i - y_i|^p}.$$

The triangle inequality for $\|\cdot\|_p$ is known as **Minkovski's inequality**

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p}, \quad x, y \in l_p.$$

Closely related is **Hölder's inequality** for sequences

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \cdot \left(\sum_{i=1}^{\infty} |y_i|^q\right)^{1/q},$$
$$x \in l_p, \;\; y \in l_q, \;\; \frac{1}{p} + \frac{1}{q} = 1 \;\; (p, q > 1).$$

Exercise 1.1.21. *Check that the following define metrics on the set of all positive integers* $\mathbb{N} := \{1, 2, 3, \ldots\}$*:*

(i) $\rho(n,m) = \frac{|m-n|}{mn}$;

(ii) $\rho(n,m) = \begin{cases} 0, & m = n; \\ 1 + \frac{1}{m+n}, & m \neq n. \end{cases}$

1.2 Sequences and Convergence

Definition 1.2.1. *Let* (X,d) *be a metric space and* $(x_n)_{n\in\mathbb{N}}$ *be a sequence in* X*. We say that* x_n ***converges*** *to some* $x \in X$ *if*

$$\forall \epsilon > 0, \;\; \exists N(\epsilon) \in \mathbb{N}, \;\; \forall n > N(\epsilon): \;\; d(x_n, x) < \epsilon.$$

Notation 1.2.2. *We write* $x_n \to x$ *or* $\lim_{n\to\infty} x_n = x$.

In other words: $x_n \in B_\epsilon(x)$ for all $n > N(\epsilon)$;
or equivalently, $x_n \to x \iff d(x_n, x) \to 0$ (in the sense of real numbers)

Theorem 1.2.3 (Uniqueness of limits)**.** *A sequence* $(x_n)_{n\in\mathbb{N}}$ *has at most one limit.*

Proof. Suppose not, and

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} x_n = y, \quad x \neq y.$$

Let

$$\epsilon := d(x,y)/2 > 0.$$

Then:
$\exists N_1(\epsilon) : d(x_n, x) < \epsilon \ \forall n > N_1(\epsilon)$;
$\exists N_2(\epsilon) : d(x_n, y) < \epsilon \ \forall n > N_2(\epsilon)$.
Put $N := \max\{N_1(\epsilon), N_2(\epsilon)\}$. Then for $n > N : d(x_n, x),\ d(x_n, y) < \epsilon$.

$$d(x,y) \leq d(x_n, x) + d(x_n, y) < 2\epsilon = d(x,y),$$

which is a contradiction with $d(x,y) > 0$. □

Theorem 1.2.4. *If* $(x_n)_{n\in\mathbb{N}}$ *converges, then* $(x_n)_{n\in\mathbb{N}}$ *is bounded.*

Proof. Choose $\epsilon = 1$ in the definition of convergence. There exists $N_0 \in \mathbb{N}$ such that $d(x_n, x) < 1$ for $n > N_0$. Let

$$R := 1 + \max_{1\leq n\leq N_0} d(x_n, x).$$

$$\begin{aligned} \text{Then } d(x_n, x) &\leq R \text{ for all } n \in \mathbb{N} \text{ i.e.,} \\ (x_n)_{n\in\mathbb{N}} &\subseteq \overline{B_R(x)}. \end{aligned}$$

□

Remark 1.2.5. *In the discrete space* X *(see Example 1.1.2), any convergent sequence is* 'eventually stationary'*:*

$$\lim_{n\to\infty} x_n = x \iff \exists N_0 \in \mathbb{N} : \quad x_n = x \ \forall n > N.$$

Definition 1.2.6. $x \in X$ *is a* ***cluster point*** *of* $(x_n)_{n\in\mathbb{N}}$ *if for any (small) open ball with center at* x*, the sequence returns infinitely often to the ball.*

In other words,

$$\forall \epsilon > 0, \ \forall N \in \mathbb{N}, \exists n = n(\epsilon) > N : d(x_n, x) < \epsilon.$$

Remark 1.2.7. *This is a weaker condition than convergence: we may have an infinite number of terms **outside** the ball $B_\epsilon(x)$. The limit of the sequence is always a cluster point, but the converse **need not** be true.*

Example 1.2.8. *Let $X = \mathbb{R}$ and*

$$x_n = \begin{cases} 0, & n \text{ even } (n = 2m), \\ 1, & n \text{ odd } (n = 2m+1). \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ has 2 cluster points, 0 and 1.

Example 1.2.9. *Let $X = \mathbb{R}$ and*

$$x_n = (0, 0, 1, 0, 1, 2, 0, 1, 2, 3, 0, 1, 2, 3, 4, 0, 1, 2, 3, 4, 5, \ldots)$$

$(x_n)_{n\in\mathbb{N}}$ has infinitely many cluster points, $0, 1, 2, 3, \ldots$.

Exercise 1.2.10. *Show that the sequence*

$$x_n = (-1)^n \left(1 - \frac{1}{n}\right)$$

does not converge in $\mathbb{R}$. Find all cluster points of the sequence.

Theorem 1.2.11. *x is a cluster point of $(x_n)_{n\in\mathbb{N}}$ if and only if there exists a subsequence $(x_{n_m})_{m\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $\lim_{m\to\infty} x_{n_m} = x$.*

Proof. We give only a sketch.

($\Longleftarrow$) is trivial.

($\Longrightarrow$) Choose a sequence of balls $B_{\epsilon_m}(x)$ with $\epsilon_m \downarrow 0$. For each m find some $x_{n_m} \in B_{\epsilon_m}$. Make sure that $n_{m+1} > n_m$ (this is always possible since $B_{\epsilon_m}(x)$ contains infinitely many elements of $(x_n)_{n\in\mathbb{N}}$). Then $x_{n_m} \to x$ as $m \to \infty$. □

Exercise 1.2.12. *Find cluster points of the following sequences in $\mathbb{R}$:*

(i) $x_n = (-1)^n$;

(ii) $x_n = \sin\left(\frac{\pi}{2} n\right)$;

(iii) $x_n = \begin{cases} 2 - \frac{1}{n}, & n \text{ even}, \\ 1/n, & n \text{ odd}; \end{cases}$

(iv) $x_n = n \mod 4$.

1.3 Open and Closed Sets

Recall the definition of an open set from Section 1.1.

Theorem 1.3.1. *Let (X, d) be a metric space. Then:*

(i) $\emptyset$ and X are ***open*** *and* ***closed*** *(simultaneously);*

(ii) The intersection of any ***finite*** *collection of open sets is open;*

(iii) The union of any ***arbitrary*** *(possibly* ***infinite****) collection of open sets is open.*

Exercise 1.3.2. *Prove this theorem.*

Example 1.3.3. *We'll show why point (ii) of Theorem 1.3.1 does not necessarily apply to* ***infinite*** *intersections. Consider open balls in $\mathbb{R}^n$, $n \geq 1$,*

$$B_{\epsilon_n}(x_0) \text{ with some } x_0 \in \mathbb{R}^n \text{ and } \epsilon_n = 1/n.$$

Then the countable intersection

$$\bigcap_{n\in\mathbb{N}} B_{\epsilon_n}(x_0) = \{x_0\}$$

is not *open in $\mathbb{R}^n$.*

Definition 1.3.4. *Let (X, d) be a metric space, and $A \subseteq X$ (a subset of X).*

(i) A point $x \in A$ is an ***interior point*** *of A if $B_\epsilon(x) \subseteq A$ for some $\epsilon > 0$. The* ***interior*** *of A, which will be denoted by $\operatorname{int} A$ or $\mathring{A}$, is the set of all interior points of A.*

(ii) $x \in X$ (possibly, $x \notin A$) is a ***closure point*** *of A if*

$$B_\epsilon(x) \cap A \neq \emptyset \text{ for all } \epsilon > 0.$$

The ***closure*** *of A, which will be denoted by $\bar{A}$, consists of all closure points of A. Obviously, $A \subseteq \bar{A}$ (since $B_\epsilon(x) \cap A \ni \{x\}$ for each $x \in A$).*

(iii) $x \in X$ is a ***boundary point*** *of A if for all $\epsilon > 0$:*

$$B_\epsilon(x) \cap A \neq \emptyset \text{ and } B_\epsilon(x) \cap A^c \neq \emptyset.$$

The ***boundary*** *of A, denoted by ∂A, consists of all boundary points of A.*

$$x \in \partial A \iff x \in \bar{A} \cap \overline{A^c}.$$

From the above definitions

$$\mathring{A} \subseteq A \subseteq \bar{A}, \quad \bar{A} = A \cup \partial A = \mathring{A} \cup \partial A, \quad \partial A = \bar{A} \backslash \mathring{A} = \bar{A} \cap \overline{A^c}.$$

In particular, the closure of the open ball $B_\epsilon(x)$ is always contained in the closed ball $\overline{B_\epsilon(x)}$, see Definition 1.1.11. Furthermore, in the Euclidean space $\mathbb{R}^n$ they coincide.

Remark 1.3.5. *In general, the closure of $B_\epsilon(x)$* ***does not coincide*** *with the closed ball $\overline{B_\epsilon(x)}$! A typical counterexample is the discrete metric with the unit open ball $B_1(x) = \{x\}$ (which is an open and closed set simultaneously) and the closed ball $\overline{B_1(x)} = X$.*

Theorem 1.3.6. *Let (X, d) be a metric space. Then:*

(i) $\mathring{A}$ is the ***largest open*** *set contained in A;*

(ii) A is ***open*** *$\iff A = \mathring{A}$;*

(iii) $\bar{A}$ is the ***smallest closed*** *set containing A;*

(iv) A is ***closed*** *$\iff A = \bar{A}$*

Proof. We will prove only (i), (ii).

(i) Show first that $\mathring{A}$ is **open**. By definition, for any $x \in \mathring{A}$ there exists $B_\epsilon(x) \subset A$. We claim that indeed $B_\epsilon(x) \subset \mathring{A}$. Pick an arbitrary $y \in B_\epsilon(x)$. Since $B_\epsilon(x)$ is an open set, there exists $\eta > 0$ such that

$$B_\eta(y) \subset B_\epsilon(x) \subset A.$$

This means that any $y \in B_\epsilon(x)$ is an interior point of A, i.e., $B_\epsilon(x) \subset \mathring{A}$.

Show next that $\mathring{A}$ is the **largest open** subset of A: If $B \subseteq A$ and B is open, then $B \subseteq \mathring{A}$. Indeed, for every point $x \in B$ one finds a ball $B_\epsilon(x) \subset B \subset A$, which means $x \in \mathring{A}$.

(ii) If $A = \mathring{A}$, then A is open, because $\mathring{A}$ is always open by Claim (i). If A is open, then any $x \in A$ is an interior point, which implies $A \subseteq \mathring{A} \subseteq A$. □

Exercise 1.3.7. *Prove Claims (iii) and (iv).* ***Hint:*** *use that A open $\iff A^c$ closed.*

The next theorem is **very important**!

Theorem 1.3.8 (Characterization of Closed Sets)**.** *A set $A \subseteq X$ is* ***closed*** *if and only if any convergent sequence in X, $(x_n)_{n\in\mathbb{N}} \subset A$ has its limit* ***inside*** *A, i.e.,*

$$\exists \lim_{n\to\infty} x_n := x \in X \implies x \in A.$$

Proof. ($\Longrightarrow$) Let A be closed and $x := \lim_{n\to\infty} x_n \in X$. Suppose $x \notin A$, i.e., $x \in A^c := X \backslash A$, which is an open set. Therefore, $B_\epsilon(x) \subseteq A^c$ for some $\epsilon > 0$. By the definition of convergence,

$$x_n \in B_\epsilon(x) \subseteq A^c, \ \ \forall n > N(\epsilon) \in \mathbb{N}.$$

This contradicts the assumption $(x_n)_{n\in\mathbb{N}} \subseteq A$.

($\Longleftarrow$) We show that A^c is open. So, let $x \in A^c$, i.e., $x \notin A$. If $B_\epsilon(x) \nsubseteq A^c$ for any $\epsilon > 0$, i.e., $B_\epsilon(x) \cap A \neq \emptyset$, we can pick a sequence $x_n \in B_{\epsilon_n}(x) \cap A$ with $\epsilon_n \downarrow 0$, $n \to \infty$. By construction, $(x_n)_{n\in\mathbb{N}} \subset A$ and $\lim_{n\to\infty} x_n = x$. By assumption we should have that $x \in A$, which is a contradiction with $x \notin A$. □

Exercise 1.3.9. *Show that any closed ball $\overline{B_\epsilon(x)}$ is a closed set.* ***Hint:*** *use Theorem 1.3.8. Indeed, let $(x_n)_{n\in\mathbb{N}} \subset \overline{B_\epsilon(x)}$ be convergent to some $y \in X$. Then by the triangle inequality, for any $n \in \mathbb{N}$*

$$d(x, y) \leq d(x, x_n) + d(x_n, y) \leq R + d(x_n, y).$$

Since $d(x_n, y) \to 0$ as $n \to \infty$, this yields $d(x, y) \leq R$.

Exercise 1.3.10. *Let $A \subseteq X$ be closed and $x \notin A$. Show that*

$$d(x, A) := \inf_{y\in A} d(x, y) > 0$$

Hint: *what happens if we suppose the converse is true?*

1.4 Limits of Functions, Continuity

Definition 1.4.1 (Continuous functions)**.** *Let (X, d) and (Y, ρ) be two metric spaces. Let $f\colon X \to Y$ be a function (also called a map or mapping). We say that f is* ***continuous at a point*** *$x_0 \in X$ if*

$$\forall \epsilon > 0, \ \exists \delta > 0 : d(x, x_0) < \delta \implies \rho(f(x), f(x_0)) < \epsilon.$$

We say that f is ***continuous*** *if it is continuous at* ***every*** *point $x_0 \in X$.*

Intuition: you can control the changes of f: a "*small*" change in x away from x_0 will not change $f(x_0)$ too much.

The following is an equivalent definition of continuity.

Definition 1.4.2 (Continuous functions in terms of open balls)**.**

$$\forall \epsilon > 0 \ \exists \delta > 0 : f(x) \in B_\epsilon(f(x_0)) \text{ for all } x \in B_\delta(x_0).$$

Note that in the above definition $\delta := \delta(\epsilon, x_0)$, i.e., it depends on ϵ as well as on the point x_0.

History: This (ϵ, δ)-definition is due to A.-L. Cauchy (1789-1857), the French mathematician who was an early pioneer of analysis. He had more than 800 research papers and became a full professor of École Polytechnique at 28 years old. Cauchy's inspiration for continuity came from properties of differentiable functions. It was thought for a very long time that all continuous functions were differentiable (except possibly on a set of isolated points). Weierstrass proved that this was **not true** in 1872 when he gave an example of a continuous function that was differentiable **nowhere**.

Exercise 1.4.3. *Let $y \in X$ be fixed. Show that the distance function f defined by*

$$(X, d) \ni x \to f(x) := d(x, y) \in \mathbb{R}$$

is continuous.

Theorem 1.4.4 (Sequential Characterization of Continuity)**.** *The function $f\colon X \to Y$ is continuous at the point $x_0 \in X$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ converging to x_0, the sequence $(f(x_n))_{n\in\mathbb{N}} \subseteq Y$ is convergent to $f(x_0) \in Y$.*

$$\textbf{\textit{Symbolically:}} \quad \lim_{n\to\infty} x_n = x_0 \implies \lim_{n\to\infty} f(x_n) = f(x_0).$$

Proof. ($\Longrightarrow$) Suppose f is continuous at a point $x_0 \in X$. Let $x_n \to x_0$ as $n \to \infty$. For a given $\epsilon > 0$, choose $\delta = \delta(\epsilon) > 0$ such that $\rho(f(x), f(x_0)) < \epsilon$ for all $x \in X$ with $d(x, x_0) < \delta$. By the definition of $\lim_{n\to\infty} x_n = x_0$,

$$\exists N(\delta) \in \mathbb{N} : d(x_n, x_0) < \delta \text{ for all } n > N(\delta).$$

Hence,

$$\rho(f(x_n), f(x_0)) < \epsilon \text{ for all } n > N(\delta).$$

Therefore, $f(x_n) \longrightarrow f(x_0)$ as $n \to \infty$.

($\Longleftarrow$) Proof by contradiction: Suppose f is not continuous at $x_0 \in X$. Then $\exists \epsilon > 0$ such that $\forall \delta > 0$ one finds $x_\delta \in X$ with $d(x_\delta, x_0) < \delta$ but $\rho(f(x_\delta), f(x_0)) \geq \epsilon$. Now let $\delta_n = 1/n$ and choose the corresponding sequence

$$x_n := x_{\delta_n}, \;\; n \in \mathbb{N}.$$

Then $x_n \to x_0$, but $f(x_n) \nrightarrow f(x_0)$ as $n \to \infty$. ☐

So, the (ϵ, δ)-definition of continuity is **equivalent** to the sequential definition.

History: The notion of sequential continuity is due to the German mathematician Heinrich Heine (1821-81).

Theorem 1.4.5 (Composition of Continuous Functions)**.** *Let* (X, d_1)*,* (Y, d_2)*, and* (Z, d_3) *be metric spaces. Let* $f\colon X \to Y$ *and* $g\colon Y \to Z$ *be functions such that* f *is continuous at some point* $x_0 \in X$ *and* g *is continuous at* $y_0 := f(x_0) \in Y$*. Then, their composition*

$$h := g \circ f\colon X \to Z$$

is continuous at x_0*.*

Proof. We use the characterization of continuity in terms of sequences. Let $x_n \to x_0$ as $n \to \infty$. As f is continuous, $y_n := f(x_n) \to f(x_0) =: y_0$. As g is continuous, $z_n := g(y_n) \to g(y_0) = g(f(x_0)) =: z_0 \in Z$ as $n \to \infty$. □

The next theorem is one of the most important characterisations of continuity and it should be memorised!

Theorem 1.4.6 (Global Continuity)**.** *The function* $f\colon X \to Y$ *is continuous if and only if the* preimage

$$f^{-1}(U) := \{x \in X \mid f(x) \in U\}$$

of any ***open*** *set* $U \subseteq Y$ *is* ***open*** *in* X*.*

Proof. ($\Longrightarrow$) Let $f\colon X \to Y$ be continuous on X. Let U be an open subset of Y, we check that that $f^{-1}(U)$ is open. Take any $x_0 \in f^{-1}(U)$, then $f(x_0) \in U$. As U is open, there exists $\epsilon > 0$ such that $B_\epsilon(f(x_0)) \subseteq U$. As f is continuous at x_0, there exists $\delta > 0$ such that

$$f(B_\delta(x_0)) \subseteq B_\epsilon(f(x_0)) \subseteq U,$$

or

$$B_\delta(x_0) \subseteq f^{-1}(U).$$

This means that $f^{-1}(U)$ is open.

($\Longleftarrow$) Let $x_0 \in X$ and $\epsilon > 0$. By assumption, $f^{-1}\left[B_\epsilon(f(x_0))\right] \subseteq X$ is open as the preimage of the open ball $B_\epsilon(f(x_0))$. Hence $\exists \delta > 0$ such that $B_\delta(x_0) \subseteq f^{-1}\left[B_\epsilon(f(x_0))\right]$, or $f(B_\delta(x_0)) \subseteq B_\epsilon(f(x_0))$. This gives the required continuity of f. □

Corollary 1.4.7. *Let* $f\colon (X, d) \to (\mathbb{R}, |\cdot|)$ *be continuous. Then for any* $C \in \mathbb{R}$

$$\begin{aligned} &\{x \in X \mid f(x) > C\} \quad \textit{is open,} \\ &\{x \in X \mid f(x) \le C\} \quad \textit{is closed.} \end{aligned}$$

There is an equivalent formulation of Theorem 1.4.6 in terms of closed sets.

Corollary 1.4.8. *The function $f\colon X \to Y$ is continuous $\iff$ the preimage $f^{-1}(B)$ of any* ***closed set*** *$B \subseteq Y$ is* ***closed*** *in X.*

Exercise 1.4.9. *Prove the above corollary using Theorem 1.4.6 and the equality*

$$[f^{-1}(U)]^c = f^{-1}(U^c)$$

for any subset $U \subseteq Y$.

Warning: A similar statement for ***images*** of open (resp. closed) sets is in general **wrong**! If $V \subseteq X$ is open (resp. closed) in X, then

$$f(V) := \{y \in Y \mid y := f(x),\ x \in V\}$$

is **not** necessarily open (resp. closed) in Y.

Exercise 1.4.10. *Find an example of a continuous function $f\colon X \to Y$ and an open subset $U \subseteq X$ such that $f(U)$ is not an open subset of Y.*

Definition 1.4.11. *Let (X, d) and (Y, ρ) be two metric spaces. A function $f\colon X \to Y$ is* ***uniformly continuous*** *if*

$$\forall \epsilon > 0,\ \exists \delta > 0 :\ \ \rho(f(x_1), f(x_2)) < \epsilon,\ \ \forall x_1, x_2 \in X \ \text{with}\ d(x_1, x_2) < \delta.$$

Remark 1.4.12. *$\delta := \delta(\epsilon)$ is* ***independent*** *of points $x \in X$, unlike in the (ϵ, δ)-definition of continuity.*

The following generalisations of continuity are important but will not be examinable:

Definition 1.4.13. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces. A function $f\colon X \to Y$ is* ***Lipschitz-continuous*** *if there exists a (Lipschitz) constant $L > 0$ such that*

$$\|f(x_1) - f(x_2)\|_Y \le L\|x_1 - x_2\|_X, \quad \forall x_1, x_2 \in X.$$

Lemma 1.4.14. *Lipschitz-continuous functions are* ***uniformly continuous****.*

Proof. Let $\epsilon > 0$ and set $\delta := \epsilon/L > 0$. Then for any $x_1, x_2 \in X$ with $\|x_1 - x_2\|_X < \delta$

$$d_Y(x_1, x_2) := \|f(x_1) - f(x_2)\|_Y \le L\|x_1 - x_2\|_X < L\delta = \epsilon.$$

□

The following definition is a generalisation of Lipschitz continuity.

Definition 1.4.15. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces. A function $f\colon X \to Y$ is* ***Hölder-continuous*** *(with exponent α) if there exist $\alpha, L > 0$ such that*

$$\|f(x_1) - f(x_2)\|_Y \le L\|x_1 - x_2\|_X^\alpha, \quad \forall x_1, x_2 \in X.$$

Hölder-continuous functions are also uniformly continuous.

1.5 Cauchy Sequences, Complete Metric Spaces

Definition 1.5.1. *Let (X,d) be a metric space. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ is a* ***Cauchy sequence*** *if*

$$\forall \epsilon > 0,\ \exists N(\epsilon) \in \mathbb{N},\ \forall m,n > N(\epsilon): \ d(x_m, x_n) < \epsilon.$$

Intuition: A sequence is Cauchy if its terms get *closer* and *closer* to each other.

Theorem 1.5.2. *Every convergent sequence in (X,d) is Cauchy.*

Proof. Let $x_n \to x$ as $n \to \infty$. Then $\forall \epsilon > 0$, $\exists N(\epsilon) > 0$ such that

$$d(x_n, x) < \epsilon/2,\ \forall n > N(\epsilon).$$

Hence, for any $n, m > N(\epsilon)$

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon.$$

□

Warning: In general metric spaces, the converse claim is false!

Definition 1.5.3. *A metric space (X,d) is called* ***complete*** *if every Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in X converges to some $x \in X$.*

A normed space $(X, \|\cdot\|_X)$ that is complete w.r.t. the metric $d_X(x_1,x_2) := \|x_1 - x_2\|_X$ is called a ***Banach*** *space.*

Exercise 1.5.4. *Every Cauchy sequence in (X,d) is bounded (compare with Theorem 1.2.4).*

Exercise 1.5.5. *Let (X,d) be a complete metric space and $Y \subseteq X$. Then*

$$(Y,d) \text{ is complete } \iff Y \text{ is a closed set in } (X,d).$$

1.6 Banach Contraction Mapping Theorem

Definition 1.6.1. *Let (X,d) be a metric space. A function*

$$T\colon X \to X$$

sending X into X is called an ***operator*** *in X. An operator T is a* ***contraction*** *of* ***modulus*** *$\beta \in (0,1)$ if*

$$d(Tx, Ty) \leq \beta \cdot d(x,y),\ \ \forall x, y \in X.$$

Exercise 1.6.2. *Show that every contraction is a uniformly continuous function.*

The next theorem is possibly the most important result in this course and it should be memorised! It is used throughout analysis, the study of differential equations, matrix analysis, game theory, dynamical systems and many other areas of mathematics and economics.

Theorem 1.6.3 (Banach Contraction Mapping Theorem)**.** *Let (X, d) be a* ***complete*** *metric space and $T\colon X \to X$ be a* ***contraction*** *of modulus $\beta \in (0,1)$. Then:*

(i) T has exactly one *fixed point $x^* \in X$ solving the equation*

$$Tx = x.$$

(ii) ***How to construct*** *x^*: For* ***any*** *starting point $x_0 \in X$, the sequence $(x_n)_{n\in\mathbb{N}}$ defined by*

$$x_1 := Tx_0,\ \ x_2 := Tx_1,\ \ \ldots\ \ ,\ \ x_{n+1} := Tx_n,\ \ \ldots \tag{$*$}$$

converges to x^*.

Proof. It is useful to prove the existence of the fixed point and its uniqueness separately.

(ia) **Existence:** Take any $x_0 \in X$ and define $(x_n)_{n\in\mathbb{N}}$ as in $(*)$. We show that this sequence is Cauchy:

$$\begin{aligned} d(x_{n+1}, x_n) &= d(Tx_n, Tx_{n-1}) \le \beta d(Tx_n, Tx_{n-1}) \\ &\le \ldots \le \beta^n d(Tx_1, Tx_0),\ n \in \mathbb{N}. \end{aligned}$$

Then, by the triangle inequality, for any $n > m$

$$d(x_n, x_m) \le \sum_{i=m}^{n-1} d(x_{i+1}, x_i) \le \left(\sum_{i=m}^{n-1} \beta^i\right) \cdot d(x_1, x_0) \le \left(\sum_{i=m}^{\infty} \beta^i\right) \cdot d(x_1, x_0)$$
$$\text{(geometric series)} \ \le \frac{\beta^m}{1-\beta} d(x_1, x_0) \to 0 \text{ as } m \to \infty \text{ (since } \beta^m \to 0).$$

Since (X, d) is complete $\Longrightarrow \exists x^* := \lim_{n\to\infty} x_n$. We show that this x^* is a fixed point of T. Recall that T is continuous as it is a contraction (see Exercise 1.6.2 and Theorem 1.4.4). Indeed,

$$Tx^* = T(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} x_{n+1} = x^*.$$

(ib) **Uniqueness:** Let $x, y \in X$ be two fixed points. T is a contraction with $\beta < 1 \Longrightarrow$

$$d(x, y) = d(Tx, Ty) \leq \beta d(x, y),$$
$$\text{i.e., } d(x, y) = 0 \iff x = y.$$

(ii) As was shown in (ia), any sequence of the form $(*)$ converges to some fixed point x^* (possibly depending on the initial point x_0). But according to (ib), the fixed point is unique, which means that each approximating sequence $(x_n)_{n \in \mathbb{N}}$ has the limit x^*, which is the same for all $x_0 \in X$.

□

Exercise 1.6.4. *Let (X, d) be a complete metric space, and let $T \colon X \to X$ be such that, for some $n \in \mathbb{N}$, the operator T^n is a contraction. Show that T has a unique fixed point.*

Hint: *(i) Prove that T^n has a unique fixed point, say x^*. (ii) Check that this x^* is also the unique fixed point for T.*

The Banach fixed point theorem is one of the most important theorems in all of mathematics! Many problems can be formulated as equations $F(x) = 0$ and can be rewritten in the form $f(x) = x$ with $f(x) := F(x) + x$.

Standard applications of the Banach fixed point theorem include:

— The Picard-Lindelöf theorem about unique solvability of ordinary differential equations.

— The Page rank algorithm used by Google: One computes a fixed point of a linear operator in $\mathbb{R}^N$ (with huge $N \to \infty$), which is a contraction. This fixed point $x^* \in \mathbb{R}^N$ gives ordering of pages.

— Image compression: Digital encoding of images in the JPEG format is also a mathematical algorithm based on the Banach fixed point theorem.

Theorem 1.6.5 (Continuous Dependence of the Fixed Point on Parameters)**.** *Let (X, d) and (Ω, ρ) be metric spaces, and $T(x, \omega)$ be a mapping $X \times \Omega \to X$. Furthermore, let (X, d) be **complete**. Suppose that for each $x \in X$*

$$\Omega \ni \omega \to T(x, \omega) \in X \textit{ is } \textbf{\textit{continuous}},$$

and for each $\omega \in \Omega$

$$X \ni x \to T(x, \omega) \in X \textit{ is a } \textbf{contraction} \textit{ with (the same) } \beta \in (0, 1).$$

Then the solution $x^*(\omega) \in X$ of the fixed point problem

$$T(x, \omega) = x$$

*is a **continuous** function of the parameter $\omega \in \Omega$.*

Proof. Let $\omega_n \to \omega$ in (Ω, ρ). We need to show that

$$d(x^*(\omega_n), x^*(\omega)) \to 0, \text{ as } n \to \infty.$$

Denote the corresponding fixed points

$$\begin{aligned} x^* \ &: = x^*(\omega) = T(x^*(\omega), \omega) = T(x^*, \omega), \\ x_n^* \ &: = x^*(\omega_n) = T(x^*(\omega_n), \omega_n) = T(x_n^*, \omega_n). \end{aligned}$$

Thus, by the triangle inequality

$$\begin{aligned} d(x_n^*, x^*) &= d(T(x^*, \omega), T(x_n^*, \omega_n)) \\ &\leq d(T(x^*, \omega), T(x^*, \omega_n)) + d(T(x^*, \omega_n), T(x_n^*, \omega_n)) \\ &\leq d(T(x^*, \omega), T(x^*, \omega_n)) + \beta \cdot d(x^*, x_n^*), \end{aligned}$$

which can be rewritten as

$$(1 - \beta) d(x_n^*, x^*) \leq d(T(x^*, \omega), T(x^*, \omega_n))$$

or

$$d(x_n^*, x^*) \leq \frac{1}{1 - \beta} d(T(x^*, \omega), T(x^*, \omega_n)).$$

Since $T(x^*, \omega)$ is continuous in ω, the right-hand side tends to zero as $\omega_n \to \omega$. Thus, $d(x_n^*, x^*)$ as $n \to \infty$. □

Exercise 1.6.6. *Let (X, d) be a complete metric space. Suppose we have two contractions $A \colon X \to X$ and $B \colon X \to X$ such that*

$$d(Ax, Ay) \leq \alpha d(x, y), \quad d(Bx, By) \leq \beta d(x, y), \ \textit{with } \alpha, \beta \in (0, 1).$$

Prove that, if for some $\epsilon > 0$

$$d(Ax, Bx) < \epsilon \textit{ for all } x \in X,$$

then the fixed points x^* and y^* obey

$$d(x^*, y^*) \leq \frac{\epsilon}{1 - \gamma}, \quad \gamma := \max\{\alpha, \beta\} < 1.$$

Exercise 1.6.7. *On a calculator, if you enter any number x_0 and then apply the* cos *function to x_0, you get a new number $x_1 := \cos x_0$. If we iterate this process $x_{n+1} := \cos x_n$, you'll find that the sequence of values x_n converges to some other value $x_\infty \simeq 0.739085133\ldots$. Show that the function $g\colon [0,1] \to [0,1]$ given by $g(x) = \cos x$ is a contraction (you might need to use the Mean Value Theorem). Interpret this property of the cosine function in terms of the Banach fixed point theorem. Why does this justify the method of iterating the* cos *function on a calculator to show that the sequence x_n converges? How many roots does the function $f(x) = \cos x - x$ have in the interval $[0,1]$?*

1.7 Cantor Intersection Theorem

Theorem 1.7.1 (Cantor Intersection Theorem[1])**.** *A metric space (X,d) is* ***complete*** *if and only if any* ***decaying*** *sequence of non-empty* ***closed*** *sets*

$$\emptyset \neq A_n = \bar{A}_n \subseteq X, \quad A_{n+1} \subseteq A_n, \quad n \in \mathbb{N},$$

with ***diameters decaying to zero***

$$\mathrm{diam} A_n := \sup\{d(x,y) \mid x,y \in A_n\} \to 0, \;\; n \to \infty,$$

has ***exactly one*** *common point*

$$x \in \bigcap_{n\in\mathbb{N}} A_n \;\; (i.e., \bigcap_{n\in\mathbb{N}} A_n = \{x\}).$$

Proof. ($\Longrightarrow$) Suppose X is complete and consider any decaying sequence $(A_n)_{n\in\mathbb{N}} \subseteq X$ as described above. Since $A_n \neq \emptyset \implies \exists x_n \in A_n$. Obviously, $x_m \in A_m \subseteq A_n$ for all $m \geq n$, and hence

$$d(x_n, x_m) \leq \mathrm{diam} A_n, \; m \geq n.$$

So, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence as $\mathrm{diam} A_n \to 0$. Since X is complete $\Longrightarrow$

$$\exists \lim_{n\to\infty} x_n =: x \in X.$$

We claim that $x \in \bigcap_{n\in\mathbb{N}} A_n$. Indeed, $x_m \in A_n$ for all $m \geq n$ and A_n is closed, thus by the characterisation of closed sets (Theorem 1.3.8), $x = \lim_{m\to\infty,\, m\geq n} x_m = x \in A_n$. So, $x \in \bigcap_{n\in\mathbb{N}} A_n$.

[1] Georg Cantor (1845–1918), German mathematician, inventor of set theory

Finally, we observe that $\bigcap_{n\in\mathbb{N}} A_n$ consists only of the unique point $\{x\}$. If $y \neq x$ is some other point from $\bigcap_{n\in\mathbb{N}} A_n$, then

$$d(x,y) \leq \mathrm{diam} A_n \to 0,\ n \to \infty,$$

which yields $x = y$.

($\Longleftarrow$) This is left as an optional exercise. It is not so trivial. Argue by contradiction.

□

1.8 Separable Spaces

Definition 1.8.1. *A subset $A \subseteq X$ is called **dense** in (X,d) if its closure is the whole of X, i.e.,*

$$\bar{A} = X.$$

Example 1.8.2. *Let $X = [0,1]$ and let $A = (0,1)$. Clearly $A \subset \bar{A}$. Also, the points 0 and 1 are in the closure of A, as the sequence $x_n := \frac{1}{n} \subset A$ converges to 0 and the sequence $x_m := 1 - \frac{1}{m} \subset A$ converges to 1. It follows that $(0,1)$ is dense in $[0,1]$.*

Definition 1.8.3. *A space (X,d) is called **separable** if there exists a **countable** set (i.e., sequence)*

$$A = (x_n)_{n\in\mathbb{N}} \subseteq X,$$

which is dense in X.

In other words Definition 1.8.3 says that *for each $x \in X$* there exists a *subsequence* $(x_{n_m})_{m\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that

$$x = \lim_{m\to\infty} x_{n_m},$$

which means that x is a *cluster point* of $(x_n)_{n\in\mathbb{N}}$. Or equivalently,

$$\forall x \in X,\ \forall \epsilon > 0,\ \exists x_n \in A:\ d(x_n,x) < \epsilon;$$

$$\text{or equivalently, } \forall x \in X,\ \forall \epsilon > 0:\ A \cap B_\epsilon(x) \neq \emptyset.$$

Definition 1.8.4. *A metric space is called a **Polish space** if it is* complete *and* separable*.*

1.9 Basic Examples revisited

(i) **Euclidean space** $(\mathbb{R}^n, |\cdot|)$ of vectors $x = (x_i)_{i=1}^n$ with metric

$$d(x,y) = |x-y| := \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

This is a Polish (normed) space with the dense set $\mathbb{Q}^n$ (consisting of vectors with rational components $x_i \in \mathbb{Q}$).

(ii) **Manhattan** (or taxicab, or city-block, or l^1-) metric (norm) on $\mathbb{R}^n$ (in particular, $n = 2$):

$$d(x,y) = \|x-y\|_{l^1} := \sum\nolimits_{i=1}^{n} |x_i - y_i|.$$

(We cannot cut the corners and walk along the streets). Again this is a separable Banach space.

(iii) **British rail** (or post) metric on $\mathbb{R}^n$ (with centre 0:=London):

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

Exercise: Show that Euclidean space with the British rail metric is complete but **not** separable. (Hint: recall that for every $x \neq 0$, the set $\{x\}$ is open).

(iii) **French Metro** metric on $\mathbb{R}^n$ (with centre 0:=Paris):

$$d(x,y) = \begin{cases} |x-y|, & x = cy, \ \ c \in \mathbb{R} \\ |x| + |y|, & x \neq y. \end{cases}$$

Exercise: Show that Euclidean space with the French Metro metric is **not** separable. (Hint: think in polar coordinates)

(iv) $C([a,b])$ – Banach space of all **continuous** functions on a **bounded** interval $[a,b]$ with the maximum norm

$$\|f-g\|_\infty := \max_{t\in[a,b]} |f(t) - g(t)|.$$

For completeness of $C([a,b])$ see Lemma 1.9.3 below. This space is **separable**: for instance, $\bar{A} = C([a,b])$, where A is the set of all polynomials with rational coefficients;

$$P(t) := a_0 t^N + a_1 t^{N-1} + \cdots + a_{N-1} t + a_N, \ \ a_i \in \mathbb{Q}, \ \ N \in \mathbb{N}.$$

(v) $\mathbb{R}^\infty$ – space of **all real sequences** $x = (x_i)_{i\geq 1}$ with $x_i \in \mathbb{R}$. The metric (which cannot be generated by any norm) is given by

$$d(x,y) := \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}.$$

This is a Polish space, a countable dense set A consists e.g. of all *finite* sequences with *rational* coefficients $(q_1, q_2, \ldots, q_N, 0, 0, \ldots)$, $q_i \in \mathbb{Q}$, $N \in \mathbb{N}$. For a sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}^\infty$ with $x_n = (x_{n,i})_{i\geq 1} = (x_{n,1}, x_{n,2}, \ldots, x_{n,i}, \ldots)$,

$$d(x_n, x) \underset{n\to\infty}{\to} 0 \iff x_{n,i} \underset{n\to\infty}{\to} x_i, \ \forall i \in \mathbb{N},$$

i.e., the convergence in the above metric d is equivalent to the **coordinate** convergence for each fixed i.

(vi) l_∞ – Banach space of all **bounded sequences**

$$l_\infty := \left\{ x = (x_i)_{i\geq 1} \in \mathbb{R}^\infty \mid \sup_{i\geq 1} |x_i| < \infty \right\}$$
$$\text{with } d(x,y) := \|x - y\|_\infty := \sup_{i\geq 1} |x_i - y_i|.$$

$$x_n \underset{n\to\infty}{\to} x \text{ in } l_\infty \iff \text{coordinate convergence } x_{n,i} \underset{n\to\infty}{\to} x_i$$
$$\textbf{uniformly} \text{ w.r.t } i \in \mathbb{N},$$
$$\text{i.e., } \sup_{i\in\mathbb{N}} |x_{n,i} - x_i| \to 0 \text{ as } n \to \infty.$$

Warning! This space is complete but **not separable**! Define its subset

$$B := \{ x = (x_i)_{i\in\mathbb{N}} \in l_\infty \mid x_i = 0 \text{ or } 1 \text{ for each } i \in \mathbb{N} \}.$$

This subset is **not countable** (its cardinality is that of the continuum). But $d(x,y) = 1$ for all $x, y \in B$, $x \neq y$. If there exists a set A that is dense in l_∞, then in each of the balls $B_{1/2}(y)$, $y \in B$, there should be **at least one** point $x \in A$. Such balls do not intersect, which means that A is also uncountable. This contradicts the assumption that l_∞ is separable.

(vii) Polish spaces of p-**summable sequences** l_p, $1 \leq p < \infty$,

$$l_p := \left\{ x = (x_i)_{i\geq 1} \in \mathbb{R}^\infty \;\middle|\; \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}$$
$$\text{with the norm } \|x\|_p := \sqrt[p]{\sum_{i=1}^{\infty} |x_i - y_i|^p}.$$

Exercise 1.9.1. (not trivial) *Check completeness of* l_p.

(viii) $L_p([a,b])$ – Banach space of all **Lebesgue** p**-integrable functions**, $1 \le p < \infty$, with

$$d(x,y) = \|x-y\|_{L_p} := \left(\int_a^b |x(t)-y(t)|^p dt\right)^{1/p}.$$

A dense set A is the set of all polynomials with rational coefficients.

Minkovski's inequality for functions

$$\left(\int_a^b |x(t)+y(t)|^p dt\right)^{1/p} \le \left(\int_a^b |x(t)|^p dt\right)^{1/p} + \left(\int_a^b |y(t)|^p dt\right)^{1/p}.$$

Hölder's inequality for functions

$$\int_a^b |x(t)y(t)|dt \le \left(\int_a^b |x(t)|^p dt\right)^{1/p} \cdot \left(\int_a^b |y(t)|^q dt\right)^{1/q}$$
$$x \in L_p,\ \ y \in L_q,\ \ \frac{1}{p}+\frac{1}{q} = 1\ \ (p,q>1).$$

(ix) $C^k([a,b])$ for $k = 1,2,...$, Banach space of k-**times continuously differentiable** functions $x\colon [a,b] \to \mathbb{R}$, with the norm

$$\|x\|_{C^k} := \sum_{i=0}^{k} \max_{t\in[a,b]} |x^{(k)}(t)|,\ \ x^{(0)}(t) := x(t).$$

Again, a countable dense set in $C^k([a,b])$ are polynomials with rational coefficients.

General spaces of bounded continuous functions

Let (X,d) be a metric space. Define

$$C_b(X) := \{f\colon X \to \mathbb{R} \mid f \text{ is continuous and bounded on } X\}$$

with the **supremum** (not maximum) norm

$$\|f\|_{C_b} = \|f\|_\infty := \sup_{x\in X} |f(x)|.$$

Compare:

(a) $C_b((-\infty,\infty)) = C_b(\mathbb{R})$ with the sup-norm; but

(b) $C_b([a,b])$ with the max-norm (since $\max f = \sup f$ on a bounded closed interval $[a,b]$).

In general, we do not assume that X is compact (as will be introduced in Section 1.12).

The convergence of $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ to $f \in C_b(X)$ in the sup-norm is equivalent to the **uniform** convergence.

Definition 1.9.2. *Let f_n, f be functions on X. We say that $f_n \to f$ **uniformly on** X (**notation:** $f_n \rightrightarrows f$) as $n \to \infty$ if*

$$\forall \epsilon > 0,\ \exists N(\epsilon) \in \mathbb{N}:\ |f_n(x) - f(x)| < \epsilon,\ \forall x \in X \text{ as } n > N(\epsilon);$$
$$\text{which implies } \|f_n - f\|_\infty := \sup_{x\in X} |f_n(x) - f(x)| \le \epsilon \text{ (here } \le \text{ and not } <\text{).}$$

Lemma 1.9.3. *Let a sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ converge **uniformly** to some function $f\colon X \to \mathbb{R}$. Then certainly $f \in C_b(X)$.*

Proof. From the uniform convergence $f_n \rightrightarrows f$ we have

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N}:\ \sup_{x\in X} |f_N(x) - f(x)| < \epsilon/3.$$

Fix now some $x \in X$. Since $f_N \in C_b(X)$,

$$\exists \delta > 0:\ \ d(x,y) < \delta \implies |f_N(x) - f_N(y)| < \epsilon/3.$$

Thus for all such $y \in B_\delta(x)$ and for $N \in \mathbb{N}$ as above

$$|f(x) - f(y)| \le |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon,$$

which means that f is continuous at each $x \in X$. □

Corollary 1.9.4. $(C_b(X), \|\cdot\|_\infty)$ *is complete.*

Proof. Let $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ be Cauchy w.r.t. $\|\cdot\|_\infty$. Recall that any Cauchy sequence is bounded, i.e.,

$$\sup_{n\ge 1} \|f_n\|_\infty := C < \infty.$$

Obviously $(f_n(x))_{n\in\mathbb{N}} \subset \mathbb{R}$ is Cauchy for each fixed $x \in X$ and hence

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

Furthermore, for each $x \in X$

$$|f(x)| \leq \sup_{n \geq 1} |f_n(x)| \leq \sup_{n \geq 1} \|f_n\|_\infty,$$
$$\text{i.e., } \|f\|_\infty \leq C.$$

The required continuity of $f\colon X \to \mathbb{R}$ then follows from Lemma 1.9.3 as soon as we check that $f_n \rightrightarrows f$ on X.

So, it remains to prove that $\|f - f_n\|_\infty \to 0$ as $n \to \infty$. Indeed, by the Cauchy property of $(f_n)_{n\in\mathbb{N}}$

$$\forall \epsilon > 0,\ \exists N(\epsilon) \in \mathbb{N}:\ \ |f_n(x) - f_m(x)| < \epsilon/2,\ \forall n, m > N(\epsilon),\ \forall x \in X.$$

Thus, for each fixed $n > N(\epsilon)$ and $x \in X$, in the above estimate we can pass to the limit $f_m(x) \to f(x)$ as $m \to \infty$ and get

$$\forall \epsilon > 0,\ \exists N(\epsilon) \in \mathbb{N}:\ \ |f_n(x) - f(x)| \leq \epsilon/2,\ \forall n > N(\epsilon),\ \forall x \in X,$$
$$\text{i.e., } \forall \epsilon > 0,\ \exists N(\epsilon) \in \mathbb{N}:\ \ \|f - f_n\|_\infty \leq \epsilon/2 < \epsilon,\ \forall n > N(\epsilon).$$

□

Warning: Pointwise convergence $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ to $f\colon X \to \mathbb{R}$ does **not** guarantee that $f \in C_b(X)$.

Example 1.9.5. $X = [0,1], f_n(x) := x^n$,

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 0, & 0 \leq x < 1, \\ 1, & x = 1. \end{cases}$$

So, $f_n \to f$ pointwise, but $f \notin C[0,1]$. This immediately says us that f_n cannot converge uniformly on $[0,1]$, otherwise by Lemma 1.9.3 we should have $f \in C[0,1]$.

1.10 Linear Mappings (Operators) in Normed Spaces

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and let the mapping $L\colon X \to Y$ be **linear**, i.e.,

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y),\ \forall x, y \in X,\ \forall \alpha,\ \beta \in \mathbb{R},$$
$$\text{in particular, } L\underbrace{0}_{\in X} = \underbrace{0}_{\in Y}.$$

Theorem 1.10.1 (Characterisation of Continuous Linear Mappings)**.** *The following are equivalent:*

(i) L is ***uniformly continuous*** *on X;*

(ii) L is ***continuous*** *on X;*

(iii) L is ***continuous only at*** *$0 \in X$ (or* ***at some*** *$x \in X$);*

(iv) L is ***bounded****, in the sense it has the bounded* ***operator norm***

$$\|L\| = \|L\|_{X \to Y} := \sup_{\|x\|_X \leq 1} \|Lx\|_Y < \infty;$$

(v) $\exists C \in (0, \infty)$: $\|Lx\|_Y \leq C\|x\|_X$, $\forall x \in X$.

Proof. (i) → (ii) → (iii) are all trivial.

(iii) → (iv) By continuity of L at $x = 0$

$$\exists \delta > 0 : \ \|x\|_X < \delta \implies \|Lx\|_Y < 1.$$

Then $\forall x \in X$, $\|x\|_X \leq 1$,

$$\left\|\frac{\delta}{2}x\right\|_X \leq \frac{\delta}{2} < \delta \implies \|Lx\|_Y = \left\|L\left(\frac{2}{\delta} \cdot \frac{\delta}{2}x\right)\right\|_Y = \frac{2}{\delta}\left\|L\left(\frac{\delta}{2}x\right)\right\|_Y < \frac{2}{\delta}.$$

(iv) → (v) Let $x \neq 0$, then

$$\|Lx\|_Y = \|x\|_X \cdot \left\|L\left(\frac{x}{\|x\|_X}\right)\right\|_Y \leq \|x\|_X \cdot \|L\|.$$

(v) → (i) Let $\|Lx\|_Y \leq C\|x\|_X$, $\forall x \in X$. Then

$$\|Lx - Ly\|_Y = \|L(x - y)\|_Y \leq C\|x - y\|_X, \ \forall x, y \in X.$$

The definition of uniform continuity follows with any $\epsilon > 0$ and $\delta = \epsilon/L > 0$. □

Example 1.10.2. *Let $X = C([a, b])$ with $\|f\|_\infty := \max_{t \in [a,b]} |f(t)|$.*
Define $I: C([a, b]) \to \mathbb{R}$ by

$$If := \int_a^b f(t)dt \ \textit{(Riemannian integral).}$$

The linear map I is ***continuous*** *since*

$$|I(f)| \leq \int_a^b |f(t)|dt \leq (b - a)\|f\|_\infty, \ \textit{i.e.,}\ \|I\| \leq (b - a) < \infty.$$

Actually, by taking $f_0(t) \equiv 1$ with $\|f_0\|_\infty = 1$, we see that $I(f_0) = (b - a)$ and hence $\|I\| = (b - a)$.

Example 1.10.3. *Let $X = C^1([0,1])$ with the same sup-norm $\|f\|_\infty$. Define $D\colon C^1([0,1]) \to C([0,1])$ by*

$$Df := f' \text{ (derivative)}.$$

Then D is linear but ***not continuous!***

Proof. Take

$$f_n(t) := t^n, t \in [0,1].$$

Obviously, $f_n \in C^1([0,1])$ and $\|f_n\|_\infty = 1$. But

$$Df_n(t) = nt^{n-1} = nf_{n-1}(t)$$

and

$$\|Df_n\|_\infty = n.$$

So,

$$\|D\| := \sup_{f \in C^1,\ \|f\|_\infty \leq 1} \|Df\|_\infty \geq \sup_n \|Df_n\|_\infty = \infty.$$

□

Aside: Equivalent definitions of the operator norm

Above we have defined the **operator norm** of L as

$$\|L\| = \|L\|_{X \to Y} := \sup_{\|x\|_X \leq 1} \|Lx\|_Y. \tag{$*$}$$

In the literature you can also meet the following definitions

$$\|L\| := \sup_{\|y\|=1} \|Ly\|_Y, \tag{$**$}$$

$$\|L\| := \sup_{z \neq 0} \frac{\|Lz\|_Y}{\|z\|_X}. \tag{$***$}$$

We claim all three definitions are **equivalent**.

Indeed, by the linearity of L

$$\sup_{z \neq 0} \frac{\|Lz\|_Y}{\|z\|_X} = \sup_{z \neq 0} \frac{\left\| L\left(\frac{z}{\|z\|_X} \cdot \|z\|_X\right) \right\|_Y}{\left\| \frac{z}{\|z\|_X} \cdot \|z\|_X \right\|_X} = \sup_{z \neq 0} \frac{\|z\|_X \cdot \left\| L\left(\frac{z}{\|z\|_X}\right) \right\|_Y}{\|z\|_X \cdot \left\| \frac{z}{\|z\|_X} \right\|_X}$$

$$= \sup_{\|y\|_X = 1} \frac{\|Ly\|_Y}{\|y\|_X} = \sup_{\|y\|_X = 1} \|Ly\|_Y \leq \sup_{0 \leq \|x\|_X \leq 1} \|Lx\|_Y = \sup_{0 < \|x\|_X \leq 1} \|Lx\|_Y$$

(since $L0 = 0$). Note that for each $x \in X$ with $0 < \|x\|_X \leq 1$

$$\begin{aligned} \|Lx\|_Y &= \left\| L\left(\|x\|_X \cdot \frac{x}{\|x\|_X} \right) \right\|_Y = \underbrace{\|x\|_X}_{\leq 1} \left\| L\left(\tfrac{x}{\|x\|_X} \right) \right\|_Y \\ &\leq \left\| L\left(\tfrac{x}{\|x\|_X} \right) \right\|_Y = \|Ly\|_Y , \end{aligned}$$

where

$$y := \frac{x}{\|x\|_X} \in X \text{ obeys } \|y\|_X = 1.$$

Hence,

$$\sup_{\|y\|=1} \|Ly\|_Y \leq \sup_{0<\|x\|_X\leq 1} \|Lx\|_Y \leq \sup_{\|y\|=1} \|Ly\|_Y.$$

So, we have proved that

$$\sup_{z\neq 0} \frac{\|Lz\|_Y}{\|z\|_X} = \sup_{\|x\|_X\leq 1} \|Lx\|_Y = \sup_{\|y\|=1} \|Ly\|_Y.$$

1.11 Compact Sets (Heine-Borel, Bolzano-Weierstrass)

Definition 1.11.1. *Let $A \subseteq X$*

*(i) An **open cover** of A is a family of open sets $(U_i)_{i\in I} \subseteq X$ (indexed by an arbitrary set I) such that*

$$A \subseteq \bigcup_{i\in I} U_i.$$

*(ii) A is **compact** if every open cover $(U_i)_{i\in I}$ of A has a finite subcover*

$$A \subseteq \bigcup_{1\leq k\leq N} U_{i_k} = U_{i_1} \cup \cdots \cup U_{i_N}.$$

Example 1.11.2. *Let $(x_n)_{n\geq 1}$ be a convergent sequence in X with $\lim_{n\to\infty} x_n = x$. Then*

$$A := \{x_n\}_{n\geq 1} \cup \{x\}$$

is compact.

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

$$x \in A \subseteq \bigcup_{i\in I} U_i \implies x \in U_{i_0} \text{ for some } i_0 \in I.$$

As U_{i_0} is open, we can find $\epsilon > 0$ such that $x \in B_\epsilon(x) \subseteq U_{i_0}$. As $x_n \to x$, there exists $N_0 \in \mathbb{N}$ such that

$$x_n \in B_\epsilon(x) \subseteq U_{i_0} \text{ for all } n > N_0.$$

Now choose some sets from the open cover such that $U_{i_n} \ni x_n$, $1 \le n \le N_0$. Then

$$A \subseteq \bigcup_{0 \le k \le N_0} U_{i_k} = U_{i_0} \cup U_{i_1} \cup \cdots \cup U_{i_N}.$$

□

Warning: Without the limit point $\{x\}$ the argument would not work! Indeed, see the counterexample below:

Exercise 1.11.3. *Show that the set $A = \{1/n\}_{n \ge 1}$ is **not** compact in $\mathbb{R}$.*

Hint: *Consider e.g. the following open cover $\bigcup_{n \in \mathbb{N}} U_n \supseteq A$:*

$$U_1 = \left(\frac{1}{2}, 2\right) \ni 1,\ U_n = \left(\frac{1}{n+1}, \frac{1}{n-1}\right) \ni 1/n,\ n \ge 2.$$

Theorem 1.11.4. *Any compact set is closed and bounded.*

Proof. We first show that $A^c := X \setminus A$ is **open**—hence that A is closed. Take any $x \in A^c$ and define a family of open sets

$$U_n := \{y \in X \mid d(y, x) > 1/n\} = \left[\overline{B_{1/n}(x)}\right]^c,\ n \in \mathbb{N}.$$

By construction, $U_1 \subseteq \cdots \subseteq U_n \subseteq U_{n+1} \subseteq \cdots$ and

$$A \subseteq X \backslash \{x\} = \bigcup_{n=1}^{\infty} U_n.$$

Since A is compact, there exists a *finite* subcover

$$A \subseteq \bigcup_{k=1}^{N} U_{n_k}.$$

Thus,

$$A \subseteq U_K = \left[\overline{B_{1/K}(x)}\right]^c \text{ with } K = \max\{n_1, \dots, n_N\},$$
$$\text{or } B_{1/K}(x) \subseteq \overline{B_{1/K}(x)} \subseteq A^c.$$

(Note that $A \subseteq B \iff B^c \subseteq A^c$.) Since $x \in A^c$ is arbitrary, this means that A^c is open.

We now show that A is **bounded**. Take any $x \in A$, then

$$A \subseteq X = \bigcup_{n=1}^{\infty} B_n(x).$$

By compactness,

$$A \subseteq B_{n_1}(x) \cup \cdots \cup B_{n_N}(x).$$

Then

$$A \subseteq B_K(x) \text{ with } K = \max\{n_1, \ldots, n_N\},$$

which means that A is bounded. $\square$

Theorem 1.11.5. *Let $A \subseteq K \subseteq X$, where K is compact and $A = \bar{A}$ is closed in X. Then A is also compact.*

Proof. By our assumption $A^c := X \backslash A$ is open. Let $(U_i)_{i \in I}$ be an open cover of A. Then $(U_i)_{i \in I}$ and A^c together constitute an open cover of K. Since K is compact, there exists a finite subcover

$$K \subseteq U_{i_1} \cup \cdots \cup U_{i_N} \cup A^c.$$

As $A \subseteq K$ and $A \cap A^c = \emptyset$, this yields $A \subseteq U_{i_1} \cup \cdots \cup U_{i_N}$. $\square$

The next theorem is a famous and very important result which allows one to more easily determine exactly when a subset of Euclidean space is compact. It should be memorised!

Theorem 1.11.6 (Heine-Borel)**.** *Let $A \subseteq \mathbb{R}^n$. Then*

$$A \text{ is } \textbf{compact} \iff A \text{ is } \textbf{closed} \text{ and } \textbf{bounded}.$$

Proof. ($\Longrightarrow$) already done in Theorem 1.11.4.

($\Longleftarrow$) Idea: Let$A = \bar{A}$ and A be bounded. Then there exists a quader $[a, b]^n \supseteq A$. By Th. 1.18 it would suffice to show that this quader is a compact set in $\mathbb{R}^n$. We omit the proof here (which is based on Cantor's intersection theorem, cf. Th. 1.14). $\square$

Warning: This theorem (more precisely, its sufficient part) holds only in $\mathbb{R}^n$ or in **finite dimensional** spaces, see Rieszcz Theorem below.

Definition 1.11.7. *A set $A \subseteq X$ is called **sequentially compact** if every sequence $(x_n)_{n \geq 1} \subseteq A$ has a convergent subsequence $(x_{n_k})_{k \geq 1}$ whose limit belongs to A:*

$$\exists \lim_{k \to \infty} x_{n_k} =: x \in A.$$

In other words, every $(x_n)_{n \geq 1} \subseteq A$ has at least one cluster point $x \in A$.

Definition 1.11.8. *A set $A \subseteq X$ is **totally bounded** if for any $\epsilon > 0$ it can be covered by a finite family of balls*

$$B_\epsilon(x_1), \dots, B_\epsilon(x_N) \text{ with } x_1, \dots, x_N \in X,\ N \in \mathbb{N}.$$

*The set $\{x_1, \dots, x_N\} \subset X$ is called an ϵ-**net** for A, i.e., $\{x_1, \dots, x_N\} \subset A$.*

Exercise 1.11.9. *Show that in Definition 1.11.8 one can always choose an ϵ-net consisting of points from A.*

The total boundedness is **much stronger** than the usual boundedness, but in $\mathbb{R}^n$ they are equivalent! (Since $\forall R, \epsilon > 0$: $B_R(0) \subset \cup_{k=1}^N B_\epsilon(x_k)$ with proper $N = N(R, \epsilon) \in \mathbb{N}$ and $x_k \in \mathbb{R}^n$, $1 \le k \le N$.)

Theorem 1.11.10 (Characterisation of Compact Sets)**.** *Let (X, d) be a metric space. For any $A \subseteq X$, the following claims are equivalent:*

(i) A is compact;

(ii) A is sequentially compact;

(iii) (A, d) is complete and A is totally bounded in X.

Remark 1.11.11. *If (X, d) is **complete**, then for each $A \subseteq X$*

$$\textit{metric space } (A, d) \textit{ is } \textbf{complete} \iff \textit{the set } A \textit{ is } \textbf{closed} \textit{ in } X,\ A = \bar{A}.$$

*(i) $\iff$ (ii) is known as the **Bolzano-Weierstrass theorem**;*

*(ii) $\iff$ (iii) is known as the **Hausdorff criterion**.*

Proof. We prove here only some of the above implications.

(i) $\Longrightarrow$ (ii) Let A be compact, and consider any sequence $(x_n)_{n\ge1} \subseteq A$. Suppose that $(x_n)_{n\ge1}$ does not have cluster points in A. Thus, for any $y \in A$ we can find a ball $B_{\epsilon(y)}(y)$ of radius $\epsilon(y) > 0$ containing only **finitely** many (possibly, even zero) x_n's. By compactness,

$$A \subseteq B_{\epsilon_1}(y_1) \cup \dots \cup B_{\epsilon_N}(y_N), \text{ for some } N \in \mathbb{N},$$

where we denote $\epsilon_1 := \epsilon(y_1) > 0, \dots, \epsilon_N := \epsilon(y_N) > 0$. Thus, A contains only **finitely** many terms of the sequence $(x_n)_{n\ge1}$, i.e., $x_n \notin A$ for all n larger than some $N_0 \in \mathbb{N}$. This contradicts the initial assumption that $(x_n)_{n\ge1} \subseteq A$.

Related Claim: Any sequentially compact set A is closed (compare with Theorem 1.11.4 saying that any compact set is closed).

Let us prove this by contradiction. Suppose that A is not closed, then by Theorem 1.11.4 there exists $(x_n)_{n\geq 1} \subseteq A$ such that

$$x_n \to x \in A^c \text{ as } n \to \infty.$$

Then any subsequence $(x_{n_k})_{k\geq 1}$ also converges to this x. This contradicts with the sequential compactness of A claiming that $x \in A$.

(ii) $\implies$ (iii) (a) We show that the metric space (A, d) is **complete**. Take any Cauchy sequence $(x_n)_{n\geq 1} \subseteq A$. By sequential compactness, there exists a subsequence $(x_{n_k})_{k\geq 1}$ that converges to some $x \in A$. So, $(x_n)_{n\geq 1}$ has a cluster point in A. But $(x_n)_{n\geq 1}$ is Cauchy, and from the very definition (see Exercise 1.11.12 below) any Cauchy sequence can have **at most one** cluster point which (provided it exists) will also be its limit point. This means that $(x_n)_{n\geq 1}$ is **convergent** to $x \in A$.

(b) Let us show that A is **totally bounded**. Suppose not, i.e., for some $\epsilon > 0$ we cannot find a finite ϵ-net for A. Take any $x_1 \in A$ and let $U_1 := B_\epsilon(x_1)$. By assumption $\exists x_2 \in A$ with $x_2 \notin U_1$. Let $U_2 := B_\epsilon(x_2)$, then $\{U_1, U_2\}$ is still not a cover of A, which implies $\exists x_3 \in A$ with $x_2 \notin U_1 \cup U_2$. Put $U_3 := B_\epsilon(x_3)$, and so on...

Consider the sequence $(x_n)_{n\geq 1}$, by construction $d(x_n, x_m) \geq \epsilon$ for all n, m. Clearly, $(x_n)_{n\geq 1}$ is not a Cauchy sequence, and hence it cannot contain some convergent subsequence $(x_{n_k})_{k\geq 1}$. □

Exercise 1.11.12 (see also Tutorials 3)**.** *Every Cauchy sequence $(x_n)_{n\geq 1} \subseteq X$ has **at most one** cluster point. If such cluster point exists, it would also be the limit point.*

Corollary 1.11.13 (already stated as Theorem 1.11.6, **Heine-Borel**)**.** *Let $X = \mathbb{R}^n$. Then*

$$A \subseteq \mathbb{R}^n \text{ is compact } \iff A \text{ is closed and bounded.}$$

Proof. In $(\mathbb{R}^n, |\cdot|)$, which is a complete space, we have

$$A \text{ is bounded } \iff A \text{ is totally bounded.}$$
$$A = \bar{A} \iff (A, |\cdot|) \text{ is complete.}$$

So, Theorem 1.20 applies. □

Remark 1.11.14. *The following statement is known as the **Theorem of F. Riescz:** For any normed space $(X, \|\cdot\|)$ it holds:*

$$\overline{B_1(0)} \text{ is compact } \iff \dim X < \infty.$$

The **dimension** $N = \dim X < \infty$ is the **smallest** number $N \in \mathbb{N}$ such that there exists a **basis** of vectors $e_1, \ldots, e_N \in X$ allowing the presentation (as a linear combination)

$$x = \sum_{i=1}^{N} \alpha_i e_i,\ \alpha_i \in \mathbb{R},$$

for all $x \in X$.

Some examples of infinite dimensional spaces include l_p, L_p with $1 \leq p \leq \infty$; $C([0,1])$.

For instance, the unit ball $\overline{B_1(0)}$ in l_p is **not (sequentially) compact**. This ball contains a sequence of basis vectors

$$e_n := (\underbrace{0, \ldots, 0}_{n-1}, 1, 0, 0, \ldots) = (\delta_{n,i})_{i=1}^{\infty},\ n \in \mathbb{N},$$

so that

$$\|e_n\|_{l_p} = 1,\ \|e_n - e_m\|_{l_p} = \sqrt[p]{2} > 0,\ \ n \neq m,$$

and hence there are no convergent subsequences in $(e_n)_{n=1}^{\infty}$.

1.12 Continuous Functions on Compact Sets

Continuous functions defined on a compact metric space have especially useful properties.

Theorem 1.12.1. *Let (X, d) and (Y, ρ) be metric spaces, and $f\colon X \to Y$ a continuous function. Then, for each **compact** set $K \subseteq X$, its image $f(K) \subseteq Y$ is **compact**.*

Proof. We use sequential compactness. Let $(y_n)_{n\geq 1} \subseteq f(K)$, which means that $y_n = f(x_n)$ for some $x_n \in K$. By *sequential compactness* of K

$$\exists \lim_{k\to\infty} x_{n_k} =: x \in K.$$

By continuity of f and its characterisation by Theorem 1.4.4.

$$y_{n_k} = f(x_{n_k}) \to f(x) =: y \in f(K).$$

□

Theorem 1.12.2. *A continuous function $f\colon (X, d) \to (Y, \rho)$ is **uniformly** continuous on every compact set $K \subseteq X$.*

Proof. Let $\epsilon > 0$ be arbitrary. Due to continuity of f, for every $x \in K$ we can choose $\delta(x) > 0$ such that

$$\rho(f(y), f(x)) < \epsilon/2, \forall y \in B_{\delta(x)}(x). \tag{$*$}$$

The family $\left\{B_{\delta(x)/2}(x)\right\}_{x\in K}$ constitutes a (trivial) *open cover* of K. By compactness of K, it holds that

$$K \subseteq B_{\delta_1/2}(x_1) \cup ... \cup B_{\delta_N/2}(x_N), \tag{$**$}$$

for some $N \in \mathbb{N}$ and $\delta_1 := \delta(x_1) > 0, \ldots, \delta_N := \delta(x_N) > 0$. Set $\delta := \min\{\delta_1, \ldots, \delta_N\} > 0$. Now let $x, y \in K$ with $d(x, y) < \delta/2$. Because of $(**)$, there exists some x_i with $1 \leq i \leq N$, such that $d(x_i, x) < \delta_i/2$. Moreover,

$$d(x_i, y) \leq d(x_i, x) + d(x, y) < \delta_i/2 + \delta/2 \leq \delta_i := \delta(x_i).$$

Hence, both $x, y \in B_{\delta(x_i)}(x_i)$ and by $(*)$

$$\begin{aligned} \rho(f(y), f(x)) &\leq \rho(f(x), f(x_i)) + \rho(f(x_i), f(y)) \\ &< \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

□

We have the following important corollary from Theorem 1.12.2.

Theorem 1.12.3 (K. Weierstrass)**.** *Given a metric space (X, d), a nonmpty compact set $K \subseteq X$, and a continuous function $f : K \to \mathbb{R}$, then:*

(i) $f(K)$ is ***bounded****;*

(ii) f attains both its ***maximum*** *and* ***minimum*** *on K. That is, there exist points $x_{\max}$, $x_{\min} \in K$ such that*

$$f(x_{\max}) = \sup_{x\in K} f(x) = \max_{x\in K} f(x), \quad f(x_{\min}) = \inf_{x\in K} f(x) = \min_{x\in K} f(x).$$

Proof. From Theorem 1.12.2, $f(K)$ is a *compact* set in $\mathbb{R}$. Therefore, by Theorem 1.11.4, $f(K)$ is *closed* and *bounded* in $\mathbb{R}$. Thus, $\sup_{x\in K} f(x)$ and $\inf_{x\in K} f(x)$ exist (i.e, are finite) and, since $f(K)$ is closed, they belong to $f(K)$. Indeed, $\sup_{x\in K} f(x) = \lim_{n\to\infty} f(x_n)$ for some sequence $(x_n)_{n\in\mathbb{N}} \subset K$ and hence $\sup_{x\in K} f(x) \in \overline{f(K)} = f(K)$, i.e., there exists some (not necessarily unique) $x_{\max} \in K$ such that $\sup_{x\in K} f(x) = f(x_{\max})$. The same argument works for $\inf_{x\in K} f(x)$. □

We are now in a position to state what an optimisation problem is. The goal of this course will be to study when such problems have solutions and, if they do have a solution, how many solutions are there and what are their values?

Optimization Problems for $f\colon K \to \mathbb{R}$

f – **objective function**;

K – **constraint set**.

$$\left.\begin{array}{c}\text{Maximize}\\ \text{Minimize}\end{array}\right\} f(x) \text{ subject to } x \in K.$$

Notation:

$$\max\{f(x) \mid x \in K\},\ \min\{f(x) \mid x \in K\}.$$

The Weierstrass extreme value theorem is a powerful tool. But it says nothing about *how* to find these extrema. Concrete (e.g., numerical) ways to do this will be the subject of Part 3 of this course.

Some Applications: Best Approximation

Problem 1.12.4. *Let (X, d) be a metric space with nonempty subset K. Given some $x \notin K$, find the "closest" element to x in K.*

Proposition 1.12.5. *Let K be a compact set. Then for every $x \notin K$ there exists $y_0 \in K$ (not necessarily unique) such that*

$$d(x, y_0) = d(x, K) := \inf\{d(x, y) \mid y \in K\}.$$

Proof. For a fixed $x \notin K$, consider a function $f\colon K \to \mathbb{R}$ defined by

$$f(y) := d(x, y),\ y \in K.$$

Note that $|f(y) - f(z)| \leq d(y, z)$ for any $y, z \in K$, thus f is (Lipschitz) continuous on K. But K is compact which implies that f attains its min on K. Hence, $\exists y_0 \in K$ such that

$$f(y_0) = d(x, y_0) = \inf\{d(x, y) \mid y \in K\} = d(x, K).$$

□

Problem 1.12.6. *Let $(X, \|\cdot\|)$ be a normed space and let $L \subseteq X$ be a linear subspace generated by a finite system of vectors $\{e_1, \ldots, e_N\}$, $N \in \mathbb{N}$, i.e.,*

$$L := \left\{ y \in X \;\middle|\; y = \sum_{i=1}^{N} \alpha_i e_i \text{ with } \alpha_1, \ldots, \alpha_N \in \mathbb{R} \right\},\ \dim L \leq N.$$

Given some $x \notin L$, find in $L \subset X$ the "best" approximation of x.

Proposition 1.12.7. *For every $x \notin L$ there exists $y_0 \in L$ such that*

$$\|x - y_0\| := \inf\{\|x - y\| \mid y \in L\}.$$

Proof. Let us fix some $x \notin K$. Since L is a closed set,

$$\inf\{\|x - y\| \mid y \in L\} =: \delta > 0.$$

By the definition of inf, $\forall n \in \mathbb{N}$ there exists $y_n \in L$ such that

$$\delta \leq \|x - y_n\| < \delta + 1/n. \tag{$*$}$$

The sequence $\{y_n\}_{n\in\mathbb{N}}$ is bounded, more precisely $\|y_n\| \leq \|x\| + \|x - y_n\| < \|x\| + \delta + 1 =: R$. The closed ball $\overline{B_R(0)}$ in the (finite dimensional) space $(L, \|\cdot\|)$ is a compact set which implies that $\exists y_{n_k} \to_{k\to\infty} y_0 \in \overline{B_R(0)} \subseteq L$. Passing to the limit in $(*)$ as $n_k \to \infty$, we conclude that $\|x - y_0\| = \delta$. □

Remark 1.12.8. *Such elements y_0 are **not** necessary unique!*

1.13 Equivalent Metrics and Norms

Definition 1.13.1. *There is an important notion of equivalence for metrics and norms:*

*(i) Given a set $X \neq \emptyset$, the metrics d_1 and d_2 are (topologically) **equivalent** if for any $x \in X$ and $(x_n)_{n\geq1} \subseteq X$, the sequence $x_n \to x$ in (X, d_1) if and only if $x_n \to x$ in (X, d_2).*

*(ii) Given a linear space X, the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are **equivalent** if the metrics generated by $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

Theorem 1.13.2 (Equivalence of all norms in $\mathbb{R}^n$)**.** *Let $\|\cdot\|$ be **any** norm on $\mathbb{R}^n$. Then there exist constants $m, M \in (0, \infty)$ such that*

$$m|x| \leq \|x\| \leq M|x|, \ \forall x \in \mathbb{R}^n,$$

where $|\cdot| := |\cdot|_{\mathbb{R}^n}$ is the Euclidean norm on $\mathbb{R}^n$.

Proof. Note that the norm function $(\mathbb{R}^n, \|\cdot\|) \ni x \to \|x\| \in \mathbb{R}$ is always continuous. But we claim that

$$(\mathbb{R}^n, \ |\cdot|) \ni x \to f(x) := \|x\| \ \in \mathbb{R}$$

is also **continuous** (although we now consider another norm on $\mathbb{R}^n$!). Indeed, each vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ is uniquely represented as a linear combination

$$x = x_1 e_1 + \cdots + x_n e_n,$$

where

$$e_i = (0, \dots, 0, \underbrace{1}_{i}, 0, \dots, 0),\ 1 \le i \le n,$$

is the **canonical basis** in $\mathbb{R}^n$. Then for $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$

$$\begin{aligned} |f(x) - f(y)| = |\ \|x\| - \|y\|\ | &\underbrace{\le}_{\text{inverse } \Delta-\text{inequ}} \|x - y\| = \left\| \sum_{i=1}^{n} (x_i - y_i) e_i \right\| \\ \le \sum_{i=1}^{n} |x_i - y_i| \cdot \|e_i\| &\underbrace{\le}_{\text{Cauchy inequ}} \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2} \left(\sum_{i=1}^{n} \|e_i\|^2 \right)^{1/2} \\ &= C|x - y|_{\mathbb{R}^n},\ \text{with } C := \left(\sum_{i=1}^{n} \|e_i\|^2 \right)^{1/2} < \infty, \end{aligned}$$

which means the uniform continuity of f. By Theorem 1.12.3 $x \to f(x)$ achieves its maximum M and minimum m on the unit sphere

$$S_1(0) := \{x \in \mathbb{R}^n \mid |x|_{\mathbb{R}^n} = 1\},$$

which is a compact set in $(\mathbb{R}^n, |\cdot|_{\mathbb{R}^n})$. It is easy to see that $m > 0$ (since $f(x_{\min}) = \|x_{\min}\| = 0$ implies $x_{\min} = 0 \notin S_1(0)$).

Consider now any $x \neq 0$, then $|x|_{\mathbb{R}^n} =: \alpha > 0$ and

$$\|x\| = \|x \cdot \alpha \cdot \alpha^{-1}\| = \alpha \|y\|,\ y := \alpha^{-1} x \in S_1(0).$$

Thus,

$$\begin{aligned} \alpha m &\le \|x\| \le \alpha M,\ \text{if } \alpha := |x|_{\mathbb{R}^n} > 0,\ \text{or} \\ m|x| &\le \|x\| \le M|x|,\ \forall x \in \mathbb{R}^n. \end{aligned}$$

□

Theorem 1.13.3. *In vector spaces, the norms* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are (topologically) equivalent if and only if there exist* $m, M \in (0, \infty)$ *such that*

$$m\|x\|_1 \le \|x\|_2 \le M\|x\|_1,\ \forall x \in X. \qquad (**)$$

Proof. ($\Longleftarrow$) is obvious.

($\Longrightarrow$) If $\|\cdot\|_1$ and $\|\cdot\|_2$ are (topologically) equivalent, the embedding operators

$$\begin{aligned}(X, \|\cdot\|_1) &\ni x \to \mathbb{I}_1 x := x \in (X, \|\cdot\|_2),\\ (X, \|\cdot\|_2) &\ni x \to \mathbb{I}_2 x := x \in (X, \|\cdot\|_1),\end{aligned}$$

are continuous. By Theorem 1.10.1 we immediately get $(**)$. □

Remark 1.13.4. *Obviously, $(**)$ implies the equivalence of $\|\cdot\|_1$ and $\|\cdot\|_2$, but if the metrics d_1 and d_2 are equivalent, in general one could* ***not*** *expect that there exist some $m, M \in (0, \infty)$ such that*

$$md_1(x, y) \le d_2(x, y) \le Md_1(x, y),\ \forall x, y \in X.$$

Equivalent metrics preserves continuity of functions and generate the same system of open sets.

Addendum: Schauder basis (not examinable)

Definition 1.13.5. *Let $(X, \|\cdot\|)$ be a separable Banach (i.e., complete normed) space. A sequence $(e_n)_{n\le\infty}$ (finite or countable) is called a* ***Schauder basis*** *of X if every element $x \in X$ has a* ***unique*** *presentation as a linear combination*

$$x = \sum_{n=1}^{\le\infty} \alpha_n e_n \text{ with some coefficients } \alpha_n \in \mathbb{R},$$

where the series above is convergent in the norm $\|\cdot\|$.

The uniqueness is equivalent to $\sum_{n=1}^{\le\infty} \alpha_n e_n = 0$ iff all $\alpha_n = 0$.

Example 1.13.6. *The space of sequences l_p, $1 \le p < \infty$, has a canonical basis*

$$e_n = (0, \ldots, 0, \underbrace{1}_{n}, 0, \ldots),\ n \in \mathbb{N}.$$

Clearly, each Banach space with a Schauder basis is necessarily *separable*. As a dense set, one can take all finite sums $\sum_{n=1}^{N} \alpha_n e_n$ with $\alpha_n \in \mathbb{Q}$ and $1 \le n \le N \in \mathbb{N}$.

Problem 1.13.7 (The Basis problem)**.** *Does every separable Banach space have a Schauder basis?*

The basis problem was posed by S. Banach in the 1930s. It remained open for more than 40 years and was finaly solved in 1973 by Per Enflo (born in 1944); a Norwegian mathematician (and concert pianist!). Surprisingly, the answer is *negative* as Enflo constructed a counterexample.

1.14 Back to the Reals: $\mathbb{R}^d$ as a Banach Space

We will first review of some basic facts in $\mathbb{R}$ ($d = 1$).

$\mathbb{R}$ = the real line, with the norm $|x|$ = absolute value of $x \in \mathbb{R}$.

- $(\mathbb{R}, |\cdot|)$ is a *Banach* space, which means that every Cauchy sequence is convergent.
- For a set $A \subset \mathbb{R}$, the **supremum** $\sup A \in \mathbb{R} \cup \{+\infty\}$ is the **least upper** bound for A. That is, (i) $\forall x \in A$, $x \le \sup A$; (ii) $\forall y < \sup A$, $\exists x \in A$ such that $x > y$.
- If $\sup A \in A$, we call this number the **maximum** of A: $\max A \in \mathbb{R}$.
- Analogously we define the **infimum** $\inf A \in \mathbb{R} \cup \{-\infty\}$ and the **minimum** $\min A \in \mathbb{R}$.
- The supremum property (one of basic axioms for $\mathbb{R}$; cannot be proved or disproved):

  Every nonempty set $A \subset \mathbb{R}$ which is 'bounded above' has its $\sup A \in \mathbb{R}$. That is:

  $$x \le M < \infty \textit{ for all } x \in A \iff \exists \sup A \in \mathbb{R}.$$
- The supremum property is equivalent to the *completeness* of $(\mathbb{R}, |\cdot|)$.
- Clearly, if A is bounded and closed, then it has both $\max A$ and $\min A$.

Theorem 1.14.1 (Bolzano-Weierstrass)**.** *Every bounded sequence $(x_n)_{n\ge1} \subset \mathbb{R}$ contains a convergent subsequence $(x_{n_k})_{k\ge1}$.*

Proof. Just combine Theorems 1.11.6 and 1.11.10

Theorem 1.14.2. *Every bounded above, increasing sequence $(x_n)_{n\ge1} \subset \mathbb{R}$ (such that $x_n \le x_{n+1} \le M < \infty$, $\forall n \ge 1$) converges to its supremum*

$$\exists \lim_{n\to\infty} x_n = \sup_{n\ge1} x_n \; (\le M).$$

Proof. Exercise for you to try at home. □

Proposition 1.14.3 (The algebra of limits in $\mathbb{R}$)**.** *Let $(x_n)_{n\ge1}$, $(y_n)_{n\ge1}$ be convergent sequences in $\mathbb{R}$,*

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} y_n = y.$$

Then:

(i) $\lim_{n\to\infty}(x_n + y_n) = x + y$;

(ii) $\lim_{n\to\infty}(x_n \cdot y_n) = x \cdot y$*;*

(iii) $\lim_{n\to\infty}(x_n/y_n) = x/y$ *if* $y \neq 0$*;*

(iv) *if* $x_n \leq y_n$ *for all* $n \geq 1$*, then* $x \leq y$*.*

Remark 1.14.4. *Note that (iv) is not true for "<": If* $x_n < y_n$ *for all* $n \geq 1$*, then in general* $x \leq y$*.*

Definition 1.14.5. $(x_n)_{n\geq 1} \subset \mathbb{R}$ *tends to* ***infinity*** *if*

$$\forall K > 0\, \exists N(K) \in \mathbb{N} : \ |x_n| > K \text{ for all } n > N(K).$$

Lemma 1.14.6. $x_n \underset{n\to\infty}{\to} \infty \iff 1/x_n \underset{n\to\infty}{\to} 0.$

Proof. Exercise for you to try at home. □

We now look at the multidimensional case: $\mathbb{R}^d$, $d \geq 2$.

Euclidean norm $|x| := \sqrt{\sum_{i=1}^d |x_i|^2}$, $x \in \mathbb{R}^d$.

By Theorem 1.13.2 all norms in $\mathbb{R}^d$ are *equivalent.*

Lemma 1.14.7. *A sequence* $x_n := (x_{n,1}, \ldots, x_{n,d})$*,* $n \in \mathbb{N}$*, converges in* $\mathbb{R}^d$ *to some limit* $x := (x_1, \ldots, x_d) \in \mathbb{R}^d$ *if and only if each coordinate sequence* $(x_{n,i})_{n\in\mathbb{N}}$ *converges to* x_i*,* $1 \leq i \leq d$*.*

Proof. Straightforward - (Exercise to do at home). □

Corollary 1.14.8. $(\mathbb{R}^d, |\cdot|)$ *is a Banach space.*

Proof. If $(x_n)_{n\geq 1} \subset \mathbb{R}^d$ is Cauchy, then every $(x_{n,i})_{n\in\mathbb{N}} \subset \mathbb{R}$ is Cauchy. As $\mathbb{R}$ is complete, $\exists \lim_{n\to\infty} x_{n,i} =: x_i \in \mathbb{R}$, $1 \leq i \leq d$. By Lemma 1.14.7, $x_n \to x := (x_1, \ldots, x_d) \in \mathbb{R}^d$. □

Corollary 1.14.9. *A function* $f\colon \mathbb{R} \to \mathbb{R}^d$ *is continuous if and only if each coordinate function* $f_i\colon \mathbb{R} \to \mathbb{R}$ *is continuous,* $1 \leq i \leq d$*.*

Proof. Exercise, see Lemma and Theorem 1.4.4. □

Lemma 1.14.10. *The following mappings are continuous:*

- $\mathbb{R} \times \mathbb{R} \ni (x, y) \to x + y \in \mathbb{R}$,
- $\mathbb{R} \times \mathbb{R} \ni (x, y) \to x \cdot y \in \mathbb{R}$,
- $\mathbb{R} \times \mathbb{R}\backslash\{0\} \ni (x, y) \to x/y \in \mathbb{R}$.

Proof. This follows from the algebra of limits and Theorem 1.4.4. □

Corollary 1.14.11. *Let (X, d) be a metric space and $f, g\colon X \to \mathbb{R}$ be continuous. Then the functions*

$$f + g\colon X \to \mathbb{R},\ f \cdot g\colon X \to \mathbb{R}$$

are continuous. If $g(x) \neq 0$ for all $x \in X$, then also $f/g\colon X \to \mathbb{R}$ is continuous.

Proof. $f + g$ is composition of (f, g) and $\mathbb{R} \times \mathbb{R} \ni (x, y) \to x + y \in \mathbb{R}$, which are continuous. Similarly, for $f \cdot g$ and f/g. □

Corollary 1.14.12. *Let $f\colon [a, b] \to \mathbb{R}$ be a continuous function defined on a closed bounded interval $[a, b] \subset \mathbb{R}$. Then there exist $x_{\min}$, $x_{\max} \in [a, b]$ such that*

$$f(x_{\min}) = \min_{x \in [a,b]} f(x),\ f(x_{\max}) = \max_{x \in [a,b]} f(x).$$

Proof. This follows from the Weierstrass Theorem, since $[a, b]$ is compact. □

Theorem 1.14.13 (Intermediate-Value Theorem)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then for each y_0 strictly between $f(a)$ and $f(b)$ there exists $x_0 \in (a, b)$ such that $f(x_0) = y_0$.*

Proof. (Idea) For concreteness, let $f(a) < y_0 < f(b)$. Define

$$x_0 := \sup \{x \in [a, b] \mid f(x) \leq y_0\}.$$

Check that $f(x_0) = y_0$. (see the proof of Th. 6.24 in de la Fuente). □

1.15 Summary: Structures on Vector Spaces

We have the following hierarchy of 'structures on spaces':

Topological spaces $(X, (U_i)_{i \in I})$

$\downarrow$

Metric spaces (X, d)

$\downarrow$

Banach spaces $(X, \| \cdot \|)$

$\downarrow$

Hilbert spaces $(X, \langle \cdot, \cdot \rangle)$

Definition 1.15.1 (Hilbert spaces)**.** *Let X be a vector space. The most restrictive structure on X is that of the **inner product**. By definition, this is the mapping*

$$X \times X \ni (x, y) \to \langle x, y \rangle \in \mathbb{R}$$

with the following properties:

(i) Symmetry: $\langle x, y \rangle = \langle y, x \rangle$;

(ii) Linearity: $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;

(iii) Positivity: $\langle x, x \rangle =: \|x\|^2 \geq 0$, $\langle x, x \rangle = 0 \Leftrightarrow x = 0$;

for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$.

It is easy to check that $\|x\| := \sqrt{\langle x, x \rangle}$ is a norm on X.

*The space $(X, \langle \cdot, \cdot \rangle)$ is called a **Hilbert space** if $(X, \| \cdot \|)$ is a Banach space (i.e., we have completeness of X w.r.t. $\| \cdot \|$).*

Remark 1.15.2. *The usual* dot product *for vectors $x, y \in \mathbb{R}^n$ given by $x \cdot y = \sum_{i=1}^{n} x_i \cdot y_i$ is an inner product: $\langle x, y \rangle := x \cdot y$.*

By means of $\langle x, y \rangle$ we introduce the notion of **orthogonality** for a pair of vectors $x, y \in X$:

$$x \perp y \text{ if } \langle x, y \rangle = 0.$$

This generlises the usual notion of orthogonality for vectors in $\mathbb{R}^n$ where we substitute the usual dot product $x \cdots y$ with the inner product $\langle x, y \rangle$.

Theorem 1.15.3 (Cauchy-Schwarz inequality)**.** *For all $x, y \in X$,*

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}.$$

Example 1.15.4. *The space of **square-summable sequences***

$$l_2 := \left\{ x = (x_i)_{i \geq 1} \in \mathbb{R}^\infty \;\middle|\; \sum_{i=1}^{\infty} x_i^2 < \infty \right\}$$

$$\text{with the norm } \|x\|_2 := \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2}.$$

In this space, the Cauchy-Schwarz inequality takes the form:

$$\left| \sum_{i=1}^{\infty} x_i y_i \right| \leq \sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{i=1}^{\infty} x_i^2 \right)^{1/2} \cdot \left(\sum_{i=1}^{\infty} y_i^2 \right)^{1/2}.$$

*An **Orthonomal basis** in l_2 consists of the vectors*

$$e_i = (0, \dots, 0, \underbrace{1}_{i}, 0, 0 \dots,), \quad 1 \leq i < \infty,$$

$$\langle e_i, e_j \rangle = \delta_{i,j} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

The most general notion is a **topological space**. Such spaces are described by a system of open sets $(U_i)_{i \in I}$, which is called its **topology**. But we cannot quantitatively measure the distance between two point $x, y \in X$. To do this we need some metric d on X which induces the topology. Not all topologies are induced by a metric!

1.16 Multivalued Mappings (Correspondences)

This section is not examinable.

Definition 1.16.1. *Let (X, d) and (Y, ρ) be two metric spaces. A **correspondence***

$$f \colon X \rightarrow\!\!\!\rightarrow Y$$

is a set valued mapping which assigns to each point $x \in X$ a subset $f(x) \subseteq Y$.

*Such $f \colon X \rightarrow\!\!\!\rightarrow Y$ are also called **multivalued** (i.e., "point-to-set") mappings.*

*The **graph** of a correspondence $f \colon X \rightarrow\!\!\!\rightarrow Y$ is a subset in $X \times Y$*

$$\mathrm{Graph}(f) := \{(x, y) \in X \times Y \mid y \in f(x)\} \subset X \times Y.$$

The **image** of the correspondence

$$\mathrm{Im}(f) := \{y \in Y \mid \exists x \in X : \; y \in f(x)\} \subset Y.$$

Example 1.16.2. *Some simple examples are given by:*

(i) Let $g \colon X \to Y$ be a (usual) function sending X to the whole Y. Then the inverse $f := g^{-1} \colon Y \rightarrow\!\!\!\rightarrow X$ can be thought of as a multivalued mapping;

(ii) Suppose we are given a family of functions $F(\,\cdot\,, u) \colon X \to Y$, where $u \in U$ runs over some parameter set U. Then

$$f(x) := \{F(x, u)\}_{u \in U} \subset Y$$

defines a correspondence $f \colon X \to Y$.

We would like to generalize the concept of continuity to correspondences by using the topological characterisation of continuity. A correspondence would be continuous if every preimage of an open set is open. The problem is how to define "preimage". One has two choices.

Definition 1.16.3. *Let $f\colon X \to\to Y$ be a correspondence, and let $V \subseteq Y$.*
*The **strong (or upper) inverse** of V under f is*

$$f_{\mathrm{str}}^{-1}(V) := \{x \in X \mid f(x) \subseteq V,\ f(x) \neq \emptyset\}.$$

*The **weak (or lower) inverse** of V under f is*

$$f_{\mathrm{weak}}^{-1}(V) := \{x \in X \mid f(x) \cap V \neq \emptyset\}.$$

Obviously, $f_{\mathrm{str}}^{-1}(V) \subseteq f_{\mathrm{weak}}^{-1}(V)$ for any $V \subseteq Y$.

Definition 1.16.4. *We now have some choices about how to define continuity of a correspondence:*

*(i) A correspondence $f\colon X \to\to Y$ is **upper hemicontinuous** (sometimes called **semicontinuous**) if the strong inverse $f_{\mathrm{str}}^{-1}(V)$ of every open set $V \subseteq Y$ is open;*
*(ii) A correspondence $f\colon X \to\to Y$ is **lower hemicontinuous** if the weak inverse $f_{\mathrm{weak}}^{-1}(V)$ of every open set $V \subseteq Y$ is open;*
*(iii) A correspondence $f\colon X \to\to Y$ is **continuous** if it has both properties.*

Definition 1.16.5. *Some other important properties that a correspondence can satisfy:*

*(i) A correspondence $f\colon X \to\to Y$ is **compact-valued** if every $f(x)$ is a compact set in Y;*
*(ii) A correspondence $f\colon X \to\to Y$ is called **closed** if its graph* Graph(f) *is a closed set in $X \times Y$. In more words, f is closed whenever*

$$\begin{cases} x_n \to x, \\ y_n \in f(x_n),\ y_n \to y \end{cases} \implies y \in f(x).$$

Theorem 1.16.6. *Let $f\colon X \to\to Y$ be a compact-valued correspondence.*

(i) f is upper hemicontinuous if and only if Graph(f) *is closed.*
(ii) If f is upper hemicontinuous, then for each compact $K \subseteq X$, the image

$$f(K) := \cup_{x \in K} f(x) \text{ is compact in } Y.$$

Theorem 1.16.7 (Sequential Characterisation of Continuity)**.** *As with normal functions on metric spaces, we can give a sequential charcterisations of continuity for correspondences:*

(i) A compact-valued correspondence $f\colon X \to\to Y$ is upper hemicontinuous iff for any convergent sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$, $\lim_{n\to\infty} x_n = x \in X$, every sequence $\{y_n\}_{n\in\mathbb{N}} \subset Y$, $y_n \in f(x_n)$, has a convergent subsequence $y_{n_k} \to y \in f(x)$ as $k \to \infty$.

(ii) A compact-valued correspondence $f\colon X \to\to Y$ is lower hemicontinuous iff for any convergent sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$, $\lim_{n\to\infty} x_n = x \in X$, there exists *a sequence $\{y_n\}_{n\in\mathbb{N}} \subset Y$, $y_n \in f(x_n)$, such that $y_n \to y \in f(x)$.*

Main problem in applications: *To construct* ***continuous realisations*** *(so-called sections), which are (single-valued) functions $X \ni x \to \varphi(x) \in Y$, $\varphi \in C(X,Y)$, such that $\varphi(x) \in f(x)$, $\forall x \in X$.*

Further reading: *See Section 2.11 in [A. de la Fuente] and references therein.*