Introduction to functional analysis

Lecture notes for 18.102, Spring 2025

Christoph Kehle

May 16, 2025

Lectures: Monday and Wednesday from 9.30–11.00 am in room 4-237.
Prerequisites: Linear algebra (18.06, 18.700, or 18.701) and Introduction to analysis (18.100A, 18.100B, 18.100P, 18.100Q) or permission by the lecturer
Problem sets: There will be 7 problem sets in total. Only the best 6 problem sets count towards your final grade. No late homework will be accepted. Exceptions are cases approved by S3.
Exams: There will be two midterms on **3 March** and on **9 April**. The final exam will take place on **21 May**.
Grading: The final grade is the weighted average of the problem sets (30%, i.e. 5% for each problem set), the midterms (30%, i.e. 15% each), and the final (40%).
Office hours: I will hold my office hours on Tuesdays from 1.30–2.30pm in 2-277.
References: I will provide lecture notes. These lecture notes do not follow a particular book. However, several references listed below will be useful.

References

[Ada75] R. A. Adams. *Sobolev spaces.* Vol. Vol. 65. Pure and Applied Mathematics. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975, pp. xviii+268.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011, pp. xiv+599.

[Kre89] E. Kreyszig. *Introductory functional analysis with applications.* Wiley Classics Library. John Wiley & Sons, Inc., New York, 1989, pp. xvi+688.

[Mel20] R. Melrose. "Functional analysis lecture notes for 18.102" (2020). Click here.

[Roy88] H. L. Royden. *Real analysis.* Third. Macmillan Publishing Company, New York, 1988, pp. xx+444.

[Rud91] W. Rudin. *Functional analysis.* Second. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991, pp. xviii+424.

Contents

1 Basic notions, topological and metric spaces **3**
1.1 Zorn's lemma . . . 3
1.2 Vector spaces . . . 4
1.3 Topological spaces . . . 5
1.4 Metric spaces . . . 9
1.5 Baire category theorem . . . 15

2 Normed and Banach spaces **17**
2.1 Linear operators . . . 18
2.2 Hahn–Banach theorem . . . 21
2.3 Cornerstones of functional analysis . . . 23

3 Measure theory **27**
3.1 Measure spaces . . . 28
3.2 Outer measures and Caratéodory's extension theorem . . . 29
3.3 The Lebesgue measure . . . 31
3.4 The Hausdorff measure . . . 34
3.5 Measurable functions . . . 35
3.6 Integration . . . 38
3.7 Lebesgue spaces $L^p(\Omega)$ . . . 43

4 Dual spaces and weak topologies **46**
4.1 Locally convex spaces and weak topologies . . . 46
4.2 Banach–Alaoglu theorem . . . 48
4.3 Riesz representation for L^p spaces . . . 49

5 Hilbert spaces **52**
5.1 Orthogonality and projection in Hilbert spaces . . . 54
5.2 Riesz' representation for Hilbert spaces . . . 57
5.3 Bessel's inequality and Parseval identity . . . 58
5.4 Hilbert space adjoints . . . 61

6 Spectral theory **63**
6.1 Spectrum in Banach spaces . . . 64
6.2 Compact operators . . . 67
6.3 Spectral theorem for compact self-adjoint operators . . . 70
6.4 Fredholm alternative . . . 73

Functional anlaysis

This course provides an introduction to functional analysis. This is an advanced undergraduate course for which knowledge of real analysis and linear algebra, as well as a certain degree of mathematical maturity, is required.

Functional analysis has a rich and interesting history for which I recommend the beautiful book *History of Functional Analysis* by J. Dieudonné. Functional analysis arose in the late 19th century from questions in classical analysis such as the convergence of Fourier series, the existence of minimizers in the calculus of variations, or the solvability of integral equations as well as ordinary and partial differential equations. The solution spaces of these problems are typically spaces of functions which admit a natural vector space structure. An integral or differential operator acting on functions can be seen as a linear operator acting on elements of a vector space, and the question of solvability can now be abstractly formulated in terms of the invertibility of linear operators between vector spaces. Typically, these vector spaces of functions are *infinite*-dimensional. At its core, functional analysis studies the interplay between the linear structure of infinite dimensional vector spaces and topological concepts such as openness, compactness, continuity, as well as analytical concepts like differentiation and integration, and additional structures such as norms and inner products. In particular, the goal is to develop the abstract frameworks of metric spaces, normed spaces, Banach spaces, Hilbert spaces and establish general results that apply to concrete realizations of these spaces. This has far-reaching applications to many modern areas of mathematics and more generally in science. Indeed, functional analysis plays a key role in partial differential equations, quantum mechanics, numerical simulations, approximation theory, probability and statistics, machine learning, data science, control theory, optimization, and many other fields.

1 Basic notions, topological and metric spaces

1.1 Zorn's lemma

We begin with the famous Zorn's lemma which will appear in several places in the course, possibly most prominently in the proof of the Hahn–Banach theorem and the proof of Tychonoff's theorem. It is equivalent to the axiom of choice, which is a standard axiom in modern mathematics. Thus, we will treat Zorn's lemma as an axiom. However, because of the non-constructive nature of arguments involving the axiom of choice or Zorn's lemma, in applications of functional analysis to PDE or mathematical physics, one is often interested in circumventing the usage of Zorn's lemma. Typically, Zorn's lemma is used to infer the existence of some highly non-unique objects.

Definition 1.1. Let P be a set. A *partial order*, denoted by $\leq$, on P is a binary relation on P (i.e. a subset of the Cartesian product) such that for all $a, b, c \in P$

- $a \leq a$,
- $a \leq b$ and $b \leq c$ imply $a \leq c$,

- $a \leq b$ and $b \leq a$ imply $a = b$.

The tuple $(P, \leq)$ is called a *partially ordered set* or *poset*.

Definition 1.2. Let $(P, \leq)$ be a non-empty partially ordered set.

- An element $x \in P$ is an *upper bound* for a subset $A \subset P$ if $a \leq x$ for all $a \in A$.
- A subset $C \subset P$ is a *chain* if it is totally ordered (i.e. $x \leq y$ or $y \leq x$ for all $x, y \in C$.)
- An element $x \in P$ is *maximal* if $y \in P$ and $x \leq y$ implies $x = y$.

Axiom 1.3 (Zorn's lemma)**.** Let $(P, \leq)$ be a non-empty partially ordered set such that every non-empty chain C has an upper bound in P. Then, $(P, \leq)$ has a maximal element.

1.2 Vector spaces

All vector spaces treated in the course are over the fields $\mathbb{R}$ or $\mathbb{C}$. We will use the symbol $\mathbb{K}$ (the German word for field is "Körper") to mean either $\mathbb{R}$ or $\mathbb{C}$, i.e. $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 1.4. Let V be a vector space over the field $\mathbb{K}$.

- $A \subset V$ is *linearly independent* if $\sum_{i \in F} \lambda_i v_i = 0$ implies $\lambda_i = 0$ for all $i \in F$, where $\lambda_i \in \mathbb{K}$, $a_i \in A$ and F a finite index set.
- $A \subset V$ *spans* V if for every $v \in V$, there exists a finite index set F and $\lambda_i \in \mathbb{K}, v_i \in A$ for $i \in F$ such that $v = \sum_{i \in F} \lambda_i a_i$.
- A set $B \subset V$ which is linearly independent and spans V is called a *Hamel basis*.
- A subset $W \subset V$ is a *subspace* of V if W is itself a vector space with the operations induced by V.

Definition 1.5. A vector space V is called *finite dimensional* if it admits a Hamel basis with finite cardinality. Otherwise, it is called *infinite dimensional*.

Theorem 1.6. Every vector space V admits a Hamel basis B.

Proof. If $V = \{0\}$ choose $B = \emptyset$. If $V \neq \{0\}$ let $0 \neq v \in V$. Define the non-empty set $\mathcal{P} = \{C \in \mathcal{P}(V) : C \text{ is linearly independent and } v \in C\}$. $\mathcal{P}$ is partially ordered by set inclusion and for any chain $\mathcal{C} \subset \mathcal{P}$ the union $\bigcup_{A \in \mathcal{C}} A$ is an upper bound. Thus, $\mathcal{P}$ has an upper bound B. Since $B \in \mathcal{P}$, the set B is linearly independent. It remains to show that B spans V. Assume for a contradiction that $w \notin \operatorname{span}(B)$. Then $B \cup \{w\}$ is linearly independent and contains v so $B \cup \{w\} \in \mathcal{P}$. Since B is maximal, $B \cup \{w\} \subset B$ so $w \in B$ giving a contradiction. This shows that B is linearly independent and spans V, i.e. B is a Hamel basis. $\square$

Definition 1.7 (Algebraic dual space)**.** Let V be a vector space. We define the algebraic dual of V as the vector space $V' = \{f : V \to \mathbb{K}, f \text{ linear}\}$.

1.3 Topological spaces

A core aspect of functional analysis is the interplay between linear algebra and analysis. More precisely, it is about how the linear structure of vector spaces plays together with analytical structures of open and closed sets and the regularity of functions between such spaces. Especially for infinite-dimensional vector spaces, it will be important in what sense different vectors are close to each other and what exactly it means for functions to be continuous. This is captured by the structure of a topology.

Definition 1.8. Let X be a non-empty set. A *topology* $\mathcal{T}$ on X is collection of sets $\mathcal{T} \subset \mathcal{P}(X)$ such that

1. $\emptyset, X \in \mathcal{T}$.

2. If I is an index set and $U_i \in \mathcal{T}$ for all $i \in I$, then $\cup_{i \in I} U_i \in \mathcal{T}$.

3. If $U_1, \dots U_n \in \mathcal{T}$, then $\cap_{1 \leq i \leq n} U_i \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ or just X is called a *topological space*. Elements of $\mathcal{T}$ are called *open* sets.

- A set $A \subset X$ with $A^c = X \setminus A \in \mathcal{T}$ is called *closed*.

- If $\mathcal{T}_1, \mathcal{T}_2$ are two topologies on X, we say that $\mathcal{T}_1$ is *finer* than $\mathcal{T}_2$ if $\mathcal{T}_2 \subset \mathcal{T}_1$. Similarly, we say that $\mathcal{T}_1$ is *coarser* than $\mathcal{T}_2$ if $\mathcal{T}_1 \subset \mathcal{T}_2$.

Example.

- For any set X the finest and coarsest topologies are given by the discrete topology $\mathcal{T}_{\text{dis}} = \mathcal{P}(X)$ and the trivial topology $\mathcal{T}_{\text{tri}} = \{\emptyset, X\}$, respectively.

- The Euclidean topology on $\mathbb{R}^n$ is defined as follows. $O \subset \mathbb{R}^n$ is open if for all $x \in O$, there exists an $\varepsilon > 0$ such that $B_\varepsilon(x) \subset O$, where $B_\varepsilon(x) = \{y \in \mathbb{R}^n : |x - y| < \varepsilon\}$.

- Let $(X, \mathcal{T})$ be a topological space and $\emptyset \neq A \subset X$. The *relative topology* $\mathcal{T}_A \subset \mathcal{P}(A)$ on A is defined as $\mathcal{T}_A = \{A \cap U : U \in \mathcal{T}\}$. This makes $(A, \mathcal{T}_A)$ a topological space.

Definition 1.9. Let X be a topological space.

- $U \subset X$ is an *open neighborhood* of $x \in X$ if $x \in U$ and $U \in \mathcal{T}$.

- $A \subset X$ is a *neighborhood* of $x \in X$ if there exists an open neighborhood U of x such that $U \subset A$.

- X is a *Hausdorff space* if for all $x, y \in X$ with $x \neq y$, there exist neighborhoods U_x of x and U_y of y such that $U_x \cap U_y = \emptyset$.

- $x \in A$ is an *interior point* of A, if there exists a neighborhood U of x such that $U \subset A$. We define the *interior of* A, denoted by $\text{int}(A)$, as the union of all interior points.

- $x \in X$ is a *boundary point* of A if for all neighborhoods U of x, $U \cap A \neq \emptyset$ and $U \cap A^c \neq \emptyset$. We denote the set of boundary points of A as ∂A and the *closure of* A as $\overline{A} = A \cup \partial A$. Note that $\text{int}(A) = A \setminus \partial A$.

- $A \subset X$ is *dense* if $\overline{A} = X$.

- A sequence $(x_n)_{n\in\mathbb{N}} \subset X$ is convergent if there exists an $x \in X$ such that for all neighborhoods U of x, there exists a $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N$. We write $x = \lim_{n\to\infty} x_n$ or $x_n \to x$ as $n \to \infty$.

Remark 1.10. It is useful to convince yourself that the definition of convergence on $\mathbb{R}$ from 18.100 agrees with the more general definition for topological spaces.

Proposition 1.11. Let X be a topological space and $A \subset X$. Then:

1. A open $\Leftrightarrow A = \text{int}(A)$.
2. $\text{int}(A)$ is the largest open set contained in A.
3. $\overline{A}$ is the smallest closed set containing A.
4. A closed $\Leftrightarrow A = \overline{A}$.
5. ∂A is closed.

Proof. 1) "$\Rightarrow$" Clearly $\text{int}(A) \subset A$, and since A is open we also have $A \subset \text{int}(A)$. "$\Leftarrow$" It suffices to show that $\text{int}(A)$ is open. Note that $\text{int}(A) = \cup_{x\in\text{int}(A)} U_x$, where $U_x \subset A = \text{int}(A)$ is an open neighborhood of x. This shows that $\text{int}(A)$ is open. 2) – 5): Problem set 1. $\square$

Definition 1.12. Let X, Y be topological spaces and $f : X \to Y$ be a function.

- f is said to be *continuous* if the preimage of every open set is open.

- f is said to be *sequentially continuous* if $\lim_{n\to\infty} x_n = x$ implies $\lim_{n\to\infty} f(x_n) = f(x)$.

- f is said to be *open* if the image of every open set is open.

- f is said to be a *homeomorphism* if f is bijective, continuous and open.

Definition 1.13. Let X be a non-empty set and $(X_i)_{i\in I}$ be a family of topological spaces.

- Let $f_i : X \to X_i$ be a family of functions indexed by $i \in I$. The *initial topology* on X induced by f_i is the coarsest topology on X such that all f_i are continuous. In the context of topological vector spaces, the initial topology is also called *weak* topology.

- Let $g_i : X_i \to X$ be a family of functions indexed by $i \in I$. The *final topology* on X induced by g_i is the finest topology on X such that all g_i are continuous.

Definition 1.14. Let X_i, $i \in I$ be topological spaces, where I is an index set. We define the *product topology* on $X = \Pi_{i\in I} X_i$ as the initial topology induced by the projections $\pi_j : X \to X_j, (x_i)_{i\in I} \mapsto x_j$.

Definition 1.15. Let $(X, \mathcal{T})$ be a topological space.

- $\mathcal{B} \subset \mathcal{T}$ is called a *base* if every element of $\mathcal{T}$ can be written as a union of sets from $\mathcal{B}$.

- $\mathcal{S} \subset \mathcal{T}$ is called a *subbase* if finite intersections of sets from $\mathcal{S}$ form a base.

- $\mathcal{N} \subset \mathcal{T}$ is a *neighborhood base of* $x \in X$ if every element of $\mathcal{N}$ is a neighborhood of x and for every neighborhood U of x there exists an element $V \in \mathcal{N}$ such that $V \subset U$.

- $\mathcal{N} \subset \mathcal{T}$ is a *neighborhood subbase of* $x \in X$ if finite intersections of sets from $\mathcal{N}$ form a neighborhood base of $x \in X$.

- X is called *first countable* if every point $x \in X$ has a countable neighborhood base.

- X is called *second countable* if X has a countable base.

- X is called *separable* if there exists a countable set D such that $\overline{D} = X$.

Remark 1.16. If X, Y are topological spaces and $f : X \to Y$. Since taking the preimage f^{-1} commutes with taking unions and intersections, we obtain that f is continuous if and only if the preimage of every subbasis element is open.

Example.

- Let $X = \Pi_{i\in I} X_i$. Then, the collection of sets of the form $\pi_i^{-1}(U_i)$ where U_i is open in X_i form a subbase of the product topology on X and sets of the form $\pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_n}^{-1}(U_{i_n})$ form a base of the product topology. In mild abuse of notation, we can also write that the sets $\Pi_{i\in I} U_i$ form a basis, where $U_i = X_i$ for all but finitely many $i \in I$. For instance if $X = X_1 \times X_2$, then sets of the form $U_1 \times U_2 = \pi_1^{-1}(U_1) \cap \pi_2^{-1}(U_2)$, where $U_1 \subset X_1, U_2 \subset X_2$ are open, form a basis.

- Let $\mathbb{R}^n$ be equipped with the Euclidean topology. Then, for $x \in \mathbb{R}^n$, $\{B_{1/k}(x)\}_{k\in\mathbb{N}}$ is a neighborhood base at x and $\{B_{1/k}(q) : q \in \mathbb{Q}, k \in \mathbb{N}\}$ is a countable base for $\mathbb{R}^n$. Thus, $\mathbb{R}^n$ is second countable.

Remark 1.17. For any family $\mathcal{C} \subset \mathcal{P}(X)$ there exists a unique topology $\mathcal{T}_\mathcal{C}$, defined as the coarsest topology on X containing $\mathcal{C}$. If the union of elements in $\mathcal{C}$ covers X, then $\mathcal{C}$ forms a subbase for $\mathcal{T}_\mathcal{C}$.

Proposition 1.18. Let X be a topological space. If X is second countable, then it is first countable and separable.

Proof. Let $\mathcal{B}$ be a countable base. First countability is clear. For the separability, let D be a countable set such that for all $\emptyset \neq B \in \mathcal{B}$, there exists a $b \in B$ with $b \in D$. Let $x \in X$ and let U be an open neighborhood around x. Since $\mathcal{B}$ is a base, there exists a $B \in \mathcal{B}$ such that $B \subset U$. Thus, there exists a $b \in D \cap B$ such that $b \in U$. Hence, $\overline{D} = X$. □

Proposition 1.19. Let X, Y be topological spaces and $f : X \to Y$ be a map.

1. If f is continuous, then f is sequentially continuous.

2. If f is sequentially continuous and X is first countable, then f is continuous.

Proof. 1) If f is continuous and $x_n \to x$ in X. Let V be an open neighborhood of $f(x)$ in Y. Since f is continuous $U = f^{-1}(V)$ is open and since $x_n \to x$ in X, eventually all x_n are in U. Hence, $f(x_n) \in V$ eventually, which shows that $f(x_n) \to f(x)$.

2) Let $V \subset Y$ open. If $f^{-1}(V) = \emptyset$, then $f^{-1}(V)$ is open, so we can assume that $f^{-1}(V) \neq \emptyset$. Let $x \in f^{-1}(V)$. For a contradiction, assume that x is not an interior point of $f^{-1}(V)$. Let $(U_n)_n$ be a countable, nested neighborhood base around x. Since x is not an interior point, there exists a sequence $x_n \in U_n \setminus f^{-1}(V)$ which in particular satisfies $x_n \to x$ as $n \to \infty$. Since f is sequentially continuous, we have that $f(x_n) \to f(x)$. Hence, $f(x_n) \in V$ for $n \geq N$ and thus, $x_n \in f^{-1}(V)$ for $n \geq N$ which is a contradiction. □

Definition 1.20. Let X be a topological space and $K \subset X$.

- K is *compact* if every open covering of K has a finite subcovering.

- K is *sequentially compact* if every sequence in K has a convergent subsequence in K.

- K is *relatively compact* if $\overline{K}$ is compact.

- A collection of sets $\mathcal{A} \subset \mathcal{P}(X)$ is said to have the *finite intersection property (FIP)* if $\cap_{1 \leq j \leq n} U_j \neq \emptyset$ for any finite subfamily $\{U_j\}_{1 \leq j \leq n} \subset \mathcal{A}$.

Proposition 1.21. Let X be a topological space and $K \subset X$. Then, K is compact if and only if every family of closed sets having the finite intersection property has non-empty intersection.

Proof. Problem set 1. □

Theorem 1.22 (Tychonoff)**.** Let X_i, $i \in I$ be topological spaces. $X = \Pi_{i \in I} X_i$ equipped with the product topology is compact if and only if X_i is compact for all $i \in I$.

Proof. "$\Rightarrow$" Clear since π_i are continuous and the compact set X is mapped to a compact set X_i. So we will show "$\Leftarrow$" in the following. Let C be a non-empty family of closed sets of X having the finite intersection property. We want to show that $\bigcap_{c \in C} c \neq \emptyset$.

The collection $\mathcal{P} = \{D \subset \mathcal{P}(X) : C \subset D, D \text{ has the FIP}\}$ is partially ordered by set inclusion. Let $\mathcal{D}$ be a non-empty chain in $\mathcal{P}$, then $\cup_{E \in \mathcal{D}} E \in \mathcal{P}$ is an upper bound. By

Zorn's lemma, there exists a maximal element $B \in \mathcal{P}$ and we note that $C \subset B$. The relevance of B is that it is maximal which means that (i) $e, f \in B \Rightarrow e \cap f \in B$ and (ii) $e \cap b \neq \emptyset$ for all $b \in B \Rightarrow e \in B$.

Since the elements of C are closed and $C \subset B$, we have $\bigcap_{b \in B} \bar{b} \subset \bigcap_{c \in C} c$, so it suffices to show that $\cap_{b \in B} \bar{b} \neq \emptyset$.

We now pull back to each product and note that $\{\pi_i(b) : b \in B\}$ has the FIP so in particular, $\{\overline{\pi_i(b)} : b \in B\}$ has the FIP for each $i \in I$. By compactness for each X_i, we obtain $x = (x_i)_{i \in I}$ with $x_i \in \bigcap_{b \in B} \overline{\pi_i(b)}$.

Finally, we will show that $\cap_{b \in B} \bar{b} \neq \emptyset$. To do set let v be a neighborhood of x, where we assume without loss of generality that $v = \bigcap_{j=1}^{n} \pi_{i_j}^{-1}(v_{i_j})$ for v_{i_j} open neighborhoods of x_{i_j}.

Since $x_{i_j} \in v_{i_j} \cap \overline{\pi_{i_j}(b)}$ for all $b \in B$, we have that $v_{i_j} \cap \pi_{i_j}(b) \neq \emptyset$ for all $b \in B$ and thus, $\pi_{i_j}^{-1}(v_{i_j}) \cap b \neq \emptyset$ for all $b \in B$. By (ii), this means that $\pi_{i_j}^{-1}(v_{i_j}) \in B$ for each $j = 1, \ldots, n$ and hence, by (ii), $v \in B$, i.e. $v \cap b \neq \emptyset$ for all $b \in B$. Since v was an arbitrary neighborhood of x, we have that $x \in \bar{b}$ for all $b \in B$ and thus $x \in \cap_{b \in B} \bar{b} \neq \emptyset$. □

Definition 1.23. A *topological vector space* $(V, \mathcal{T})$ is a vector space V equipped with a Hausdorff topology $\mathcal{T}$ such that scalar multiplication $\mathbb{K} \times V \to V$ and addition $V \times V \to V$ are continuous.

Definition 1.24 (Continuous dual)**.** Let V be a topological vector space. Define the *continuous dual* V^* as $V^* = \{f \in V' : f \text{ continuous}\}$.

Example. There are examples (e.g. $L^p(\mathbb{R}), 0 < p < 1$) of topological vector spaces which have trivial continuous dual.

Many important results in functional analysis can be formulated in the context of topological vector spaces. Many interesting spaces such as $C^\infty(\mathbb{R})$ or spaces of distributions are merely topological vector spaces that cannot be equipped with a norm. Nevertheless, in this course, we will be mostly concerned with more rigid topological structures on our vector spaces, namely those of normed (or metric) spaces.

1.4 Metric spaces

Definition 1.25. Let X be a non-empty set. A mapping $d : X \times X \to [0, \infty)$ is a *metric* if it satisfies

1. $d(x, y) = 0 \Leftrightarrow x = y$,
2. $d(x, y) = d(y, x)$,
3. $d(x, y) \leq d(x, z) + d(z, y)$

for all $x, y, z \in X$. A pair (X, d) is called a metric space. We denote $B_\varepsilon(x) = \{y \in X : d(x, y) < \varepsilon\}$ and $\overline{B}_\varepsilon(x) = \{y \in X : d(x, y) \leq \varepsilon\}$.

Remark 1.26. A metric space (X, d) carries a natural topology, the *metric topology* induced by the subbase $\{B_{1/n}(x)\}_{x \in X, n \in \mathbb{N}}$. It is easy to verify that

- the topology is Hausdorff and first countable (i.e. the topology is characterized by convergent and divergent sequences),
- $A \subset X$ is closed if and only if $(x_n)_n \subset A$ with $x_n \to x$ implies $x \in A$,
- $x_n \to x$ as $n \to \infty$ if and only if for all $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$ for all $n \geq N$,
- and if Y is a topological space then $f : X \to Y$ is continuous if and only if f is sequentially continuous (recall Proposition 1.19),
- the family $\{B_{1/n}(x)\}_{x \in X, n \in \mathbb{N}}$ is a base for the topology.

Proposition 1.27. Let X be a metric space, and let $x \in X$ and $\varepsilon > 0$. Then:

1. $B_\varepsilon(x)$ is open.
2. $\bar{B}_\varepsilon(x)$ is closed.
3. $d : X \times X \to \mathbb{R}$ is continuous.

Proof. Let $y \in B_\varepsilon(x)$ and let fix $k \in \mathbb{N}$ with $0 < \frac{1}{k} < \varepsilon - d(x, y)$. Then, $B_{\frac{1}{k}}(y) \subset B_\varepsilon(x)$ because for $z \in B_{\frac{1}{k}}(y)$ we have $d(z, x) \leq d(z, y) + d(y, x) < \frac{1}{k} + d(x, y) < \varepsilon$. A similar argument also shows that $\bar{B}_\varepsilon(x)$ is closed. The last property is also a consequence of the triangle inequality and is left as an exercise. $\square$

Remark 1.28. Note that $\overline{B_\varepsilon(x)} \subset \bar{B}_\varepsilon(x)$. For example, the above inclusion is strict for $X = \mathbb{R}$ with the discrete metric ($d(x, y) = 1$ for $x \neq y$ and $d(x, x) = 0$) which induces the discrete topology.

Definition 1.29. Let (X, d) be a metric space. A sequence $(x_n)_{n \in \mathbb{N}} \subset X$ is called a *Cauchy sequence* if for all $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq N$.

Definition 1.30. X is *complete* if every Cauchy sequence in X converges.

Remark 1.31. Completeness is a notion of metric spaces. There exist metric spaces (X, d) and (X, d') with the same topology but such that (X, d) is complete while (X, d') is incomplete, see problem set 1.

Proposition 1.32. Let (X, d) be a metric space. Then, X is separable if and only if it is second countable.

Proof. "$\Leftarrow$" See Proposition 1.18. "$\Rightarrow$" See problem set 1. $\square$

Proposition 1.33. Let X be a complete metric space. A subspace $W \subset X$ is complete if and only if it is closed.

Proof. "$\Rightarrow$" Let $(x_n)_n \subset W$ satisfy $x_n \to x$ as $n \to \infty$. Then x_n is a Cauchy sequence and since W is complete, $x \in W$ which shows that W is closed. (In a metric space A is closed if and only if $x_n \to x$, where $x_n \in A$ implies $x \in A$.) "$\Leftarrow$" Let $(x_n)_n \subset W$ be a Cauchy sequence. Since X is complete, $x_n \to x$ as $n \to \infty$ for come $x \in X$. Moreover, $x \in W$ because W is closed. $\square$

Proposition 1.34. Let X be a metric space. Then $K \subset X$ is compact if and only if K is sequentially compact.

Proof. See problem set 2. $\square$

Remark 1.35. For first countable topological spaces only the direction "$\Rightarrow$" in Proposition 1.34 is true. For second countable topological spaces the equivalence in Proposition 1.34 is also true.

Definition 1.36. Let (X, d) and (Y, d') be metric spaces.

- A map $\Phi : X \to Y$ is an *isometry* if $d(x, y) = d'(\Phi(x), \Phi(y))$ for all $x, y \in X$.
- We say (X, d) and (Y, d') are *isometric* if there exists a bijective isometry $\Phi : X \to Y$.

Theorem 1.37 (Completion of a metric space)**.** Let (X, d) be a metric space. There exists a complete metric space (X', d') in which (X, d) embeds isometrically as a dense set, i.e. there exists a map $\Phi : X \to X'$ such that $\Phi(X) \subset X'$ is dense and $d'(\Phi(x), \Phi(y)) = d(x, y)$ for all $x, y \in X$. Moreover, the space (X', d')—the completion of (X, d)—is unique up to isometry.

Proof. 1) **Construction of (X', d').** Consider the set

$$\{(x_n)_n \subset X : x_n \text{ is a Cauchy sequence}\}$$

and define $(x_n)_n \sim (x'_n)_n$ if $d(x_n, x'_n) \to 0$ as $n \to \infty$. Denote the set of equivalence classes as X' and define a metric on X' by $d'([(x_n)_n], [(y_n)_n]) = \lim_{n\to\infty} d(x_n, y_n)$, where x_n, y_n are representatives of the equivalence class. The above limit exists because we will show below that $(d(x_n, y_n))_n$ is a Cauchy sequence in $\mathbb{R}$ which is complete. Indeed, using the triangle inequality and the reverse triangle inequality we have

$$\begin{aligned}
|d(x_n, y_n) - d(x_m, y_m)| &= |d(x_n, y_n) - d(x_m, y_n) + d(x_m, y_n) - d(x_m, y_m)| \\
&\le |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)| \\
&\le d(x_n, x_m) + d(y_n, y_m) \to 0 \text{ as } n, m \to \infty.
\end{aligned}$$

2) d' **is a well-defined metric.** We have to show that d' is independent of the representative of the equivalence class. Let $[(x_n)_n] = [(x'_n)_n]$ and $[(y_n)_n] = [(y'_n)_n]$. Then, as before,

$$\begin{aligned}
\lim_{n\to\infty} |d(x_n, y_n) - d(x'_n, y'_n)| &\le \limsup_{n\to\infty} |d(x_n, y_n) - d(x'_n, y_n) + d(x'_n, y_n) - d(x'_n, y'_n)| \\
&\le \limsup_{n\to\infty} |d(x_n, y_n) - d(x'_n, y_n)| + |d(x'_n, y_n) - d(x'_n, y'_n)| \\
&\le \limsup_{n\to\infty} d(x_n, x'_n) + d(y'_n, y_n) = 0,
\end{aligned}$$

where we used that $[(x_n)_n] = [(x'_n)_n]$ and $[(y_n)_n] = [(y_n)'_n]$. It is clear that d' satisfies the properties of a metric and in particular, (X', d') is a metric space.

3) $X \to X'$**,** $x \mapsto [(x, x, x, \dots)]$ **is an isometry with dense image.** The map

$$\Phi : X \to X', x \mapsto [(x, x, x, \dots)]$$

is manifestly an isometry and hence continuous and injective. In order to show that the image of Φ is dense let $[(x_n)_n] \in X'$ and $\varepsilon > 0$. Since $(x_n)_n$ is Cauchy, there exists an $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/2$ for all $n, m \geq N$. Thus, $d'([(x_n)_n], \Phi(x_N)) = \lim_{n\to\infty} d(x_n, x_N) \leq \varepsilon/2 < \varepsilon$. Thus, $\Phi(X)$ is dense in X'.

4) (X', d') **is complete.** It suffices to check the completeness with Cauchy sequences in the dense set $\Phi(X) \subset X'$ (see problem set 1). Let $(\Phi(x_m))_m$ be a Cauchy sequence in (X', d'). Then, $(x_m)_m$ is a Cauchy sequence in X because $d(x_m, x_n) = d'(\Phi(x_m), \Phi(x_n))$. In particular, $[(x_n)_n] \in X'$. We will show that $(\Phi(x_m))_m \to [(x_n)_n]$ as $m \to \infty$.

Let $\varepsilon > 0$. Since $(x_m)_m$ is a Cauchy sequence, there exists $N \in \mathbb{N}$ such that $d(x_m, x_n) < \varepsilon/2$ for all $n, m \geq N$. Hence, $d'(\Phi(x_m), [(x_n)_n]) = \lim_{n\to\infty} d(x_m, x_n) < \varepsilon$ for all $m \geq N$.

5) **Uniqueness up to isometry.** Let $(\tilde{X}, \tilde{d})$ be another completion of (X, d), i.e. there exists an isometry $\tilde{\Phi} : X \to \tilde{X}$ with dense image. Define the map

$$\Psi = \Phi \circ \tilde{\Phi}^{-1} : \tilde{\Phi}(X) \to \Phi(X)$$

which is an isometry defined on the dense set $\tilde{\Phi}(X) \subset \tilde{X}$. The map Ψ extends uniquely to an isometry $\tilde{\Psi} : \tilde{X} \to X'$ by setting $\tilde{\Psi}(x) = \lim_{n\to\infty} \Psi(x_n)$, where $x_n \to x$ in $\tilde{X}$ and $(x_n)_n \subset \tilde{\Phi}(X)$. Since Ψ is an isometry, the sequence $(\Psi(x_n))_n$ is a Cauchy sequence and thus the limit $\lim_{n\to\infty} \Psi(x_n)$ converges. Using an analogous argument to the arguments before, the limit is independent of the sequence $(x_n)_n$. Similarly, one can show that $\tilde{\Psi}$ is onto which shows that $\tilde{X}$ and X' are isometric. We leave these details as an exercise.

□

Theorem 1.38 (Young's inequality)**.** For $a, b \geq 0$ and $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. Without loss of generality $a, b \neq 0$. In the case $p = q = 2$ this follows from $(a - b)^2 \geq 0$. For the general case, we use the *concavity* of the logarithm:

$$\log(\lambda x + (1 - \lambda)y) \geq \lambda \log(x) + (1 - \lambda) \log(y)$$

for $x, y > 0$ and $\lambda \in [0, 1]$. For $\lambda \in (0, 1)$ we write

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \lambda \log\left(\frac{1}{\lambda}\frac{a^p}{p}\right) + (1 - \lambda) \log\left(\frac{1}{1 - \lambda}\frac{b^q}{q}\right).$$

Choosing $\lambda = \frac{1}{p}$, $1 - \lambda = 1 - \frac{1}{p} = \frac{1}{q}$ and exponentiation gives the result. □

Remark 1.39. Note that the condition $1 = \frac{1}{p} + \frac{1}{q}$ on the powers is necessary. Indeed, assume that for all $a, b > 0$ we have $ab \leq c_1 a^p + c_2 b^q$ for some $p, q > 1$ and $c_1, c_2 > 0$. Then, setting $a = x^{\frac{q}{p}}$ and $b = x$ gives $x^{1+\frac{q}{p}} \leq (c_1 + c_2)x^q$. Then, sending $x \to \infty$ gives $1 + \frac{q}{p} \leq q$ and sending $x \to 0$ gives $1 + \frac{q}{p} \geq q$ so $\frac{1}{q} + \frac{1}{p} = 1$.

Theorem 1.40 (Hölder's inequality)**.** Assume that $p, q > 1$ are Hölder conjugates, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Let $(x_n)_n, (y_n)_n \subset \mathbb{K}$ be sequences such that

$$\sum_{n=1}^{\infty} |x_n|^p < \infty \text{ and } \sum_{n=1}^{\infty} |y_n|^q < \infty.$$

Then,

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} \left(\sum_{n=1}^{\infty} |y_n|^q \right)^{1/q}.$$

Proof. We define the normalized quantities

$$\tilde{x}_n = \frac{x_n}{\left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}}, \text{ and } \tilde{y}_n = \frac{y_n}{\left(\sum_{n=1}^{\infty} |y_n|^q\right)^{1/q}}.$$

Then, using Young's inequality, we obtain

$$\sum_{n=1}^{N} |\tilde{x}_n \tilde{y}_n| \leq \sum_{n=1}^{N} \frac{|\tilde{x}_n|^p}{p} + \sum_{n=1}^{N} \frac{|\tilde{y}_n|^q}{q} \to \frac{1}{p} + \frac{1}{q} = 1$$

as $N \to \infty$. □

Remark 1.41. The condition $1 = \frac{1}{p} + \frac{1}{q}$ is necessary.

Remark 1.42. For the limiting case $p = 1$ and $q = \infty$, the inequality

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \sup_{n \in \mathbb{N}} |y_n| \cdot \sum_{n=1}^{\infty} |x_n|$$

is also often included in the family of Hölder's inequalities.

Proposition 1.43. The following spaces of sequences are examples of metric spaces.

1. For $1 \leq p < \infty$, we set $\ell^p = \{(x_n)_n \subset \mathbb{K} : \sum_{i=1}^{\infty} |x_n|^p < \infty\}$ with metric $d(x, y) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^p\right)^{1/p}$.
2. For $p = \infty$, we set $\ell^\infty = \{(x_n)_n \subset \mathbb{K} : \sup_{n \in \mathbb{N}} |x_n| < \infty\}$ with metric $d(x, y) = \sup_{n \in \mathbb{N}} |x_n - y_n|$.
3. $c_0 = \{(x_n)_n \in \ell^\infty : \lim_{n \to \infty} x_n = 0\}$ with metric $d(x, y) = \sup_{n \in \mathbb{N}} |x_n - y_n|$.
4. $c_c = \{(x_n)_n \in \ell^\infty : x_n = 0 \text{ for all but finitely many } n \in \mathbb{N}\}$ with metric $d(x, y) = \sup_{n \in \mathbb{N}} |x_n - y_n|$.

Proof. The fact that 2)–4) are metric spaces follows directly from the triangle inequality. The fact that ℓ^p is a metric space for $1 \leq p < \infty$ follows from the Minkowski inequality below. $\square$

Lemma 1.44 (Minkowski inequality)**.** Let $x, y \in \ell^p$ for $1 \leq p < \infty$. Then,

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{\frac{1}{p}} \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{1}{p}}$$

Proof. See problem set. $\square$

Proposition 1.45. The space c_0 and the space ℓ^p for any $p \in [1, \infty]$ are complete.

Proof. We begin with the space ℓ^p for $1 \leq p < \infty$. Let $((x_n^m)_n)_m$ be a Cauchy sequence in ℓ^p. Clearly, for fixed $n \in \mathbb{N}$ the sequence $(x_n^m)_m$ is a Cauchy sequence in $\mathbb{K}$ and hence convergent. By completeness of $\mathbb{K}$ we have a limit x_n for each $n \in \mathbb{N}$. We will first show that $(x_n)_n \in \ell^p$. Indeed,

$$\sum_{n=1}^{N} |x_n|^p = \lim_{m\to\infty} \sum_{n=1}^{N} |x_n^m|^p \leq \limsup_{m\to\infty} \sum_{n=1}^{\infty} |x_n^m|^p = \limsup_{m\to\infty} d(0, x^m) < \infty$$

since Cauchy sequences are bounded. Sending $N \to \infty$ shows that $(x_n)_n \in \ell^p$. To show that $x^m \to x$ in ℓ^p we let $\varepsilon > 0$ and let $M \in \mathbb{N}$ such that $d(x^m, x^n) < \varepsilon$ for $n, m \geq M$. Then,

$$\sum_{k=1}^{N} |x_k^m - x_k^n|^p \leq \varepsilon^p \Rightarrow \sum_{k=1}^{N} |x_k^m - x_k|^p \leq \varepsilon^p \Rightarrow \sum_{k=1}^{\infty} |x_k^m - x_k|^p \leq \varepsilon^p,$$

from which we obtain that $x^m \to x$ as $m \to \infty$. For the cases $p = \infty$ and c_0, see problem set 2. $\square$

Proposition 1.46. The space ℓ^p is separable for $1 \leq p < \infty$. ℓ^∞ is not separable.

Proof. See problem set 2. $\square$

Proposition 1.47. Let K be a compact topological space. Consider the space $C(K)$ of real-valued, continuous functions. The assignment $d(f, g) = \sup_{x\in K} |f(x) - g(x)|$ defines a metric on $C(K)$ which makes $C(K)$ a complete metric space.

Proof. Since K is compact, d is well-defined so that $C(K)$ is a metric space. It remains to show that it is complete. Let $(f_n)_n$ be a Cauchy sequence in $C(K)$. For each $x_0 \in K$, the sequence $(f_n(x_0))_n$ is Cauchy in $\mathbb{R}$ and we define f as the pointwise limit $f(x) := \lim_{n\to\infty} f_n(x)$. We will now show that $f_n \to f$ with respect to d. Let $\varepsilon > 0$. Then $\sup_{x\in K} |f_n(x) - f_m(x)| \leq \varepsilon$ for $n, m > N$. Thus, for every $x \in K$ and $n > N$ we have

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \varepsilon \quad \Rightarrow \quad \sup_{x\in K} |f_n(x) - f(x)| \leq \varepsilon.$$

This shows that $f_n \to f$ uniformly.

Finally, we want to show that f is continuous. Let $x \in K$ and let $\varepsilon > 0$. Then, choose $N \geq 1$ such that $d(f, f_N) < \varepsilon/3$. Moreover, since f_N is continuous, there exists an open neighborhood $U_{x,N}$ of x such that $|f_N(x) - f_N(y)| < \varepsilon/3$ for all $y \in U_{x,N}$. Then, for $y \in U_{x,N}$

$$|f(x)-f(y)| \leq |f(x)-f_N(x)|+|f_N(x)-f_N(y)|+|f_N(y)-f(y)| < 2\varepsilon/3+|f_N(x)-f_N(y)| < \varepsilon.$$

Since x was arbitrary, this shows that f is continuous and concludes the proof. $\square$

Remark 1.48. The last part of the above argument also shows the *uniform limit theorem*: If a sequence f_n of continuous functions on a topological space converges uniformly to f, then f also has to be continuous.

Remark 1.49. If K is not compact, the same argument as above shows that $C_b(K)$, the space of bounded continuous functions forms a Banach space.

1.5 Baire category theorem

Definition 1.50. Let X be a topological space and $A \subset X$. A is said to be

- *nowhere dense* if $\bar{A}$ contains no interior points.
- *meager* (or of *first category*) if $A \subset \cup_{i \in \mathbb{N}} B_i$, where B_i are nowhere dense.
- *coarse* (or of *second category*) if A is not meager.
- *Baire-generic* (or *residual*) if A^c is meager.

Theorem 1.51 (Baire category theorem)**.** Let X be a complete metric space and let $(A_n)_n$ be a family of open and dense sets. Then, $\cap_{n \in \mathbb{N}} A_n$ is dense.

Proof. Let $x_0 \in X$ and $B_\varepsilon(x_0)$ be an open ball around x_0. Since A_1 is open and dense and $B_\varepsilon(x_0)$ is open, there exists $x_1 \in X$ and $\varepsilon_1 > 0$ such that $\overline{B}_{\varepsilon_1}(x_1) \subset A_1 \cap B_\varepsilon(x_0)$. For $n \geq 2$ we use that A_n is open and dense, to find $x_n \in X$ and $0 < \varepsilon_n < \frac{1}{n}$ such that

$$\overline{B}_{\varepsilon_n}(x_n) \subset B_{\varepsilon_{n-1}}(x_{n-1}) \cap A_n \cap B_\varepsilon(x_0).$$

Since the balls $B_{\varepsilon_n}(x_n)$ are nested we have that $d(x_n, x_m) \leq \frac{1}{\min(n,m)} \to 0$ as $n, m \to \infty$. Thus, $(x_n)_n$ is a Cauchy sequence and by completeness of X, there exists an $x \in X$ such that $x_n \to x$ as $n \to \infty$. By construction $x_n \in \overline{B}_{\epsilon_m}(x_m) \subset A_m$ for all $n \geq m$ and since $\overline{B}_{\varepsilon_m}(x_m)$ is closed we have that

$$x \in \overline{B}_{\epsilon_m}(x_m) \subset A_m \cap B_\varepsilon(x_0)$$

for all $m \in \mathbb{N}$ which concludes the proof. $\square$

Corollary 1.52. A complete metric space is not the countable union of nowhere dense sets.

Proof. Define $A = \cup_{n\in\mathbb{N}} A_n$, where A_n are nowhere dense sets. In particular, $\overline{A_n}$ is closed and nowhere dense for every n. Hence, $\overline{A_n}^c$ is open and dense for every n so by Baire's theorem, $\emptyset \neq \cap_{n\in\mathbb{N}} \overline{A_n}^c = (\cup_{n\in\mathbb{N}} \overline{A_n})^c \subset A^c$ so $A \neq X$. □

Theorem 1.53. A Baire-generic continuous function on [0,1] is nowhere differentiable. More precisely, the set $A = \{f \in C[0,1] : f \text{ is nowhere differentiable}\}$ is a Baire-generic subset of $C[0,1]$.

Proof. We define for $n, k \in \mathbb{N}$ the sets

$$A_{n,k} := \{f \in C[0,1] : \text{for all } x \in [0,1], \exists y \in B_{1/k}(x) \text{ such that } |f(x) - f(y)| > n|x-y|\}.$$

We note that $\cap_{n,k\in\mathbb{N}} A_{n,k} \subset A$. We claim that $A_{n,k}$ is open and dense for each $n, k \in \mathbb{N}$.

To show that $A_{n,k}$ is open, we will show that $A_{n,k}^c$ is closed. Let $(f_m)_m$ be a sequence in $A_{n,k}^c$ such that $f_m \to f$ in $C[0,1]$. We will show that $f \in A_{n,k}^c$. Since $f_m \in A_{n,k}^c$, there is a sequence x_m such that $|f_m(x_m) - f(y)| \leq n|x_m - y|$ for all $y \in B_{1/k}(x_m)$. By compactness, x_m has a convergent subsequence with limit x. In mild abuse of notation, we denote the subsequence again by x_m. Then, for $y \in B_{1/k}(x)$ we have

$$\begin{aligned}|f(x_m) - f(y)| &\leq |f_m(x_m) - f(x_m)| + |f_m(x_m) - f_m(y)| + |f_m(y) - f(y)| \\ &\leq 2 \sup_{x\in[0,1]} |f_m(x) - f(x)| + n|x_m - y|.\end{aligned}$$

Taking the limit $m \to \infty$ on both sides shows $|f(x) - f(y)| \leq n|x-y|$ for all $y \in B_{1/k}(x)$ which shows that $f \in A_{n,k}^c$ is closed and $A_{n,k}$ is open.

In order to show that $A_{n,k}$ is dense, we let $f \in C[0,1]$ and $\varepsilon > 0$ be given. Note that piecewise linear functions are dense in $C[0,1]$ (recall problem set 1) so we can assume without loss of generality that f is piecewise linear. Let $\alpha_f \geq 0$ be the largest slope (in magnitude) of f. Let g be a piecewise linear function such that $\sup_{x\in[0,1]} |g(x)| < \varepsilon$, and $|g'(x)| > \alpha_f + n$ for all $x \in [0,1]$ where g is differentiable. We claim that $f + g \in A_{n,k}$. Indeed, let $x \in [0,1]$. Then, there exists a $y \in B_{1/k}(x)$ such that $|g(x) - g(y)| > (n + \alpha_f)|x-y|$. In particular,

$$\begin{aligned}|g(x) + f(x) - g(y) - f(y)| &\geq |g(x) - g(y)| - |f(x) - f(y)| \\ &\geq (n+\alpha_f)|x-y| - \alpha_f|x-y| > n|x-y|\end{aligned}$$

so $f + g \in A_{n,k}$. To conclude we note that by Baire's theorem the set $A = \bigcap_{n,k\in\mathbb{N}} A_{n,k}$ is dense (in fact Baire generic). □

Remark 1.54. There is a saying among mathematicians: "Most continuous functions are nowhere differentiable, but nobody has ever seen one." While this is mostly true in practice for an analyst, you may have encountered such a function in fractal geometry or in probability theory: for instance the Wiener process[1] is a famous example of a continuous but nowhere differentiable curve.

[1]Norbert Wiener (1894-1964) is somewhat of a legend at MIT because he spent essentially his whole professional career at MIT. The common room in the mathematics department is named after him.

The first rigorous construction of such a continuous but nowhere differentiable function was given by Karl Weierstraß in 1872. Weierstraß can be seen as the father of modern analysis and played a key role in establishing the firm foundations of analysis. For mathematicians like Hermite or Poincaré, who practiced a more intuitive approach to analysis, the construction came as a surprise. Poincaré famously called it a "monster".

2 Normed and Banach spaces

Definition 2.1. A *normed space* is a vector space V equipped with a *norm* $\|\cdot\|$, i.e. a map $\|\cdot\| : V \to [0, \infty)$ such that

1. $\|\alpha v\| = |\alpha| \|v\|$
2. $\|v + w\| \leq \|v\| + \|w\|$
3. $\|v\| = 0$ implies $v = 0$.

A map $p : V \to [0, \infty)$ for which condition 3. (positive definiteness) is dropped is called a *seminorm*.

Remark 2.2. If clear from the context, we will often use the notation $\|\cdot\|$ for different norms on different spaces. We will use the initiation $\|\cdot\|_X$ to explicitly mention that the norm is with respect to the normed space X.

Remark 2.3. A norm on V induces a natural metric, $d(v, w) = \|v - w\|$, and hence a natural topology on V. From the triangle inequality, it is easy to verify that addition, multiplication and the map $x \mapsto \|x\|$ are continuous (Exercise).

Definition 2.4. A complete normed space is called a *Banach space*.

Theorem 2.5. Let X be a normed space. Then there exists a Banach space $\tilde{X}$ and an isometry from X onto a dense subspace of $\tilde{X}$. The space $\tilde{X}$ is unique up to isometric isomorphism.

Proof. The proof is similar to the metric space competition and will be left to the reader. $\square$

Example. All finite dimensional normed space V (e.g. $\mathbb{R}^n$ or $\mathbb{C}^n$) are Banach spaces. Moreover, ℓ^p for $p \in [1, \infty]$ and $C(K)$ are Banach spaces. For $x \in \ell^p$ we use the notation $\|x\|_p = (\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}}$.

Definition 2.6. Let X be a normed space. A set $\{e_n, n \in \mathbb{N}\}$ is a Schauder basis of X if for every $x \in X$ there exists a unique sequence $x_n \in \mathbb{K}$ such that $\sum_{n=1}^{N} x_n e_n \to x$ as $N \to \infty$ or equivalently, $\|\sum_{n=1}^{N} x_n e_n - x\| \to 0$ as $N \to \infty$.

Lemma 2.7. Let X be a normed space. If X admits a Schauder basis, then it is separable.

Proof. If X is a complex vector space consider the countable set $A_N = \{\sum_{i=1}^N \lambda_i e_i, \lambda_i \in \mathbb{Q} + i\mathbb{Q}\}$. It is easy to see that $\cup_{n\in\mathbb{N}} A_n$ is dense in X. $\square$

Remark 2.8. Note that the reverse implication in Lemma 2.7 is not true.

We will now proceed to a central lemma for infinite dimensional normed spaces. It can be seen as a statement about "orthogonality" in normed spaces up to $\varepsilon > 0$.

Lemma 2.9 (Riesz's lemma)**.** Let X be a normed space and Y be a proper closed subspace. For every $\varepsilon > 0$ there exists a vector $x \in X \setminus Y$ with $\|x\| = 1$ such that $\inf_{y\in Y} \|x - y\| \geq 1 - \varepsilon$.

Proof. Let $z \in X \setminus Y$. Since Y is closed $d := \inf_{y\in Y} \|z - y\| > 0$. Choose $y \in Y$ such that $d \leq \|z - y\| \leq d/(1 - \varepsilon)$. Set $x = (z - y)/(\|z - y\|)$ and note that $\|x\| = 1$. For $w \in Y$ we compute

$$\|x - w\| = \frac{1}{\|z - y\|} \|z - y - \|z - y\|w\| \geq \frac{1 - \varepsilon}{d} d = 1 - \varepsilon$$

since $-y - \|z - y\|w \in Y$. $\square$

Theorem 2.10. Let X be a normed space. Then X is finite-dimensional if and only if $\bar{B}_1(0)$ is compact.

Proof. "$\Rightarrow$" Let X be finite dimensional and consider a sequence $(x_n)_n \subset \bar{B}_1(0)$. Fix a basis $e_1, \ldots, e_N$ and write $x_n = \sum_{m=1}^N x_n^m e_m$. Then, since all norms on finite-dimensional spaces are equivalent we have that

$$1 \geq \|x_n\| = \|\sum_{m=1}^N x_n^m e_m\| \geq c \sup_{1\leq m\leq N} |x_n^m|$$

for some constant $c > 0$. Thus, $|x_n^m| \leq c^{-1}$. By the Heine–Borel property of $\mathbb{R}^N$, there exists a subsequence such that $x_{n_k}^m$ are all convergent. Hence, x_{n_k} is convergent and thus $\bar{B}_1(0)$ is sequentially compact and hence compact (because a normed space is a metric space).

"$\Leftarrow$" We will show that if X is infinite-dimensional, then the unit ball is not sequentially compact, i.e. not compact. To do so let $x_1 \in X$ with $\|x_1\| = 1$ and define $Y_1 = \operatorname{span}(x_1)$ which is a closed subspace of X. Now, using Riesz's lemma, choose x_2 with $\|x_2\| = 1$ and $\inf_{y\in Y_1} \|y - x_2\| \geq \frac{1}{2}$. Define $Y_2 = \operatorname{span}(x_1, x_2)$ and note that Y_2 is a closed subspace. Defining x_n and Y_n iteratively such that $\|x_n\| = 1$ and $\inf_{y\in Y_{n-1}} \|x_n - y\| \geq \frac{1}{2}$ gives us a bounded sequence which clearly has no convergent subsequence. $\square$

2.1 Linear operators

Definition 2.11. Let V, W be normed spaces and $\operatorname{dom}(T) \subset V$ a subspace. A linear map $T : \operatorname{dom}(T) \to W$ with *domain* $\operatorname{dom}(T)$ is called *bounded* if

$$\|T\|_{\operatorname{dom}(T)\to Y} = \sup_{0\neq x\in\operatorname{dom}(T)} \frac{\|Tx\|}{\|x\|} < \infty.$$

We denote the *range* of T as $\operatorname{range}(T) = \{T(v) : v \in \operatorname{dom}(T) \subset V\} \subset W$ and the *kernel* of T as $\ker(T) = \{v \in \operatorname{dom}(T) : T(v) = 0\}$. Linear maps between normed spaces are also denoted as *linear operators*.

Remark 2.12. From linear algebra we know for a linear operator T:

- $\operatorname{range}(T)$ and $\ker(T)$ are subspaces,
- $\dim \operatorname{range}(T) \leq \dim \operatorname{dom}(T)$,
- $\ker(T) = \{0\}$ if and only if there exists a linear inverse $T^{-1} : \operatorname{range}(T) \to \operatorname{dom}(T)$ such that $T^{-1}T = I_{\operatorname{dom}(T)}$ and $TT^{-1} = I_{\operatorname{range}(T)}$.

Example. Consider $V = W = C[0,1]$.

- The differentiation operator $T_D : C^1[0,1] \to C[0,1], f \mapsto f'$ with domain $C^1[0,1] \subset C[0,1]$ is a linear operator.
- The integration operator $T_I : C(\mathbb{R}) \to C(\mathbb{R}), f \mapsto \int_0^x f(y)dy$ is a linear operator.

Proposition 2.13. The integration operator T_I defined above is bounded with $\|T_I\| = 1$ but the differentiation operator defined above is unbounded.

Proof. Let T_I be the integration operator from above and $f \in C[0,1]$. Then, by the fundamental theorem of calculus, $T_I f \in C^1[0,1] \subset C[0,1]$ and moreover,

$$|(T_I f)(x)| = |\int_0^x f(y)dy| \leq \sup_{y\in[0,1]} |f(y)||x| \leq \|f\|.$$

Hence, $\|T_I\| \leq 1$. The inequality is sharp for $f(x) = 1$. Hence, $\|T_I\| = 1$.

For the differentiation operator, we consider the sequence of function $f_n(x) = \sin(nx)$ for which $\|f_n\| \leq 1$. Thus, we estimate

$$\|T_D\| \geq \|f_n'\| \geq \sup_{x\in[0,1]} |n\cos(nx)| = n \to \infty$$

as $n \to \infty$. □

A crucial property of linear operators is that boundedness is *equivalent* to continuity, which we will prove in the following.

Theorem 2.14. Let V, W be normed spaces. A linear map $T : V \supset \operatorname{dom}(T) \to W$ is continuous if and only if it is bounded.

Proof. "$\Leftarrow$": Without loss of generate $T \neq 0$. Let $\varepsilon > 0$ and $x \in \operatorname{dom}(T)$. Then for $x_0 \in \operatorname{dom}(T)$ with $\|x - x_0\| < \delta$ we have

$$\|T(x - x_0)\| \leq \|T\|\|x - x_0\| < \|T\|\delta < \varepsilon$$

for $\delta = \frac{\varepsilon}{\|T\|}$.

“$\Rightarrow$”: Assume that T is continuous at $x \in \operatorname{dom}(T)$. Then, for every $\varepsilon > 0$ there is $\delta > 0$ such that $\|x - x_0\| \le \delta$ implies $\|T(x) - T(x_0)\| \le \varepsilon$. For $y \in \operatorname{dom}(T)$ with $\|y\| = 1$ set $z = \delta y + x_0$ and note that $\|z - x_0\| \le \delta$. Now we compute

$$\|T(y)\| = \frac{1}{\delta}\|T(z - x_0)\| = \frac{1}{\delta}\|T(z) - T(x_0)\| \le \frac{\varepsilon}{\delta}.$$

This shows that $\|T\| \le \frac{\varepsilon}{\delta}$. □

Remark 2.15. In fact the above shows that continuity of T (and hence boundedness T) is equivalent to continuity of T at one point.

Definition 2.16. Let V, W be normed spaces. Define $BL(V, W)$ as the set of linear bounded maps $T : V \to W$ and denote $BL(V) = BL(V, V)$.

Proposition 2.17. Let V, W be normed spaces. Let $T : V \to W$ be linear and $\dim(V) < \infty$, then T is bounded.

Proof. This is again a consequence of the equivalence of norms in finite dimensional normed spaces. □

Theorem 2.18. $BL(V, W)$ is a normed space equipped with the operator norm $\|\cdot\|_{V\to W}$. It is a Banach space if W is a Banach space.

Proof. Clearly $BL(V, W)$ is a normed space so it remains to show that $BL(V, W)$ is complete if W is complete. To this end let $(T_n)_n \subset BL(V, W)$ be a Cauchy sequence. For fixed $x \in V$, the sequence $T_n x$ is Cauchy because $\|T_n x - T_m x\| \le \|T_n - T_m\|\|x\| \to 0$ and $(T_n)_n \subset BL(V, W)$ is Cauchy. Since W is complete we define $Tx = \lim_{n\to\infty} T_n x$ which clearly is a bounded linear operator because for any x, $\|Tx\| \le \lim_{n\in\mathbb{N}} \|T_n x\| \le \limsup_{n\to\infty} \|T_n\|\|x\| < \infty$.

Finally, in order to show that $T_n \to T$ we fix $\varepsilon > 0$. Then,

$$\sup_{\|x\|=1} \|T_n x - Tx\| = \sup_{\|x\|=1} \lim_{m\to\infty} \|T_n x - T_m x\| \le \lim_{m\to\infty} \|T_n - T_m\| < \varepsilon$$

for $n \ge N$ because $(T_n)_n$ is Cauchy. □

Corollary 2.19. Let V be a normed space. Then the continuous dual $V^* = BL(V, \mathbb{K})$ is a Banach space.

Theorem 2.20 (Bounded linear extension)**.** Let X be a normed space and Y be Banach space. Let $T : X \supset \operatorname{dom}(T) \to Y$ be a bounded linear operator. Then T has an extension $\tilde{T} : \overline{\operatorname{dom}(T)} \to Y$ (i.e. $\tilde{T}|_{\operatorname{dom}(T)} = T$) such that $\|T\| = \|\tilde{T}\|$.

Proof. Let $x \in \overline{\operatorname{dom}(T)}$ and consider a sequence $(x_n)_n \subset \operatorname{dom}(T)$ with $x_n \to x$ as $n \to \infty$. Since T is bounded, $T(x_n)$ is a Cauchy sequence in Y and thus has a limit y. Again by boundedness, this limit is independent of the approximating sequence so we define $\tilde{T}(x) = \lim_{n\to\infty} T(x_n)$ which is well-defined. It is clearly a linear extension of T with norm bounded by T. □

Remark 2.21. For general topological vector space X, the continuous dual X^* may be empty, e.g. $L^p([0, 1]), 0 < p < 1$. A central result in functional analysis is that for normed spaces, the continuous dual is always non-empty.

2.2 Hahn–Banach theorem

Theorem 2.22 (Hahn–Banach extension, real version)**.** Let V be a real vector space. Let $p : V \to \mathbb{R}$ be sublinear, i.e.

$$p(x+y) \leq p(x) + p(y) \text{ and } p(\lambda x) = \lambda p(x)$$

for all $\lambda \geq 0$ and $x \in V$. Let $W \subset V$ be a subspace of V and $f : W \to \mathbb{R}$ be a linear map such that $f(x) \leq p(x)$ for all $x \in W$. Then, there exists a linear map $f^* : V \to \mathbb{R}$ such that $f^*|_W = f$ and $f(x) \leq p(x)$ for all $x \in V$.

Proof. This is again an application of Zorn's lemma.

Step 1. We define the poset

$$\mathcal{P} = \{g : \mathrm{dom}(g) \to \mathbb{R} \text{ linear with } W \subset \mathrm{dom}(g), g|_W = f, g \leq p \text{ on } \mathrm{dom}(g)\},$$

where "$\subset$" is the restriction/extension. We note that $\mathcal{P}$ is non-empty as $f \in \mathcal{P}$. Let $\mathcal{C} \subset \mathcal{P}$ be a chain. Define the subspace $\bigcup_{g\in\mathcal{C}} \mathrm{dom}(g)$ and the map

$$g_{\mathcal{C}} : \bigcup_{g\in\mathcal{C}} \mathrm{dom}(g) \to \mathbb{R}, x \mapsto g_x(x),$$

where $g_x \in \mathcal{C}$ such that $x \in \mathrm{dom}(g_x)$. By construction, $g_{\mathcal{C}}$ is independent of the choice of g_x so it is easy to see that $g_{\mathcal{C}}$ is an upper bound of $\mathcal{C}$. Hence, by Zorn's lemma, there exists a maximal element $f^* : \mathrm{dom}(f^*) \to \mathbb{R}$ of $\mathcal{P}$.

Step 2. It remains to show that $\mathrm{dom}(f^*) = V$. For the sake of a contradiction argument, suppose that $0 \neq v \in V \setminus \mathrm{dom}(f^*)$. Define

$$h : \mathrm{span}(\mathrm{dom}(f^*), v) \to \mathbb{K}, w = x + \alpha v \mapsto f^*(x) + \alpha\lambda,$$

where any $w \in \mathrm{span}(\mathrm{dom}(f^*), v)$ is uniquely decomposed into $w = x+\alpha v$ for $x \in \mathrm{dom}(f^*)$ and some $\alpha \in \mathbb{R}$. Here $\lambda = h(v)$ which we are free to choose. We will choose λ to find an h that is dominated by p. We need to show that

$$h(x+\alpha v) = f^*(x) + \alpha\lambda \leq p(x+\alpha v) \text{ for all } x \in W, \alpha \in \mathbb{R}.$$

Since p is positively homogeneous, this is equivalent to

$$\begin{aligned} h(x+v) &= f^*(x) + \lambda \leq p(x+v) \\ h(x-v) &= f^*(x) - \lambda \leq p(x-v) \end{aligned}$$

for all $x \in W$. Thus, we will have to choose λ such that

$$f^*(y) - p(y-v) \leq \lambda \leq p(x+v) - f^*(x)$$

for all $x, y \in W$. This is always possible because

$$f^*(y) + f^*(x) = f^*(x+y) \leq p(x+y) = p(x+v+y-v) \leq p(x+v) + p(y-v)$$

for all $x, y \in W$. $\square$

For complex vector spaces, the expression $f(x) \leq p(x)$ does not make sense so we will replace this condition with $|f(x)| \leq p(x)$, where p is now assumed to be a seminorm.

Theorem 2.23 (Hahn–Banach extension, complex version)**.** Let V be a complex vector space. Let $p : V \to \mathbb{C}$ be a seminorm, i.e.

$$p(x+y) \leq p(x) + p(y) \text{ and } p(\lambda x) = |\lambda| p(x)$$

for all $\lambda \in \mathbb{C}$ and $x \in V$. Let $W \subset V$ be a subspace of V and $f : W \to \mathbb{C}$ be a linear map such that $|f(x)| \leq p(x)$ for all $x \in W$. Then, there exists a linear map $f^* : V \to \mathbb{C}$ such that $f^*|_W = f$ and $|f(x)| \leq p(x)$ for all $x \in V$.

Proof. We split f into its real and imaginary part $f = f_1 + i f_2$ and note that f_1 and f_2 are $\mathbb{R}$-linear and satisfy $f_1(ix) = -f_2(x)$, hence f is uniquely determined by its real part. We note that $f_1(x) \leq p(x)$ so by considering V as a real vector space (note that in this case, the elements x and ix are linearly independent) we can extend f_1 to f_1^* on all of V such that $f_1^* \leq p(x)$, where f_1^* is $\mathbb{R}$-linear. We now define $f^*(x) = f_1^*(x) - i f_1^*(ix)$ and note that f^* is $\mathbb{C}$-linear. To show the bound we compute

$$|f^*(x)| = e^{i\theta} f^*(x) = f^*(e^{i\theta}x) = f_1^*(e^{i\theta}x) \leq p(e^{i\theta}x) = p(x),$$

where we used that the left hand side is real. □

Corollary 2.24. Let V be a non-trivial normed space. For all $0 \neq v \in V$, there exists a $f \in V^*$ such that $\|f\| = 1$ and $f(v) = \|v\|$. In particular, V^* is non-empty.

Proof. For $0 \neq v \in V$ the linear functional $f_0 : \operatorname{span}(v) \to \mathbb{K}$ by setting $f_0(v) = \|v\|$ and extend by linearity to $\operatorname{span}(v)$. Using the Hahn–Banach theorem we can extend f_0 to the whole space. □

The above tells us that the dual is always non-empty and has a rather rich structure as it separates points. We recall the definition of X^* and we define $X^{**} = (X^*)^*$ and similarly X^{***} etc. We first note the following direct consequence of Corollary 2.24.

Corollary 2.25. Let X be a normed space and $x \in X$. Then $\|x\| = \sup_{f \in X^* : \|f\|=1} |f(x)|$.

Corollary 2.26. Let X be a normed space and U be a proper, closed subspace. Then, for any $x \in X \setminus U$, there exists a function $f \in X^*$ such that $f(x) \neq 0$ but $f|_U = 0$.

Proof. We define the quotient map $\omega : X \mapsto X/U$ with induced norm on X/U defined as $\|[v]\|_{X/U} = \inf_{u \in U} \|v - u\|$. This makes X/U a normed space. (Check and see how closedness enters.) In particular, $\omega(u) = 0$ for all $u \in U$ and $\omega(x) \neq 0$. Use Corollary 2.24 to find a function $f \in (X/U)^*$ such that $f(x) \neq 0$. Then, the map $f \circ \omega \in X^*$ has the desired properties. □

Proposition 2.27. Let X be a normed space. The canonical embedding $i : X \to X^{**}$, defined by $i(x)(f) = f(x)$ for $f \in X^*$, is a linear isometry.

Proof. Clearly, i is well-defined and linear. Moreover, it is an isometry because

$$\|i(x)\| = \sup_{f \in X^*:\|f\|=1} |i(x)(f)| = \sup_{f \in X^*:\|f\|=1} |f(x)| = \|x\|.$$

$\square$

Spaces for which the canonical embedding in its double dual is surjective have special properties. Since X^{**} is always complete, any such space has to be complete by itself.

Definition 2.28. A Banach space X is called *reflexive* if the canonical embedding $i : X \to X^{**}$ is surjective.

Example.

- Every finite dimensional space is reflexive.
- ℓ^p is reflexive for $1 < p < \infty$.
- Note that $\ell^1 \cong (c_0)^*$ and $\ell^\infty = (\ell^1)^*$ so none of these spaces are reflexive.

Proposition 2.29. Let X be a normed space and assume that X^* is separable. Then X is separable.

Proof. See problem set. $\square$

Proposition 2.30. A Banach space X is reflexive if and only if X^* is reflexive.

Proof. See problem set. $\square$

2.3 Cornerstones of functional analysis

One often regards the *Banach–Steinhaus theorem*, *open mapping theorem* and the *closed graph theorem* as the cornerstones[2] of functional analysis. An important ingredient in the following three theorems is completeness, as we will make fundamental use of the Baire category theorem for complete metric spaces.

We will begin with the proof of the Banach–Steinhaus theorem which is a result under which pointwise boundedness of a sequence of operators implies uniform boundedness.

Theorem 2.31 (Principle of uniform boundedness, Banach–Steinhaus)**.** Let X be a Banach space and Y be a normed space. Let $(T_n)_n \subset BL(X,Y)$ be a sequence of linear operators and assume that $\sup_{n\in\mathbb{N}} \|T_n(x)\| < \infty$ for all $x \in X$. Then, $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$.

Proof. Define the set $A_m = \{x \in X : \sup_{n\in\mathbb{N}} \|T_n(x)\| \le m\}$ and note that each $A_m = \cap_{n\in\mathbb{N}}\{x \in X : \|T_n(x)\| \le m\}$ is closed and $\cup_{m\in\mathbb{N}} A_m = X$. By Baire's theorem (Corollary 1.52) there exists an A_M with non-empty interior, i.e. there exists a ball $B_{2\varepsilon}(x_0) \subset A_M$ for some $\varepsilon > 0$ and $x_0 \in A_M$. Let $x \in X$ with $\|x\| \le 1$. Then,

$$\|T_n(x)\| = \varepsilon^{-1}\|T_n(\varepsilon x)\| \le \varepsilon^{-1}\|T_n(\varepsilon x - x_0)\| + \varepsilon^{-1}\|T_n(x_0)\| \le 2\varepsilon^{-1}M.$$

Thus, $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$. $\square$

[2]Some authors also include the Hahn–Banach theorem in that list.

We will now move on to the open mapping theorem. We recall that the dual concept of a continuous map is an open map. It is defined as a map such that the image of every open set is open. (Note however that this is not equivalent to a map that maps closed sets to closed sets.) It is easy to see that an open, linear map between two normed spaces is necessarily surjective (Check!).

Our next cornerstone gives a converse result to this: It shows under which condition a surjective bounded linear operator between normed spaces is an open mapping. The key ingredient is the completeness of both, the domain and the codomain.

Theorem 2.32 (Open mapping theorem)**.** Let X, Y be Banach spaces and $T : X \to Y$ be a surjective bounded linear operator. Then T is open.

Proof. Since T is linear, it suffices to show that $T(B_1^X(0))$ contains an open ball around the origin in Y.

Step 1. We will first show that $\overline{T(B_{1/2}^X(0))}$ contains an interior point. Since T is surjective, $Y = \cup_{n\in\mathbb{N}} T(B_n^X(0))$ so by the Baire category theorem, there exists a $N \in \mathbb{N}$ such that $\overline{T(B_N^X(0))}$ contains an interior point and by linearity so does $\overline{T(B_{1/2}^X(0))}$. In particular, there exists an $\varepsilon > 0$ and $y_0 \in \overline{T(B_{1/2}^X(0))}$ such that $B_\varepsilon^Y(y_0) \subset \overline{T(B_{1/2}^X(0))}$.

Step 2. We will now show that $B_\varepsilon^Y(0) \subset \overline{T(B_1^X(0))}$ which follows once we show that

$$\overline{T(B_{1/2}^X(0))} - y_0 \subset \overline{T(B_1^X(0))}.$$

Let $y \in \overline{T(B_{1/2}^X(0))} - y_0$, i.e. $y + y_0 \in \overline{T(B_{1/2}^X(0))}$. For $y + y_0$ and y_0 we can write

$$y_0 = \lim_{n\to\infty} Tx_n, \quad y + y_0 = \lim_{n\to\infty} Tz_n$$

for $\|x_n\| < \frac{1}{2}$ and $\|z_n\| < \frac{1}{2}$. Thus, $y = \lim_{n\to\infty} T(z_n - x_n)$, where $\|z_n - x_n\| < 1$ for each $n \in \mathbb{N}$. This shows that $y \in \overline{T(B_1^X(0))}$ and concludes the step.

Step 3. Using the linearity, we obtain

$$V_n \doteq B_{\varepsilon 2^{-n}}^Y(0) \subset \overline{T(B_{2^{-n}}^X(0))}. \tag{1}$$

Step 4. We will finally show that $V_1 = B_{\varepsilon/2}^Y(0) \subset T(B_1^X(0))$. Let $y \in V_1 \subset \overline{T(B_{1/2}^X(0))}$ and choose $Tx_1 \in T(B_{1/2}^X(0))$ such that $\|y - Tx_1\| \leq \varepsilon/4$. From (1) for $n = 2$ we obtain that $y - Tx_1 \in V_2 \subset \overline{T(B_{2^{-2}}^X(0))}$. As before we find $x_2 \in B_{2^{-2}}^X(0)$ such that $\|y - Tx_1 - Tx_2\| \leq \varepsilon/8$. Continuing inductively, we obtain a sequence $(x_n)_n$ with $\|x_n\| < 2^{-n}$ and for which

$$\|y - \sum_{n=1}^N Tx_n\| \leq \varepsilon 2^{-N-1}. \tag{2}$$

We note that $w_N = \sum_{n=1}^N x_n$ is a Cauchy sequence because

$$\|w_N - w_M\| \leq \sum_{n=\min(N,M)}^\infty \|x_n\| \leq \sum_{n=\min(N,M)}^\infty 2^{-n} \to 0$$

as $\min(N,M) \to \infty$. Hence, by the completeness of X we have $w_N \to w$ as $N \to \infty$. From (2) we obtain that $y = Tw$ and we are done once we show that $\|w\| < 1$. Indeed,

$$\|w\| \leq \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} 2^{-n} = 1.$$

This concludes the proof. □

Remark 2.33. Note that the completeness of X and Y were used in the proof.

Corollary 2.34 (Inverse mapping theorem)**.** Let X, Y be Banach spaces and $T : X \to Y$ be a bijective bounded linear operator. Then $T^{-1} \in BL(Y,X)$.

Proof. By the open mapping theorem T is an open map, i.e. T^{-1} is continuous. □

Corollary 2.35. Let X and Y be Banach spaces and $T \in BL(X,Y)$ be injective. Then $T^{-1} : \mathrm{range}(T) \to X$ is continuous if and only if $\mathrm{range}(T)$ is closed.

Proof. "⇒" If T^{-1} is continuous, then T is an isomorphism between X and $\mathrm{range}(T)$. Since X is complete so is $\mathrm{range}(T)$ and thus closed.

"⇐" If $\mathrm{range}(T)$ is closed it is complete so $T : X \to \mathrm{range}(T)$ is a bijective bounded linear operator between Banach spaces. By Corollary 2.34, it is a homeomorphism. □

We will now proceed to prove the last of the cornerstones, the closed graph theorem. The essence of the theorem is that closedness of the graph together with completeness of the spaces gives us regularity of the linear operator. For a map $f : X \to Y$ between sets we recall that the graph is defined as $\mathrm{graph}(f) = \{(x, f(x)) : x \in X\} \subset X \times Y$. We recall the following topological closed graph theorem from the first problem set.

Proposition 2.36. Let X and Y be compact Hausdorff spaces and $f : X \to Y$ be a map. Then, the $\mathrm{graph}(f) \subset X \times Y$ is closed if and only if f is continuous.

Proof. (See also problem set 1) "⇐": We have that $f \times \mathrm{id} : X \times Y \to Y \times Y$ is continuous. Hence, $\mathrm{graph}(f) = (f \times \mathrm{id})^{-1}(\Delta_Y)$ is closed, because the diagonal $\Delta_Y = \{(y,y) : y \in Y\}$ is closed in Y^2 because Y is Hausdorff.

"⇒" We have that $\mathrm{graph}(f) \subset X \times Y$ is a closed subset of a compact Hausdorff space. Hence, $\mathrm{graph}(f)$ is compact. In particular, the projection map $\pi_X \colon \mathrm{graph}(f) \to X$ is a bijective continuous map between compact Hausdorff spaces, hence a homeomorphism. Let now $V \subset Y$ be closed. Then, $X \times V \cap \mathrm{graph}(f)$ is closed. Thus, $\pi_X(X \times V \cap \mathrm{graph}(f)) = f^{-1}(V)$ is closed. □

Remark 2.37. The above result even holds if X is merely a topological space. Y however has to be a compact Hausdorff space.

In the functional analysis setup, we deal with the linear operators and we can relax the assumption of compactness. However, we will still need the "weaker" assumption of completeness. We will now introduce the notion of a closed operator which just means that its graph is closed.

Definition 2.38. Let V, W be normed spaces and $T : \mathrm{dom}(T) \to W$ be a linear operator with domain $\mathrm{dom}(T) \subset V$. We say that T is a *closed linear operator* if $\mathrm{graph}(T) = \{(x, Tx) : x \in \mathrm{dom}(T)\} \subset V \times W$ is closed.

For normed space V, W as above, we equip the product space $V \times W$ with the norm $\|(v, w)\|_{V\times W} = \|v\|_V + \|w\|_W$ which induces the product topology. It is easy to check that if V, W are Banach spaces, then so is $V \times W$. Note that the name *closed operator* is somewhat unfortunate as a closed operator does not have to be a closed map (which is defined as mapping closed sets to closed sets). Analogously to the topological closed graph theorem, we will now prove that closed operators which are defined on a closed domain of a Banach space are bounded.

Theorem 2.39 (Closed graph theorem)**.** Let X, Y be Banach spaces and $T : X \supset \mathrm{dom}(T) \to Y$ a closed linear operator. If $\mathrm{dom}(T)$ is closed in X, then T is bounded.

Proof. We consider again the projection $\pi_X : \mathrm{graph}(T) \to \mathrm{dom}(T)$ which maps $\pi_X(x, Tx) = x$. Clearly, π_X is a bounded linear operator between the Banach spaces $\mathrm{graph}(T)$ and $\mathrm{dom}(T)$, where we use that closed subspaces of Banach spaces are Banach spaces themselves. Moreover, π_X is bijective so by the inverse mapping theorem (Corollary 2.34) a homeomorphism. In particular, the inverse $\pi_X^{-1} : \mathrm{dom}(T) \to \mathrm{graph}(T), x \mapsto (x, Tx)$ is bounded from which we conclude that T is bounded. □

This tells us that a closed linear operator defined on a Banach space has to be continuous.

Proposition 2.40. Let V, W be normed spaces. A linear operator $T : V \supset \mathrm{dom}(T) \to W$ is closed if and only if the following implication holds:

$$(x_n)_n \subset \mathrm{dom}(T) \text{ with } x_n \to x \in V \text{ and } T(x_n) \to y \in Y \Rightarrow x \in \mathrm{dom}(T) \text{ and } y = Tx.$$

Proof. Follows from the definition of a closed linear operator. □

Remark 2.41. In the case of a $\mathrm{dom}(T) = V$ we consider the following statements:

(i) $x_n \to x$,

(ii) $Tx_n \to y$,

(iii) $Tx = y$.

We note that the T being continuous means (i) $\Rightarrow$ (ii), (iii) while T being closed means (i), (ii) $\Rightarrow$ (iii). So in the cases of Banach spaces, the closed graph theorem gives us an "easier" condition to check to determine whether an operator is bounded.

Remark 2.42. We note that bounded operators do not have to be closed and closed operators do not have to be bounded. For instance consider $X = C[0, 1]$ and $\mathrm{dom}(T) = C^1[0, 1] \subset X$. Then it is easy to see that the differentiation operator $Tf \mapsto f'$ is closed but not bounded (problem set). On the other hand, consider the identity map on a proper dense subspace $I : V \supset \mathrm{dom}(T) \to V$ which is clearly bounded and not closed if $\dim(V) = \infty$. However, this operator is the restriction of the closed operator I on the whole space. Indeed, this motivates the notion of a closure $\tilde{T}$ of an operator $T : \mathrm{dom}(T) \to X$ which is a closed operator $\tilde{T}$ such that $\tilde{T}|_{\mathrm{dom}(T)} = T$ and such that $\mathrm{graph}(\tilde{T}) = \overline{\mathrm{graph}(T)}$. While many important operators T are unbounded (e.g. differentiation operators), essentially all operators of practical relevance are closed (or closable) linear operators. The property of being closed plays an important role in the study of self-adjoint unbounded operators which appear prominently in applications to PDEs and quantum mechanics.

3 Measure theory

The goal of this section is to introduce the abstract measure spaces, but then focus mostly on the Lebesgue measure and the Lebesgue integral on $\mathbb{R}^n$. This is a generalization of the Riemann integral you encountered in 18.100. There are several reasons why we would like to have another version of integration theory beyond the Riemann integral.

1. (Generalization of intervals to measurable sets) It turns out that both the Riemann and Lebesgue integral rely on approximating the integral by simple functions, i.e. functions of the form $f_n = \sum_{i=1}^n \lambda_i 1_{A_i}$. In the case of the Riemann integral, the allowed sets for A_i are intervals which makes the f_n step functions. In this case, the length of the interval $l(A_i)$ gives us an intuitive volume of the set A_i and the (Riemann) integral of f_n is then naturally computed as $\int f_n = \sum_{i=1}^n \lambda_i l(A_i)$. One aspect of Lebesgue's approach to integration is to relax this condition on the A_i and allow A_i to be more general subsets of $\mathbb{R}$. In particular, functions of the form $1_{\mathbb{Q}}$ which are not Riemann integrable can then be integrated. The "downside" of this approach is that we will have to define a notion of length or volume for more general sets than intervals. This requires the development of measure theory, which turns out to be an incredibly rich and rewarding theory that also forms the foundation of modern probability.

2. (Completeness) If we consider the space of Riemann integrable functions on (let's say) $[0, 1]$. Let us equip the space with the norm $\int_0^1 |f(x)|dx$, where the integral is taken in the sense of Riemann. Strictly speaking, in order to make this a normed space, we actually have to identify $f \sim g$ if $\int_0^1 |f(x) - g(x)|dx = 0$ but let us ignore that for the moment. We saw already that it would be extremely useful to have a Banach spaces, not only a normed space. While we can abstractly define a completion, it will turn out that the Lebesgue integral and the Lebesgue integrable functions precisely give an explicit construction of the completion.

3. (Convergence theorems) A further disadvantage of Riemann integrals is that it has rather poor commutation properties with pointwise limits. In analysis we are often interested in showing $\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 \lim_{n\to\infty} f_n(x)dx$ for instance. In general, the pointwise limit of Riemann integrable functions might not be Riemann integrable. Even if it is Riemann integrable, the only real condition which allows

us to interchange the limit is uniform convergence. Here might lie the biggest practical advantage of the Lebesgue integral. It will turn out that pointwise limits of measurable functions are always measurable! Moreover, the famous results *dominated convergence*, *monotone convergence* and *Fatou's lemma* have turned out to be very useful in applications giving conditions under which interchanging limits and integration is allowed.

3.1 Measure spaces

Definition 3.1. A *measure space* is a triplet $(\Omega, \mathcal{F}, \mu)$ where Ω is a set, $\mathcal{F}$ is a *σ-algebra* on Ω, i.e. $\mathcal{F} \subset \mathcal{P}(\Omega)$ such that

1. $\emptyset \in \mathcal{F}$
2. $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$ (closed under complementation)
3. $A_i \in \mathcal{F}$ for $i \in \mathbb{N}$ implies $\cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$ (closed under countable unions),

and μ is a *measure*, i.e. a function $\mu : \mathcal{F} \to [0, \infty]$ such that

1. $\mu(\emptyset) = 0$
2. If $A_i \in \mathcal{F}$, $i \in \mathbb{N}$ is a collection of disjoint sets, then $\mu(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i)$.

The elements of $\mathcal{F}$ are called *measurable sets*. A set $A \in \mathcal{F}$ with $\mu(A) = 0$ is called a *null set*.

Definition 3.2. A property of points $x \in \Omega$ is said to hold *almost everywhere* if it holds for all $x \in \Omega \setminus N$, where N is a null set.

Example. Let X be a set. Set $\mathcal{F} = \mathcal{P}(X)$ and define $\mu(A)$ to be the number of elements of A if A is finite and otherwise set $\mu(A) = \infty$. Then, $(X, \mathcal{F}, \mu$ is a measure space and μ is called the counting measure. Note that the empty set $\emptyset$ is the only null set.

Proposition 3.3. Note that if μ is a measure then we have the following monotonicities:

1. Let $E_1 \subset E_2 \subset \dots$ be measurable. Then,
$$\mu(\cup_{n\in\mathbb{N}} E_n) = \lim_{n\to\infty} \mu(E_n) = \sup_n \mu(E_n).$$
2. Let $E_1 \supset E_2 \supset \dots$ be measurable and $\mu(E_n) < \infty$ for at least one n. Then,
$$\mu(\cap_{n\in\mathbb{N}} E_n) = \lim_{n\to\infty} \mu(E_n) = \inf_n \mu(E_n).$$

Proof. Monotonicity 1.) follows by observing $\cup_{n\in\mathbb{N}} E_n = \cup_{n\in\mathbb{N}} \tilde{E}_n$, where $\tilde{E}_n$ are disjoint and defined as $\tilde{E}_n = E_n \setminus E_{n-1}$ with $E_0 = \emptyset$. For 2.) we refer to the problem set. □

Definition 3.4. We say that the measure μ is *finite* if $\mu(\Omega) < \infty$ and μ is *σ-finite* if $\Omega = \cup_{n\in\mathbb{N}} A_n$, where $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.

Proposition 3.5. Let Ω be a set and $\mathcal{A} \subset \mathcal{P}(\Omega)$. There exists a unique smallest σ-algebra $\mathcal{F}_{\mathcal{A}}$ containing $\mathcal{A}$ which is constructed as the intersection of all σ-algebras containing $\mathcal{A}$. $\mathcal{F}_{\mathcal{A}}$ is called the σ-algebra *generated* by $\mathcal{A}$.

Proof. One can easily check that arbitrary intersections of σ-algebras are again σ-algebras. Thus, $\mathcal{F}_{\mathcal{A}}$ defined as the intersection of all σ-algebras containing $\mathcal{A}$, is the unique smallest σ-algebra containing $\mathcal{A}$. $\square$

Definition 3.6. Let $(X, \mathcal{T})$ be a locally compact Hausdorff space.

- The σ-algebra $\mathcal{B}$ generated by $\mathcal{T}$ is called the *Borel σ-algebra* and elements of $\mathcal{B}$ are called *Borel sets*.

- If μ is a measure defined on $\mathcal{B}$ such that $\mu(K) < \infty$ for all compact sets K, then μ is called a *Borel measure*.

3.2 Outer measures and Caratéodory's extension theorem

In the important example of $\Omega = \mathbb{R}$, it turns out that one **cannot** define a map $\mu : \mathcal{F} \to [0, \infty]$ on a σ-algebra $\mathcal{F} \subset \mathcal{P}(\mathbb{R})$ such that the following holds true.

1. $\mathcal{F} = \mathcal{P}(\mathbb{R})$,

2. $\mu((a, b)) = b - a$,

3. $\mu(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu(A_i)$ for pairwise disjoint sets $(A_i)_{i\in\mathbb{N}}$.

The demand for the above properties go back to Borel and Lebesgue who were searching for a satisfactory theory of assigning a volume to subsets of $\mathbb{R}^n$. It was G. Vitali in 1905 who proved that the above problem cannot be solved for $n = 1$. You will develop Vitali's construction in the problem set.

To give a reasonable notion of a natural measure on $\mathbb{R}$, we have to relax one of the conditions. First, since we want to construct a measure, we would like to keep property 3). Moreover, from our intuition, property 2) should also hold true. In order to construct the Lebesgue measure we will therefore have to give up on property 1) and indeed consider the σ-algebra of Lebesgue measurable sets which turns out to be a proper subset of $\mathcal{P}(\mathbb{R})$. Before we do this, we however first define an *outer measure* which keeps 1) and 2) but relaxes property 3).

Definition 3.7. Let Ω be a set. A function $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ is called an *outer measure* if

1. $\mu^*(\emptyset) = 0$,

2. $A \subset B$ implies $\mu^*(A) \leq \mu^*(B)$,

3. $\mu^*(\cup_{i\in\mathbb{N}} A_i) \le \sum_{i\in\mathbb{N}} \mu^*(A_i)$.

Definition 3.8. Let Ω be a set and μ^* be an outer measure on Ω. A set $A \subset \mathcal{P}(\Omega)$ is said to be *μ^*-measurable* if it satisfies *Carathéodory's criterion*:

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c) \text{ for all } B \in \mathcal{P}(\Omega). \tag{3}$$

Remark 3.9. The above criterion can be motivated as follows. It defines measurable sets exactly as such sets, which can be used to slice other sets respecting the additivity. In Lebesgue's original approach, he also defined an *inner measure* as $\mu_*(A) = \mu^*(B) - \mu^*(B \setminus A)$ for some $A \subset B$. Then Lebesgue defined A to be measurable if the inner and the outer measure agree, i.e. $\mu_*(A) = \mu^*(A)$. This gives $\mu^*(A \cap B) = \mu^*(A) = \mu^*(B) - \mu^*(B \setminus A)$ and can then be seen as a motivation for (3). In fact, it was the Carathéodory who introduced criterion (3) which is significantly more elegant and efficient compared to Lebesgue's original construction.

We remark already an important class of measurable sets.

Lemma 3.10. Let Ω be a set and μ^* be an outer measure. Any set A with $\mu^*(A) = 0$ or $\mu^*(A^c) = 0$ is μ^*-measurable.

Proof. Let $\mu^*(A) = 0$ and $B \subset \Omega$. Then $\mu^*(A^c \cap B) + \mu^*(A \cap B) \le \mu^*(B) + \mu^*(A) = \mu^*(B)$ so A is measurable. The proof is similar for A^c. $\square$

The following general result shows the power of Carathéodory's criterion of measurability.

Theorem 3.11 (Carathéodory's extension theorem)**.** Let Ω be a set and μ^* be an outer measure on Ω. The family $\mathcal{F} \subset \mathcal{P}(\Omega)$ of μ^*-measurable sets is a σ-algebra and $\mu = \mu^*|_{\mathcal{F}}$ is a measure on $\mathcal{F}$. Moreover, μ is *complete*, i.e. if $A \subset N \in \mathcal{F}$ and $\mu(N) = 0$, then $A \in \mathcal{F}$ with $\mu(A) = 0$.

Remark 3.12. There is even a more general version of the Carathédory extension theorem which only assumes the existence of a *pre-measure* from which an outer measure is constructed.

Proof of Theorem 3.11. In order to show that $\mathcal{F}$ is a σ-algebra, we have to check the three properties from Definition 3.1 for sets satisfying (3). Clearly $\emptyset \in \mathcal{F}$. Let $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ since $(A^c)^c = A$.

Checking the last property takes a bit more time. We will first show closedness with respect to finite unions. Let $A_1, A_2 \in \mathcal{F}$. The inequality " $\le$ " in (3) follows from the subadditivity of outer measures. Thus, it suffices to show " $\ge$ ". Let $B \subset \Omega$.

$$\begin{aligned}
\mu^*(B \cap (A_1 \cup A_2)) + \mu^*(B \cap (A_1 \cup A_2)^c)& \\
&\le \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) \\
&\quad + \mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) \\
&= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) = \mu^*(B).
\end{aligned}$$

This also shows closeness under finite intersections since $A_1 \cap A_2 = (A_1 \cup A_2)^c$.

We will now upgrade to countably many A_i. Let $A_i \in \mathcal{F}$ be a countable collection of measurable sets. We have to show that $A := \cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$. We assume without loss of generality that the A_i are pairwise disjoint. (Indeed, if not note that $A = \cup_{i=1}^{\infty} \tilde{A}_i$, where for $i \geq 2$, $\tilde{A}_i = A_i \cap \cup_{j=1}^{i-1} A_j^c \in \mathcal{F}$ because we have already shown closedness under taking the complement, as well as taking finite intersections and unions.)

We will first show that $\mu^*(B \cap \cup_{i=1}^n A_i) = \sum_{i=1}^n \mu^*(B \cap A_i)$ for all $B \subset \Omega$. We argue by induction and note that the case $n = 1$ is trivial. For the induction step we have

$$
\begin{aligned}
\mu^*(B \cap \cup_{i=1}^{n+1} A_i) &= \mu^*(B \cap \cup_{i=1}^{n+1} A_i \cap B_{n+1}) + \mu^*(B \cap (\cup_{i=1}^{n+1} A_i) \cap B_{n+1}^c) \\
&= \mu^*(B \cap B_{n+1}) + \mu^*(B \cap \cup_{i=1}^{n} A_i) = \sum_{i=1}^{n+1} \mu^*(B \cap A_i).
\end{aligned}
$$

Finally, we will show that $A \in \mathcal{F}$. We estimate

$$
\begin{aligned}
\mu^*(B) &= \mu^*(B \cap \cup_{i=1}^n A_i) + \mu^*(B \cap \cap_{i=1}^n A_i^c) \geq \mu^*(B \cap \cup_{i=1}^n A_i) + \mu^*(B \cap A^c) \\
&= \sum_{i=1}^n \mu^*(B \cap A_i) + \mu^*(B \cap A^c) \to \sum_{i=1}^\infty \mu^*(B \cap A_i) + \mu^*(B \cap A^c) \text{ as } n \to \infty.
\end{aligned}
$$

Thus,

$$
\mu^*(B) \geq \sum_{i=1}^\infty \mu^*(B \cap A_i) + \mu^*(B \cap A^c) \geq \mu^*(B \cap \cup_{i\in\mathbb{N}} A_i) + \mu^*(B \cap A^c)
$$

from which we conclude that $A = \cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$. Hence, $\mathcal{F}$ is a σ-algebra and the estimate above shows that $\mu^*(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mu^*(A_i)$ for pairwise disjoint $A_i \in \mathcal{F}$. Since $\mu^*(\emptyset) = 0$, we have that $\mu = \mu^*|_{\mathcal{F}}$ is a measure.

Finally, note that μ is complete because all null sets of μ^* are measurable in view of Lemma 3.10. □

3.3 The Lebesgue measure

Definition 3.13. A set $C \subset \mathbb{R}^n$ of the form $C = I_1 \times \ldots I_n$, where $I \subset \mathbb{R}$ are intervals, is called a *cuboid*. We define $\mathrm{vol}(C) = \Pi_{i=1}^n l(I_i)$, where $l(I)$ is the length of the interval. C is called an *open/closed* cuboid if all the intervals are open/closed, respectively.

Definition 3.14 (Lebesgue outer measure)**.** On $\mathbb{R}^n$ define the *Lebesgue outer measure* $\lambda^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ as

$$
\lambda^*(A) = \inf \left\{ \sum_{i\in\mathbb{N}} \mathrm{vol}(C_i), \text{where } (C_i)_{i\in\mathbb{N}} \text{ are open cuboids such that } A \subset \cup_{i\in\mathbb{N}} C_i \right\}.
$$

Proposition 3.15. The Lebesgue outer measure λ^* defines an outer measure. Moreover, λ^* is translation invariant and $\lambda^*(C) = \mathrm{vol}(C)$ for any cuboid C.

Proof. Clearly, λ^* is monotonic and $\lambda^*(\emptyset) = 0$. Let us give more details for the countable subadditivity property. Let $\varepsilon > 0$ and let A_i, $i \in \mathbb{N}$ be a family of subsets. For $i \in \mathbb{N}$, there exists a sequence of open cuboids $C_{i,n}$ such that $\lambda^*(A_i) \geq \sum_{n\in\mathbb{N}} \mathrm{vol}(C_{i,n}) - \varepsilon 2^{-i}$ and $A_i \subset \cup_{n\in\mathbb{N}} C_{i,n}$. Note that $\cup_{i\in\mathbb{N}} A_i \subset \cup_{i\in\mathbb{N}} \cup_{n\in\mathbb{N}} C_{i,n}$ and hence,

$$\lambda^*(\cup_{i\in\mathbb{N}} A_i) \leq \sum_{i\in\mathbb{N}} \sum_{n\in\mathbb{N}} \mathrm{vol}(C_{i,n}) \leq \sum_{i\in\mathbb{N}} \left(\lambda^*(A_i) + \varepsilon 2^{-i}\right) \leq \sum_{i\in\mathbb{N}} \lambda^*(A_i) + \varepsilon.$$

Since $\varepsilon > 0$, this shows the first claim.

Clearly, λ^* is translation invariant by translating cuboids. Let us now show the last property and only consider the case for $n = 1$. We first observe that for an open interval $\mathring{I}$ we have $\lambda^*(\mathring{I}) \leq \mathrm{vol}(\mathring{I})$.

We will now show that $\mathrm{vol}(I) \leq \lambda^*(I)$ for a closed interval $I = [a, b]$. Assume that $(I_i)_{i\in\mathbb{N}}$ is an open cover of I with open intervals. Then, there exists a finite subcover, i.e. finitely many I_i which cover $I = [a, b]$. Clearly, $b - a \leq \sum_{i=1}^{n} \mathrm{vol}(I_i)$ and hence $\mathrm{vol}(I) \leq \lambda^*(I)$. For the other direction, we consider an open interval $J_n = (a - 1/n, b + 1/n)$. Then, $\lambda^*(I) \leq \lambda^*(J_n) = b - a + 2/n \to b - a$ as $n \to \infty$. For an open interval $\mathring{I}$ it still remains to show that $\mathrm{vol}(\mathring{I}) \leq \lambda^*(\mathring{I})$. However, this follows by noting that $[a + 1/n, b - 1/n] \subset (a, b)$ and performing a similar argument as before. The cases of a half-open interval and higher dimensions work analogously. We will omit these steps here. $\square$

Definition 3.16. The family $\mathcal{L} \subset \mathcal{P}(\mathbb{R}^n)$ of λ^*-measurable sets of $\mathbb{R}^n$ is called the Lebesgue σ-algebra and $\lambda = \lambda^*|_{\mathcal{L}}$ is called the *Lebesgue measure* on $\mathcal{L}$.

If not specified otherwise, we equip $\mathbb{R}^n$ and $\mathbb{C}^n$ (seen as $\mathbb{R}^{2n}$) with the Lebesgue measure.

Proposition 3.17. All open and closed sets of $\mathbb{R}^n$ are measurable.

Proof. We will restrict to the case $n = 1$; the cases $n > 1$ work analogously. We will first show that sets of the form $A = (a, \infty)$ for $a \in \mathbb{R}$ are measurable. Let $B \subset \mathbb{R}$ and $\varepsilon > 0$. Let I_n be a collection of open intervals such that $B \subset \cup_{n\in\mathbb{N}} I_n$ and $\lambda^*(B) \geq \sum_{n\in\mathbb{N}} \mathrm{vol}(I_n) - \varepsilon$. Since I_n are open intervals, we have that $I_n \cap A$ is an open interval for every $n \in \mathbb{N}$. Further, for $n \in \mathbb{N}$, note that $I_n \cap A^c \subset I_n \cap (-\infty, a + \varepsilon 2^{-n})$.

Since $B \cap A \subset \cup_{n\in\mathbb{N}} I_n \cap A$ and $B \cap A^c \subset \cup_{n\in\mathbb{N}} I_n \cap (-\infty, a + \varepsilon 2^{-n})$, we estimate

$$\begin{aligned}
\lambda^*(B \cap A) + \lambda^*(B \cap A^c) &\leq \lambda^*(\cup_{n\in\mathbb{N}} I_n \cap A) + \lambda^*\left(\cup_{n\in\mathbb{N}} I_n \cap (-\infty, a + \varepsilon 2^{-n})\right) \\
&\leq \sum_{n\in\mathbb{N}} \lambda^*(I_n \cap A) + \sum_{n\in\mathbb{N}} \lambda^*\left(I_n \cap (-\infty, a + \varepsilon 2^{-n})\right) \\
&= \sum_{n\in\mathbb{N}} \left[\mathrm{vol}(I_n \cap A) + \mathrm{vol}(I_n \cap A^c) + \mathrm{vol}((a, a + \varepsilon 2^{-n}))\right] \\
&= \sum_{n\in\mathbb{N}} \left[\mathrm{vol}(I_n) + \varepsilon 2^{-n})\right] \leq \lambda^*(B) + 2\varepsilon,
\end{aligned}$$

where we have used that for intervals I we have $\lambda^*(I) = \mathrm{vol}(I)$. Since ε was arbitrary, we conclude that sets of the form (a, ∞) are measurable. By taking approximations, intersections and complements, all intervals are measurable. Since all open sets can be written as a countable union of intervals, we conclude that all open sets are measurable. The case $n > 1$ is similar. $\square$

Proposition 3.18. The Lebesgue measure λ on $\mathbb{R}^n$ has the following properties.

- If C be a cuboid, then $\lambda(C) = \mathrm{vol}(C)$.
- It is translation invariant, i.e. $\lambda(A) = \lambda(A + x)$ for all $A \in \mathcal{L}$ and $x \in \mathbb{R}^n$.

Proof. The first statement follows from Proposition 3.15 and the fact that any cuboid is measurable. Since the Lebesgue outer measure is translation-invariant, this property also holds for the Lebesgue measure. $\square$

Theorem 3.19. Let $A \subset \mathbb{R}^n$ be a set. Then A is Lebesgue measurable if and only if the following two statements hold:

1. For every $\varepsilon > 0$ there exist an open set U with $A \subset U$ such that $\lambda^*(U \setminus A) < \varepsilon$.
2. For every $\varepsilon > 0$ there exist a closed set C with $C \subset A$ such that $\lambda^*(A \setminus C) < \varepsilon$.

Remark 3.20. Note that this implies that A is measurable if and only if every $\varepsilon > 0$ there exists an open U and closed sets C such that $C \subset A \subset U$ such that $\lambda(U \setminus C) \leq \varepsilon$.

Proof. "$\Rightarrow$": Let A be Lebesgue measurable. Assume first that $\lambda^*(A) < \infty$. By definition of λ^* there exists an open set $U := \cup_{i\in\mathbb{N}} U_i \supset A$ such that

$$\lambda^*(U) = \lambda^*(\cup_{i\in\mathbb{N}} U_i) \leq \sum_{i\in\mathbb{N}} \lambda^*(U_i) \leq \lambda^*(A) + \varepsilon/2$$

and thus

$$\lambda^*(U \setminus A) = \lambda^*(U) - \lambda^*(A) \leq \varepsilon/2,$$

where we used that A is measurable. If $\lambda^*(A) = \infty$, then consider disjoints sequence $(A_i)_i$ of measurable sets with $A = \cup_{i\in\mathbb{N}} A_i$ and $\lambda^*(A_i) < \infty$, e.g. $A_i = A \cap ([i-1, i) \cup (-i, -i+1])$ and similar in $\mathbb{R}^n$. Apply the above argument to each A_i and obtain open sets U_i with $\lambda^*(U_i \setminus A_i) < \varepsilon 2^{-i}$. Then, set $U = \cup_{i\in\mathbb{N}} U_i$ which satisfies $U \setminus A = \cup_{i\in\mathbb{N}}(U_i \setminus A) \subset \cup_{i\in\mathbb{N}}(U_i \setminus A_i)$. Hence,

$$\lambda^*(U \setminus A) \leq \lambda^*(\cup_{i\in\mathbb{N}} U_i \setminus A_i) = \sum_{i\in\mathbb{N}} \lambda^*(U_i \setminus A_i) < \sum_{i\in\mathbb{N}} \varepsilon 2^{-i} = \varepsilon.$$

For 2.) do the same argument but with A replaced by A^c and note that $\tilde{U} \setminus A^c = A \setminus \tilde{U}^c$, where $\tilde{U}^c$ is closed if $\tilde{U}$ was open.

"$\Leftarrow$": By 1.) and 2.), choose for each n an open U_n and a closed C_n such that $\lambda^*(U_n \setminus C_n) < 1/n$ and define $U = \cap_{n\in\mathbb{N}} U_n$, $C = \cup_{n\in\mathbb{N}} C_n$ which are measurable. Then,

$$\lambda^*(U \setminus C) \leq \lambda^*(U_n \setminus C_n) \leq 1/n$$

so $\lambda^*(U \setminus C) = 0$. Now $A = C \cup (A \setminus C)$ is measurable because C is measurable and $A \setminus C$ is measurable as a null set as $\lambda^*(A \setminus C) \leq \lambda^*(U \setminus C) = 0$.

$\square$

Definition 3.21. A Borel measure μ is called

- *inner regular* if $\mu(A) = \sup_{K \subset A: K \text{ compact}} \mu(K)$ for all $A \in \mathcal{A}$,
- *outer regular* if $\mu(A) = \inf_{O \supset A: O \text{ open}} \mu(K)$ for all $A \in \mathcal{A}$,
- *regular* if μ is both inner and outer regular.

Theorem 3.22. The Lebesgue measure, restricted to the Borel sets, is a regular Borel measure.

Proof. We have shown above that λ is outer regular. The inner regularity of λ is left as an exercise. □

Proposition 3.23. A set is Lebesgue measurable if and only if it can be written as a union of a Borel set and a Lebesgue null set.

This means that the Lebesgue σ-algebra is the *completion* of the Borel σ-algebra.

Proposition 3.24. There exist non-measurable sets in $\mathbb{R}$.

Proof. Problem set. □

Remark 3.25. While the above shows that not all sets are measurable, "most" sets which one encounters in practice are however measurable. Indeed, as a rule of thumb, any set whose existence does not rely on the axiom of choice (or some equivalent formulation thereof) is measurable.

Remark 3.26. The famous Banach–Tarski paradox states that the unit ball in $\mathbb{R}^3$ can be decomposed into finitely (as few as five) many subsets which, after reassembling them via translation and rotation, form two disjoint copies of the unit ball. The sets involved cannot be measurable and the construction relies on the axiom of choice.

3.4 The Hausdorff measure

Another interesting measure which can be seen as a generalization of the Lebesgue measure to fractional dimensions is the *Hausdorff measure*. For a set $A \subset \mathbb{R}^n$, we will use the shorthand notation for its diameter $d(A) = \sup_{a,b \in A} |a - b|$.

Definition 3.27 (Hausdorff measure)**.** Let $A \subset \mathbb{R}^n$. Then, the s-dimensional Hausdorff outer measure $\mathcal{H}_s(A)$ is defined as

$$\mathcal{H}_s(A) := \lim_{\delta \to 0} \mathcal{H}_{s,\delta}, \tag{4}$$

where

$$\mathcal{H}_{s,\delta}(A) := \inf\{\sum_{i=1}^{\infty} (d(U_i))^s : d(U_i) \leq \delta, A \subset \cup_{i \in \mathbb{N}} U_i\}. \tag{5}$$

The Hausdorff measure which we denote again by $\mathcal{H}_s$ is then the restriction of the outer measure to $\mathcal{H}_s$-measurable sets.

We note that $\mathcal{H}_s$ scales like a s-dimensional volume $\mathcal{H}_s(\alpha A) = \alpha^s \mathcal{H}_s(A)$ and it agrees (up to a constant) with the Lebesgue measure on $\mathbb{R}^n$ if $s = n$.

Remark 3.28. Even though for each $\delta > 0$, $\mathcal{H}_{s,\delta}$ is an outer measure it is not very well behaved because the corresponding measure has few measurable sets, e.g. $[0,1]$ is not measurable for $s = 1/2$ and $\delta > 0$ for instance.

It is not too difficult to show that for each $A \subset \mathbb{R}^n$, the function $s \mapsto H_s(A)$ is increasing and that there exists a unique $s_0 \geq 0$ such that

$$\mathcal{H}_s(A) = \begin{cases} \infty & \text{if } s < s_0 \\ 0 & \text{if } s > s_0. \end{cases}$$

This observation leads to the following definition of the *Hausdorff dimension*

$$\dim_{\mathcal{H}}(A) = \inf\{s \geq 0 : \mathcal{H}_s(A) = 0\}.$$

Remark 3.29. The Hausdorff measure is of particular relevance in the study of fractals, i.e. sets which are defined through some self-similar procedure. In the problem set, you will encounter other sets arising in the subject of Diophantine approximation (i.e. the subject of approximating real numbers by rational numbers) which has a fractional Hausdorff dimension. We will end our discussion of the Hausdorff measure here.

3.5 Measurable functions

To motivate the definition of a measurable function, let us briefly recall the theory of Riemann integration. For a function f (let us say on $[0,1]$), one partitions the integration domain $[0,1]$ into subintervals $I_i = (x_i, x_{i+1})$ and then assigns upper and lower Riemann sums as $L_R = \sum_{i=1}^n (x_{i+1} - x_i) \inf_{I_i} f(x)$ and the upper Riemann sum as $U_R = \sum_{i=1}^n (x_{i+1} - x_i) \sup_{I_i} f$. To define the Riemann integral, one then takes the limit of the upper and lower sum as the partition gets finer. If this limits agree and are independent of the partition, it is called the Riemann integral of f and f is said to be Riemann integrable.

The idea of Lebesgue was different. Instead of decomposing the domain (the x-axis), he proposed to decompose the codomain (the y-axis). Let us say that the function satisfies $0 \leq f \leq M$. The idea of Lebesgue is to consider a partition of the codomain by considering intervals $0 = y_1 < y_2 < \cdots < y_{n+1} = M$ and then defining the lower and upper Lebesgue sum as $L_L = \sum_{i=1}^n y_i \lambda(\{x : y_i \leq f(x) < y_{i+1}\})$ and the upper Lebesgue sum as $U_L = \sum_{i=1}^n y_{i+1} \lambda(\{x : y_i \leq f(x) < y_{i+1}\})$. This then naturally leads to the requirement that f should satisfy that the set $f^{-1}(I)$ is Lebesgue measurable for intervals I. This motivates Definition 3.30, the concept of a *measurable function*.

Notation. We will make use of the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Note that $\bar{\mathbb{R}}$ has a natural topology by defining $O \subset \bar{\mathbb{R}}$ to be open if $O \cap \mathbb{R}$ is open in $\mathbb{R}$. Similar considerations apply to the extended non-negative real axis $[0, \infty]$. We also declare $\pm\infty + x = x \pm \infty = \pm\infty$ for all $x \in \mathbb{R}$, $x \cdot (\pm\infty) = (\pm\infty) \cdot x = \pm\infty$ for $x \in (0, +\infty]$, $\pm\infty \pm \infty = \pm\infty$ but $+\infty - \infty$ is not well-defined. Further $0 \cdot (\pm\infty) = \pm\infty \cdot 0 = 0$. Note that $\bar{\mathbb{R}}$ is a compact topological space. We note that any monotone increasing sequence in $\bar{\mathbb{R}}$ has a limit. We will use this convention only in the chapter on measure theory and integration.

In the following, we will only consider (extended) real-valued functions. The case of complex-valued functions is analogous by considering the real and imaginary part independently.

Throughout this section, $(\Omega, \mathcal{F}, \mu)$ denotes a measure space. We note that the Borel σ-algebra induced on $\bar{\mathbb{R}}$ is denoted by $\bar{\mathcal{B}}$.

Definition 3.30. A function $f : \Omega \to \bar{\mathbb{R}}$ is *measurable* if the preimage of any Borel set in $\bar{\mathbb{R}}$ is measurable.

Proposition 3.31. Let $f : \Omega \to \bar{\mathbb{R}}$ be a function. Then, the following are equivalent.

1. f is measurable.
2. $f^{-1}((\alpha, \infty])$ is measurable for every $\alpha \in \mathbb{R}$.
3. $f^{-1}([\alpha, \infty])$ is measurable for every $\alpha \in \mathbb{R}$.
4. $f^{-1}([-\infty, \alpha))$ is measurable for every $\alpha \in \mathbb{R}$.
5. $f^{-1}([-\infty, \alpha])$ is measurable for every $\alpha \in \mathbb{R}$.

Moreover, for each $\alpha \in \bar{\mathbb{R}}$, the set $\{x \in \Omega : f(x) = \alpha\}$ is measurable.

Proof. The proof is as in Proposition 3.17 using suitable intersections and approximations by intervals as well as $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$ as well as $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$. $\square$

Remark 3.32. More generally, one often defines measurable functions between measure spaces as functions such that preimages of measurable sets are measurable. However, note there exist continuous functions $f : \mathbb{R} \to \mathbb{R}$ which are not Lebesgue measurable if the codomain $\mathbb{R}$ is equipped with the Lebesgue σ-algebra.

Proposition 3.33. Let Ω carry a topology and assume that all Borel sets of Ω are measurable. Then, every continuous function $f : \Omega \to \bar{\mathbb{R}}$ is measurable.

Proof. It suffices to check that $f^{-1}((\alpha, \infty])$ is measurable. Since $(\alpha, \infty]$ is open, its preimage under f is open (and in particular Borel) and hence a measurable set. $\square$

Theorem 3.34. Let f and g be measurable and finite almost everywhere and $a, b \in \mathbb{R}$, then $af + bg$, fg, $|f|$, and $\min(f, g), \max(f, g)$ are measurable. Moreover, if $(f_n)_n$ is a sequence of real-valued measurable functions, then $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$, $\liminf_{n\to\infty} f_n$, $\limsup_{n\to\infty} f_n$ are measurable. In particular, if $f_n \to f$ almost everywhere, then f is measurable.

Proof. Let $a > 0$ and $f : \Omega \to \overline{\mathbb{R}}$ be measurable. Clearly af is measurable because $(af)^{-1}((\alpha, \infty)) = f^{-1}((\alpha/a, \infty))$. An analogous argument shows it for $a < 0$ or $a = 0$. For $f+g$ we note that $\{f+g > \alpha\} = \cup_{q\in\mathbb{Q}}\{f > q\}\cap\{g > \alpha-q\}$. Indeed, if $f(x)+g(x) > \alpha$, then $f(x) > \alpha - g(x)$, i.e. there exists a $q \in \mathbb{Q}$ such that $f(x) > q > \alpha - g(x)$. The other direction is immediate.

We now note that f^2 is measurable because $\{f^2 > \alpha\} = \{f > \sqrt{\alpha}\} \cup \{f < -\sqrt{\alpha}\}$ for $\alpha \geq 0$.

For fg we note that $fg = 1/4((f+g)^2 - (f-g)^2)$. A similar argument applies to $|f|$.

Let $(f_n)_n$ be real-valued and measurable. We note that $\{\sup_{n\in\mathbb{N}} f_n > \alpha\} = \cup_{n\in\mathbb{N}}\{f_n > \alpha\}$ which shows that the sup is measurable. A similar argument also applied to inf and thus $\limsup_{n\to\infty} f_n = \inf_{n\geq N}\sup_{j\geq N} f_j$, $\liminf_{n\to\infty} f_n = \sup_{n\geq N}\inf_{j\geq N} f_j$ are measurable. □

Definition 3.35. A *simple function* $f : \Omega \to \mathbb{R}$ is a measurable function which only takes finite values. It can be canonically written in the form $f = \sum_{i=1}^n \lambda_i \mathbf{1}_{A_i}$, where $\lambda_i \in \mathbb{R}$ and A_i are measurable and A_i are pairwise disjoint and λ_i are pairwise different.

The following lemma is a fundamental ingredient in the theory of Lebesgue integration. It will then allow us to define the Lebesgue integral.

Lemma 3.36. A non-negative function f is measurable if and only if there exists an increasing sequence of simple functions $0 \leq f_1 \leq f_2 \leq \dots$ such that $\sup_{n\in\mathbb{N}} f_n(x) = f(x)$ for all $x \in \Omega$.

Proof. If f is the limit of simple functions, then it is measurable by Theorem 3.34.

Let us now prove the converse and assume f is non-negative and measurable. We define the sets

$$A_{j,n} = \begin{cases} \{j2^{-n} \leq f < (j+1)2^{-n}\} & \text{if } 0 \leq j \leq n2^n - 1 \\ \{f \geq n\} & \text{if } j = n2^n. \end{cases}$$

For fixed n, the sets $A_{j,n}$ are measurable, disjoint and their union over $j \in \{0, 1, \dots, n2^n\}$ is all of Ω. Thus,

$$f_n = \sum_{j=0}^{n2^n} \frac{j}{2^n}\mathbf{1}_{A_{j,n}}$$

is a sequence of increasing simple functions as indeed $A_{j,n}$ is the disjoint union of $A_{2j,n+1}$ and $A_{2j+1,n+1}$ for $j \leq n2^n - 1$ and $A_{n2^n,n}$ is the disjoint union of of the sets $A_{j,n+1}$, $j = n2^{n+1}, \dots, (n+1)2^{n+1} - 1$ and $A_{(n+1)2^{n+1},n+1}$.

If $f(x) = \infty$ at some point $x \in \Omega$, then $f_n(x) = n \to \infty = f(x)$. Moreover, by construction

$$f_n(x) \leq f(x) < f_n(x) + 2^{-n}$$

if $f(x) < n$. In particular, $\lim_{n\to\infty} f_n(x) = \sup_{n\in\mathbb{N}} f_n(x) = f(x)$. □

Remark 3.37. From the above proof we note that if f is a bounded non-negative measurable function, then the constructed f_n above converges uniformly to f.

3.6 Integration

It turns out that the monotone approximation of measurable functions by simple functions is the key towards a satisfactory definition of the Lebesgue integral of a measurable function. We will indeed first define the Lebesgue integral for simple functions and then extend (using the monotone approximation) this definition to measurable functions.

Definition 3.38. Let $f = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}$ be a non-negative simple function. Then, we define

$$\text{Simp} \int_{\Omega} f d\mu = \sum_{i=1}^{n} \lambda_i \mu(A_i) \tag{6}$$

which is a number in $[0, \infty]$.

Remark 3.39. Note that Definition 3.38 is well-defined and independent of the representation of f as a simple function. In particular, f does not have to be written its canonical form. This is left as an exercise for the reader. We also observe that (6) is linear and monotone.

Definition 3.40. Let f be a non-negative measurable function. We define

$$\int_{\Omega} f d\mu = \lim_{n\to\infty} \text{Simp} \int_{\Omega} f_n d\mu,$$

where f_n is a monotonically increasing sequence of simple functions with $f_n \to f$.

In order for the previous definition to be well-defined, we have to show that it is indeed independent of the choice of sequence of simple functions. To this end we will first show the following proposition.

Lemma 3.41. Let f be a non-negative simple function and g_n be an increasing sequence of simple functions such that $f \leq \lim_{n\to\infty} g_n$, then

$$\text{Simp} \int_{\Omega} f d\mu \leq \lim_{n\to\infty} \text{Simp} \int_{\Omega} g_n d\mu.$$

Proof. Let $f = \sum_{j=1}^{m} \lambda_j \mathbf{1}_{A_j}$ be written in its canonical form. Let $\varepsilon > 0$ and define the sets

$$B_n = \{(1+\varepsilon) g_n \geq f\}.$$

If $f(x) = 0$ then, $x \in B_n$ for all $n \in \mathbb{N}$ and if $f(x) > 0$ then, $x \in B_n$ for all n sufficiently large. Since g_n is increasing so is B_n and $\cup_{n\in\mathbb{N}} B_n = \Omega$. Thus, by upwards monotonicity

$$\begin{aligned}
\text{Simp} \int_{\Omega} f d\mu &= \sum_{j=1}^{m} \lambda_j \mu(A_j) = \lim_{n\to\infty} \sum_{j=1}^{m} \lambda_j \mu(A_j \cap B_n) = \lim_{n\to\infty} \text{Simp} \int_{\Omega} \sum_{j=1}^{m} \lambda_j \mathbf{1}_{A_j \cap B_n} d\mu \\
&= \lim_{n\to\infty} \int_{\Omega} f \mathbf{1}_{B_n} d\mu \leq \lim_{n\to\infty} \text{Simp} \int_{\Omega} (1+\varepsilon) g_n \mathbf{1}_{B_n} d\mu \\
&\leq (1+\varepsilon) \lim_{n\to\infty} \text{Simp} \int_{\Omega} g_n d\mu,
\end{aligned}$$

where we have used the linearity of Simp $\int_\Omega$ and the monotonicity

$$\text{Simp}\int_\Omega s_1 d\mu \leq \text{Simp}\int_\Omega s_2 d\mu$$

for simple functions s_1, s_2 with $s_1 \leq s_2$. These properties follow directly from the definition. Taking $\varepsilon \to 0$ gives the desired result. $\square$

We obtain the following corollary which establishes that Definition 3.40 is well-defined.

Corollary 3.42. Let f_n and g_n be increasing sequences of simple functions and $\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n$. Then,

$$\lim_{n\to\infty} \text{Simp}\int f_n d\mu = \lim_{n\to\infty} \text{Simp}\int g_n d\mu.$$

Proof. For some fixed n_0 note that $f_{n_0} \leq \lim_{n\to\infty} g_n$. Thus,

$$\text{Simp}\int f_{n_0} d\mu \leq \lim_{n\to\infty} \text{Simp}\int g_n d\mu.$$

Now taking the limit $n_0 \to \infty$ gives "$\leq$". By reversing the roles of f_n and g_n we obtain "$\geq$". $\square$

From the corollary above we deduce that the expression of the integral in Definition 3.40 is well-defined and that it agrees with the definition Simp $\int$ for simple functions.

Proposition 3.43. Let f, g be non-negative measurable functions and $\alpha, \beta \geq 0$. Then

$$\int_\Omega \alpha f + \beta g d\mu = \alpha\int_\Omega f d\mu + \beta\int_\Omega g d\mu$$

and if $f \leq g$ then

$$\int_\Omega f d\mu \leq \int_\Omega g d\mu.$$

Proof. This follows directly from the definition. $\square$

Proposition 3.44. If f is a non-negative measurable function. Then, $f = 0$ almost everywhere if and only if $\int_\Omega f d\mu = 0$.

Proof. If $\{f > 0\}$ is a null set, then $f \leq g = \infty \cdot \mathbf{1}_{\{f>0\}}$ which however integrates to zero. By monotonicity, $\int_\Omega f d\mu = 0$. For the other direction let $A = \{f > 0\}$ and write $A_n = \{f > 1/n\}$ and note that $A_n \to A$ as $n \to \infty$, where the limit is increasing. Since $n^{-1}\mathbf{1}_{A_n} \leq f$ we have

$$0 \leq \frac{1}{n}\mu(A_n) = \int \frac{1}{n}\mathbf{1}_{A_n} d\mu \leq \int_\Omega f d\mu = 0.$$

Since $A_n \to A$ we have by Proposition 3.3 that $\mu(A) = 0$. $\square$

In particular, this means that if two non-negative measurable functions f, g agree almost everywhere, their integrals agree. We will now show our first fundamental convergence theorem in the Lebesgue integration theory.

Theorem 3.45 (Monotone convergence)**.** Let $(f_n)_n$ be an increasing sequence of non-negative measurable functions and define $f = \lim_{n\to\infty} f_n$. Then,

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Proof. Since f_n are monotone, we obtain

$$\lim_{n\to\infty} \int f_n d\mu \le \int f d\mu$$

and it remains to show the reverse inequality. For the reverse inequality we let $s \le f$ be a simple function and $\varepsilon > 0$. Consider the set $B_n = \{(1-\varepsilon)s \le f_n\}$. Clearly, B_n are measurable, monotone and satisfy $\Omega = \cup_{n\in\mathbb{N}} B_n$ and thus $s\mathbf{1}_{B_n} \to s$ monotonically. But then using Lemma 3.41 we obtain

$$(1-\varepsilon)\int_\Omega s d\mu \le (1-\varepsilon)\lim_{n\to\infty}\int_\Omega s\mathbf{1}_{B_n} d\mu \le \lim_{n\to\infty}\int_\Omega f_n d\mu.$$

Since $\varepsilon > 0$ was arbitrary and the above holds for all simple function $s \le f$ we obtain the claim. □

We will now extend our integral to functions which can also take negative values. In order to obtain a meaningful theory we will have to limit ourself to integrable functions.

Definition 3.46. A function $f : \Omega \to \overline{\mathbb{R}}$ is said to be *integrable* if it is measurable and

$$\int_\Omega |f| d\mu < \infty.$$

In this case, we define

$$\int_\Omega f d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu,$$

where $f = f^+ - f^-$, $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$.

Proposition 3.47. Let $f : \Omega \to \overline{\mathbb{R}}$ be integrable. Then $\{|f| = \infty\}$ is a null set.

Proof. Set $A_n = \{|f| \ge n\}$ and note that $A := \{|f| = \infty\} \subset A_n$. Moreover, $n\mathbf{1}_{A_n} \le |f|$. So if $|f|$ is integrable then $n\mu(A) \le n\mu(A_n) = \int_\Omega n\mathbf{1}_{A_n} d\mu \le \int_\Omega |f| d\mu < \infty$. Hence, $\mu(A) = 0$. □

Proposition 3.48. Let f, g be integrable and $\alpha_1, \alpha_2 \in \mathbb{R}$. Then, the function $\alpha_1 f + \alpha_2 g$ is integrable and

$$\int_\Omega \alpha_1 f d\mu + \alpha_2 g \, d\mu = \alpha_1 \int_\Omega f d\mu + \alpha_2 \int_\Omega g d\mu.$$

Moreover, if $f \leq g$. Then

$$\int f d\mu \leq \int g d\mu$$

and in particular,

$$\left| \int_\Omega f d\mu \right| \leq \int_\Omega |f| d\mu.$$

Proof. This follows by decomposing $f = f^+ - f^-$ and using the linearity for non-negative measurable functions. This is left as an exercise. □

Remark 3.49. One can verify that the statements of monotone convergence Theorem 3.45 and Proposition 3.48 remain true if the respective assumptions only hold almost everywhere. Indeed, any construction invling

We have already encountered the monotone convergence theorem for increasing sequences of functions. For sequences that fail to be increasing, we still have the following result.

Lemma 3.50 (Fatou's lemma)**.** Let $(f_n)_n$ be a sequence of non-negative measurable functions. Then,

$$\int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Proof. Define $f = \liminf_{n\to\infty} f_n$ and $g_n = \inf_{k\geq n} f_k$ so g_n is increasing and $g_n \leq f_n$. Moreover, $f = \lim_{n\to\infty} g_n$. Thus, using that $g_n \leq g_{n+1}$ and monotone convergence, we obtain

$$\int f d\mu = \int \lim_{n\to\infty} g_n d\mu = \lim_{n\to\infty} \int g_n d\mu = \liminf_{n\to\infty} \int g_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

□

The following theorem on dominated convergence and the monotone convergence theorem are the most used convergence theorems. The theorem is quite general because it allows for $\mu(X) = \infty$ and moreover only requires $\lim_{n\to\infty} f_n(x) = f(x)$ pointwise (or pointwise a.e.).

Theorem 3.51 (Dominated convergence)**.** Let $(f_n)_n$ and f be measurable functions such that $\lim_{n\to\infty} f_n = f$ almost everywhere. Let g be a non-negative integrable function and assume that $0 \leq |f_n| \leq g$ almost everywhere. Then, all f_n and f are integrable and

$$\int_\Omega |f_n - f| d\mu \to 0 \text{ and consequently } \int_\Omega f_n d\mu \to \int_\Omega f d\mu.$$

Proof. By modifying f_n and f on a set of measure zero we may assume that $f_n \to f$ and $|f_n| \leq g$ pointwise everywhere. In particular, $0 \leq |f_n - f| \leq 2g$ such that the function $2g - |f_n - f|$ is non-negative. By Fatou's lemma we have

$$\liminf_{n\to\infty} \int -|f_n - f| + 2g d\mu \geq \int 2g d\mu$$

from which we conclude (using that $\int g d\mu < \infty$) and linearity that

$$\limsup_{n\to\infty} \int |f_n - f| d\mu = 0.$$

The second statement follows from the monotonicity of the integral. □

We note that a Riemann integrable function f on $\mathbb{R}^n$ is Lebesgue integrable (and in particular measurable) as the approximation from below in the Riemann lower sum constitutes an increasing sequence of simple functions converging to f. In particular, all results for the Riemann integral also extend to the Lebesgue integral. In fact the following result holds true

Theorem 3.52. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if the set of points of discontinuity is a null set. Moreover, in this case, the Riemann and Lebesgue integral agree.

Remark 3.53. Note however that there exists *improper* Riemann integrable functions like $\frac{\sin(x)}{x}$ on $[0, \infty)$ which are Riemann integrable as an improper integral but not Lebesgue integrable.

The example below is an extension of the fundamental theorem of calculus where we only assume that f' is differentiable and not necessarily continuously differentiable. Note that f' is not necessarily Riemann integrable.

Example. Let $f : [a, b] \to \mathbb{R}$ be differentiable and f' be uniformly bounded. Then, f' is Lebesgue-integrable and

$$\int_a^b f' d\lambda = f(b) - f(a).$$

Proof. Without loss of generally let $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $|f'| \leq M$ on $\mathbb{R}$. Set $g_n(x) = n(f(x + 1/n) - f(x))$ and note that $g_n(x) \to f'(x)$ pointwise by assumption. Hence, f' is measurable and since $|f'| \leq M$ it is integrable. From the mean value theorem, we know that $|g_n| \leq M$. Hence, using the dominated convergence theorem we conclude

$$\int_a^b f'(x) dx = \lim_{n\to\infty} \int_a^b g_n(x) dx.$$

Since f is continuous we have that $F(x) = \int_a^x f(t) dt$ is differentiable with $F' = f$ (Exercise!) and

$$\begin{aligned}\int_a^b g_n(x) dx &= n \int_a^b f(x + 1/n) - f(x) dx \\ &= n(F(b + 1/n) - F(b)) - n(F(a + 1/n) - F(a)) \to f(b) - f(a).\end{aligned}$$

as $n \to \infty$. □

Remark 3.54. Another important result in Lebesgue integration is Fubini's theorem. To state it, we note that given two measure spaces there is a a natural product measure $\mu_1 \otimes \mu_2$ such that $\mu_1 \otimes \mu_2(A \times B) = \mu_1(A)\mu_2(B)$. Fubini's theorem tells us that if f is either non-negative or f is integrable with respect to $\mu_1 \otimes \mu_2$, then one can interchange the order of integration as follows:

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f(x,y) d\mu_1 \otimes d\mu_2(x,y) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x,y) d\mu_2(y) \right) d\mu_1(x) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x,y) d\mu_1(x) \right) d\mu_1(y)\end{aligned}$$

3.7 Lebesgue spaces $L^p(\Omega)$

In this section we let Ω be a measure space.

Definition 3.55. Let $1 \leq p < \infty$. We define the vector space $\mathcal{L}^p(\Omega)$ as the space of measurable functions such that $|f|^p$ is integrable. We define f, g to be equivalent denoted as $f \sim g$, if $f = g$ almost everywhere and $L^p(\Omega) = \mathcal{L}^p(\Omega)/\sim$. The norm

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}} \tag{7}$$

makes $L^p(\Omega)$ a normed space (the fact that (7) is a norm follows from Proposition 3.44 and Proposition 3.57 below).

We define the vector space $\mathcal{L}^\infty(\Omega)$ as the space of measurable functions such that $|f|$ is finite almost everywhere and

$$\operatorname*{ess\,sup}_\Omega |f| = \inf\{c \in \mathbb{R} : \mu(\{|f| \geq c\}) = 0\} < \infty.$$

We define f, g to be equivalent denoted as $f \sim g$, if $f = g$ almost everywhere and define $L^\infty(\Omega) = \mathcal{L}^\infty(\Omega)/\sim$. This makes $L^\infty(\Omega)$ a normed space with the norm

$$\|f\|_{L^\infty} = \operatorname*{ess\,sup}_\Omega |f|,$$

where we again already appeal to Proposition 3.44 and Proposition 3.57 below.

Remark 3.56. In practice, we will always work with representatives of the equivalence class. However, for $\|\cdot\|_p$ to be a norm (and not only a seminorm) we need to identify functions that agree almost everywhere.

Example.

- If $\Omega = \mathbb{N}$ and $\mu = \delta$, where $\delta(A) = |A|$, then $L^p(\mu) = \ell^p$.
- Another important example is the Lebesgue spaces on the reals: $L^p(\mathbb{R})$.

Proposition 3.57. Let $f, g : \Omega \to \mathbb{K}$ be measurable. Then,

1. $\|fg\|_1 \leq \|f\|_p \|g\|_q$ for $1 = \frac{1}{p} + \frac{1}{q}$ and $1 \leq p, q \leq \infty$. (Hölder inequality)

2. $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ for $1 \leq p \leq \infty$. (Minkowski inequality)

Proof. This is similar to the proof for the ℓ^p spaces and will be omitted. $\square$

Theorem 3.58 (Riesz–Fischer)**.** The space $L^p(\Omega)$ for $p \in [1, \infty]$ is complete.

Proof. **Case 1:** $p = \infty$**.** Let f_n be Cauchy in L^∞. For all $n, m \in \mathbb{N}$, there exists a null set $N_{n,m}$ such that $\sup_{x \in N_{n,m}^c} |f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$. Then, $N = \cup_{n,m \in \mathbb{N}} N_{n,m}$ is null and for each $x \in N^c$, the sequence $f_n(x)$ is Cauchy in $\mathbb{K}$ and hence convergent with limit $f(x)$. Now, let $\varepsilon > 0$. Then choose N sufficiently large that $\|f_m - f_n\|_\infty < \varepsilon$ for all $m, n \geq N$. Then,

$$\sup_{x \in N^c} |f(x) - f_n(x)| \leq \sup_{x \in N^c} \lim_{m \to \infty} |f_m(x) - f_n(x)| < \varepsilon.$$

Hence $f_n \to f$ in L^∞ as $n \to \infty$.

Case 2: $1 \leq p < \infty$**.** Let f_n be Cauchy in L^p. Then it has a fast-Cauchy subsequence f_{n_k}, i.e. a subsequence such that $\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k}$. We define the function

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| \tag{8}$$

which is in L^p because $\|g\|_p \leq \|f_{n_1}\|_p + \sum_{k=1}^\infty \|f_{n_{k+1}} - f_{n_k}\|_p$ by monotone convergence. In particular, this means that $\sum_{k=1}^\infty |f_{n_{k+1}}(x) - f_{n_k}(x)|$ is finite almost everywhere and hence $(f_{n_k}(x))_k$ is Cauchy in $\mathbb{K}$ for almost every $x \in \Omega$. Define f as the almost everywhere limit. We note that $|f| \leq g$ because $f_{n_l} = f_{n_1} + \sum_{k=1}^{l-1}(f_{n_{k+1}} - f_{n_k})$. In particular, $|f|^p$ and $|f_{n_k}|^p$ are integrable. From the dominated convergence theorem we obtain that

$$\|f - f_{n_k}\|_p \to 0 \tag{9}$$

as $k \to \infty$. Hence, f_n converges to f. $\square$

The above proof also gives the following important result.

Theorem 3.59. If $f_n \to f$ as $n \to \infty$ in $L^p(\Omega)$ for $1 \leq p \leq \infty$. Then, there exists a subsequence f_{n_k} such that $f_{n_k} \to f$ pointwise almost everywhere as $k \to \infty$.

Remark 3.60. Passing to a subsequence is necessary in the above theorem as you constructed a sequence of functions which does not converge pointwise anywhere but which converges in $L^1([0,1])$.

Proposition 3.61. Let Ω be a finite measure space. Then $L^p(\Omega) \subset L^q(\Omega)$ for $1 \leq q \leq p \leq \infty$. Moreover, the inclusions are dense.

Proof. Problem set 4. $\square$

Theorem 3.62. Let Ω be a locally compact (for all $x \in \Omega$ there exists compact neighborhood) and σ-compact (there exists K_n compact such that $\Omega = \cup_{n \in \mathbb{N}} K_n$). Let μ be a regular Borel measure on Ω. Then, $L^p(\Omega)$ is separable for $1 \leq p < \infty$.

Proof. We have shown it for ℓ^p spaces in the problem set. For the general result, see [Rud91]. $\square$

Remark 3.63. In particular, the spaces $L^p(\mathbb{R}^n)$ are separable for $1 \leq p < \infty$. Moreover, C_c^∞ is dense $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$.

Theorem 3.64. The space $L^p(\Omega)$ is *uniformly convex* for $1 < p < \infty$, i.e. for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $f, g \in L^p$ with $\|f\|_p = \|g\|_p = 1$ and $\|f - g\|_p \geq \varepsilon$ we have $\|\frac{f+g}{2}\|_p \leq 1 - \delta$.

The above theorem is a consequence of the following inequalities due to Clarkson.

Lemma 3.65 (Clarkson inequality)**.** Let $1 < p < \infty$, q its Hölder conjugate and $f, g \in L^p$.

1. If $2 \leq p < \infty$, then

$$\left\|\frac{f+g}{2}\right\|_p^p + \left\|\frac{f-g}{2}\right\|_p^p \leq \frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p \tag{10}$$

$$\left\|\frac{f+g}{2}\right\|_p^q + \left\|\frac{f-g}{2}\right\|_p^q \geq \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{q/p} \tag{11}$$

2. If $1 < p < 2$, then

$$\left\|\frac{f+g}{2}\right\|_p^q + \left\|\frac{f-g}{2}\right\|_p^q \leq \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{q/p} \tag{12}$$

$$\left\|\frac{f+g}{2}\right\|_p^p + \left\|\frac{f-g}{2}\right\|_p^p \geq \frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p. \tag{13}$$

Proof. We will begin by showing (10). It is enough to show the following inequality for $a, b \in \mathbb{R}$:

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \frac{1}{2}|a|^p + \frac{1}{2}|b|^p.$$

We first note that for $c, d \geq 0$ we have

$$c^q + d^q \leq (c+d)^q$$

for any $q \geq 1$. Indeed, without loss of generality, we can set $d = 1$ and then consider the function

$$f(x) = (x+1)^q - x^q - 1$$

which satisfies $f(0) = 0$ and $f'(x) = q(x+1)^{q-1} - qx^{q-1} \geq 0$. Hence, $f(x) \geq 0$ for $x \geq 0$.

Apply this observation to $c = (|a+b|/2)^2$ and $d = (|a-b|/2)^2$ and $q = p/2$ to obtain

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leq \left(\frac{|a-b|^2}{4} + \frac{|a+b|^2}{4}\right)^{p/2} = \left(\frac{|a|^2 + |b|^2}{2}\right)^{p/2} \leq \frac{1}{2}|a|^p + \frac{1}{2}|b|^p,$$

where in the last step we used the convexity of $x \mapsto x^{p/2}$. The other inequalities follow using similar arguments, see e.g. [Ada75, Theorem 2.38].

$\square$

Proof of Theorem 3.64. If $1 < p < 2$, then from (12) we obtain

$$\left\| \frac{f+g}{2} \right\|_p^q \le 1 - \left\| \frac{f-g}{2} \right\|_p^q \le 1 - 2^{-q}\varepsilon^q. \tag{14}$$

Similarly, the case $2 \le p < \infty$ follows from (10). □

An interesting consequence of uniform convexity is the following.

Theorem 3.66 (Milman–Pettis theorem)**.** Every uniformly convex Banach space is reflexive.

Proof. See e.g. [Bre11]. □

In particular, this shows that L^p is reflexive for $1 < p < \infty$. In the next section, we will give a direct characterization of the dual of L^p and directly show that L^p is reflexive.

4 Dual spaces and weak topologies

In this section we will explore more consequences of the rich structure of duals of normed space and introduce the central *weak topologies* induced by the dual space. We will then give an explicit characterization of the duals of the L^p spaces.

4.1 Locally convex spaces and weak topologies

Definition 4.1. A *locally convex space (LCS)* X is a topological vector space for which the topology is induced by a family of seminorms $(p_i)_{i\in I}$ on X (i.e. the sets $V_{i,\varepsilon}(x) = \{y \in X : p_i(x-y) < \varepsilon\}$ form a neighborhood subbase) which separates points (i.e. for $x \in X$ there exists p_i such that $p_i(x) \neq 0$).

Remark 4.2. More precisely, if X is a LCS then a set $U \subset X$ is open if and only if for all $x \in U$ there exists a finite set $F \subset I$ and $\varepsilon > 0$ such that

$$\bigcap_{i\in F} \{y \in X : p_i(x-y) < \varepsilon\} = \{y \in X : p_i(x-y) < \varepsilon \text{ for all } i \in F\} \subset U.$$

Moreover, from the assumption that the family of seminorms separates points, we obtain that every locally convex space is Hausdorff.

Definition 4.3. Let X be a normed space.

- The *weak topology* on X is the locally convex topology on X induced by the family of seminorms $\{x \mapsto |f(x)|\}_{f\in X^*}$. Notations: $\mathcal{T}_w$, w_X or $\sigma(X, X^*)$.
- The *weak* topology* on X^* is the locally convex topology on X^* induced by the family of seminorms $\{f \mapsto |f(x)|\}_{x\in X}$. Notations: $\mathcal{T}_{w*}$, $w{*}_{X^*}$ or $\sigma(X^*, X)$.

Proposition 4.4. Let X be a normed space. The weak topology on X is the initial topology induced by X^* (i.e. the weak topology is the coarsest topology such that for all $f \in X^*$, the map $(X, \mathcal{T}_w) \to \mathbb{K}, x \mapsto f(x)$ is continuous).

Proof. The sets $\{y \in X : |f(x-y)| < \varepsilon\}$ for $f \in X^*$ and $\varepsilon > 0$ form a neighborhood subbase (around $x \in X$) for the LCS topology. Note that sets of the form $f^{-1}(U)$ for $f \in X^*$ and $U \subset \mathbb{K}$ open form a subbase for the initial topology. In particular, $\{y \in X : |f(x-y)| < \varepsilon\}$ for $f \in X^*$ and $\varepsilon > 0$ form a neighborhood subbase (around $x \in X$) for the initial topology. Since both topologies have the same neighborhood subbase around every point, the topologies agree. $\square$

Analogously, we obtain the following.

Proposition 4.5. Let X be a normed space and X^* be its dual. The weak* topology on X^* is the initial topology induced by X (i.e. the weak* topology is the coarsest topology such that for all $x \in X$, the map $(X^*, \mathcal{T}_{w*}) \to \mathbb{K}, f \mapsto f(x)$ is continuous).

Remark 4.6.

- The weak topology is coarser than the norm (strong) topology. They agree for finite dimensional normed spaces.
- The weak* topology is coarser than the weak topology on X^*. They agree if and only if X is reflexive. Indeed, if $\sigma(X^*, X) \subset \sigma(X^*, X^{**})$, where the equality holds if and only if $X = X^{**}$.
- We also note that a function $g : Y \to (X, \mathcal{T}_w)$ is continuous if and only if $f \circ g : Y \to \mathbb{K}$ is continuous for all $f \in X^*$. Similarly, $g : Y \to (X^*, \mathcal{T}_{w*})$ is continuous if and only if $i(x) \circ g : Y \to \mathbb{K}$ is continuous for all $x \in X$, where $i : X \to X^{**}$ is the canonical embedding.

Proposition 4.7. Let X be a normed space.

1. A sequence $(x_n)_n \subset X$ converges to x in the weak topology, denoted as $x_n \rightharpoonup x$ as $n \to \infty$ if and only if $f(x_n) \to f(x)$ as $n \to \infty$ for all $f \in X^*$.
2. A sequence $(f_n)_n \subset X^*$ converges to f in the weak* topology, denoted as $f_n \overset{*}{\rightharpoonup} f$ as $n \to \infty$ if and only if $f_n(x) \to f(x)$ as $n \to \infty$ for all $x \in X$.

Proof. We have that $x_n \rightharpoonup x$ if and only if for any member $U_{f,\varepsilon,x} = \{y : |f(x-y)| < \varepsilon\}$ of the neighborhood subbase of x, we have $x_n \in U_{f,\varepsilon,x}$ eventually (i.e. for all $n \geq N_{f,\varepsilon}$). But this is equivalent to the statement that $|f(x_n) - f(x)| < \varepsilon$ eventually, i.e. $f(x_n) \to f(x)$ for all $f \in X^*$. $\square$

Theorem 4.8. Let $(x_n)_n \subset \ell^1$. Then, $x_n \to x$ if and only if $x_n \rightharpoonup x$ as $n \to \infty$.

Proof. Problem set. $\square$

Remark 4.9. The reason why it is true is that $(\ell^1)^* = \ell^\infty$ is a very large space so testing only against ℓ^∞ elements is enough. Note however that the weak topology is not the same as the norm topology. In fact, the weak topology (or weak* topology) on an infinite dimensional space is never metrizable.

Theorem 4.10. Let X be a normed space and $x_n \rightharpoonup x$. Then,

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\| \leq \sup_{n\in\mathbb{N}} \|x_n\| < \infty.$$

Proof. See problem set. $\square$

Theorem 4.11. Let X be a normed space and X^* its dual. If $f_n \overset{*}{\rightharpoonup} f$. Then,

$$\|f\| \leq \liminf_{n\to\infty} \|f_n\| \leq \sup_{n\in\mathbb{N}} \|f_n\| < \infty.$$

Proof. See problem set. $\square$

4.2 Banach–Alaoglu theorem

We have seen that the unit ball in an infinite dimensional normed space is never compact in the norm topology. This is somewhat unfortunate because compactness arguments are incredibly useful in applications. However, we will see that we can obtain compactness when we consider coarser topologies.

Theorem 4.12 (Banach–Alaoglu)**.** Let X be a normed space. Then the closed unit ball

$$\bar{B}_{X^*} = \{f \in X^* : \|f\| \leq 1\} \subset X^*$$

is weak* compact.

Remark 4.13. A word of caution: Even though $\bar{B}_{X^*}$ is compact, the space X^* with the weak* topology is not locally compact (i.e. there does not exist a compact neighborhood around any point). In particular, $\bar{B}_{X^*}$ has empty interior in the weak* topology.

Proof. For $x \in X$ define $K_x = \{\lambda \in \mathbb{K} : |\lambda| \leq \|x\|\}$ and set $K = \Pi_{x\in X} K_x$ with projections $\pi_x : K \to K_x$. By Tychonoff's theorem (Theorem 1.22), K is compact in the product topology. Note that (in mild abuse of notation) $\bar{B}_{X^*} \subset K$ because $f(x) \in K_x$ and thus $\Pi_{x\in X}\{f(x)\} \in K$. Note that the weak* topology on $\bar{B}_{X^*}$ is the initial topology induced by the maps $\{\bar{B}_{X^*} \ni f \mapsto f(x)\}_{x\in X}$ and the subspace topology on $\bar{B}_{X^*} \subset K$ is the initial topology induced by the maps $\{\pi_x\}_{x\in X} = \{\bar{B}_{X^*} \ni f \to f(x)\}_{x\in X}$. Thus, the subspace topology on $\bar{B}_{X^*} \subset K$ induced by the product topology on K agrees with the weak* topology on $\bar{B}_{X^*}$. Thus, it suffices to show that $\bar{B}_{X^*} \subset K$ is closed. To see this we write

$$\begin{aligned}\bar{B}_{X^*} = \{f \in K : f \text{ linear}\} &= \bigcap_{x,y\in X,\lambda,\mu\in\mathbb{K}} \{f \in K : f(\lambda x + \mu y) - \lambda f(x) - \mu f(y) = 0\} \\ &= \bigcap_{x,y\in X,\lambda,\mu\in\mathbb{K}} \{f \in K : (\pi_{\lambda x+\mu y} - \pi_{\lambda x} - \pi_{\mu y})(f) = 0\}\end{aligned}$$

which is an intersection of closed sets. Thus, $\bar{B}_{X^*}$ is closed and hence compact. $\square$

Of particular interest in PDEs for the direct method of calculus of variations is sequential compactness. Under the additional assumption of separability of X we can improve the Banach–Alaoglu result to obtain weak* sequential compactness.

Theorem 4.14 (Banach–Alaoglu, separable normed space)**.** Let X be a separable normed space. Then, $\bar{B}_{X^*}$in X^* is weak* sequentially compact.

Proof. Proof 1. On the problem set, you will show that if X is separable, then the weak* topology restricted to $\bar{B}_{X^*}$ is metrizable. Hence, sequential compactness is equivalent to compactness (problem set 2), so $\bar{B}_{X^*}$ is sequentially compact.

Proof 2. We will also give a direct proof using a diagonal argument, a standard technique in analysis. Let $D = (x_n)_n \subset X$ be dense and let $(f_k)_k \subset \bar{B}_{X^*}$. We have to show that f_k has a subsequence which converges in the weak* topology. For $n = 1$ consider the sequence $(f_k(x_1))_k$ in $\mathbb{K}$. The sequence is uniformly bounded by $\|x_1\|$ and thus has a convergence subsequence $(f_{k_{j,1}}(x_1))_j$. Now, consider the sequence $(f_{k_{j,1}}(x_2))_j$ in $\mathbb{K}$ which is again uniformly bounded and thus has a convergent subsequence $(f_{k_{j,2}}(x_2))_j$, where we emphasize that $(k_{j,2})_j$ is a subsequence of $(k_{j,1})_j$. We iterate this procedure to obtain that $(k_{j,n})_j$ is a subsequence of any $(k_{j,m})_j$ for $m \leq n$. We now take the diagonal sequence $(k_{j,j})_j$.

For $x_l \in D$ we define $\bar{f}(x_l) = \lim_{j\to\infty} f_{k_{j,j}}(x_l)$ which converges by the construction above. By linearity, we extend $\bar{f}$ to $\operatorname{span}(D)$ which is a dense subspace. Since $\bar{f}$ is bounded, we can extend $\bar{f}$ uniquely to f on X by the bounded linear extension theorem Theorem 2.20. □

Remark 4.15. We will end our discussion on locally convex space here, but there are many interesting results one can show. Especially, metrizable LCS (called *Fréchet spaces*), which are complete, constitute the foundation of distribution theory.

4.3 Riesz representation for L^p spaces

Our next theorem will be the famous characterizations for the dual spaces of the L^p spaces. We will consider a measure space Ω with measure μ. We let $p \in [1, \infty)$ and $q \in (1, \infty]$ be Hölder conjugates, i.e. $1 = \frac{1}{p} + \frac{1}{q}$. We define the map

$$\phi : L^q \to (L^p)^*, g \mapsto \phi_g, \text{ where } \phi_g(f) = \int_\Omega g f d\mu.$$

This is a linear, well-defined map because $|\phi_g(f)| \leq \|g\|_q \|f\|_p$ by the Hölder inequality. This also shows that the map ϕ is bounded by $\|\phi\| \leq 1$.

Theorem 4.16 (Riesz representation)**.** Let Ω, μ, p, q and ϕ as above.

1. If $1 < p < \infty$, then $(L^p(\Omega))^* \simeq L^q(\Omega)$.
2. If $p = 1$ and Ω is σ-finite, then $(L^1(\Omega))^* \simeq L^\infty(\Omega)$.

Remark 4.17. An alternative and possibly more standard proof of the L^p duality can be given using the Radon–Nikodym theorem. In the following, we will however give a more elementary proof via a direct minimization argument exploiting the uniform convexity of L^p for $1 < p < \infty$. The case $p = 1$ will then follow using a limiting argument.

Proof. Note that from Höder's inequality, we obtain that ϕ is a well-defined linear map with norm $\|\phi\| \leq 1$. We will first show that ϕ is in fact an isometry.

Step 1: ϕ is an isometry for $p > 1$. Let $0 \not\equiv g \in L^q$. Define $f = e^{i\theta}|g|^{q-1}$, where θ is chosen such that $fg = |g|^q$ almost everywhere. Then,

$$\|\phi_g\| \geq \frac{|\phi_g(f)|}{\|f\|_p} = \frac{\int |g|^q}{\left(\int |g|^{pq-p}\right)^{\frac{1}{p}}} = \left(\int |g|^q\right)^{\frac{1}{q}} = \|g\|_q.$$

Together with the bound $\|\phi_g\| \leq \|g\|_q$ from above we obtain that ϕ is an isometry.

Step 2: ϕ is an isometry for $p = 1$. Let $0 \not\equiv g \in L^\infty$ and fix $0 < \alpha < \|g\|_{L^\infty}$. By definition, there exists a measurable set K such that $|g|_K| \geq \alpha$ and $\mu(K) > 0$. Since μ is σ-finite we find a collection of sets A_n such that $\cup_{n\in\mathbb{N}} A_n = \Omega$. We set $K_m = K \cap \cup_{n=1}^m A_n$ and note that $\cup_{m\in\mathbb{N}} K_m = K$ from which we conclude that there exists a $M \in \mathbb{N}$ such that $\mu(K_M) > 0$ as $0 < \mu(K) = \mu(\cup_m K_m) \leq \sum_{m\in\mathbb{N}} \mu(K_m)$ by the σ-additivity of μ. Now we define $f = \lambda \mathbf{1}_{K_M}$, where λ is chosen such that $\lambda g = |g|$ on K_M. Then,

$$\|\phi_g\| \geq \frac{|\phi_g(f)|}{\|f\|_1} = \frac{\int |g|\mathbf{1}_{K_M}}{\mu(K_M)} \geq \alpha.$$

Since $\alpha < \|g\|_{L^\infty}$ was arbitrary, we obtain $\|\phi_g\| = \|g\|_{L^\infty}$.

Step 3: ϕ is surjective for $1 < p < \infty$. Let $F \in (L^p)^*$ and assume without loss of generality that $\|F\| = 1$.

Some intuition: We want to construct a g with $\|g\|_q = 1$ such that $F(f) = \int fg d\mu$ for all $\|f\|_p = 1$. If there exists a F-normalizing f such that $1 = \|F\| = F(f) = \int gf \leq \|g\|_q \|f\|_p \leq 1$, then we would have equality in Hölder's inequality so in particular $|f|^p = |g|^q$ so this would mean we should set $g = \lambda|f|^{p/q} = \lambda|f|^{p-1}$ s.t. $gf = |f|^p$. In order to construct such an f (and then also g), we will exploit the convexity of the unit sphere and the completeness.

There exists a sequence $(f_n)_n \subset L^p$ with $\|f_n\|_p = 1$ such that $\frac{1}{2} \leq F(f_n) \to 1$ as $n \to \infty$. We will show that $(f_n)_n$ is a Cauchy sequence. Let $\varepsilon > 0$. Since L^p is uniformly convex, there exists a $\delta > 0$ such that if

$$\left\|\frac{1}{2}(f+g)\right\| > 1 - \delta, \text{then } \|f - g\| < \varepsilon. \tag{15}$$

Choose N large such that $F(f_n) > 1 - \delta$ for all $n \geq N$. Thus, by linearity and since $\|F\| = 1$, we estimate

$$\left\|\frac{1}{2}\left(f_n + f_m\right)\right\| \geq F\left(\frac{1}{2}(f_n + f_m)\right) > 1 - \delta$$

for all $n, m \geq N$ from which we conclude using (15) that $\|f_n - f_m\|_p < \varepsilon$ for all $n, m \geq N$. Hence, $(f_n)_n$ is Cauchy with limit f satisfying $\|f\|_p = 1$ and $F(f) = 1$. We have now constructed our normalizing $f \in L^p$. From our intuition gained before, we set $g = \lambda|f|^{p/q}$ for some function λ with $|\lambda| = 1$ chosen such that $gf = |f|^p$. Note that $g \in L^q$ with

$$\|g\|_q^q = \int |f|^p d\mu = 1 = \int gf d\mu = \phi_g(f).$$

We want to show that $F = \phi_g$. To do so, we will prove the following claim. If $F_1, F_2 \in (L^p)^*$ with $\|F_1\| = \|F_2\| = 1$ and there exists a $f \in L^p$ with $\|f\|_p = 1$ such that $F_1(f) = F_2(f) = 1$, then $F_1 = F_2$. For the sake of a contradiction, suppose that $F_1 \neq F_2$. Hence there exists a function h such that $F_1(h) \neq F_2(h)$. By multiplying h with a constant we can assume that $F_1(h) - F_2(h) = 2$. By replacing h with $h + \alpha f$ for a suitable constant α, we can assume in addition that $F_1(h) = 1$ and hence, $F_2(h) = -1$. For $t > 0$ we have $F_1(f + th) = 1 + t$ and $F_2(f - th) = 1 + t$ and thus

$$1 + t \leq \|f + th\| \text{ and } 1 + t \leq \|f - th\|.$$

For $1 < p < 2$ we estimate using (13)

$$\begin{aligned} 1 + t^p\|h\|_p^p &= \left\| \frac{(f+th)+(f-th)}{2} \right\|_p^p + \left\| \frac{(f+th)-(f-th)}{2} \right\|_p^p \\ &\geq \frac{1}{2}\|f+th\|_p^p + \frac{1}{2}\|f-th\|_p^p \geq (1+t)^p \geq 1 + tp \end{aligned}$$

which gives a contradiction by sending $t \to 0$. Similarly, we argue for $2 \leq p < \infty$ using (11). This establishes the result for $1 < p < \infty$.

Step 4: ϕ is surjective for $p = 1$. We will approximate the construction for $p > 1$. First, we will consider the case of a finite measure, i.e. $\mu(\Omega) < \infty$. Let $F \in (L^1)^*$. Since L^p (for $p > 1$) is continuously embedded in L^1, we have that $F \in (L^p)^*$. Thus, by the above, there exists a $g_q \in L^q$ with $\phi_{g_q} = F$ and

$$\|g_q\|_q = \|F\|_{(L^p)^*} \leq \sup_{\|g\|_p = 1} |F(g)| \leq \sup_{\|g\|_\infty \leq (\mu(X))^{1/p}} |F(g)| \leq \mu(X)^{1/p}$$

which is uniformly bounded for $p \in [1, 2]$.

We note that g_q is independent of q because for q_1, q_2 define the test function $f \in L^\infty$ such that $f(g_{q_1} - g_{q_2}) = |g_{q_1} - g_{q_2}|$. Then,

$$0 = F(f) - F(f) = \phi_{g_{q_1}}(f) - \phi_{g_{q_2}}(f) = \int |g_{q_1} - g_{q_2}| d\mu,$$

hence $g_{q_1} = g_{q_2}$. Moreover, $g \in L^\infty$ because for $g_F := \min(|g|, \|F\| + 1)$ we note that (exercise!) $\|g_F\|_\infty = \lim_{q\to\infty} \|g_F\|_q \leq \|F\|$ so $\min(|g|, \|F\| + 1) \leq \|F\|$ almost everywhere and hence $|g| \leq \|F\|$ almost everywhere. Finally, F and ϕ_g agree on the dense subspace L^2 (problem set 4), and since they are continuous, they have to agree everywhere.

In order to relax the assumption of a finite measure space, we write $\Omega = \bigsqcup_{n\in\mathbb{N}} A_n$ for pairwise disjoint sets A_n with $\mu(A_n) < \infty$. For $F \in (L^1)^*$ define $F_n \in (L^1(A_n))^*$ as $F_n(f) = F(f)$, where $f \in L^1(\Omega)$ is continued by zero outside A_n. Then, we obtain a $g_n \in L^\infty(A_n)$ such that $\phi_{g_n} = F_n$ and $\|g_n\|_\infty = \|F_n\| \leq \|F\|$ by the previous step. Then, define $g = \sum_{n\in\mathbb{N}} g_n$, where g_n is continued by zero outside of A_n. Since the g_n have disjoint support, $\|g\|_\infty \leq \|F\|$.

In order to show that $\phi_g = F$ we let $f \in L^1(\Omega)$. Then, $f = \sum_{n=1}^\infty f_n$ pointwise, where $f_n = f\mathbf{1}_{A_n}$. Since $f = \lim_{N\to\infty} \sum_{n=1}^N f_n$ in $L^1(\Omega)$ by dominated convergence, and F is continuous we have

$$F(f) = \sum_{n=1}^{\infty} F(f_n) = \sum_{n=1}^{\infty} F_n(f_n) = \sum_{n=1}^{\infty} \int f_n g_n \mathbf{1}_{A_n} = \sum_{n=1}^{\infty} \int g f \mathbf{1}_{A_n} = \int g f = \phi_g(f),$$

where we used dominated convergence in the last step again which is allowed because

$$\left| \sum_{n=1}^{\infty} g f \mathbf{1}_{A_n} \right| \leq |gf| \in L^1(\Omega)$$

by Hölders inequality.

□

We immediately obtain the following corollary.

Corollary 4.18. The space $L^p(\Omega)$ is reflexive for $1 < p < \infty$.

Theorem 4.19. The spaces ℓ^1, ℓ^∞, $L^1(\mathbb{R})$, and $L^\infty(\mathbb{R})$ are not reflexive.

Proof. We have established these facts for ℓ^1 and ℓ^∞ before.

If $L^1(\mathbb{R})$ was reflexive, then $L^1(\mathbb{R}) = (L^1(\mathbb{R}))^{**} = (L^\infty(\mathbb{R}))^*$ was separable (since $L^1(\mathbb{R})$ is separable) so $L^\infty(\mathbb{R})$ would have to be separable which is a contradiction (note that all different members of the uncountable family $(\mathbf{1}_{[0,x]})_{x\in\mathbb{R}}$ have unit distance from each other). But if $L^1(\mathbb{R})$ is not reflexive then $L^\infty(\mathbb{R}) = (L^1(\mathbb{R}))^*$ cannot be reflexive (pset 3). □

Remark 4.20. An example of a functional on $L^\infty(\mathbb{R})$ that is not in $L^1(\mathbb{R})$ can be given as follows. Consider the function $F : C_b(\mathbb{R}) \to \mathbb{K}, f \mapsto f(0)$. Note that F is a well-defined functional on the space of bounded continuous functions $C_b(\mathbb{R}) \subset L^\infty(\mathbb{R})$. Moreover, $\|F\| = 1$. By Hahn–Banach it can be continued to a functional on $(L^\infty(\mathbb{R}))^*$. If F was represented by $g \in L^1(\mathbb{R})$, one can verify that then $g = 0$ almost everywhere which is a contradiction.

5 Hilbert spaces

In the previous section we encountered the $L^p(\Omega)$ spaces and characterized their dual space as $L^q(\Omega)$. A special case is the case $p = 2$ for which we have $L^2(\Omega) \cong (L^2(\Omega))^*$ and moreover Hölder's inequality gives

$$\int_\Omega |fg| d\mu \leq \|f\|_{L^2} \|g\|_{L^2}.$$

This is special because f and g lie in the same space and there is a natural "inner product" defined on $L^2(\Omega)$ as

$$\langle f, g \rangle = \int_\Omega \bar{f} g d\mu.$$

Based on Hilbert's work on the $L^2(\Omega)$ and the additional structure of the inner product, John von Neumann introduced the abstract concept of a *Hilbert spaces*. The additional structure of an inner product allows us to define orthogonality and orthogonal projections. In particular, many geometric features of the Euclidean space have generalizations to Hilbert spaces. They also play a central role in applications such as in PDEs, quantum mechanics, etc.

Definition 5.1. Let X be a vector space. A map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ is called a *scalar product* (or *inner product*) if

1. $\langle x, x \rangle \geq 0$ for every $x \in X$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.
2. $\langle x, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle$ for every $\lambda, \mu \in \mathbb{K}$ and $x, y, z \in X$.
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$

The pair $(X, \langle \cdot, \cdot \rangle$ is called an *inner product space* (or pre-Hilbert space).

Remark 5.2. In the Hilbert space axioms, we choose the convention of linearity with respect to the second component. Be aware that many other textbooks impose linearity in the first component.

Proposition 5.3 (Cauchy–Schwarz inequality)**.** Let H be a pre-Hilbert space. Then

$$|\langle v, w \rangle| \leq \|v\| \|w\|$$

with equality if and only if v and w are linearly dependent.

Proof. For $t \in \mathbb{K}$ define $p(t) = \|v + tw\|^2 = \|v\|^2 + |t|^2 \|w\|^2 + 2\mathrm{Re}\,(t\langle v, w \rangle)$. Assume WLOG that $w \neq 0$, and set

$$t_{\min} = -\frac{\overline{\langle v, w \rangle}}{\|w\|^2}$$

Hence,

$$0 \leq p(t_{\min}) = \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2}$$

from which we obtain the conclusion. Clearly, if v, w are linearly dependent, we have equality. If equality holds, then $0 = p(t_{\min})$ but then v and w are dependent. $\square$

Lemma 5.4. An inner product space is also a normed space with norm $\|v\| = \langle v, v \rangle^{\frac{1}{2}}$ and the inner product $\langle \cdot, \cdot \rangle \to \mathbb{K}$ is continuous.

Proof. It is easy to verify that $\|v\|$ fulfills the axioms of a norm. For the continuity we let $v_n \to v$ and $w_n \to w$. Then, using the bilinearity of the inner product and the Cauchy–Schwarz inequality we obtain

$$\begin{aligned} |\langle v_n, w_n \rangle - \langle v, w \rangle| &\leq |\langle v_n, w_n \rangle - \langle v_n, w \rangle| + |\langle v_n, w \rangle - \langle v, w \rangle| \\ &\leq \sup_m \|v_m\| \|w_n - w\| + \|w\| \|v_n - v\| \to 0 \end{aligned}$$

as $n \to \infty$. $\square$

Definition 5.5. A complete inner product space is called a *Hilbert space*.

Definition 5.6. A linear map $T : H_1 \to H_2$ between two inner product spaces is called an *(Hilbert space) isomorphism* (or a unitary map) if T is bijective and $\langle Tx, Ty \rangle = \langle x, y \rangle$ for all $x, y \in H_1$.

Theorem 5.7 (Hilbert space completion)**.** Let H_0 be an inner product space. Then there exists a Hilbert space H, a dense set $D \subset H$ and an isomorphism $T : H_0 \to D$. The space H is unique except for (Hilbert space) isomorphisms.

Proof. The proof is similar to the completion of metric and normed spaces, except that one has to additionally take the Hilbert space structure into account. The details of the proof are left to the reader. □

Remark 5.8. We also note that it is easy to verify that the norm of an inner produce space satisfies the *parallelogram law*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{16}$$

In fact, for any real inner product space, one has the polarization identity

$$\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$$

and similarly, for complex inner product spaces we have

$$\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2)$$

We note that not all normed spaces satisfy the parallelogram law and all normed spaces which satisfies the parallelogram inequality are inner product spaces.

Example. The space ℓ^2 is a Hilbert space and more generally, the space $L^2(\Omega)$ for a general measurable space Ω is a Hilbert space with inner product

$$\langle f, g \rangle = \int_\Omega \bar{f} g d\mu.$$

We note that the Clarkson inequalities of Lemma 3.65 reduce to the parallelogram law in the case $p = 2$. It is easy to verify that for $p \neq 2$, the parallelogram law is not satisfied for ℓ^p or, more generally, L^p. Thus, L^p and ℓ^p are Hilbert spaces if and only if $p = 2$.

5.1 Orthogonality and projection in Hilbert spaces

We will now introduce the concept of orthogonality.

Definition 5.9. Let H be an inner product space and $U \subset H$ and $x, y \in H$.

1. We say that x is *orthogonal* to y denoted as $x \perp y$ if $\langle x, y \rangle = 0$. In particular, for such x, y the Pythagorean theorem $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ holds.

2. We define the *orthogonal complement* of U as $U^\perp = \{x \in H : \langle x, y\rangle = 0 \text{ for all } y \in U\}$

3. We say that U is an *orthogonal set* if $\langle x, y\rangle = 0$ for all $x \neq y \in U$.

4. We say that U is an *orthonormal set* if it is an orthogonal set and $\|y\| = 1$ for all $y \in U$.

5. An orthonormal set U is an *orthonormal basis* (ONB) if it is maximal, i.e. if $U \subset V$, where V is an orthonormal set, then $U = V$.

Remark 5.10. Instead of an orthogonal/orthonormal set U we will often consider an orthogonal/orthonormal family $(e_\alpha)_{\alpha\in I}$ for an index set I. Given such a set, the index set can be U itself and the map is the identity map.

Lemma 5.11. An orthogonal set is linearly independent.

Proof. Let $e_1, \dots, e_n$ be orthonormal and $\sum_{i=1}^n \alpha_i e_i = 0$. Multiplying by e_j gives

$$\sum_{i=1}^n \alpha_i \langle e_j, e_i\rangle = 0,$$

i.e. $\alpha_j = 0$ for all $j \in \{1, \dots, n\}$. □

Lemma 5.12. Let X be an inner product space and $A \subset X$ be a subset. Then

- $A^\perp$ is closed.
- $A \subset (A^\perp)^\perp$
- $A^\perp = (\overline{\text{span}(A)})^\perp$

Proof. Let $b_n \to b$, where $b_n \in A^\perp$. Since the inner product is continuous we obtain $\langle b, a\rangle = \lim_{n\to\infty}\langle b_n, a\rangle = 0$ for all $a \in A$, i.e. $b \in A^\perp$. Clearly, $A \subset (A^\perp)^\perp$. Finally, if $b \in A^\perp$. Then, for all $a \in \text{span}(A)$, we have $\langle a, b\rangle = 0$ and by continuity we obtain that then $b \in (\overline{\text{span}(A)})^\perp$. The other direction is trivial. □

We now come to an important property of Hilbert spaces which reminds us of projection properties in Euclidean geometry.

For this, we first recall from linear algebra that a vector space X is an algebraic sum of two subspaces Y and Z of X, denoted as $X = Y \oplus_{alg} Z$ ($X \simeq Y \oplus_{alg} Z$) if for each $x \in X$ there exists unique $y \in Y$ and $z \in Z$ such that $x = y + z$. If the map $(Y, Z) \mapsto X, (y, z) \mapsto y + z$ is continuous, then we say that it is a topological direct sum, which we denote as $X = Y \oplus Z$. In the case of a Hilbert space and a closed subspace we will show that every closed subspace $V \subset H$ can be *complemented*, i.e. there exists a linear subspace (which will actually be $V^\perp$) such that $H = V \oplus V^\perp$. This is in general, not true for Banach spaces, e.g. $c_0 \subset \ell^\infty$ is a linear subspace which cannot be complemented. Before that we will state the following observation.

Lemma 5.13. Let X be a normed space and $P : X \to X$ be a linear continuous projection, i.e. $P^2 = P$. Then

- Either $P = 0$ or $\|P\| \geq 1$.
- Both $\ker(P)$ and $\mathrm{range}(P)$ are closed.
- $X = \ker(P) \oplus \mathrm{range}(P)$.

Proof. We have $\|P\| = \|P^2\| \leq \|P\|^2$ so either $P = 0$ or $\|P\| \geq 1$. Since P is continuous we have that its kernel is closed. Moreover, $(I - P)^2 = I - P$ so $I - P$ is a continuous projection with closed kernel. However, $\ker(I - P) = \mathrm{range}(P)$ because $x \in \mathrm{range}(P)$ if and only $x = Pz$ for some z which holds true if and only if $Px = P^2z = Pz = x$ so $x \in \ker(P - I)$. $\square$

Theorem 5.14. Let $V \neq \{0\}$ be a closed subspace of a Hilbert space H and $x \in H$. Then there exists a unique element $y \in V$ such that

$$\|x - y\| = \mathrm{dist}(x, V) := \inf_{z \in V} \|x - z\|.$$

In addition, the element $y \in V$ is the unique element such that $y - x \in V^\perp$.

The map

$$P_V : H \to V, x \mapsto y$$

is called the *orthogonal projection* onto V and is a bounded linear operator satisfying $P_V^2 = P_V$, $\mathrm{range}(P_V) = V$, and $\ker(P_V) = V^\perp$ and $\|P_V\| = 1$.

Moreover, $1 - P_V : H \to V^\perp$ is a projection with $\|1 - P_V\| = 1$ (except for $V = H$) and in particular,

$$H = V \oplus V^\perp.$$

Proof. **Step 1: Existence and uniqueness.** Without loss of generality $x \notin V$. Consider a minimizing sequence $(y_n)_n \subset V$ such that $\|x - y_n\| \to \mathrm{dist}(x, V)$ as $n \to \infty$. Note that y_n is a Cauchy sequence because

$$\begin{aligned}\|y_n - y_m\|^2 = \|y_n - x + x - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|x - y_m\|^2 - 4\|\frac{1}{2}(y_n + y_m) - x\|^2 \\ &\leq 2\|y_n - x\|^2 + 2\|x - y_m\|^2 - 4(\mathrm{dist}(x, V))^2.\end{aligned}$$

Taking $n, m \to \infty$ shows that $(y_n)_n$ is a Cauchy sequence with limit $y \in V$ since V is complete. An analogous argument using the parallelogram identity also gives uniqueness.

Step 2: Orthonality of $y - x$. To show the orthogonality we use a linearization argument. Let $z \in V$ and consider for $\varepsilon \in \mathbb{R}$ the vector $y + \varepsilon z \in V$. Then, since y is a minimizer we have

$$\|x - y\|^2 \leq \|x - y - \varepsilon z\|^2 = \|x - y\|^2 - 2\varepsilon \mathrm{Re}\langle x - y, z\rangle + \varepsilon^2 \|z\|^2$$

and hence

$$2\mathrm{Re}(\langle x - y, z\rangle) \leq \varepsilon \|z\|^2.$$

Considering both cases $\varepsilon > 0$ and $\varepsilon < 0$ and $|\varepsilon| \to 0$ shows that $\mathrm{Re}\langle x-y, z\rangle = 0$. By considering $i\varepsilon$ instead of ε, we analogously obtain $\mathrm{Im}\langle x-y, z\rangle$. For the uniqueness let $y, y' \in V$ be such that $y-x \in V^\perp$ and $y'-x \in V^\perp$. But then, $y-y' = y-x-(y'-x) \in V^\perp$ so $y = y'$.

Step 3: Linearity of P_V. In order to show that the map P_V is linear we note that for $\lambda x + \mu x' \in H$, we have that $\lambda P_V(x) + \mu P_V(x') - \lambda x + \mu x' \in V^\perp$ since $V^\perp$ is a linear space. Thus, by the characterization proved above, P_V is linear. The characterization also implies that $P_V^2 = P_V$ as well as $\mathrm{range}(P_V) = V$ and $\ker(P_V) = V^\perp$.

Step 4: Boundedness of P_V. In order to show the boundedness statement, we let $x \in H$, then $P_V(x) \in V$ and $x - P_V(x) \in V^\perp$. In particular, we obtain

$$\|x\|^2 = \|P_V x + (1-P_V x)\|^2 = \|P_v x\|^2 + \|(1-P_V)x\|^2$$

which implies $\|P_V\| \le 1$ and $\|1-P_V\| \le 1$.

From Lemma 5.13 we obtain that $H = V \oplus V^\perp$. □

Remark 5.15. The first conclusion of Theorem 5.14 is also true for uniformly convex Banach spaces and for closed convex sets. More precisely, let V be a closed convex subset of a uniformly convex Banach space X and $x \in X$. Then there exists a unique element $y \in V$ such that

$$\|x-y\| = \mathrm{dist}(x, V) := \inf_{z \in V} \|x-z\|.$$

Theorem 5.16. Let H be a Hilbert space and V be a subspace. Then

$$\overline{V} = (V^\perp)^\perp$$

Proof. The inclusion $\subset$ is clear from Lemma 5.12. Assume first that V is a closed subspace. If $v \in (V^\perp)^\perp$, then $0 = P_{V^\perp} v = (1-P_V)v =$ so $v = P_V(v)$ which means $v \in V$. If V is not closed, then apply the result to $\overline{V}$ and note that $V^\perp = \overline{V}^\perp$. □

5.2 Riesz' representation for Hilbert spaces

Theorem 5.17 (Riesz's representation for Hilbert spaces)**.** Let $f \in H^*$ be a linear functional on a Hilbert space H. There exists a unique $z_f \in H$ such that

$$f(x) = \langle z_f, x\rangle \tag{17}$$

for all $x \in H$ and $\|f\| = \|z_f\|$. In particular, the map

$$\mathcal{A} : H^* \to H, f \mapsto z_f$$

is an anti-linear isometric bijection with inverse

$$\mathcal{A}^{-1} : H \to H^*, x \mapsto \langle x, \cdot\rangle.$$

Remark 5.18. In quantum mechanics, one uses the bra-ket notation: $\langle\psi|$ is an element of H^* and $|\psi\rangle$ is an element of H. Then the map $\mathcal{A}^{-1} : H \to H^*$ is the map $|\psi\rangle \mapsto \langle\psi|$.

Proof. Without loss of generality $f \neq 0$ so let $0 \neq x_0 \in \ker(f)^\perp$, where used that $\ker(f)$ is closed. Now define

$$z_f = \overline{f(x_0)} \frac{x_0}{\|x_0\|^2}$$

and note that $f(x) = \langle z_f, x\rangle = 0$ for all $x \in \ker(f)$. Moreover, $\ker(f)^\perp$ is 1-dimensional[3] so for any $y \in \ker(f)^\perp$, we have $y = \alpha x_0$ so

$$\langle z_f, y\rangle = \alpha\langle z_f, x_0\rangle = \alpha f(x_0)\frac{1}{\|x_0\|^2}\langle x_0, x_0\rangle = f(y).$$

For the uniqueness of z_f, let z_f and z_f' satisfy (17), then $\langle z_f - z_f', x\rangle = 0$ for all $x \in X$ so in particular for $x = z_f - z_f'$ which shows $z_f = z_f'$. For the norm, we note that $\|f\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} |\langle z_f, x\rangle| \leq \|z_f\|$. Moreover, $\|f\| \geq \frac{1}{\|z_f\|}|f(z_f)| = \|z_f\|$.

To see the anti-linearity of $\mathcal{A}$ we note that

$$\langle z_{\alpha f+\beta g}, x\rangle = (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) = \alpha\langle z_f, x\rangle + \beta\langle z_g, x\rangle = \langle \bar{\alpha} z_f + \bar{\beta} z_g, x\rangle.$$

Since $\mathcal{A}$ is isometric it is injective. It is also surjective as the $z \in H$, the map $f_z : x \mapsto \langle z, x\rangle$ is a bounded linear functional. $\square$

We directly obtain the following corollary from the Riesz representation theorem and the canonical map $\mathcal{A} : H^* \to H$.

Corollary 5.19. Let H be a Hilbert space.

- A sequence $(x_n)_n$ converges weakly to x if and only if $\langle y, x_n - x\rangle \to 0$ as $n \to \infty$ for every $y \in H$.
- H is reflexive.
- Every bounded sequence in H contains a weakly convergent subsequence.

Proof. The only non-trivial statement is the third statement. For this, we note that if $(x_n)_n$ is a convergent sequence, then define $H_0 := \overline{\operatorname{span}}\{x_n : n \in \mathbb{N}\}$. Note that H_0 is a Hilbert space as a closed subspace of H and is separable. But since weakly convergent subsequences are bounded, by Banach–Alaoglu, they are compact in the weak topology. Since H_0 is separable, the unit ball is metrizable in the weak topology and is thus sequentially compact. Hence, $(x_n)_n$ has a convergent subsequence in $H_0 \subset H$. $\square$

5.3 Bessel's inequality and Parseval identity

Throughout this subsection we let H be a Hilbert space. Some statements clearly also extend to inner product spaces.

[3] Here we use that $H/\ker(f) \cong \ker(f)^\perp$ and $H/\ker(f) \cong \operatorname{range}(f)$ so $\ker(f)^\perp \cong \operatorname{range}(f)$.

Lemma 5.20 (Bessel inequality)**.** Let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal set and $x \in H$. Then,

$$\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 \leq \|x\|^2. \tag{18}$$

Proof. For $\{e_1, \dots e_N\}$ we consider $x = u + v$, where $u := \sum_{i=1}^{N} \langle e_i, x\rangle e_i$ and $v := x - u$. We note that $u \perp v$ because

$$\langle u, v\rangle = \sum_{i=1}^{N} \langle x, e_i\rangle \left(\langle e_i, x\rangle - \sum_{j=1}^{N} \langle e_j, x\rangle \langle e_i, e_j\rangle \right) = \sum_{i=1}^{N} \langle x, e_i\rangle \left(\langle e_i, x\rangle - \langle e_i, x\rangle \right) = 0$$

Similarly,

$$\langle u, u\rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle x, e_i\rangle \langle e_j, x\rangle \langle e_i, e_j\rangle = \sum_{i=1}^{N} |\langle x, e_i\rangle|^2$$

and thus

$$\|x\|^2 = \|u\|^2 + \|v\|^2 \geq \|u\|^2 = \sum_{i=1}^{N} |\langle x, e_i\rangle|^2.$$

Sending $N \to \infty$ proves the claim. □

Lemma 5.21. Let $S \subset H$ be an orthonormal set. Let $x \in H$. Then the set $S_x = \{e \in S : \langle e, x\rangle \neq 0\}$ is (at most) countable.

Proof. From Bessel's inequality we obtain for every $n \in \mathbb{N}$ the set $S_{x,n} = \{e \in S : |\langle e, x\rangle| \geq \frac{1}{n}\}$ is finite. Hence $S_x = \cup_{n\in\mathbb{N}} S_{x,n}$ is countable. □

In order to deal with non-separable Hilbert space we have to introduce the conecept of unconditional convergence of series. Recall that a series $\sum_{x=1}^{\infty} x_n$ is absolutely convergent if $\sum_{x=1}^{\infty} \|x_n\| < \infty$.

Definition 5.22. Let X be a normed space and I an uncountable index set. Let $x_i \in X$ for all $i \in I$. We say that the series $\sum_{i\in I} x_i$ converges *unconditionally* to $x \in X$ if

- $I_0 = \{i \in I : x_i \neq 0\}$ is countable.
- for all enumerations of $I_0 = \{i_1, i_2, \dots\}$ we have $x = \sum_{n=1}^{\infty} x_{i_n}$.

Remark 5.23. For finite dimensional normed spaces a series converges unconditionally if and only if it converges absolutely. In infinite dimensions, absolutely convergent series still converge unconditionally but the converse is not necessarily true, e.g. the series $\sum_{n=1}^{\infty} \frac{1}{n} e_n$ does not converge absolutely in ℓ^∞. However, $\sum_{n\in\mathbb{N}} \frac{1}{n} e_n$ converges unconditionally to $x = (1, 1/2, 1/3, \dots)$. In fact, every infinite-dimensional Banach space admits series which are unconditional convergent but which do not converge absolutely. This is a theorem of Dvoretzky and Rogers.

Lemma 5.24 (Bessel inequality)**.** Let $(e_i)_{i\in I}$ be an orthonormal set. Then, for every $x \in H$

$$\sum_{i\in I} |\langle x, e_i\rangle|^2 \leq \|x\|^2.$$

Proof of Lemma 5.24. This follows directly from Lemma 5.21 and from Lemma 5.20. □

Remark 5.25. For $\alpha_i \in [0,\infty)$, $i \in I$ the uncountable sum $\sum_{i\in I} \alpha_i$ is defined as

$$\sum_{i\in I} \alpha_i := \sup_{J\subset I, |J|<\infty} \sum_{j\in J} \alpha_j.$$

If the sum is finite, we can verify that $\alpha_i = 0$ for all but countably many $i \in I$.

Lemma 5.26. Let $S \subset H$ be an orthonormal set. Then the following statements are equivalent.

1. S is an orthonormal basis.
2. If $x \in X$ and $\langle x, y\rangle = 0$ for all $y \in S$, then $x = 0$.
3. $\operatorname{span}(S)$ is dense in H.

Proof. 1) ⇒ 2) If $\langle x, y\rangle = 0$ for all $y \in S$ and $x \neq 0$, then the set $S \cup \{x/\|x\|\}$ would be an ONB. However, S is maximal so $x = 0$.

2) ⇒ 3) The closure $Y = \overline{\operatorname{span}(S)}$ is a closed subspace of H. If $x \in Y^\perp$, then by Lemma 5.12, $x \in S^\perp$ so by 2) $x = 0$. Hence, $Y^\perp = \{0\}$ so $Y = H$.

3) ⇒ 1) If $S' \supset S$ is an ONS and $x \in S' \setminus S$, then $x \perp S$, and hence $x \perp \overline{\operatorname{span}}(S) = H$ so $x = 0$ which gives a contradiction. So S is an orthonormal basis.

□

Lemma 5.27 (Parseval identity)**.** Let H be a Hilbert space and $(e_i)_{i\in I}$ be an ONB. Then,

$$x = \sum_{i\in I} \langle e_i, x\rangle e_i,$$

where the series converges unconditionally. Moreover,

$$\|x\|^2 = \sum_{i\in I} |\langle e_i, x\rangle|^2.$$

Proof. We let $(e_n)_{n\in\mathbb{N}} \subset (e_i)_{i\in I}$ be an enumeration of the subset for which the coefficients $\langle e_n, x\rangle$ are non-zero. We define $x_N = \sum_{n=1}^{N} \langle e_n, x\rangle e_n$ and verify that $(x_N)_N$ is a Cauchy sequence. Indeed, for $M \geq N$, we compute using the orthogonality that

$$\|x_N - x_M\|^2 \leq \sum_{n=N}^{M} |\langle e_n, x\rangle|^2 \to 0$$

as $M, N \to \infty$ since the sum converges by Bessel's inequality. Thus, $x_N \to y$ as $N \to \infty$ and we are left to show that $y = x$. Clearly, $x, y \perp (e_i)_{i\in I} \setminus (e_n)_{n\in\mathbb{N}}$ and since $\langle e_n, y\rangle = \langle e_n, x\rangle$ for all n we have that $x = y$ by Lemma 5.26. The second statement also follows from the continuity of the inner product and orthogonality and is left to the reader. □

Theorem 5.28. Every Hilbert space H admits an orthonormal basis. Moreover, H is separable if and only if every orthonormal basis of H is countable.

Proof. This is an application of Zorn's lemma. Consider $P = \{O \subset H : O \text{ is orthonormal}\}$. Clearly, P is a non-empty poset by set inclusion. For any chain C the set $\cup_{c \in C} c$ is an upper bound so P has a maximal element O. Clearly, O is orthonormal and complete.

If O is a countable ONB, then it is a countable Schauder basis so by Lemma 2.7, H is separable. Now assume that H is separable and let $(d_n)_{n\in\mathbb{N}}$ be a countable dense set and $(e_i)_{i\in I}$ be an ONB. Consider the collection of balls $(B_{1/2}(e_i))_{i\in I}$. Since $\|e_i - e_j\| = \sqrt{2}$ for $i \neq j$, these balls are disjoint. Since $(d_n)_n$ are dense, for each $i \in I$ there exists a $n(i) \in \mathbb{N}$ such that $d_{n(i)} \in B_{1/2}(e_i)$. Since the balls are disjoint, the map $I \to \mathbb{N}, i \mapsto n(i)$ is injective so I is countable. $\square$

Remark 5.29. In practice, virtually all Hilbert spaces of relevance and interest are separable.

In the separable case, there is essentially only one Hilbert space, ℓ^2.

Theorem 5.30. Let H be a separable Hilbert space. If $\dim(H) = n < \infty$, then H is isomorphic to $\mathbb{K}^n$ and when $\dim(H) = \infty$, then H is isomorphic to ℓ^2.

Proof. Let us only consider the case $\dim(H) = \infty$. Let $(e_i)_{i\in\mathbb{N}}$ be a ONB and define the map $T : H \to \ell^2$ by defining

$$x \mapsto Tx = (\langle e_i, x\rangle)_{i\in\mathbb{N}} \in \ell^2.$$

T is well-defined linear operator by the Bessel inequality. Moreover, T is injective because e_i is an ONB. To show that T is surjective we note that for $(x_i)_{i\in\mathbb{N}} \in \ell^2$, the sequence $\sum_{i=1}^{\infty} x_i e_i$ is well-defined using the same argument as in the proof of Lemma 5.27. $\square$

5.4 Hilbert space adjoints

Definition 5.31. Let X, Y be normed spaces and $T \in BL(X,Y)$. The *dual operator* T' of T is defined as

$$T' : Y^* \to X^*, f \mapsto f \circ T.$$

Proposition 5.32. The map $'$: $BL(X,Y) \to BL(Y^*, X^*)$ is a linear and bounded isometry.

Proof. Clearly, $'$ is linear. Moreover, it is an isometry because

$$\begin{aligned} \|T'\| &= \sup_{\|f\|=1} \|T'(f)\| = \sup_{\|f\|=1} \sup_{\|x\|=1} |T'(f)(x)| \\ &= \sup_{\|f\|=1} \sup_{\|x\|=1} |f(Tx)| = \sup_{\|x\|=1} \sup_{\|f\|=1} |f(Tx)| = \sup_{\|x\|=1} \|Tx\| = \|T\|, \end{aligned}$$

where we used Corollary 2.24. $\square$

Recall that in a Hilbert space the adjoint map $\mathcal{A} : H^* \to H$ is an isomorphism.

Definition 5.33. Let H be a Hilbert space and $T \in BL(H)$. We define the Hilbert adjoint T^* as

$$T^* = \mathcal{A}T'\mathcal{A}^{-1}$$

so that T^* is uniquely characterized by

$$\langle x, Ty\rangle = \langle T^*x, y\rangle$$

for all $x, y \in H$. Moreover, $\|T^*\| = \|T\|$.

Remark 5.34. A word of caution: Often the dual operator is also denoted with * instead of $'$.

Proposition 5.35. Let H be a Hilbert space and $S, T \in BL(H)$, $\lambda, \mu \in \mathbb{C}$. Then the following holds true.

- $(\lambda T + \mu S)^* = \bar{\lambda}T^* + \bar{\mu}S^*$.
- $(ST)^* = T^*S^*$
- $T^{**} = T$
- $\|T^*T\| = \|T\|^2$
- $\ker(S) = \mathrm{range}(S^*)^\perp$
- $\ker(S^*) = \mathrm{range}(S)^\perp$
- If $S^{-1} \in BL(H)$ exist, then S^* is invertible with $(S^*)^{-1} = (S^{-1})^*$.

Proof. The first three statements are clear from the definition. For the fourth statement, we note that

$$\begin{aligned}
\|T^*T\| = \sup_{\|x\|=\|y\|=1} |\langle x, T^*Ty\rangle| &= \sup_{\|x\|=\|y\|=1} |\langle Tx, Ty\rangle| \leq \|T\|^2 \\
&= \sup_{\|x\|=1} |\langle Tx, Tx\rangle| \leq \sup_{\|x\|=\|y\|=1} |\langle Tx, Ty\rangle| = \|T^*T\|.
\end{aligned}$$

For the fifth statement, we observe that $x \in \ker(S)$ if and only if $\langle y, Sx\rangle = 0$ for all $y \in H$ which is true if and only if $\langle S^*y, x\rangle = 0$ for all $y \in H$, i.e. if and only if $x \perp \mathrm{range}(S^*)$.

For the last statement we note that $(S^{-1})^*S^* = (SS^{-1})^* = I = (S^{-1}S)^* = S^*(S^{-1})^*$. □

Definition 5.36. Let H be a Hilbert space and $T \in BL(H)$. We say that T is

- *self-adjoint* (or *hermitian*) if $T^* = T$,
- *unitary* if $T^*T = TT^* = I$,

- *normal* if $TT^* = T^*T$.

Example.

- If H is $\mathbb{C}^n$ and $M \in BL(H)$ be represented by the matrix $(M_{ij})_{i,j}$ then M^* is represented by $(\overline{M_{ji}})_{i,j}$.
- Consider the shift operator $S : \ell^2 \to \ell^2$ which acts as $(x_1, x_2, \dots) \mapsto (x_2, x_3, \dots)$. Then, $S^* : \ell^2 \to \ell^2$ is given as $(x_1, x_2, \dots) \mapsto (0, x_1, x_2, \dots)$. T is not normal because $TT^* = id$ but $T^*T = P_U$ for $U = \{(x_i)_i : x_1 = 0\}$.
- T^*T and TT^* are always self-adjoint.

Theorem 5.37 (Hellinger–Toeplitz)**.** Let H be a Hilbert space and let $T : H \to H$ be a linear opeator which satisfies

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

for all $x, y \in H$. Then, $T \in BL(H)$ and moreover T is self-adjoint.

Proof. This is a consequence of the closed graph theorem. Let $x_n \to x$ and $Tx_n \to y$. Then,

$$\langle z, y\rangle = \lim_{n\to\infty} \langle z, Tx_n\rangle = \lim_{n\to\infty} \langle Tz, x_n\rangle = \langle Tz, x\rangle = \langle z, Tx\rangle.$$

Since $z \in H$ was arbitrary we have that $Tx = y$. Thus, the graph of T is closed and by the closed graph theorem, T is bounded. □

Proposition 5.38. Let $T : H \to H$ be a self-adjoint operator. Then

$$\|T\| = \sup_{\|x\|=1} |\langle x, Tx\rangle|.$$

Proof. The non-trivial direction is "$\leq$". It suffices to show that

$$\|T\| = \sup_{\|x\|,\|y\|=1} |\langle x, Ty\rangle| \leq \sup_{\|x\|=1} |\langle x, Tx\rangle|.$$

Moreover, it suffices to consider only the case $\langle x, Ty\rangle \in \mathbb{R}$ where $\|x\| = \|y\| = 1$. Using that T is self-adjoint, we can decompose

$$\langle x, Ty\rangle = \frac{1}{4}\langle x + y, T(x + y)\rangle - \frac{1}{4}\langle x - y, T(x - y)\rangle.$$

Hence,

$$|\langle x, Ty\rangle| \leq \frac{1}{4}(\|x + y\|^2 + \|x - y\|^2) \sup_{\|x\|=1} |\langle x, Tx\rangle = \sup_{\|x\|=1} |\langle x, Tx\rangle|.$$

□

6 Spectral theory

For linear operators (i.e. matrices) $A : V \to V$ in finite dimensions, we say that λ is an eigenvalue of A if $A - \lambda$ (we will use this instead of $A - \lambda I$) is not injective. An eigenvector is a non-zero element in the kernel of $A - \lambda$. In infinite dimensions, $A - \lambda$ can be injective but fail to be surjective. This will lead us to define the concept of the *spectrum* of A which is a generalization of an eigenvalue.

6.1 Spectrum in Banach spaces

Throughout this subsection we assume that X is a complex Banach space.

Definition 6.1. Let $A \in BL(X)$ be a bounded linear operator. We define

- the *resolvent set* of A as
$$\rho(A) = \{\lambda \in \mathbb{C} : (A - \lambda)^{-1} \text{ exists in } BL(X)\}.$$
- the *spectrum* of A as
$$\sigma(A) = \operatorname{spec}(A) = \mathbb{C} \setminus \rho(A).$$
- the *point spectrum* of A as
$$\sigma_p(A) = \{\lambda \in \mathbb{C} : A - \lambda \text{ is not injective}\}$$
- the *continuous spectrum* of A as
$$\sigma_c(A) = \{\lambda \in \mathbb{C} : A - \lambda \text{ injective, not surjective with dense image}\}$$
- the *residual spectrum* of A as
$$\sigma_r(A) = \{\lambda \in \mathbb{C} : A - \lambda \text{ injective, not surjective, image not dense}\}.$$

For $\lambda \in \rho(A)$, we define the *resolvent* operator as
$$R_\lambda(A) = (A - \lambda)^{-1}.$$

Typically, of course not always, σ_p consists of isolated points, σ_c of unions of intervals and σ_r is empty. This also explains the origins of the names. An element $\lambda \in \sigma_p(A)$ is called an *eigenvalue* of A and a non-zero element (which we often normalize to have unit norm) in the kernel of $A - \lambda$ is called *eigenvector* of A. This agrees of course with the usual terminology for finite-dimensional operators (matrices).

Remark 6.2. We note that the spectrum can be written as the disjoint union $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. Indeed, by the inverse mapping theorem Corollary 2.34, if $A - \lambda$ is bijective, then its inverse is a bounded linear operator. (Here we use that X is a Banach space.)

Proposition 6.3. Let $A \in BL(X)$. Then $\sigma(A) = \sigma(A')$. If X is a Hilbert space, then $\sigma(A^*) = \{\bar{\lambda} : \lambda \in \sigma(A)\}$.

Proof. In the case of the Banach space, we note that $A - \lambda I$ is an isomorphism if and only if $(A - \lambda I)' = A' - \lambda I$ is (see problem set). The case for Hilbert spaces follows from $(A^* - \lambda I)^{-1} = \big((A - \bar{\lambda} I)^*\big)^{-1} = \big((A - \bar{\lambda} I)^{-1}\big)^*$. □

Example.
- If $\dim(X) < \infty$, then $\sigma_c = \sigma_r = \emptyset$.

- Let H be a Hilbert space and $A \in BL(H)$ self-adjoint. Then $\sigma(A) \subset \mathbb{R}$ (see already Corollary 6.22) and $\sigma_r(A) = \emptyset$ (note that range$(A-\lambda)^\perp = \ker(A-\lambda)$ so if $A-\lambda$ is injective, then range$(A-\lambda)$ is dense.)

- Let $X = C[0,1]$ and $Af(x) = \int_0^x f(t)dt$. Then, $\sigma(A) = \sigma_r(A) = \{0\}$. (problem set)

Theorem 6.4. Let $A \in BL(X)$. Then:

1. $\rho(A)$ is open.
2. The map $\rho(A) \to BL(X), z \mapsto R_z(A)$ is analytic, i.e. it can be written locally as an absolutely convergent power series.
3. $\sigma(A)$ is compact and non-empty. More, precisely, $\sigma(A) \subset \{z \in \mathbb{C} : |z| \le \|A\|\}$.

Before we prove the theorem above, we will need the following additional lemma.

Lemma 6.5. Let $A \in BL(X)$ and assume that $\|A\| < 1$. Then, $(I-A)^{-1}$ exists and is a bounded linear operator. Moreover,

$$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

Proof. The idea is to use the Neumann series. Recall that for $q \in \mathbb{C}$ with $|q| < 1$ we have $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$. Since $\|A^n\| \le \|A\|^n$ we have (recall problem set 2) that $S = \sum_{n=0}^{\infty} A^n$ converges. To show that indeed the limit is the inverse, we compute

$$(I-A)(I + A + A^2 + \cdots + A^n) = I - A^{n+1} \to I$$

as $n \to \infty$ as $\|A^{n+1}\| \le \|A\|^{n+1} \to 0$ as $n \to \infty$. Thus, $(I-A)S = S(I-A) = I$. $\square$

Proof of Theorem 6.4. 1) Let $\lambda_0 \in \rho(A)$. Then $A-\lambda_0$ is invertible and $c := \|(A-\lambda_0)^{-1}\| < \infty$. Let $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < 1/c$. Then,

$$A - \lambda = A - \lambda_0 + (\lambda_0 - \lambda) = (A-\lambda_0)(I + (\lambda_0 - \lambda)(A-\lambda_0)^{-1})$$

Since $A - \lambda_0 I$ is invertible, we have that $A - \lambda I$ is invertible if $B = I + (\lambda_0 - \lambda)(A-\lambda_0)^{-1}$ is invertible. But B is invertible by Lemma 6.5 because $|\lambda_0 - \lambda| \, \|(A-\lambda_0 I)^{-1}\| < 1$. In particular, from Lemma 6.5 we have that

$$(A-\lambda)^{-1} = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n ((A-\lambda_0)^{-1})^{n+1}$$

which shows 2).

For 3) we first note that for $|\lambda| > \|A\|$ we have that $(A-\lambda) = \lambda(\lambda^{-1}A - I)$ is invertible in $BL(H)$ by Lemma 6.5. Hence, $\sigma(A)$ is closed and bounded (i.e. compact) and in addition $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \le \|A\|\}$.

It remains to show that the spectrum is non-empty. For a contradiction, assume that $\rho(A) = \mathbb{C}$. Then, the map $\mathbb{C} \to BL(X) : \lambda \mapsto R_\lambda(A)$ is analytic on $\mathbb{C}$ as it can be written locally as a power series. Consider now a functional $f \in (BL(X))^*$. Then, $f(R_\lambda(A)) : \mathbb{C} \to \mathbb{C}$ is entire because it can be locally written as

$$f(R_\lambda(A)) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n f\left(R_{\lambda_0}(A)^{n+1}\right).$$

Moreover, it is bounded because for $|\lambda| > 2\|A\|$ we have, using $(A - \lambda) = \lambda(\lambda^{-1}A - I)$ that

$$|f(R_\lambda(A))| \leq \|f\||\lambda|^{-1} \sum_{n=0}^{\infty} \left(\frac{\|A\|}{|\lambda|}\right)^n \leq 2\|f\||\lambda|^{-1} \tag{19}$$

and for $|\lambda| \leq 2\|A\|$, it is bounded because it is a compact set. Hence, by Liouville's theorem, $f(R_\lambda(A))$ is constant in λ for all f. Thus, by the Hahn–Banach theorem, $(A - I)^{-1} = A^{-1}$ but this gives a contradiction. $\square$

Definition 6.6. We define the *spectral radius* of a bounded operator $A \in BL(X)$ as

$$r(A) := \inf_{n \in \mathbb{N}} \|A^n\|^{\frac{1}{n}} = \lim_{n \to \infty} \|A^n\|^{\frac{1}{n}}.$$

The above limit exists due to the following

Lemma 6.7. Let $(a_n)_n$ be a sequence of real numbers and assume that $0 \leq a_{n+m} \leq a_n a_m$ for all $n, m \in \mathbb{N}$. Then $\lim_{n\to\infty} a_n^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} a_n^{\frac{1}{n}}$.

Proof. Clearly we have $\leq$ so it suffices to show $\geq$. Set $a \doteq \inf_{n\in\mathbb{N}} a_n^{\frac{1}{n}}$. For $\varepsilon > 0$ choose $N \in \mathbb{N}$ such that $a_N^{\frac{1}{N}} < a + \varepsilon$ and set $b := \max\{a_1, \dots a_N\}$. For a natural number n we write $n = kN + r$ for $1 \leq r \leq N$. Then,

$$a_n^{\frac{1}{n}} = a_{kN+r}^{\frac{1}{n}} \leq (a_N^k a_r)^{\frac{1}{n}} \leq (a+\varepsilon)^{kN/n} b^{\frac{1}{n}} = (a+\varepsilon)^{1-r/n} b^{\frac{1}{n}} \to a + \varepsilon$$

as $n \to \infty$. $\square$

The following theorem justifies the name spectral radius for $r(A)$.

Theorem 6.8. Let $A \in BL(X)$ and $\lambda \in \sigma(A)$. Then

$$r(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Proof. "$\geq$": We will first show that $|\lambda| \leq r(A)$ for all $\lambda \in \sigma(A)$. To this end we have to show that for $|\lambda| > r(A)$, $A - \lambda = \lambda(\lambda^{-1}A - 1)$ is invertible. Similar to the argument before, it suffices to show that $\sum_{n=1}^{\infty} \left(\frac{A}{\lambda}\right)^n$ converges. But this converges because for all n sufficiently large $q|\lambda| > \|A^n\|^{\frac{1}{n}}$ for some $q < 1$, i.e. $\|\left(\frac{A}{\lambda}\right)^n\| \leq q^n$ for n sufficiently large, which shows that the series converges.

"$\le$": Let $r_0 = \max\{|\lambda| : \lambda \in \sigma(A)\}$ (which is attained by compactness of the spectrum). Let $|\mu| > r_0$. We will then show that $|\mu| \ge r(A)$ which shows that $r_0 \ge r(A)$, i.e. $r_0 = r(A)$. For $f \in (BL(X))^*$, consider in the set $\{\lambda : |\lambda| > r_0\}$ the analytic function

$$F_f(\lambda) = f((A - \lambda)^{-1}).$$

For $|\lambda| > r(T)$ we have

$$F_f(\lambda) = -\sum_{n=0}^{\infty} f(A^n)\lambda^{-n-1}.$$

This Laurent series converges in the largest open annulus in which F_f is analytic. In particular, it converges uniformly at μ. In particular $\lim_{n\to\infty} f(A^n/\mu^{n+1}) = 0$. Since f was arbitrary, we have that A^n/μ^{n+1} converges weakly to zero. Thus, it is bounded. In particular, there is a $C > 0$ such that

$$\|A^n\|^{1/n} \le C^{1/n}|\mu|^{1+1/n} \to |\mu|$$

which shows that $r(A) \le |\mu|$. $\square$

Corollary 6.9. If H is a Hilbert space and $A \in BL(H)$ is normal. Then $\lim_{n\to\infty} \|A^n\|^{1/n} = \|A\|$ and thus $r(A) = \|A\|$.

Proof. We have $||A||^2 = \|AA^*\|$ and $\|A^2\|^2 = \|A^2(A^2)^*\| = \|AA^*(AA^*)^*\| = \|AA^*\|^2 = \|A\|^4$, i.e. $\|A^2\| = \|A\|^2$. By induction, we have $\|A^{2k}\| = \|A\|^{2k}$ and thus $\lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \|A^{2k}\|^{1/2k} = \|A\|$. $\square$

6.2 Compact operators

Throughout this section, X, Y are Banach spaces and H is a Hilbert space

Definition 6.10. Let $K \in BL(X, Y)$. We say that K is *compact* if for every bounded subset $B \subset X$, the closure of the image $\overline{K(B)}$ is compact, i.e. $K(B)$ is relatively compact. We define the set of compact operators as $K(X, Y) \subset BL(X, Y)$.

Proposition 6.11. An operator $K \in BL(X, Y)$ is compact if and only if for every bounded sequence $(x_n)_n \subset X$, the image $K(x_n)$ has a convergent subsequence.

Proof. "$\Rightarrow$": This follows from the equivalence of compactness and sequential compactness. "$\Leftarrow$": Assume that M is bounded. We want to show that $T(M)$ is relatively compact. Let $(y_n)_n$ be a sequence in $T(M)$. Then $y_n = Tx_n$ for a bounded sequence x_n. Thus, y_n has a convergent subsequence. Since $(y_n)_n$ was arbitrary, we have that $T(M)$ is relatively compact. $\square$

In order to present an important example of a compact operator, namely that of certain integral operators, we will prove an important compactness criterion for elements of the Banach space $C(K)$ for a compact Hausdorff space K.

Theorem 6.12 (Arzelà–Ascoli)**.** Let (K, d) be a compact metric space and $M \subset C(K)$ be a subset of continuous functions. Then M is compact if and only if M satisfies

1. M is closed,
2. M is uniformly equicontinuous, i.e. for all $\varepsilon > 0$, there is an $\delta > 0$ such that for all $x, y \in K$, $d(x, y) < \delta$ implies $|f(x) - f(y)| \leq \varepsilon$ for all $f \in M$.
3. M is pointwise bounded, i.e. $\sup_{f \in M} |f(x)| < \infty$, for all $x \in K$.

Proof. "$\Leftarrow$" This is an application of a Cantor diagonal argument. Since $C(K)$ is a metric space, it suffices to show sequential compactness. We let $(f_n)_n$ be a sequence of functions in M. Since K is a compact metric space, it is separable. We let $(x_m)_m$ be a countable dense set in K. We will show that $(f_n)_n$ has a subsequence which is a Cauchy sequence. For $m = 1$ we consider the sequence $(f_n(x_1))_n$. This sequence is a bounded sequence in $\mathbb{K}$ and thus has a convergent subsequence $(f_{n_{k_1}}(x_1))_{k_1}$. For $m = 2$ we consider the sequence $(f_{n_{k_1}}(x_2))_{k_1}$ in $\mathbb{K}$. Again, this has a convergent subsequence which we denote by $(f_{n_{k_2}}(x_2))_{k_2}$. We continue inductively to construct subsequences $(f_{n_{k_i}})_{k_i}$ for $i \geq 1$. We then choose the diagonal sequence $(f_{n_{\{k_1=1\}}}, f_{n_{\{k_2=2\}}}, \dots)$ which we denote by $(f_{n_j})_j$. By construction, the diagonal sequence $(f_{n_j})_j$ converges for all x_m.

We will now show that the diagonal sequence is a Cauchy sequence. Let $\varepsilon > 0$ and let $\delta > 0$ be as in the statement of the uniform equicontinuity. Consider the coverings $B_\delta(x_m)$ of δ-balls around the dense set $(x_m)_m$. By compactness of K, we can extract a finite subcover with δ balls centered at $(x_{m_1}, \dots x_{m_r})$. Since $f_{n_j}(x_{m_1}), \dots f_{n_j}(x_{m_r})$ are a Cauchy sequences, there exists a $j_0(\varepsilon)$ such that $|f_{n_j}(x_{m_1}) - f_{n_l}(x_{m_1})|, \dots, |f_{n_j}(x_{m_r}) - f_{n_l}(x_{m_r})| < \varepsilon$ for all $j, l \geq j_0(\varepsilon)$.

Let $x \in K$, and let an x_{m_o}, $o = 1, \dots, r$ such that $d(x, x_{m_0}) < \delta$. Then,

$$|f_{n_j}(x) - f_{n_i}(x)| \leq |f_{n_j}(x) - f_{n_j}(x_{m_0})| + |f_{n_j}(x_{m_0}) - f_{n_i}(x_{m_0})| + |f_{n_i}(x_{m_0}) - f_{n_i}(x)| \leq 3\varepsilon$$

for all $i, j \geq j_0(\varepsilon)$. Hence, f_{n_j} is a Cauchy sequence and thus convergent as $C(K)$ is complete.

"$\Rightarrow$" Since M is compact, then it is closed because $C(K)$ is Hausdorff and moreover bounded as $C(K)$ is a metric space. It remains to show the uniform equicontinuity. Let $\varepsilon > 0$. Since K is compact, we have that every $f \in M$ is uniformly continuous. Thus, for every $f \in M$, there exists a $\delta > 0$ such that $d(x, y) < \delta$ implies $|f(x) - f(y)| < \varepsilon$. Now cover M by $(B_\varepsilon(f))_{f \in M}$ and extract a finite subcover with center at $f_1, \dots f_n$. Then, choose $\delta > 0$ sufficiently small such that $|f_j(x) - f_j(y)| \leq \varepsilon$ whenever $d(x, y) < \delta$ and $j = 1, \dots n$. Then, let $x, y \in K$ with $d(x, y) < \delta$ and $f \in M$. Choose $f_j \in M$ such that $\|f - f_j\| \leq \varepsilon$.. Then,

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| \leq 3\varepsilon.$$

This shows that M is equicontinuous. $\square$

Example.
- Let $k : [0,1]^2 \to \mathbb{R}$ be continuous. Define an operator $T_k : L^2([0,1]) \to L^2([0,1])$ as
$$f \mapsto (T_k f)(x) := \int_0^1 f(y)k(x,y)dy.$$
You will show in the problem set that T_k is a compact operator.

- Similarly, $T_k : C[0,1] \to C[0,1]$ as defined above is also compact, which follows from the Arzelà–Ascoli theorem.

- Let $T \in BL(X,Y)$. Then T is compact if T has finite rank, i.e. if $\dim(\text{range}(T)) < \infty$.

A feature of compact operators is that they map weakly convergent sequences to strongly convergent sequences.

Proposition 6.13. Let $K \in BL(X,Y)$ be a compact operator. If $x_n \rightharpoonup x$ as $n \to \infty$ then $K(x_n) \to K(x)$ as $n \to \infty$.

Proof. Let $x_n \rightharpoonup x$. We want to show that $K(x_n) \to K(x)$ as $n \to \infty$. It suffices to show that every subsequence of $K(x_n)$ has a subsequence which converges to $K(x)$. Let $(x_{n_k})_k$ be a subsequence. Since $(x_{n_k})_k$ is weakly convergent, we have by the compactness of K that $K(x_{n_{k_j}}) \to y$ for some subsequence $(x_{n_{k_j}})_j$. Let $f \in Y^*$. Then, $f \circ K \in X^*$ so $f(K(x_{n_{k_j}})) \to f(K(x))$ but also $f(K(x_{n_{k_j}})) \to f(y)$ so $K(x) = y$. This shows that $K(x_n)_n \to K(x)$ for the original sequence. $\square$

Theorem 6.14. The set $K(X,Y) \subset BL(X,Y)$ is a closed subspace.

Proof. Clearly, $K(X,Y)$ is a subspace. The proof of the closedness of $K(X,Y)$ requires a diagonal argument. Let $T \in BL(X,Y)$ and $T_n \to T$ where $T_n \in K(X,Y)$. Let $(x_m)_m$ be a bounded sequence and assume WLOG that $\|x_m\| \leq 1$. Since T_1 is compact, we have that $(T_1(x_m))_m$ has a convergent subsequence which we denote with $(T_1 x_{m_1})_{m_1}$. Since T_2 is compact, we can again extract a further subsequence $(x_{m_2})_{m_2}$. We continue inductively and will denote the diagonal sequence by $(x_{m_j})_j$. In particular, we have that $(T_n x_{m_j})_j$ converges for all n. Thus, for $\varepsilon > 0$ let $N \in \mathbb{N}$ such that $\|T - T_N\| \leq \varepsilon/3$. Then
$$\|Tx_{m_j} - Tx_{m_i}\| \leq \|Tx_{m_j} - T_N x_{m_j}\| + \|T_N x_{m_j} - T_N x_{m_i}\| + \|T_N x_{m_i} - Tx_{m_i}\| \leq \varepsilon$$
for i, j sufficiently large only depending on ε. Thus, $(Tx_{m_j})_j$ is Cauchy and thus convergent. $\square$

Corollary 6.15. The limit of a sequence of finite rank operators is a compact operator.

Remark 6.16. The converse to the above corollary also holds true in the case of Hilbert spaces and has been a long-standing open problem for Banach spaces. Enflo showed in 1973 that this property is false in general Banach spaces. He received a living goose as an award from Mazur. Banach spaces with that property are said to have the approximation property.

Proposition 6.17. If $K \in K(X)$ and $T \in BL(X)$, then $KT, TK \in K(X,Y)$. In particular, this makes $K(X) \subset BL(X)$ a closed, 2-sided ideal.

Proof. If x_n is a bounded sequence, then so is Tx_n and thus KTx_n has a convergent subsequence. Furthermore, Kx_n has a convergent subsequence but then also TKx_n has a convergent subsequence. □

Theorem 6.18 (Schauder's theorem)**.** An operator $T \in BL(X,Y)$ is compact if and only if $T' \in BL(Y^*, X^*)$ is compact.

Proof. "$\Rightarrow$": Let f_n be a bounded sequence in Y^*. Since T is compact we have that $Z := \overline{T(B_1^X(0))}$ is a compact set. Consider the sequence $f_n|_Z$ restricted to Z and note that f_n is bounded and equicontinuous $|f_n(z) - f_n(\tilde{z})| \leq \sup_n \|f_n\| \|z - \tilde{z}\| \leq C\|z - \tilde{z}\|$. Thus, the closure of $\{f_n : n \in \mathbb{N}\}$ is compact and hence, f_n has a convergent subsequence (f_{n_k}) by Arzelà–Ascoli. Hence,

$$\|T' f_{n_k} - T' f_{n_l}\| = \sup_{\|x\|=1\}} \|f_{n_k}(T(x)) - f_{n_l}(T(x))\| \leq \|f_{n_k} - f_{n_l}\|_\infty \to 0$$

as $k, l \to \infty$. Hence, T' is compact.

"$\Leftarrow$": If T' is compact so is T'' and thus also $T'' \circ i_X$, where $i_X : X \to X^{**}$ is the canonical embedding. By a direct computation, we observe that $T'' \circ i_X = i_Y \circ T$. Thus, the operator $i_Y \circ T : X \to Y''$ is compact. Since Y is closed within Y'' we have that $T : X \to Y$ is compact. □

Corollary 6.19. The set of compact operators $K(H) \subset BL(H)$ on a Hilbert space is a closed, $*$-closed, 2-sided ideal in the set of bounded operators.

Proof. If $K \in K(H)$ is compact, then so it K'. But then also $K^* = \mathcal{A}T'\mathcal{A}^{-1}$ is compact which follows from the same argument as in the proof of Proposition 6.17. □

Example. Let T be a diagonal operator on a separable Hilbert space, i.e. there exists an orthonormal basis $(e_i)_{i\in\mathbb{N}}$ such that $Te_i = \alpha_i e_i$. Assume that $\alpha_i \to 0$ as $i \to \infty$. T is compact because it is the uniform limit of T_n which is defined as $T_n e_i = \alpha_i e_i$ for $i \leq n$ and zero otherwise. Indeed, $\|T - T_n\| \leq \sup_{j\geq n} |\alpha_j| \to 0$ as $n \to \infty$.

6.3 Spectral theorem for compact self-adjoint operators

In the following, we will focus on compact self-adjoint operators on Hilbert spaces. We let H be a Hilbert space.

Definition 6.20. For $T \in BL(H)$ we define the *numerical range* as $W(T) = \{\langle x, Tx\rangle : \|x\| = 1\}$.

Theorem 6.21. For $T \in BL(H)$ we have $\sigma(T) \subset \overline{W(T)}$.

Proof. Let $\lambda \notin \overline{W(T)}$. Thus, there exists $d > 0$ such that for all $\|x\| = 1$ we have

$$d \le |\lambda - \langle x, Tx\rangle| \le \|(\lambda - T)x\|.$$

Thus $\lambda - T$ is injective and an isomorphism onto its image with the norm of the inverse bounded by $1/d$. Thus, range$(\lambda - T)$ is closed (check it!). It suffices to show that the range of $\lambda - T$ has trivial orthogonal complement. Indeed, if $x \perp$ range$(\lambda - T)$, then $\langle x, (\lambda - T)y\rangle$ for all $y \in H$. Thus, for $y = x$ we have

$$\lambda\langle x, x\rangle = \langle x, Tx\rangle,$$

but this can only happen if $x = 0$ by definition of the numerical range of T. □

Corollary 6.22. Let $T \in BL(H)$ be self-adjoint. Then, $\sigma(T) \subset \mathbb{R}$.

Theorem 6.23. Let $T \in BL(H)$ be a compact self-adjoint operator. Then at least one of $-\|T\|$ and $\|T\|$ is an eigenvalue of T.

Proof. WLOG $T \neq 0$. From Proposition 5.38 we can find unit vectors $(x_n)_n$ such that $\|T\| = \lim_{n\to\infty} |\langle x_n, Tx_n\rangle|$. After passing to a subsequence and since T is self-adjoint, we have that $\langle x_n, Tx_n\rangle \to \lambda$, where $\lambda = \|T\|$ or $\lambda = -\|T\|$. We compute

$$\|Tx_n - \lambda x_n\|^2 = \|Tx_n\|^2 - 2\lambda\langle x_n, Tx_n\rangle + \lambda^2 \le 2\lambda^2 - 2\lambda\langle x_n, Tx_n\rangle \to 0$$

as $n \to \infty$. Since T is compact, Tx_n has a convergent subsequence to say y, and thus λx_n also has a convergent subsequence to y. Then, $Tx_{n_k} = \lambda^{-1}T\lambda x_{n_k} \to \lambda^{-1}Ty$. But then $y = \lambda^{-1}Ty$ or $Ty = \lambda y$. To conclude that $y \neq 0$ we note that $\|y\| = \lim_{n\to\infty} \|\lambda x_{n_k}\| = |\lambda| \neq 0$. □

Theorem 6.24. Let $T \in BL(H)$ be self-adjoint. Then, the eigenvectors for distinct eigenvalues are orthogonal.

Proof. Let x, y be unit eigenvectors for different eigenvalues λ, μ, i.e. $Tx = \lambda x$ and $Ty = \mu y$. Then,

$$\mu\langle x, y\rangle = \langle x, \mu y\rangle = \langle x, Ty\rangle = \langle Tx, y\rangle = \lambda\langle x, y\rangle$$

and hence $(\mu - \lambda)\langle x, y\rangle = 0$ from which we obtain the claim. □

Theorem 6.25. Let $T \in BL(H)$ be a compact self-adjoint operator. Then, the set of eigenvalues of T is a finite or countably infinite set of real numbers. If it is infinite, then the eigenvalues form a sequence that converges to zero.

Proof. Assume that the set of eigenvalues is infinite. Let $\varepsilon > 0$ and we will show that at most finitely many eigenvalues exists with absolute value larger than ε. Assume for a contradiction that this is not the case and there exists an infinite sequence $(\lambda_i)_i$ of distinct eigenvalues with $|\lambda_i| \ge \varepsilon$. Let x_i be a corresponding sequence of unit eigenvectors. Then, using the orthogonality of the eigenvectors, we have

$$\|Tx_i - Tx_j\|^2 = \|\lambda_i x_i - \lambda_j x_j\|^2 = |\lambda_i|^2\|x_i\|^2 + |\lambda_j|^2\|x_j\|^2 \ge 2\varepsilon^2.$$

However, this cannot happen as Tx_i has a convergent subsequence by compactness of T. Clearly, this also shows that $\{\lambda : \lambda \text{ is an eigenvalue}\}$ is countable. □

Proposition 6.26. Let $T \in BL(H)$ and $M \subset H$ be a closed subspace. If M is an *invariant subspace* of T (i.e. $Tm \in M$ for all $m \in M$), then $M^\perp$ is an invariant subspace for T^*. Conversely, if $M^\perp$ is an invariant subspace for T^*, then M is an invariant subspace for T. In particular, if T is self-adjoint and M is an invariant subspace of T, then $M^\perp$ is also an invariant subspace of T.

Proof. Let $x \in M^\perp$. Let $m \in M$. Then,

$$\langle m, T^*x \rangle = \langle Tm, x \rangle = 0$$

because $Tm \in M$ and $x \in M^\perp$. Hence, $T^*x \in M^\perp$. The second statement follows from $T^{**} = T$ and $M^{\perp\perp} = M$. $\square$

Remark 6.27. The question of whether every bounded linear operator on a Banach space admits a non-trivial invariant subspace has sparked interest. It is known as the invariant subspace problem. Counterexamples to this statement have been constructed for Banach spaces by Enflo. Whether the statement is true for Hilbert spaces is still open; however, a recent preprint of Enflo suggests an affirmative answer: arXiv:2305.15442.

Theorem 6.28 (Spectral theorem for compact self-adjoint operators)**.** Let $T \in BL(H)$ be a compact, self-adjoint operator. There exist a countable orthonormal sequence $(e_n)_n$ of eigenvectors of T and sequence $(\lambda_n)_n$ in $\mathbb{C} \setminus \{0\}$ which converges to zero such that

$$Tx = \sum_n \lambda_n \langle e_n, x \rangle e_n$$

and

$$H = \ker(T) \oplus \overline{\operatorname{span}}(e_1, e_2, \dots).$$

In particular, the λ_i are the non-zero eigenvalues and the e_i are the corresponding eigenvectors. Moreover, $\|T\| = \sup_n |\lambda_n|$.

Proof. Let $(\lambda_n)_n$ be the sequence of non-zero eigenvalues counted by multiplicities, e.g. $\lambda_1 = \lambda_2 = \cdots = \lambda_{d_1}$, where d_1 is the dimension of $\ker(T - \lambda_1)$, i.e. the geometric multiplicity of λ_1. Note that d_1, is indeed finite because if $(x_n)_n$ is a sequence in the unit ball of $\ker(T - \lambda_1)$, then $x_n = \frac{1}{\lambda_1} T x_n$ has a convergent subsequence because T is compact. Hence, then unit ball of $\ker(T - \lambda_1)$ is compact and thus $\ker(T - \lambda_1)$ is finite dimensional (Theorem 2.10). The same argument also extends to $d_2, \dots$. For each d_i-dimensional subspace we choose an orthonormal basis which we gather as a sequence as $(e_n)_n = (e_1^1, \dots, e_{d_1}^1, e_1^2 \dots)$. By Theorem 6.24 this sequence defines an orthonormal set and moreover

$$Te_k = \lambda_k e_k$$

for all $k \in \mathbb{N}$. Moreover, $\ker(T) \perp e_k$ for all $k \in \mathbb{N}$ and thus $H_1 := \ker(T) \oplus \overline{\operatorname{span}}(e_1, e_2, \dots)$ is a closed subspace of H which is an invariant subspace of T. It remains to show that $H = H_1$. Define $H_2 = H_1^\perp$. By Proposition 6.26 we have that H_2 is an invariant subspace of T because H_1 is an invariant subspace of T. Hence, $T|_{H_2}$ is a compact operator on H_2.

If $T|_{H_2} = 0$ then, $H_2 \subset \ker(T) \subset H_1$. If $T|_{H_2} \neq 0$, then, by Theorem 6.23, there exist $x \in H_2$ such that x is an eigenvector to a non-zero eigenvalue with magnitude $\|T|_{H_2}\|$. But then $x \in H_1 \cap H_2$ and thus $x = 0$, giving a contradiction. Hence $H = H_1$ and every $x \in H$ can be decomposed into

$$x = y + \sum_k \langle e_k, x \rangle e_k,$$

where $y \in \ker(T)$. Since T is continuous, we obtain

$$Tx = \sum_k \lambda_k \langle e_k, x \rangle e_k.$$

The claim about the norm follows from Theorem 6.23 and Theorem 6.4. □

6.4 Fredholm alternative

In this section, we will study operators of the form $T - I$, where T is a compact operator on a Hilbert space H. Here, we will not assume that T is self-adjoint. We will set

$$S := T - I.$$

Proposition 6.29. We have that the operator S as defined above satisfies:

1. $\ker(S)$ is finite dimensional.
2. S has closed range and $\dim(\mathrm{range}(S)^\perp) < \infty$.

Proof. Note that $\ker(S)$ has a compact unit ball because for any sequence $(x_n)_n$ in the unit ball of $\ker(S)$ we have $x_n = Tx_n$, which has a convergent subsequence as T is compact. Thus, $\ker(S)$ is finite-dimensional by Theorem 2.10.

In order to show that S has closed range we consider $\tilde{S} : H/\ker(S) \to H$ which is an injective bounded linear operator. The range of S is the same as the range $\tilde{S}$. We note that an injective operator has closed range if and only if it is bounded from below.

Assume not, then there exists a sequence $([x_n])_n$ in $H/\ker(S)$ such that $\|[x_n]\| = 1$ and such that $\tilde{S}[x_n] \to 0$. Thus, there exist a $y_n \in \ker(S)$ such that $\|x_n + y_n\| \leq 2$ and $S(x_n + y_n) \to 0$. Since T is compact, $x_n + y_n = T(x_n + y_n)$ has a convergent subsequence. Thus, (upon relabeling the sequence) $x_n + y_n \to y$ and $y = Ty$, hence, $y \in \ker(S)$. This however is a contradiction because $0 = \|[y]\| = \lim_{n\to\infty} \|[x_n + y_n]\| = \lim_{n\to\infty} \|[x_n]\| = 1$, where we used that the projection map is continuous. This shows that S has closed range.

In order to show that $\dim(\mathrm{range}(S)^\perp)$ is finite, we note that $\mathrm{range}(S)^\perp = \ker(S^*)$ by Proposition 5.35 and that $\ker(S^*) = \ker(T^* - I)$ is finite dimensional by the previous argument. □

Proposition 6.30. Define $M_j = \mathrm{range}((T - I)^j)$ for $j \in \mathbb{N}$. Then, there exists a positive integer j_0 such that $M_{j_0} = M_j$ for all $j \geq j_0$.

Proof. By expanding $(T-I)^j$ we observe that $(T-I)^j$ satisfies the assumptions of Proposition 6.29. Thus, M_j is closed for all $j \geq 1$. Clearly, $M_j \supset M_{j+1}$. Suppose that for every step, the containment is strict. Then, the quotients M_j/M_{j+1} have dimensions greater or equal to one. In particular, we can find an $x_j \in M_j$ such that $\|[x_j]\|_{M_j/M_{j+1}} = 1$ and we can assume without loss of generality that $\|x_j\| \leq 2$ for each j. We will show that $\|Tx_j - Tx_l\| \geq 1$ for $j \neq l$ contradicting the fact that T is compact. Let $j < k$, then $x_k \in M_k \subset M_{j+1}$, $(T-I)x_j \in M_{j+1}$ and $(T-I)x_k \in M_{k+1} \subset M_{j+1}$. But then

$$Tx_j - Tx_k = (T-I)x_j - x_k - (T-I)x_k + x_j$$

Since $(T-I)x_j - x_k - (T-I)x_k \in M_{j+1}$ and $x_j \in M_j$ we have that $\|Tx_j - Tx_l\| \geq 1$. This gives a contradiction and concludes the proof. □

Theorem 6.31 (Fredholm alternative)**.** Let $T \in BL(H)$ be a compact operator and $\lambda \in \mathbb{C} \setminus \{0\}$. Then:

1. If $T - \lambda$ is injective, then $T - \lambda$ is invertible with $(T-\lambda)^{-1} \in BL(H)$.
2. If $T - \lambda$ is surjective, then $T - \lambda$ is invertible with $(T-\lambda)^{-1} \in BL(H)$.

Proof. We will prove the first statement first. By dividing by λ we can assume WLOG that $\lambda = 1$. If $T-I$ is injective, then $\ker(T-I) = \{0\}$. We note that $T-I$ has closed range by Proposition 6.29. If the range of $T-I$ was not all of H, we have that $\operatorname{range}(T-I) \subset H$ is a closed proper subspace. But this then means that $\operatorname{range}((T-I)^2) \subset \operatorname{range}(T-I)$ is again a closed proper subspace. Continuing inductively, however, contradicts Proposition 6.30. Thus, $T-I$ is surjective and by the inverse mapping theorem, we have that $T-I$ is invertible.

For the second statement we note that if $T-\lambda$ is surjective, then $(T-\lambda)^* = T^* - \bar{\lambda}$ is injective. But this also means that $T-\lambda$ is invertible. □

Remark 6.32.

- The statements can be interpreted as follows. If we consider the equation $(T-\lambda)x = y$ for an unknown x, then 1) says that "uniqueness" implies "existence" and 2) says that "existence" implies "uniqueness".
- It is called "alternative" because by 1), either λ is an eigenvalue or $T-\lambda$ is invertible in $BL(H)$.
- Moreover, the results can be extended to Banach spaces.

Remark 6.33. One can extend the result and show that $\dim\ker(T-\lambda) = \dim(\operatorname{range}(T-\lambda))^{\perp}$. An operator S satisfying

1. $\ker(S)$ is finite dimensional.
2. S has closed range and $\operatorname{codim}(\operatorname{range}(S)) := \dim(X/\operatorname{range}(S)) < \infty$.

is called a *Fredholm operator* and the number $\operatorname{ind}(S) = \dim(\ker(S)) - \operatorname{codim}(\operatorname{range}(S))$ is called the *Fredholm index*. Thus, according to Remark 6.32, the operator $S = T - I$ for T compact is a Fredholm operator with index 0. For such an operator with index zero, we have that injectivity is equivalent to surjectivity.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 1

Due date: 2/14/2025

1. Let X be a topological space and $A \subset X$. Prove the following:
   (a) $\mathrm{int}(A) = A \setminus \partial A$.
   (b) $\mathrm{int}(A)$ is the largest open set contained in A.
   (c) $\overline{A}$ is the smallest closed set containing A.
   (d) A closed $\Leftrightarrow A = \overline{A}$.
   (e) $\partial A = \bar{A} \setminus \mathrm{int}(A)$ and deduce that ∂A is closed.
   (f) If X is Hausdorff, then limits are unique, i.e. $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} x_n = y$ implies $x = y$. Is it still true if the Hausdorff assumption is dropped?
   (g) A collection of sets $\mathcal{A} \subset \mathcal{P}(X)$ is said to have the *finite intersection property (FIP)* if $\cap_{1\leq j\leq n} U_j \neq \emptyset$ for any finite subfamily $\{U_j\}_{1\leq j\leq n} \subset \mathcal{A}$. Prove that X is compact if and only if every family of closed sets having the FIP has non-empty intersection.

2. Let X, Y be topological spaces and $f : X \to Y$ be a function. Prove the following:
   (a) If X is Hausdorff and $K \subset X$ compact, then K is closed.
   (b) If f is continuous and $K \subset X$ is compact, then $f(K)$ is compact.
   (c) If X is compact, Y is Hausdorff and f continuous and bijective, then f is a homeomorphism.
   (d) X is Hausdorff if and only if the diagonal $\Delta = \{(x,x) : x \in X\} \subset X \times X$ is closed.
   (e) The projection $\pi_X : X \times Y \to X$ is an open map, i.e. it maps open sets to open sets.
   (f) If Y is Hausdorff and f is continuous, then the graph $\mathrm{graph}(f) = \{(x, f(x)) : x \in X\} \subset X \times Y$ is closed in the product topology.
   (g) If X and Y are compact Hausdorff spaces and f has a closed graph, then f is continuous.
   Note: The statement even holds true if X is merely a topological space. It is known as the *topological closed graph theorem.*

3. Let (X, d) be a metric space. Recall that $B_\varepsilon(x) = \{y : d(x,y) < \varepsilon\}$ and $\bar{B}_\varepsilon(x) = \{y : d(x,y) \leq \varepsilon\}$.
   (a) Show that there exists another metric d' on X which satisfies $d'(x,y) \leq 1$ for all $x, y \in X$ and which induces the same topology as (X, d).
   (b) Construct two different metrics d_1 and d_2 on $X = \mathbb{R}$ such that (X, d_1) and (X, d_2) induce the same topology but such that (X, d_1) is complete, whereas (X, d_2) is not.
   (c) Let $D \subset X$ be dense. Prove that if every Cauchy sequence $(x_n)_n \subset D$ has a limit in X, then X is complete.
   (d) Prove that if X is separable, then it is second countable.
   Note: This shows that for metric spaces separability and second countability are equivalent. However, there exist topological spaces (e.g. the Sorgenfrey line) which are separable but not second countable.

4. Consider the vector space of continuous functions on the unit interval $X = C([0,1])$.
   (a) Show that $d(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$ defines a metric on X.
   (b) Show that piecewise linear functions are dense in (X, d). Note that a function $f \in X$ is piecewise linear (or piecewise affine) if there exists $0 = x_1 < x_2 < \cdots x_n = 1$ such that $f|_{(x_i, x_{i+1})}$ is affine for all $1 \leq i \leq n-1$.
   (c) Show that X is separable. *Hint:* Construct a countable family of piecewise linear functions which are dense in X.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 2

Due date: 2/24/2025

1. Recall the definition of ℓ^p, c_c and c_0 and the Hölder inequality from the lectures.
   (a) Prove the Minkowski inequality: $\|x+y\|_p \leq \|x\|_p + \|y\|_p$ for $x, y \in \ell^p, 1 \leq p < \infty$.
   (b) Show that $\ell^p \subsetneq \ell^q$ for $1 \leq p < q \leq \infty$. In particular, show that $\|x\|_q \leq \|x\|_p$ for all $x \in \ell^p$.
   (c) For which $q \in (1, \infty]$ is the following true: $\bigcup_{p\in[1,q)} \ell^p = \ell^q$.
   (d) Show that ℓ^∞ and c_0 are complete and that c_c is not complete.
   (e) Show that $c_c \subsetneq \ell^p$ for all $1 \leq p \leq \infty$ and characterize the closure $\overline{c_c} \subset \ell^p$ in the respective metrics.
   (f) Show that ℓ^p is separable if $p \in [1, \infty)$ and show that ℓ^∞ is not separable.
   (g) Is $c_0 = \cup_{p\in[1,\infty)}\ell^p$?

2. (a) Let (X, d) be a metric space. Prove that $K \subset X$ is compact if and only if K is sequentially compact.
   Hint: For the direction "$\Leftarrow$" you may want to show the following intermediate steps:
   (i) For any open covering $(U_i)_{i\in I}$ of K there exists a $\varepsilon > 0$ (depending only on K and $(U_i)_{i\in I}$) such that for all $x \in K$: $B_\varepsilon(x) \subset U_{i_x}$ for some $i_x \in I$. (The number ε is also called the *Lebesgue number* of the covering $(U_i)_{i\in I}$.)
   (ii) Finitely many of such $B_\varepsilon(x)$ balls already cover K.
   (b) Let X be a non-empty, complete, and countable metric space. Show that there exist a $x \in X$ and $\varepsilon > 0$ such that $\{x\} = B_\varepsilon(x)$.
   (c) Let $f : [0, \infty) \to \mathbb{R}$ be a continuous function such that for all $x_0 > 0$: $\lim_{n\to\infty} f(nx_0) = 0$. Prove that $\lim_{x\to\infty} f(x) = 0$.
   (d) *Optional ungraded exercise*: Let $f \in C^\infty(\mathbb{R})$ have the property that for every $x \in \mathbb{R}$ there exists a $k \in \mathbb{N}$ depending on x such that $\frac{d^k f}{dx^k}(x) = 0$. Show that f is a polynomial.

3. Prove that $C^1[0, 1] = \{f \in C[0, 1] : f \text{ is continuously differentiable}\}$ is incomplete with respect to the norm $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$ but complete with respect to the norm $\|f\|_\infty + \|f'\|_\infty$.

4. (a) Construct a discontinuous linear map $T : c_c \to c_c$.
   (b) Let X be an infinite dimensional Banach space. Show that there exists a discontinuous linear map $T : X \to X$. *Hint:* Use a Hamel basis.

5. (a) Prove that a normed space X is complete if and only if $\sum_{n=1}^\infty \|x_n\| < \infty$ implies that the series $(\sum_{n=1}^N x_n)_{N\in\mathbb{N}}$ is convergent in X.
   (b) Let V be a normed space and $U, W \subset V$ dense linear subspaces. Is $U \cap W$ is dense in V? *Hint:* Recall problem 4 from problem set 1 and Weierstrass's approximation theorem from 18.100.

Please provide an estimate of the time you spent on the problem set and list your collaborators.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 3

Due date: 3/10/2025

1. Let $k : [0,1] \times [0,1] \to \mathbb{R}$ be continuous. Define the operator $T_k : C[0,1] \to C[0,1]$, $f \mapsto T_k f$, where
$$(T_k f)(x) = \int_0^1 k(x,y) f(y) dy.$$
   (a) Prove that T_k is a well-defined operator, i.e. that $T_k f \in C[0,1]$ for all $f \in C[0,1]$.
   (b) Prove that T_k is bounded.
   (c) We now relax the assumption of continuity of k and define $k(x,y) = |x-y|^{-\alpha}$ for $x \neq y$ and $k(x,x) = 0$. Show that for $\alpha < 1$, the operator $T_k : C[0,1] \to C[0,1]$ is well-defined and bounded. Compute the operator norm $\|T_k\|$.

2. (a) Consider $X = C([0,1])$ and recall the unbounded differentiation operator with domain $C^1([0,1]) \subset C([0,1])$ defined as $T_D : C^1([0,1]) \to C([0,1])$, $f \mapsto f'$. Show that T_D is closed.
   (b) Let X be a Banach space and let $A : X \to X^*$ be a linear operator such that $(Ax)(y) = (Ay)(x)$ for all $x, y \in X$. Prove that A is bounded.
   (c) Let X be a normed space and assume that X^* is separable. Show that X is separable.
   (d) *Optional ungraded exercise:* Let X be a reflexive Banach space. Show that any closed subspace of X is reflexive.
   (e) *Optional ungraded exercise:* Let X be a Banach space. Show that X is reflexive if and only if X^* is reflexive.

3. For $x \in \ell^p$ with $p \in [1,\infty]$ we define the map $F_x : \ell^q \to \mathbb{R}$, where $F_x(y) = \sum_{n=1}^\infty x_n y_n$ and q is the Hölder conjugate of p.
   (a) Show that for each $x \in \ell^p$, the map F_x is well-defined, linear and bounded, i.e. $F_x \in (\ell^q)^*$.
   (b) For $p \in [1,\infty]$, show that the linear map $F : \ell^p \to (\ell^q)^*$, $x \mapsto F_x$ is a bounded isometry.
   (c) Show that F is surjective if and only if $p \in (1,\infty]$ and conclude that
      i. $(\ell^q)^* \cong \ell^p$ if and only if $p \in (1,\infty]$,
      ii. ℓ^p is reflexive if and only if $p \in (1,\infty)$.
   (d) Show that $(c_0)^* \cong \ell^1$.
   Remark: The dual space of ℓ^∞ can be identified with the *ba space* $ba(2^{\mathbb{N}})$.

Please provide an estimate of the time you spent on the problem set and list your collaborators.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 4

Due date: 3/31/2025

1. (a) Show the downwards monotonicity of measures: Let $E_1 \supset E_2 \supset \dots$ be a sequence of measurable sets in a measure space and assume that there exist an $i \in \mathbb{N}$ such that $\mu(E_i) < \infty$. Show that $\mu(\cap_{n\in\mathbb{N}} E_n) = \lim_{n\to\infty} \mu(E_n) = \inf_{n\in\mathbb{N}} \mu(E_n)$.
   (b) Let $f : \mathbb{R} \to \mathbb{R}$ be a C^1-function and define $A = \{x \in \mathbb{R} : f'(x) = 0\}$. Show that $\lambda(f(A)) = 0$, where λ is the Lebesgue measure.
   Hint: Consider the sets $A_{n,\varepsilon} = \{x \in [-n, n] : |f'(x)| < \varepsilon 2^{-n}\}$.
   (c) Let Ω be a finite measure space. Show that $L^p(\Omega) \subset L^q(\Omega)$ for $1 \le q \le p \le \infty$ and show that the inclusions are dense.
   (d) Assume that there exist a constant $C > 0$ and $p, q \in (1, \infty)$ such that $\|fg\|_{L^1(\mathbb{R})} \le C\|f\|_{L^p(\mathbb{R})}\|g\|_{L^q(\mathbb{R})}$ for all $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$. Show that $\frac{1}{p} + \frac{1}{q} = 1$.

2. In this exercise, we will explore the different ways of losing mass. These examples show that the assumptions in the dominated convergence and monotone convergence theorems are necessary.
   (a) Construct a sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ with pointwise limit $\lim_{n\to\infty} f_n = 0$ such that $0 \le f_n \le 1$, $\lambda(\text{support}(f_n)) \le 1$ and $\liminf_{n\to\infty} \int f_n(x)dx \neq 0$.
   (b) Construct a sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ with uniform limit $\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |f_n(x)| \to 0$ such that $0 \le f_n \le 1$ and $\liminf_{n\to\infty} \int f_n(x)dx \neq 0$.
   (c) Construct a sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ with pointwise limit $\lim_{n\to\infty} |f_n(x)| \to 0$ such that $\text{support}(f_n) \subset [0, 1]$ and $\liminf_{n\to\infty} \int f_n(x)dx \neq 0$.
   (d) Construct a sequence of functions $f_n : [0, 1] \to [0, 1]$ such that $\lim_{n\to\infty} \int f_n(x)dx = 0$ but such that $(f_n)_n$ does not converge pointwise almost everywhere.

3. In this exercise, we will explore the problem of Diophantine approximation, which deals with the approximation of real numbers by rational numbers. We will define a set D_γ below which is the set of reals which can be approximated by rationals with accuracy parameter γ. In this exercise we will explore rather surprising facts about the genericity of D_γ.
   (a) Let $x \in [0, 1]$ and $N \in \mathbb{N}$. Show that there exist $m \in \{0, 1, , \dots , N\}$ and $n \in \{1, , \dots , N\}$ such that $|x - \frac{m}{n}| < \frac{1}{Nn}$. *Hint:* Use the pigeonhole principle.
   Remark: This can be improved to $\le \frac{1}{(N+1)n}$ instead of $< \frac{1}{Nn}$ on the right-hand side.
   (b) Consider the set $D_\gamma = \{x \in [0, 1] : |x - \frac{p}{q}| < q^{-\gamma}$ for infinitely many $(p, q) \in \mathbb{N}_0 \times \mathbb{N}\}$.
      i. If $\gamma \le 2$, show that $D_\gamma = [0, 1]$, i.e. $\lambda(D_\gamma) = 1$.
      ii. If $\gamma > 2$ show that $\lambda(D_\gamma) = 0$. *Hint:* Write $D_\gamma = \bigcap_{n\in\mathbb{N}} \bigcup_{q\ge n, p\in\mathbb{N}_0} \{x \in [0, 1] : |x - \frac{p}{q}| < q^{-\gamma}\}$.
      iii. Show that D_γ is Baire-generic for any $\gamma \in \mathbb{R}$.
      iv. Is $\cap_{\gamma\in\mathbb{R}} D_\gamma = \mathbb{Q}$? If not, can you explicitly construct an element in $\cap_{\gamma\in\mathbb{R}} D_\gamma \cap \mathbb{Q}^c$?
      v. *Optional (not so difficult) ungraded problem:* Show that the Hausdorff dimension satisfies $\dim_{\mathcal{H}}(D_\gamma) \le \frac{2}{\gamma}$.
      vi. *Optional (difficult) ungraded problem:* Show that $\dim_{\mathcal{H}}(D_\gamma) = \frac{2}{\gamma}$ if $\gamma \ge 2$.
   Remark: We just showed that the set $D_\gamma \subset \mathbb{R}$ is generic in the topological sense but exceptional in the measure theoretical sense.

4. *Optional ungraded problem:* Let μ be a translation-invariant measure on $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ such that $\mu((0, 1]) < \infty$. Show that $\mu \equiv 0$ and conclude that there exist subsets of $\mathbb{R}$ which are not Lebesgue measurable.
   Hint: Use the axiom of choice to consider a set of representatives $R \subseteq [0, 1]$ for $[0, 1]/\sim$, where $x \sim y$ if $y - x \in \mathbb{Q}$. Then consider the disjoint sets $A_n = q_n + R$, where $(q_n)_{n\in\mathbb{N}}$ is an enumeration of $\mathbb{Q}$.
   This construction goes back to work of Vitali in 1905.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 5

Due date: 4/14/2025

1. For $f \in L^1(\mathbb{R})$ define the Fourier transform $\mathcal{F}(f)$ as

$$\mathcal{F}(f)(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx.$$

   (a) Show that $\mathcal{F} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ is a bounded linear operator.

   (b) Show that $\mathcal{F}(f)(\xi)$ is a continuous function of ξ.

   (c) Show that $\operatorname{span}\{\mathbf{1}_I : I \text{ is a bounded interval}\}$ is dense in $L^1(\mathbb{R})$.

   (d) Show that $\lim_{\xi\to\pm\infty} \mathcal{F}(f)(\xi) = 0$ and thus $\mathcal{F} : L^1 \to C_0(\mathbb{R})$, where $C_0(\mathbb{R}) \subset L^\infty(\mathbb{R})$ is the set of continuous functions which converge to zero at infinity.
   Hint: Show it first for characteristic functions $\mathbf{1}_I$ and then use (c).
   Remark: Using Fourier inversion, one can show that the map $\mathcal{F} : L^1(\mathbb{R}) \to C_0(\mathbb{R})$ is injective but not surjective. It is an open problem to classify the image $\mathcal{F}(L^1(\mathbb{R})) \subset C_0(\mathbb{R})$.

2. (a) Let $x_n \rightharpoonup x$ in a normed space X. Show that $\|x\| \leq \liminf_{n\to\infty} \|x_n\| \leq \sup_{n\in\mathbb{N}} \|x_n\| < \infty$.

   (b) Consider the space c_0. Prove that there does not exist a normed space X such that $X^* = c_0$.
   Hint: Use Banach–Alaoglu.

3. (a) Let $1 < p < \infty$. Construct a sequence $(x_n)_n \subset \ell^p$ such that $x_n \rightharpoonup x$ but $x_n \not\to x$ as $n \to \infty$.

   (b) Let X be a separable normed space. Show that $\bar{B}_{X^*}$ equipped with the weak* topology is metrizable (i.e. there exists a metric on $\bar{B}_{X^*}$ which induces the weak* topology).

   (c) *Optional ungraded exercise:* Let $(x_n)_n \subset \ell^1$ be such that $x_n \rightharpoonup x$ as $n \to \infty$. Show that $x_n \to x$ as $n \to \infty$.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 6

Due date: 4/28/2025

1. (a) Let H be a Hilbert space. Show that $x_n \to x$ in H if and only if $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ in H.
   (b) Let X be a uniformly convex Banach space. Show that $x_n \to x$ in X if and only if $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ in X. In particular, the above conclusion holds for all L^p for $1 < p < \infty$.
   (c) Find an example in $L^1([0,1])$ such that $f_n \rightharpoonup f$ and $\|f_n\| \to \|f\|$ but $\|f_n - f\| \not\to 0$.

2. In this problem we will find an application of the duality of Hilbert space to solve the Dirichlet problem

$$\begin{cases} \Delta u = f & \text{on } D \\ u = 0 & \text{on } \partial D \end{cases} \tag{1}$$

   on a bounded open domain $D \subset \mathbb{R}^n$, where $f \in L^2(D)$ is a continuous real-valued function.

   (a) Show that for all $\phi \in C_c^\infty(D)$, the following Poincaré inequality holds:

$$\int_D \phi^2 dx \le 4d^2 \int_D \sum_{i=1}^n (\partial_{x_i}\phi)^2 dx, \tag{2}$$

   where d is such that $D \subset [-d,d]^n$. *Hint:* Extend ϕ by trivially to $[-d,d]^n$. Then, use the fundamental theorem of calculus to show $\phi(x) = \int_{-d}^x \partial_{x_1}\phi dx_1$.

   (b) Deduce that $\langle \phi_1, \phi_2 \rangle_{H_0^1} = \int_D \sum_{i=1}^n \partial_{x_i}\phi_1 \partial_{x_i}\phi_2 dx$ defines an inner product on $C_c^\infty(D)$. We denote its completion by $H_0^1(D)$.

   (c) Let $g \in H_0^1(D)$. Then, show that $g \in L^2(D)$ and that g has *weak partial derivatives*, denoted by $\partial_{x_j} g$, $1 \le j \le n$, which are defined by the conditions $\partial_{x_j} g \in L^2(D)$ and $\langle \phi, \partial_{x_j} g \rangle_{L^2} = -\langle \partial_{x_j}\phi, g \rangle_{L^2}$ for all $\phi \in C_c^\infty(D)$ and $1 \le j \le n$.

   (d) Show that there exists $u \in H_0^1(D)$ such that

$$\langle f, v \rangle_{L^2} = -\langle u, v \rangle_{H_0^1}$$

   for all $v \in H_0^1(D)$. *Hint:* Consider the functional $F : H_0^1(D) \to \mathbb{R}, F(v) = -\langle f, v \rangle_{L^2}$.

   (e) Deduce that $u \in H_0^1(D)$ is a *weak solution* to (1), i.e.

$$\int_D u\Delta\phi = \int_D f\phi$$

   for all $\phi \in C_c^\infty(D)$. Note that the boundary condition is encoded in the fact that $u \in H_0^1(D)$.

3. Consider the sequence $x_n = \sqrt{n} e_n$ in ℓ^2. Show that 0 is in the weak closure of $\{x_n : n \in \mathbb{N}\}$ but there is no subsequence of x_n converging to zero.
   Remark: Note that this is a manifestation of the fact that the weak topology on ℓ^2 is not metrizable.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.

18.102 Introduction to Functional Analysis

Problem set 7

Due date: 5/9/2025

1. Let X, Y be normed spaces and $A \in BL(X, Y)$.

   (a) Let A be an isomorphism. Show that $A' : Y^* \to X^*$ is an isomorphism.

   (b) Assume that X, Y are Banach spaces. Show that if A' is an isomorphism, then so is A.

2. Let $T : C[0,1] \to C[0,1]$ defined as
$$Tf(x) = \int_0^x f(y)dy.$$
Prove that $\sigma(T) = \sigma_r(T) = \{0\}$.

3. Let $k : [0,1]^2 \to \mathbb{C}$ be continuous. Consider the operator $T_k : L^2([0,1]) \to L^2([0,1])$ defined as
$$(T_k f)(x) := \int_0^1 k(x,y)f(y)dy.$$

   (a) Show that T_k is compact.
   Hint: Show first that T_k can be approximated by operators with finite rank.

   (b) Assume that $k(x,y) = \overline{k(y,x)}$. Show that $T_k : L^2([0,1]) \to L^2([0,1])$ is self-adjoint.

   (c) Assume again that $k(x,y) = \overline{k(y,x)}$ and let $\lambda \in \rho(T)$ and $g \in L^2([0,1])$. Show that the solution $f \in L^2([0,1])$ to the equation $T_k f - \lambda f = g$, i.e. the unique solution to the integral equation
$$\int_0^1 k(x,y)f(y)dy - \lambda f(x) = g(x)$$
   can be written as
$$f = \sum_n \frac{1}{\lambda_n - \lambda}\langle e_n, g\rangle e_n - \frac{1}{\lambda}\sum_{e \in O}\langle e, g\rangle e,$$
   where $(e_n)_n$ and $(\lambda_n)_n$ are as in the proof of the spectral theorem and O is an orthonormal basis of $\ker T_k$.

4. *Ungraded optional exercise:* The goal of the exercise is to show that imposing the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$ on self-adjoint operators $\hat{x}, \hat{p}$ in quantum mechanics necessarily requires unbounded operators.

   (a) Let X, P be matrices on $\mathbb{C}^n$. Show that $[X, P] = i\hbar$ implies $\hbar = 0$.

   (b) Let X, P be bounded linear operators on a Hilbert space H. Show that $[X, P] = i\hbar$ implies $\hbar = 0$.
   Hint: Show first that $[X^n, P] = in\hbar X^n$.

If you have questions or find mistakes/typos please email me: `kehle@mit.edu`.