2 Normed spaces

Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different.

— Johann Wolfgang von Goethe

When speaking about the complex numbers $\mathbb{C}$, we already observed that basically everything regarding convergence that can be done in $\mathbb{R}$ can be transferred to $\mathbb{C}$ by using the modulus in the complex numbers instead of the modulus in the real numbers. The notion of a *norm* further abstracts the essential properties of the modulus. Moreover, we have (at least as a set) identified $\mathbb{C}$ with $\mathbb{R}^2$ and equipped it with componentwise operations. This is a very elementary construction for finite-dimensional vector spaces. In this chapter we study *normed spaces* which generalise these concepts in the following sense: normed spaces are vector spaces equipped with a map called the norm, which plays the role of the modulus.

There are many examples of normed spaces, the simplest being $\mathbb{R}^N$ and $\mathbb{K}^N$. We will be particularly interested in the infinite-dimensional normed spaces, like the sequence spaces ℓ^p or function spaces like $C(K)$. Also the important Lebesgue spaces $L^p(\Omega, \Sigma, \mu)$ and the abstract Hilbert spaces that we will study later on will be examples of normed spaces.

2.1 Vector spaces

In this section we give a brief reminder of vector spaces and associated notions. In what follows, $\mathbb{K}$ denotes either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$.

Definition 2.1. A **vector space** E over $\mathbb{K}$ is a set E together with two maps $+\colon E \times E \to E$ (addition) and $\cdot\colon \mathbb{K} \times E \to E$ (scalar multiplication) such that the following properties are satisfied. Firstly, the pair $(E, +)$ is an **Abelian group**:

(VA1) For all $x, y, z \in E$, one has $(x + y) + z = x + (y + z)$.

(VA2) There exists an element $\mathbf{0}$, such that $x + \mathbf{0} = x$ for all $x \in E$.

(VA3) For all $x \in E$, there exists a $-x \in E$ such that $x + (-x) = \mathbf{0}$.

(VA4) For all $x, y \in E$, one has $x + y = y + x$.

Furthermore, the scalar multiplication satisfies the following properties.

(VS1) For all $x \in E$ and $\lambda, \mu \in \mathbb{K}$, one has $(\lambda + \mu) \cdot x = (\lambda \cdot x) + (\mu \cdot x)$ and $(\lambda\mu) \cdot x = \lambda \cdot (\mu x)$

(VS2) For all $\lambda \in \mathbb{K}$ and $x, y \in E$, one has $\lambda \cdot (x + y) = (\lambda \cdot x) + (\lambda \cdot y)$.

This set of axioms has several consequences, most notably, the neutral element $\mathbf{0}$ from (VA2) is unique and for every $x \in E$ its inverse element $-x$ from (VA3) is unique and equal to $(-1) \cdot x$.

In what follows, we will denote scalar multiplication by mere concatenation and write λx rather than $\lambda \cdot x$. As is customary, we will insist that scalar multiplications are carried out before additions, thus $\lambda x + y$ should be interpreted as $(\lambda x) + y$ rather than $\lambda(x + y)$. Moreover, we will write $x - y := x + (-y)$. Finally, we will usually simply write 0 for the neutral element $\mathbf{0}$ in a vector space.

*To say that a set is a **vector subspace** of a given vector space means that it is a subset that contains the identity element that is a vector space when equipped with the restriction of the same operations. It is not hard to see that to prove that a subset is a vector subspace it suffices to check that addition and scalar multiplication of elements in the subset yield elements in the subset.*

Important examples of vector spaces are the spaces $\mathbb{K}^N$, endowed with component-wise addition and scalar multiplication, i.e.

$$(x_1, \ldots, x_N) + (y_1, \ldots, y_N) = (x_1 + y_1, \ldots, x_N + y_N)$$

and

$$\lambda(x_1, \ldots, x_N) = (\lambda x_1, \ldots, \lambda x_N)$$

for all $x, y \in \mathbb{K}^N$ and $\lambda \in \mathbb{K}$. In particular, $\mathbb{R}$ is a vector space over $\mathbb{R}$ and $\mathbb{C}$ is a vector space both over $\mathbb{R}$ and $\mathbb{C}$. Another example is the vector space of all functions from a set A to $\mathbb{R}$ with respect to pointwise addition and scalar multiplication of functions. More specifically, one can consider the vector space of all functions $[0,1] \to \mathbb{R}$, which can also be written as $\mathbb{R}^{[0,1]}$. It is easily observed that the continuous functions from $[0,1]$ to $\mathbb{R}$ are a vector subspace of this space, and that the polynomial functions from $[0,1]$ to $\mathbb{R}$ are a vector subspace of the vector space of the continuous functions. Analogous statements hold for $\mathbb{C}$-valued functions.

In other words, a collection of vectors is linearly independent if there is no redundancy in it, in the sense that none of its vectors can be written as a linear combination of (finitely) many other vectors in the collection.

Definition 2.2. Let V be a vector space (over $\mathbb{K}$). A collection $(x_j)_{j \in J}$ of elements in V (where J is an arbitrary index set) is called **linearly independent** if

$$\alpha_{j_1} x_{j_1} + \ldots + \alpha_{j_N} x_{j_N} = 0$$

with $N \in \mathbb{N}$ and $\alpha_{j_k} \in \mathbb{K}$ for $k \in \{1, \ldots, N\}$ implies $\alpha_{j_1} = \ldots = \alpha_{j_N} = 0$. The **linear span** of a subset $A \subset V$ is the set

It is worthwhile pointing out that a linear combination is always a finite *sum of scaled vectors.*

$$\operatorname{span} A := \{\alpha_1 x_1 + \ldots + \alpha_N x_N : N \in \mathbb{N}, \alpha_k \in \mathbb{K}, x_k \in A, k = 1, \ldots, N\}.$$

The above elements of span A are called **linear combinations**. A linearly independent collection of elements $(x_j)_{j\in J}$ is called a **(Hamel) basis** of V if $\operatorname{span}\{x_j : j \in J\} = V$. In this case the cardinality of J is called the **dimension** of V.

We have not properly introduced the cardinality of sets. We use it here informally as 'the possibly infinite number that describes the size of the set'. For a finite set it simply is the number of its elements.

We note without proof that the dimension of a vector space is well-defined, i.e. every basis has the same cardinality. Moreover, the dimension is the largest cardinality a linearly independent collection of vectors can have. For example, the elements $(1,0,0),(0,5,0),(0,1,1)$ form a basis of $\mathbb{R}^3$. So $\dim \mathbb{R}^3 = 3$ and any collection of 4 or more vectors in $\mathbb{R}^3$ must be linearly dependent.

Vector spaces are a very suitable setting for basic geometry. Frequently the elements of vector spaces are called *points* or *vectors*. In a vector space one can speak about *lines*, *line segments* and *convex sets*.

Definition 2.3. Let V be a vector space. A **line** is a set of the form $\{\alpha x + y : \alpha \in \mathbb{K}\}$ with $x, y \in V$ and $x \neq 0$. If $x, y \in V$, the **(closed) line segment** between x and y is the set $\{\lambda x + (1-\lambda y) : \lambda \in [0,1]\}$. A subset $A \subset V$ is called **convex** if for all $x, y \in A$ the closed line segment between x and y is contained in A. A linear combination $\sum_{k=1}^{N} \lambda_k x_k$ such that $\lambda_k \in [0,1]$ with $\sum_{k=1}^{N} \lambda_k = 1$ is called a **convex combination** of $x_1, \ldots, x_N \in V$.

Exercise 2.4. Draw convex and nonconvex sets of $\mathbb{R}^2$. Think geometrically about linear combinations and convex combinations. Observe that convexity is a very strong geometric property that implies that the set cannot have holes and must be connected. Show that the set under the graph of $\log\colon (0,\infty) \to \mathbb{R}$ is convex.

2.2 Definition and basic properties of a normed space

We introduce a notion of *length* for elements of a vector space. Note that a 'length' is something that a single element on its own should have, whereas 'distance' is something that only makes sense for a pair of elements.

Definition 2.5. Let X be a vector space over $\mathbb{K}$. A **norm** on X is a map $\|\cdot\|\colon X \to [0,\infty)$ that satisfies the following three properties.

(N1) $\|x\| = 0$ implies $x = \mathbf{0}$. (definiteness)

(N2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ for all $x \in X$ and $\lambda \in \mathbb{K}$. (homogeneity)

(N3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$. (triangle inequality)

A **normed space** is a pair $(X, \|\cdot\|)$, where X is a vector space and $\|\cdot\|$ is a norm on X.

It is clear that $(\mathbb{R}, |\cdot|)$ is a normed space (over $\mathbb{R}$). In the following section we shall encounter more interesting examples of normed spaces. To practice dealing with complex numbers, we give the following example.

Example 2.6. We shall verify that $(\mathbb{C}, |\cdot|)$ is a normed space over both $\mathbb{C}$ and $\mathbb{R}$, where $|z| = \sqrt{z \cdot \overline{z}}$. It follows straight from the field axioms of $\mathbb{R}$ and the definition of the operations in $\mathbb{C}$ that $\mathbb{C}$ is a vector space over $\mathbb{C}$ and $\mathbb{R}$. So it remains to show that $|\cdot|$ is a norm on $\mathbb{C}$ (both over $\mathbb{C}$ and $\mathbb{R}$). First of all $|\cdot| \colon \mathbb{C} \to [0, \infty)$ as $|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2} \geq 0$. If $|z| = 0$, then $(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 = 0$; consequently $\operatorname{Re} z = 0$ and $\operatorname{Im} z = 0$, hence $z = 0$. So (N1) is satisfied.

To check (N2), we first let $\lambda, z \in \mathbb{C}$. Then $\lambda = \alpha + i\beta$ and $z = x + iy$. Note that

$$\overline{\lambda \cdot z} = \alpha x - \beta y - i(\alpha y + \beta x) = \overline{\lambda} \cdot \overline{z}.$$

So

$$\|\lambda \cdot z\| = \sqrt{\lambda \cdot z \cdot \overline{\lambda \cdot z}} = \sqrt{\lambda \cdot \overline{\lambda} \cdot z \cdot \overline{z}} = \sqrt{\lambda \cdot \overline{\lambda}}\sqrt{z \cdot \overline{z}} = |\lambda||z|,$$

and (N2) is satisfied for $\mathbb{K} = \mathbb{C}$. If $\lambda \in \mathbb{R}$, then $|\lambda|_{\mathbb{C}} = |\lambda|_{\mathbb{R}}$. So (N2) also holds for $\mathbb{K} = \mathbb{R}$.

Finally, let $w, z \in \mathbb{C}$. Observe that

$$\operatorname{Re}(w\overline{z}) \leq \sqrt{\operatorname{Re}(w\overline{z})^2 + \operatorname{Im}(w\overline{z})^2} = |w\overline{z}| \leq |w||\overline{z}| = |w||z|,$$

where we also used that (N2) holds. Hence

$$\begin{aligned} |z+w|^2 &= (z+w)\overline{(z+w)} = z\overline{z} + w\overline{w} + w\overline{z} + z\overline{w} \\ &= |z|^2 + |w|^2 + 2\operatorname{Re}(w\overline{z}) \\ &\leq |z|^2 + |w|^2 + 2|w||z| \\ &= (|z| + |w|)^2. \end{aligned}$$

Taking the square root on both sides yields (N3).

In a normed space the norm quantifies the length of a vector. To quantify how far a point x is from a point y in a normed space, one takes the norm of $x - y$ (which is equal to the norm of $y - x$). This allows to extend all the definitions regarding convergence from $\mathbb{R}$ to general normed spaces. It is important to note that this is a consistent extension, i.e., in $\mathbb{R}$ convergence in the sense of normed spaces agrees with the previously defined notion of convergence.

When dealing with metric spaces (or topological spaces), one encounters further consistent extensions of convergence.

Definition 2.7. Let $(X, \|\cdot\|)$ be a normed space. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is said to **converge** to $a \in X$ if

$$\forall \varepsilon > 0 \; \exists N_0 \in \mathbb{N} : \forall n \geq N_0 : \|x_n - a\| < \varepsilon.$$

In this case one writes $\lim_{n\to\infty} x_n = a$ or $x_n \to a$ as $n \to \infty$. A sequence $(x_n)_{n\in\mathbb{N}}$ in X is called a **Cauchy sequence** if

$$\forall \varepsilon > 0\ \exists N_0 \in \mathbb{N} : \forall n, m \geq N_0 : \|x_n - x_m\| < \varepsilon.$$

A subset $A \subset X$ is called **bounded** in X if there exists an $M > 0$ such that $\|x\| \leq M$ for all $x \in A$. Similarly, a sequence (x_n) in X is called **bounded** if $\sup_{n\in\mathbb{N}}\|x_n\| < \infty$. The **open ball** about $x \in X$ with radius $r > 0$ is the set

$$B(x,r) := \{y \in X : \|y - x\| < r\}.$$

In a nontrivial normed space there are always unbounded subsets due to properties (N1) and (N2).

We shall see later that the open ball deserves to be called 'open'. Also note that the notation $B(x,r)$ does not specify the normed space. As usual, if confusion is possible, one might choose to write $B_X(x,r)$, for example.

In addition, let $(Y, \|\cdot\|_Y)$ be a normed space. Let $A \subset X$ and $f\colon A \to Y$ a map. Then f is called **continuous** at $a \in A$ (as a map from $(X, \|\cdot\|_X)$ to $(Y, \|\cdot\|_Y)$), if for all sequences $(x_n)_{n\in\mathbb{N}}$ in A such that $x_n \to a$ in $(X, \|\cdot\|_X)$ one has $f(x_n) \to f(a)$ in $(Y, \|\cdot\|_Y)$. If f is continuous at all points in A, then f is called **continuous**. If there exists an $L \geq 0$ such that

$$\|f(x) - f(y)\|_Y \leq L\|x - y\|_X$$

for all $x, y \in A$, then f is called **Lipschitz continuous**, and L is called the **Lipschitz constant**.

Exercise 2.8. Show that a Lipschitz continuous function is continuous, but that there are continuous functions that are not Lipschitz continuous.

So a function is continuous if an approximation of the inputs yields a sequence of outputs that approximate the desired output. In other words, one can swap limits with the continuous function, i.e., $f(\lim x_n) = \lim f(x_n)$. In applications often only approximations are available, which makes continuity a very desirable property of functions.

Exercise 2.9. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, $A \subset X$ and $f\colon A \to Y$. Show that f is continuous at $a \in A$ if and only if

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall y \in B_X(x, \delta) : f(y) \in B_Y(f(a), \varepsilon).$$

Definition 2.10 (Product spaces)**.** Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces over the same field $\mathbb{K}$. Then $X \times Y$ is made into a vector space using the componentwise operations

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

and

$$\lambda \cdot (x, y) := (\lambda x, \lambda y)$$

for $\lambda \in \mathbb{K}$. Define $\|\cdot\|_{X\times Y}\colon X \times Y \to [0, \infty)$ by

$$\|(x,y)\|_{X\times Y} := \|x\|_X + \|y\|_Y.$$

Then $(X \times Y, \|\cdot\|_{X\times Y})$ is a normed space called the **product space** of $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$.

Exercise 2.11. Show that $(X \times Y, \|\cdot\|_{X\times Y})$ in the previous definition really is a normed space. Prove that a sequence $((x_n, y_n))_{n\in\mathbb{N}}$ converges in $(X \times Y, \|\cdot\|_{X\times Y})$ if and only if $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge in $(X, \|\cdot\|_X)$

and $(Y, \|\cdot\|_Y)$, respectively. Use induction to show that analogously the Cartesian product of normed spaces $X_1, \ldots, X_N$ for $N \in \mathbb{N}$ can be made into a normed space.

Proposition 2.12. *Let $(X, \|\cdot\|)$ be a normed space.*

1. *Every Cauchy sequence in $(X, \|\cdot\|)$ (and therefore every convergent sequence) is bounded.*

This is sometimes called the 'reverse triangle inequality'.

2. *One has*
$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \tag{2.1}$$
for all $x, y \in X$.

3. *The map $X \times X \to X$ given by $(x, y) \mapsto x + y$, the map $\mathbb{K} \times X \to X$ given by $(\alpha, x) \mapsto \alpha x$ and the map $X \mapsto \mathbb{R}$ given by $x \mapsto \|x\|$ are continuous. Here $X \times X$ and $\mathbb{K} \times X$ are to be understood as product spaces.*

Proof. 1. Let (x_n) be a Cauchy sequence in X. So for $\varepsilon = 1$ there exists an $n_0 \in \mathbb{N}$ such that $\|x_n - x_m\| \leq \varepsilon = 1$ for all $n, m \geq n_0$. Note that $\|x_n\| \leq \|x_{n_0}\| + 1$ for all $n \geq n_0$. Hence
$$\sup_{n \in \mathbb{N}} \|x_n\| \leq \max\{\|x_1\|, \ldots, \|x_{n_0-1}\|, \|x_{n_0}\| + 1\} < \infty.$$
Therefore (x_n) is bounded.

2. Observe that
$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|.$$
Therefore $\|x\| - \|y\| \leq \|x - y\|$. Swapping the role of x and y, we also obtain $-\|x\| + \|y\| \leq \|x - y\|$. This implies (2.1).

3. Recall that a sequence (x_n, y_n) converges to (x, y) in the product space $X \times X$ if and only if $x_n \to x$ in X and $y_n \to y$ in X for $n \to \infty$. Similarly, (α_n, x_n) converges to (α, x) in the product space $\mathbb{K} \times X$ if and only if $\alpha_n \to \alpha$ in $\mathbb{K}$ and $x_n \to x$ in X. So suppose (x_n, y_n) converges to (x, y) in $X \times X$. Then $\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0$ for $n \to \infty$. This shows that the map $X \times X \to X$ given by $(x, y) \mapsto x + y$ is continuous. One argues similarly for the map $\mathbb{K} \times X \to X$ given by $(\alpha, x) \mapsto \alpha x$. Note that (2.1) implies that $\|\cdot\|$ is Lipschitz continuous with Lipschitz constant 1. $\square$

Definition 2.13. A complete normed space is called a **Banach space**.

This exercise is somewhat more difficult.

Exercise 2.14. Let $(X, \|\cdot\|)$ be a normed space and (x_k) be a sequence in X. Then the partial sums $s_n := \sum_{k=1}^{n} x_k$ are well-defined. If the sequence (s_n) is convergent in $(X, \|\cdot\|)$, then we say that the series $\sum_{k=1}^{\infty} x_k$ converges. If the (real-valued) series $\sum_{k=1}^{\infty} \|x_k\|$ converges, we say that the series $\sum_{k=1}^{\infty} x_k$ is absolutely convergent.

Show that $(X, \|\cdot\|)$ is a Banach space if and only if every absolutely convergent series converges.

Hint: Given a Cauchy sequence (x_n), consider a suitable subsequence (x_{n_k}) and telescopic sums like $\sum_{k=M}^{N}(x_{n_{k+1}} - x_{n_k})$.

2.3 Examples of normed spaces

2.3.1 Finite dimensional spaces

Example 2.15. Let $1 \le p < \infty$. Then $\|\cdot\|_p$ defined by

$$\|x\|_p := \Big(\sum_{k=1}^{N} |x_k|^p \Big)^{1/p}$$

is a norm on the vector space $\mathbb{K}^N$. Similarly, $\|x\|_\infty := \max\{|x_1|, \dots, |x_N|\}$ defines a norm on $\mathbb{K}^N$. In fact, the norm properties (N1) and (N2) follow straight from the definition. The verification of property (N3) consists in establishing Minkowski's inequality, which we shall prove in the following.

Exercise 2.16. Sketch the open unit balls $B(0,1)$ in the normed spaces $(\mathbb{R}^2, \|\cdot\|_p)$ for $p = 1, 2, \infty$. How do the the open unit balls look like for general $1 < p < \infty$.

For the proof of Minkowski's inequality we first need the following inequality which is interesting in its own right.

Theorem 2.17 (Hölder's inequality). *Let $x, y \in \mathbb{K}^N$. Let $p, q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, where $\frac{1}{\infty} := 0$. Then*

$$\sum_{k=1}^{N} |x_k y_k| \le \|x\|_p \|y\|_q. \tag{2.2}$$

Proof. We only give the proof in the case where $p, q \in (1, \infty)$. The cases where p, q might be 1 or ∞ are easier and left as an exercise.

We first establish an auxiliary inequality. Let $\lambda := \frac{1}{p} \in (0,1)$. Then $1 - \lambda = \frac{1}{q}$. Since $\log\colon (0, \infty) \to \mathbb{R}$ is concave (i.e., the set under the graph of the function is convex), we obtain

$$\lambda \log(a) + (1 - \lambda) \log(b) \le \log(\lambda a + (1 - \lambda) b)$$

for all $a, b > 0$. Applying the monotonically increasing exponential function exp to both sides yields

$$a^\lambda b^{1-\lambda} \le \lambda a + (1 - \lambda) b. \tag{2.3}$$

Note that (2.3) holds for all $a, b \geq 0$.

We may assume that $x, y \neq 0$. Using the homogeneity and after rescaling, we may assume that $\|x\|_p = \|y\|_q = 1$. So it remains to show that

$$\sum_{k=1}^{N} |x_k y_k| \leq 1.$$

It follows from (2.3) that

$$|x_k y_k| = |x_k||y_k| = (|x_k|^p)^\lambda (|y_k|^q)^{1-\lambda} \leq \lambda |x_k|^p + (1-\lambda)|y_k|^q$$

for all $k = 1, \ldots, N$. Summing up and using that $\|x\|_p = \|y\|_q = 1$ gives

$$\sum_{k=1}^{N} |x_k y_k| \leq \lambda \sum_{k=1}^{N} |x_k|^p + (1-\lambda) \sum_{k=1}^{N} |y_k|^q = \lambda - (1-\lambda) = 1.$$

This establishes the inequality. $\square$

Corollary 2.18 (Minkowski's inequality)**.** *For $p \in [1, \infty]$ we have*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

for all $x, y \in \mathbb{K}^N$.

Proof. Again we only give the proof in the case where $p \in (1, \infty)$. The remaining cases are an easy exercise.

Let $q := \frac{p}{p-1}$. Then $1 = \frac{1}{p} + \frac{1}{q}$. Let $x, y \in \mathbb{K}^N$. Using Hölder's inequality, we obtain

$$\begin{aligned}
\|x + y\|_p^p &= \sum_{k=1}^{N} |x_k + y_k||x_k + y_k|^{p-1} \\
&\leq \sum_{k=1}^{N} |x_k||x_k + y_k|^{p-1} + \sum_{k=1}^{N} |y_k||x_k + y_k|^{p-1} \\
&\leq \|x\|_p \Big(\sum_{k=1}^{N} |x_k + y_k|^p \Big)^{1/q} + \|y\|_p \Big(\sum_{k=1}^{N} |x_k + y_k|^p \Big)^{1/q} \\
&= (\|x\|_p + \|y\|_p) \|x + y\|_p^{p/q}.
\end{aligned}$$

The assertion follows since $p - \frac{p}{q} = 1$. $\square$

Remark 2.19. There are many more norms in $\mathbb{K}^N$ than only positive multiples of the p-norms $\|\cdot\|_p$ for $p \in [1, \infty]$. For example, also $\|\cdot\|_p + \|\cdot\|_q$ is a norm in $\mathbb{K}^N$. We shall see later, however, that all norms in $\mathbb{K}^N$ lead to the same notion of convergence.

Finally, note that $1 = \|e_k\|_p = \|e_k\|_q$ for all $k \in \{1, \ldots, N\}$, where $e_k =$

$(0, \ldots, 1, 0, \ldots) \in \mathbb{K}^N$ has a single 1 in the kth component. Consequently a norm is not determined by its behaviour on a basis.

2.3.2 Sequence spaces

We now generalise the finite dimensional spaces $(\mathbb{K}^N, \|\cdot\|_p)$ to infinite-dimensional sequence spaces.

Definition 2.20 (The spaces ℓ^p)**.** For $p \in [1, \infty)$ we let ℓ^p be the set of all $\mathbb{K}$-valued sequences $x = (x_k)_{k \in \mathbb{N}}$ such that

$$\|x\|_p := \Big(\sum_{k=1}^{\infty} |x_k|^p \Big)^{1/p} < \infty,$$

i.e., $\ell^p = \{x = (x_k) \text{ sequence in } \mathbb{K} : \|x\|_p < \infty\}$. The set ℓ^∞ is the set of all bounded $\mathbb{K}$-valued sequences and we set

$$\|x\|_\infty := \sup_{k \in \mathbb{N}} |x_k|$$

for all $x \in \ell^\infty$.

Proposition 2.21. *Let $p \in [1, \infty]$. The set ℓ^p is an infinite-dimensional vector space over $\mathbb{K}$ with respect to scalar multiplication and componentwise addition, i.e., for $x, y \in \ell^p$ and $\alpha \in \mathbb{K}$ we set*

$$x + y := (x_k + y_k)_{k \in \mathbb{N}} \qquad \text{and} \qquad \alpha x := (\alpha x_k)_{k \in \mathbb{N}}.$$

Moreover, $\|\cdot\|_p$ is a norm on ℓ^p. Furthermore, $(\ell^p, \|\cdot\|_p)$ is a Banach space.

Proof. We already know that $\mathbb{K}^{\mathbb{N}}$ is a vector space with respect to scalar multiplication and componentwise addition. So it suffices to show that ℓ^p are vector subspaces. To this end, it suffices to show that $\alpha x, x + y \in \ell^p$ for $\alpha \in \mathbb{K}$ and $x, y \in \ell^p$. Suppose that $1 \le p < \infty$. Then applying the Minkowsi inequality to the first N components we obtain

$$\begin{aligned} \Big(\sum_{k=1}^{N} |x_k + y_k|^p \Big)^{1/p} &\le \Big(\sum_{k=1}^{N} |x_k|^p \Big)^{1/p} + \Big(\sum_{k=1}^{N} |y_k|^p \Big)^{1/p} \\ &\le \|x\|_p + \|y\|_p < \infty. \end{aligned} \tag{2.4}$$

In the following we will frequently encounter the vectors e_n. Note that the collection $(e_n)_{n \in \mathbb{N}}$ is linearly independent, but not a Hamel basis of ℓ^p. In fact, it only spans the proper subspace $\mathbf{c_{00}}$, which will be introduced later.

As this holds independently of N, it follows that $x + y \in \ell^p$. It is obvious that $\alpha x \in \ell^p$. So ℓ^p is a vector space. For all $n \in \mathbb{N}$ let $e_n = (0, \ldots, 1, 0, \ldots)$, so e_n is a sequence with a single 1 in the nth component. Clearly $e_n \in \ell^p$ for all $n \in \mathbb{N}$. Moreover, the collection $(e_n)_{n \in \mathbb{N}}$ is linearly independent in ℓ^p, which implies that the dimension of ℓ^p is at least as large as the (infinite) cardinality of $\mathbb{N}$.

Moreover, it is easily verified that $\|\cdot\|_p$ satisfies (N1) and (N2). For $1 \leq p < \infty$ property (N3) follows from (2.4) after taking the limit $N \to \infty$. Hence $(\ell^p, \|\cdot\|_p)$ is a normed space for $1 \leq p < \infty$. The corresponding statement for $p = \infty$ is an exercise.

It remains to prove that $(\ell^p, \|\cdot\|_p)$ is complete. We again suppose that $1 \leq p < \infty$ and leave the case $p = \infty$ as an exercise. So let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in ℓ^p. We write $x_n = (x_1^{(n)}, x_2^{(n)}, \ldots)$. Let $k \in \mathbb{N}$ and $\varepsilon > 0$. Then for all $n, m \geq N_0$ one has

$$\big|x_k^{(n)} - x_k^{(m)}\big| = \Big(\big|x_k^{(n)} - x_k^{(m)}\big|^p\Big)^{1/p} \leq \|x_n - x_m\|_p < \varepsilon.$$

This proof also shows that convergence in ℓ^p implies componentwise convergence. Show that the opposite implication is not true, not even in ℓ^∞!

Hence $(x_k^{(n)})_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, we define the sequence $(y_k)_{k\in\mathbb{N}}$ via $y_k := \lim_{n\to\infty} x_k^{(n)}$.

It remains to prove that $y \in \ell^p$ and that $x_n \to y$ in ℓ^p. Note that

$$\begin{aligned}\Big(\sum_{k=1}^{N} |y_k|^p\Big)^{1/p} &= \lim_{n\to\infty} \Big(\sum_{k=1}^{N} \big|x_k^{(n)}\big|^p\Big)^{1/p} \\ &\leq \limsup_{n\to\infty} \Big(\sum_{k=1}^{\infty} \big|x_k^{(n)}\big|^p\Big)^{1/p} \\ &= \limsup_{n\to\infty} \|x_n\|_p \leq M,\end{aligned}$$

where the bound $M > 0$ exists as (y_k) is a Cauchy sequence. Since the bound is independent of N, we obtain $y \in \ell^p$. Next, observe that for $\varepsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that

$$\sum_{k=1}^{N} \big|x_k^{(n)} - y_k\big|^p = \lim_{m\to\infty} \sum_{k=1}^{N} \big|x_k^{(n)} - x_k^{(m)}\big|^p \leq \limsup_{m\to\infty} \|x_n - x_m\|_p^p \leq \varepsilon$$

for all $n, m \geq N_0$. As this is independent of N it follows that $\|x_n - y\|_p^p < \varepsilon$ for all $n \geq N_0$. We have shown that the Cauchy sequence (x_n) converges in ℓ^p. □

Exercise 2.22. Let $1 \leq p \leq \infty$. Show that $(\mathbb{K}^N, \|\cdot\|_p)$ is a Banach space.

Remark 2.23. The ℓ^p spaces that we consider here are frequently also denoted by $\ell^p(\mathbb{N})$ since the elements are sequences indexed by $\mathbb{N}$. It is possible to more generally consider $\ell^p(J)$ for arbitrary index sets J, but for uncountable index sets this requires a few technical adjustments. We point out that $\ell^p(\{1, \ldots, N\})$ would directly correspond to $(\mathbb{K}^N, \|\cdot\|_p)$. Later in the chapter about measure theory we will study the so-called L^p spaces which will generalise the $\ell^p(J)$ spaces even further.

2.4 Topological notions in normed spaces

We have already introduced open balls in normed spaces.

Definition 2.24. Let $(X, \|\cdot\|)$ be a normed space. A subset $G \subset X$ is called **open** (in $(X, \|\cdot\|)$) if for all $x \in G$ there exists an $\varepsilon > 0$ such that $B(x, \varepsilon) \subset G$. A subset $F \subset X$ is called **closed** if $F^c := X \setminus F$ is open. The **interior** of $A \subset X$ is the union of all open subsets contained in A, i.e.,

Intuitively, open sets allow for some 'wiggle-room' around all of their elements.

$$\operatorname{int} A := \overset{\circ}{A} := \bigcup_{G \subset A,\, G \text{ open}} G.$$

The **closure** of $A \subset X$ is the intersection of all closed supersets of A, i.e.,

$$\operatorname{cl} A := \overline{A} := \bigcap_{A \subset F,\, F \text{ closed}} F.$$

The following properties now follow straight from the definitions and de Morgan's laws.

Proposition 2.25. *Let $(X, \|\cdot\|)$ be a normed space.*

1. *The open balls in X are open.*
2. *$\emptyset$ and X are open.*
3. *If $U_1, \dots, U_N$ are open, then $\bigcap_{k=1}^N U_k$ is open.*
4. *If $(U_j)_{j \in J}$ is a collection of open sets, then $\bigcup_{j \in J} U_j$ is open.*
5. *$\emptyset$ and X are closed.*
6. *If $A_1, \dots A_N$ are closed, then $\bigcup_{k=1}^N A_k$ is closed.*
7. *If $(A_j)_{j \in J}$ is a collection of closed sets, then $\bigcap_{j \in J} A_j$ is closed.*
8. *Let $A \subset X$. Then the interior of A is open, the closure of A is closed, and $\operatorname{int} A \subset A \subset \operatorname{cl} A$. Moreover, A is open if and only if $A = \operatorname{int} A$, and A is closed if and only if $A = \operatorname{cl} A$.*

The definition of closed sets and the closure is inconvenient for practical purposes. The following result connects these notions with the convergence of sequences.

Proposition 2.26. *Let $(X, \|\cdot\|)$ be a normed space and $A \subset X$. Then $x \in \operatorname{cl} A$ if and only if there exists a sequence $(x_n)_{n \in \mathbb{N}}$ in A such that $x_n \to x$ in X.*

Proof. Let $x \in \operatorname{cl} A$. We claim that then for all $n \in \mathbb{N}$ one has $B(x, \frac{1}{n}) \cap A \neq \emptyset$. In fact, assume for contradiction that there exists an $n_0 \in \mathbb{N}$ such that $B(x, \frac{1}{n_0}) \cap A = \emptyset$. Then $B(x, \frac{1}{n_0})^c$ is closed and contains A. Hence

$\operatorname{cl} A \subset B(x, \frac{1}{n_0})^c$, and hence $x \notin A$, which is a contradiction. So there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in A such that $x_n \in B(x, \frac{1}{n})$ for all $n \in \mathbb{N}$. Then $x_n \to x$ in X.

Conversely, suppose that $(x_n)_{n\in\mathbb{N}}$ is a sequence in A that converges to $x \in X$. We need to show that $x \in \operatorname{cl} A$. Let F be closed such that $A \subset F$. It suffices to show $x \in F$ as then by generalisation x is in the intersection of all closed supersets of A. Assume for contradiction that $x \in F^c$. As F^c is open, there exists an $\varepsilon > 0$ such that $B(x, \varepsilon) \subset F^c$. But there exists an $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ one has $x_n \in B(x, \varepsilon) \subset F^c$. This is a contradiction since $B(x, \varepsilon) \cap F = \emptyset$, but $x_{N_0} \in B(x, \varepsilon) \cap F$. We have proved that $x \in \operatorname{cl} A$. □

We obtain the following consequence.

Corollary 2.27. *Let $(X, \|\cdot\|)$ be a normed space and $A \subset X$. Then A is closed if and only if it contains the limit of every convergent sequence of elements in A, i.e., if $(x_n)_{n\in\mathbb{N}}$ is a sequence in A such that $x_n \to x$ in X, then $x \in A$.*

Exercise 2.28. Show that $\{y \in X : \|x - y\| \leq r\}$ is the closure of $B(x, r)$.

Exercise 2.29. Consider the set

$$A := \{x \in \ell^\infty : |x_n| < 1 \text{ for all } n \in \mathbb{N}\}.$$

Is A open in $(\ell^\infty, \|\cdot\|_\infty)$?

Definition 2.30. Let $(X, \|\cdot\|)$ be a normed space. A subset $A \subset X$ is called **dense** in X if $\operatorname{cl} A = X$. The normed space $(X, \|\cdot\|)$ is called **separable** if there exists a countable subset $A \subset X$ such that $\operatorname{cl} A = X$. A normed space that is not separable is called **inseparable**.

Example 2.31. As $\mathbb{Q}$ is dense in $\mathbb{R}$, the normed space $(\mathbb{R}, |\cdot|)$ is separable. Moreover, $\mathbb{C}$ is separable as $\mathbb{Q} + i\mathbb{Q}$ is dense in $\mathbb{C}$. Similarly $\mathbb{K}^N$ is separable.

Proposition 2.32. *A normed space $(X, \|\cdot\|)$ is separable if and only if there exists a countable set A such that the linear span of A is dense in X, i.e., if $X = \operatorname{cl}(\operatorname{span}(A))$.*

Proof. If X is separable, then there exists a countable set $A \subset X$ such that $\operatorname{cl}(A) = X$. As $A \subset \operatorname{span} A$, it is trivial that $X = \operatorname{cl}(\operatorname{span}(A))$.

So suppose now that $A \subset X$ is countable such that $X = \operatorname{cl}(\operatorname{span}(A))$. Let $\mathbb{L} = \mathbb{Q}$ if $\mathbb{K} = \mathbb{Q}$ and $\mathbb{L} = \mathbb{Q} + i\mathbb{Q}$ if $\mathbb{K} = \mathbb{C}$. For all $N \in \mathbb{N}$, consider the set

$$D_N := \{\alpha_1 x_1 + \ldots + \alpha_N x_N : \alpha_k \in \mathbb{L} \text{ and } x_k \in A \text{ for } k = 1, \ldots, N\}.$$

Note that D_N is countable for all $N \in \mathbb{N}$. It follows that $D := \bigcup_{N\in\mathbb{N}} D_N$ is countable.

Any element in $\mathrm{span}(A)$ can be approximated by elements in D. Here we make use of Proposition 2.12.3. It follows that $\mathrm{span}(A) \subset \mathrm{cl}\, D$. Therefore $X = \mathrm{cl}(\mathrm{span}(A)) \subset \mathrm{cl}\, D$. This shows that X is separable. □

We introduce three other vector spaces of sequences that are related to the ℓ^p spaces.

Definition 2.33. By $\mathbf{c}$ one denotes the vector subspace of $\mathbb{K}^{\mathbb{N}}$ consisting of all convergent $\mathbb{K}$-valued sequences, i.e.,

$$\mathbf{c} := \{(x_k) : (x_k) \text{ convergent sequence in } \mathbb{K}\}.$$

The vector subspace of $\mathbf{c}$ consisting of all sequences converging to zero is denoted by

$$\mathbf{c_0} := \{(x_k) : \lim_{k\to\infty} x_k = 0\}.$$

Finally, let $\mathbf{c_{00}}$ be the vector subspace of $\mathbf{c_0}$ consisting of all sequences (x_k) where $x_k \neq 0$ for only finitely many indices, i.e.,

$$\mathbf{c_{00}} := \{(x_k) : x_k = 0 \text{ for almost all indices } k \in \mathbb{N}\}.$$

Exercise 2.34. Show that while $\mathbf{c_0}$ is not a vector subspace of $(\ell^p, \|\cdot\|_p)$ for any $p \in [1, \infty)$, both $\mathbf{c}$ and $\mathbf{c_0}$ are closed vector subspaces of $(\ell^\infty, \|\cdot\|_\infty)$. Show that $\mathbf{c_{00}}$ is a nonclosed vector subspace of $(\ell^p, \|\cdot\|_p)$ for all $p \in [1, \infty]$, and that $\mathbb{K}^N$, after extending the vectors by 0 to sequences, is a closed vector subspace of $(\ell^p, \|\cdot\|_p)$ for all $p \in [1, \infty]$.

We give the proof that $\mathbf{c}$ is a closed subspace of ℓ^∞.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{c}$ such that $x_n \to y$ in ℓ^∞. We need to show that $y \in \mathbf{c}$. Firstly, for all $n \in \mathbb{N}$ there exists an $\alpha_n \in \mathbb{K}$ such that $x_n = (x_k^{(n)})_{k\in\mathbb{N}} \to \alpha_n$ as $k \to \infty$. Let $\varepsilon > 0$. As (x_n) is a Cauchy sequence, there exists an $N_0 \in \mathbb{N}$ such that

$$|\alpha_n - \alpha_m| = \lim_{k\to\infty} \big|x_k^{(n)} - x_k^{(m)}\big| \leq \sup_{k\in\mathbb{N}} \big|x_k^{(n)} - x_k^{(m)}\big| = \|x_n - x_m\|_\infty < \varepsilon$$

for all $n, m \geq N_0$. So $(\alpha_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$. As $\mathbb{K}$ is complete, let $\alpha \in \mathbb{K}$ be the limit of (α_n).

Let $\varepsilon > 0$ and $N_0 \in \mathbb{N}$ such that $|\alpha_n - \alpha| < \varepsilon$ and $\|y - x_n\|_\infty < \varepsilon$ for all $n \geq N_0$. Let $N_1 \in \mathbb{N}$ be such that $\big|x_k^{(N_0)} - \alpha_{N_0}\big| < \varepsilon$ for all $k \geq N_1$. Then

$$|y_k - \alpha| \leq \big|y_k - x_k^{(N_0)}\big| + \big|x_k^{(N_0)} - \alpha_{N_0}\big| + |\alpha_{N_0} - \alpha| < 3\varepsilon$$

for all $k \geq N_1$. We have proved that $y = (y_k) \to \alpha$ as $k \to \infty$. Hence $y \in \mathbf{c}$. □

We note a few further properties of the ℓ^p spaces.

Proposition 2.35. *The normed space $(\ell^p, \|\cdot\|_p)$ is separable for $1 \leq p < \infty$, and nonseparable for $p = \infty$.*

Proof. Let $1 \leq p < \infty$. By Proposition 2.32 it suffices to show that $\mathbf{c_{00}} = \operatorname{span}\{e_n : n \in \mathbb{N}\}$ is dense in ℓ^p, where $e_n := (0, \dots, 1, 0, \dots)$ with a single 1 in the nth component. So let $x \in \ell^p$ and define $y_n = (x_1, \dots, x_n, 0, 0, \dots) \in \mathbf{c_{00}}$ for all $n \in \mathbb{N}$. Then, as the partial sums of x are a Cauchy sequence, it follows that

The remainders of a convergent series have to go to 0 since the partial sums form a Cauchy sequence.

$$\|y_n - x\|_p^p = \sum_{k=n+1}^{\infty} |x_k|^p = \lim_{N \to \infty} \sum_{k=n+1}^{N} |x_k|^p < \varepsilon$$

for $n, N \geq N_0$. This shows that $\mathbf{c_{00}}$ is dense in ℓ^p.

Now suppose $p = \infty$. Note that $\{0,1\}^{\mathbb{N}} \subset \ell^\infty$. Moreover, if $x, y \in \{0,1\}^{\mathbb{N}}$ and $x \neq y$, then $\|x - y\|_\infty = 1$. Suppose $A \subset \ell^\infty$ is dense. Then $A \cap B(x, \frac{1}{2}) \neq \emptyset$ for all $x \in \{0,1\}^{\mathbb{N}}$ as otherwise $x \notin \operatorname{cl} A$. So for every $x \in \{0,1\}^{\mathbb{N}}$ there exists an $a_x \in A \cap B(x, \frac{1}{2})$. Suppose $x, y \in \{0,1\}^{\mathbb{N}}$ such that $x \neq y$. Then

$$\begin{aligned} \|a_x - a_y\|_\infty &= \|a_x - x + x - y + y - a_y\|_\infty \\ &\geq \|x - y\|_\infty - \|x - a_x\|_\infty - \|y - a_y\|_\infty \\ &> 1 - \frac{1}{2} - \frac{1}{2} = 0. \end{aligned}$$

So $a_x \neq a_y$. This shows that A has at least the cardinality of $\{0,1\}^{\mathbb{N}}$, which is uncountable. Thus ℓ^∞ is nonseparable. $\square$

We continue with a very important topological concept.

Definition 2.36. Let $(X, \|\cdot\|)$ be a normed space. A subset $A \subset X$ is called **compact** if every sequence in A has a convergent subsequence with limit in A, i.e. for all sequences $(x_n)_{n\in\mathbb{N}}$ in A there exists a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} x_{n_k} \in A$.

Exercise 2.37. Show that a compact subset is always closed and bounded. Moreover, show that in $\mathbb{R}$ the converse holds due to Bolzano–Weierstraß.

Sets that can be covered by finitely many ε-balls are called ***totally bounded****. In finite dimensions this can be seen to be the same as boundedness. In general, however, it is a much stronger property than boundedness.*

Exercise 2.38. Let $(X, \|\cdot\|)$ be a normed space and $K \subset X$ compact. Then for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ and $x_1, \dots, x_N \in K$ such that $K \subset \bigcup_{k=1}^{N} B(x_k, \varepsilon)$. Deduce that there exists a countable dense subset of K.

In the following result compactness is essential.

Exercise 2.39. Let $(X, \|\cdot\|)$ be a normed space, $A \subset X$ compact and $f\colon A \to \mathbb{R}$ continuous. Then f attains its minimum and maximum on A, i.e. there exists $a_{\min}, a_{\max} \in A$ such that

$$f(a_{\min}) \leq f(x) \leq f(a_{\max})$$

for all $x \in A$.

Proposition 2.40. *Let $A \subset \mathbb{K}^N$. Then A is compact in $(\mathbb{K}^N, \|\cdot\|_\infty)$ if and only if A is closed and bounded.*

Proof. We have already observed the necessity.

It remains to show that $A \subset \mathbb{K}^d$ is compact if it is closed and bounded. For simplicity we assume that $\mathbb{K} = \mathbb{R}$, otherwise we need to apply the following arguments to both the real and imaginary parts. Let (x_n) be a sequence in A. Then the first coordinates $(x_1^{(n)})_{n\in\mathbb{N}}$ form a bounded sequence. By Bolzano–Weierstraß there exists a convergent subsequence $(x_1^{(n_{1,k})})_{k\in\mathbb{N}}$. Then the subsequence of second coordinates $(x_2^{(n_{1,k})})_{k\in\mathbb{N}}$ is bounded. So there exists a convergent subsequence $(x_2^{(n_{2,k})})_{k\in\mathbb{N}}$ of this subsequence. Note that the subsequence $(x_1^{(n_{2,k})})_{k\in\mathbb{N}}$ still converges. Proceeding inductively, after N steps of taking subsequences of subsequences, we find a subsequence $(x_{n_{N,k}})_{k\in\mathbb{N}}$ that converges in all components. It is readily observed that componentwise convergence implies convergence in $(\mathbb{K}^N, \|\cdot\|_\infty)$.

As A is closed in $(\mathbb{K}^N, \|\cdot\|_\infty)$, we obtain that the limit is contained in A. This shows that A is compact. □

Exercise 2.41. Consider the set

$$F := \{x \in \ell^p : \|x\|_p \le 1\}.$$

Show that F is bounded and closed, but not compact.

As an application of compactness, we will obtain the following result that shows that in $\mathbb{K}^N$ all norms essentially behave the same.

Proposition 2.42. *Let $\|\cdot\|$ be a norm on $\mathbb{K}^N$. Then there exist constants $m, M > 0$ such that*

$$m\|x\|_\infty \le \|x\| \le M\|x\|_\infty$$

for all $x \in \mathbb{K}^N$.

Proof. Observe that

$$\|x\| \le \sum_{k=1}^{N} |x_k|\|e_k\| \le \sum_{k=1}^{N} \|e_k\|\|x\|_\infty = M\|x\|_\infty,$$

with $M := \sum_{k=1}^{N} \|e_k\| > 0$. So the inequality on the right is established.

Define $K := \{x \in \mathbb{K}^N : \|x\|_\infty = 1\}$. Then K is closed and bounded, and therefore compact in $(\mathbb{K}^N, \|\cdot\|_\infty)$ by Proposition 2.40. Next define $f\colon K \to [0,\infty)$ by $f(x) := \|x\|$. It follows that

$$|f(x) - f(y)| = |\|x\| - \|y\|| \le \|x - y\| \le M\|x - y\|_\infty$$

for all $x, y \in K$. This implies that f is continuous as a map from $(\mathbb{K}^N, \|\cdot\|_\infty)$ to $\mathbb{R}$. Hence f attains its minimum on K by Exercise 2.39. Suppose the minimum is attained in $x^* \in K$. As $\|x^*\|_\infty = 1$, it follows that $m := \|x^*\| > 0$. As the inequality is trivial for $x = 0$, suppose $x \neq 0$. Then

$$\left\| \frac{x}{\|x\|_\infty} \right\| \geq \|x^*\| = m,$$

and hence $m\|x\|_\infty \leq \|x\|$, which establishes the inequality on the left. □

Exercise 2.43. Let X be a vector space and let $\mathcal{N}$ be the set of all norms on X. We define a relation $\sim$ on $\mathcal{N}$ by saying $\|\cdot\|_1 \sim \|\cdot\|_2$ if and only if there exist $m, M > 0$ such that

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1 \tag{2.5}$$

for all $x \in X$. Show that $\sim$ is an equivalence relation.

Definition 2.44. Two norms on the same vector space that dominate each other like in (2.5) are called **equivalent**.

Remark 2.45. Equivalent norms yield the same notion of convergence, i.e. a sequence converges in the first norm if and only if converges in the second norm (and then with the same limit), and a sequence is Cauchy with respect to the first norm if and only if it is Cauchy with respect to the second. Consequently equivalent norms give rise to the same continuous functions, the same closed, bounded or compact sets, and a space is complete or separable either with respect to both norms or with respect to none.

Corollary 2.46. *Any two norms on $\mathbb{K}^N$ are equivalent. More specifically, let $\|\cdot\|$ be a norm on $\mathbb{K}^N$. Then $(\mathbb{K}^N, \|\cdot\|)$ is complete and separable, a sequence converges with respect to $\|\cdot\|$ if and only if it converges componentwise and a set is compact in $(\mathbb{K}^N, \|\cdot\|)$ if and only if it is closed and bounded. In particular, the closed unit ball $\{x \in \mathbb{K}^N : \|x\| \leq 1\}$ is compact.*

The above corollary could have also been formulated with respect to any finite-dimensional vector space X instead of $\mathbb{K}^N$. However, it is essential that the space is finite dimensional. We already observed that in the infinite-dimensional ℓ^p spaces things are completely different. For example, ℓ^p is separable for $1 \leq p < \infty$, but ℓ^∞ is not. Moreover, in the ℓ^p spaces a set does not need to be compact if it is closed and bounded, see Example 2.41. In fact, with a little more effort one can prove the following characterisation of finite-dimensional normed spaces.

Theorem 2.47. *Let $(X, \|\cdot\|)$ be a normed space. The following are equivalent:*

The implication (iii)⇒(i) is the one that requires the work.

(i) $\dim X < \infty$.

(ii) *Every bounded closed set in X is compact.*

(iii) *The closed unit ball in X is compact.*

2.5 Bounded linear operators

In this section, we study the continuity of *linear* maps between normed (vector) spaces. The following proposition gives the main characterisation of such maps. First we shall give the following definition.

Definition 2.48. Let E, F be vector spaces over $\mathbb{K}$. A map $T\colon E \to F$ is called **linear** if $T(x + \lambda y) = T(x) + \lambda T(y)$ for all $x, y \in E$ and $\lambda \in \mathbb{K}$.

Exercise 2.49. Show that a matrix in $\mathbb{K}^{M\times N}$, where M is the number of rows and N is the number of columns, can be interpreted as a linear map from $\mathbb{K}^N$ to $\mathbb{K}^M$ using matrix multiplication.

Proposition 2.50. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T\colon X \to Y$ be a linear map. The following are equivalent:*

(i) *T is Lipschitz continuous.*

(ii) *T is continuous at* 0.

(iii) *There exists a constant $C > 0$ such that $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$.*

(iv) $\sup\{\|Tx\|_Y : \|x\|_X \le 1\} < \infty$.

Proof. (i)⇒(ii) Trivial.

(ii)⇒(iii) Let $\varepsilon = 1$. Then by Exercise 2.9 there exists a $\delta > 0$ such that for all $x \in X$ such that $\|x\|_X \le \delta$ one has $\|Tx\|_Y < 1$. Let $C > 0$ be such that $\frac{1}{C} < \delta$. Then

$$\left\| T\frac{x}{C\|x\|_X} \right\|_Y < 1$$

for all $x \in X \setminus \{0\}$. Hence $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$.

(iii)⇒(iv) Using $C > 0$ as in (iii), it follows that $\|Tx\|_Y \le C$ for all $x \in X$ such that $\|x\|_X \le 1$. Hence the supremum is finite.

(iv)⇒(i) Let $x, y \in X$. Then

$$\begin{aligned} \|Tx - Ty\|_Y &= \|x - y\|_X \left\| T\frac{x - y}{\|x - y\|_X} \right\|_Y \\ &\le \|x - y\|_X \sup\{\|Tz\|_Y : \|z\|_X \le 1\}. \end{aligned}$$

As the supremum is finite, the map T is Lipschitz continuous. □

Definition 2.51. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. A **bounded operator** from X to Y is a continuous linear map $T\colon X \to Y$. We write $\mathscr{L}(X, Y)$ for the set of all bounded operators from X to Y and set $\|T\|_{\mathscr{L}(X,Y)} := \sup\{\|Tx\|_Y : \|x\|_X \le 1\}$. If $T \in \mathscr{L}(X, Y)$ is bijective such that $T^{-1} \in \mathscr{L}(Y, X)$, then T is called an **isomorphism** and one says that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are **isomorphic**. An operator $T \in \mathscr{L}(X, Y)$ is called **isometric** if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$.

The word 'isomorphic' is of Greek origin and means 'of same form or shape'. In fact, isomorphic spaces not only behave the same topologically, but also with respect to completeness, for example.

If it is clear which norms we use on X and Y, we also write $\|T\|$ instead of $\|T\|_{\mathscr{L}(X,Y)}$. If $(X, \|\cdot\|_X) = (Y, \|\cdot\|_Y)$, we write $\mathscr{L}(X)$ instead of $\mathscr{L}(X, X)$. If $Y = \mathbb{K}$, we write X' instead of $\mathscr{L}(X, \mathbb{K})$. We call X' the **dual space** of X.

Exercise 2.52. For all $a \in \mathbb{R}$ let $f_a \colon \mathbb{R} \to \mathbb{R}$ be given by $f_a(x) = ax$. Show that the dual of $(\mathbb{R}, |\cdot|)$ is given by $\{f_a : a \in \mathbb{R}\}$. Moreover, show that $\|f_a\|_{\mathbb{R}'} = |a|$. Deduce that there exists an isomorphism $T \colon (\mathbb{R}, |\cdot|) \to (\mathbb{R}, |\cdot|)'$ such that $\|Ta\|_{\mathbb{R}'} = |a|$ for all $a \in \mathbb{R}$. In other words, $(\mathbb{R}, |\cdot|)$ is **isometric isomorphic** to its dual.

In fact, we shall see later that it follows from the Riesz–Fréchet representation theorem that every Hilbert space over $\mathbb{R}$ is in a natural way isometric isomorphic to its dual, while every Hilbert space over $\mathbb{C}$ is isometric isomorphic to its ***bidual****, i.e., the dual of its dual.*

We shall prove that $\mathscr{L}(X, Y)$ is a vector space and that $\|\cdot\|_{\mathscr{L}(X,Y)}$ defines a norm on that space. We first prove the following alternative description of $\|\cdot\|_{\mathscr{L}(X,Y)}$.

Lemma 2.53. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T \in \mathscr{L}(X, Y)$. Then $\|Tx\|_Y \leq \|T\|_{\mathscr{L}(X,Y)} \|x\|_X$ for all $x \in X$. Moreover,*

$$\|T\|_{\mathscr{L}(X,Y)} = \inf\{C > 0 : \|Tx\|_Y \leq C\|x\|_X \text{ for all } x \in X\}.$$

Proof. We write $\|T\|$ instead of $\|T\|_{\mathscr{L}(X,Y)}$. Define

$$A := \{C > 0 : \|Tx\|_Y \leq C\|x\|_X \text{ for all } x \in X\}.$$

By rescaling and using homogeneity, it follows that $\|Tx\|_Y \leq \|T\| \|x\|_X$ for all $x \in X$. Therefore $\|T\| \in A$ and $\inf A \leq \|T\|$.

Conversely, we give a proof by contradiction. Suppose that $\inf A < \|T\|$. So there exists a $C \in A$ such that $C < \|T\|$. Since $\|Tx\|_Y \leq C$ for all $x \in X$ such that $\|x\|_X \leq 1$, it follows that $\|T\| \leq C < \|T\|$. This is a contradiction. Hence $\inf A = \|T\|$, which was to be shown. □

Theorem 2.54. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Then $\mathscr{L}(X, Y)$ is a vector space (with respect to pointwise addition and scalar multiplication) and $\|\cdot\|_{\mathscr{L}(X,Y)}$ defines a norm on $\mathscr{L}(X, Y)$. If Y is complete, then also the space $(\mathscr{L}(X, Y), \|\cdot\|_{\mathscr{L}(X,Y)})$ is complete. In particular, the dual space X' is always complete.*

Proof. It is easy to see that the set of all (not necessarily linear or continuous) maps from X to Y form a vector space Y^X with respect to pointwise addition and scalar multiplication. By Proposition 2.12.3 the sum and scalar multiple of continuous maps are again continuous. Moreover, it is obvious that linear combinations of linear maps from X to Y remain linear. This shows that $\mathscr{L}(X, Y)$ is a vector subspace of the vector space of Y^X.

We next verify that $\|\cdot\|_{\mathscr{L}(X,Y)}$ is a norm on the space $\mathscr{L}(X, Y)$. Let $T, S \in \mathscr{L}(X, Y)$ and $\alpha \in \mathbb{K}$. If $\|T\|_{\mathscr{L}(X,Y)} = 0$ then $Tx = 0$ for all $x \in X$.

So (N1) is satisfied. Moreover, one has

$$\|\alpha T\|_{\mathscr{L}(X,Y)} = \sup_{\|x\|_X \le 1} \|\alpha Tx\|_Y = |\alpha| \sup_{\|x\|_X \le 1} \|Tx\|_Y = |\alpha| \|T\|_{\mathscr{L}(X,Y)}.$$

So (N2) is satisfied. The triangle inequality (N3) follows from

$$\begin{aligned}
\|T+S\|_{\mathscr{L}(X,Y)} &= \sup_{\|x\|_X \le 1} \|(T+S)x\|_Y \\
&\le \sup_{\|x\|_X \le 1} (\|Tx\|_Y + \|Sx\|_Y) \\
&\le \sup_{\|x\|_X \le 1} \|Tx\|_Y + \sup_{\|x\|_X \le 1} \|Sx\|_Y = \|T\|_{\mathscr{L}(X,Y)} + \|S\|_{\mathscr{L}(X,Y)}.
\end{aligned}$$

This shows that $(\mathscr{L}(X,Y), \|\cdot\|_{\mathscr{L}(X,Y)})$ is a normed space.

It remains to prove that $(\mathscr{L}(X,Y), \|\cdot\|_{\mathscr{L}(X,Y)})$ is complete if $(Y, \|\cdot\|_Y)$ is complete. Let $(T_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{L}(X,Y)$. Then, for all $x \in X$, we have

$$\|T_n x - T_m x\|_Y \le \|T_n - T_m\|_{\mathscr{L}(X,Y)} \|x\|_X,$$

proving that $(T_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. So $(T_n x)$ converges in Y. Define $T\colon X \to Y$ by setting $Tx = \lim_{n\to\infty} T_n x$ for all $x \in X$. We first show that T is linear. Let $x, y \in X$ and $\alpha \in \mathbb{K}$. Then

$$T(\alpha x + y) = \lim_{n\to\infty} T_n(\alpha x + y) = \lim_{n\to\infty} \alpha T_n x + T_n y = \alpha Tx + Ty.$$

Here we have used the linearity of the T_n in the second step and Proposition 2.12.3 for the third equality.

We next prove that $T \in \mathscr{L}(X,Y)$ and $\|T_n - T\|_{\mathscr{L}(X,Y)} \to 0$. Given $\varepsilon > 0$, pick n_0 such that $\|T_n - T_m\|_{\mathscr{L}(X,Y)} \le \varepsilon$ for all $n, m \ge n_0$. Now let $x \in X$ with $\|x\|_X \le 1$. Then $\|T_n x - T_m x\|_Y \le \varepsilon$ for all $n, m \ge n_0$. Letting $m \to \infty$, it follows that $\|T_n x - Tx\|_Y \le \varepsilon$ for all $n \ge n_0$. On the one hand, this proves that $T \in \mathscr{L}(X,Y)$ since

$$\|Tx\|_Y \le \|T_{n_0} x\|_Y + \|Tx - T_{n_0} x\|_Y \le \|T_{n_0}\|_{\mathscr{L}(X,Y)} + \varepsilon$$

for all $\|x\|_X \le 1$. On the other hand, by taking the supremum over $x \in X$ with $\|x\|_X \le 1$, it follows that $\|T_n - T\|_{\mathscr{L}(X,Y)} \le \varepsilon$ for all $n \ge n_0$. So $T_n \to T$ in $(\mathscr{L}(X,Y), \|\cdot\|_{\mathscr{L}(X,Y)})$. □

Exercise 2.55. Let $(X_k, \|\cdot\|_k)$ be normed spaces for $k = 1, 2, 3$. Show that if $T \in \mathscr{L}(X_1, X_2)$ and $S \in \mathscr{L}(X_2, X_3)$, then $ST \in \mathscr{L}(X_1, X_3)$ and

$$\|ST\|_{\mathscr{L}(X_1,X_3)} \le \|S\|_{\mathscr{L}(X_2,X_3)} \|T\|_{\mathscr{L}(X_1,X_2)}.$$

Remark 2.56. In finite dimensions, bounded linear operators correspond to matrices. Moreover, note that the operator norm $\|\cdot\|_{\mathscr{L}(X,Y)}$ depends on the specific norms chosen on X and Y. So by choosing different norms on $\mathbb{K}^N$ one obtains different operator norms for matrices. Note that the same matrix can be considered as a bounded linear operator for different norms on $\mathbb{K}^N$. However, since the space of matrices of the form $\mathbb{K}^{M\times N}$ for $N, M \in \mathbb{N}$ is finite dimensional, different norms on this space turn out equivalent thanks to Corollary 2.46.

*The operator T_m in this example is commonly called a **multiplication operator**. Such operators correspond to diagonal matrices in the finite-dimensional case.*

Example 2.57. Consider the normed space $(\ell^p, \|\cdot\|_p)$ for $p \in [1, \infty]$. For $m \in \ell^\infty$, we define $T_m : \ell^p \to \ell^p$ by

$$T_m x = (m_1 x_1, m_2 x_2, m_3 x_3, \ldots).$$

Then $T_m \in \mathscr{L}(\ell^p)$ and $\|T_m\|_{\mathscr{L}(\ell^p)} = \|m\|_{\ell^\infty}$.

Proof. We give the proof for $1 \leq p < \infty$ and leave the case $p = \infty$ as an exercise.

For $x \in \ell^p$, we have

$$\|T_m x\|_p^p = \sum_{k=1}^{\infty} |m_k x_k|^p = \sum_{k=1}^{\infty} |m_k|^p |x_k|^p \leq \sum_{k=1}^{\infty} \|m\|_\infty^p |x_k|^p = \|m\|_\infty^p \|x\|_p^p.$$

This proves that $T_m \in \mathscr{L}(\ell^p)$ and that $\|T_m\|_{\mathscr{L}(\ell^p)} \leq \|m\|_\infty$. To see that equality holds, consider $e_k \in \ell^p$, where $e_k = (0, \ldots, 0, 1, 0, \ldots)$ with a 1 at the kth position. Then $\|e_k\|_p = 1$ and $\|T_m e_k\|_p = |m_k|$. Thus $\|T_m\|_{\mathscr{L}(\ell^p)} \geq |m_k|$ for all $k \in \mathbb{N}$. Hence $\|T_m\|_{\mathscr{L}(\ell^p)} \geq \|m\|_\infty$. □

Exercise 2.58. Let $1 \leq p, q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Given $y \in \ell^q$, define

$$\varphi_y(x) = \sum_{k=1}^{\infty} y_k x_k.$$

Note that this is well-defined by Hölder's inequality. Show that φ_y is a continuous linear map from ℓ^p to $\mathbb{K}$, hence $\varphi_y \in (\ell^p)'$. Moreover, show that $\|\varphi_y\|_{(\ell^p)'} = \|y\|_q$.

Remark 2.59. The previous exercise suggests that there might be a connection between $(\ell^p)'$ and ℓ^q for conjugate indices p and q. In fact, it is not hard to show that for $p \in [1, \infty)$ the map $y \mapsto \varphi_y$ is an isomorphism between ℓ^q and the dual of ℓ^p. However, it follows from the set theoretic *axiom of choice* that the dual of ℓ^∞ is 'strictly larger' than ℓ^1.

The axiom of choice also ensures the existence of a basis for every vector space.

We end this section with the following useful result.

Proposition 2.60. *Let $(X, \|\cdot\|_X)$ be a normed space, $(Y, \|\cdot\|_Y)$ be a complete normed space and X_0 be a dense vector subspace of X. Given $T \in \mathscr{L}(X_0, Y)$ there exists a unique operator $\tilde{T} \in \mathscr{L}(X, Y)$ with $Tx = \tilde{T}x$ for all $x \in X_0$. Moreover, $\|T\|_{\mathscr{L}(X_0,Y)} = \|\tilde{T}\|_{\mathscr{L}(X,Y)}$.*

Proof. Let $x \in X$. By density, there exists a sequence (x_n) in X_0 such that $x_n \to x$. Since $\|Tx_n - Tx_m\|_Y \leq \|T\|_{\mathscr{L}(X_0,Y)} \|x_n - x_m\| \to 0$ as $n, m \to \infty$, it follows that (Tx_n) is a Cauchy sequence in Y. By completeness, Tx_n converges to some $\tilde{T}x \in Y$. Note that $\tilde{T}x$ does not depend on the approximating sequence (x_n). Indeed, if (y_n) was another sequence in X_0 converging to x and $Ty_n \to z$, then

$$\|z - \tilde{T}x\|_Y \leq \|z - Ty_n\|_Y + \|T\|_{\mathscr{L}(X_0,Y)} \|y_n - x_n\|_X + \|Tx_n - \tilde{T}x\|_Y.$$

The right hand side tends to 0 as $n \to \infty$ since $\|y_n - x_n\| \to 0$. Hence $z = \tilde{T}x$. Now it is easy to see that $\tilde{T} \colon X \to Y$ is linear, cf. the proof of Theorem 2.54. Moreover, $\tilde{T}x = Tx$ for all $x \in X_0$. To see that $\tilde{T} \in \mathscr{L}(X, Y)$, let $x \in X$ and (x_n) be a sequence in X_0 converging to x. Then

$$\begin{aligned} \|\tilde{T}x\|_Y &= \|\lim_{n\to\infty} Tx_n\|_Y = \lim_{n\to\infty} \|Tx_n\|_Y \\ &\leq \lim_{n\to\infty} \|T\|_{\mathscr{L}(X_0,Y)} \|x_n\|_X = \|T\|_{\mathscr{L}(X_0,Y)} \|x\|_X, \end{aligned}$$

where we have used the continuity of the norm. This proves that $\tilde{T} \in \mathscr{L}(X, Y)$ and $\|\tilde{T}\|_{\mathscr{L}(X,Y)} \leq \|T\|_{\mathscr{L}(X_0,Y)}$. The other inequality is trivial. □

2.6 Spaces of continuous functions

Function spaces are a particularly important class of normed spaces. As the name suggests, the elements of a function space are functions of a certain class, for example, continuous or suitably integrable functions. This is a most convenient setting if one is looking for solutions of differential equations, for example, since function spaces allow to formulate a problem in way that mimics the elementary one-dimensional case of real analysis and to put the problem in a framework that allows to apply the abstract functional analytic machinery. A striking example will be given in the following Section 2.7.

In this section we focus on spaces of continuous functions. It will be convenient to first introduce the following.

Definition 2.61 (Space of bounded functions)**.** Let Ω be a set. A $\mathbb{K}$-valued function $f \colon \Omega \to \mathbb{K}$ is called bounded if

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

The set of all bounded functions from Ω to $\mathbb{K}$ is denoted by

$$\mathcal{F}_\mathrm{b}(\Omega) := \{f \colon \Omega \to \mathbb{K} : f \text{ is bounded}\}.$$

Proposition 2.62. *Let Ω be a set. The set $\mathcal{F}_\mathrm{b}(\Omega)$ is a vector space with respect to the pointwise addition of functions and the pointwise scalar multiplication.*

Moreover,

$$\|f\|_\infty = \sup_{x\in\Omega} |f(x)|$$

defines a norm on $\mathcal{F}_b(\Omega)$. The space $(\mathcal{F}_b(\Omega), \|\cdot\|_\infty)$ is complete.

*This norm $\|\cdot\|_\infty$ is usually called the **sup-norm**.*

Proof. We leave the proof as an exercise. The completeness follows analogously as in the proof of Theorem 2.54, where the completeness of $\mathscr{L}(X,Y)$ was shown for a Banach space Y. □

We now introduce the following spaces of continuous functions.

Definition 2.63. Let $(X, \|\cdot\|)$ be a normed space and $M \subset X$. By $\mathrm{C}(M) = \mathrm{C}(M;\mathbb{K})$ one denotes the **vector space of continuous functions** from M to $\mathbb{K}$ with usual pointwise operations. The subspace of **bounded continuous functions** in $\mathrm{C}(M)$ is denoted by $\mathrm{C_b}(M) = \mathrm{C_b}(M;\mathbb{K})$. So $\mathrm{C_b}(M) = \mathrm{C}(M) \cap \mathcal{F}_b(M)$.

More generally, one could consider continuous functions from M into a normed space $(Y, \|\cdot\|_Y)$. The corresponding function spaces would be denoted by $\mathrm{C}(M;Y)$ and $\mathrm{C_b}(M;Y)$.

We point out that if K is compact in $(X, \|\cdot\|)$ then $\mathrm{C_b}(K) = \mathrm{C}(K)$ by Exercise 2.39.

Let $(X, \|\cdot\|)$ be a normed space and $M \subset X$. We study further structural properties of the vector spaces $\mathrm{C}(M)$ and $\mathrm{C_b}(M)$. Let $f, g \in \mathrm{C}(M)$. We define the **product** of f and g pointwise by $(fg)(x) := f(x)g(x)$. Clearly $fg \in \mathrm{C}(M)$ and, if $f, g \in \mathrm{C_b}(M)$, then $fg \in \mathrm{C_b}(M)$. Moreover, it is easily observed that polynomials are continuous functions from $\mathbb{K}$ to $\mathbb{K}$ and therefore elements of $\mathrm{C}(\mathbb{K})$.

The following result shows that there exist many continuous functions.

Exercise 2.64 (Urysohn's lemma)**.** Let $(X, \|\cdot\|)$ be a normed space and $A, B \subset X$ be closed sets such that $A \cap B = \emptyset$. Show that there exists a continuous function $f \in \mathrm{C}(X)$ such that $f(x) = 0$ for all $x \in A$ and $f(x) = 1$ for all $x \in B$.

Hint: First show that $x \mapsto \inf\{\|x-y\| : y \in A\}$ is Lipschitz continuous if $A \neq \emptyset$. For f use a suitable combination of these functions.

Proposition 2.65. *Let $(X, \|\cdot\|)$ be a normed space and $M \subset X$. The space $\mathrm{C_b}(M)$ is a closed subspace of $(\mathcal{F}_b(M), \|\cdot\|_\infty)$. Hence, $(\mathrm{C_b}(M), \|\cdot\|_\infty)$ is a Banach space.*

Proof. It suffices to prove that $\mathrm{C_b}(M)$ is closed in $\mathcal{F}_b(M)$. To that end, let a sequence (f_n) in $\mathrm{C_b}(M)$ be given such that (f_n) converges with respect to $\|\cdot\|_\infty$ to some $f \in \mathcal{F}_b(M)$. We have to prove that f is continuous. So let $x_k \to x$ in M. Then

$$\begin{aligned}|f(x) - f(x_k)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(x_k)| + |f_n(x_k) - f(x_k)| \\ &\le 2\|f - f_n\|_\infty + |f_n(x) - f_n(x_k)|.\end{aligned}$$

Given $\varepsilon > 0$, we first pick an $n_0 \in \mathbb{N}$ such that $\|f - f_{n_0}\|_\infty \le \varepsilon/4$. Since f_{n_0} is continuous, we can pick a $k_0 \in \mathbb{N}$ such that $|f_{n_0}(x) - f_{n_0}(x_k)| \le \frac{\varepsilon}{2}$

for all $k \geq k_0$. Now the above estimate yields $|f(x) - f(x_k)| \leq \varepsilon$ for all $k \geq k_0$. It follows that f is continuous and hence an element of $\mathrm{C_b}(M)$. $\square$

Remark 2.66. Let $(X, \|\cdot\|)$ be a normed space and $M \subset X$. A sequence of functions $f_n \colon M \to \mathbb{K}$ is said to **converge uniformly** on M to $f \colon M \to \mathbb{K}$, if for all $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq n_0$ and all $x \in M$.

Note that the same $n_0 \in \mathbb{N}$ is used simultaneously for all $x \in M$, hence the 'uniformly'.

Obviously, $f_n \to f$ with respect to $\|\cdot\|_\infty$ if and only if $f_n \to f$ uniformly on M. Note that by Proposition 2.65 the uniform limit of bounded continuous functions is continuous. This is not true for the pointwise limit of a sequence of bounded continuous functions. For example, let $M = [0,1]$ and $f_n(x) = x^n$ for all $x \in [0,1]$ and $n \in \mathbb{N}$. Then the pointwise limit of (f_n) is not continuous.

We content ourselves with presenting the *Stone–Weierstrass theorem* without proof. It will allow us to deduce that $\mathrm{C}(K)$ is separable if K is compact.

Definition 2.67. Let $(X, \|\cdot\|)$ be a normed space and $M \subset X$. Let $\mathscr{A}$ be a subset of $\mathrm{C}(M)$.

1. A vector subspace $\mathscr{A}$ of $\mathrm{C}(M)$ is called an **algebra** if $\mathscr{A}$ is closed under multiplication, i.e., $f, g \in \mathscr{A}$ implies $fg \in \mathscr{A}$.
2. An algebra $\mathscr{A}$ is called **unital** if the function $\mathbb{1} \colon M \to \mathbb{K}$, defined by $\mathbb{1}(x) = 1$ for all $x \in M$, belongs to $\mathscr{A}$.
3. $\mathscr{A}$ is said to **separate points** (in M), if for all $x, y \in M$ with $x \neq y$, there exists an $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.
4. $\mathscr{A}$ is said to be **closed under conjugation**, if $f \in \mathscr{A}$ implies $\bar{f} \in \mathscr{A}$.

Note that by Urysohn's lemma the unital algebra $\mathscr{A} = \mathrm{C}(M)$ separates points. Moreover, it is closed under conjugation.

Theorem 2.68. *Let $(X, \|\cdot\|)$ be a normed space and $K \subset X$ be compact. Suppose $\mathscr{A} \subset \mathrm{C}(K)$ is an unital algebra that separates points. If $\mathbb{K} = \mathbb{C}$, then suppose in addition that $\mathscr{A}$ is closed under conjugation. Then $\mathscr{A}$ is dense in $(\mathrm{C}(K), \|\cdot\|_\infty)$.*

Sketch of proof. We only consider the case $\mathbb{K} = \mathbb{R}$. One first shows that the closure of $\mathscr{A}$ contains $|f|$ for all $f \in \mathscr{A}$. It follows that the functions $\min\{f, g\} = \frac{1}{2}(f + g - |f - g|)$ and $\max\{f, g\} = \frac{1}{2}(f + g + |f - g|)$ are contained in the closure of $\mathscr{A}$.

In the lecture this proof is skipped.

Now let $f \in C(K)$ and $\varepsilon > 0$ be given. For all points $a, b \in K$ such that $a \neq b$ we find a function $g_{a,b} \in \mathscr{A}$ such that $g_{a,b}(a) = f(a)$ and $g_{a,b}(b) = f(b)$. For example,

$$g_{a,b}(x) := f(a)\mathbb{1} + (f(b) - f(a))\frac{h(x) - h(a)}{h(b) - h(a)},$$

where $h \in \mathscr{A}$ is such that $h(a) \neq h(b)$. Let $b \in K$ be fixed. For all $a \in K$ define the open set

$$U_a := \{x \in K : g_{a,b}(x) < f(x) + \varepsilon\}.$$

Note that on U_a the function $g_{a,b}$ is good in that it is not much larger than f. Moreover, as $a \in U_a$ the family $(U_a)_{a \in K}$ is an open cover of K. Compactness allows to select a finite subcover such that $K \subset U_{a_1} \cup \ldots \cup U_{a_N}$. By defining

In fact, a compact set is compact if and only if every open cover admits a finite subcover. We use this here without proof.

$$g_b := \min\{g_{a_k,b} : k = 1, \ldots, N\}$$

we obtain a g_b in the closure of $\mathscr{A}$ such that $g_b(x) < f(x) + \varepsilon$ for all $x \in K$.

Similarly, for all $b \in K$ one can define the open set

$$V_b := \{x \in K : g_b(x) > f(x) - \varepsilon\}.$$

Note that on V_b the function g_b is good in that it is not much smaller than f. Moreover, as $b \in V_b$ the family $(V_b)_{b \in K}$ is an open cover of K. Compactness allows to select a finite subcover $K \subset V_{b_1} \cup \ldots \cup V_{b_M}$. Now define

$$g := \max\{g_{b_k} : k = 1, \ldots, M\}.$$

It is easily checked that g is in the closure of $\mathscr{A}$ and $\|f - g\|_\infty < \varepsilon$. $\square$

Corollary 2.69. *Let $(X, \|\cdot\|)$ be a normed space and $K \subset X$ be compact. Then $\mathrm{C}(K)$ is separable.*

Proof. By Exercise 2.38, there exists a countable set $\{x_n \in K : n \in \mathbb{N}\}$ that is dense in K. For all $n, m \in \mathbb{N}$, the sets $\operatorname{cl} B(x_n, \frac{1}{2m})$ and $B(x_n, \frac{1}{m})^{\mathrm{c}}$ are closed and disjoint. Hence, by Urysohn's lemma there exist continuous (real-valued) functions $f_{n,m} \colon K \to [0,1]$ with $f_{n,m}(x) = 0$ for all $x \in \operatorname{cl} B(x_n, \frac{1}{2m})$ and $f_{n,m}(x) = 1$ for all $x \in B(x_n, \frac{1}{m})^{\mathrm{c}}$. We define $\mathscr{P}$ as the set of all finite products of functions $f_{n,m}$ including the function $\mathbb{1}$. This is a countable set. We then define

$$\mathscr{A} := \operatorname{span} \mathscr{P} = \Big\{ \sum_{k=1}^{N} \alpha_k g_k : N \in \mathbb{N}, \alpha_k \in \mathbb{K} \text{ and } g_k \in \mathscr{P} \Big\}.$$

Then $\mathscr{A}$ is unital algebra which is closed under conjugation. Moreover, it separates points. Indeed if $x \neq y$, then $\rho := \|x - y\| > 0$. Pick $m \in \mathbb{N}$ such that $\frac{1}{2m} < \rho/4$ and then $n \in \mathbb{N}$ such that $\|x - x_n\| \leq \frac{1}{2m}$. Then $f_{n,m}(x) = 0$ and, since

$$\|y - x_n\| \geq \|y - x\| - \|x - x_n\| > \rho - \frac{1}{2m} > \frac{3}{4}\rho > \frac{1}{m},$$

one has $f_{n,m}(y) = 1$.

By Theorem 2.68, $\mathscr{A}$ is dense in $C(K)$. Since $\mathscr{A}$ is the linear span of the countable set $\mathscr{P}$, it follows that $C(K)$ is separable by Proposition 2.32. $\square$

The Stone–Weierstrass theorem also allows us to deduce approximation results:

Corollary 2.70. *For every $f \in C[0,1]$, there exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ which converges uniformly on $[0,1]$ to f.*

Proof. It is an easy exercise to show that the assumptions of Theorem 2.68 are satisfied. $\square$ *Important!*

Consider $[0,2\pi]$. A **trigonometric polynomial** is a function $p\colon [0,2\pi] \to \mathbb{C}$ of the form $p(t) = \sum_{k=-N}^{N} \alpha_k e^{ikt}$ where $N \in \mathbb{N}$ and $\alpha_{-N}, \ldots, \alpha_N \in \mathbb{C}$. Using that $e^{ikt}e^{ilt} = e^{i(k+l)t}$, it is easy to see that the trigonometric polynomials form a unital algebra. Since $\overline{e^{ikt}} = e^{-ikt}$, the trigonometric polynomials are closed under conjugation. However, the trigonometric polynomials do not separate points in $[0,\pi]$, since $p(0) = p(2\pi)$ for all trigonometric polynomials p. Nevertheless, the trigonometric polynomials separate points in $(0,2\pi)$. Indeed, if $e^{it} = e^{is}$, then it follows that $t - s = 2\pi n$ for some $n \in \mathbb{Z}$. Hence, if $t,s \in (0,2\pi)$ and $t \neq s$, then $e^{it} \neq e^{is}$.

Corollary 2.71. *For every $f \in C[0,2\pi]$ with $f(0) = f(2\pi)$ there exists a sequence of trigonometric polynomials that converges uniformly to f.*

Proof. Note that the set

$$K := \left\{z \in \mathbb{C} : z = e^{it} \text{ with } t \in [0,2\pi]\right\} = \{z \in \mathbb{C} : |z| = 1\}$$

is compact in $\mathbb{C}$. For all $k \in \mathbb{Z}$ let $\tilde{p}_k\colon K \to \mathbb{C}$ be given by $\tilde{p}_k(z) = z^k$. Then it follows from a straightforward application of Theorem 2.68 that $\tilde{\mathscr{A}} := \operatorname{span}\{\tilde{p}_k : k \in \mathbb{Z}\}$ is dense in $C(K)$. Here the closedness under conjugation of $\tilde{\mathscr{A}}$ follows from the identity $\overline{z^k} = z^{-k}$ for $z \in K$. Now let $\varepsilon > 0$ and define $\tilde{f} \in C(K)$ by setting $\tilde{f}(z) = f(t)$ where $z = e^{it}$. Note that this is well-defined. Let $\tilde{p} \in \tilde{\mathscr{A}}$ such that $\|\tilde{f} - \tilde{p}\|_\infty < \varepsilon$. Then $p(t) := \tilde{p}(e^{it})$ defines a trigonometric polynomial such that $\|f - p\|_\infty < \varepsilon$. $\square$

2.7 Banach's fixed point theorem

In this section, we present a result which is of great importance in applications. Many problems can be reformulated as so-called *fixed point problems*.

Definition 2.72. Let M be a set and $\varphi\colon M \to M$ be a map. A **fixed point** of φ is an element $x^* \in M$ with $\varphi(x^*) = x^*$. We define the **iterates** of φ inductively by $\varphi^1 = \varphi$ and $\varphi^{n+1} = \varphi \circ \varphi^n$ for $n \geq 1$.

Theorem 2.73 (Banach's fixed point theorem)**.** *Let $(X, \|\cdot\|)$ be a Banach space and $M \subset X$ a closed subset. Let $f\colon M \to M$ be a map such that there exists a sequence $q_n \geq 0$ with $\sum_{n=1}^{\infty} q_n < \infty$ with*

$$\|\varphi^n(x) - \varphi^n(y)\| \leq q_n \|x - y\|$$

for all $x, y \in M$ and $n \in \mathbb{N}$. Then φ has a unique fixed point $x^* \in M$.

Proof. We first prove existence of a fixed point. To that end, let $x_0 \in M$ be arbitrary and define a sequence $(x_n)_{n\in\mathbb{N}}$ inductively by setting $x_n := \varphi(x_{n-1})$ for all $n \in \mathbb{N}$. Suppose $n, m \in \mathbb{N}$ such that $n \geq m$. Then

$$\begin{aligned}\|x_n - x_m\| &= \left\| \sum_{k=m}^{n-1} (x_{k+1} - x_k) \right\| \\ &\leq \sum_{k=m}^{n-1} \|\varphi^k(\varphi(x_0)) - \varphi^k(x_0)\| \\ &\leq \Big(\sum_{k=m}^{\infty} q_k \Big) \|\varphi(x_0) - x_0\|,\end{aligned}$$

which goes to zero as $m \to \infty$ since $\sum_{k=1}^{\infty} q_k < \infty$. This shows that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X and hence convergent. Because M is closed, the limit x^* of $(x_n)_{n\in\mathbb{N}}$ lies in M.

We now prove that x^* is a fixed point of φ. To that end, first observe that φ is Lipschitz continuous on M with Lipschitz constant q_1. Hence

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

It remains to prove uniqueness of the fixed point. To that end, assume that x^* and y^* are fixed points of φ, i.e. $\varphi(x^*) = x^*$ and $\varphi(y^*) = y^*$. Then, by induction, $\varphi^n(x^*) = x^*$ and $\varphi^n(y^*) = y^*$ for all $n \in \mathbb{N}$. Therefore

$$\|x^* - y^*\| = \|\varphi^n(x^*) - \varphi^n(y^*)\| \leq q_n \|x^* - y^*\| \to 0$$

as $n \to \infty$, since $\sum_{n=1}^{\infty} q_n < \infty$. Thus $\|x^* - y^*\| = 0$ and hence $x^* = y^*$. ☐

Corollary 2.74 (Classical Banach fixed point theorem)**.** *Let $(X, \|\cdot\|)$ be a Banach space and $\varphi\colon X \to X$ be a* ***strict contraction****, i.e., Lipschitz continuous with Lipschitz constant $L \in [0, 1)$. Then φ has a unique fixed point.*

Proof. The assumptions of Theorem 2.73 are satisfied with $q_n = L^n$ which is summable since $L \in [0, 1)$. ☐

Exercise 2.75. Consider the closed set $M := [1, \infty)$ in the Banach space $(\mathbb{R}, |\cdot|)$. Show that $\varphi\colon M \to M$ given by $\varphi(x) = x + \frac{1}{x}$ satisfies

$$|\varphi(x) - \varphi(y)| < |x - y|.$$

However, φ does not have a fixed point. Why does this not contradict Theorem 2.73?

The usefulness of Banach's fixed point theorem arises from the fact that many problems can be reformulated as fixed point problems. As an illustration, we now solve very general ordinary differential equations using the Banach fixed point theorem. We note that a similar approach also works for stochastic differential equations.

Definition 2.76. Let $I = [0, T]$ be a compact interval, $f\colon I \times \mathbb{R} \to \mathbb{R}$ be a continuous function and $x_0 \in \mathbb{R}$. A **solution** to the ordinary differential equation

$$\text{(ODE)} \qquad \begin{cases} u'(t) = f(t, u(t)) & \text{for } t \in [0, T], \\ u(0) = x_0, \end{cases}$$

is a continuously differentiable function $u^*\colon [0, T] \to \mathbb{R}$ such that (ODE) holds for $u = u^*$.

Note that the problem (ODE) is of course dependent on f and x_0.

We now reformulate (ODE) as a fixed point problem.

Lemma 2.77. *Given a compact interval I, $x_0 \in \mathbb{R}$ and $f\colon I \times \mathbb{R} \to \mathbb{R}$ continuous, define $\Phi\colon \mathrm{C}(I) \to \mathrm{C}(I)$ by*

$$(\Phi u)(t) = x_0 + \int_0^t f(s, u(s))\, ds.$$

Then u^* solves (ODE) if and only if $\Phi u^* = u^*$.

*The map Φ is commonly called the **Picard operator**. Note that Φ is not linear in general.*

Proof. If u^* solves (ODE), then, by the fundamental theorem of calculus, we have

$$u^*(t) - x_0 = u^*(t) - u^*(0) = \int_0^t (u^*)'(s)\, ds = \int_0^t f(s, u^*(s))\, ds$$

for all $t \in I$. Thus $\Phi u^* = u^*$.

Conversely, if $\Phi u^* = u^*$, then, by the fundamental theorem of calculus, u^* is continuously differentiable and

$$(u^*)'(t) = \frac{d}{dt} \int_0^t f(s, u^*(s))\, ds = f(t, u^*(t))$$

for all $t \in I$. Since $u^*(0) = x_0$, it follows that u^* solves (ODE). □

Theorem 2.78 (Picard–Lindelöf)**.** *Let $I = [0, T]$ be a compact interval and $f\colon I \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that there exists an $L \geq 0$ with*

$$|f(t, x) - f(t, y)| \leq L|x - y|$$

for all $x, y \in \mathbb{R}$ and $t \in I$. Then, for every $x_0 \in \mathbb{R}$ there exists a unique solution $u^*\colon [0, T] \to \mathbb{R}$ of the differential equation (ODE).

Proof. Let Φ be as in Lemma 2.77. By this Lemma it suffices to show that Φ has a unique fixed point. Since $(\mathrm{C}(I), \|\cdot\|_\infty)$ is complete, it suffices to show that the hypothesis of Banach's fixed point theorem (Theorem 2.73) is satisfied.

We claim that

$$|(\Phi^n u)(t) - (\Phi^n v)(t)| \leq \frac{L^n t^n}{n!} \|u - v\|_\infty$$

for all $u, v \in \mathrm{C}(I)$ and $t \in I$. Note that $\sum_{n=1}^{\infty} \frac{L^n T^n}{n!}$ is summable and hence the hypothesis of Banach's fixed point theorem is satisfied. Thus, proving the claim finishes the proof.

We proceed by induction. For $n = 1$, we have

$$\begin{aligned}
|(\Phi u)(t) - (\Phi v)(t)| &\leq \int_0^t |f(s, u(s)) - f(s, v(s))| \, ds \\
&\leq \int_0^t L|u(s) - v(s)| \, ds \\
&\leq Lt\|u - v\|_\infty.
\end{aligned}$$

Now assume that $|(\Phi^n u)(t) - (\Phi^n v)(t)| \leq \frac{L^n t^n}{n!} \|u - v\|_\infty$. Then

$$\begin{aligned}
|(\Phi^{n+1} u)(t) - (\Phi^{n+1} v)(t)| &\leq \int_0^t |f(s, (\Phi^n u)(s)) - f(s, (\Phi^n v)(s))| \, ds \\
&\leq \int_0^t L|(\Phi^n u)(s) - (\Phi^n v)(s)| \, ds \\
&\leq \int_0^t L \frac{L^n s^n}{n!} \|u - v\|_\infty \, ds = \frac{L^{n+1} t^{n+1}}{(n+1)!} \|u - v\|_\infty.
\end{aligned}$$

This finishes the proof. □

Exercise 2.79. The proof of Banach's fixed point theorem shows that for any initial function $u \in \mathrm{C}(I)$ the sequence $\Phi^n u$ converges to the unique fixed-point u^* of Φ that is the unique solution of (ODE). This allows to approximate the solution u^* by simply iterating Φ. This method is called **Picard iteration**.

For the ordinary differential equation $u'(t) = tu(t)$ on $[0, 1]$, use Picard iteration to construct the solution of the equation for the initial value $u(0) = 1$.