Pseudo-differential operators and symmetries

Michael Ruzhansky and Ville Turunen

September 2009

Part I

Foundations of analysis

Part I of the monograph contains preliminary material that could be useful for anyone working in the theory of pseudo-differential operators.

The material of the book is on the intersection of the classical analysis with the representation theory of Lie groups. Aiming at making the presentation self-sufficient we include the preliminary material that may be used as a reference for concepts developed later. In any case, the material presented in this part may be used either as a reference or as an independent textbook on the foundations of analysis.

Throughout the book, we assume that the reader has survived undergraduate calculus courses, so that concepts like *partial derivatives* and the *Riemann integral* are familiar. Otherwise, the prerequisites for understanding the material in this book are quite modest. We shall start with a naive version of a set theory, metric spaces, topology, functional analysis, measure theory and integration in Lebesgue's sense.

Chapter A

Sets, topology and metrics

First, we present the basic notations and properties of sets, used elsewhere in the book. The set theory involved is "naive", sufficient for our purposes; for a thorough treatment, see e.g. [46]. The sets of integer, rational, real or complex numbers will be taken for granted, we shall not construct them.

Let us first list some abbreviations that we are going to use:

- "P *and* Q" means that both properties P and Q are true.
- "P *or* Q" means that at least one of the properties P and Q is true.
- "$P \Rightarrow Q$" reads "*If* P *then* Q", meaning that "P is false or Q is true". Equivalently "$Q \Leftarrow P$", i.e. "Q *only if* P".
- "$P \iff Q$" is "$P \Rightarrow Q$ *and* $P \Leftarrow Q$", reading "P *if and only if* Q".
- "$\exists x$" reads "*There exists* x".
- "$\exists! x$" reads "*There exists a unique* x".
- "$\forall x$" reads "*For every* x".
- "$P := Q$" or "$Q =: P$" reads "P *is defined to be* Q".

A.1 Sets, collections, families

Naively, a *set* (or a *collection* or a *family*) A consists of *points* (or *elements* or *members*) x.

Example. Sets of points, like a *collection of coins*, a *family of two parents and three children*, a *flock of sheep*, a *pack of wolves*, or a *crowd of protesters*.

Example. Points in a set, like the *members of a parliament*, the *flowers in a bundle*, or the *stars in a constellation*.

We denote $x \in A$ if the element x *belongs to the set* A, and $x \notin A$ if x does not belong to A. A set A is a *subset* of a set B, denoted by $A \subset B$ or $B \supset A$, if

$$\forall x : \ x \in A \Rightarrow x \in B.$$

Sets A, B are *equal*, denoted by $A = B$, if $A \subset B$ and $B \subset A$, i.e.

$$\forall x : \ x \in A \iff x \in B.$$

If $A \subset B$ and $A \neq B$ then A is called a *proper subset* of B.

Remark A.1.1 (**Notation for numbers**). The sets of *integer, rational, real and complex numbers* are respectively $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$; let $\mathbb{N} = \mathbb{Z}^+$ and $\mathbb{R}^+$ stand for the corresponding subsets of (strictly) positive numbers. Then

$$\mathbb{Z}^+ \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

We also denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

There are various ways for expressing sets. Sometimes all the elements can be listed:

- The *empty set* $\emptyset = \{\}$ is the unique set without elements: $\forall x : x \notin \emptyset$.
- Set $\{x\}$ consists of a single element $x \in \{x\}$.
- Set $\{x, y\} = \{y, x\}$ consists of elements x and y. And so on. Yet $\{x\} = \{x, x\} = \{x, x, x\}$ etc.

A set consisting of those elements for which property P holds can be denoted by

$$\{x : \ P(x)\} = \{x \mid P(x)\}.$$

A set consisting of finitely many elements $x_1, \ldots, x_n$ could be denoted by

$$\begin{aligned} \{x_1, \ldots, x_n\} &= \{x_k : \ k \in \{1, \ldots, n\}\} \\ &= \left\{x_k \mid k \in \mathbb{Z}^+ : \ k \leq n\right\} \\ &= \{x_k\}_{k=1}^n, \end{aligned}$$

and the infinite set of positive integers by

$$\mathbb{Z}^+ = \{1, 2, 3, 4, 5, \cdots\}.$$

The *power set* $\mathcal{P}(X)$ consists of all the subsets of X,

$$\mathcal{P}(X) = \{A : \ A \subset X\}$$

Example. For the set $X = \{1\}$, we have

$$\begin{aligned} \mathcal{P}(X) &= \{\emptyset, \{1\}\}, \\ \mathcal{P}(\mathcal{P}(X)) &= \{\emptyset, \{\emptyset\}, \{1\}, \{\emptyset, \{1\}\}\}, \end{aligned}$$

and we leave it as an exercise to find out $\mathcal{P}(\mathcal{P}(\mathcal{P}(X)))$, which contains $2^4 = 16$ elements in this case.

Example. Always at least $\emptyset, X \in \mathcal{P}(X)$. If $x \in X$, then $\{x\} \in \mathcal{P}(X)$ and $\{\{x\}\} \in \mathcal{P}(\mathcal{P}(X))$,

$$\begin{aligned} &x \neq \{x\} \neq \{\{x\}\} \neq \cdots, \\ &x \in \{x\} \in \{\{x\}\} \in \cdots. \end{aligned}$$

However, we shall not allow $x \in x$ nor $x \notin x$; consider *Russell's paradox*: given $x = \{a : a \notin a\}$, is $x \in x$?

For $A, B \subset X$, let us define the *union* $A \cup B$, the *intersection* $A \cap B$ and the *difference* $A \setminus B$ by

$$\begin{aligned} A \cup B &:= \{x : x \in A \text{ or } x \in B\}, \\ A \cap B &:= \{x : x \in A \text{ and } x \in B\}, \\ A \setminus B &:= \{x : x \in A \text{ and } x \notin B\}. \end{aligned}$$

The *complement* A^c of A in X is defined by $A^c := X \setminus A$.

Example. If $A = \{1, 2\}$ and $B = \{2, 3\}$ then $A \cup B = \{1, 2, 3\}$, $A \cap B = \{2\}$ and $A \setminus B = \{1\}$.

Example. $\mathbb{R} \setminus \mathbb{Q}$ is the set of irrational numbers.

Exercise A.1.2. Show that

$$\begin{aligned} (A \cup B) \cup C &= A \cup (B \cup C), \\ (A \cap B) \cap C &= A \cap (B \cap C), \\ (A \cup B) \cap C &= (A \cap C) \cup (B \cap C), \\ (A \cap B) \cup C &= (A \cup C) \cap (B \cup C). \end{aligned}$$

Notice that in the latter two cases above, the order of the parentheses is essential. On the other hand, the associativity in the first two equalities allows us to abbreviate $A \cup B \cup C := (A \cup B) \cup C$ and $A \cap B \cap C := (A \cap B) \cap C$ and so on.

Definition A.1.3 (Index sets). Let I be any set and assume that for every $i \in I$ we are given a set A_i. Then I is an *index set* for the collection of sets A_i.

Definition A.1.4 (Unions and intersections of families). For a family $\mathcal{A} \subset \mathcal{P}(X)$, the *union* $\bigcup \mathcal{A}$ and the *intersection* $\bigcap \mathcal{A}$ are defined by

$$\bigcup \mathcal{A} = \bigcup_{B \in \mathcal{A}} B \quad := \quad \{x \mid \exists B \in \mathcal{A} : \; x \in B\},$$
$$\bigcap \mathcal{A} = \bigcap_{B \in \mathcal{A}} B \quad := \quad \{x \mid \forall B \in \mathcal{A} : \; x \in B\}.$$

Example. If $\mathcal{A} = \{B, C\}$ then $\bigcup \mathcal{A} = B \cup C$ and $\bigcap \mathcal{A} = B \cap C$.

Notice that if $\mathcal{A} \subset \mathcal{B} \subset \mathcal{P}(X)$ then

$$\emptyset \subset \bigcup \mathcal{A} \subset \bigcup \mathcal{B} \subset X \quad \text{and} \quad \emptyset \subset \bigcap \mathcal{B} \subset \bigcap \mathcal{A} \subset X.$$

Especially, for $\emptyset \subset \mathcal{P}(X)$ we have

$$\bigcup \emptyset = \emptyset \quad \text{and} \quad \bigcap \emptyset = X. \tag{A.1}$$

Notice that $A \cup B = \bigcup \{A, B\}$ and $A \cap B = \bigcap \{A, B\}$. For unions (and similarly for intersections), the following notations are also commonplace:

$$\bigcup_{j \in K} A_j \quad := \quad \bigcup \{A_j \mid j \in K\},$$
$$\bigcup_{k=1}^{n} A_k \quad := \quad \bigcup \{A_k \mid k \in \mathbb{Z}^+ : \; 1 \leq k \leq n\},$$
$$\bigcup_{k=1}^{\infty} A_k \quad := \quad \bigcup \{A_k \mid k \in \mathbb{Z}^+\}.$$

Example. $\bigcap_{k=1}^{3} A_k = A_1 \cap A_2 \cap A_3$.

Exercise A.1.5. Prove *de Morgan's rules*:

$$X \setminus \bigcup_{j \in K} A_j \quad = \quad \bigcap_{j \in K} (X \setminus A_j),$$
$$X \setminus \bigcap_{j \in K} A_j \quad = \quad \bigcup_{j \in K} (X \setminus A_j).$$

A.2 Relations, function, equivalences and orders

The *Cartesian product* of sets A and B is

$$A \times B = \{(x, y) : \; x \in A, \; y \in B\},$$

where the elements $(x, y) := \{x, \{x, y\}\}$ are ordered pairs: if $x \neq y$ then $(x, y) \neq (y, x)$, whereas $\{x, y\} = \{y, x\}$. A *relation from A to B* is a subset $R \subset A \times B$. We denote xRy if $(x, y) \in R$, saying "*x is in relation R to y*"; analogously, $x \not R\, y$ means $(x, y) \notin R$ ("*x is not in relation R to y*").

Functions

A relation $f \subset X \times Y$ is called a *function* (or a *mapping*) from X to Y, denoted by

$$f : X \to Y \quad \text{or} \quad X \xrightarrow{f} Y,$$

if for each $x \in X$ there exists a unique $y \in Y$ such that $(x, y) \in f$:

$$\forall x \in X\ \exists! y \in Y : \quad (x, y) \in f;$$

in this case, we denote

$$y := f(x) \quad \text{or} \quad x \mapsto f(x) = y.$$

Intuitively, a function $f : X \to Y$ is a rule taking $x \in X$ to $f(x) \in Y$. Functions $X \xrightarrow{f} Y$ and $Y \xrightarrow{g} Z$ yield a *composition* $X \xrightarrow{g \circ f} Z$ by $g \circ f(x) := g(f(x))$. The *restriction* of $f : X \to Y$ to $A \subset X$ is $f|_A : A \to Y$ defined by $f|_A(x) := f(x)$.

Example. The *characteristic function* of a set $E \in \mathcal{P}(X)$ is $\chi_E : X \to \mathbb{R}$ defined by

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Definition A.2.1 (Injections, surjections, bijections). A function $f : X \to Y$ is

- an *injection* if $f(x_1) = f(x_2)$ implies $x_1 = x_2$,
- a *surjection* if for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$,
- and a *bijection* if it is both injective and surjective, and in this case we may define the *inverse function* $f^{-1} : Y \to X$ such that $f(x) = y$ if and only if $x = f^{-1}(y)$.

Definition A.2.2. (Image and preimage) A function $f : X \to Y$ begets functions

$$\begin{aligned} f^+ &: \mathcal{P}(X) \to \mathcal{P}(Y), \quad & f^+(A) = f(A) := \{f(x) \in Y : x \in A\}, \\ f^- &: \mathcal{P}(Y) \to \mathcal{P}(X), \quad & f^-(B) = f^{-1}(B) := \{x \in Y : f(x) \in B\}. \end{aligned}$$

Sets $f(A)$ and $f^{-1}(B)$ are called the *image* of $A \subset X$ and the *preimage* of $B \subset Y$, respectively.

Exercise A.2.3. Let $f : X \to Y$, $A \subset X$ and $B \subset Y$. Show that

$$A \subset f^{-1}(f(A)) \quad \text{and} \quad f(f^{-1}(B)) \subset B.$$

Give examples showing that these subsets can be proper.

Exercise A.2.4. Let $f : X \to Y$, $A_0 \subset X$, $B_0 \subset Y$, $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$. Show that

$$\begin{cases} f(\bigcup \mathcal{A}) & = & \bigcup_{A \in \mathcal{A}} f(A), \\ f(\bigcap \mathcal{A}) & \subset & \bigcap_{A \in \mathcal{A}} f(A), \\ f(X \setminus A_0) & \supset & Y \setminus f(A_0), \end{cases}$$

where the subsets can be proper, while

$$\begin{cases} f^{-1}(\bigcup \mathcal{B}) & = & \bigcup_{B \in \mathcal{B}} f^{-1}(B), \\ f^{-1}(\bigcap \mathcal{B}) & = & \bigcap_{B \in \mathcal{B}} f^{-1}(B), \\ f^{-1}(Y \setminus B_0) & = & X \setminus f^{-1}(B_0). \end{cases}$$

These set-operation-friendly properties of $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ will be encountered later in topology and measure theory.

Definition A.2.5 (Induced and co-induced families). Let $f : X \to Y$, $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$. Then f is said to *induce* the family $f^{-1}(\mathcal{B}) \subset \mathcal{P}(X)$ and to *co-induce* the family $\mathcal{D} \subset \mathcal{P}(Y)$, where

$$\begin{aligned} f^{-1}(\mathcal{B}) &:= \left\{ f^{-1}(B) \mid B \in \mathcal{B} \right\}, \\ \mathcal{D} &:= \left\{ B \subset Y \mid f^{-1}(B) \in \mathcal{A} \right\}. \end{aligned}$$

Equivalences

Definition A.2.6 (Equivalence relation). A subset $\sim$ of $X \times X$ is an *equivalence relation* on X if it is

1. *reflexive*: $x \sim x$ (for all $x \in X$);
2. *symmetric*: if $x \sim y$ then $y \sim x$ (for all $x, y \in X$);
3. *transitive*: if $x \sim y$ and $y \sim z$ then $x \sim z$ (for all $x, y, z \in X$).

The *equivalence class* of $x \in X$ is

$$[x] := \{ y \in X \mid x \sim y \},$$

and the equivalence classes form the *quotient space*

$$X/\sim \quad := \quad \{ [x] \mid x \in X \}.$$

Notice that $x \in [x] \subset X$, that $[x] \cap [y] = \emptyset$ if $[x] \neq [y]$, and that $X = \bigcup_{x \in X} [x]$.

Example. Clearly, the identity relation $=$ is an equivalence relation on X, and $f(x) := \{x\}$ defines a natural bijection $f : X \to X/=$.

Example. Let X and Y denote the sets of all women and men, respectively. For simplicity, we may assume the disjointness $X \cap Y = \emptyset$. Let $Isolde, Juliet \in X$ and $Romeo, Tristan \in Y$. For $a, b \in X \cup Y$, let $x \sim y$ if and only if a and b are of the same gender. Then

$$\begin{aligned} Y = [Tristan] = [Romeo] &\neq [Juliet] = [Isolde] = X, \\ X \cup Y &= [Romeo] \cup [Juliet], \\ (X \cup Y)/\sim &= \{[Romeo], [Juliet]\}. \end{aligned}$$

Exercise A.2.7. Let us define a relation $\sim$ in the Euclidean plane $\mathbb{R}^2$ by setting $(x_1, x_2) \sim (y_1, y_2)$ if and only if $x_1 - y_1, x_2 - y_2 \in \mathbb{Z}$. Show that $\sim$ is an equivalence relation. What is the equivalence class of the origin $(0, 0) \in \mathbb{R}^2$? What is common between a doughnut and the quotient space here?

Exercise A.2.8. Let us define a relation $\sim$ in the punctured Euclidean space $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ by setting $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$ if and only if $(x_1, x_2, x_3) = (ty_1, ty_2, ty_3)$ for some $t \in \mathbb{R}^+$. Prove that $\sim$ is an equivalence relation. What is common between a sphere and the quotient space here?

Orders

Definition A.2.9 (Partial order). A non-empty set X is *partially ordered* if there is a *partial order* $\leq$ on X. That is, $\leq$ is a relation from X to X, such that it is

1. *reflexive*: $x \leq x$ (for all $x \in X$);
2. *anti-symmetric*: if $x \leq y$ and $y \leq x$ then $x = y$ (for all $x, y \in X$);
3. *transitive*: if $x \leq y$ and $y \leq z$ then $x \leq z$ (for all $x, y, z \in X$).

We say that y *is greater than* x (or x *is less than* y), denoted by $x < y$, if $x \leq y$ and $x \neq y$.

Example. The set $\mathbb{R}$ of real numbers has the usual order $\leq$. Naturally, any of its non-empty subsets, e.g. $\mathbb{Z}^+ \subset \mathbb{R}$, inherits the order. The set $[-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$ has the order $\leq$ extended from $\mathbb{R}$, with conventions $-\infty \leq x$ and $x \leq +\infty$ for every $x \in [-\infty, +\infty]$.

Example. Let us order $X = \mathcal{P}(S)$ by inclusion. That is, for $A, B \subset S$, let $A \leq B$ if and only if $A \subset B$.

Example. Let X, Y be sets, where Y has a partial order $\leq$. We may introduce a new partial order for all functions $f, g : X \to Y$ by setting

$$f \leq g \quad \overset{\text{definition}}{\iff} \quad \forall x \in X : \ f(x) \leq g(x).$$

This partial order is commonplace especially when $Y = \mathbb{R}$ or $Y = [-\infty, \infty]$.

Definition A.2.10 (Chains and total order). A non-empty subset $K \subset X$ is a *chain* if $x \leq y$ or $y \leq x$ for all $x, y \in K$. The partial order is *total* (or *linear*) if the whole set X is a chain.

Example. $[-\infty, +\infty]$ is a chain with the usual partial order. Thereby also its subsets are chains, e.g. $\mathbb{R}$ and $\mathbb{Z}^+$. If $\{A_j : \ j \in J\} \subset \mathcal{P}(S)$ is a chain then $A_j \subset A_k$ or $A_k \subset A_j$ for each $j, k \in J$. Moreover, $\mathcal{P}(S)$ is not a chain if S has more than one element.

Definition A.2.11 (Bounds). Let $\leq$ be a partial order on X. The sets of *upper and lower bounds* of $A \subset X$ are defined, respectively, by

$$\begin{aligned} \uparrow A &:= \{x \in X \mid \forall a \in A : \ a \leq x\}, \\ \downarrow A &:= \{x \in X \mid \forall a \in A : \ x \leq a\}. \end{aligned}$$

If $x \in A \cap \uparrow A$ then it is the *maximum of* A, denoted by $x = \max(A)$. If $x \in A \cap \downarrow A$ then it is the *minimum of* A, denoted by $x = \min(A)$. If $A \cap \uparrow \{z\} = \{z\}$ then the element $z \in A$ is called *maximal in* A. Similarly, if $A \cap \downarrow \{z\} = \{z\}$ then the element $z \in A$ is called *minimal in* A. If $\sup(A) := \min(\uparrow A) \in X$ exists, it is called the *supremum of* A, and if $\inf(A) := \max(\downarrow A) \in X$ exists, it is the *infimum of* A.

Remark A.2.12. Notations like

$$\sup_{k \geq 1} x_k = \sup_{k \in \mathbb{Z}^+} x_k = \sup\{x_k : \ k \in \mathbb{Z}^+\}$$

are quite common.

Example. The minimum in $\mathbb{Z}^+$ is 1, but there is no maximal element. For each $A \subset [-\infty, \infty]$, the infimum and the supremum exist.

Example. Let $X = \mathcal{P}(S)$. Then $\max(X) = S$ and $\min(X) = \emptyset$. If $\mathcal{A} \subset X$ then $\sup(\mathcal{A}) = \bigcup \mathcal{A}$ and $\inf(\mathcal{A}) = \bigcap \mathcal{A}$. For each $x \in S$, element $S \setminus \{x\} \in X$ is maximal in the subset $X \setminus \{S\}$.

Definition A.2.13 (lim sup and lim inf). Let $x_k \in X$ for each $k \in \mathbb{Z}^+$. If the following supremums and infimums exist, let

$$\begin{aligned} \limsup_{k \to \infty} x_k &:= \inf\left\{\sup\{x_k : \ j \leq k\} \mid j \in \mathbb{Z}^+\right\}, \\ \liminf_{k \to \infty} x_k &:= \sup\left\{\inf\{x_k : \ j \leq k\} \mid j \in \mathbb{Z}^+\right\}. \end{aligned}$$

Example. Let $E_k \in \mathcal{P}(X)$ for each $k \in \mathbb{Z}^+$. Then

$$\begin{aligned} \limsup_{k \to \infty} E_k &= \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} E_k, \\ \liminf_{k \to \infty} E_k &= \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} E_k. \end{aligned}$$

Exercise A.2.14. Let $A = \limsup_{k\to\infty} E_k$ and $B = \liminf_{k\to\infty} E_k$ as in the example above. Show that

$$\chi_A = \limsup_{k\to\infty} \chi_{E_k} \quad \text{and} \quad \chi_B = \liminf_{k\to\infty} \chi_{E_k},$$

where $\chi_E : X \to \mathbb{R}$ is the characteristic function of $E \subset X$.

A.3 Dominoes tumbling, and transfinite induction

The principle of mathematical induction can be compared to a sequence of dominoes, falling over one after another when the first tumbles down. More precisely,

if $1 \in S \subset \mathbb{Z}^+$ *and* $n \in S \Rightarrow n+1 \in S$ *for every* $n \in \mathbb{Z}^+$,
then $S = \mathbb{Z}^+$.

The Transfinite Induction Principle generalises this, working on any well-ordered set.

Definition A.3.1. A partially ordered set X is said to be *well-ordered*, if $\min(A)$ exists whenever $\emptyset \neq A \subset X$.

Example. With its usual order, $\mathbb{Z}^+$ is well-ordered. With their usual orders, $\mathbb{Z}$, $\mathbb{R}$ and $[-\infty, +\infty]$ are not well-ordered. With the inclusion order, $\mathcal{P}(S)$ is not well-ordered, if there is more than one element in S.

Theorem A.3.2 (Transfinite Induction Principle). *Let X be well-ordered and $S \subset X$. Assume that for each $x \in X$ it holds that $x \in S$ if*

$$\{y \in X : \ y < x\} \subset S.$$

Then $S = X$.

Exercise A.3.3. Prove the Transfinite Induction Principle.

Exercise A.3.4 (Transfinite $\Longrightarrow$ mathematical induction). Check that in the case $X = \mathbb{Z}^+$, the Transfinite Induction Principle is the usual *mathematical induction*.

The value of the Transfinite Induction Principle might be limited, as we have to assume the well-ordering of the underlying set. Actually, many (but not all) working mathematicians assume that

every non-empty set can be well-ordered,

which is so called **Well-Ordering Principle**. Is such a principle likely to be true? After all, for example on sets $\mathbb{R}$ or $\mathcal{P}(\mathbb{Z}^+)$, can we imagine what well-orderings might look like? All the elementary tools which we use in our mathematical reasoning should be at least believable, so maybe the Well-Ordering Principle does not appear as a satisfying set theoretic axiom. Could we perhaps prove or disprove it from other, intuitively more reliable principles? We shall return to this question later.

A.4 Axiom of choice: equivalent formulations

In this section we shall consider how to calculate the number of points in a set, and what infinity might mean in general.

Choosing

We may always choose one point out of a non-empty set, no matter how many points there are around. But sometimes we need infinitely many tasks done at once. For instance, we might want to choose a point from each of the non-empty subsets $A \subset X$ in no time at all: as a tool, we need the Axiom of Choice for X.

Definition A.4.1. Let $X \neq \emptyset$. A mapping $f : \mathcal{P}(X) \to X$ is called a *choice function* on X if $f(A) \in A$ whenever $\emptyset \neq A \subset X$.

Example. Let $X = \{p, q\}$ where $p \neq q$. Let $f : \mathcal{P}(X) \to X$ such that $f(X) = p = f(\{p\})$ and $f(\{q\}) = q$. Then f is a choice function on X.

The following Axiom of Choice should be considered as an axiom or a fundamental principle. In this section we discuss its implications.

Axiom A.4.2 (Axiom of Choice). For every non-empty set there exists a choice function.

Exercise A.4.3. Prove that a choice function exists on a well-ordered set. Thus the Well-Ordering Principle implies the Axiom of Choice.

The Axiom of Choice might look more convincing than the Well-Ordering Principle. Yet, we should be careful, as we are dealing with all kinds of sets, about which our intuition might be deficient. The Axiom of Choice might be plausible for $X = \mathbb{Z}^+$, or maybe even for $X = \mathbb{R}$, but can we be sure whether it is true in general? Nevertheless, let us add Axiom A.4.2 to our set-theoretic tool box.

There are plenty of equivalent formulations for the Axiom of Choice. In the sequel, we present some variants, starting with the "Axiom of Choice for Cartesian Products", to be presented soon.

Definition A.4.4 (Cartesian product). Let X_j be a set for each $j \in J$. The *Cartesian product* is defined to be

$$\prod_{j \in J} X_j := \left\{ f \quad | \quad f : J \to \bigcup_{j \in J} X_j \quad \text{and} \quad \forall j \in J : \ f(j) \in X_j \right\}.$$

If $X_j = X$ for each $j \in J$, we write $X^J := \prod_{j \in J} X_j$. The elements $f \in X^J$ are then functions $f : J \to X$. Moreover, let $X^n := X^{\mathbb{Z}_n}$, where

$$\mathbb{Z}_n := \{k \in \mathbb{Z}^+ \mid k < n\}.$$

Exercise A.4.5. Give an example of a bijection

$$g : X_1 \times X_2 \to \prod_{j \in \{1,2\}} X_j,$$

especially in the case $X \times X \to X^2$. Thereby $X_1 \times X_2$ can be identified with the Cartesian product $\prod_{j \in \{1,2\}} X_j$.

Exercise A.4.6. Give a bijection $g : \mathcal{P}(X) \to \{0,1\}^X$.

Theorem A.4.7 (Axiom of Choice for Cartesian Products). *The Cartesian product of non-empty sets is non-empty.*

Exercise A.4.8. Show that the Axiom of Choice is equivalent to the Axiom of Choice for Cartesian Products.

Theorem A.4.9 (Hausdorff Maximal Principle). *Any chain is contained in a maximal chain.*

Proof. Let $(X, \leq)$ be a partially ordered set with a chain $C_0 \subset X$. Let

$$\mathcal{T} := \{C \mid C \subset X \text{ is a chain such that } C_0 \subset C\}.$$

Now $C_0 \in \mathcal{T}$, so $\mathcal{T} \neq \emptyset$. Let $f : \mathcal{P}(X) \to X$ be a choice function for X. Let us define $s : \mathcal{T} \to \mathcal{T}$ such that $s(C) = C$ if $C \in \mathcal{T}$ is maximal, and if $C \in \mathcal{T}$ is not maximal then

$$s(C) := C \cup \{f(\{x \in X \setminus C : \ C \cup \{x\} \in \mathcal{T}\})\};$$

in this latter case, the chain $s(C)$ is obtained by adding one element to the chain C. The claim follows if we can show that $C = s(C)$ for some $C \in \mathcal{T}$. Let $\mathcal{U} \subset \mathcal{T}$ be a *tower* if

- $C_0 \in \mathcal{U}$,
- $\bigcup \mathcal{K} \in \mathcal{U}$ for any chain $\mathcal{K} \subset \mathcal{U}$,
- $s(\mathcal{U}) \subset \mathcal{U}$. In other words: if $A \in \mathcal{U}$ then $s(A) \in \mathcal{U}$.

For instance, $\mathcal{T}$ is a tower. Let $\mathcal{V}$ be the intersection of all towers. Clearly, $\mathcal{V}$ is a tower, in fact the minimal tower. It will turn out that $\bigcup \mathcal{V} \in \mathcal{T}$ is a maximal chain. This follows if we can show that $\mathcal{V}' \subset \mathcal{V}$ is a tower, where

$$\mathcal{V}' := \{C \in \mathcal{V} \mid \forall B \in \mathcal{V} : B \subset C \text{ or } C \subset B\},$$

since the minimality would imply $\mathcal{V} = \mathcal{V}'$. Clearly, $C_0 \in \mathcal{V}'$, and if $\mathcal{K} \subset \mathcal{V}'$ is a chain then $\bigcup \mathcal{K} \in \mathcal{V}'$. Let $C \in \mathcal{V}'$; we have to show that $s(C) \in \mathcal{V}'$. This follows, if we can show that $\langle C \rangle \subset \mathcal{V}$ is a tower, where

$$\langle C \rangle := \{A \in \mathcal{V} \mid A \subset C \text{ or } s(C) \subset A\}.$$

Clearly, $C_0 \in \langle C\rangle$, and if $\mathcal{K} \subset \langle C\rangle$ is a chain then $\bigcup \mathcal{K} \in \langle C\rangle$. Let $A \in \langle C\rangle$; we have to show that $s(A) \in \langle C\rangle$, i.e. show that $s(A) \subset C$ or $s(C) \subset s(A)$. Since $C \in \mathcal{V}'$, we have $s(A) \subset C$ or $C \subset s(A)$. Suppose the non-trivial case "$C \subset s(A)$ *and* $A \subset C$". Since $s(A) = A \cup \{x\}$ for some $x \in X$, we must have $s(A) = C$ or $C = A$. The proof is complete. □

Theorem A.4.10 (Zorn's Lemma). *A partially ordered set where every chain has an upper bound has a maximal element.*

Exercise A.4.11 (Hausdorff Maximal Principle $\iff$ Zorn's Lemma). Show that the Hausdorff Maximal Principle is equivalent to Zorn's Lemma.

Theorem A.4.12 (Zorn's Lemma $\implies$ Axiom of Choice). *Zorn's Lemma implies the Axiom of Choice.*

Proof. Let X be a non-empty set. Let

$$P := \{f \mid f : \mathcal{P}(A) \to A \text{ is a choice function for some } A \subset X\}.$$

Now $P \neq \emptyset$, because $(\{x\} \mapsto x) : \mathcal{P}(\{x\}) \to \{x\}$ belongs to P for any $x \in X$. Let us endow P with the partial order $\leq$ by inclusion:

$$f \leq g \quad \stackrel{\text{definition}}{\iff} \quad f \subset g$$

(here recall that $f \in P$ is a subset $f \subset \mathcal{P}(A) \times A$ for some $A \subset X$). Suppose $C = \{f_j : j \in J\} \subset P$ is a chain. Then it is easy to verify that

$$\bigcup C = \bigcup_{j \in J} f_j \in P$$

is an upper bound for C, so according to **Zorn's Lemma** there exists a maximal element $f \in P$, which is a choice function for some $A \subset X$. We have to show that $A = X$. On the contrary, suppose $B \subset X$ such that $B \notin \mathcal{P}(A)$. Take $x \in B$. Then $f \subset f \cup \{(B, x)\} \in P$, which would contradict the maximality of f. Hence f must be a choice function for $A = X$. □

How many points?

Intuitively, cardinality measures the number of the elements in a set. Cardinality is a relative concept: sets A, B are compared by whether there is an injection, a surjection or a bijection from one to another. The most interesting results concern infinite sets.

Definition A.4.13 (Cardinality). Sets A, B have the same *cardinality*, denoted by

$$|A| = |B| \quad (\text{or } A \sim B),$$

if there exists a bijection $f : A \to B$. If there exists $C \subset B$ such that $|A| = |C|$, we denote

$$|A| \leq |B|.$$

Moreover, $|A| \leq |B| \neq |A|$ is abbreviated by

$$|A| < |B|.$$

The cardinality of a set A is often also denoted by $\mathrm{card}(A)$.

Exercise A.4.14. Let $|A| = |B|$. Show that $|\mathcal{P}(A)| = |\mathcal{P}(B)|$.

Exercise A.4.15. Show that $|\mathbb{Z}^+| = |\mathbb{Z}|$.

Remark A.4.16. Clearly for every set A, B, C we have

$$\begin{aligned} &A \sim A, \\ &A \sim B \iff B \sim A, \\ &A \sim B \text{ and } B \sim C \implies A \sim C; \end{aligned}$$

formally this is an equivalence relation, though we may have difficulties when discussing the "set of all sets". Notice that $|A| \leq |B|$ means that there is an injection $f : A \to B$, and in this case we may identify set A with $f(A) \subset B$. Obviously,

$$|A| \leq |B| \leq |C| \implies |A| \leq |C|.$$

It is less obvious whether $|A| = |B|$ when $|A| \leq |B| \leq |A|$:

Theorem A.4.17 (Schröder–Bernstein). *Let $|X| \leq |Y|$ and $|Y| \leq |X|$. Then $|X| = |Y|$.*

Proof. Let $f : X \to Y$ and $g : Y \to X$ be injections. Let $X_0 := X$ and $X_1 := g(Y)$. Define inductively $\{X_k : k \in \mathbb{Z}^+\} \subset \mathcal{P}(X)$ by

$$X_{k+2} := g(f(X_k)).$$

Let $X_\infty := \bigcap_{k=0}^\infty X_k$. Now $X_\infty \subset X_{k+1} \subset X_k$ for each $k \geq 0$. Moreover,

$$X_k \setminus X_{k+1} \sim \begin{cases} X_0 \setminus X_1, & \text{if } k \text{ is odd}, \\ X_1 \setminus X_2, & \text{if } k \text{ is even}, \end{cases}$$

so that

$$\begin{aligned} X &= X_\infty \cup \bigcup_{k=0}^\infty (X_k \setminus X_{k+1}) \\ &\sim X_\infty \cup \bigcup_{k=0}^\infty (X_{k+1} \setminus X_{k+2}) \\ &= X_1 \\ &\sim Y. \end{aligned}$$

Thus $X \sim Y$. □

The following Law of Trichotomy is equivalent to the Axiom of Choice, though we derive it only as a corollary to Zorn's Lemma:

Theorem A.4.18 (The Law of Trichotomy). *Let X, Y be sets. Then exactly one of the following holds:*

$$|X| < |Y|, \quad |X| = |Y|, \quad |Y| < |X|.$$

Proof. Assume the non-trivial case $X, Y \neq \emptyset$. Let us define

$$\mathcal{J} := \{f \mid A \subset X,\ f : A \to Y \text{ injective}\}.$$

Clearly, $\mathcal{J} \neq \emptyset$. Thus we may define a partial order $\leq$ on $\mathcal{J}$ by

$$g \leq h \iff g \subset h;$$

notice here that $g, h \subset X \times Y$. Let $\mathcal{K} \subset \mathcal{J}$ be a *chain*. Then it has an upper bound $\bigcup \mathcal{K} \in \mathcal{J}$. Hence by **Zorn's Lemma**, there exists a maximal element $f \in \mathcal{J}$. Now $f : A \to Y$ is injective, where $A \subset X$. If $A = X$ then

$$|X| \leq |Y|.$$

If $f(A) = Y$ then

$$|Y| = |A| \leq |X|.$$

So let us suppose that $A \neq X$ and $f(A) \neq Y$. Then take $x_0 \in X \setminus A$ and $y_0 \in Y \setminus f(A)$. Define

$$g : A \cup \{x_0\} \to Y, \quad g(x) := \begin{cases} f(x), & \text{if } x \in A, \\ y_0, & \text{if } x = x_0. \end{cases}$$

Then $g \in \mathcal{J}$ and $f \leq g \neq f$, which contradicts the maximality of f. Thereby $A = X$ or $f(A) = Y$, meaning

$$|X| \leq |Y| \quad \text{or} \quad |Y| \leq |X|.$$

Finally, if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$ by Theorem A.4.17. □

There is no greatest cardinality:

Theorem A.4.19. *Let X be a set. Then $|X| < |\mathcal{P}(X)|$.*

Proof. If $X = \emptyset$ then $\mathcal{P}(X) = \{\emptyset\}$, and the only injection from X to $\mathcal{P}(X)$ is then the empty relation, which is not a bijection. Assume that $X \neq \emptyset$. Then function

$$f : X \to \mathcal{P}(X), \quad f(x) := \{x\}$$

is an injection, establishing $|X| \leq |\mathcal{P}(X)|$. To get a contradiction, assume that $X \sim \mathcal{P}(X)$, so that there exists a bijection $g : X \to \mathcal{P}(X)$. Let

$$A := \{x \in X : \ x \notin g(x)\}.$$

Let $x_0 := g^{-1}(A)$. Now $x_0 \in A$ if and only if

$$x_0 \notin g(x_0) = A,$$

which is a contradiction. □

Definition A.4.20 (Counting). Let A, B, C, D be sets. For $n \in \mathbb{Z}^+$, let

$$\mathbb{Z}_n := \{k \in \mathbb{Z}^+ \mid k \leq n\} = \{1, \ldots, n\}.$$

We say that $|\emptyset| = 0$, $|\mathbb{Z}_n| = n$,
A is *finite* if $|A| = n$ for some $n \in \mathbb{Z}^+ \cup \{0\}$.
B is *infinite* if it is not finite.
C is *countable* if $|C| \leq |\mathbb{Z}^+|$.
D is *uncountable* if it is not countable.

Remark A.4.21. To strive for transparency in the proofs in this section, let us forget the Law of Trichotomy, which would provide short-cuts like

$$|X| < |Y| \iff |Y| \not\leq |X|.$$

The reader may easily simplify parts of the reasoning using this. The reader is also encouraged to find out where we use the Axiom of Choice or some other nontrivial tools.

Proposition A.4.22. *Let A, B be sets. Then $|A| < |\mathbb{Z}^+| \leq |B|$ if and only if A is finite and B is infinite.*

Proof. Let $A \neq \emptyset$ be finite, so $A \sim \mathbb{Z}_n \subset \mathbb{Z}^+$ for some $n \in \mathbb{Z}^+$. Hence $|A| \leq |\mathbb{Z}^+|$. If $f : \mathbb{Z}^+ \to A$ then $f(n+1) \in f(\mathbb{Z}_n)$, so f is not injective, especially not bijective. Thus $|\mathbb{Z}^+| \not\leq |A|$ and $|A| < |\mathbb{Z}^+|$. Consequently, if $|\mathbb{Z}^+| \leq |B|$ then B is infinite.

Let B be infinite. Take $x_1 \in B \neq \emptyset$. Let $A_n = \{x_1, \cdots, x_n\} \subset B$ be a finite set. Inductively, take $x_{n+1} \in B \setminus A_n \neq \emptyset$. Define

$$\begin{cases} g : \mathbb{Z}^+ \to B, \\ g(n) := x_n. \end{cases}$$

Now g is injective. Hence $|\mathbb{Z}^+| \leq |B|$.

Let $B \subset \mathbb{Z}^+$ be infinite. Define $h : \mathbb{Z}^+ \to B$ inductively by

$$\begin{cases} h(1) := \min(B), \\ h(n+1) := \min\left(B \setminus \{h(1), \cdots, h(n)\}\right). \end{cases}$$

Now h is a bijection: $|B| = |\mathbb{Z}^+|$. So if $|A| < |\mathbb{Z}^+|$ then A is finite. □

Proposition A.4.23. *Let C, D be sets. Then $|C| \leq |\mathbb{Z}^+| < |D|$ if and only if C is countable and D is uncountable.*

Proof. Property $|C| \leq |\mathbb{Z}^+|$ is just the definition of countability. Let D be uncountable, i.e. $|D| \not\leq |\mathbb{Z}^+|$. By Proposition A.4.22, D is not finite, i.e. it is infinite, i.e. $|\mathbb{Z}^+| \leq |D|$. Because of $|\mathbb{Z}^+| \neq |D|$, we have $|\mathbb{Z}^+| < |D|$.

Let $|\mathbb{Z}^+| < |D|$. By Proposition A.4.22, D is infinite, i.e. $|D| \not< |\mathbb{Z}^+|$. Because of $|\mathbb{Z}^+| \neq |D|$, we have even $|D| \not\leq |\mathbb{Z}^+|$, i.e. D is uncountable. □

Remark A.4.24. Let us collect the results from Propositions A.4.22 and A.4.23: For sets A, B, C, D,

$$\begin{cases} |A| < |\mathbb{Z}^+| \leq |B|, \\ |C| \leq |\mathbb{Z}^+| < |D| \end{cases}$$

if and only if A is finite, B is infinite, C is countable, and D is uncountable. In the proofs, we used induction, i.e. well-ordering for $\mathbb{Z}^+$.

Proposition A.4.25. *Let $A_k \subset X$ be a countable subset for each $k \in \mathbb{Z}^+$. Then $\bigcup_{k=1}^{\infty} A_k$ is countable.*

Proof. We may enumerate the elements of each countable A_k:

$$A_k := \left\{a_{kj} : \ j \in \mathbb{Z}^+\right\},$$

$$\begin{aligned} A_1 &= \{a_{11}, a_{12}, a_{13}, a_{14}, \cdots\}, \\ A_2 &= \{a_{21}, a_{22}, a_{23}, a_{24}, \cdots\}, \\ A_3 &= \{a_{31}, a_{32}, a_{33}, a_{34}, \cdots\}, \\ A_4 &= \{a_{41}, a_{42}, a_{43}, a_{44}, \cdots\}, \\ &\vdots \end{aligned}$$

Their union is enumerated by

$$\begin{aligned} \cup_{k=1}^{\infty} A_k &= \{a_{11}, \\ &\quad a_{21}, a_{12}, \\ &\quad a_{31}, a_{22}, a_{13}, \\ &\quad a_{41}, a_{32}, a_{23}, a_{14}, \cdots\} \\ &= \left\{a_{k-j+1,j} : \ 1 \leq j \leq k, \ k \in \mathbb{Z}^+\right\}. \end{aligned}$$

□

Exercise A.4.26. Show that the set $\mathbb{Q}$ of rational numbers is countably infinite.

Exercise A.4.27 (Algebraic numbers). A number $\lambda \in \mathbb{C}$ is called *algebraic* if $p(\lambda) = 0$ for some non-zero polynomial p with integer coefficients, i.e. of some polynomial

$$p(z) = \sum_{k=0}^{n} a_k z^k,$$

where $n \in \mathbb{Z}^+$, $\{a_k\}_{k=0}^n \subset \mathbb{Z}$ and $a_n \neq 0$. Let $\mathbb{A} \subset \mathbb{C}$ be the set of algebraic numbers. Show that $\mathbb{Q} \subset \mathbb{A}$, that $\mathbb{A}$ is countable, and give an example of a number $\lambda \in (\mathbb{R} \cap \mathbb{A}) \setminus \mathbb{Q}$.

Proposition A.4.28. $|\mathbb{R}| = |\mathcal{P}(\mathbb{Z}^+)|$.

Proof. Let us define

$$f : \mathbb{R} \to \mathcal{P}(\mathbb{Q}), \quad f(x) := \{r \in \mathbb{Q} : \ r < x\} .$$

Obviously f is injective, hence $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{Q})|$. By Exercise A.4.26, $|\mathbb{Q}| = |\mathbb{Z}^+|$, implying $|\mathcal{P}(\mathbb{Q})| = |\mathcal{P}(\mathbb{Z}^+)|$. On the other hand, let us define

$$g : \mathcal{P}(\mathbb{Z}^+) \to \mathbb{R}, \quad g(A) := \sum_{k \in A} 10^{-k}.$$

For instance, $0 = g(\emptyset) \leq g(A) \leq g(\mathbb{Z}^+) = 1/9$. Nevertheless, g is injective, implying $|\mathcal{P}(\mathbb{Z}^+)| \leq |\mathbb{R}|$. This completes the proof. □

Exercise A.4.29. Let X be an uncountable set. Show that there exists an uncountable subset $S \subset X$ such that $X \setminus S$ is also uncountable.

A.5 Well-ordering principle revisited

Trivially, the Well-Ordering Principle implies the Axiom of Choice. Actually, there is the reverse implication, too:

Theorem A.5.1 (Well-Ordering Principle). *Every non-empty set can be well-ordered.*

Proof. Let $X \neq \emptyset$. Let

$$P := \{(A_j, \leq_j) \mid j \in J,\ A_j \subset X,\ (A_j, \leq_j) \text{ well-ordered}\} .$$

Clearly, $P \neq \emptyset$. Define a partial order $\leq$ on P by inclusion:

$$(A_j, \leq_j) \leq (A_k, \leq_k) \overset{\text{definition}}{\Longleftrightarrow} \leq_j \subset \leq_k .$$

Take a chain $C \subset P$. Let

$$B := \bigcup_{(A_j, \leq_j) \in C} A_j, \quad \leq_B := \bigcup_{(A_j, \leq_j) \in C} \leq_j .$$

Then $(B, \leq_B) \in P$ is an upper bound for the chain $C \subset P$, so there exists a maximal element $(A, \leq_A) \in P$ by **Zorn's Lemma** A.4.10. Now, if there was $x \in X \setminus A$, then we easily see that $A \cup \{x\}$ could be well-ordered by $\leq_x$ for which $\leq_A \subset \leq_x$, which would contradict the maximality of $(A, \leq_A)$. Therefore $A = X$ has been well-ordered. □

Although we already know that the Well-Ordering Principle and the Hausdorff Maximal Principle are equivalent, let us demonstrate how to use transfinite induction in a related proof:

Proposition A.5.2 (Well-Ordering Principle $\Longrightarrow$ Hausdorff Maximal Principle). *The Well-Ordering Principle implies the Hausdorff Maximal Principle.*

Proof. Let $(X, \leq)$ be well-ordered, i.e. there exists $\min(A) \in A$ whenever $\emptyset \neq A \subset X$. Let $\leq_0$ be a partial order on X. Let us define $f : X \to \mathcal{P}(X)$ by transfinite induction in the following way:

$$f(x) := \begin{cases} \{x\}, & \text{if } \{x\} \cup f(\{y : \ y < x\}) \text{ is a chain with respect to } \leq_0, \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $f(X) \subset \mathcal{P}(X)$ is a maximal chain. □

Exercise A.5.3. Fill in the details in the proof of Proposition A.5.2.

Remark A.5.4 (**Formulations of the Axiom of Choice**). Collecting earlier results and exercises, we see that the following claims are equivalent: the Axiom of Choice, the Axiom of Choice for Cartesian Products, the Hausdorff Maximal Principle, Zorn's Lemma, and the Well-Ordering Principle. The Law of Trichotomy was derived as a corollary to these, but it is actually another equivalent formulation for the Axiom of Choice (see e.g. [121]).

Remark A.5.5 (**Continuum Hypothesis**). When working in analysis, one does not often pay much attention to the underlying set theoretic foundations. Yet, there are many deep problems involved. For instance, it can be shown that there is the smallest uncountable cardinality $|\Omega|$, i.e. whenever S is uncountable then

$$|\mathbb{Z}^+| < |\Omega| \leq |S|.$$

So $|\Omega| \leq |\mathbb{R}|$. A natural question is whether $|\Omega| = |\mathbb{R}|$? Actually, in year 1900, David Hilbert proposed the so called *Continuum Hypothesis*

$$|\Omega| = |\mathbb{R}|.$$

The *Generalised Continuum Hypothesis* is that if X, Y are infinite sets and $|X| \leq |Y| \leq |\mathcal{P}(X)|$ then $|X| = |Y|$ or $|Y| = |\mathcal{P}(X)|$. Without going into details, let (ZF) denote the *Zermelo–Fraenkel* axioms for set theory, (AC) the Axiom of Choice, and (CH) the Generalised Continuum Hypothesis. From 1930s to 1960s, Kurt Gödel and Paul Cohen discovered that:

1. Within (ZF) one cannot prove whether (ZF) is consistent.
2. (ZF+AC+CH) is consistent if (ZF) is consistent.
3. (AC) is independent of (ZF).
4. (CH) is independent of (ZF+AC).

The reader will be notified, whenever we apply (AC) or its equivalents (which is not that often); in this book, we shall not need (CH) at all.

A.6 Metric spaces

Definition A.6.1 (**Metric space**). A function $d : X \times X \to [0, \infty)$ is called a *metric* on the set X if for every $x, y, z \in X$ we have

$$\begin{aligned} d(x,y) = 0 &\iff x = y && \text{(non-degeneracy);} \\ d(x,y) &= d(y,x) && \text{(symmetry);} \\ d(x,z) &\le d(x,y) + d(y,z) && \text{(triangle inequality).} \end{aligned}$$

Then (X, d) (or simply X when d is evident) is called a *metric space*. Sometimes a metric is called a *distance function*. When $x \in X$ and $r > 0$,

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

is called the *x-centered open ball of radius r*. If we want to emphasise that the ball is taken with respect to metric d, we will write $B_d(x, r)$.

Remark A.6.2. In a metric space (X, d),

$$\bigcup_{k=1}^{\infty} B_k(x) = X \quad \text{and} \quad \bigcap_{k=1}^{\infty} B_{1/k}(x) = \{x\}.$$

Example (**Discrete metric**). The mapping $d : X \times X \to [0, \infty)$ defined by

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

is called the *discrete metric* on X. Here

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

Example. Normed vector spaces form a very important class of metric spaces, see Definition B.4.1.

Exercise A.6.3. For $1 \le p < \infty$,

$$d_p(x,y) = \|x - y\|_p := \left(\sum_{j=1}^{n} |x_j - y_j|^p \right)^{1/p}$$

defines a metric $d_p : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$. Function

$$d_\infty(x,y) = \max_{1 \le j \le n} |x_j - y_j|$$

also turns $\mathbb{R}^n$ into a metric space. Unless otherwise mentioned, the space $\mathbb{R}^n$ is endowed with the *Euclidean metric* d_2 (distance "as the crow flies").

Exercise A.6.4 (Sup-metric). Let $a < b$ and let $B([a,b])$ be the space of all bounded functions $f : [a,b] \to \mathbb{R}$. Function

$$d_\infty(f,g) = \sup_{y\in[a,b]} |f(y) - g(y)|$$

turns $B([a,b])$ into a metric space. It is called the *sup-metric*.

Remark A.6.5 (**Metric subspaces**). If $A \subset X$ and $d : X \times X \to [0,\infty)$ is a metric then the restriction

$$d|_{A\times A} : A \times A \to [0,\infty)$$

is a metric on A, with $B_{d|_{A\times A}}(x,r) = A \cap B_d(x,r)$.

Exercise A.6.6. Let $a < b$ and let $C([a,b])$ be the space of all continuous functions $f : [a,b] \to \mathbb{R}$. Show the following statements. The function $d_\infty(f,g) = \sup_{y\in[a,b]} |f(y)-g(y)|$ turns $(C([a,b]), d_\infty)$ into a metric subspace of $(B([a,b]), d_\infty)$. The space $C([a,b])$ also becomes a metric space with metric $d_p(f,g) = \left(\int_a^b |f(y) - g(y)|^p \,\mathrm{d}y\right)^{1/p}$, for any $1 \le p < \infty$. However, $B([a,b])$ with these d_p is not a metric space.

Definition A.6.7 (Diameter and bounded sets). The *diameter* of a set $A \subset X$ in a metric space (X,d) is

$$\operatorname{diam}(A) := \sup\{d(x,y) \mid x,y \in A\},$$

with convention $\operatorname{diam}(\emptyset) = 0$. A set $A \subset X$ is said to be *bounded*, if $\operatorname{diam}(A) < \infty$.

Example. $\operatorname{diam}(\{x\}) = 0$, $\operatorname{diam}(\{x,y\}) = d(x,y)$, and
$\operatorname{diam}(\{x,y,z\}) = \max\{d(x,y), d(y,z), d(x,z)\}$.

Exercise A.6.8. Show that $\operatorname{diam}(B_r(x)) \le 2r$, so that balls are bounded.

Definition A.6.9 (Distance between sets). The *distance* between sets $A,B \subset X$ is

$$\operatorname{dist}(A,B) := \inf\{d(x,y) \mid x \in A,\ y \in B\},$$

with convention that $\operatorname{dist}(A,\emptyset) = \infty$.

Exercise A.6.10. Show that $A \cap B_r(x) \ne \emptyset$ if and only if $\operatorname{dist}(\{x\},A) < r$.

We note that the function $\operatorname{dist}(A,B)$ does not define a metric on subsets of X. For example:

Exercise A.6.11. Give an example of sets $A,B \subset \mathbb{R}^2$ for which $\operatorname{dist}(A,B) = 0$ even though $A \cap B = \emptyset$. Here we consider naturally the Euclidean metric.

Exercise A.6.12. Show that set S in a metric space (X,d) is bounded if and only if there exist some $a \in X$ and $r > 0$ such that $S \subset B_r(a)$.

Lemma A.6.13. *Let S be a bounded set in a metric space (X, d) and let $c \in X$. Then $S \subset B_R(c)$ for some $R > 0$.*

Proof. Since S is a bounded set, there exist some $a \in X$ and $r > 0$ such that $S \subset B_r(a)$. Consequently, for all $x \in S$ we have

$$d(x, c) \leq d(x, a) + d(a, c) < r + d(a, c),$$

so the statement follows with $R = r + d(a, c)$. □

Proposition A.6.14. *The union of finitely many bounded sets in a non-empty metric space is bounded.*

Proof. Let $S_1, \ldots, S_n$ be bounded sets in a non-empty metric space (X, d). Let us take some $c \in X$. Then by Lemma A.6.13 there exists some R_i, $i = 1, \ldots, n$, such that $S_i \subset B_{R_i}(c)$. If we take $R = \max\{R_1, \ldots, R_n\}$, then we have $S_i \subset B_{R_i}(c) \subset B_R(c)$, which implies that $\cup_{i=1}^n S_i \subset B_R(c)$ is bounded. □

Remark A.6.15. We note that the union of infinitely many bounded sets does not have to be bounded. For example, the union of sets $S_i = (0, i) \subset \mathbb{R}$, $i \in \mathbb{N}$, is not bounded in $(\mathbb{R}, d_\infty)$.

Usually, the topological properties can be characterised with *generalised sequences* (or *nets*). Now, we briefly study this phenomenon in metric topology, where ordinary sequences suffice.

Definition A.6.16 (Sequences). A *sequence* in a set A is a mapping $x : \mathbb{Z}^+ \to A$. We denote $x_k := x(k)$ and

$$x = (x_k)_{k \in \mathbb{Z}^+} = (x_k)_{k=1}^\infty = (x_1, x_2, x_3, \cdots).$$

Notice that $x \neq \{x_1, x_2, x_3, \cdots\} = \{x_k : \ k \in \mathbb{Z}^+\}$.

Definition A.6.17 (Convergence). Let (X, d) be a metric space. A sequence $x : \mathbb{Z}^+ \to X$ *converges* to a point $p \in X$, if $\lim_{k \to \infty} d(x_k, p) = 0$, i.e.

$$\forall \varepsilon > 0 \ \exists k_\varepsilon \in \mathbb{Z}^+ : \quad k \geq k_\varepsilon \ \Rightarrow \ d(x_k, p) < \varepsilon.$$

In such a case, we write $\lim_{k \to \infty} x_k = p$ or $x_k \to p$ or $x_k \xrightarrow[k \to \infty]{d} p$ etc.

Clearly, $x_k \to p$ as $k \to \infty$ if and only if

$$\forall \varepsilon > 0 \quad \exists N : \quad k \geq N \Rightarrow \ x_k \in B_\varepsilon(p).$$

We now collect some properties of limits.

Proposition A.6.18 (Uniqueness of limits in metric spaces). *Let (X, d) be a metric space. If $x_k \to p$ and $x_k \to q$ as $k \to \infty$, then $p = q$.*

Proof. Let $\varepsilon > 0$. Since $x_k \to p$ and $x_k \to q$ as $k \to \infty$, it follows that there are some numbers N_1, N_2 such that $d(x_k, p) < \varepsilon$ for all $k > N_1$ and such that $d(x_k, q) < \varepsilon$ for all $k > N_2$. Hence by the triangle inequality for all $k > \max\{N_1, N_2\}$ we have $d(p, q) \leq d(p, x_k) + d(x_k, q) < 2\varepsilon$. Since this conclusion if true for any $\varepsilon > 0$, it follows that $d(p, q) = 0$ and hence $p = q$. □

A.7 Topological spaces

Previously, metric provided a way of measuring distances between sets. The branch of mathematics called topology can be thought as a way to describe "qualitative geography of a set" without referring to specific numerical distance values. We begin by considering properties of metric spaces that motivate the definition of topology which follows after them.

Definition A.7.1 (**Open sets and neighbourhoods).** A set $U \subset X$ in a metric space X is said to be *open* if for every $x \in U$ there is some $\varepsilon > 0$ such that $B_\varepsilon(x) \subset U$. For a point $x \in X$, any open set containing x is called an *open neighbourhood* of x.

Proposition A.7.2. *Every ball $B_r(a)$ in a metric space (X, d) is open.*

Proof. Let $x \in B_r(a)$. Then the number $\varepsilon = r - d(x, a) > 0$ is positive, and $B_\varepsilon(x) \subset B_r(a)$. Indeed, for any $y \in B_\varepsilon(x)$ we have $d(y, a) \leq d(y, x) + d(x, a) < \varepsilon + d(x, a) = r$. □

Proposition A.7.3. *Let (X, d) be a metric space. Then $x_k \to p$ as $k \to \infty$ if and only if every open neighbourhood of x contains all but finitely many of the points x_k.*

Proof. "If" implication is immediate because balls are open. On the other hand, let $p \in U$ where U is an open set. Then there is some $\varepsilon > 0$ such that $B_\varepsilon(p) \subset U$. Now, if $x_k \to p$ as $k \to \infty$, there is some N such that for all $k > N$ we have $x_k \in B_\varepsilon(p) \subset U$, implying the statement. □

Definition A.7.4 (**Continuous mappings in metric spaces).** Let (X_1, d_1) and (X_2, d_2) be two metric spaces, let $f : X_1 \to X_2$, and let $a \in X_1$. Then f is said to be *continuous at a* if for every $\varepsilon > 0$ there is some $\delta > 0$ such that $d_1(x, a) < \delta$ implies $d_2(f(x), f(a)) < \varepsilon$. The mapping f is said to be *continuous* (on X_1) if it is continuous at all points of X_1.

Example. Let $X_1 = C([a, b])$ and $X_2 = \mathbb{R}$ be equipped with the sup-metrics d_1 and d_2, respectively. Then mapping $\Phi : X_1 \to X_2$ defined by $\Phi(h) = \int_a^b h(y)dy$ is continuous.

Definition A.7.5 (**Preimage).** Let $f : X_1 \to X_2$ be a mapping and let $S \subset X_2$ be any subset of X_2. Then the preimage of S under f is defined by

$$f^{-1}(S) = \{x \in X_1 : f(x) \in S\}.$$

Theorem A.7.6. *Let $(X_1, d_1), (X_2, d_2)$ be metric spaces and let $f : X_1 \to X_2$. Then the following statements are equivalent:*

(i) *f is continuous on X_1;*

(ii) *for every $a \in X_1$ and every ball $B_\varepsilon(f(a)) \subset X_2$ there is a ball $B_\delta(a) \subset X_1$ such that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$;*

(iii) *for every open set $U \subset X_2$ its preimage $f^{-1}(U)$ is open in X_1.*

Proof. First, let us show that equivalence of (i) and (ii). Condition (i) is equivalent to saying that for every $\varepsilon > 0$ there is $\delta > 0$ such that $d_1(x, a) < \delta$ implies $d_2(f(a), f(x)) < \varepsilon$. In turn this is equivalent to saying that for every $\varepsilon > 0$ there is $\delta > 0$ such that $x \in B_\delta(a)$ implies $f(x) \in B_\varepsilon(f(a))$, which means that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$.

To show that (ii) implies (iii), let us assume that f is continuous and that $U \subset X_2$ is open. Take $x \in f^{-1}(U)$. Then $f(x) \in U$ and since U is open there is some $\varepsilon > 0$ such that $B_\varepsilon(f(a)) \subset U$. Consequently, by (ii), there exists $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a))) \subset f^{-1}(U)$, implying that $f^{-1}(U)$ is open.

Finally, let us show that (iii) implies (ii). We observe that by (iii) for every $a \in X_1$ and every $\varepsilon > 0$ the set $f^{-1}(B_\varepsilon(f(a)))$ is an open set containing a. Hence there is some $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$, completing the proof. □

Theorem A.7.7. *Let X be a metric space. We have the following properties of open sets in X:*

(T1) *$\emptyset$ and X are open sets in X.*

(T2) *The union of any collection of open subsets of X is open.*

(T3) *The intersection of a finite collection of open subsets of X is open.*

Proof. It is obvious that the empty set $\emptyset$ is open. Moreover, for any $x \in X$ and any $\varepsilon > 0$ we have $B_\varepsilon(x) \subset X$, implying that X is also open.

To show (T2), suppose that we have a collection $\{A_i\}_{i \in I}$ of open sets in X, for an index set I. Let $a \in \cup_{i \in I} A_i$. Then there is some $j \in I$ such that $a \in A_j$ and since A_j is open there is some $\varepsilon > 0$ such that $B_\varepsilon(a) \subset A_j \subset \cup_{i \in I} A_i$, implying (T2).

To show (T3), assume that $A_1, \dots A_n$ is a finite collection of open sets and let $a \in \cap_{i=1}^n A_i$. It follows that for every $i = 1, \dots, n$ we have $a \in A_i$ and hence there is $\varepsilon_i > 0$ such that $B_{\varepsilon_i}(a) \subset A_i$. Let now $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_n\}$. Then $B_\varepsilon(a) \subset A_i$ for all i and hence $B_\varepsilon(a) \subset \cap_{i=1}^n A_i$ implying that the intersection of A_i's is open. □

Definition A.7.8 (**Topology**)**.** A family of sets $\tau \subset \mathcal{P}(X)$ is called a *topology* on the set X if

1. $\bigcup \mathcal{U} \in \tau$ for every collection $\mathcal{U} \subset \tau$, and
2. $\bigcap \mathcal{U} \in \tau$ for every finite collection $\mathcal{U} \subset \tau$.

Then (X, τ) (or simply X when τ is evident) is called a *topological space*; a set $A \subset X$ is called *open* (or *τ-open*) if $A \in \tau$, and *closed* (or *τ-closed*) if $X \setminus A \in \tau$. Let the collection of τ-closed sets be denoted by

$$\tau^* = \{X \setminus U : \; U \in \tau\}.$$

Then the axioms of the topology become naturally complemented:

1. $\bigcap \mathcal{A} \in \tau^*$ for every collection $\mathcal{A} \subset \tau^*$, and
2. $\bigcup \mathcal{A} \in \tau^*$ for every finite collection $\mathcal{A} \subset \tau^*$.

Remark A.7.9. Recall our natural conventions (A.1) for the union and the intersection of the empty family. Thereby $\tau \subset \mathcal{P}(X)$ is a topology if and only if the following conditions hold:

(T1) $\emptyset, X \in \tau$,

(T2) $\bigcup \mathcal{U} \in \tau$ for every non-empty collection $\mathcal{U} \subset \tau$, and

(T3) $U \cap V \in \tau$ for every $U, V \in \tau$.

Consequently, for any topology of X, the subsets $\emptyset \subset X$ and $X \subset X$ are always both open and closed.

Proposition A.7.3 motivates the following notion of convergence in topological spaces.

Definition A.7.10 (Convergence in topological spaces). Let (X, τ) be a topological space. We say that a sequence x_k *converges* to p as $k \to \infty$, and write $x_k \to p$ as $k \to \infty$, if every open neighbourhood of p contains all but finitely many of points x_k.

Proposition A.7.11. *Let X and Y be topological spaces and let $f : X \to Y$ be continuous. If $x_k \to p$ in X as $k \to \infty$ then $f(x_k) \to f(p)$ in Y as $k \to \infty$.*

Proof. Let U be an open set in Y containing $f(p)$. Then $p \in f^{-1}(U)$ and $f^{-1}(U)$ is open in X, implying that there is N such that $x_k \in f^{-1}(U)$ for all $k > N$. Consequently, $f(x_k) \in U$ for all $k > N$ implying that $f(x_k) \to f(p)$ in Y as $k \to \infty$. □

Corollary A.7.12. *Any metric set is a topological space by Theorem A.7.7. The canonical topology of a metric space (X, d) is the set τ consisting of all sets in (X, d) which are open according to Definition A.7.1. This canonical metric topology will be denoted by τ_d or by $\tau(d)$. Metric convergence in (X, d) is equivalent to the topological convergence in the canonical metric topology (X, τ_d).*

Remark A.7.13. Notice that the intersection of any finite collection of τ-open sets is τ-open. On the other hand, it may well be that a countably infinite intersection of open sets is not open. In a metric space (X, d),

$$\bigcap_{k=1}^{\infty} B_{1/k}(x) = \{x\}.$$

Now $\{x\} \in \tau_d$ if and only if $\{x\} = B_r(x)$ for some $r > 0$.

Corollary A.7.14 (Properties of closed sets). *Let X be a topological space. We have the following properties of closed sets in X:*

(C1) $\emptyset$ *and* X *are closed in* X.

(C2) *The intersection of any collection of closed subsets of* X *is closed.*

(C3) *The union of a finite collection of closed subsets of* X *is closed.*

Proof. Let A_i, $i \in I$, be any collection of subsets of X. The corollary follows immediately from Remark A.7.9 and de Morgan's rules

$$X \backslash \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X \backslash A_i), \quad X \backslash \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X \backslash A_i),$$

see Exercise A.1.5. □

Definition A.7.15 (Comparing metric topologies). Let d_1, d_2 be two metrics on a set X. The topology $\tau(d_1)$ defined by d_1 is said to be *stronger* than topology $\tau(d_2)$ defined by d_2 if $\tau(d_1) \supset \tau(d_2)$. In this case the topology $\tau(d_2)$ is also said to be *weaker* than $\tau(d_1)$. Metrics d_1, d_2 on a set X are said to be *equivalent* if they define the same topology $\tau(d_1) = \tau(d_2)$.

Proposition A.7.16 (Criterion for comparing metric topologies). *Let d_1, d_2 be two metrics on a set X such that there is a constant $C > 0$ such that $d_2(x,y) \leq Cd_1(x,y)$ for all $x, y \in X$. Then $\tau(d_2) \subset \tau(d_1)$, i.e. every d_2-open set is also d_1-open.*

Consequently, if there is a constant $C > 0$ such that

$$C^{-1} d_1(x,y) \leq d_2(x,y) \leq C d_1(x,y), \tag{A.2}$$

for all $x, y \in X$, then metrics d_1 and d_2 are equivalent. Such metrics are called Lipschitz equivalent.

Sometimes such metrics are called just *equivalent*, however we use the term "Lipschitz" to distinguish this equivalence from the one in Definition A.7.15.

Proof. Fixing the constant $C > 0$ from (A.2), we observe that $d_1(x,y) < r$ implies $d_2(x,y) < Cr$, which means that $B_{d_1}(x,r) \subset B_{d_2}(x,Cr)$. Let now $U \in \tau(d_2)$ and let $x \in U$. Then there is some $\varepsilon > 0$ such that $B_{d_2}(x,\varepsilon) \subset U$ implying that $B_{d_1}(x, \varepsilon/C) \subset U$. Hence $U \in \tau(d_1)$. □

Exercise A.7.17. Prove that the metrics d_p, $1 \leq p \leq \infty$, from Exercise A.6.3, are all Lipschitz equivalent. The corresponding topology is called the Euclidean metric topology on $\mathbb{R}^n$.

Definition A.7.18 (Relative topology). Let (X, τ) be a topological space and let $A \subset X$. Then we define the *relative topology* on A by

$$\tau_A = \{U \cap A : U \in \tau\}.$$

Proposition A.7.19 (Relative topology is a topology). *Any subset A of a topological space (X, τ) when equipped with the relative topology τ_A is a topological space.*

Proof. We have to check the properties (T1)–(T3) of Remark A.7.9. It is easy to see that $\emptyset = \emptyset \cap A \in \tau_A$ and that $A = X \cap A \in \tau_A$. To show (T2), let $V_i \in \tau_A$, $i \in I$, be a family of sets from τ_A. Then there exist sets $U_i \in \tau$ such that $V_i = U_i \cap A$. Consequently, we have

$$\bigcup_{i\in I} V_i = \bigcup_{i\in I}(U_i \cap A) = \left(\bigcup_{i\in I} U_i\right) \cap A \in \tau_A.$$

To show (T3), let $V_1, \ldots, V_n$ be a family of sets from τ_A. It follows that there exist sets $U_i \in \tau$ such that $V_i = U_i \cap A$. Consequently, we have

$$\bigcap_{i=1}^{n} V_i = \bigcap_{i=1}^{n}(U_i \cap A) = \left(\bigcap_{i=1}^{n} U_i\right) \cap A \in \tau_A,$$

completing the proof. □

Remark A.7.20 (**Metric subspaces**). Let (X, d) be a metric space with canonical topology $\tau(d)$. Let $Y \subset X$ be a subset of X and let us define $d_Y = d|_{Y\times Y}$. Then $\tau(d_Y) = \tau(d)_Y$, i.e. the canonical topology of the metric subspace coincides with the relative topology of the metric space.

Definition A.7.21 (**Product topology).** Let (X_1, τ_1) and (X_2, τ_2) be topological spaces. A subset of $X_1 \times X_2$ is said to be *open in the product topology* if it is a union of sets of the form $U_1 \times U_2$, where $U_1 \in \tau_1$, $U_2 \in \tau_2$. The collection of all such open sets is denoted by $\tau_1 \otimes \tau_2$.

Proposition A.7.22 (**Product topology is a topology).** *The set* $X_1 \times X_2$ *with the collection* $\tau_1 \otimes \tau_2$ *is a topological space.*

Proof. We have to check properties (T1)–(T3) of Remark A.7.9. It is easy to see that $\emptyset = \emptyset \times \emptyset \in \tau_1 \otimes \tau_2$ and that $X_1 \times X_2 \in \tau_1 \otimes \tau_2$.

To show (T2), assume that $A_\alpha \in \tau_1 \otimes \tau_2$ for all $\alpha \in I$. Then each A_α is a union of sets of the form $U_1 \times U_2$ with $U_1 \in \tau_1$, $U_2 \in \tau_2$. Consequently, the union $\cup_{\alpha\in I} A_\alpha$ is a union of sets of the same form and does, therefore, also belong to $\tau_1 \otimes \tau_2$.

To show (T3), even for n sets, assume that $A_i \in \tau_1 \otimes \tau_2$, for all $i = 1, \ldots, n$. By definition there exist collections $U^i_{\alpha_i} \in \tau_1$, $V^i_{\alpha_i} \in \tau_2$, $\alpha_i \in I_i$, $i = 1, \ldots, n$, such that

$$A_i = \bigcup_{\alpha_i \in I_i} (U^i_{\alpha_i} \times V^i_{\alpha_i}),\ i = 1, \ldots, n.$$

Consequently,

$$\bigcap_{i=1}^{n} A_i = \bigcup_{\alpha_i\in I_i,\, 1\le i\le n} \left((\cap_{j=1}^n U^j_{\alpha_i}) \times (\cap_{j=1}^n V^j_{\alpha_i})\right) \in \tau_1 \otimes \tau_2,$$

completing the proof. □

Theorem A.7.23 (Topologies on $\mathbb{R}^2$). *The product topology on $\mathbb{R} \times \mathbb{R}$ is the Euclidean metric topology of $\mathbb{R}^2$.*

Proof. We start by proving that every set open in the product topology of $\mathbb{R}^2$ is also open in the Euclidean topology of $\mathbb{R}^2$. First we note that any open set in $\mathbb{R}$ in the Euclidean topology is a union of open intervals, i.e. every open set U can be written as $U = \cup_{x\in U} B_{\varepsilon_x}(x)$, where $B_{\varepsilon_x}(x)$ is an open ball centred at x with some $\varepsilon_x > 0$. Then we note that every open rectangle in $\mathbb{R}^2$ is open in the Euclidean topology. Indeed, any rectangle $R = (a, b) \times (c, d)$ in $\mathbb{R}^2$ can be written as a union of balls, i.e.

$$R = \cup_{x\in R} B_{\varepsilon_x}(x),$$

with balls $B_{\varepsilon_x}(x)$ taken with respect to d_2, with some $\varepsilon_x > 0$, implying that R is open in the Euclidean topology of $\mathbb{R}^2$. Finally, we note that any open set A in the product topology is a union of sets of the form $U_1 \times U_2$, where U_1, U_2 are open in $\mathbb{R}$. Consequently, writing both U_1 and U_2 as unions of open intervals, we obtain A is a union of open rectangles in $\mathbb{R}^2$, which we showed to be open in the Euclidean topology, implying in turn that A is also open in the Euclidean topology of $\mathbb{R}^2$.

Conversely, let us prove that every set open in the Euclidean topology of $\mathbb{R}^2$ is also open in the product topology of $\mathbb{R}^2$. First we note that clearly every disc $B_\varepsilon(x)$ in $\mathbb{R}^2$ can be written as a union of open rectangles and is, therefore, open in the product topology of $\mathbb{R}^2$. Consequently, every open set U in the Euclidean topology can be written as $U = \cup_{x\in U} B_{\varepsilon_x}(x)$ for some $\varepsilon_x > 0$, so that it is also open in the product topology as a union of open sets. □

A.8 Kuratowski's closure

In this section we describe another approach to topology based on Kuratowski's closure operator. This provides another (and perhaps more intuitive) approach to some notions of the previous section.

Definition A.8.1 (Metric interior, closure, boundary, etc.). In a metric space (X, d), the *metric closure* of $A \subset X$ is

$$\overline{A} = \mathrm{cl}_d(A) := \{x \in X \mid \forall r > 0 : \; A \cap B_r(x) \neq \emptyset\}.$$

In other words, $x \in \mathrm{cl}_d(A) \iff \mathrm{dist}(\{x\}, A) = 0$ (i.e. "x is close to A"). This is also equivalent to saying that every ball around x contains point(s) of A.

The *metric interior* $\mathrm{int}_d(A)$, the *metric exterior* $\mathrm{ext}_d(A)$ and the *metric boundary* $\partial_d(A)$ are defined by

$$\begin{aligned} \mathrm{int}_d(A) &:= X \setminus \mathrm{cl}_d(X \setminus A), \\ \mathrm{ext}_d(A) &:= X \setminus \mathrm{cl}_d(A), \\ \partial_d(A) &:= \mathrm{cl}_d(A) \cap \mathrm{cl}_d(X \setminus A). \end{aligned}$$

Notice that in this way, we have defined mappings

$$\mathrm{cl}_d, \mathrm{int}_d, \mathrm{ext}_d, \partial_d : \mathcal{P}(X) \to \mathcal{P}(X).$$

Exercise A.8.2. Let (X, d) be a metric space and $A \subset X$. Prove the following claims:

$$\begin{aligned}
\mathrm{int}_d(A) &= \{x \in X \mid \exists r > 0 : \ B_r(x) \subset A\}, \\
\partial_d(A) &= \mathrm{cl}_d(A) \setminus \mathrm{int}_d(A), \\
X &= \mathrm{int}_d(A) \cup \partial_d(A) \cup \mathrm{ext}_d(A).
\end{aligned}$$

Consequently, prove that $\mathrm{cl}_d(A)$ is closed for any set $A \subset X$.

Definition A.8.3. Let (X, d) be a metric space. Then

$$\tau_d := \mathrm{int}_d(\mathcal{P}(X)) = \{\mathrm{int}_d(A) \mid A \subset X\}$$

is called the *metric topology* or the family of *metrically open sets*. The corresponding family of *metrically closed sets* is

$$\tau_d^* := \mathrm{cl}_d(\mathcal{P}(X)) = \{\mathrm{cl}_d(A) \mid A \subset X\}.$$

By the following Lemma A.8.4, we have

- a set $C \subset X$ is metrically closed if and only if $C = \mathrm{cl}_d(C)$,
- a set $U \subset X$ is metrically open if and only if $U = \mathrm{int}_d(U)$.

Lemma A.8.4. *Let (X, d) be a metric space and $A \subset X$. Then*

$$\mathrm{cl}_d\left(\mathrm{cl}_d(A)\right) = \mathrm{cl}_d(A), \tag{A.3}$$

$$\mathrm{int}_d\left(\mathrm{int}_d(A)\right) = \mathrm{int}_d(A). \tag{A.4}$$

Proof. Let $C = \mathrm{cl}_d(A)$. Trivially, $C \subset \mathrm{cl}_d(C)$. Let $x \in \mathrm{cl}_d(C)$. Let $r > 0$. Take $y \in C \cap B_r(x)$, and then $z \in A \cap B_r(y)$. Hence

$$d(x, z) \leq d(x, y) + d(y, z) < 2r,$$

so $x \in C$. Thus (A.3) is obtained. By the definition of the metric interior, (A.3) implies (A.4). □

Definition A.8.5 (Topological interior, closure, boundary, etc.). Let τ be a topology on X. For $A \subset X$, the *interior* $\mathrm{int}_\tau(A)$ is the largest open subset of A, and the *closure* $\overline{A} = \mathrm{cl}_\tau(A)$ is the smallest closed set containing A. That is,

$$\begin{aligned}
\overline{A} = \mathrm{int}_\tau(A) &:= \bigcup \{U \in \tau \mid U \subset A\}, \\
\mathrm{cl}_\tau(A) &:= \bigcap \{S \in \tau^* \mid A \subset S\}.
\end{aligned}$$

These define mappings $\mathrm{int}_\tau, \mathrm{cl}_\tau : \mathcal{P}(X) \to \mathcal{P}(X)$. The *boundary* $\partial_\tau(A)$ of a set $A \subset X$ is defined by

$$\partial_\tau(A) := \mathrm{cl}_\tau(A) \cap \mathrm{cl}_\tau(X \setminus A).$$

A set $A \subset X$ is *dense* if $\mathrm{cl}_\tau(A) = X$. The topological space (X, τ) is *separable* if it has a countable dense subset. A point $x \in X$ is an *isolated point* of a set $A \subset X$ if $A \cap U = \{x\}$ for some $U \in \tau$. A point $y \in X$ is an *accumulation point* of a set $B \subset X$ if $(B \cap V) \setminus \{y\} \neq \emptyset$ for every $V \in \tau$. A *neighbourhood of* $x \in X$ is any open set $U \subset X$ containing x. The family of neighbourhoods of $x \in X$ is denoted by

$$\mathcal{V}_\tau(x) := \{U \in \tau \mid x \in U\}$$

(or simply $\mathcal{V}(x)$, when τ is evident).

Remark A.8.6. Intuitively, the closure $\mathrm{cl}_\tau(A) \subset X$ contains those points that are *close to* A. Clearly,

$$\begin{aligned} \tau &= \{\mathrm{int}_\tau(A) \mid A \subset X\}, \\ \tau^* &= \{\mathrm{cl}_\tau(A) \mid A \subset X\}. \end{aligned}$$

Moreover, $U \in \tau$ if and only if $U = \mathrm{int}_\tau(U)$, and $C \in \tau^*$ if and only if $C = \mathrm{cl}_\tau(C)$.

Exercise A.8.7. Prove that

$$\partial_\tau(A) = \mathrm{cl}_\tau(A) \setminus \mathrm{int}_\tau(A).$$

Exercise A.8.8. Let τ_d be the metric topology of a metric space (X, d). Show that $\mathrm{int}_d = \mathrm{int}_{\tau_d}$ and that $\mathrm{cl}_d = \mathrm{cl}_{\tau_d}$.

Proposition A.8.9 (A characterisation of open sets). *Let A be a subset of a topological space X. Then A is open if and only if for every $x \in A$ there is an open set U_x containing x such that $U_x \subset S$.*

Proof. If A is open we can take $U_x = A$ for every $x \in A$. Conversely, writing $A = \cup_{x \in A} U_x$ by property (T2) of open sets we get that A is open if all U_x are open. ☐

Proposition A.8.10 (A characterisation of closures). *Let A be a subset of a topological space X. Then $x \in \overline{A}$ if and only if every open set containing x contains a point of A.*

Proof. We will prove that $x \notin \overline{A}$ if and only if there is an open set U such that $x \in U$ but $A \cap U = \emptyset$. Since $\overline{A}$ is defined as the intersection of all closed sets containing A, it follows that $x \notin \overline{A}$ means that there is a closed set C such that $A \subset C$ and $x \notin C$. Set $U = X \backslash U$ is then the required set. ☐

Definition A.8.11 (Closure operator). Let X be a set. A *closure operator* on X is a mapping $c : \mathcal{P}(X) \to \mathcal{P}(X)$ satisfying *Kuratowski's closure axioms*

1. $c(\emptyset) = \emptyset$,

2. $A \subset c(A)$,
3. $c(c(A)) = c(A)$,
4. $c(A \cup B) = c(A) \cup c(B)$.

Instead of a closure operator $c : \mathcal{P}(X) \to \mathcal{P}(X)$, we could study an *interior operator* $i : \mathcal{P}(X) \to \mathcal{P}(X)$, related to each other by

$$\begin{aligned} i(S) &= X \setminus c(X \setminus S), \\ c(A) &= X \setminus i(X \setminus A). \end{aligned}$$

Kuratowski's closure axioms become interior axioms:

1. $i(X) = X$,
2. $i(S) \subset S$,
3. $i(i(S)) = i(S)$,
4. $i(S \cap T) = i(S) \cap i(T)$.

Theorem A.8.12. *Let (X, τ) be a topological space. Then the mappings $\mathrm{int}_\tau, \mathrm{cl}_\tau : \mathcal{P}(X) \to \mathcal{P}(X)$ are interior and closure operators, respectively.*

Proof. Obviously, $\mathrm{int}_\tau(X) = X$ and $\mathrm{int}_\tau(A) \subset A$. Moreover, $\mathrm{int}_\tau(U) = U$ for $U \in \tau$, and $\mathrm{int}_\tau(A) \in \tau$, because τ is a topology. Hence $\mathrm{int}_\tau(\mathrm{int}_\tau(A)) = \mathrm{int}_\tau(A)$. Finally,

$$\begin{cases} \mathrm{int}_\tau(A \cap B) \subset \mathrm{int}_\tau(A) \subset A, \\ \mathrm{int}_\tau(A \cap B) \subset \mathrm{int}_\tau(B) \subset B, \end{cases}$$

yielding

$$\mathrm{int}_\tau(\mathrm{int}_\tau(A \cap B)) \subset \mathrm{int}_\tau(\mathrm{int}_\tau(A) \cap \mathrm{int}_\tau(B)) \subset \mathrm{int}_\tau(A \cap B),$$

where $\mathrm{int}_\tau(A) \cap \mathrm{int}_\tau(B) \in \tau$, so that $\mathrm{int}_\tau(A \cap B) = \mathrm{int}_\tau(A) \cap \mathrm{int}_\tau(B)$. □

Theorem A.8.13. *Let $i : \mathcal{P}(X) \to \mathcal{P}(X)$ be an interior operator. Then the family $\tau_i = i(\mathcal{P}(X)) = \{i(A) : A \subset X\}$ is a topology. Moreover, $i = \mathrm{int}_{\tau_i}$.*

Proof. First,

$$\emptyset = i(\emptyset) \in \tau_i, \quad X = i(X) \in \tau_i.$$

Second, if $A, B \in \tau_i$ then $A \cap B = i(A) \cap i(B) = i(A \cap B) \in \tau_i$. Third, let $\mathcal{A} = \{A_j : j \in J\} \subset \tau_i$. Now

$$\bigcup \mathcal{A} \overset{A_j = i(A_j)}{=} \bigcup_{j \in J} i(A_j) \overset{i(A_j) \subset i(\bigcup \mathcal{A})}{\subset} i(\bigcup \mathcal{A}) \subset \bigcup \mathcal{A}.$$

Thus $\bigcup \mathcal{A} = i(\bigcup \mathcal{A}) \in \tau_i$. Next,

$$\begin{aligned} \mathrm{int}_{\tau_i}(A) &= \bigcup \{U \in \tau_i \mid U \subset A\} \\ &= \bigcup \{i(B) \mid i(B) \subset A,\ B \subset X\}. \end{aligned}$$

Here we see that $i(A) \subset \mathrm{int}_{\tau_i}(A)$. Moreover, if $i(A) \subset i(B) \subset A$ then $i(A) = i(i(A)) \subset i(i(B)) = i(B) \subset i(A)$. Hence $i(A) = \mathrm{int}_{\tau_i}(A)$. □

Remark A.8.14. Above we have seen how topologies and closure operators (or interior operators) on a set are in bijective correspondence.

Exercise A.8.15. For each $j \in J$, let τ_j be a topology on X. Prove that $\tau = \bigcap_{j \in J} \tau_j$ is a topology. Give an example, where $\bigcup_{j \in J} \tau_j$ is not a topology.

Definition A.8.16 (Base of topology). Let (X, τ) be a topological space. A family $\mathcal{B} \subset \mathcal{P}(X)$ is called a *base* (or *basis*) for the topology τ if any open set is a union of some members of $\mathcal{B}$, i.e.

$$\tau = \left\{ \bigcup \mathcal{B}' : \ \mathcal{B}' \subset \mathcal{B} \right\}.$$

A family $\mathcal{A} \subset \mathcal{P}(X)$ is called a *subbase* (or *subbasis*) for the topology τ if

$$\left\{ \bigcap \mathcal{A}' : \ \mathcal{A}' \subset \mathcal{A} \text{ is finite} \right\}$$

is a base for the topology. A topology is called *second countable* if it has a countable base.

Example. Trivially a topology τ is a base for itself, as $U = \bigcup \{U\}$ for every $U \in \tau$. If (X, d) is a metric space then

$$\mathcal{B} := \{B_r(x) \mid x \in X,\ r > 0\}$$

constitutes a base for τ_d.

Exercise A.8.17. Let $\mathcal{A} \subset \mathcal{P}(X)$. Show that there is the minimal topology $\tau_{\mathcal{A}}$ on X such that $\mathcal{A} \subset \tau_{\mathcal{A}}$: more precisely, if σ is a topology on X for which $\mathcal{A} \subset \sigma$, then $\tau_{\mathcal{A}} \subset \sigma$.

Exercise A.8.18. Let $\tau_{\mathcal{A}}$ be as in the previous exercise. Prove that a base for this topology is provided by

$$\mathcal{B} = \left\{ \bigcap \mathcal{A}' : \ \mathcal{A}' \subset \mathcal{A} \cup \{X\} \text{ is finite} \right\}.$$

Finally, we give another proof of Corollary A.7.12 that metric spaces are topological spaces using the introduced notions of interior and closure.

Theorem A.8.19 (Metric topology is a topology). *Any metric topology is a topology.*

Proof. Let τ_d be the metric topology of (X, d). By Lemma A.8.4, $U \in \tau_d$ if and only if $U = \mathrm{int}_d(U)$. Now $\emptyset, X \in \tau_d$, because

$$\begin{cases} \mathrm{int}_d(\emptyset) = \{x \in X \mid \exists r > 0 : \ B_r(x) \subset \emptyset\} = \emptyset, \\ \mathrm{int}_d(X) = \{x \in X \mid \exists r > 0 : \ B_r(x) \subset X\} = X. \end{cases}$$

Next, if $B_r(x) \subset U$ and $B_s(x) \subset V$ then $B_{\min\{r,s\}}(x) \subset U \cap V$. Thus if $U, V \in \tau_d$ then $U \cap V \in \tau_d$. Finally, if $B_r(x) \subset U_k$ for some $k \in J$ then $B_r(x) \subset \bigcup_{j \in J} U_j$. Thus if $\{U_j : \ j \in J\} \subset \tau_d$ then $\bigcup_{j \in J} U_j \in \tau_d$. □

Exercise A.8.20 (Product topology). Let X, Y be topological spaces with bases $\mathcal{B}_X, \mathcal{B}_Y$, respectively. Show that sets

$$\{U \times V \mid U \in \mathcal{B}_X, \ V \in \mathcal{B}_Y\}$$

form a base for the product topology of $X \times Y = \{(x, y) \mid x \in X, \ y \in Y\}$ from Definition A.7.21.

The metric topology (but not only, cf. topological spaces with countable topology bases) can be characterised by the limits of sequences:

Theorem A.8.21. *Let (X, d) be a metric space, $p \in X$ and $A \subset X$. Then $p \in c_d(A)$ if and only if some sequence $x : \mathbb{Z}^+ \to A$ converges to p.*

Proof. Let $x_k \to p$, where $x_k \in A$ for each $k \in \mathbb{Z}^+$. That is,

$$\forall \varepsilon > 0 \ \exists k_\varepsilon \in \mathbb{Z}^+ : \quad k \geq k_\varepsilon \ \Rightarrow \ x_k \in B_d(p, \varepsilon).$$

Thus $A \cap B_d(p, \varepsilon) \neq \emptyset$ for every $\varepsilon > 0$. Thereby $p \in c_d(A)$.

Let $p \in c_d(A)$, that is $A \cap B_r(x) \neq \emptyset$ for all $r > 0$. For each $k \in \mathbb{Z}^+$, take $x_k \in A \cap B_d(p, 1/k)$. Now $(x_k)_{k=1}^\infty$ is a sequence in A, converging to p, because $d(x_k, p) < 1/k$. □

A.9 Complete metric spaces

In this section we discuss complete metric spaces, give a sample application to Fredholm integral equations using Banach's Fixed Point Theorem, and show that every metric space can be "completed" and such a completion is essentially unique. Later, we will revisit this topic again to show completeness of $\mathbb{R}$ and $\mathbb{R}^n$ in Theorem A.13.10 and Corollary A.13.11.

Definition A.9.1 (Cauchy sequences and completeness). Let (X, d) be a metric space. A sequence $x : \mathbb{Z}^+ \to X$ is a *Cauchy sequence* if

$$\forall \varepsilon > 0 \ \exists k_\varepsilon \in \mathbb{Z}^+ : \ i, j \geq k_\varepsilon \ \Rightarrow \ d(x_i, x_j) < \varepsilon.$$

A metric space is called *complete* if all Cauchy sequences converge.

Example. The Euclidean metric space $(\mathbb{R}^n, d)$ is complete (see Corollary A.13.11), but its dense subset $\mathbb{Q}^n$ is not (metric of course inherited from d). For instance, *Napier's constant* $\mathrm{e} \in \mathbb{R} \setminus \mathbb{Q}$ is obtained as the limit of numbers $\sum_{j=0}^{k} 1/j! \in \mathbb{Q}$.

Lemma A.9.2 (**Properties of Cauchy sequences**)**.** *We have the following properties:*

(1) *Every convergent sequence is a Cauchy sequence.*

(2) *Every Cauchy sequence is bounded.*

(3) *If a Cauchy sequence has a convergent subsequence, it converges to the same limit.*

Proof. We assume that a metric space (X, d) is non-empty. To prove (1), let $x_k \to p$. We want to show that $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence. Let $\varepsilon > 0$. Take $k_\varepsilon \in \mathbb{Z}^+$ such that $d(x_k, p) < \varepsilon$ if $k \geq k_\varepsilon$. Let $i, j \geq k_\varepsilon$. Then

$$d(x_i, x_j) \leq d(x_i, p) + d(p, x_j) < 2\varepsilon.$$

To prove (2), let $(x_k)_{k=1}^{\infty}$ be a Cauchy sequence. Take $\varepsilon = 1$. Then there is some k such that for $i, j \geq k$ we have $d(x_i, x_j) < 1$. Let us now fix some $a \in X$. Then for $i > k$ we have

$$d(a, x_i) \leq d(a, x_{k+1}) + d(x_{k+1}, x_i) < \rho + 1,$$

with $\rho = d(a, x_{k+1})$. Setting $R := \max\{d(a, x_1), \ldots, d(a, x_k), \rho\}$, we get that $x_i \in B_{R+1}(a)$ for all i.

To prove (3), let $(x_n)_{n=1}^{\infty}$ be a Cauchy sequence, with a convergent subsequence $x_{n_i} \to p \in X$. Fix some $\varepsilon > 0$. Then there is some k such that for all $n, m \geq k$ we have $d(x_n, x_m) < \varepsilon$. At the same time, there is some N such that for $n_i > N$, we have $d(x_{n_i}, p) < \varepsilon$. Consequently, for $n \geq \max\{k, N\}$, we get

$$d(x_n, p) \leq d(x_n, x_{n_i}) + d(x_{n_i}, p) < 2\varepsilon,$$

which means that $x_n \to p$ as $n \to \infty$. □

Theorem A.9.3. *Let (X, d) be a complete metric space, and $A \subset X$. Then $(A, d|_{A \times A})$ is complete if and only if $A \subset X$ is closed.*

Proof. Let $A \subset X$ be closed. Take a Cauchy sequence $x : \mathbb{Z}^+ \to A$. Due to the completeness of (X, d), x converges to a point $p \in X$. Now $p \in A$, because A is closed. Thus $(A, d|_{A \times A})$ is complete.

Suppose $(A, d|_{A \times A})$ is complete. We have to show that $c_d(A) = A$. Take $p \in c_d(A)$. For each $k \in \mathbb{Z}^+$, take $x_k \in A \cap B_d(p, 1/k)$. Clearly, $x_k \to p$, so $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence in A. Due to the completeness of $(A, d|_{A \times A})$, $x_k \to a$ for some $a \in A$. Because the limits in X are unique, $p = a \in A$. Thus $A = c_d(A)$ is closed. □

We now show one application of the notion of completeness to solving integral equations.

Definition A.9.4 (Pointwise convergence of functions). Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of functions and let $f : [a,b] \to \mathbb{R}$. Then we say that f_n converges to f *pointwise* on $[a,b]$ if $f_n(x) \to f(x)$ as $n \to \infty$ for all $x \in [a,b]$. In other words, this means that

$$\forall x \in [a,b] \quad \forall \varepsilon > 0 \quad \exists N = N(\varepsilon, x) : \quad n > N \Longrightarrow |f_n(x) - f(x)| < \varepsilon.$$

As before, by $C([a,b])$ we denote the space of all continuous functions $f : [a,b] \to \mathbb{R}$. By default we always equip it with the sup-metric d_∞.

Exercise A.9.5. Find a sequence of continuous functions $f_n \in C([0,1])$ such that $f_n \to f$ pointwise on $[0,1]$, but $f : [0,1] \to \mathbb{R}$ is not continuous on $[0,1]$.

To remedy this situation, we introduce another notion of convergence of functions:

Definition A.9.6 (Uniform convergence of functions). Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of functions and let $f : [a,b] \to \mathbb{R}$. Then we say that f_n converges to f *uniformly* on $[a,b]$ if

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) : \quad \forall n > N \quad x \in [a,b] \Longrightarrow |f_n(x) - f(x)| < \varepsilon.$$

The difference with the pointwise convergence here is that the same index N works for all $x \in [a,b]$.

Theorem A.9.7. *Let $f_n \in C([a,b])$ be a sequence of continuous functions, let $f : [a,b] \to \mathbb{R}$, and suppose that f_n converges to f uniformly on $[a,b]$. Then f is continuous on $[a,b]$.*

Proof. Fix $\varepsilon > 0$. Since $f_n \to f$ uniformly, there is some $N = N(\varepsilon)$ such that for all $n > N$ and all $x \in [a,b]$ we have $|f_n(x) - f(x)| < \varepsilon$. Let $c \in [a,b]$. We will show that f is continuous at c. Since every function f_n is continuous at c, there is some $\delta = \delta(n) > 0$ such that $|x - c| < \delta$ implies $|f_n(x) - f_n(c)| < \varepsilon$. Taking some $n > N$, we get

$$|f(x) - f(c)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(c)| + |f_n(c) - f(c)| < 3\varepsilon$$

for all $|x - c| < \delta$, implying that f is continuous at c. □

This result extends to uniform limits of continuous functions on general topological spaces, see Exercise C.2.18.

Proposition A.9.8 (Metric uniform convergence). *We have $f_n \to f$ in metric space $(C([a,b]), d_\infty)$ if and only if $f_n \to f$ uniformly on $[a,b]$.*

Proof. Convergence $f_n \to f$ in metric space $(C([a,b]), d_\infty)$ means that for every $\varepsilon > 0$ there is N such that for all $n > N$ we have $\sup_{x \in [a,b]} |f_n(x) - f(x)| < \varepsilon$. But this means that $|f_n(x) - f(x)| < \varepsilon$ for all $x \in [a,b]$, which is the uniform convergence. □

Theorem A.9.9 (Completeness of continuous functions). *Space $C([a,b])$ with sup-metric d_∞ is complete.*

Proof. Let $f_n \in C([a,b])$ be a Cauchy sequence. Fix $\varepsilon > 0$. Then there is some N such that for all $m, n > N$ we have

$$\sup_{x\in[a,b]} |f_n(x) - f_m(x)| < \varepsilon. \tag{A.5}$$

Therefore, for each $x \in [a,b]$ the sequence $(f_n(x))_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. If we use that $\mathbb{R}$ is complete (see Theorem A.13.10), it converges to some point in $\mathbb{R}$, which we call $f(x)$. Thus, for every $x \in [a,b]$ we have $f_n(x) \to f(x)$ as $n \to \infty$. Passing to the limit as $n \to \infty$ in (A.5), we obtain $\sup_{x\in[a,b]} |f(x) - f_m(x)| \le \varepsilon$, which means that $d_\infty(f, f_m) \le \varepsilon$, completing the proof. □

Theorem A.9.10 (Banach's Fixed Point Theorem). *Let (X,d) be a non-empty complete metric space, let $k < 1$ be a constant, and let $f : X \to X$ be such that*

$$d(f(x), f(y)) \le k\ d(x,y) \tag{A.6}$$

for all $x, y \in X$. Then there exists a unique point $a \in X$ such that $a = f(a)$.

A mapping f satisfying (A.6) with some constant $k < 1$ is called a *contraction.* A point a such that $a = f(a)$ is called a *fixed point* of f.

Exercise A.9.11. Show that the conditions of Theorem A.9.10 are indispensable. For example, the conclusion of Theorem A.9.10 fails if X is not complete. Show that it also fails if $k \ge 1$. Finally, show that if f satisfies

$$d(f(x), f(y)) < d(x,y)$$

instead of (A.6), the conclusion also fails.

Proof of Theorem A.9.10. First we observe that f is continuous. Indeed, if $d(x,y) < \varepsilon$, it follows that $d(f(x), f(y)) \le kd(x,y) < k\varepsilon < \epsilon$. We now construct a certain Cauchy sequence, whose limit will the the required fixed point of f. Take any $x_0 \in X$. For all $n \ge 0$, define $x_{n+1} = f(x_n)$. Then for all $n \ge 1$ we have

$$d(x_{n+1}, x_n) = d(f(x_{n+1}), f(x_n)) \le kd(x_n, x_{n-1}),$$

implying that $d(x_{n+1}, x_n) \le k^n d(x_1, x_0)$. Consequently, for $n > m \ge 1$, we have

$$\begin{aligned}
d(x_n, x_m) &\le d(x_n, x_{n-1}) + \cdots + d(x_{m+1}, x_m) \\
&\le (k^{n-1} + \cdots + k^m) d(x_1, x_0) \\
&\le k^m \sum_{i=0}^{\infty} k^i\ d(x_1, x_0) \\
&= \frac{k^m}{1-k} d(x_1, x_0).
\end{aligned}$$

Since $k < 1$ it follows that $d(x_n, x_m) \to 0$ as $n, m \to \infty$ which means that x_n is a Cauchy sequence. Since X is complete, $x_n \to a$ for some $a \in X$. We claim that a is a fixed point of f. Indeed, since $x_n \to a$ and since f is continuous, we have $f(x_n) \to f(a)$ by Proposition A.7.11. Therefore, $x_{n+1} \to f(a)$ as $n \to \infty$, and by the uniqueness of limits in metric spaces Proposition A.6.18 we have $f(a) = a$.

Finally, let us show that there is only one fixed point. Suppose that $f(a) = a$ and $f(b) = b$. It follows that $d(a, b) = d(f(a), f(b)) \leq kd(a, b)$ and since $k < 1$, we must have $d(a, b) = 0$ and hence $a = b$. □

Corollary A.9.12 (Fredholm integral equations). *Let $p : [0, 1] \to \mathbb{R}$ be continuous, $p \geq 0$, and such that $\int_0^1 p(t)\ \mathrm{d}t < 1$. Let $g \in C([0, 1])$. Then there exists unique function $f \in C([0, 1])$ such that*

$$f(x) = g(x) - \int_0^x f(t)\ p(t)\ \mathrm{d}t.$$

Proof. As usual, let us equip $C([0, 1])$ with the sup-metric d_∞, and let us define $T : C([0, 1]) \to C([0, 1])$ by

$$(Tf)(x) = g(x) - \int_0^x f(t)\ p(t)\ \mathrm{d}t.$$

We claim that Y is a contraction, which together with the completeness of $C([0, 1])$ in Theorem A.9.9 and Banach's Fixed Point Theorem A.9.10 would imply the statement. We have

$$\begin{aligned} d_\infty(Tf, Tg) &= \sup_{x\in[0,1]} \left| \int_0^x (f(t) - g(t))\ p(t)\ \mathrm{d}t \right| \\ &\leq \sup_{x\in[0,1]} \int_0^x |f(t) - g(t)|\ p(t)\ \mathrm{d}t \\ &= \int_0^1 |f(t) - g(t)|\ p(t)\ \mathrm{d}t \\ &\leq \sup_{x\in[0,1]} |f(x) - g(x)| \int_0^1 p(t)\ \mathrm{d}t \\ &\leq kd(f, g), \end{aligned}$$

where $k = \int_0^1 p(t)\ \mathrm{d}t < 1$. □

Finally we will show that every metric space can be "completed" to become a complete metric space and such "completion" is essentially unique.

Definition A.9.13 (Completion). Let (X, d) be a metric space. A complete metric space X^* is said to be a *completion* of X if X is a topological subspace of X^* and if $\overline{X} = X^*$ (i.e. if X is dense in X^*).

Theorem A.9.14 (**Completions of metric spaces).** *Every metric space* (X, d) *has a completion. This completion is unique up to an isometry leaving* X *fixed.*

Proof. **Existence.** We will construct a completion as a space of equivalence classes of Cauchy sequences in X. Thus, we will call Cauchy sequences $(x_n)_{n=1}^{\infty}$ and $(x'_n)_{n=1}^{\infty}$ equivalent if $d(x_n, x'_n) \to 0$ as $n \to \infty$. One can readily see that this is an equivalence relation as in Definition A.2.6, and we define X^* to be the space of equivalence classes of such Cauchy sequences. Space X^* has a metric d^* defined as follows. For $x^*, y^* \in X^*$, pick some representatives $(x_n)_{n=1}^{\infty} \in x^*$ and $(y_n)_{n=1}^{\infty} \in y^*$, and set

$$d^*(x^*, y^*) := \lim_{n\to\infty} d(x_n, y_n). \tag{A.7}$$

We first check that d^* is a well-defined function on X^*, namely that the limit in (A.7) exists and that it is independent of the choice of representatives on equivalence classes x^* and y^*. To check that the limit exists, we use the fact that x_n and y_n are Cauchy sequences, so for n and m sufficiently large we can estimate

$$\begin{aligned} & |d(x_n, y_n) - d(x_m, y_m)| \\ = \; & |d(x_n, y_n) - d(x_n, y_m) + d(x_n, y_m) - d(x_m, y_m)| \\ \le \; & |d(x_n, y_n) - d(x_n, y_m)| + |d(x_n, y_m) - d(x_m, y_m)| \\ \le \; & d(y_n, y_m) + d(x_n, x_m), \end{aligned}$$

and the latter goes to zero as $n, m \to \infty$. It follows that the sequence of real numbers $(d(x_n, y_n))_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$, and hence converges because $\mathbb{R}$ is complete by Theorem A.13.10 (which will be proved later).

Let us now show that $d^*(x^*, y^*)$ is independent of the choice of representatives from x^* and y^*. Let us take $(x_n)_{n=1}^{\infty}, (x'_n)_{n=1}^{\infty} \in x^*$ and $(y_n)_{n=1}^{\infty}, (y'_n)_{n=1}^{\infty} \in y^*$. Then by a calculation similar to the one before we can show that

$$|d(x_n, y_n) - d(x'_n, y'_n)| \le d(x_n, x'_n) + d(y_n, y'_n),$$

which implies that $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n)$.

We now claim that (X^*, d^*) is a metric space. Non-degeneracy and symmetry in Definition A.6.1 are straightforward. The triangle inequality for d^* follows from that for d. Indeed, passing to the limit as $n \to \infty$ in the inequality $d(x_n, z_n) \le d(x_n, y_n) + d(y_n, z_n)$, we get $d^*(x^*, z^*) \le d^*(x^*, y^*) + d^*(y^*, z^*)$.

Next we will verify that X^* is a completion of X. We first have to check that (X, d) is a topological subspace of (X^*, d^*). We observe that for every $x \in X$ its equivalence class contains the convergent constant sequence $(x_n = x)_{n=1}^{\infty}$, and hence any equivalent Cauchy sequence must be also convergent. Thus, the class x^* consists of all sequences $(x_n)_{n=1}^{\infty}$ convergent to x. Now, if $x, y \in X$ and $(x_n)_{n=1}^{\infty} \in x^*, (y_n)_{n=1}^{\infty} \in y^*$, we have $x_n \to x$ and $y_n \to y$ as $n \to \infty$, and hence $d(x, y) = \lim_{n\to\infty} d(x_n, y_n) = d^*(x^*, y^*)$. Therefore, the mapping $x \mapsto x^*$ is an isometry from X to X^* and hence X is a topological subspace of X^* if we identify it with its image under this isometry. Thus, in the sequel we will no longer distinguish between X and its image in X^*.

We next show that X is dense in X^*. Let $x^* \in X^*$, let $\varepsilon > 0$, and let $(x_n)_{n=1}^\infty \in x^*$. Since x_n is a Cauchy sequence, there is some N such that for all $n, m > N$ we have $d(x_n, x_m) < \varepsilon$. Letting $m \to \infty$, we get $d^*(x_n, x^*) = \lim_{m\to\infty} d(x_n, x_m) \leq \varepsilon$. Therefore, any neighbourhood of x^* contains a point of X, which means that $\overline{X} = X^*$ by Proposition A.8.10.

Finally, we show that (X^*, d^*) is complete. First we observe that by the construction of X^* any Cauchy sequence $(x_n)_{n=1}^\infty$ of points of X converges to $x^* \in X^*$, for $x^* \ni (x_n)_{n=1}^\infty$. Second, for any Cauchy sequence x_n^* of points in X^* there is an equivalent sequence x_n of points of X because $\overline{X} = X^*$. Indeed, for every n there is a point $x_n \in X$ such that $d^*(x_n, x_n^*) < \frac{1}{n}$. Sequence x_n is then a Cauchy sequence and by the first part of this argument it converges to its equivalence class x^* in X^*. Therefore, x_n^* also converges to x^* in (X^*, d^*).

Uniqueness. We want to show that if (X^*, d^*) and (X^{**}, d^{**}) are two completions of X then there is a bijection $f : X^* \to X^{**}$ such that $f(x) = x$ for all $x \in X$, and such that $f(x^*) = x^{**}$, $f(y^*) = y^{**}$ implies that $d^*(x^*, y^*) = d^{**}(x^{**}, y^{**})$. We define f in the following way. For $x^* \in X^*$, in view of the density of X in X^*, there exists a sequence $x_n \in X$ such that $x_n \to x^*$ in (X^*, d^*). Therefore, x_n is a Cauchy sequence in X, and since X^{**} is also a completion of X and is complete, it has some limit in X^{**}, so that $x_n \to x^{**}$ in (X^{**}, d^{**}). One can readily see that this x^{**} is independent of the choice of sequence x_n convergent to x^*. We define f by setting $x^{**} = f(x^*)$.

By construction it is clear that $f(x) = x$ for all $x \in X$. Moreover, let $x_n \to x^*$ and $y_n \to y^*$ in (X^*, d^*) and let $x_n \to x^{**}$ and $y_n \to y^{**}$ in (X^{**}, d^{**}). Consequently,

$$d^*(x^*, y^*) = \lim_{n\to\infty} d^*(x_n, y_n) = \lim_{n\to\infty} d(x_n, y_n) = \\ = \lim_{n\to\infty} d^{**}(x_n, y_n) = d^{**}(x^{**}, y^{**}),$$

completing the proof. □

A.10 Continuity and homeomorphisms

Recall that an expression like "(X, τ) *is a topological space*" is often abbreviated by "X *is a topological space*". In the sequel, to simplify notation, we may use the same letter c for the closure operators of different topological spaces: that is, if $A \subset X$ and $B \subset Y$, $c(A)$ is the closure in the topology of X, and $c(B)$ is the closure in the topology of Y. If needed, we shall express which topologies are meant. In reading the following definition, recall how we have interpreted $x \in c(A)$ as "$x \in X$ *is close to* $A \subset X$":

Definition A.10.1 (Continuous mappings). A mapping $f : X \to Y$ is *continuous at point* $x \in X$ if

$$x \in c(A) \implies f(x) \in c(f(A))$$

for every $A \subset X$. A mapping $f : X \to Y$ is *continuous* if it is continuous at every point $x \in X$, i.e.

$$f(c(A)) \subset c(f(A))$$

for every $A \subset X$. If precision is needed, we may emphasize the topologies involved and, instead of mere *continuity*, speak specifically about (τ_X, τ_Y)*-continuity*. The set of continuous functions from X to Y is often denoted by $C(X, Y)$, with convention $C(X) = C(X, \mathbb{R})$ (or $C(X) = C(X, \mathbb{C})$).

Exercise A.10.2. Let $c \in \mathbb{R}$. Let $f, g : X \to \mathbb{R}$ be continuous, where we use the Euclidean metric topology on $\mathbb{R}$. Show that the following functions $X \to \mathbb{R}$ are then continuous: cf, $f+g$, fg, $|f|$, $\max\{f, g\}$, $\min\{f, g\}$ (here e.g. $\max\{f, g\}(x) := \max\{f(x), g(x)\}$ etc.). Moreover show that if $g(x) \neq 0$ then f/g is continuous at $x \in X$.

Exercise A.10.3. Let (X_1, τ_1) and (X_2, τ_2) be topological spaces. Show that a mapping $f : X_1 \to X_2$ is continuous at $x \in X_1$ if and only if

$$\forall V \in \mathcal{V}_{\tau_2}(f(x)) \quad \exists U \in \mathcal{V}_{\tau_1}(x) : \quad f(U) \subset V.$$

Exercise A.10.4. Let (X, d_X) and (Y, d_Y) be metric spaces, $p \in X$ and $f : X \to Y$. Show that the following conditions are equivalent:

1. f is continuous at $p \in X$ (with respect to the metric topologies).
2. $\forall \varepsilon > 0\ \exists \delta > 0\ \forall w \in X : \ d_X(p, w) < \delta \ \Rightarrow \ d_Y(f(p), f(w)) < \varepsilon$.
3. $f(x_k) \to f(p)$ whenever $x_k \to p$.

Theorem A.10.5. *Let $f : X \to Y$. Then f is continuous if and only if $f^{-1}(V) \in \tau_X$ for every $V \in \tau_Y$.*

Remark A.10.6. The continuity criterion here might be read as: "*preimages of open sets are open*". Sometimes this condition is taken as the definition of continuity of f. Equivalently, by taking complements, this means "*preimages of closed sets are closed*".

Proof. Let us assume that "preimages of closed sets are closed". Then $A' = f^{-1}(c(f(A)))$ is closed, and $A \subset A'$, so $c(A) \subset c(A') = A'$. Hence

$$f(c(A)) \subset f(A') \subset c(f(A)).$$

Property $f(c(A)) \subset c(f(A))$ means the continuity of $f : X \to Y$.

Conversely, let $f : X \to Y$ be continuous. Let $A = f^{-1}(c(B))$, where $B \subset Y$. Then $f(c(A)) \subset c(f(A)) \subset c(c(B)) = c(B)$, so

$$c(A) \subset f^{-1}(f(c(A))) \subset f^{-1}(c(B)) = A.$$

Therefore $c(A) = A$, i.e. A is closed. $\square$

Corollary A.10.7. *Let $f : X \to Y$, and let τ_Y be a topology on Y. Then $f^{-1}(\tau_Y) = \{f^{-1}(V) \mid V \in \tau_Y\}$ is a topology on X. Moreover, f is (τ_X, τ_Y)-continuous if and only if $f^{-1}(\tau_Y) \subset \tau_X$.*

Exercise A.10.8. Prove Corollary A.10.7. The topology $f^{-1}(\tau_Y)$ is called the topology *induced* from τ_Y by f. Show that the relative topology on a subset $A \subset X$ of a topological space X in Definition A.7.18 is induced by the identity mapping $A \to X$.

Definition A.10.9 (Induced topology). Let $\mathcal{F}$ be a family of mappings $f : X \to Y$, where (Y, τ_Y) is a topological space. Then

$$\bigcap_{f \in \mathcal{F}} f^{-1}(\tau_Y) \subset \mathcal{P}(X)$$

is the topology *induced* from τ_Y by $\mathcal{F}$.

Proposition A.10.10. *Let X, Y, Z be topological spaces and let $f : X \to Y$ and $g : Y \to Z$ be continuous. Then $g \circ f : X \to Z$ is continuous.*

Proof. We will use Theorem A.10.5. Let U be open in Z. Then $g^{-1}(U)$ is open in Y and hence $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$ is open in X, implying that $g \circ f$ is continuous. □

Exercise A.10.11. Prove Proposition A.10.10 directly from Definition A.10.1.

Definition A.10.12 (Homeomorphisms and topological equivalence). A bijective mapping $f : X \to Y$ is a *homeomorphism* if both f and f^{-1} are continuous. In this case we say that the corresponding topological spaces (X, τ_X) and (Y, τ_Y) are *homeomorphic*. Homeomorphic spaces are also called *topologically equivalent*. A property which holds in all topologically equivalent spaces is called a *topological property*.

Example. Any two open intervals in $\mathbb{R}$ are topologically equivalent. For a set X, properties "X has five elements" or "all subsets of X are open" are topological properties.

Remark A.10.13. A homeomorphism is a topological isomorphism: homeomorphic spaces are topologically the same. As the saying goes, a topologist is a person who does not know the difference between a doughnut and a coffee cup. Let us denote briefly $X \approx Y$ when (X, τ_X) and (Y, τ_Y) are homeomorphic. It is easy to see that have an equivalence

$$\begin{aligned} & X \approx X, \\ X \approx Y \quad \Longrightarrow \quad & Y \approx X, \\ X \approx Y \text{ and } Y \approx Z \quad \Longrightarrow \quad & X \approx Z. \end{aligned}$$

Analogously, there is a concept of metric space isomorphisms: a bijective mapping $f : X \to Y$ between metric spaces $(X, d_X), (Y, d_Y)$ is called an *isometric isomorphism* if $d_Y(f(a), f(b)) = d_X(a, b)$ for every $a \in X$ and $b \in Y$.

Example. The reader may check that $(x \mapsto x/(1+|x|)) : \mathbb{R} \approx (-1,1)$. Using algebraic topology, one can prove that $\mathbb{R}^m \approx \mathbb{R}^n$ if and only if $m = n$ (this is not trivial!).

Example. Any isometric isomorphism is a homeomorphism. Clearly the unbounded $\mathbb{R}$ and the bounded $(-1,1)$ are not isometrically isomorphic. An orthogonal linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$ is an isometric isomorphism, when $\mathbb{R}^n$ is endowed with the Euclidean norm. The forward shift operator on $\ell^p(\mathbb{Z})$ is an isometric isomorphism, but the forward shift operator on $\ell^p(\mathbb{N})$ is only a non-surjective isometry.

Exercise A.10.14. Let (X, d_X) and (Y, d_Y) be metric spaces. Recall that $f : X \to Y$ is continuous if and only if

$$\forall a \in X\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall b \in X : \quad d_X(a,b) < \delta \implies d_Y(f(a), f(b)) < \varepsilon.$$

A function $f : X \to Y$ is *uniformly continuous* if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall a, b \in X : \quad d_X(a,b) < \delta \implies d_Y(f(a), f(b)) < \varepsilon,$$

and *Lipschitz-continuous* if

$$\exists C < \infty\ \forall a, b \in X : \quad d_Y(f(a), f(b)) \le C\ d_X(a,b).$$

Prove that the Lipschitz-continuity implies the uniform continuity, and that the uniform continuity implies the continuity; give examples showing that these implications cannot be reversed.

Theorem A.10.15. *Two metrics d_1, d_2 on a set X are equivalent if and only if the identity mapping from (X, d_1) to (X, d_2) is a homeomorphism.*

Proof. Let $id(x) = x$ be the identity mapping from (X, d_1) to (X, d_2). Since $id^{-1}(U) = U$ for any set U, the forward implication follows from the definition of a continuous mapping and that of equivalent metrics. On the other hand, suppose the identity map is a homeomorphism. Again, since $id^{-1}(U) = U$ we get that every set open in (X, d_2) is open in (X, d_1) since id is continuous. The converse is true since id^{-1} is also continuous. □

A.11 Compact topological spaces

Eventually, we will mainly concentrate on compact Hausdorff spaces, but in this section we deal with more general classes of topological spaces.

Definition A.11.1 (Coverings). Let X be a set and $K \subset X$. A family $\mathcal{U} \subset \mathcal{P}(X)$ is called a *cover of* K if

$$K \subset \bigcup \mathcal{U};$$

if the cover $\mathcal{U}$ is a finite set, it is called a *finite cover*. A cover $\mathcal{U}$ of $K \subset X$ has a *subcover* $\mathcal{U}' \subset \mathcal{U}$ if $\mathcal{U}'$ itself is a cover of K. In a topological space, an *open cover* refers to a cover consisting of open sets.

Definition A.11.2 (Compact sets). Let (X, τ) be a topological space. A subset $K \subset X$ is *compact* (more precisely *τ-compact*) if every open cover of K has a finite subcover. We say that (X, τ) is a *compact space* if X itself is τ-compact. A topological space is *locally compact* if each of its points has a neighbourhood whose closure is compact.

Remark A.11.3. Briefly, in a topological space (X, τ), $K \subset X$ is compact if and only if the following holds: given any family $\mathcal{U} \subset \tau$ such that $K \subset \bigcup \mathcal{U}$, there exists a finite subfamily $\mathcal{U}' \subset \mathcal{U}$ such that $K \subset \bigcup \mathcal{U}'$.

Remark A.11.4. Let us consider a tongue-in-cheek geography-zoological analogue for compactness: In a space or universe (X, τ), let non-empty open sets correspond to territories of angry animals; recall the metaphor that a point $x \in U \in \tau$ is "far away from (i.e. not close to) the set $X \setminus U$". Compactness of an island $K \subset X$ means that any given territorial cover $\mathcal{U}$ has a finite subcover $\mathcal{U}'$: already a finite number of beasts governs the whole island.

Example.

1. If τ_1 and τ_2 are topologies of X, $\tau_1 \subset \tau_2$, and (X, τ_2) is a compact space then (X, τ_1) is a compact space.
2. $(X, \{\emptyset, X\})$ is a compact space.
3. If $|X| = \infty$ then $(X, \mathcal{P}(X))$ is not a compact space, but it is locally compact. Clearly any space with a finite topology is compact. Even though a compact topology can be of *any* cardinality, it is in a sense "not far away from being finite".
4. A metric space is compact if and only if it is sequentially compact (i.e. every sequence contains a converging subsequence, see Theorem A.13.4).
5. A subset $X \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded (Heine–Borel Theorem A.13.7).
6. Theorem B.4.21 due to Frigyes Riesz asserts that a closed ball in a normed vector space over $\mathbb{C}$ (or $\mathbb{R}$) is compact (i.e. the space is locally compact) if and only if the vector space is finite-dimensional.

Of course, we may work with a complemented version of the compactness criterion in terms of closed sets:

Proposition A.11.5 (Finite intersection property). *A topological space X is compact if and only if the closed sets in X have the finite intersection property, which means that any collection $\{F_\alpha\}_\alpha$ of closed sets in X with $\cap_\alpha F_\alpha = \emptyset$ has a finite subcollection $\{F_i\}_{i=1}^n \subset \{F_\alpha\}_\alpha$ such that $\cap_{i=1}^n F_i = \emptyset$.*

Proof. Defining $U_\alpha = X \backslash F_\alpha$, we observe that condition $\cap_\alpha F_\alpha = \emptyset$ means that $\{U_\alpha\}_\alpha$ is an open covering of X. The condition that X is compact means that any such covering has some finite subcollection $\{U_i\}_{i=1}^n$ with $\cup_{i=1}^n U_i = X$, which in turn means that $\cap_{i=1}^n F_i = \emptyset$. □

Proposition A.11.6 (**Characterisation of compact subspaces).** *Let (X,τ) be a topological space and let $Y \subset X$. Topological subspace (Y,τ_Y) is compact if and only if every collection $\{U_\alpha\}_{\alpha\in I}$ of sets $U_\alpha \in \tau$ with $\bigcup_{\alpha\in I} U_\alpha \supset Y$ has a finite subcollection that covers Y.*

Proof. Assume that (Y,τ_Y) is compact and let $\{U_\alpha\}_{\alpha\in I}$ be a collection of sets $U_\alpha \in \tau$ with $\bigcup_{\alpha\in I} U_\alpha \supset Y$. Then the collection $\{U_\alpha \cap Y\}_{\alpha\in I}$ is an open cover of (Y,τ_Y) and hence has a finite subcover $\{U_i \cap Y\}_{i=1}^n$. The corresponding collection $\{U_i\}_{i=1}^n$ is a finite subcollection of $\{U_\alpha\}_{\alpha\in I}$ that covers Y.

Conversely, let $\{V_\alpha\}_{\alpha\in I} \subset \tau_Y$ be an open cover of Y. Then there exist sets $U_\alpha \in \tau$ such that $V_\alpha = U_\alpha \cap Y$. Consequently, $\{U_\alpha\}_{\alpha\in I} \subset \tau$ is a cover of Y, and by assumption it has a finite subcollection $\{U_i\}_{i=1}^n$ that covers Y. The corresponding collection $\{V_i\}_{i=1}^n$ is then a finite open cover of Y. □

Exercise A.11.7. Show that a finite set in a topological space is compact.

Exercise A.11.8. Let $x \in \mathbb{R}^n$ and $r > 0$. Show that the open ball $B_r(x) \subset \mathbb{R}^n$ is not compact in the Euclidean metric topology.

Exercise A.11.9. Prove that a union of two compact sets is compact.

Proposition A.11.10. *Let (X,τ) be a topological space, $K \subset X$ compact and $S \subset X$ closed. Then $K \cap S$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $K \cap S$. Then $\mathcal{U} \cup \{X \setminus S\}$ is an open cover of K, thus having a finite subcover $\mathcal{U}'$. Then $\mathcal{U}' \cap \mathcal{U} \subset \mathcal{U}$ is a finite subcover of $K \cap S$. □

Proposition A.11.11 (**Some properties of compact sets).** *We have the following properties:*

(1) *A closed subset of a compact topological space is compact.*

(2) *A compact subset of a metric space is bounded (and closed).*

Proof. To prove (1), let Y be a closed subset of a compact topological space (X,τ). Let $\{U_\alpha\}_{\alpha\in I} \subset \tau$ be an open cover of Y. Since Y is closed, its complement $X\backslash Y$ is open, and collection $\{X\backslash Y, U_\alpha\}_{\alpha\in I}$ is an open cover of X. Since X is compact, it has a finite subcover and since $X\backslash Y$ is disjoint from Y, removing $X\backslash Y$ (if necessary) from this subcover we obtain a finite subcover of Y.

To prove (2), let Y be a compact subspace of a metric space (X,d). Collection of unit balls $\{B_1(y)\}_{y\in Y}$ is an open cover of Y, and hence it has a finite subcover, say $\{B_1(y_i)\}_{i=1}^n$. Applying Proposition A.6.14 we obtain that Y must be bounded. □

Proposition A.11.12. *Let X be a compact space and $f : X \to Y$ continuous. Then $f(X) \subset Y$ is compact.*

Proof. Let $\mathcal{V}$ be an open cover of $f(X)$. Then $\mathcal{U} := \{f^{-1}(V) \mid V \in \mathcal{V}\}$ is an open cover of X, thus having a finite subcover $\mathcal{U}' = \{f^{-1}(V) \mid V \in \mathcal{V}'\}$, where $\mathcal{V}' \subset \mathcal{V}$ is a finite collection. Then $f(X)$ is covered by $\mathcal{V}' \subset \mathcal{V}$: if $y \in f(X)$ then $y = f(x)$ for some $x \in X$, so $x \in f^{-1}(V_0)$ for some $V_0 \in \mathcal{V}'$, so $y = f(x) \in f(f^{-1}(V_0)) \subset V_0$. □

Corollary A.11.13. *Let $f : X \to \mathbb{R}$ be a continuous mapping from a compact topological space X to $\mathbb{R}$ equipped with the Euclidean topology. Then $f(X)$ is a bounded subset of $\mathbb{R}$.*

Theorem A.11.14 (Product of compact spaces is compact). *Let X, Y be compact topological spaces. Then $X \times Y$ in the product topology is compact.*

Proof. Let $\mathcal{C} = \{W_\alpha\}_{\alpha \in I}$ be an open cover of $X \times Y$ in the product topology. In particular, it means that each W_α is a union of "rectangles" of the form $U \times Y$ where U and V are open in X and Y, respectively. For every (x, y) there is a "rectangle" $U_x^y \times V_x^y$ and the corresponding set W_x^y such that

$$(x, y) \in U_x^y \times V_x^y \subset W_x^y \in \mathcal{C}.$$

For every $x \in y$, collection $\{V_x^y\}_{y \in Y}$ is an open covering of Y which then must have some finite subcover, which we denote by $\{V_x^{y_i(x)}\}_{i=1}^{n(x)}$. Set $U_x = \cap_{i=1}^{n(x)} U_x^{y_i(x)}$ is open in X and collection $\{W_x^{y_i(x)}\}_{i=1}^{n(x)}$ is a cover of $U_x \times Y$.

In turn, the collection $\{U_x\}_{x \in X}$ is an open cover of X which then must have some finite subcover, which we denote by $\{U_{x_j}\}_{j=1}^m$. We now claim that the collection $\{W_{x_j}^{y_i(x_j)}\}_{ij} \subset \mathcal{C}$ is a finite cover of $X \times Y$. Indeed, for every $(x, y) \in X \times Y$ there is some U_{x_j} that contains x, and then there is some $V_{x_j}^{y_j(x_j)}$ that contains y, implying that $(x, y) \in W_{x_j}^{y_j(x_j)}$. □

Lemma A.11.15. *Let (X, τ) be a compact space and $S \subset X$ infinite. Then S has an accumulation point.*

Proof. Recall that $x \in X$ is an accumulation point of $S \subset X$ if

$$\forall U \in \tau : \quad x \in U \implies (S \cap U) \setminus \{x\} \neq \emptyset.$$

Suppose $S \subset X$ has no accumulation points, i.e.

$$\forall x \in X \ \exists U_x \in \tau : \quad x \in U_x \text{ and } S \cap U_x \subset \{x\}.$$

Now $\mathcal{U} = \{U_x : \ x \in X\}$ is an open cover of X, having a finite subcover $\mathcal{U}' \subset \mathcal{U}$ by compactness. Then

$$S = S \cap \left(\bigcup \mathcal{U}'\right) = \bigcup_{U_x \in \mathcal{U}'} (S \cap U_x).$$

Here the union is finite, and $S \cap U_x \subset \{x\}$ in each case. Thus S is finite. □

A.12 Compact Hausdorff spaces

Next we are going to witness how beautiful compact Hausdorff topologies are. Among topological spaces, Hausdorff spaces are those where points are distinctively separated by open neighbourhoods; this happens especially in metric topology. Roughly, Hausdorff spaces have enough open sets to distinguish between any two points, while compact spaces "do not have too many open sets". Combining these two properties, compact Hausdorff spaces form a useful class of topological spaces.

Definition A.12.1 (**Hausdorff spaces).** A topological space (X, τ) is called a *Hausdorff space* if for each $a, b \in X$, where $a \neq b$, there exists $U, V \in \tau$ such that $a \in U$, $b \in V$ and $U \cap V = \emptyset$.

Example.

1. If τ_1 and τ_2 are topologies of X, $\tau_1 \subset \tau_2$, and (X, τ_1) is a Hausdorff space then (X, τ_2) is a Hausdorff space.
2. $(X, \mathcal{P}(X))$ is a Hausdorff space.
3. If X has more than one point and $\tau = \{\emptyset, X\}$ then (X, τ) is not Hausdorff.
4. Clearly any metric space (X, d) is a Hausdorff space; if $x, y \in X$, $x \neq y$, then $B_r(x) \cap B_r(y) = \emptyset$, when $r \leq d(x, y)/2$.
5. The distribution spaces $\mathcal{D}'(\mathbb{R}^n)$, $\mathcal{S}'(\mathbb{R}^n)$ and $\mathcal{E}'(\mathbb{R}^n)$ are non-metrisable Hausdorff spaces.

Theorem A.12.2. *In Hausdorff spaces, we have the following properties:*

(1) *Every convergent sequence has a unique limit.*

(2) *All finite sets are closed.*

(3) *Every topological subspace is also Hausdorff.*

(4) *A compact subspace of a Hausdorff space is closed.*

(5) *A subset of a compact Hausdorff space is compact if and only if it is closed.*

Proof. To prove (1), let x_n be a sequence such that $x_n \to p$ and $x_n \to q$ as $n \to \infty$. Assume $p \neq q$. Then there exist open sets U, V such that $p \in U, q \in V$ and $U \cap V = \emptyset$. Consequently, there are numbers N and M such that for all $n > N$ we have $x_n \in U$ and for all $n > M$ we have $x_n \in V$, which yields a contradiction.

To prove (2), in view of property (C3) of Corollary A.7.14 it is enough to show that one-point sets $\{x\}$ in a Hausdorff topological space X are closed. For every $y \in X \backslash \{x\}$ there exist open disjoint sets $U_y \ni x$ and $V_y \ni y$. Since $x \notin V_y$ it follows that $V_y \subset X \backslash \{x\}$ and hence $X \backslash \{x\} = \bigcup_{y \in X \backslash \{x\}} V_y$, implying that $X \backslash \{x\}$ is open.

To prove (3), let Y be a subset of a Hausdorff topological space (X, τ) and let τ_Y be the relative topology on Y. Let $a, b \in Y$ be such that $a \neq b$. Since (X, τ) is Hausdorff there exist open disjoint sets $U, V \in \tau$ such that $a \in U$ and $b \in V$. Consequently, $a \in U \cap Y \in \tau_Y$ and $b \in V \cap Y \in \tau_Y$, and $U \cap Y$ and $V \cap Y$ are disjoint, implying that (Y, τ_Y) is Hausdorff.

To prove (4), let Y be a compact subspace of a topological space X. If $Y = X$ the statement is trivial. Assuming that $Y \neq X$, let us take some $x \in X \backslash Y$. Then for every $y \in Y$ there are open disjoint sets $U_y \ni x$ and $V_y \ni y$. The collection $\{V_y\}_{y \in Y}$ is a covering of Y, and hence by Proposition A.11.6 there is a finite collection $V_{y_1}, \dots, V_{y_n}$ still covering Y. Then set $U_x = \cap_{i=1}^n U_{y_i}$ is open, $x \in U_x$, and $U_x \cap Y = \emptyset$. Therefore, $X \backslash Y = \cup_{x \in X \backslash Y} U_x$ is open and hence Y is closed.

Statement (5) follows immediately from (4) and property (1) of Proposition A.11.11. □

Theorem A.12.3 (**Hausdorff property is a topological property**). *Let $f : X_1 \to X_2$ be an injective and continuous mapping between topological spaces (X_1, τ_1) and (X_2, τ_2). If (X_2, τ_2) is Hausdorff then (X_1, τ_1) is also Hausdorff. Consequently, the Hausdorff property is a topological property.*

Proof. Let $x, y \in X_1$ be such that $x \neq y$. Since f is injective, we have $f(x) \neq f(y)$ and since (X_2, τ_2) is Hausdorff there exist open disjoint sets $U, V \in \tau_2$ such that $f(x) \in U$ and $f(y) \in V$. Since f is continuous, sets $f^{-1}(U)$ and $f^{-1}(V)$ are open disjoint neighbourhoods of x and y in X_1, respectively, implying that (X_1, τ_1) is also Hausdorff. That the Hausdorff property is a topological property follows immediately from this. □

Exercise A.12.4 (**Product of Hausdorff spaces**). If (X_1, τ_1) and (X_1, τ_1) are Hausdorff topological spaces, show that $(X_1 \times X_2, \tau_1 \otimes \tau_2)$ is a Hausdorff topological space.

Theorem A.12.5. *Let X be a Hausdorff space, $A, B \subset X$ compact subsets, and $A \cap B = \emptyset$. Then there exist open sets $U, V \subset X$ such that $A \subset U$, $B \subset V$, and $U \cap V = \emptyset$.*

Proof. The proof is trivial if $A = \emptyset$ or $B = \emptyset$. So assume $x \in A$ and $y \in B$. Since X is a Hausdorff space and $x \neq y$, we can choose neighbourhoods $U_{xy} \in \mathcal{V}(x)$ and $V_{xy} \in \mathcal{V}(y)$ such that $U_{xy} \cap V_{xy} = \emptyset$. The collection $\mathcal{P} = \{V_{xy} \mid y \in B\}$ is an open cover of the compact set B, so that it has a finite subcover

$$\mathcal{P}_x = \{V_{xy_j} \mid 1 \leq j \leq n_x\} \subset \mathcal{P}$$

for some $n_x \in \mathbb{N}$. Let

$$U_x := \bigcap_{j=1}^{n_x} U_{xy_j}.$$

Now $\mathcal{O} = \{U_x \mid x \in A\}$ is an open cover of the compact set A, so that it has a finite subcover

$$\mathcal{O}' = \{U_{x_i} \mid 1 \leq i \leq m\} \subset \mathcal{O}.$$

Then define

$$U := \bigcup \mathcal{O}', \quad V := \bigcap_{i=1}^{m} \bigcup \mathcal{P}_{x_i}.$$

It is an easy task to check that U and V have desired properties. □

Corollary A.12.6. *Let X be a compact Hausdorff space, $x \in X$, and $W \in \mathcal{V}(x)$. Then there exists $U \in \mathcal{V}(x)$ such that $\overline{U} \subset W$.*

Proof. Now $\{x\}$ and $X \setminus W$ are closed sets in a compact space, thus they are compact. Since these sets are disjoint, there exist open disjoint sets $U, V \subset X$ such that $x \in U$ and $X \setminus W \subset V$; i.e.

$$x \in U \subset X \setminus V \subset W.$$

Hence $x \in U \subset \overline{U} \subset X \setminus V \subset W$. □

Proposition A.12.7. *Let (X, τ_X) be a compact space and (Y, τ_Y) a Hausdorff space. Any bijective continuous mapping $f : X \to Y$ is a homeomorphism.*

Proof. Let $U \in \tau_X$. Then $X \setminus U$ is closed, hence compact. Consequently, $f(X \setminus U)$ is compact, and due to the Hausdorff property $f(X \setminus U)$ is closed. Therefore $(f^{-1})^{-1}(U) = f(U)$ is open. □

Corollary A.12.8. *Let X be a set with a compact topology τ_2 and a Hausdorff topology τ_1. If $\tau_1 \subset \tau_2$ then $\tau_1 = \tau_2$.*

Proof. The identity mapping $(x \mapsto x) : X \to X$ is a continuous bijection from (X, τ_2) to (X, τ_1). □

A more direct proof of the Corollary. Let $U \in \tau_2$. Since (X, τ_2) is compact and $X \setminus U$ is τ_2-closed, $X \setminus U$ must be τ_2-compact. Now $\tau_1 \subset \tau_2$, so that $X \setminus U$ is τ_1-compact. (X, τ_1) is Hausdorff, implying that $X \setminus U$ is τ_1-closed, thus $U \in \tau_1$; this yields $\tau_2 \subset \tau_1$. □

Definition A.12.9 (Separating points). A family $\mathcal{F}$ of mappings $X \to \mathbb{C}$ is said to *separate the points of the set X* if there exists $f \in \mathcal{F}$ such that $f(x) \neq f(y)$ whenever $x \neq y$.

Definition A.12.10 (Support). The *support* of a function $f \in C(X)$ is the set

$$\mathrm{supp}(f) := \overline{\{x \in X \mid f(x) \neq 0\}}.$$

Let $f \in C(X)$ such that $0 \leq f \leq 1$. Notations

$$K \prec f, \quad f \prec U$$

mean, respectively, that $K \subset X$ is compact and $\chi_K \leq f$, and that $U \subset X$ is open and $\mathrm{supp}(f) \subset U$.

Theorem A.12.11 (Urysohn's Lemma). *Let X be a compact Hausdorff space, $A, B \subset X$ closed non-empty sets, $A \cap B = \emptyset$. Then there exists $f \in C(X)$ and $U \subset X \setminus A$ such that $B \prec f \prec U$. Especially, we find f such that*

$$0 \leq f \leq 1, \quad f(A) = \{0\}, \quad f(B) = \{1\}.$$

Proof. The set $\mathbb{Q} \cap [0,1]$ is countably infinite; let $\phi : \mathbb{N} \to \mathbb{Q} \cap [0,1]$ be a bijection satisfying $\phi(0) = 0$ and $\phi(1) = 1$. Choose open sets $U_0, U_1 \subset X$ such that

$$A \subset U_0 \subset \overline{U_0} \subset U_1 \subset \overline{U_1} \subset X \setminus B.$$

Then we proceed inductively as follows: Suppose we have chosen open sets $U_{\phi(0)}, U_{\phi(1)}, \ldots, U_{\phi(n)}$ such that

$$\phi(i) < \phi(j) \Rightarrow \overline{U_{\phi(i)}} \subset U_{\phi(j)}.$$

Let us choose an open set $U_{\phi(n+1)} \subset X$ such that

$$\phi(i) < \phi(n+1) < \phi(j) \Rightarrow \overline{U_{\phi(i)}} \subset U_{\phi(n+1)} \subset \overline{U_{\phi(n+1)}} \subset U_{\phi(j)}$$

whenever $0 \leq i, j \leq n$. Let us define

$$r < 0 \Rightarrow U_r := \emptyset, \quad s > 1 \Rightarrow U_s := X.$$

Hence for each $q \in \mathbb{Q}$ we get an open set $U_q \subset X$ such that

$$\forall r, s \in \mathbb{Q} : \ r < s \Rightarrow \overline{U_r} \subset U_s.$$

Let us define a function $f : X \to [0,1]$ by

$$f(x) := \inf\{r : \ x \in U_r\}.$$

Clearly $0 \leq f \leq 1$, $f(A) = \{0\}$ and $f(B) = \{1\}$.

Let us prove that f is continuous. Take $x \in X$ and $\varepsilon > 0$. Take $r, s \in \mathbb{Q}$ such that

$$f(x) - \varepsilon < r < f(x) < s < f(x) + \varepsilon;$$

then f is continuous at x, since $x \in U_s \setminus \overline{U_r}$ and for every $y \in U_s \setminus \overline{U_r}$ we have $|f(y) - f(x)| < \varepsilon$. Thus $f \in C(X)$. □

Corollary A.12.12. *Let X be a compact space. Then $C(X)$ separates the points of X if and only if X is Hausdorff.*

Exercise A.12.13. Prove the previous Corollary.

Definition A.12.14 (Partition of unity). A *partition of unity* on $K \subset X$ in a topological space (X, τ) is a family $\mathcal{F} = \{\phi_j : X \to [0,1] \mid j \in J\}$ of continuous functions such that

$$\chi_K \leq \sum_{j \in J} \phi_j \leq 1,$$

where the sum is required to be *locally finite*: for each $x \in X$ there exists $U \in \mathcal{V}(x)$ such that $\mathrm{supp}(\phi_j) \subset X \setminus U$ for all but finitely many $\phi_j \in \mathcal{F}$. Moreover, if now $\phi_j \prec U_j$ for all $j \in J$, where $\mathcal{U} = \{U_j : \ j \in J\}$ is an open cover of X, then $\mathcal{F}$ is called a *partition of unity on K subordinate to $\mathcal{U}$*.

Corollary A.12.15 (**Partition of Unity**). *Let $\mathcal{U}$ be an open cover of a compact set $K \subset X$ in a Hausdorff space (X, τ). Then there exists a partition of unity on K subordinate to $\mathcal{U}$.*

Proof. Assume the non-trivial case $K \neq \emptyset$. Take a finite subcover $\mathcal{U}' = \{U_j \mid 1 \leq j \leq n\} \subset \mathcal{U}$. For $x \in K$, take $j \in \{1, \ldots, n\}$ such that $x \in U_j$; then choose $V_x \in \mathcal{V}(x)$ such that $\overline{V_x} \subset U_j$. Then $\mathcal{O} = \{V_x \mid x \in K\}$ is an open cover of K, thus having a finite subcover $\mathcal{O}' \subset \mathcal{O}$. Let

$$K_j := \bigcup \{\overline{V} \in \mathcal{O}' : \ \overline{V} \subset U_j\}.$$

Urysohn's Lemma provides functions $f_j \in C(X)$ satisfying $K_j \prec f_j \prec U_j$. Again by Urysohn's Lemma, there exists $g \in C(X)$ such that

$$\bigcup_{j=1}^{n} K_j \prec g \prec \left\{ x \in X : \ \sum_{k=1}^{n} f_k(x) > 0 \right\}.$$

Notice that $K \subset \bigcup_{j=1}^{n} K_j$. Let

$$\phi_j := f_j / (1 - g + \sum_{k=1}^{n} f_k).$$

Then $\{\phi_j \in C(X)\}_{j=1}^{n}$ provides a desired partition of unity. □

Exercise A.12.16. In a compact metric space (X, d), Urysohn's Lemma is much easier to obtain: When $A, B \subset X$ are closed and non-empty such that $A \cap B = \emptyset$, define $f : X \to \mathbb{R}$ by

$$f(x) := \min \left\{ 1, \frac{\operatorname{dist}(A, \{x\})}{\operatorname{dist}(A, B)} \right\}.$$

Show that f is continuous, $0 \leq f \leq 1$, $f(A) = \{0\}$ and $f(B) = 1$.

Definition A.12.17 (**Equicontinuity**). Let X be a topological space. A family $\mathcal{F}$ of mappings $f : X \to \mathbb{C}$ is called *equicontinuous at* $p \in X$ if for every $\varepsilon > 0$ there exists a neighbourhood $U \subset X$ of p such that $|f(x) - f(p)| < \varepsilon$ whenever $f \in \mathcal{F}$ and $x \in U$.

Exercise A.12.18. Prove the following Theorem A.12.19. (Hint: a bounded sequence of numbers has a convergent subsequence...)

Theorem A.12.19 (**Arzelà-Ascoli Theorem**). *Let $K \subset \mathbb{R}^n$ be compact. For each $j \in \mathbb{Z}^+$, let $f_j : K \to \mathbb{C}$ be continuous, and assume that $\mathcal{F} = \{f_j \mid j \in \mathbb{Z}^+\}$ is equicontinuous on K. If $\mathcal{F}$ is bounded, i.e.*

$$\sup_{x \in K, j \in \mathbb{Z}^+} |f_j(x)| < \infty,$$

then there is a subsequence $\{f_{j_k} \mid k \in \mathbb{Z}^+\}$ that converges uniformly on K.

A.13 Sequential compactness

In this section, a metric space (X, d) is endowed with its canonical metric topology τ_d.

Proposition A.13.1. *Let (X, d) be a metric space, and let $K \subset X$ be compact. Then K is closed and bounded.*

Proof. Let us assume $K \neq \emptyset$, to avoid a triviality. Let $x_0 \in X$. Then $\mathcal{U} = \{B_k(x_0) \mid k \in \mathbb{Z}^+\}$ is an open cover of K. Due to compactness of K, there is a subcover $\mathcal{U}' = \{B_k(x_0) \mid k \in S\}$, where $S \subset \mathbb{Z}^+$ is finite. Now

$$K \subset \bigcup \mathcal{U}' = \bigcup_{k \in S} B_k(x_0) = B_{\max(S)}(x_0).$$

Therefore $\mathrm{diam}(K) \leq 2\max(S) < \infty$, so K is bounded.

We have to prove that K is closed. Let $x \in X \setminus K$. Then

$$\mathcal{V} := \left\{B_{d(x,y)/2}(y) \mid y \in K\right\}$$

is an open cover of K. By compactness, there is a finite subcover

$$\mathcal{V}' = \left\{B_{d(x,y_j)/2}(y_j)\right\}_{j=1}^{n}.$$

Let $r := \min\{d(x, y_j)/2\}_{j=1}^{n}$. Then

$$B_r(x) \cap K \subset \bigcup_{j=1}^{n} \left(B_r(x) \cap B_{d(x,y_j)/2}(y_j)\right) = \emptyset,$$

so $x \notin c_d(K)$. Thereby $K = c_d(K)$ is closed. □

Exercise A.13.2. Give an example of a bounded non-compact metric space.

Definition A.13.3 (Sequential compactness). A metric space is *sequentially compact* if each of its sequences has a converging subsequence. That is, given a sequence $(x_k)_{k=1}^{\infty}$ in a sequentially compact metric space (X, d), there is a converging sequence $(x_{k_j})_{j=1}^{\infty}$, where $k_{j+1} > k_j \in \mathbb{Z}^+$ for each $j \in \mathbb{Z}^+$.

Theorem A.13.4 (Compact ⇔ sequentially compact in metric spaces). *A metric space (X, d) is compact if and only if it is sequentially compact.*

Proof. Let us assume that $X \neq \emptyset$ is compact. Take a sequence $(x_k)_{k=1}^{\infty}$ in X. If the set $\{x_k : k \in \mathbb{Z}^+\}$ is finite, there exists $y \in X$ such that $y = x_k$ for infinitely many $k \in \mathbb{Z}^+$. Then a desired convergent subsequence is given by $(y, y, y, \ldots)$. Now assume that the set $S := \{x_k : k \in \mathbb{Z}^+\}$ is infinite, so it has an accumulation point $p \in X$ by Lemma A.11.15. Take $k_1 \in \mathbb{Z}^+$ such that $x_{k_1} \in S \cap B_1(p)$. Inductively, take $k_{j+1} > k_j \in \mathbb{Z}^+$ such that $x_{k_{j+1}} \in S \cap B_{1/j}(p)$. Then $d(p, x_{k_{j+1}}) < 1/j \to_{j\to\infty}$

0, so $x_{k_j} \to_{j\to\infty} p$. We have proven that a compact metric space is sequentially compact.

Now let (X,d) be sequentially compact. We want to show that its metric topology is compact. Take an open cover $\mathcal{U} = \{U_\alpha : \alpha \in A\}$ of X. We claim that

$$\exists \varepsilon_0 > 0 \ \forall x \in X \ \exists \alpha \in A : \quad B_{\varepsilon_0}(x) \subset U_\alpha. \tag{A.8}$$

Let us prove this by deducing a contradiction from the logically negated assumption

$$\forall \varepsilon > 0 \ \exists x \in X \ \forall \alpha \in A : \quad B_\varepsilon(x) \not\subset U_\alpha.$$

This would especially imply

$$\forall k \in \mathbb{Z}^+ \ \exists x_k \in X \ \forall \alpha \in A : \quad B_{1/k}(x_k) \not\subset U_\alpha.$$

This gives us a sequence $(x_k)_{k=1}^\infty$, which by sequential compactness has a subsequence $(x_{k_j})_{j=1}^\infty$ converging to a point $p \in X$. Since $\mathcal{U}$ covers X, we have $p \in U_{\alpha_p}$ for some $\alpha_p \in A$. Since U_{α_p} is open, $B_\varepsilon(p) \subset U_{\alpha_p}$ for some $\varepsilon > 0$. But for large enough j,

$$B_{1/k_j}(x_{k_j}) \subset B_\varepsilon(p) \subset U_{\alpha_p}.$$

This is a contradiction, so (A.8) must be true. Now we claim that

$$X \text{ can be covered with finitely many open balls of radius } \varepsilon_0. \tag{A.9}$$

What happens if (A.9) is not true? Then take $x_1 \in X$, and inductively

$$x_{k+1} \in X \setminus \bigcup_{j=1}^{k} B_{\varepsilon_0}(x_j) \neq \emptyset,$$

where the non-emptiness of the set is due to the counter-assumption. Now $d(x_j, x_k) \geq \varepsilon_0 > 0$ if $j \neq k$, so the sequence $(x_k)_{k=1}^\infty$ does not have a convergent subsequence. But this would contradict the sequential compactness. Hence (A.9) must be true. □

Exercise A.13.5. Think why the compactness of X follows from (A.8) and (A.9).

Exercise A.13.6. Show that a compact metric space is complete.

Corollary A.13.7 (Heine–Borel Theorem). *Let $\mathbb{R}^n$ be endowed with its Euclidean topology. Then $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. In any metric topology, compactness implies closedness and boundedness, see Proposition A.13.1. So let $S \subset \mathbb{R}^n$ be non-empty, closed and bounded. We shall prove that it is sequentially compact. Take a sequence $(x_k)_{k=1}^\infty$ in S. By boundedness, there exist $a, b \in \mathbb{R}$ such that $S \subset [a,b]^n =: Q_1$. That is, Q_1 is a closed cube of sidelength $b - a$.

Now we chop Q_1 inductively into pieces. When the cube Q_j has been chosen, we decompose Q_j "dyadically" to a union of 2^n cubes $Q_{j+1,m}$ (here $m \in \{1, \cdots, 2^n\}$), whose interiors are disjoint and whose sidelengths are $2^{-j}(b-a)$. Choose $Q_j \in \{Q_{j+1,m} : j \in \{1, \cdots, 2^n\}\}$ such that $x_k \in Q_{j+1}$ for infinitely many $k \in \mathbb{Z}^+$.

We construct the convergent subsequence $(x_{k_j})_{j=1}^{\infty}$ inductively. Let $k_1 := 1$. Take $k_{j+1} > k_j \in \mathbb{Z}^+$ such that $x_{k_{j+1}} \in Q_{j+1}$. Now $(x_{k_j})_{j=1}^{\infty}$ is a Cauchy sequence, because

$$\begin{cases} Q_1 \supset Q_2 \supset Q_3 \supset \cdots \supset Q_j \supset Q_{j+1} \supset \cdots, \\ \operatorname{diam}(Q_{j+1}) = \sqrt{n}\, 2^{-j}(b-a) \to_{j\to\infty} 0. \end{cases}$$

Due to the completeness of $\mathbb{R}^n$, the Cauchy sequence $(x_{k_j})_{j=1}^{\infty}$ of $S \subset \mathbb{R}^n$ converges to a point $p \in \mathbb{R}^n$. But $p \in S$, because S is closed. Thus S is sequentially compact. □

Corollary A.13.8. *Let (X, τ) be a compact topological space and $f : X \to \mathbb{R}$ continuous. Then there exist* $\max(f(X)), \min(f(X)) \in \mathbb{R}$.

Proof. Assume that $X \neq \emptyset$. By Proposition A.11.12, $f(X) \subset \mathbb{R}$ is compact. By the Heine–Borel Theorem A.13.7, equivalently $f(X) \subset \mathbb{R}$ is closed and bounded. Thereby $\sup(f(X)), \inf(f(X)) \in f(X)$. □

We note that the Heine–Borel theorem can also be proved without referring to the sequential compactness. For simplicity, we show this in the one-dimensional case.

Theorem A.13.9 (Heine-Borel Theorem in 1D). *Closed intervals $[a, b]$ are compact in $\mathbb{R}$ in the Euclidean topology.*

Proof. We will assume $a < b$ since otherwise the statement is trivial. For an open covering $\mathcal{C} = \{U_\alpha\}_{\alpha \in I}$ of $[a, b]$ let $S \subset [a, b]$ be defined by

$$S = \{x \in [a, b] : [a, x] \text{ can be covered by finitely many sets from } \mathcal{C}\}.$$

The statement of the theorem will follow if we show that $b \in S$. Since $S \neq 0$ in view of $a \in S$ and since $S \subset [a, b]$ is bounded, we can define $c = \sup S$ so that $c \in [a, b]$. The statement of the theorem will follow if we show that $c \in S$ and that $c = b$.

To show that $c \in S$, we observe that since $c \in [a, b]$, there is some set $U_c \in \mathcal{C}$ such that $c \in U_c$. Since U_c is open, there is some $\varepsilon > 0$ such that $(c - \varepsilon, c] \subset U_c$. At the same time, since $c - \varepsilon < c = \sup S$, the closed interval $[a, c - \varepsilon]$ can be covered by finitely many sets from $\mathcal{C}$ by the definition of S and c. Consequently, adding U_c to this finite collection of sets from $\mathcal{C}$ we obtain a finite covering of $[a, c]$, implying that $c \in S$.

To show that $c = b$, let us assume that $c < b$. As before, let U_c be such that $c \in U_c \in \mathcal{C}$. Since U_c is open and $c < b$, there is $\varepsilon > 0$ such that $[c, c + 2\varepsilon) \subset U_c$. Since $c \in S$, closed interval $[a, c]$ can be covered by finitely many sets from $\mathcal{C}$, and adding U_c to this finite collection we obtain a finite covering of $[a, c+\varepsilon]$ which is a contradiction with $c = \sup S$. □

Theorem A.13.10 (**$\mathbb{R}$ is complete).** *The real line $\mathbb{R}$ with the Euclidean metric is complete.*

Proof. Let x_n be a Cauchy sequence in $\mathbb{R}$. By Lemma A.9.2, (2), the set $\{x_n\}_{n=1}^{\infty}$ is bounded, i.e. there are some $a, b \in \mathbb{R}$ such that $\{x_n\}_{n=1}^{\infty} \subset [a, b]$. By Heine–Borel Theorem in Corollary A.13.7 or in Theorem A.13.9 the interval $[a, b]$ is compact, and by Exercise A.13.6 it must be complete. Therefore, x_n must have a convergent subsequence, and since it is a Cauchy sequence, the whole sequence is convergent by Lemma A.9.2, (3). □

Corollary A.13.11 (**$\mathbb{R}^n$ is complete).** *The space $\mathbb{R}^n$ is complete with respect to any of the Lipschitz equivalent metrics d_p, $1 \le p \le \infty$.*

Proof. Since all metrics d_p are Lipschitz equivalent, it is enough to take one, e.g. d_∞. Writing $\bar{x}_k = (x_k^{(1)}, \ldots, x_k^{(n)})$ and $d_\infty(\bar{x}_k, \bar{x}_l) = \max_{1\le i\le n} |x_k^{(i)} - x_l^{(i)}|$, we have that $d_\infty(\bar{x}_k, \bar{x}_l) < \varepsilon$ implies $|x_k^{(i)} - x_l^{(i)}| < \varepsilon$ for all $i = 1, \ldots, n$. Thus, if $\bar{x}_k \in \mathbb{R}^n$ is a Cauchy sequence in $\mathbb{R}^n$, it follows that $x_k^{(i)}$ is a Cauchy sequence in $\mathbb{R}$ for all i, and hence it has a limit, say $x^{(i)}$, for all i, by Theorem A.13.10. Writing $\bar{x} = (x^{(1)}, \ldots, x^{(n)})$, we claim that $\bar{x}_k \to \bar{x}$ as $k \to \infty$. Indeed, let $\varepsilon > 0$. Then for all i there is a number N_i such that $k > N_i$ implies $|x_k^{(i)} - x^{(i)}| < \varepsilon$. Therefore, for $k > \max_{1\le i\le n} N_i$, we have $|x_k^{(i)} - x^{(i)}| < \varepsilon$ for all i, which means that $d_\infty(\bar{x}_k, \bar{x}) < \varepsilon$. □

Alternative proof of Theorem A.13.4

We now state several results of independent importance that will give another proof of Theorem A.13.4.

Lemma A.13.12 (**Lebesgue's covering lemma).** *Let $\mathcal{C}$ be an open covering of a sequentially compact metric space (X, d). Then there is $\varepsilon > 0$ such that every ball with radius ε is contained in some set from the covering $\mathcal{C}$. Such ε is called a Lebesgue's number of the covering $\mathcal{C}$.*

Proof. Suppose that no such $\varepsilon > 0$ exists. It means that for every $n \in \mathbb{N}$ there is a ball $B_{1/n}(x_n)$ which is not contained in any set from $\mathcal{C}$. Let $(x_{n_j})_{j=1}^{\infty}$ be a convergent subsequence of $(x_n)_{n=1}^{\infty}$ with some limit $x \in X$, so that $x_{n_j} \to X$ as $j \to \infty$. Let $U \in \mathcal{C}$ be a set in $\mathcal{C}$ containing x. Since U is open, there is some $\delta > 0$ such that $B_{2\delta}(x) \subset U$. Now let N be one of indices n_j such that $d(x_N, x) < \delta$ and such that $\frac{1}{N} < \delta$. We claim that $B_{1/N}(x_N) \subset B_{2\delta}(x)$ which would be a contradiction with our choice of the sequence x_n and the fact that $B_{2\delta}(x) \subset U$. Indeed, if $y \in B_{1/N}(x_N)$, we have

$$d(y, x) \le d(y, x_N) + d(x_N, x) < \frac{1}{N} + \delta < 2\delta,$$

which means that $y \in B_{2\delta}(x)$, completing the proof. □

Lemma A.13.13 (Totally bounded metric spaces). *Let X be a sequentially compact metric space. Then X is totally bounded, which means that for every $\varepsilon > 0$ there are finitely many balls in X with radius ε that cover X.*

Proof. Suppose that there is $\varepsilon > 0$ such that no finitely many balls in X with radius ε cover X. We will now construct a sequence of points in X with no convergent subsequence. Let $x_1 \in X$ be an arbitrary point. Let x_2 be any point in $X \backslash B_\varepsilon(x_1)$. Inductively, suppose we have points $x_1, \ldots, x_n \in X$ such that $x_j \in X \backslash \cup_{i=1}^{j-1} B_\varepsilon(x_i)$. Since the collection $\{B_\varepsilon(x_i)\}_{i=1}^n$ does not cover X, we can always choose some $x_n \in X \backslash \cup_{i=1}^{n-1} B_\varepsilon(x_i)$. All points in this sequence have the property that $d(x_n, x_k) \geq \varepsilon$ for any $n, k \in \mathbb{N}$ which means that sequence $(x_n)_{n=1}^\infty$ can not have any convergent subsequence. □

Exercise A.13.14. In general, a metric space is said to be *totally bounded* if

$$\forall \varepsilon > 0 \ \exists \{x_j \mid j \in \{1, \cdots, n_\varepsilon\}\} \subset X : \quad X = \bigcup_{j=1}^{n_\varepsilon} B_\varepsilon(x_j).$$

Show that a metric space (X, d) is compact if and only if it is bounded and totally bounded.

Alternative proof of Theorem A.13.4. Let (X, d) be a metric space. First we will prove that if X is compact it is sequentially compact. Let $(x_n)_{n=1}^\infty$ be a sequence of points in x. Define $A_n = \{x_n, x_{n+1}, x_{n+2}, \cdots\}$, so that $A_1 = \{x_n\}_{n=1}^\infty$. Let $F_n = \overline{A_n}$. Clearly F_n is a closed set and the intersection of any finite number of sets F_n is non-empty since it contains A_N for some N. Since X is compact, by the finite intersection property in Proposition A.11.5 we have $\cap_{n=1}^\infty F_n \neq \emptyset$. Let now $x \in \cap_{n=1}^\infty F_n$, so that $x \in F_n = \overline{A_n}$ for all n. Using a characterisation of closures in Proposition A.8.10, it follows that every open ball $B_{1/j}(x)$ contains a point $x_{n_j} \in A_{n_j}$ with n_j as large as we want. Therefore, we have a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $d(x_{n_j}, x) < 1/j$, which means that it is a convergent subsequence of $\{x_n\}$.

Let us now prove that if X is sequentially compact it is compact. Let $\mathcal{C}$ be an open cover of X and let $\varepsilon > 0$ be its Lebesgue's number according to Lemma A.13.12. By Lemma A.13.13, X is totally bounded, so that it is covered by finitely many balls $\{B_\varepsilon(x_i)\}_{i=1}^n$. Since ε is a Lebesgue number, for every $i = 1, \ldots, n$, there is some $U_i \in \mathcal{C}$ such that $B_\varepsilon(x_i) \subset U_i$. Consequently, $\{U_i\}_{i=1}^n$ must be a cover for X. □

A.14 Stone–Weierstrass theorem

In the sequel we study densities of subalgebras in $C(X)$. These results will be applied in characterising function algebras among Banach algebras. For material concerning algebras we refer to Chapter D. First we study continuous functions on $[a,b] \subset \mathbb{R}$:

Theorem A.14.1 (Weierstrass Theorem (1885)). *Polynomials are dense in* $C([a,b])$.

Proof. Evidently, it is enough to consider the case $[a,b]=[0,1]$. Let $f \in C([0,1])$, and let $g(x) = f(x) - (f(0) + (f(1)-f(0))x)$; then $g \in C(\mathbb{R})$ if we define $g(x)=0$ for $x \in \mathbb{R} \setminus [0,1]$. For $n \in \mathbb{N}$ let us define $k_n : \mathbb{R} \to [0,\infty)$ by

$$k_n(x) := \begin{cases} \frac{(1-x^2)^n}{\int_{-1}^{1}(1-t^2)^n \, \mathrm{d}t}, & \text{when } |x|<1, \\ 0, & \text{when } |x| \geq 1. \end{cases}$$

Then define $P_n := g * k_n$ (*convolution* of g and k_n), that is

$$\begin{aligned} P_n(x) &= \int_{-\infty}^{\infty} g(x-t)\; k_n(t)\; \mathrm{d}t \;=\; \int_{-\infty}^{\infty} g(t)\; k_n(x-t)\; \mathrm{d}t \\ &= \int_{0}^{1} g(t)\; k_n(x-t)\; \mathrm{d}t, \end{aligned}$$

and from this last expression we see that P_n is a polynomial on $[0,1]$. Notice that P_n is real-valued if f is real-valued. Take any $\varepsilon > 0$. The function g is uniformly continuous, so that there exists $\delta > 0$ such that

$$\forall x, y \in \mathbb{R}: \; |x-y| < \delta \Rightarrow |g(x)-g(y)| < \varepsilon.$$

Let $\|g\| = \max_{t \in [0,1]} |g(t)|$. Take $x \in [0,1]$. Then

$$\begin{aligned} & |P_n(x) - g(x)| \\ = \; & \left| \int_{-\infty}^{\infty} g(x-t)\; k_n(t)\; \mathrm{d}t - g(x) \int_{-\infty}^{\infty} k_n(t)\; \mathrm{d}t \right| \\ = \; & \left| \int_{-1}^{1} (g(x-t) - g(x))\; k_n(t)\; \mathrm{d}t \right| \\ \leq \; & \int_{-1}^{1} |g(x-t) - g(x)|\; k_n(t)\; \mathrm{d}t \\ \leq \; & \int_{-1}^{-\delta} 2\|g\|\; k_n(t)\; \mathrm{d}t + \int_{-\delta}^{\delta} \varepsilon\; k_n(t)\; \mathrm{d}t + \int_{\delta}^{1} 2\|g\|\; k_n(t)\; \mathrm{d}t \\ \leq \; & 4\|g\| \int_{\delta}^{1} k_n(t)\; \mathrm{d}t + \varepsilon. \end{aligned}$$

The reader may verify that $\int_\delta^1 k_n(t)\; \mathrm{d}t \to_{n\to\infty} 0$ for every $\delta > 0$. Hence $\|Q_n - f\| \to_{n\to\infty} 0$, where $Q_n(x) = P_n(x) + f(0) + (f(1)-f(0))x$. □

Exercise A.14.2. Show that $\int_\delta^1 k_n(t)\; \mathrm{d}t \to_{n\to\infty} 0$ in the proof of the Weierstrass Theorem A.14.1.

Definition A.14.3 (Involutive subalgebras). For $f : X \to \mathbb{C}$ let us define $f^* : X \to \mathbb{C}$ by $f^*(x) := \overline{f(x)}$, and define $|f| : X \to \mathbb{C}$ by $|f|(x) := |f(x)|$. A subalgebra $\mathcal{A} \subset \mathcal{F}(X)$ is called *involutive* if $f^* \in \mathcal{A}$ whenever $f \in \mathcal{A}$.

Theorem A.14.4 (Stone–Weierstrass Theorem (1937)). *Let X be a compact space. Let $\mathcal{A} \subset C(X)$ be an involutive subalgebra separating the points of X. Then $\mathcal{A}$ is dense in $C(X)$.*

Proof. If $f \in \mathcal{A}$ then $f^* \in \mathcal{A}$, so that the real part $\mathrm{Re} f = \dfrac{f + f^*}{2}$ belongs to $\mathcal{A}$. Let us define

$$\mathcal{A}_{\mathbb{R}} := \{\mathrm{Re} f \mid f \in \mathcal{A}\};$$

this is a $\mathbb{R}$-subalgebra of the $\mathbb{R}$-algebra $C(X, \mathbb{R})$ of continuous real-valued functions on X. Then

$$\mathcal{A} = \{f + ig \mid f, g \in \mathcal{A}_{\mathbb{R}}\},$$

so that $\mathcal{A}_{\mathbb{R}}$ separates the points of X. If we can show that $\mathcal{A}_{\mathbb{R}}$ is dense in $C(X, \mathbb{R})$ then $\mathcal{A}$ would be dense in $C(X)$.

First we have to show that $\overline{\mathcal{A}_{\mathbb{R}}}$ is closed under taking maximums and minimums. For $f, g \in C(X, \mathbb{R})$ we define

$$\max(f, g)(x) := \max(f(x), g(x)), \quad \min(f, g)(x) := \min(f(x), g(x)).$$

Notice that $\overline{\mathcal{A}_{\mathbb{R}}}$ is an algebra over the field $\mathbb{R}$. Since

$$\max(f, g) = \frac{f + g}{2} + \frac{|f - g|}{2}, \quad \min(f, g) = \frac{f + g}{2} - \frac{|f - g|}{2},$$

it is enough to prove that $|h| \in \overline{\mathcal{A}_{\mathbb{R}}}$ whenever $h \in \overline{\mathcal{A}_{\mathbb{R}}}$. Let $h \in \overline{\mathcal{A}_{\mathbb{R}}}$. By the Weierstrass Theorem A.14.1 there is a sequence of polynomials $P_n : \mathbb{R} \to \mathbb{R}$ such that

$$P_n(x) \to_{n \to \infty} |x|$$

uniformly on the interval $[-\|h\|, \|h\|]$. Thereby

$$\||h| - P_n(h)\| \to_{n \to \infty} 0,$$

where $P_n(h)(x) := P_n(h(x))$. Since $P_n(h) \in \overline{\mathcal{A}_{\mathbb{R}}}$ for every n, this implies that $|h| \in \overline{\mathcal{A}_{\mathbb{R}}}$. Now we know that $\max(f, g), \min(f, g) \in \overline{\mathcal{A}_{\mathbb{R}}}$ whenever $f, g \in \overline{\mathcal{A}_{\mathbb{R}}}$.

Now we are ready to prove that $f \in C(X, \mathbb{R})$ can be approximated by elements of $\mathcal{A}_{\mathbb{R}}$. Take $\varepsilon > 0$ and $x, y \in X$, $x \neq y$. Since $\mathcal{A}_{\mathbb{R}}$ separates the points of X, we may pick $h \in \mathcal{A}_{\mathbb{R}}$ such that $h(x) \neq h(y)$. Let $g_{xx} = f(x)\mathbf{1}$, and let

$$g_{xy}(z) := \frac{h(z) - h(y)}{h(x) - h(y)} f(x) + \frac{h(z) - h(x)}{h(y) - h(x)} f(y).$$

Here $g_{xx}, g_{xy} \in \mathcal{A}_{\mathbb{R}}$, since $\mathcal{A}_{\mathbb{R}}$ is an algebra. Furthermore,

$$g_{xy}(x) = f(x), \quad g_{xy}(y) = f(y).$$

Due to the continuity of g_{xy}, there is an open set $V_{xy} \in \mathcal{V}(y)$ such that

$$z \in V_{xy} \;\Rightarrow\; f(z) - \varepsilon < g_{xy}(z).$$

Now $\{V_{xy} \mid y \in X\}$ is an open cover of the compact space X, so that there is a finite subcover $\{V_{xy_j} \mid 1 \le j \le n\}$. Define

$$g_x := \max_{1 \le j \le n} g_{xy_j};$$

$g_x \in \overline{\mathcal{A}_{\mathbb{R}}}$, because $\overline{\mathcal{A}_{\mathbb{R}}}$ is closed under taking maximums. Moreover,

$$\forall z \in X: \; f(z) - \varepsilon < g_x(z).$$

Due to the continuity of g_x (and since $g_x(x) = f(x)$), there is an open set $U_x \in \mathcal{V}(x)$ such that

$$z \in U_x \;\Rightarrow\; g_x(z) < f(z) + \varepsilon.$$

Now $\{U_x \mid x \in X\}$ is an open cover of the compact space X, so that there is a finite subcover $\{U_{x_i} \mid 1 \le i \le m\}$. Define

$$g := \min_{1 \le i \le m} g_{x_i};$$

$g \in \overline{\mathcal{A}_{\mathbb{R}}}$, because $\overline{\mathcal{A}_{\mathbb{R}}}$ is closed under taking minimums. Moreover,

$$\forall z \in X: \; g(z) < f(z) + \varepsilon.$$

Thus

$$f(z) - \varepsilon < \min_{1 \le i \le m} g_{x_i}(z) = g(z) < f(z) + \varepsilon,$$

that is $|g(z) - f(z)| < \varepsilon$ for every $z \in X$, i.e. $\|g - f\| < \varepsilon$. Hence $\mathcal{A}_{\mathbb{R}}$ is dense in $C(X, \mathbb{R})$ implying that $\mathcal{A}$ is dense in $C(X)$. □

Remark A.14.5. Notice that under the assumptions of the Stone–Weierstrass Theorem, the compact space is actually a compact Hausdorff space, since continuous functions separate the points.

A.15 Manifolds

We now give an example of Hausdorff spaces which is a starting point of the geometric analysis. We will come back to this topic with more details in Section 5.2.

Definition A.15.1 (Manifold). A topological space (X, τ) is called an *n-dimensional (topological) manifold* if it is second countable, Hausdorff and each of its points has a neighbourhood homeomorphic to an open set of the Euclidean space $\mathbb{R}^n$. If $\phi : U \to U'$ is a homeomorphism, where $U \in \tau$ and $U' \subset \mathbb{R}^n$ is open then the pair (U, ϕ) is called a *chart* on X.

Exercise A.15.2. Show that the sphere $\mathbb{S}^n = \left\{x \in \mathbb{R}^{n+1} : \sum_{j=1}^n x_j^2 = 1\right\}$ is an n-dimensional manifold.

Exercise A.15.3. Let X and Y be manifolds of respective dimensions m, n. Show that $X \times Y$ is a manifold of dimension $m+n$.

Definition A.15.4 (**Differentiable manifold).** Let (X, τ) be an n-dimensional manifold. A collection $\mathcal{A} = \{(U_i, \phi_i) : i \in I\}$ of charts on X is called a C^k*-atlas* if $\{U_i : i \in I\}$ is a cover of X and if the mappings

$$\left(x \mapsto \phi_j(\phi_i^{-1}(x))\right) : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

are C^k-smooth whenever $U_i \cap U_j \neq \emptyset$. If there is a C^k-atlas then X is called a C^k*-manifold* (*differentiable manifold*).

A.16 Connectedness and path-connectedness

In this section we discuss notions of connected and path-connected topological spaces and a relation between them.

Proposition A.16.1. *Let (X, τ) be a topological space. Then the following statements are equivalent*

(i) *There exist non-empty open subsets U, V of X such that $U \cap V = \emptyset$ and $U \cup V = X$.*

(ii) *There exists a non-empty subset U of X such that U is open and closed and such that $U \neq X$.*

(iii) *There exists a continuous surjective mapping from X to the set $\{0, 1\}$ equipped with the discrete topology.*

Proof. Statement (i) and (ii) are equivalent if we take $V = X \backslash U$. Let us show that (i) implies (iii). Define a mapping f by $f(x) = 0$ for $x \in U$ and $f(x) = 1$ for $x \in V$. Since U and V are non-empty, the mapping f is surjective. If W is any subset of $\{0, 1\}$, its preimage $f^{-1}(W)$ is one of the sets $\emptyset, U, V, X$. Since all of them are open, f is continuous.

Finally, to show that (iii) implies (i), we set $U = f^{-1}(0)$ and $V = f^{-1}(1)$. Since f is continuous, both sets are open. Moreover, clearly they are disjoint, $U \cup V = X$, and they are non-empty because f is surjective. □

Definition A.16.2 (**Connected topological space).** A topological space (X, τ) is said to be *disconnected* if it satisfies any of the equivalent properties of Proposition A.16.1. Otherwise, it is said to be *connected.*

Proposition A.16.3 (**"Connectedness" is a topological property).** *Let X and Y be topological spaces and let $f : X \to Y$ be continuous. If X is connected, then $f(X)$ is also connected. Consequently, "connectedness" is a topological property.*

Proof. Suppose that $f(X) = U \cup V$ with U, V as in Proposition A.16.1, (i). Then $X = f^{-1}(U) \cup f^{-1}(V)$ and sets $f^{-1}(U), f^{-1}(V)$ satisfy conditions of Proposition A.16.1, (i), yielding a contradiction. □

Exercise A.16.4. Prove that a subset A of a topological space X is disconnected (in the relative topology) if and only if there are open sets U, V in X such that $U \cap A \neq \emptyset$, $V \cap A \neq \emptyset$, $A \subset U \cup V$ and $U \cap V \cap A = \emptyset$.

Proposition A.16.5 (Closures are connected). *Let X be a topological space and let $A \subset X$. If A is connected, then its closure $\overline{A}$ is also connected.*

Proof. Let U and V be open sets in X such that $\overline{A} \subset U \cup V$ and $U \cap V \cap \overline{A} = \emptyset$. Since $A \subset \overline{A}$, we have $A \subset U \cup V$and $U \cap V \cap A = \emptyset$. Since A is connected, by Exercise A.16.4 we must then have either $U \cap A = \emptyset$ or $V \cap A = \emptyset$, which means that either $A \subset X \backslash U$ or $A \subset X \backslash V$. Since the sets $X \backslash U$ and $X \backslash V$ are closed in X, it follows that we have either $\overline{A} \subset X \backslash U$ or $\overline{A} \subset X \backslash V$, which means that either $U \cap \overline{A} = \emptyset$ or $V \cap \overline{A} = \emptyset$. By Exercise A.16.4 again, it means that $\overline{A}$ is connected. □

Definition A.16.6 (Path–connected topological spaces). A topological space X is said to be *path–connected* if for any two points $a, b \in X$ there is a *path* from a to b, i.e. a continuous mapping $\gamma : [0, 1] \to X$ such that $\gamma(0) = a$ and $\gamma(1) = b$.

Theorem A.16.7 (Path–connected $\Longrightarrow$ connected). *A path–connected topological space is connected.*

Exercise A.16.8. Show that the converse is not true. For example, prove that the set $X = \{(0, t) : -1 \leq t \leq 1\} \cup \{(t, \sin \frac{1}{t}), t > 0\}$ in the relative topology of the Euclidean space $\mathbb{R}^2$ is connected but not path–connected.

We first prove a special case of Theorem A.16.7, namely we show that intervals in $\mathbb{R}$ are connected. We then reduce the general case to this one. By an interval in $\mathbb{R}$ we understand any open or closed or half-open, finite or infinite interval.

Theorem A.16.9 (Interval in $\mathbb{R} \Longrightarrow$ connected). *Every interval I in $\mathbb{R}$ with the Euclidean topology is connected.*

Proof. We will prove it by contradiction. Suppose $I = U \cup V$, where U and V are non-empty, disjoint, and open in the relative topology of I. Let $u \in U$, $v \in V$, and assume $u < v$. Since I is an interval we have $[u, v] \subset I$, and we denote

$$A = \{x \in I : \; u \leq x \text{ and } [u, x] \subset U\}.$$

Since $u \in A$, A is non-empty, and since $v \notin U$, A is bounded above. Thus, we can define $w = \sup A$, and we have $[u, w) \subset U$. Since $w \in [u, v]$, we also have $w \in I = U \cup V$, so that either $w \in U$ or $w \in V$.

We will now show that both choices are impossible. Suppose $w \in U$. Then $w < v$ and since U is open, there is some $\delta > 0$ such that $(w - \delta, w + \delta) \cap A \subset U$.

Now, if we take some $z \in (w, w+\delta) \cap A$, we have $[w, z] \subset U$, so that also $[u, z] \in U$, contradicting $w = \sup A$.

Suppose now $w \in V$. Then $u < w$ and since V is open, there is some $\delta > 0$ such that $(w-\delta, w+\delta) \cap A \subset V$. Now, if we take some $z \in (w-\delta, w) \cap A$, we have $(z, w] \subset V$, so that for all $x \in (z, w]$ we have that $[u, x] \not\subset U$, contradicting $w = \sup A$ again. □

Proof of Theorem A.16.7. Let X be a path–connected topological space and let f be a continuous mapping from X to $\{0, 1\}$ equipped with the discrete topology. By Proposition A.16.1 it is enough to show that f must be constant. Without loss of generality, suppose that $f(x) = 0$ for some $x \in X$. Let $y \in X$ and let γ be a path from x to y. Then the composition mapping $f \circ \gamma : [0, 1] \to \{0, 1\}$ is continuous. Since $[0, 1]$ is connected by Theorem A.16.9 it follows that $f \circ \gamma$ can not be surjective, so that $f(y) = f(\gamma(1)) = f(\gamma(0)) = f(x) = 0$. Thus, $f(y) = 0$ for all $y \in X$, which means that f can not be surjective. □

Theorem A.16.9 has the converse:

Theorem A.16.10 (Connected in $\mathbb{R} \implies$ interval). *If I is a connected subset of $\mathbb{R}$ it must be an interval.*

Proof. First we show that $I \subset \mathbb{R}$ is an interval if and only if for any $a, c \in I$ and any $b \in \mathbb{R}$ with $a < b < c$ we must have $b \in I$.

If I is an interval the implication is trivial. Conversely, we will prove that if $a \in I$, then $I \cap [a, \infty)$ is $[a, \infty)$ or $[a, e]$ or $[a, e)$ for some $e \in \mathbb{R}$. If I is not bounded above, then for any $b > a$ there is $c \in I$ such that $c > b$. Hence $b \in I$ by assumption. In this case $I \cap [a, \infty) = [a, \infty)$. So we may assume that I is bounded above and let $e = \sup I$. If $e = a$, then $I \cap [a, \infty) = [a, a]$, so we may assume $e > a$. Then for any b with $a < b < e$ there is some $c \in I$ such that $a < b < c$, and hence $b \in I$ by assumption. Therefore, $I \cap [a, \infty)$ is $[a, e]$ or $[a, e)$ depending on whether $e \in I$ or not. Arguing in a similar way for $I \cap (-\infty, a]$ we get that I must be an interval.

Now, suppose I is not an interval. By the above claim, there exits some $a, c \in I$ and $b \notin I$ such that $a < b < c$. But then $U = I \cap (-\infty, b)$ and $V = I \cap (b, \infty)$ is a decomposition of I into a union of non-empty, open disjoint sets with $U \cup V = I$, contradicting the assumption that I is connected. □

We will now show a converse to Theorem A.16.7, provided that we are dealing with subsets of $\mathbb{R}^n$.

Theorem A.16.11 (Open connected in $\mathbb{R}^n \implies$ path–connected). *Every open connected subset of $\mathbb{R}^n$ with the Euclidean topology is path–connected.*

Proof. First we note that if we have a path γ_1 from a to b and a path γ_2 from b to c, we can glue them together to obtain a path from a to c, e.g. by setting $\gamma(t) = \gamma_1(2t)$ for $0 \le t \le \frac{1}{2}$, and $\gamma(t) = \gamma_2(2t-1)$ for $\frac{1}{2} \le t \le 1$.

Let A be a non-empty open connected set in $\mathbb{R}^n$ (the statement is trivial for the empty set). Take some $a \in A$ and define

$$U = \{b \in A : \text{ there is a path from } a \text{ to } b \text{ in } A\}.$$

We claim that U is open and closed. Indeed, if $b \in U$, we have $B_\varepsilon(b) \subset A$ for some $\varepsilon > 0$. Consequently, for any $c \in B_\varepsilon(b)$ we have a path from a to b in A by the definition of U, and we obviously also have a path from b to c in $B_\varepsilon(b)$ (e.g. just a straight line). Glueing these paths together, we obtain a path from a to c in A, which means that $B_\varepsilon(b) \subset U$ and hence U is open. To show that U is also closed, take some $b \in A \backslash U$. Then we have $B_\varepsilon(b) \subset A$ for some $\varepsilon > 0$. If now $c \in B_\varepsilon(b)$, there is a path from c to b in $B_\varepsilon(b)$. Consequently, there can be no path from a to c because otherwise there would be a path from a to b in A. Thus, $c \in A \backslash U$, implying that $A \backslash U$ is open.

Finally, writing $A = U \cup (A \backslash U)$ as a union of two disjoint open sets, and observing that U contains a and is, therefore, non-empty, it follows that $A \backslash U = \emptyset$ because A is connected. But this means that $A = U$ and hence A is path–connected. □

A.17 Co-induction and quotient spaces

Definition A.17.1 (Co-induced topology). Let X and J be sets, (X_j, τ_j) be topological spaces for every $j \in J$, and $\mathcal{F} = \{f_j : X_j \to X \mid j \in J\}$ be a family mappings. The *$\mathcal{F}$-co-induced topology* of X is the strongest topology τ on X such that the mappings f_j are continuous for every $j \in J$.

Exercise A.17.2. Let τ be the co-induced topology from Definition A.17.1. Show that

$$\tau = \left\{U \subset X \mid \forall j \in J : \ f_j^{-1}(U) \in \tau_j\right\}.$$

Definition A.17.3 (Quotient topology). Let (X, τ_X) be a topological space, and let $\sim$ be an equivalence relation on X. Let

$$[x] := \{y \in X \mid x \sim y\},$$

$$X/\sim := \{[x] \mid x \in X\},$$

and define the *quotient map* $\pi : X \to X/\sim$ by $\pi(x) := [x]$ The *quotient topology* of the *quotient space* $X/\sim$ is the $\{\pi\}$-co-induced topology on $X/\sim$.

Exercise A.17.4. Show that $X/\sim$ is compact if X is compact.

Example. Let $\mathcal{A}$ be a topological vector space and $\mathcal{J}$ its subspace. Let us denote $[x] := x + \mathcal{J}$ for $x \in \mathcal{A}$. Then the quotient topology of $\mathcal{A}/\mathcal{J} = \{[x] \mid x \in \mathcal{A}\}$ is the topology co-induced by the family $\{(x \mapsto [x]) : \mathcal{A} \to \mathcal{A}/\mathcal{J}\}$.

Remark A.17.5. The message of the following Exercise A.17.6 is that if our compact space X is not Hausdorff, we can factor out the inessential information that the continuous functions $f : X \to \mathbb{C}$ do not see, to obtain a compact Hausdorff space related nicely to X.

Exercise A.17.6. Let X be a topological space, and let us define a relation $R \subset X \times X$ by

$$(x, y) \in R \overset{\text{definition}}{\iff} \forall f \in C(X) : \; f(x) = f(y).$$

Prove:
(a) R is an equivalence relation on X.
(b) There is a natural bijection between the sets $C(X)$ and $C(X/R)$.
(c) X/R is a Hausdorff space.
(d) If X is a compact Hausdorff space then $X \cong X/R$.

Exercise A.17.7. For $A \subset X$, let us define the equivalence relation R_A by

$$(x, y) \in R_A \overset{\text{definition}}{\iff} x = y \text{ or } \{x, y\} \subset A.$$

Let X be a topological space, and let $\infty \subset X$ be a closed subset. Prove that the mapping

$$X \setminus \infty \to (X/R_\infty) \setminus \{\infty\}, \quad x \mapsto [x],$$

is a homeomorphism.

Finally, let us state a basic property of co-induced topologies:

Proposition A.17.8. *Let X have the $\mathcal{F}$-co-induced topology, and Y be a topological space. A mapping $g : X \to Y$ is continuous if and only if $g \circ f$ is continuous for every $f \in \mathcal{F}$.*

Proof. If g is continuous then the composed mapping $g \circ f$ is continuous for every $f \in \mathcal{F}$.

Conversely, suppose $g \circ f_j$ is continuous for every $f_j \in \mathcal{F}$, $f_j : X_j \to X$. Let $V \subset Y$ be open. Then

$$f_j^{-1}(g^{-1}(V)) = (g \circ f_j)^{-1}(V) \subset X_j \quad \text{is open;}$$

thereby $g^{-1}(V) = f_j(f_j^{-1}(g^{-1}(V))) \subset X$ is open. □

Corollary A.17.9. *Let X, Y be topological spaces, R be an equivalence relation on X, and endow X/R with the quotient topology. A mapping $f : X/R \to Y$ is continuous if and only if $(x \mapsto f([x])) : X \to Y$ is continuous.* □

A.18 Induction and product spaces

The main theorem of this section is Tihonov's theorem which is a generalisation of Theorem A.11.14 to infinitely many sets. However, we also discuss other topologies induced by infinite families, and some of their properties.

Definition A.18.1 (**Induced topology).** Let X and J be sets, (X_j, τ_j) be topological spaces for every $j \in J$ and $\mathcal{F} = \{f_j : X \to X_j \mid j \in J\}$ be a family of mappings. The $\mathcal{F}$*-induced topology* of X is the weakest topology τ on X such that the mappings f_j are continuous for every $j \in J$.

Example. Let (X, τ_X) be a topological space, $A \subset X$, and let $\iota : A \to X$ be defined by $\iota(a) = a$. Then the $\{\iota\}$-induced topology on A is

$$\tau_X|_A := \{U \cap A \mid U \in \tau_X\}.$$

This is called the *relative topology* of A, see Definition A.7.18. Let $f : X \to Y$. The restriction $f|_A = f \circ \iota : A \to Y$ satisfies $f|_A(a) = f(a)$ for every $a \in A \subset X$.

Exercise A.18.2. Prove **Tietze's Extension Theorem**: *Let X be a compact Hausdorff space, $K \subset X$ closed and $f \in C(K)$. Then there exists $F \in C(X)$ such that $F|_K = f$.* (Hint: approximate F by continuous functions that would exist by Urysohn's lemma.)

Example. Let (X, τ) be a topological space. Let σ be the $C(X) = C(X, \tau)$-induced topology, i.e. the weakest topology on X making the all τ-continuous functions continuous. Obviously, $\sigma \subset \tau$, and $C(X, \sigma) = C(X, \tau)$. If (X, τ) is a compact Hausdorff space it is easy to check that $\sigma = \tau$.

Example. Let X, Y be topological spaces with bases $\mathcal{B}_X, \mathcal{B}_Y$, respectively. Recall that the product topology for $X \times Y = \{(x, y) \mid x \in X,\ y \in Y\}$ has a base

$$\{U \times V \mid U \in \mathcal{B}_X,\ V \in \mathcal{B}_Y\}.$$

This topology is actually induced by the family

$$\{p_X : X \times Y \to X,\ p_Y : X \times Y \to Y\},$$

where the *coordinate projections* p_X and p_Y are defined by $p_X((x, y)) = x$ and $p_Y((x, y)) = y$.

Definition A.18.3 (**Product topology).** Let X_j be a set for every $j \in J$. The *Cartesian product*

$$X = \prod_{j \in J} X_j$$

is the set of the mappings

$$x : J \to \bigcup_{j \in J} X_j \quad \text{such that} \quad \forall j \in J :\ x(j) \in X_j.$$

Due to the Axiom of Choice, X is non-empty if all X_j are non-empty. The mapping

$$p_j : X \to X_j, \quad x \mapsto x_j := x(j),$$

is called the jth *coordinate projection*. Let (X_j, τ_j) be topological spaces. Let $X := \prod_{j \in J} X_j$ be the Cartesian product. Then the $\{p_j \mid j \in J\}$-induced topology on X is called the *product topology* of X.

If $X_j = Y$ for all $j \in J$, it is customary to write

$$\prod_{j \in J} X_j = Y^J = \{f \mid f : J \to Y\}.$$

Let us state a basic property of induced topologies:

Proposition A.18.4. *Let X have the $\mathcal{F}$-induced topology, and Y be a topological space. A mapping $g : Y \to X$ is continuous if and only if $f \circ g$ is continuous for every $f \in \mathcal{F}$.*

Proof. If g is continuous then the composed mapping $f \circ g$ is continuous for every $f \in \mathcal{F}$, by Proposition A.10.10.

Conversely, suppose $f_j \circ g$ is continuous for every $f_j \in \mathcal{F}$, $f : X \to X_j$. Let $y \in Y$, $V \subset X$ be open, $g(y) \in V$. Then there exist $\{f_{j_k}\}_{k=1}^n \subset \mathcal{F}$ and open sets $W_{j_k} \subset X_{j_k}$ such that such that

$$g(y) \in \bigcap_{k=1}^{n} f_{j_k}^{-1}(W_{j_k}) \subset V.$$

Let

$$U := \bigcap_{k=1}^{n} (f_{j_k} \circ g)^{-1}(W_{j_k}).$$

Then $U \subset Y$ is open, $y \in U$, and $g(U) \subset V$; hence $g : Y \to X$ is continuous at an arbitrary point $y \in Y$, i.e. $g \in C(Y, X)$. □

Remark A.18.5 (**Hausdorff preserved in products**). It is easy to see that a Cartesian product of Hausdorff spaces is always Hausdorff: if $X = \prod_{j \in J} X_j$ and $x, y \in X$, $x \neq y$, then there exists $j \in J$ such that $x_j \neq y_j$. Therefore there are open sets $U_j, V_j \subset X_j$ such that

$$x_j \in U_j, \quad y_j \in V_j, \quad U_j \cap V_j = \emptyset.$$

Let $U := p_j^{-1}(U_j)$ and $V := p^{-1}(V_j)$. Then $U, V \subset X$ are open,

$$x \in U, \quad y \in V, \quad U \cap V = \emptyset.$$

Also compactness is preserved in products; this is stated in Tihonov's Theorem (Tychonoff's Theorem). Before proving this we introduce a tool, which can be compared with Proposition A.11.5:

Definition A.18.6 (Non-Empty Finite InterSection (NEFIS) property). Let X be a set. Let $NEFIS(X)$ be the set of those families $\mathcal{F} \subset \mathcal{P}(X)$ such that every finite subfamily of $\mathcal{F}$ has a non-empty intersection. In other words, a family $\mathcal{F} \subset \mathcal{P}(X)$ belongs to $NEFIS(X)$ if and only if $\bigcap \mathcal{F}' \neq \emptyset$ for every finite subfamily $\mathcal{F}' \subset \mathcal{F}$.

Lemma A.18.7. *A topological space X is compact if and only if $\mathcal{F} \notin NEFIS(X)$ whenever $\mathcal{F} \subset \mathcal{P}(X)$ is a family of closed sets satisfying $\bigcap \mathcal{F} = \emptyset$.*

Proof. Let X be a **set**, $\mathcal{U} \subset \mathcal{P}(X)$, and $\mathcal{F} := \{X \setminus U \mid U \in \mathcal{U}\}$. Then

$$\bigcap \mathcal{F} = \bigcap_{U \in \mathcal{U}} (X \setminus U) = X \setminus \bigcup \mathcal{U},$$

so that $\mathcal{U}$ is a cover of X if and only if $\bigcap \mathcal{F} = \emptyset$. Now the claim follows the definition of compactness. □

Theorem A.18.8 (Tihonov's Theorem (1935)). *Let X_j be a compact space for every $j \in J$. Then $X = \prod_{j \in J} X_j$ is compact.*

Proof. To avoid the trivial case, suppose $X_j \neq \emptyset$ for every $j \in J$. Let $\mathcal{F} \in NEFIS(X)$ be a family of closed sets. In order to prove the compactness of X we have to show that $\bigcap \mathcal{F} \neq \emptyset$.

Let

$$P := \{\mathcal{G} \in NEFIS(X) \mid \mathcal{F} \subset \mathcal{G}\}.$$

Let us equip the set P with a partial order relation $\leq$:

$$\mathcal{G} \leq \mathcal{H} \overset{\text{definition}}{\Longleftrightarrow} \mathcal{G} \subset \mathcal{H}.$$

The **Hausdorff Maximal Principle** A.4.9 says that the chain $\{\mathcal{F}\} \subset \mathcal{P}$ belongs to a maximal chain $C \subset P$. The reader may verify that $\mathcal{G} := \bigcup C \in P$ is a maximal element of P.

Notice that the maximal element $\mathcal{G}$ may contain non-closed sets. For every $j \in J$ the family

$$\{p_j(G) \mid G \in \mathcal{G}\}$$

belongs to $NEFIS(X_j)$. Define

$$\mathcal{G}_j := \{\overline{p_j(G)} \mid G \in \mathcal{G}\}.$$

Clearly also $\mathcal{G}_j \in NEFIS(X_j)$, and the elements of $\mathcal{G}_j$ are closed sets in X_j. Since X_j is compact, we have $\bigcap \mathcal{G}_j \neq \emptyset$. Hence, by the **Axiom of Choice** A.4.2, there is an element $x := (x_j)_{j \in J} \in X$ such that

$$x_j \in \bigcap \mathcal{G}_j.$$

We shall show that $x \in \bigcap \mathcal{F}$, which proves Tihonov's Theorem.

If $V_j \subset X_j$ is a neighbourhood of x_j and $G \in \mathcal{G}$ then

$$p_j(G) \cap V_j \neq \emptyset,$$

because $x_j \in \overline{p_j(G)}$. Thus

$$G \cap p_j^{-1}(V_j) \neq \emptyset$$

for every $G \in \mathcal{G}$, so that $\mathcal{G} \cup \{p_j^{-1}(V_j)\}$ belongs to P; the maximality of $\mathcal{G}$ implies that

$$p_j^{-1}(V_j) \in \mathcal{G}.$$

Let $V \in \tau_X$ be a neighbourhood of x. Due to the definition of the product topology,

$$x \in \bigcap_{k=1}^{n} p_{j_k}^{-1}(V_{j_k}) \subset V$$

for some finite index set $\{j_k\}_{k=1}^n \subset J$, where $V_{j_k} \subset X_{j_k}$ is a neighbourhood of x_{j_k}. Due to the maximality of $\mathcal{G}$, any finite intersection of members of $\mathcal{G}$ belongs to $\mathcal{G}$, so that

$$\bigcap_{k=1}^{n} p_{j_k}^{-1}(V_{j_k}) \in \mathcal{G}.$$

Therefore for every $G \in \mathcal{G}$ and $V \in \mathcal{V}_{\tau_X}(x)$ we have

$$G \cap V \neq \emptyset.$$

Hence $x \in \overline{G}$ for every $G \in \mathcal{G}$, yielding

$$x \in \bigcap_{G \in \mathcal{G}} \overline{G} \overset{\mathcal{F} \subset \mathcal{G}}{\subset} \bigcap_{F \in \mathcal{F}} \overline{F} = \bigcap_{F \in \mathcal{F}} F = \bigcap \mathcal{F},$$

so that $\bigcap \mathcal{F} \neq \emptyset$. □

Remark A.18.9. Actually, Tihonov's Theorem A.18.8 is equivalent to the Axiom of Choice A.4.2; we shall not prove this.

A.19 Metrisable topologies

It is often very useful to know whether a topology on a space comes from some metric. Here we try to construct metrics on compact spaces. We shall learn that a compact space X is metrisable if and only if the corresponding normed algebra $C(X)$ is separable. Metrisability is equivalent to the existence of a countable family of continuous functions separating the points of the space. As a vague analogy to manifolds, the reader may view such a countable family as a set of coordinate functions on the space.

Definition A.19.1 (**Metrisable topology).** A topological space (X,τ) is called *metrisable* if there exists a metric d on X such that the topology τ is the canonical metric topology of (X,d), i.e. if there exists a metric d on X such that $\tau=\tau_d$.

Example (Discrete topology). The discrete topology on the set X is the collection τ of all subsets of X. This is a metric topology corresponding to the discrete metric.

Exercise A.19.2. Let X, Y be metrisable. Prove that $X\times Y$ is metrisable, and that

$$(x_n,y_n)\overset{X\times Y}{\to}(x,y)\quad\Leftrightarrow\quad x_n\overset{X}{\to}x \text{ and } y_n\overset{Y}{\to}y.$$

Remark A.19.3. There are plenty of non-metrisable topological spaces, the easiest example being X with more than one point and with $\tau=\{\emptyset,X\}$. If X is an infinite-dimensional Banach space then the weak*-topology[1] of $X':=\mathcal{L}(X,\mathbb{C})$ is not metrisable. The distribution spaces $\mathcal{D}'(\mathbb{R}^n)$, $\mathcal{S}'(\mathbb{R}^n)$ and $\mathcal{E}'(\mathbb{R}^n)$ are non-metrisable topological spaces. We shall later prove that for the compact Hausdorff spaces metrisability is equivalent to the existence of a countable base.

Exercise A.19.4. Show that (X,τ) is a topological space, where

$$\tau=\{U\subset X\mid U=\emptyset,\ \text{or } X\setminus U \text{ is finite}\}.$$

When is this topology metrisable?

Theorem A.19.5. *Let (X,τ) be compact. Assume that there exists a countable family $\mathcal{F}\subset C(X)$ separating the points of X. Then (X,τ) is metrisable.*

Proof. Let $\mathcal{F}=\{f_n\}_{n=0}^{\infty}\subset C(X)$ separate the points of X. We can assume that $|f_n|\le 1$ for every $n\in\mathbb{N}$; otherwise consider for instance functions $x\mapsto f_n(x)/(1+|f_n(x)|)$. Let us define

$$d(x,y):=\sup_{n\in\mathbb{N}}2^{-n}|f_n(x)-f_n(y)|$$

for every $x,y\in X$. Next we prove that $d:X\times X\to[0,\infty)$ is a metric: $d(x,y)=0\Leftrightarrow x=y$, because $\{f_n\}_{n=0}^{\infty}$ is a separating family. Clearly also $d(x,y)=d(y,x)$ for every $x,y\in X$. Let $x,y,z\in X$. We have the triangle inequality:

$$\begin{aligned}
d(x,z) &= \sup_{n\in\mathbb{N}}2^{-n}|f_n(x)-f_n(z)|\\
&\le \sup_{n\in\mathbb{N}}(2^{-n}|f_n(x)-f_n(y)|+2^{-n}|f_n(y)-f_n(z)|)\\
&\le \sup_{m\in\mathbb{N}}2^{-m}|f_m(x)-f_m(y)|+\sup_{n\in\mathbb{N}}2^{-n}|f_n(y)-f_n(z)|\\
&= d(x,y)+d(y,z).
\end{aligned}$$

Hence d is a metric on X.

[1] see Definition B.4.35

Finally, let us prove that the metric topology coincides with the original topology, $\tau_d = \tau$. Let $x \in X$, $\varepsilon > 0$. Take $N \in \mathbb{N}$ such that $2^{-N} < \varepsilon$. Define

$$U_n := f_n^{-1}(B_\varepsilon(f_n(x))) \in \mathcal{V}_\tau(x), \quad U := \bigcap_{n=0}^{N} U_n \in \mathcal{V}_\tau(x).$$

If $y \in U$ then

$$d(x,y) = \sup_{n\in\mathbb{N}} 2^{-n}|f_n(x) - f_n(y)| < \varepsilon.$$

Thus $x \in U \subset B_\varepsilon(x) = \{y \in X \mid d(x,y) < \varepsilon\}$. This proves that the original topology τ is finer than the metric topology τ_d, i.e. $\tau_d \subset \tau$. Combined with the facts that (X,τ) is compact and (X,τ_d) is Hausdorff, this implies that we must have $\tau_d = \tau$, by Corollary A.12.8. □

Corollary A.19.6. *Let X be a compact Hausdorff space. Then X is metrisable if and only if it has a countable basis.*

Proof. Suppose X is a compact space, metrisable with a metric d. Let $r > 0$. Then $\mathcal{B}_r = \{B_d(x,r) \mid x \in X\}$ is an open cover of X, thus having a finite subcover $\mathcal{B}'_r \subset \mathcal{B}_r$. Then $\mathcal{B} := \bigcup_{n=1}^{\infty} \mathcal{B}'_{1/n}$ is a countable basis for X.

Conversely, suppose X is a compact Hausdorff space with a countable basis $\mathcal{B}$. Then the family

$$\mathcal{C} := \{(B_1,B_2) \in \mathcal{B}\times\mathcal{B} \mid \overline{B_1} \subset B_2\}$$

is countable. For each $(B_1,B_2) \in \mathcal{C}$, Urysohn's Lemma (Theorem A.12.11) provides a function $f_{B_1B_2} \in C(X)$ satisfying

$$f_{B_1B_2}(\overline{B_1}) = \{0\} \quad \text{and} \quad f_{B_1B_2}(X \setminus B_2) = \{1\}.$$

Next we show that the countable family

$$\mathcal{F} = \{f_{B_1B_2} : \ (B_1,B_2) \in \mathcal{C}\} \subset C(X)$$

separates the points of X. Indeed, Take $x,y \in X$, $x \neq y$. Then $W := X \setminus \{y\} \in \mathcal{V}(x)$. Since X is a compact Hausdorff space, by Corollary A.12.6 there exists $U \in \mathcal{V}(x)$ such that $\overline{U} \subset W$. Take $B', B \in \mathcal{B}$ such that $x \in B' \subset \overline{B'} \subset B \subset U$. Then $f_{B'B}(x) = 0 \neq 1 = f_{B'B}(y)$. Thus X is metrisable. □

Corollary A.19.7. *Let X be a compact Hausdorff space. Then X is metrisable if and only if $C(X)$ is separable.*

Proof. Suppose X is a metrisable compact space. Let $\mathcal{F} \subset C(X)$ be a countable family separating the points of X (as in the proof of the previous Corollary). Let $\mathcal{G}$ be the set of finite products of functions f for which $f \in \mathcal{F} \cup \mathcal{F}^* \cup \{\mathbf{1}\}$; the set $\mathcal{G} = \{g_j\}_{j=0}^{\infty}$ is countable. The linear span $\mathcal{A}$ of $\mathcal{G}$ is the involutive algebra generated by $\mathcal{F}$ (the smallest $*$-algebra containing $\mathcal{F}$, see Definition D.5.1); due to the Stone–Weierstrass Theorem (see Theorem A.14.4), $\mathcal{A}$ is dense in $C(X)$. If $S \subset \mathbb{C}$ is a countable dense set then

$$\{\lambda_0 \mathbf{1} + \sum_{j=1}^{n} \lambda_j g_j \mid n \in \mathbb{Z}^+,\ (\lambda_j)_{j=0}^{n} \subset S\}$$

is a countable dense subset of $\mathcal{A}$, thereby dense in C(X).

Conversely, assume that $\mathcal{F} = \{f_n\}_{n=0}^{\infty} \subset C(X)$ is a dense subset. Take $x, y \in X$, $x \neq y$. By Urysohn's Lemma (Theorem A.12.11) there exists $f \in C(X)$ such that $f(x) = 0 \neq 1 = f(y)$. Take $f_n \in \mathcal{F}$ such that $\|f - f_n\| < 1/2$. Then

$$|f_n(x)| < 1/2 \quad \text{and} \quad |f_n(y)| > 1/2,$$

so that $f_n(x) \neq f_n(y)$; $\mathcal{F}$ separates the points of X. □

Exercise A.19.8. Prove that a topological space with a countable basis is separable. Prove that a metric space has a countable basis if and only if it is separable.

Exercise A.19.9. There are non-metrisable separable compact Hausdorff spaces! Prove that X,

$$X = \{f : [0,1] \to [0,1] \mid x \leq y \Rightarrow f(x) \leq f(y)\},$$

endowed with a relative topology, is such a space. Hint: Tihonov's Theorem.

A.20 Topology via generalised sequences

Definition A.20.1 (Directed set). A non-empty set J is *directed* if there exists a relation "$\leq$" such that $\leq \subset J \times J$ (where $(x, y) \in \leq$ is usually denoted by $x \leq y$) such that for every $x, y, z \in J$ it holds that

1. $x \leq x$,
2. if $x \leq y$ and $y \leq z$ then $x \leq z$,
3. there exists $w \in J$ such that $x \leq w$ and $y \leq w$.

Definition A.20.2 (Nets and convergence). A *net* (or a *generalised sequence*) in a topological space (X, τ) is a mapping $(j \mapsto x_j) : J \to X$, denoted also by $(x_j)_{j \in J}$, where J is a directed set. If $K \subset J$ is a directed set (with respect to the natural inherited relation $\leq$) then the net $(x_j)_{j \in K}$ is called a *subnet* of the net $(x_j)_{j \in J}$. A net $(x_j)_{j \in J}$ *converges to a point* $p \in X$, denoted by

$$x_j \to p \quad \text{or} \quad x_j \xrightarrow[j \in J]{} p \quad \text{or} \quad \lim_{j \in J} x_j = \lim x_j = p,$$

if for every neighbourhood U of p there exists $j_U \in J$ such that $x_j \in U$ whenever $j_U \leq j$.

Example. A *sequence* $(x_j)_{j\in\mathbb{Z}^+}$ is a net, where $\mathbb{Z}^+$ is directed by the usual partial order; sequences characterise topology in spaces of countable local bases, for instance metric spaces. But there are more complicated topologies, where sequences are not enough; for instance, the weak*-topology for the dual of an infinite-dimensional topological vector space.

Exercise A.20.3 (Nets and closure). Let X be a topological space. Show that $p \in X$ belongs to the closure of $S \subset X$ if and only if there exists a net $(x_j)_{j\in J} : J \to S$ such that $x_j \to p$.

Exercise A.20.4 (Nets and continuity). Show that a function $f : X \to Y$ is continuous at $p \in X$ if and only if $f(x_j) \to f(p)$ whenever $x_j \to p$ for nets $(x_j)_{j\in J}$ in X.

Exercise A.20.5 (Nets and compactness). Show that a topological space X is compact if and only if its every net has a converging subnet.

Exercise A.20.6. In the spirit of Exercises A.20.3, A.20.4 and A.20.5, express other topological concepts via nets.

Chapter B

Elementary functional analysis

We assume that the reader has already knowledge of (complex) matrices, determinants, etc. In this chapter, we shall present basic machinery for dealing with vector spaces, especially Banach and Hilbert spaces. We do not go into depth in this direction as there are plenty excellent specialised monographs available devoted to various aspects of the subject, see e.g. [11, 35, 53, 59, 63, 70, 84, 86, 87, 113, 131, 143, 150]. However, we still make an independent presentation of a collection of results which are indispensable for anyone working in analysis, and which are useful for other parts of this book.

B.1 Vector spaces

Definition B.1.1 (Vector space). Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. A $\mathbb{K}$-*vector space* (or a *vector space over the field* $\mathbb{K}$, or a *vector space* if $\mathbb{K}$ is implicitely known) is a set V endowed with mappings

$$\begin{aligned}((x, y) \mapsto x + y) &: V \times V \to V, \\ ((\lambda, x) \mapsto \lambda x) &: \mathbb{K} \times V \to V\end{aligned}$$

such that there exists an *origin* $0 \in V$ and such that the following properties hold:

$$\begin{aligned}&(x + y) + z = x + (y + z), \\ &x + 0 = x, \\ &x + (-1)x = 0, \\ &x + y = y + x, \\ &1x = x, \\ &\lambda(\mu x) = (\lambda\mu)x, \\ &\lambda(x + y) = \lambda x + \lambda y, \\ &(\lambda + \mu)x = \lambda x + \mu x\end{aligned}$$

for all $x, y, z \in V$ and $\lambda, \mu \in \mathbb{K}$. We may write $x + y + z := (x + y) + z$ and $-x := (-1)x$. Elements of a vector space are called *vectors*.

Definition B.1.2 (Convex and balanced sets). A subset C of a vector space is *convex* if $tx + (1-t)y \in C$ for every $x, y \in C$ whenever $0 < t < 1$. A subset B of a vector space is *balanced* if $\lambda x \in B$ for every $x \in B$ whenever $|\lambda| \leq 1$.

Example. $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ is itself a vector space over $\mathbb{K}$, likewise $\mathbb{K}^n$ with operations $(x_k)_{k=1}^n + (y_k)_{k=1}^n := (x_k + y_k)_{k=1}^n$ and $\lambda (x_k)_{k=1}^n := (\lambda x_k)_{k=1}^n$.

Example. Let V be a $\mathbb{K}$-vector space and $X \neq \emptyset$. The set V^X of mappings $f : X \to V$ is a $\mathbb{K}$-vector space with pointwise operations $(f+g)(x) := f(x) + g(x)$ and $(\lambda f)(x) := \lambda\, f(x)$. The vector space $\mathbb{K}^n$ can be naturally identified with $\mathbb{K}^X$, where $X = \{k \in \mathbb{Z}^+ : \ k \leq n\}$.

Example. Let V be a vector space such that its vector operations restricted to $W \subset V$ endow this subset with the vector space structure. Then W is called a *vector subspace*. A vector space V has always *trivial subspaces* $\{0\}$ and V. The vector space V^X has e.g. the subspace $\{f : X \to V \mid \forall x \in K : \ f(x) = 0\}$, where $K \subset X$ is a fixed subset.

Definition B.1.3 (Algebraic basis). Let V be a vector space and $S \subset V$. Let us denote

$$\sum_{x \in S} \lambda(x) x = \sum_{x \in S:\ \lambda(x) \neq 0} \lambda(x) x,$$

when $\lambda : S \to \mathbb{K}$ is finitely supported, i.e. $\{x \in S : \ \lambda(x) \neq 0\}$ is finite. The *span* of a subset S of a vector space V is

$$\operatorname{span}(S) := \left\{ \sum_{x \in S} \lambda(x) x \,\middle|\, \lambda : S \to \mathbb{K} \text{ finitely supported} \right\}.$$

Thus $\operatorname{span}(S)$ is the smallest subspace containing $S \subset V$. A subset S of a $\mathbb{K}$-vector space is said to be *linearly independent* if

$$\sum_{x \in S} \lambda(x) x = 0 \quad \Rightarrow \quad \lambda \equiv 0.$$

A subset is *linearly dependent* if it is not linearly independent. A subset $S \subset V$ is called an *algebraic basis* (or a *Hamel basis*) of V if S is linearly independent and $V = \operatorname{span}(S)$.

Remark B.1.4. Let $\mathcal{B}$ be an algebraic basis for V. Then there exists a unique set of functions $(x \mapsto \langle x, b \rangle_{\mathcal{B}}) : V \to \mathbb{K}$ such that

$$x = \sum_{b \in \mathcal{B}} \langle x, b \rangle_{\mathcal{B}}\ b$$

for every $x \in V$. Notice that $\langle x, b \rangle_{\mathcal{B}} \neq 0$ for at most finitely many $b \in \mathcal{B}$. Consider this e.g. with respect to the vector space $\mathbb{K}^X$ in the example before.

Example. The canonical algebraic basis for $\mathbb{K}^n$ is $\{e_k\}_{k=1}^n$, where $e_k = (\delta_{jk})_{j=1}^n$ and $\delta_{kk} = 1$ and $\delta_{jk} = 0$ otherwise.

Lemma B.1.5. *Every vector space $V \neq \{0\}$ has an algebraic basis. Moreover, any two algebraic bases have the same cardinality*[1].

Proof. Let $\mathcal{F}$ be the family of all linearly independent subsets of V. Now $\mathcal{F} \neq \emptyset$, because $\{x\} \in \mathcal{F}$ for every $x \in V \setminus \{0\}$. Endow $\mathcal{F}$ with a partial order by inclusion. Let $\mathcal{C} \subset \mathcal{F}$ be a chain and let $F := \bigcup \mathcal{C}$. It is easy to check that $F \in \mathcal{F}$ is an upper bound for $\mathcal{C}$. Thereby there is a maximal element $M \in \mathcal{F}$. Obviously, M is an algebraic basis for V.

Let $\mathcal{A}, \mathcal{B}$ be algebraic bases for V. The reader may prove (by induction) that $\mathrm{card}(\mathcal{A}) = \mathrm{card}(\mathcal{B})$ when $\mathcal{A}$ is finite. So suppose $\mathrm{card}(\mathcal{A}) \leq \mathrm{card}(\mathcal{B})$, where $\mathcal{A}$ is infinite. Now $\mathrm{card}(\mathcal{A}) = \mathrm{card}(S)$, where

$$S := \{(a, b) \in \mathcal{A} \times \mathcal{B} : \ \langle a, b \rangle_{\mathcal{B}} \neq \emptyset\}.$$

Assume $\mathrm{card}(\mathcal{A}) < \mathrm{card}(\mathcal{B})$. Thus

$$\exists b_0 \in \mathcal{B} \ \forall a \in \mathcal{A} : \ \langle a, b_0 \rangle_{\mathcal{B}} = 0.$$

But then

$$\begin{aligned} b_0 &= \sum_{a \in \mathcal{A}} \langle b_0, a \rangle_{\mathcal{A}} \ a \\ &= \sum_{a \in \mathcal{A}} \langle b_0, a \rangle_{\mathcal{A}} \sum_{b \in \mathcal{B}} \langle a, b \rangle_{\mathcal{B}} \ b \\ &= \sum_{b \in \mathcal{B}} \left(\sum_{a \in \mathcal{A}} \langle b_0, a \rangle_{\mathcal{A}} \ \langle a, b \rangle_{\mathcal{B}} \right) b \\ &= \sum_{b \in \mathcal{B} \setminus \{b_0\}} \left(\sum_{a \in \mathcal{A}} \langle b_0, a \rangle_{\mathcal{A}} \ \langle a, b \rangle_{\mathcal{B}} \right) b \\ &\in \mathrm{span}(\mathcal{B} \setminus \{b_0\}), \end{aligned}$$

contradicting the linear independence of $\mathcal{B}$. Thus $\mathrm{card}(A) = \mathrm{card}(B)$. □

Definition B.1.6 (Algebraic dimension). By Lemma B.1.5, we may define the *algebraic dimension* $\dim(V)$ of a vector space V to be the cardinality of any of its algebraic bases. The vector space V is said to be *finite-dimensional* if $\dim(V)$ is finite, and *infinite-dimensional* otherwise.

---

[1] Here we will use the notation $\mathrm{card}(\mathcal{A})$ for the cardinality of $\mathcal{A}$ to avoid any confusion with the notation for norms; see Definition A.4.13.

Definition B.1.7 (**Linear operators and functionals).** Let V, W be $\mathbb{K}$-vector spaces. A mapping $A : V \to W$ is called a *linear operator* (or a linear mapping), denoted $A \in L(V, W)$, if

$$\begin{cases} A(u+v) = A(u) + A(v), \\ A(\lambda v) = \lambda\ A(v) \end{cases}$$

for every $u, v \in V$ and $\lambda \in \mathbb{K}$. Then it is customary to write $Av := A(v)$, and $L(V) := L(V, V)$. A linear mapping $f : V \to \mathbb{K}$ is called a *linear functional.*

Definition B.1.8 (**Kernel and image).** The *kernel* $\mathrm{Ker}(A) \subset V$ of a linear operator $A : V \to W$ is defined by

$$\mathrm{Ker}(A) := \{u \in V : Au = 0\},$$

where 0 is the origin of the vector space W. The image $\mathrm{Im}(A) \subset W$ of A is defined by

$$\mathrm{Im}(A) := \{Au : u \in V\}.$$

Exercise B.1.9. Show that $\mathrm{Ker}(A)$ is a vector subspace of V and that $\mathrm{Im}(A)$ is a vector subspace of W.

Exercise B.1.10. Let $C \subset V$ be convex and $A \in L(V, W)$. Show that $A(C) \subset W$ is convex.

Definition B.1.11 (**Spectrum of an operator).** Let V be a $\mathbb{K}$-vector space. Let $I \in L(V)$ denote the identity operator $(x \mapsto x) : V \to V$. The *spectrum* of $A \in L(V)$ is

$$\sigma(A) := \{\lambda \in \mathbb{K} :\ \lambda I - A \text{ is not bijective}\}.$$

Exercise B.1.12. Appealing to the Fundamental Theorem of Algebra, show that $\sigma(A) \neq \emptyset$ for $A \in L(\mathbb{C}^n)$.

Exercise B.1.13. Give an example, where $\sigma(A) = \emptyset \neq \sigma(A^2)$.

Exercise B.1.14. Show that $\sigma(A) = \{0\}$ if A is nilpotent, i.e. if $A^k = 0$ for some $k \in \mathbb{Z}^+$.

Exercise B.1.15. Show that $\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}$ in general, and that $\sigma(AB) = \sigma(BA)$ if A is bijective.

Definition B.1.16 (**Quotient vector space).** Let M be a subspace of a $\mathbb{K}$-vector space V. Let us endow the quotient set $V/M := \{x + M \mid x \in V\}$ with the operations

$$\begin{aligned} ((x+M, y+M) \mapsto x+y+M) \quad &: \quad V/M \times V/M \to V/M, \\ (\lambda, (y+M)) \mapsto \lambda y + M) \quad &: \quad \mathbb{K} \times V/M \to V/M. \end{aligned}$$

Then it is easy to show that with these operations, this so-called *quotient vector space* is indeed a vector space.

Remark B.1.17. In the case of a topological vector space V (see Definition B.2.1), the quotient V/M is endowed with the quotient topology, and then V/M is a topological vector space if and only if the original subspace $M \subset V$ was closed.

B.1.1 Tensor products

The basic idea in multilinear algebra is to linearise multilinear operators. The functional analytic foundation is provided by tensor products that we condensely review here. We also introduce locally convex spaces and Fréchet spaces as well as Montel and nuclear spaces.

Definition B.1.18 (Bilinear mappings). Let X_j $(1 \le j \le r)$ and V be $\mathbb{K}$-vector spaces (that is, vector spaces over the field $\mathbb{K}$). A mapping $A : X_1 \times X_2 \to V$ is 2-*linear* (or *bilinear*) if $x \mapsto A(x, x_2)$ and $x \mapsto A(x_1, x)$ are linear mappings for each $x_j \in X_j$. The reader may guess what conditions an r-*linear* mapping

$$X_1 \times \cdots \times X_r \to V$$

satisfies.

Definition B.1.19 (Tensor product of spaces). The *algebraic tensor product* of $\mathbb{K}$-vector spaces $X_1, \ldots, X_r$ is a $\mathbb{K}$-vector space V endowed with an r-linear mapping i such that for every $\mathbb{K}$-vector space W and for every r-linear mapping

$$A : X_1 \times \cdots \times X_r \to W$$

there exists a (unique) linear mapping $\tilde{A} : V \to W$ satisfying $\tilde{A}i = A$. (The reader is encouraged to draw a commutative diagram involving the vector spaces and mappings $i, A, \tilde{A}$!) Any two tensor products for $X_1, \ldots, X_r$ can easily be seen isomorphic, so that we may denote *the* tensor product of these vector spaces by

$$X_1 \otimes \cdots \otimes X_r.$$

In fact, such a tensor product always exists. Indeed, let X, Y be $\mathbb{K}$-vector spaces. We may formally define the set $B := \{x \otimes y \mid x \in X,\ y \in Y\}$, where $x \otimes y = a \otimes b$ if and only if $x = a$ and $y = b$. Let Z be the $\mathbb{K}$-vector space with basis B, i.e.

$$\begin{aligned} Z &= \operatorname{span}\{x \otimes y \mid x \in X,\ y \in Y\} \\ &= \left\{ \sum_{j=0}^{n} \lambda_j (x_j \otimes y_j) :\ n \in \mathbb{N},\ \lambda_j \in \mathbb{K},\ x_j \in X,\ y_j \in Y \right\}. \end{aligned}$$

Let

$$\begin{aligned} [0 \otimes 0] \quad := \quad \operatorname{span}\big\{ & \alpha_1\beta_1(x_1 \otimes y_1) + \alpha_1\beta_2(x_1 \otimes y_2) \\ & + \alpha_2\beta_1(x_2 \otimes y_1) + \alpha_2\beta_2(x_2 \otimes y_2) \\ & - (\alpha_1 x_1 + \alpha_2 x_2) \otimes (\beta_1 y_1 + \beta_2 y_2) : \\ & \alpha_j, \beta_j \in \mathbb{K},\ x_j \in X,\ y_j \in Y \big\}. \end{aligned}$$

For $z \in Z$, let $[z] := z + [0 \otimes 0]$. The *tensor product* of X, Y is the $\mathbb{K}$-vector space

$$X \otimes Y := Z/[0 \otimes 0] = \{[z] \mid z \in Z\},$$

where $([z_1], [z_2]) \mapsto [z_1 + z_2]$ and $(\lambda, [z]) \mapsto [\lambda z]$ are well-defined mappings $(X \otimes Y) \times (X \otimes Y) \to X \otimes Y$ and $\mathbb{K} \times (X \otimes Y) \to X \otimes Y$, respectively.

Definition B.1.20 (Tensor product of operators). Let X, Y, V, W be $\mathbb{K}$-vector spaces, and let $A : X \to V$ and $B : Y \to W$ be linear operators. The *tensor product* of A, B is the linear operator $A \otimes B : X \otimes Y \to V \otimes W$, which is the unique linear extension of the mapping $x \otimes y \mapsto Ax \otimes By$, where $x \in X$ and $y \in Y$.

Example. Let X and Y be finite-dimensional $\mathbb{K}$-vector spaces with bases $\{x_i\}_{i=1}^{\dim(X)}$ and $\{y_j\}_{j=1}^{\dim(Y)}$, respectively. Then $X \otimes Y$ has a basis

$$\{x_i \otimes y_j \mid 1 \le i \le \dim(X),\ 1 \le j \le \dim(Y)\}.$$

Let S be a finite set. Let $\mathcal{F}(S)$ be the $\mathbb{K}$-vector space of functions $S \to \mathbb{K}$; it has a basis $\{\delta_x \mid x \in S\}$, where $\delta_x(y) = 1$ if $x = y$, and $\delta_x(y) = 0$ otherwise. Now it is easy to see that for finite sets S_1, S_2 the vector spaces $\mathcal{F}(S_1) \otimes \mathcal{F}(S_2)$ and $\mathcal{F}(S_1 \times S_2)$ are isomorphic; for $f_j \in \mathcal{F}(S_j)$, we may regard $f_1 \otimes f_2 \in \mathcal{F}(S_1) \otimes \mathcal{F}(S_2)$ as a function $f_1 \otimes f_2 \in \mathcal{F}(S_1 \times S_2)$ by

$$(f_1 \otimes f_2)(x_1, x_2) := f_1(x_1)\ f_2(x_2).$$

Definition B.1.21 (Inner product on $V \otimes W$). Suppose V, W are finite-dimensional inner product spaces over $\mathbb{K}$. The natural inner product for $V \otimes W$ is obtained by extending

$$\langle v_1 \otimes w_1, v_2 \otimes w_2 \rangle_{V \otimes W} := \langle v_1, v_2 \rangle_V\ \langle w_1, w_2 \rangle_W.$$

Definition B.1.22 (Duals of tensor product spaces). The dual $(V \otimes W)'$ of a tensor product space $V \otimes W$ is naturally identified with $V' \otimes W'$.

Alternative approach to tensor products

Now we briefly describe another approach to tensor products.

Definition B.1.23 (Algebraic tensor product). Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Let X, Y be $\mathbb{K}$-vector spaces, and X', Y' their respective algebraic duals, i.e. the spaces of linear functionals $X \to \mathbb{K}$ and $Y \to \mathbb{K}$. For $x \in X$ and $y \in Y$, define the bilinear functional $x \otimes y : X' \times Y' \to \mathbb{K}$ by

$$(x \otimes y)(x', y') := x'(x)\ y'(y).$$

Let $B(X', Y')$ denote the space of all bilinear functionals $X' \times Y' \to \mathbb{K}$. The *algebraic tensor product* (or simply the *tensor product*) $X \otimes Y$ is the vector subspace of $B(X', Y')$ which is spanned by the set $\{x \otimes y :\ x \in X,\ y \in Y\}$.

Exercise B.1.24. Show that $a : (X \otimes Y)' \to B(X,Y)$ is a linear bijection, where $a(f)(x \otimes y) := f(x,y)$ for $f \in (X \otimes Y)'$, $x \in X$ and $y \in Y$.

Exercise B.1.25. Let X, Y, Z be $\mathbb{K}$-vector spaces. Let $B(X,Y;Z)$ denote the vector space of bilinear mappings $X \times Y \to Z$. Find a linear bijection $B(X,Y;Z) \to L(X \otimes Y, Z)$, where $L(V,Z)$ is the vector space of linear mappings $V \to Z$.

B.2 Topological vector spaces

Vector spaces can be combined with topology. For reader's convenience, if one has not encountered Banach and Hilbert spaces yet, we suggest skipping the sections on topological vector spaces and locally convex spaces at this point, and return here only later.

Definition B.2.1 (Topological vector space). A *topological vector space* V over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ is both a topological space and a vector space over $\mathbb{K}$ such that $\{0\} \subset V$ is closed and such that the mappings

$$\begin{aligned} ((\lambda, x) \mapsto \lambda x) &\quad : \quad \mathbb{K} \times V \to V, \\ ((x,y) \mapsto x + y) &\quad : \quad V \times V \to V \end{aligned}$$

are continuous. The *dual space* V' of a topological vector space V consists of continuous linear functionals $f : V \to \mathbb{K}$.

Exercise B.2.2. Show that a topological vector space is a Hausdorff space.

Exercise B.2.3. Show that in a topological vector space every neighbourhood of 0 contains a balanced neighbourhood of 0.

Exercise B.2.4. Prove that a topological vector space V is metrisable if and only if it has a countable family $\{U_j\}_{j=1}^{\infty}$ of neighbourhoods of $0 \in V$ such that $\bigcap_{j=1}^{\infty} U_j = \{0\}$. Moreover, show that in this case a compatible metric $d : V \times V \to [0, \infty)$ can be chosen translation-invariant in the sense that $d(x+z, y+z) = d(x,y)$ for every $x, y, z \in V$.

Definition B.2.5 (Equicontinuity in vector space). Let X be a topological space and V a topological vector space. A family $\mathcal{F}$ of mappings $f : X \to V$ is called *equicontinuous at* $p \in X$ if for every neighbourhood $W \subset V$ of $f(p)$ there exists a neighbourhood $U \subset X$ of p such that $f(x) \in W$ whenever $f \in \mathcal{F}$ and $x \in U$.

Remark B.2.6 (**NEFIS property and compactness**). Recall the Non-Empty Finite Intersection property (NEFIS) from Definition A.18.6: that is, we denote $NEFIS(X)$ the set of those families $\mathcal{F} \subset \mathcal{P}(X)$ such that every finite subfamily of $\mathcal{F}$ has a non-empty intersection. Recall also that a topological space X is compact if and only if $\bigcap \mathcal{F} \neq \emptyset$ whenever $\mathcal{F} \in NEFIS(X)$ consists of closed sets.

Definition B.2.7 (Small sets property). Let X be a topological vector space. A family $\mathcal{F} \subset \mathcal{P}(X)$ is said to *contain small sets* if for every neighbourhood U of $0 \in X$ there exists $x \in X$ and $S \in \mathcal{F}$ such that $S \subset x + U$.

Definition B.2.8 (Completeness of topological vector spaces). A subset S of a topological vector space X is called *complete* if $\bigcap \mathcal{F} \neq \emptyset$ whenever $\mathcal{F} \in NEFIS(X)$ consists of closed subsets of S and contains small sets.

Exercise B.2.9 (Completeness and Cauchy nets). A net $(x_j)_{j\in J}$ in a topological vector space X is called *a Cauchy net* if for every neighbourhood V of $0 \in X$ there exists $k = k_V \in J$ such that $x_i - x_j \in V$ whenever $k \leq i, j$. Prove that $S \subset X$ is complete if and only if each Cauchy net in S converges to a point in S.

Exercise B.2.10. Show that a complete subset of a topological vector space is closed, and that a closed subset of a complete topological vector space is complete.

Exercise B.2.11 (Completeness and Cartesian product). Let X_j be a topological vector space for each $j \in J$. Show that the product space $X = \prod_{j\in J} X_j$ is a complete if and only if X_j is complete for every $j \in J$.

Definition B.2.12 (Total boundedness in topological vector spaces). A subset S of a topological vector space X is *totally bounded* if for every neighbourhood U of $0 \in X$ there exists a finite set $F \subset X$ such that $S \subset F + U$.

Exercise B.2.13 (Hausdorff Total Boundedness Theorem). Prove the following *Hausdorff Total Boundedness Theorem*: *A subset of a topological vector space is compact if and only if it is totally bounded and complete.*

Definition B.2.14 (Completion of a topological vector space). A *completion of a topological vector space* X is an injective open continuous linear mapping $\iota : X \to \widehat{X}$, where $\iota(X)$ is a dense subset of the complete topological vector space $\widehat{X}$.

Exercise B.2.15 (Existence and uniqueness of completion). Let X be a topological vector space. Show that it has a completion $\iota : X \to \widehat{X}$, and that this completion is unique in the following sense: if $\kappa : X \to Z$ is another completion, then the linear mapping $(\iota(x) \mapsto \kappa(x)) : \iota(X) \to Z$ has a unique continuous extension to an isomorphism $\widehat{X} \to Z$ of topological vector spaces.

Exercise B.2.16 (Extension of continuous linear operators). Let $A : X \to Y$ be continuous and linear, where the topological vector spaces X, Y have respective completions $\iota_X : X \to \widehat{X}$ and $\iota_Y : Y \to \widehat{Y}$. Show that there exists a unique continuous linear mapping $\widehat{A} : \widehat{X} \to \widehat{Y}$ such that $\widehat{A} \circ \iota_X = \iota_Y \circ A$, i.e. that the following diagram is commutative:

$$\begin{array}{ccc} X & \xrightarrow{A} & Y \\ {\scriptstyle \iota_X}\downarrow & & \downarrow{\scriptstyle \iota_Y} \\ \widehat{X} & \xrightarrow{\widehat{A}} & \widehat{Y} \end{array}$$

B.3 Locally convex spaces

A locally convex space is a topological vector space where a local base for the topology can be given by convex neighbourhoods. If the reader is not familiar with Banach and Hilbert spaces yet, we suggest first examining those concepts, and returning to this section only afterwards. In the sequel, we present some essential results for locally convex spaces in a series of exercises of widely varying difficulty, for which the reader may find help e.g. from [63], [86] and [131].

Definition B.3.1 (**Locally convex spaces).** A topological vector space V (over $\mathbb{K}$) is called *locally convex* if for every neighbourhood U of $0 \in V$ there exists a convex neighbourhood C such that $0 \in C \subset U$.

Exercise B.3.2. Show that in a locally convex space each neighbourhood of 0 contains a convex balanced neighbourhood of 0.

Exercise B.3.3. Let $\mathcal{U}$ be the family of all convex balanced neighbourhoods of 0 in a topological vector space V. For $U \in \mathcal{U}$, define a so-called *Minkowski functional* $p_U : V \to [0, \infty)$ by

$$p_U(x) := \inf \left\{ \lambda \in \mathbb{R}^+ : \ x/\lambda \in U \right\}.$$

Show that p_U is a seminorm (see Definition B.4.1). Moreover, prove that V is locally convex if and only if its topology is induced by the family

$$\{p_U : V \to [0, \infty) \mid U \in \mathcal{U}\}.$$

Definition B.3.4 (**Fréchet spaces).** A locally convex space having a complete (and translation-invariant) metric is called a *Fréchet space*.

Exercise B.3.5. Show that a locally convex space V is metrisable if and only if it has the following property: there exists a countable collection $\{p_k\}_{k=1}^{\infty}$ of continuous seminorms $p_k : V \to [0, \infty)$ such that for every $x \in V \setminus \{0\}$ there exists $k_x \in \mathbb{Z}^+$ satisfying $p_{k_x}(x) \neq 0$ (i.e. the seminorm family separates the points of V).

Exercise B.3.6. Let $k \in \mathbb{Z}^+ \cup \{0, \infty\}$ and let $U \subset \mathbb{R}^n$ be an open non-empty set. Endow space $C^k(U)$ with a Fréchet space structure.

Exercise B.3.7. Let $\Omega \subset \mathbb{C}$ be open and non-empty. Endow the space $\mathbb{H}(\Omega) \subset C(\Omega)$ of analytic functions $f : \Omega \to \mathbb{C}$ with a structure of a Fréchet space.

Definition B.3.8 (**Schwartz space).** For $f \in C^\infty(\mathbb{R}^n)$ and $\alpha, \beta \in \mathbb{N}_0^n$, let

$$p_{\alpha\beta}(f) := \sup_{x \in \mathbb{R}^n} \left| x^\beta \partial_x^\alpha f(x) \right|.$$

If $p_{\alpha\beta}(f) < \infty$ for every α, β, then f is called a *rapidly decreasing smooth function*. The collection of such functions is called the *Schwartz space* $\mathcal{S}(\mathbb{R}^n)$.

Exercise B.3.9. Show that the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space.

Definition B.3.10 (**LF-space).** A $\mathbb{C}$-vector space X is called an *LF-space* (or a *limit of Fréchet spaces*) if $X = \bigcup_{j=1}^{\infty} X_j$, where each $X_j \subset X_{j+1}$ is a subspace of X, having a Fréchet space topology τ_j such that $\tau_j = \{U \cap X_j : U \in \tau_{j+1}\}$. The topology τ of the LF-space X is then generated by the set

$$\tau := \{x + V \mid x \in X,\ V \in \mathcal{C},\ V \cap X_j \in \tau_j \text{ for every } j\},$$

where $\mathcal{C}$ is the family of those convex subsets of X that contain 0.

Exercise B.3.11. Let τ be the topology of an LF-space $X = \bigcup_{j=1}^{\infty} X_j$ as in Definition B.3.10. Prove that

$$\tau_j = \{U \cap X_j \mid U \in \tau\}.$$

Moreover, show that a linear functional $f : X \to \mathbb{C}$ is continuous if and only if the restriction $f|_{X_j} : X_j \to \mathbb{C}$ is continuous for every j.

Exercise B.3.12. Let $U \subset \mathbb{R}^n$ be an open non-empty set. Let $\mathcal{D}(U)$ consist of compactly supported C^∞-smooth functions $f : U \to \mathbb{C}$. Endow $\mathcal{D}(U)$ with an LF-space structure; this is not a Fréchet space anymore.

Definition B.3.13 (**Test functions and distributions).** The LF-space $\mathcal{D}(U)$ of Exercise B.3.12 is called the space of *test functions*, and a continuous linear functional $f : \mathcal{D}(U) \to \mathbb{C}$ is called a *distribution on* $U \subset \mathbb{R}^n$. The space of distributions on U is denoted by $\mathcal{D}'(U)$.

Exercise B.3.14 (**Locally convex Hahn–Banach Theorem).** Prove the following analogue of the Hahn–Banach Theorem B.4.25: *Let X be a locally convex space (over $\mathbb{K}$) and $f : M \to \mathbb{K}$ be a continuous linear functional on a vector subspace $M \subset X$. Then there exists a continuous extension $F : X \to \mathbb{K}$ such that $F|_M = f$.*

Exercise B.3.15. Let X be a $\mathbb{K}$-vector space, and suppose V is a vector space of linear functionals $f : X \to \mathbb{K}$ that separates the points of X. Show that V induces a locally convex topology on X, and that then the dual $X' = V$.

Definition B.3.16 (**Weak topology).** Let X be a topological vector space such that the dual $X' = \mathcal{L}(X, \mathbb{K})$ separates the points of X. The X'-induced topology is called the *weak topology of X*.

Exercise B.3.17 (**Closure of convex sets).** Let X be locally convex and $C \subset X$ convex. Show that the closure of C is the same in both the original topology and the weak topology.

Definition B.3.18 (**Weak*-topology).** Let X be a topological vector space. The *weak*-topology* of the dual X' is the topology induced by the family $\{x' \mid x \in X\}$, where $x' : X' \to \mathbb{K}$ is defined by $x'(f) := f(x)$.

Exercise B.3.19. Let $x \in X$. Show that $x' = (f \mapsto f(x)) : X' \to \mathbb{K}$ is linear. Moreover, prove that if a linear functional $f : X' \to \mathbb{K}$ is a continuous with respect to the weak*-topology, then $f = x'$ for some $x \in X$.

Exercise B.3.20 (Banach–Alaoglu theorem in topological vector spaces). Prove the following generalisation of the Banach–Alaoglu Theorem B.4.36: *Let X be a topological vector space. Let $U \subset X$ be a neighbourhood of $0 \in X$, and let*

$$K := \{f \in X' \mid \forall x \in U : \ |f(x)| \leq 1\}.$$

Then $K \subset X'$ is compact in the weak-topology.*

Definition B.3.21 (Convex hull). The *convex hull* of a subset S of a vector space X is the intersection of all convex sets that contain S. (Notice that at least X is a convex set containing S.)

Exercise B.3.22. Show that the convex hull of S is the smallest convex set that contains S.

Exercise B.3.23. Show that $x \in X$ belongs to the convex hull of S if and only if $x = \sum_{k=1}^{n} t_k x_k$ for some $n \in \mathbb{Z}^+$, where the vectors $x_k \in S$, and $t_k > 0$ are such that $\sum_{k=1}^{n} t_k = 1$.

Definition B.3.24 (Extreme set). Let K be a subset of a vector space X. A non-empty set $E \subset K$ is called an *extreme set* of K if the conditions

$$\begin{cases} x, y \in K, \\ tx + (1-t)y \in E \quad \text{for some } t \in (0,1) \end{cases}$$

imply that $x, y \in E$. A point $z \in K$ is called an *extreme point* of $K \subset X$ if $\{z\}$ is an extreme set of K (alternative characterisation: if $x, y \in K$ and $z = tx + (1-t)y$ for some $0 < t < 1$ then $x = y = z$).

Exercise B.3.25 (Krein–Milman Theorem). Prove the following *Krein–Milman Theorem*: *Let X be a locally convex space and $K \subset X$ compact. Then K is contained in the closure of the convex hull of the set of the extreme points of K.*
(Hint: The first problem is the very existence of extreme points. The family of compact extreme sets of K can be ordered by inclusion, and by the Hausdorff Maximal Principle there is a maximal chain. Notice that X' separates the points of X...)

Exercise B.3.26. Let K be a compact subset of a Fréchet space X. Show that the closure of the convex hull of K is compact.

Exercise B.3.27. Let $f : G \to X$ be continuous, where X is a Fréchet space and G is a compact Hausdorff space. Let μ be a finite positive Borel measure on G. Show that there exists a unique vector $v \in X$ such that

$$\phi(v) = \int_G \phi(f) \ \mathrm{d}\mu$$

for every $\phi \in X'$.

Definition B.3.28 (Pettis integral). Let $f : G \to X$, μ and v be as in Exercise B.3.27. Then the vector $v \in X$ is called the *Pettis integral* (or the *weak integral*) of f with respect to μ, denoted by

$$v = \int_G f \, \mathrm{d}\mu.$$

Exercise B.3.29. Let $f : G \to X$ and μ be as in Definition B.3.28. Assume that X is even a Banach space. Show that

$$\left\| \int_G f \, \mathrm{d}\mu \right\| \leq \int_G \|f\| \, \mathrm{d}\mu.$$

Definition B.3.30 (Barreled space). A subset B of a topological vector space X is called a *barrel* if it is closed, balanced, convex and $X = \bigcup_{t>0} tB$. A topological vector space is called *barreled* if its every barrel contains a neighbourhood of the origin.

Remark B.3.31. Notice that a barreled space is not necessarily locally convex.

Exercise B.3.32 (LF-spaces are barreled). Show that LF-spaces are barreled.

Definition B.3.33 (Heine–Borel property). A metric space is said to satisfy the *Heine–Borel property* if its closed and bounded sets are compact.

Definition B.3.34 (Montel space). A barreled locally convex space with the Heine–Borel property is called a *Montel space*.

Exercise B.3.35. Prove that $C^\infty(U)$ and $\mathcal{D}(U)$ are Montel spaces, where $U \subset \mathbb{R}^n$ is open and non-empty.

Exercise B.3.36. Let $\Omega \subset \mathbb{C}$ be open and non-empty. Show that the space $\mathbb{H}(\Omega)$ of analytic functions on Ω is a Montel space.

Exercise B.3.37. Prove that the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is a Montel space.

Exercise B.3.38. Let $U \subset \mathbb{R}^n$ be open and non-empty. Show that $C^k(U)$ is not a Montel space if $k \in \mathbb{N}_0$.

B.3.1 Topological tensor products

In this section we review the topological tensor products. If the reader is interested in more details on this subject we refer to [84] and to [131].

Definition B.3.39 (Projective tensor product). Let $X \otimes Y$ be the algebraic tensor product of locally convex spaces X, Y. The *projective tensor topology* or the *π-topology* of $X \otimes Y$ is the strongest topology for which the bilinear mapping $((x,y) \mapsto x \otimes y) : X \times Y \to X \otimes Y$ is continuous. This topological space is denoted by $X \otimes_\pi Y$, and its completion by $X \widehat{\otimes}_\pi Y$.

Exercise B.3.40. Let X, Y be locally convex spaces over $\mathbb{C}$. Show that the dual of $X \otimes_\pi Y$ (and also of $X \widehat{\otimes}_\pi Y$) is isomorphic to the space of continuous bilinear mappings $X \times Y \to \mathbb{C}$.

Exercise B.3.41. Let X, Y be locally convex metrisable spaces. Show that $X \widehat{\otimes}_\pi Y$ is a Fréchet space. Moreover, if X, Y are barreled, show that $X \otimes Y$ is barreled.

Exercise B.3.42. Let X, Y be locally convex metrisable barreled spaces. Show that $X \otimes Y$ is barreled.

Exercise B.3.43 (Projective Banach tensor product). Let X, Y be Banach spaces. For $f \in X \otimes Y$, define

$$\|f\|_\pi := \inf \left\{ \sum_j \|x_j\| \ \|y_j\| : \ f = \sum_j x_j \otimes y_j \right\}.$$

Show that $f \mapsto \|f\|_\pi$ is a norm on $X \otimes Y$, and that the corresponding norm topology is the projective tensor topology.

Exercise B.3.44. Let X, Y be locally convex spaces over $\mathbb{C}$. Show that the algebraic tensor product $X \otimes Y$ can be identified with the space $B(X', Y')$ of continuous bilinear functionals $X' \times Y' \to \mathbb{C}$, where X' and Y' are the dual spaces with the weak topologies.

Definition B.3.45 (Injective tensor product). Let X, Y be locally convex spaces over $\mathbb{C}$. Let $\widetilde{B}(X', Y')$ be the space of those bilinear functionals $X' \times Y' \to \mathbb{C}$ that are continuous separately in each variable. Endow $\widetilde{B}(X', Y')$ with the topology τ of uniform convergence on the products of an equicontinuous subset of X' and an equicontinuous subset of Y'. Interpreting $X \otimes Y \subset \widetilde{B}(X', Y')$ as in Exercise B.3.44, let the *injective tensor topology* be the restriction of τ to $X \otimes Y$. This topological space is denoted by $X \otimes_\varepsilon Y$, and its completion by $X \widehat{\otimes}_\varepsilon Y$.

Exercise B.3.46. Let X, Y be locally convex spaces over $\mathbb{C}$. Show that the bilinear mapping $((x, y) \mapsto x \otimes y) : X \times Y \to X \otimes_\varepsilon Y$ is continuous. From this, deduce that the injective topology of $X \otimes Y$ is coarser than the projective topology (i.e. is a subset of the projective topology).

Exercise B.3.47. Studying the mapping $((x, y) \mapsto x \otimes y) : X \times Y \to X \otimes Y$, explain how the inclusion $X \widehat{\otimes}_\pi Y \subset X \widehat{\otimes}_\varepsilon Y$ should be understood.

Exercise B.3.48 (Injective Banach tensor product). Let X, Y be Banach spaces. For $f \in X \otimes Y$, define

$$\|f\|_\varepsilon := \sup \left\{ |x' \otimes y'(f)| : \ x' \in X', \ y' \in Y', \ \|x'\| = 1 = \|y'\| \right\}.$$

Show that $f \mapsto \|f\|_\varepsilon$ is a norm on $X \otimes Y$, and that the corresponding norm topology is the injective tensor topology.

Definition B.3.49 (Nuclear space). A locally convex space X is called *nuclear* if $X\widehat{\otimes}_\pi Y = X\widehat{\otimes}_\varepsilon Y$ for every locally convex space Y (where the equality of sets is understood in the sense of Exercise B.3.47). In such a case, these completed tensor products are written simply $X\widehat{\otimes}Y$.

Exercise B.3.50. Let X, Y be nuclear spaces, and let $M, N \subset X$ be vector subspaces such that N is closed. Show that M, X/N, $X \times Y$ and $X\widehat{\otimes}Y$ are nuclear spaces.

Exercise B.3.51. Show that $C^\infty(U)$, $\mathcal{D}(U)$, $\mathcal{S}(\mathbb{R}^n)$, $\mathbb{H}(\Omega)$ and their dual spaces (of distributions) are nuclear.

Exercise B.3.52. Let X, Y be Fréchet spaces and X nuclear. Show that $\mathcal{L}(X', Y) \cong X\widehat{\otimes}Y$, $\mathcal{L}(X, Y) \cong X'\widehat{\otimes}Y$, and that $X'\widehat{\otimes}Y' \cong (X\widehat{\otimes}Y)'$.

Exercise B.3.53. Prove the following *Schwartz Kernel Theorem* B.3.55:

Remark B.3.54. In the following Schwartz Kernel Theorem B.3.55, we denote $\langle\psi, A\phi\rangle := (A\phi)(\psi)$, and $\langle\psi \otimes \phi, K_A\rangle := K_A(\psi \otimes \phi)$.

Theorem B.3.55 (Schwartz Kernel Theorem). *Let $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ be open and non-empty, and let $A : \mathcal{D}(U) \to \mathcal{D}'(V)$ be linear and continuous. Then there exists a unique distribution $K_A \in \mathcal{D}'(V)\widehat{\otimes}\mathcal{D}'(U) \cong \mathcal{D}'(V \times U)$ such that*

$$\langle\psi, A\phi\rangle = \langle\psi \otimes \phi, K_A\rangle$$

for every $\phi \in \mathcal{D}(U)$ and $\psi \in \mathcal{D}(V)$. Moreover, if $A : \mathcal{D}(U) \to C^\infty(V)$ is continuous then it can be interpreted that $K_A \in C^\infty(V, \mathcal{D}'(U))$.

Definition B.3.56 (Schwartz kernel). The distribution K_A in Theorem B.3.55 is called the *Schwartz kernel* of A, written informally as

$$A\phi(x) = \int_V K_A(x, y)\ \phi(y)\ \mathrm{d}y.$$

Exercise B.3.57. Let $A : \mathcal{D}(U) \to \mathcal{D}'(V)$ be continuous and linear as in Theorem B.3.55. Give necessary and sufficient conditions for A such that $K_A \in C^\infty(V \times U)$.

Exercise B.3.58. Find variants of the Schwartz Kernel Theorem B.3.55 for Schwartz functions and for tempered distributions.

B.4 Banach spaces

Definition B.4.1 (Seminorm and norm; normed spaces). Let X be a $\mathbb{K}$-vector space. A mapping $p : X \to \mathbb{R}$ is a *seminorm* if

$$\begin{cases} p(x + y) \le p(x) + p(y), \\ p(\lambda x) = |\lambda|\ p(x) \end{cases}$$

for every $x, y \in X$ and $\lambda \in \mathbb{R}$. If $p : X \to \mathbb{R}$ is a seminorm for which $p(x) = 0$ implies $x = 0$, then it is called a *norm*. Typically, a norm on X is written as $x \mapsto \|x\|_X$ or simply $\|x\|$. A vector space with a norm is called a *normed space*.

Example. On the vector space $\mathbb{K}$, the absolute value mapping $x \mapsto |x|$ is a norm.

Exercise B.4.2. Let $p : X \to [0, \infty)$ be a seminorm on a $\mathbb{K}$-vector space X and

$$x \sim y \quad \overset{\text{definition}}{\Longleftrightarrow} \quad p(x - y) = 0.$$

Prove the following claims:
(a) $\sim$ is an equivalence relation on X.
(b) The set $L := \{[x] : \ x \in X\}$, with $[x] := \{y \in X : \ x \sim y\}$, is an $\mathbb{R}$-vector space when endowed with operations

$$[x] + [y] := [x + y], \quad \lambda[x] := [\lambda x]$$

and the norm $[x] \mapsto p(x)$.

Exercise B.4.3. Let $w_j \geq 0$ for every $j \in J$. Define

$$\sum_{j \in J} w_j := \sup \left\{ \sum_{k \in K} w_k : \ K \subset J \text{ finite} \right\}.$$

Show that $\{j \in J : \ w_j > 0\}$ is at most countable if $\sum_{j \in J} w_j < \infty$.

Exercise B.4.4. For $x \in \mathbb{K}^J$ define

$$\|x\|_{\ell^p} := \begin{cases} \left(\sum_{j \in J} |x_j|^p \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \sup_{j \in J} |x_j|, & \text{if } p = \infty, \end{cases}$$

where $x_j := x(j)$. Show that $\ell^p(J) := \left\{ x \in \mathbb{K}^J : \ \|x\|_{\ell^p} < \infty \right\}$ is a Banach space with respect to the norm $x \mapsto \|x\|_{\ell^p}$.

Exercise B.4.5. Norms p_1, p_2 on a vector space V are called (Lipschitz) *equivalent* if $a^{-1} p_1(x) \leq p_2(x) \leq a p_1(x)$ for every $x \in V$, where $a \geq 1$ is a constant. Show that any two norms on a finite-dimensional space V are equivalent. Consequently, a finite-dimensional normed space is a Banach space.

Exercise B.4.6. Let K be a compact space. Show that

$$C(K) := \{f : K \to \mathbb{K} \mid f \text{ continuous } \}$$

is a Banach space when endowed with the norm

$$f \mapsto \|f\|_{C(K)} := \sup_{x \in K} |f(x)|.$$

Remark B.4.7. The previous exercise deals with special cases of of $L^p(\mu)$, the *Lebesgue p-spaces*. These Banach spaces are introduced in Definition C.4.6.

Definition B.4.8 (Normed and Banach spaces). Notice that the *norm metric* $((x,y) \mapsto \|x-y\|) : X \times X \to \mathbb{R}$ is a metric on X. Let τ_X denote the corresponding metric topology, called the *norm topology*, where the *open ball centered at $x \in X$ with radius $r > 0$* is

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

Ball $\mathbb{B}_X(0,1)$ is called the *open unit ball*. The *closed ball centered at $x \in X$ with radius $r > 0$* is

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

Notice that here $\overline{\mathbb{B}(x,r)} = \overline{\mathbb{B}}(x,r)$, where $\overline{S}$ refers to the norm closure of a set $S \subset X$. A *Banach space* is a normed space where the norm metric is complete.

Exercise B.4.9. Show that $V := \{x \in \ell^p(J) : \ \{j \in J : \ x_j \neq 0\} \text{ finite}\}$ is a dense normed vector subspace of $\ell^p(J)$.

Definition B.4.10 (Bounded linear operators). A linear mapping $A : X \to Y$ between normed spaces X, Y is called *bounded*, denoted $A \in \mathcal{L}(X,Y)$, if

$$\|Ax\| \le C \ \|x\|$$

for every $x \in X$, where $C < \infty$ is a constant. The *norm of* $A \in \mathcal{L}(X,Y)$ is

$$\|A\| := \sup_{x \in X: \ \|x\| \le 1} \|Ax\|.$$

This norm is also called the *operator norm* and is often denoted by $\|A\|_{op}$. We often abbreviate $\mathcal{L}(X) := \mathcal{L}(X,X)$.

Exercise B.4.11. Let $A : X \to Y$ be a linear operator between normed spaces X and Y. Show that A is bounded if and only if it is continuous.

Exercise B.4.12. Show that $\mathcal{L}(X,Y)$ is really a normed space.

Exercise B.4.13. Show that $\|AB\| \le \|A\|\|B\|$ if $B \in \mathcal{L}(X,Y)$ and $A \in \mathcal{L}(Y,Z)$.

Exercise B.4.14. Show that $\mathcal{L}(X,Y)$ is a Banach space if Y is a Banach space.

Definition B.4.15 (Duals). Let V be a Banach space over $\mathbb{K}$. The dual of V is the space

$$V' = \mathcal{L}(V,\mathbb{K}) := \{A : V \to \mathbb{K} \mid A \text{ bounded and linear}\}.$$

endowed with the (operator) norm

$$A \mapsto \|A\| := \sup_{f \in V: \ \|f\|_V \le 1} |A(f)|.$$

Exercise B.4.16. Prove that V' is a Banach space.

Definition B.4.17 (Compact linear operator). Let X, Y be normed spaces, and let $\overline{\mathbb{B}} = \overline{\mathbb{B}(0,1)} = \{x \in X : \|x\| \leq 1\}$. A linear mapping $A : X \to Y$ is called *compact*, denoted $A \in \mathcal{LC}(X,Y)$, if the closure of $A(\overline{\mathbb{B}}) \subset Y$ is compact. We also denote $\mathcal{LC}(X) := \mathcal{LC}(X,X)$.

Exercise B.4.18. Show that $\mathcal{LC}(X,Y)$ is a linear subspace of $\mathcal{L}(X,Y)$, and it is closed if Y is complete.

Exercise B.4.19. Let $B_0, C_0 \in \mathcal{L}(X,Y)$ and $B_1, C_1 \in \mathcal{L}(Y,Z)$ such that C_0, C_1 are compact. Show that $C_1 B_0, B_1 C_0$ are compact.

Lemma B.4.20. (Almost Orthogonality Lemma [F. Riesz]) *Let X be a normed space with closed subspace $Y \neq X$. For each $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that $\|x_\varepsilon\| = 1$ and $\operatorname{dist}(x_\varepsilon, Y) \geq 1 - \varepsilon$.*

Proof. Let $z \in X \setminus Y$ and $r := \operatorname{dist}(z, Y) > 0$. Take $y = y_\varepsilon \in Y$ such that $r \leq \|z - y\| < (1-\varepsilon)^{-1} r$. Let $x_\varepsilon := (z-y)/\|z-y\|$. If $u \in Y$ then

$$\begin{aligned} \|x_\varepsilon - u\| &= \left\| \frac{z-y}{\|z-y\|} - u \right\| \\ &= \frac{\|z - (y + \|z-y\| u)\|}{\|z-y\|} \\ &> \frac{r}{(1-\varepsilon)^{-1} r} \\ &= 1 - \varepsilon, \end{aligned}$$

showing that $\operatorname{dist}(x_\varepsilon, Y) \geq 1 - \varepsilon$. □

Theorem B.4.21 (Riesz's Compactness Theorem). *Let X be a normed space. Then X is finite-dimensional if and only if $\overline{\mathbb{B}(0,1)}$ is compact.*

Proof. A set in a finite-dimensional normed space is compact if and only if it is bounded, by the Heine–Borel Theorem.

Now let X be infinite-dimensional. Let $0 < \varepsilon < 1$ and $x_1 \in X$ such that $\|x_1\| = 1$. Inductively, let $Y_k := \operatorname{span}\{x_j\}_{j=1}^k \neq X$, and choose $x_{k+1} \in X \setminus Y_k \neq \emptyset$ such that $\|x_{k+1}\| = 1$ and $\operatorname{dist}(x_{k+1}, Y_k) > 1 - \varepsilon$. Then it is clear that the sequence $(x_k)_{k=1}^\infty$ does not have a converging subsequence. Hence $\overline{\mathbb{B}(0,1)}$ is not compact by Theorem A.13.4. □

Remark B.4.22 (**Is identity compact?**). Riesz's Compactness Theorem B.4.21 could also be stated: *a normed space X is finite-dimensional if and only if the identity mapping $I = (x \mapsto x) : X \to X$ is compact.* This together with the results of Exercises B.4.18 and B.4.19 proves that $\mathcal{LC}(X)$ is a closed two-sided proper ideal of $\mathcal{L}(X)$, where X is a Banach space.

Theorem B.4.23 (Baire's Theorem). *Let (X, d) be a complete metric space and $U_j \subset X$ be dense and open for each $k \in \mathbb{Z}^+$. Then $G = \bigcap_{k=1}^\infty U_k$ is dense.*

Proof. We must show that $G \cap \mathbb{B}(x_0, r_0) \neq \emptyset$ for any $x_0 \in X$ and $r_0 > 0$. Assuming $X \neq \emptyset$, take x_1 and r_1 such that

$$\overline{\mathbb{B}(x_1, r_1)} \subset U_1 \cap \mathbb{B}(x_0, r_0).$$

Inductively, we choose x_{k+1} and $r_{k+1} < 1/k$ so that

$$\overline{\mathbb{B}(x_{k+1}, r_{k+1})} \subset U_{k+1} \cap \mathbb{B}(x_k, r_k).$$

Then $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence, thus converging to some $x \in X$ by completeness. By construction, $x \in G \cap \mathbb{B}(x_0, r_0)$. □

Exercise B.4.24 (**Baire's theorem and interior points**). Clearly, Baire's Theorem B.4.23 is equivalent to the following: in a complete metric space, a countable union of sets without interior points is without interior points. Use this to prove that an algebraic basis of an infinite-dimensional Banach space must be uncountable.

Theorem B.4.25 (**Hahn–Banach Theorem**). *Let X be a real normed space and $f : M_f \to \mathbb{R}$ be bounded and linear on a vector subspace $M_f \subset X$. Then there exists extension $F : X \to \mathbb{R}$ such that $F|_{M_f} = f$ and $\|F\| = \|f\|$.*

Proof. Let

$$\begin{aligned} S \;:=\; \{h : M_h \to \mathbb{R} \quad | \quad & h \text{ linear on vector subspace } M_h \subset X, \\ & M_f \subset M_h,\ \|h\| = \|f\|\}. \end{aligned}$$

Then $f \in S \neq \emptyset$. Endow S with the partial order

$$g \leq h \iff \begin{cases} M_g \subset M_h, \\ g = h|_{M_g}. \end{cases}$$

Take a chain $(f_j)_{j \in J} \subset S$. Then $f_j \leq h$ for each $j \in J$, where $h \in S$ is defined so that $M_h = \bigcup_{j \in J} M_{f_j}$, $h|_{M_{f_j}} = f_j$. Thereby, in view of Zorn's lemma (Theorem A.4.10), there is a maximal element $F : M_F \to \mathbb{R}$ in S. Suppose $M_F \neq X$. Then take $x_0 \in X \setminus M_F$. Given $a \in \mathbb{R}$, define

$$G_a : M_F + \mathbb{R}x_0 \to \mathbb{R}, \quad G_a(u + tx_0) = F(u) - ta.$$

Then G_a is bounded, linear, and $G_a|_{M_F} = F$. Hence $\|G_a\| \geq \|F\| = \|f\|$. Could it be that $\|G_a\| = \|F\|$ (this would contradict the maximality of F)? For any $u, v \in M_F$,

$$\begin{aligned} |F(u) - F(v)| &= |F(u - v)| \\ &\leq \|F\| \, \|u - v\| \\ &\leq \|F\| \left(\|u + x_0\| + \|v + x_0\| \right). \end{aligned}$$

Hence there exists $a_0 \in \mathbb{R}$ such that

$$F(u) - \|F\| \; \|u + x_0\| \leq a_0 \leq F(v) + \|F\| \; \|v + x_0\|$$

for every $u, v \in M_F$. Thus

$$|F(w) - a_0| \leq \|F\| \; \|w + x_0\|$$

for every $w \in M_F$. From this (assuming that non-trivially $t \neq 0$), we get

$$\begin{aligned} |G_{a_0}(u + tx_0)| &= |t| \; |u/t - a_0| \\ &\leq |t| \; \|F\| \; \|u/t + x_0\| \\ &= \|F\| \; \|u + tx_0\|; \end{aligned}$$

but this means $\|G_{a_0}\| \leq \|F\|$, a contradiction. □

Exercise B.4.26 (Complex version of the Hahn–Banach Theorem). Prove the complex version of the Hahn–Banach Theorem: Let X be a complex normed space and $f : M_f \to \mathbb{C}$ be bounded and linear on a vector subspace $M_f \subset X$. Then there exists an extension $F : X \to \mathbb{C}$ such that $F|_{M_f} = f$ and $\|F\| = \|f\|$.

Corollary B.4.27. *Let X be a normed space and $x \in X$. Then*

$$\|x\| = \max \{|F(x)| : \; F \in L(X, \mathbb{K}), \; \|F\| \leq 1\} .$$

Corollary B.4.28 (Hahn–Banach $\Longrightarrow$ Riesz' Compactness Theorem). *Let X be a normed space. Then $\overline{\mathbb{B}}(0,1)$ is compact if and only if it is finite-dimensional.*

Proof. A set in a finite-dimensional normed space is compact if and only if it is bounded, by the Heine–Borel Theorem.

The proof for the converse follows [28]: Suppose X is locally compact and let $S^1 := \{x \in X : \; \|x\| = 1\}$. Then $\left\{S^1 \cap \mathrm{Ker}(f) : \; f \in \mathcal{L}(X, \mathbb{K})\right\}$ is a family of compact sets, whose intersection is empty by the Hahn–Banach Theorem. Thereby there exists $\{f_k\}_{k=1}^n \subset \mathcal{L}(X, \mathbb{K})$ such that

$$\bigcap_{k=1}^{n} S^1 \cap \mathrm{Ker}(f_k) = \emptyset, \quad \text{i.e.} \quad \bigcap_{k=1}^{n} \mathrm{Ker}(f_k) = \{0\}.$$

Since the co-dimension of $\mathrm{Ker}(f_k) \leq 1$, this implies that $\dim(X) \leq n$. □

Theorem B.4.29 (Banach–Steinhaus Theorem, or Uniform Boundedness Principle). *Let X be a Banach space, let Y be a normed space, and let $\{A_j\}_{j \in J} \subset \mathcal{L}(X, Y)$ be such that*

$$\sup_{j \in J} \|A_j x\| < \infty$$

for every $x \in X$. Then $\sup_{j \in J} \|A_j\| < \infty$.

Proof. Let $p_j(x) := \|A_j x\|$ and $p(x) := \sup\{p_j(x) \mid j \in J\}$. Clearly, $p, p_j : X \to \mathbb{R}$ are seminorms. Moreover, p_j is continuous for every $j \in J$, and we must show that also p is continuous. Since p_j is continuous for every $j \in J$, set

$$U_k := \{x \in X : \ p(x) > k\} = \bigcup_{j \in J} \{x \in X : \ p_j(x) > k\}$$

is open. Now $\bigcap_{k=1}^{\infty} U_k = \emptyset$, so that by Baire's Theorem B.4.23 there exists $k_0 \in \mathbb{Z}^+$ for which $\overline{U_{k_0}} \neq X$; actually, here $\overline{U_1} \neq X$, because $U_1 = k^{-1} U_k$. Choose $x_0 \in X$ and $r_0 > 0$ such that

$$\mathbb{B}(x_0, r_0) \subset X \setminus \overline{U_1}.$$

If $z \in \mathbb{B}(0,1)$ then

$$\begin{aligned} r_0\, p(z) &= p(r_0 z) \\ &\leq p(x_0 + r_0 z) + p(-x_0) \\ &\leq 2. \end{aligned}$$

Thus $\|A_j\| \leq 2/r_0$ for every $j \in J$. □

Definition B.4.30 (Open mappings). A mapping $f : X \to Y$ between topological spaces X, Y is said to be *open*, if $f(U) \subset Y$ is open for every open $U \subset X$.

Theorem B.4.31 (Open Mapping Theorem). *Let $A \in \mathcal{L}(X,Y)$ be surjective, where X, Y are Banach spaces. Then A is open.*

Proof. It is sufficient to show that $\mathbb{B}_Y(0,r) \subset \overline{A(\mathbb{B}_X(0,1))}$ for some $r > 0$. For each $k \in \mathbb{Z}^+$, set $U_k := Y \setminus \overline{A(\mathbb{B}_X(0,k))}$ is open. Now $\bigcap_{k=1}^{\infty} U_k = \emptyset$, because A is surjective. By Baire's Theorem B.4.23, $\overline{U_{k_0}} \neq Y$ for some $k_0 \in \mathbb{Z}^+$; actually, $\overline{U_1} \neq Y$, because $A(\mathbb{B}_X(0,1)) = k^{-1} A(\mathbb{B}_X(0,k))$. Take $y_0 \in Y$ and $r_0 > 0$ such that

$$\mathbb{B}_Y(y_0, r_0) \subset Y \setminus \overline{U_1}.$$

Now

$$\mathbb{B}_Y(y_0, r_0) \subset Y \setminus U_1 = \overline{A(\mathbb{B}_X(0,1))}.$$

Let $\varepsilon > 0$ and $y \in \mathbb{B}_Y(0, r_0)$. Take $w_1, w_2 \in \mathbb{B}_X(0,1)$ such that

$$\begin{aligned} \|y_0 - A w_1\| &< \varepsilon/2, \\ \|(y_0 + y) - A w_2\| &< \varepsilon/2. \end{aligned}$$

Then $w_1 - w_2 \in \mathbb{B}_X(0,2)$ and $\|y - A(w_1 - w_2)\| < \varepsilon$. By linearity, this yields

$$\forall \varepsilon > 0\ \forall y \in \mathbb{B}_Y(0, r_0)\ \exists x \in \mathbb{B}_X(0, 2\|y\|/r_0) : \ \|y - Ax\| < \varepsilon.$$

Thus if $z \in \mathbb{B}_Y(0, r_0)$, take $x_0 \in \mathbb{B}_X(0,2)$ such that $\|z - A x_0\| < r_0/2$. Inductively, choose $x_k \in \mathbb{B}_X(0, 2^{1-k})$ such that $\|z - A \sum_{j=0}^{k} x_j\| < 2^{1-k} r_0$. Now $\sum_{j=0}^{k} x_j \to_k$

$x \in \overline{\mathbb{B}_X(0,4)} \subset \mathbb{B}_X(0,5)$, because X is complete. We have $z = Ax$ by continuity of A. Thereby

$$\mathbb{B}_Y(0,r_0) \subset A(\mathbb{B}_X(0,5)),$$

implying $\mathbb{B}_Y(0,r_0/5) \subset A(\mathbb{B}_X(0,1))$. □

Corollary B.4.32 (**Bounded Inverse Theorem).** *Let $B \in \mathcal{L}(X,Y)$ be bijective between Banach spaces X,Y. Then B^{-1} is continuous.*

Definition B.4.33 (**Graph).** The *graph* of a mapping $f : X \to Y$ is

$$\begin{aligned}\Gamma(f) &:= \{(x, f(x)) \mid x \in X\} \\ &\subset X \times Y.\end{aligned}$$

Theorem B.4.34 (**Closed Graph Theorem).** *Let $A : X \to Y$ be a linear mapping between Banach spaces X,Y. Then A is continuous if and only if its graph is closed in $X \times Y$.*

Proof. Suppose A is continuous. Take a Cauchy sequence $((x_j, Ax_j))_{j=1}^{\infty}$ of $\Gamma(A) \subset X \times Y$. Then $(x_j)_{j=1}^{\infty}$ is a Cauchy sequence of X, thereby converging to some $x \in X$ by completeness. Then $Ax_j \to Ax$ by the continuity of A. Hence $(x_j, Ax_j) \to (x, Ax) \in \Gamma(A)$; the graph is closed.

Now assume that $\Gamma(A) \subset X \times Y$ is closed. Thus the graph is a Banach subspace of $X \times Y$. Define a mapping $B := (x \mapsto (x, Ax)) : X \to \Gamma(A)$. It is easy to see that B is a linear bijection. By the Open Mapping Theorem, B is continuous. This implies the continuity of A. □

Definition B.4.35 (**Weak*-topology).** Let $x \mapsto \|x\|$ be the norm of a normed vector space X over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The dual space $X' = \mathcal{L}(X, \mathbb{K})$ of X is set of bounded linear functionals $f : X \to \mathbb{K}$, having a norm

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

This endows X' with a Banach space structure. However, it is often better to use a weaker topology for the dual: let us define $x(f) := f(x)$ for every $x \in X$ and $f \in X'$; this gives the interpretation $X \subset X'' := \mathcal{L}(X', \mathbb{K})$, because

$$|x(f)| = |f(x)| \le \|f\| \|x\|.$$

So we may treat X as a set of functions $X' \to \mathbb{K}$, and we define the *weak*-topology* of X' to be the X-induced[2] topology of X'.

Theorem B.4.36 (**Banach–Alaoglu Theorem).** *Let X be a Banach space. Then the closed unit ball*

$$K := \overline{B_{X'}(0,1)} = \{\phi \in X' :\ \|\phi\|_{X'} \le 1\}$$

of X' is weak*-compact.

[2]see Definition A.18.1

Proof. Due to Tihonov's Theorem A.18.8,

$$P := \prod_{x\in X} \{\lambda \in \mathbb{C} : \ |\lambda| \leq \|x\|\} = \overline{\mathbb{D}(0,\|x\|)}^X$$

is compact in the product topology τ_P. Any element $f \in P$ is a mapping

$$f : X \to \mathbb{C} \quad \text{such that} \quad f(x) \leq \|x\|.$$

Hence $K = X' \cap P$. Let τ_1 and τ_2 be the relative topologies of K inherited from the weak*-topology $\tau_{X'}$ of X' and the product topology τ_P of P, respectively. We shall prove that $\tau_1 = \tau_2$ and that $K \subset P$ is closed; this would show that K is a compact Hausdorff space.

First, let $\phi \in X'$, $f \in P$, $S \subset X$, and $\delta > 0$. Define

$$\begin{aligned} U(\phi,S,\delta) &:= \{\psi \in X' : \ x \in S \Rightarrow |\psi x - \phi x| < \delta\}, \\ V(f,S,\delta) &:= \{g \in P : \ x \in S \Rightarrow |g(x) - f(x)| < \delta\}. \end{aligned}$$

Then

$$\begin{aligned} \mathcal{U} &:= \{U(\phi,S,\delta) \mid \phi \in X', \ S \subset X \text{ finite}, \ \delta > 0\}, \\ \mathcal{V} &:= \{V(f,S,\delta) \mid f \in P, \ S \subset X \text{ finite}, \ \delta > 0\} \end{aligned}$$

are bases for the topologies $\tau_{X'}$ and τ_P, respectively. Clearly

$$K \cap U(\phi,S,\delta) = K \cap V(\phi,S,\delta),$$

so that the topologies $\tau_{X'}$ and τ_P agree on K, i.e. $\tau_1 = \tau_2$.

Still we have to show that $K \subset P$ is closed. Let $f \in \overline{K} \subset P$. First we show that f is linear. Take $x, y \in X$, $\lambda, \mu \in \mathbb{C}$ and $\delta > 0$. Choose $\phi_\delta \in K$ such that

$$f \in V(\phi_\delta, \{x, y, \lambda x + \mu y\}, \delta).$$

Then

$$\begin{aligned} & |f(\lambda x + \mu y) - (\lambda f(x) + \mu f(y))| \\ \leq\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\phi_\delta(\lambda x + \mu y) - (\lambda f(x) + \mu f(y))| \\ =\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\lambda(\phi_\delta x - f(x)) + \mu(\phi_\delta y - f(y))| \\ \leq\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\lambda| \ |\phi_\delta x - f(x)| + |\mu| \ |\phi_\delta y - f(y)| \\ \leq\ & \delta \ (1 + |\lambda| + |\mu|). \end{aligned}$$

This holds for every $\delta > 0$, so that actually

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y),$$

f is linear! Moreover, $\|f\| \leq 1$, because

$$|f(x)| \leq |f(x) - \phi_\delta x| + |\phi_\delta x| \leq \delta + \|x\|.$$

Hence $f \in K$, K is closed. □

Remark B.4.37. The Banach–Alaoglu Theorem B.4.36 implies that a bounded weak*-closed subset of the dual space is a compact Hausdorff space in the relative weak*-topology. However, in a normed space norm-closed balls are compact if and only if the dimension is finite!

B.4.1 Banach space adjoint

We now come back to the adjoints of Banach spaces and of operators introduced in Definition B.4.15. Here we give a condensed treatment to acquaint the reader with the topic.

Definition B.4.38 (**Duality).** Let X be a Banach space and $X' = \mathcal{L}(X, \mathbb{K})$ its dual. For $x \in X$ and $x' \in X'$ let us denote

$$\langle x, x' \rangle := x'(x).$$

We endow X' with the norm $x' \mapsto \|x'\|$ given by

$$\|x'\| := \sup \{ |\langle x, x' \rangle| : \ x \in X, \ \|x\| \leq 1 \}.$$

Exercise B.4.39. Let X be a Banach space and $x \in X$. Show that

$$\|x\| = \sup \{ |\langle x, x' \rangle| : \ x' \in X', \ \|x'\| \leq 1 \}.$$

Exercise B.4.40. Let X, Y be Banach spaces with respective duals X', Y'. Let $A \in \mathcal{L}(X, Y)$. Show that there exists a unique $A' \in \mathcal{L}(Y', X')$ such that

$$\langle Ax, y' \rangle = \langle x, A'(y') \rangle \tag{B.1}$$

for every $x \in X$ and $y' \in Y'$. Prove also that

$$\|A'\| = \|A\|.$$

Definition B.4.41 (**Adjoint operator).** Let $A \in \mathcal{L}(X, Y)$ be as in Exercise B.4.40. Then $A' \in \mathcal{L}(Y', X')$ defined by (B.1) is called the (*Banach*) *adjoint* of A.

Exercise B.4.42. Show that $A \in \mathcal{L}(X, Y)$ is compact if and only if $A' \in \mathcal{L}(Y', X')$ is compact.

Definition B.4.43 (**Complemented subspace).** A closed subspace V of a topological vector space X is said to be *complemented* in X by a subspace $W \subset X$ if

$$\begin{cases} V + W = X & \text{and} \\ V \cap W = \{0\}. \end{cases}$$

Then we write $X = V \oplus W$, saying that X is the *direct sum* of V and W.

Exercise B.4.44. Show that a closed subspace V is complemented in X if X/V is finite-dimensional.

Exercise B.4.45. Show that a finite-dimensional subspace of a locally convex space is complemented. (Hint: Hahn–Banach.)

Exercise B.4.46. Let $A \in \mathcal{L}(X)$ be compact, where X is a Banach space. Let λ be a non-zero scalar. Show that the range set

$$(\lambda I - A)(X) = \{\lambda x - Ax : \ x \in X\}$$

is closed, $\mathrm{Ker}(\lambda I - A) = \{x \in X : Ax = \lambda x\}$ is finite-dimensional, and that

$$\begin{aligned} & \dim\left(\mathrm{Ker}(\lambda I - A)\right) \\ = \ & \dim\left(\mathrm{Ker}(\lambda I - A')\right) \\ = \ & \dim\left(X/((\lambda I - A)(X))\right) \\ = \ & \dim\left(X'/((\lambda I - A')(X'))\right). \end{aligned}$$

Definition B.4.47 (Reflexive space). Let X be a Banach space and $X' = \mathcal{L}(X, \mathbb{K})$ its dual Banach space. The *second dual* of X is $X'' := (X')' = \mathcal{L}(X', \mathbb{K})$. It is then easy to show that we can define a linear isometry $(x \mapsto x'') : X \to X''$ onto a closed subspace of X'' by

$$x''(f) := f(x).$$

Thus X can be regarded as a subspace of X''. If $X'' = \{x'' : \ x \in X\}$ then X is called *reflexive*.

Exercise B.4.48. Show that $(x \mapsto x'') : X \to X''$ in Definition B.4.47 has the claimed properties.

Exercise B.4.49. Let $1 < p < \infty$. Show that $\ell^p = \ell^p(\mathbb{Z}^+)$ is reflexive. What about ℓ^1 and ℓ^∞?

Exercise B.4.50. Show that $C([0,1])$ is not reflexive.

Exercise B.4.51. Let X be a Banach space. Prove that X is reflexive if and only if its closed unit ball is compact in the weak topology. (Hint: Hahn–Banach and Banach–Alaoglu).

Exercise B.4.52. Let V be a closed subspace of a reflexive Banach space X. Show that V and X/V are reflexive.

Exercise B.4.53. Show that X is reflexive if and only if X' is reflexive.

B.5 Hilbert spaces

Definition B.5.1 (Inner product and Hilbert spaces). Let $\mathcal{H}$ be a $\mathbb{C}$-vector space. A mapping $((x,y) \mapsto \langle x,y\rangle) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is an *inner product* if

$$\begin{aligned}
&\langle x+y, z\rangle = \langle x,z\rangle + \langle y,z\rangle,\\
&\langle \lambda x, y\rangle = \lambda\langle x,y\rangle,\\
&\langle y,x\rangle = \overline{\langle x,y\rangle},\\
&\langle x,x\rangle \geq 0,\\
&\langle x,x\rangle = 0 \quad \Rightarrow \quad x = 0
\end{aligned}$$

for every $x, y \in \mathcal{H}$ and $\lambda, \mu \in \mathbb{C}$. Then $\mathcal{H}$ endowed with the inner product is called an *inner product space*. An inner product defines the canonical norm

$$\|x\| := \langle x,x\rangle^{1/2};$$

we shall soon prove that this is a norm in the usual sense. $\mathcal{H}$ is called a *Hilbert space* (or a *complete inner product space*) if it is a Banach space with respect to the canonical norm.

Exercise B.5.2. Show that $\ell^2(J)$ is a Hilbert space, where

$$\langle x,y\rangle = \sum_{j\in J} x_j \overline{y_j}.$$

Definition B.5.3 (Orthogonality). Vectors $x, y \in \mathcal{H}$ are said to be *orthogonal* in an inner product space $\mathcal{H}$, denoted $x \perp y$, if $\langle x,y\rangle = 0$. For $S \subset \mathcal{H}$, let

$$S^\perp := \{x \in \mathcal{H} \mid \forall y \in S : \; x \perp y\}.$$

Subspaces $M, N \subset \mathcal{H}$ are called *orthogonal*, denoted $M \perp N$, if $\langle x,y\rangle = 0$ for every $x \in M$ and $y \in N$. A collection $\{x_\alpha\}_{\alpha\in I}$ is called *orthonormal* if $\|x_\alpha\| = 1$ for all $\alpha \in I$ and if $\langle x_\alpha, x_\beta\rangle = 0$ for all $\alpha \neq \beta$, $\alpha, \beta \in I$.

Exercise B.5.4. Show that $S^\perp \subset \mathcal{H}$ is a closed vector subspace, and that $S \subset (S^\perp)^\perp$. Show that if V is a closed vector subspace of $\mathcal{H}$ then $V = (V^\perp)^\perp$.

Exercise B.5.5 (Pythagoras' theorem). Let $x_1, x_2, \ldots, x_n \in \mathcal{H}$ be mutually orthogonal, i.e. assume that $x_i \perp x_j$ for all $i \neq j$. Prove that $\|\sum_{j=1}^n x_j\|^2 = \sum_{j=1}^n \|x_j\|^2$. (This generalised the famous theorem of Pythagoras of Samos on the triangles in the plane.)

Proposition B.5.6 (Cauchy–Schwarz inequality). *Let $\mathcal{H}$ be an inner product space. Then*

$$|\langle x,y\rangle| \leq \|x\| \; \|y\| \tag{B.2}$$

for every $x, y \in \mathcal{H}$.

Proof. We may assume that $x \neq 0$ and $y \neq 0$, otherwise the statement is trivial. For $t \in \mathbb{R}$,

$$\begin{aligned} 0 &\leq \|x - ty\|^2 \\ &= \langle x - ty, x - ty \rangle \\ &= \langle x, x\rangle - t\langle x, y\rangle - t\langle y, x\rangle + t^2\langle y, y\rangle \\ &= \|y\|^2 t^2 - 2t\mathrm{Re}\langle x, y\rangle + \|x\|^2 \\ &= \|y\|^2 \left(t - \frac{\mathrm{Re}\langle x, y\rangle}{\|y\|^2}\right)^2 + \left(\|x\|^2 - \left(\frac{\mathrm{Re}\langle x, y\rangle}{\|y\|}\right)^2\right). \end{aligned}$$

Taking $t = \frac{\mathrm{Re}\langle x,y\rangle}{\|y\|^2}$, we get

$$|\mathrm{Re}\langle x, y\rangle| \leq \|x\| \ \|y\|$$

for every $x, y \in \mathcal{H}$. Now $\langle x, y\rangle = |\langle x, y\rangle| \, \mathrm{e}^{\mathrm{i}\phi}$ for some $\phi \in \mathbb{R}$, and

$$|\langle x, y\rangle| = \langle \mathrm{e}^{-\mathrm{i}\phi} x, y\rangle = \left|\mathrm{Re}\langle \mathrm{e}^{-\mathrm{i}\phi} x, y\rangle\right| \leq \|\mathrm{e}^{-\mathrm{i}\phi} x\| \ \|y\| = \|x\| \ \|y\|.$$

This completes the proof. □

Corollary B.5.7 (**Triangle inequality**)**.** *Let $\mathcal{H}$ be an inner product space. Then*

$$\|x + y\| \leq \|x\| + \|y\|.$$

Consequently, the canonical norm of an inner product space is a norm in the usual sense.

Proof. Now

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &\overset{(B.2)}{\leq} \|x\|^2 + 2\ \|x\| \ \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2, \end{aligned}$$

completing the proof. □

Remark B.5.8. One may naturally study $\mathbb{R}$-Hilbert spaces, where the scalar field is $\mathbb{R}$ and the inner product takes real values. Then

$$\langle x, y\rangle = \frac{\|x\|^2 + \|y\|^2 - \|x - y\|^2}{2}.$$

Thus the inner product can be recovered from the norm here.

Exercise B.5.9. Prove this remark. In ($\mathbb{C}$-) Hilbert spaces, prove that

$$\langle x, y\rangle = \frac{\|x+y\|^2 - \|x-y\|^2 + \mathrm{i}\|x+\mathrm{i}y\|^2 - \mathrm{i}\|x-\mathrm{i}y\|^2}{4}.$$

Exercise B.5.10. Every Hilbert space is canonically a Banach space, but not vice versa: in a real Banach space, $(x,y) \mapsto (\|x\|^2+\|y\|^2-\|x-y\|^2)/2$ does not always define an inner product. Present some examples.

Lemma B.5.11. *Let $\mathcal{H}$ be a Hilbert space. Suppose $C \subset \mathcal{H}$ is closed, convex and non-empty. Then there exists unique $z \in C$ such that $\|z\| = \inf\{\|x\| :\ x \in C\}$.*

Proof. Let $r := \inf\{\|x\| :\ x \in C\}$. For any $x, y \in \mathcal{H}$, the *parallelogram identity*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \tag{B.3}$$

holds. Take a sequence $(x_k)_{k=1}^{\infty}$ in C such that $\|x_k\| \to_{k\to\infty} r$. Now $(x_j+x_k)/2 \in C$ due to convexity, so that $4r^2 \leq \|x_j+x_k\|^2$. Hence

$$\begin{aligned} 4r^2 + \|x_j - x_k\|^2 &\leq \|x_j+x_k\|^2 + \|x_j-x_k\|^2 \\ &\overset{(B.3)}{=} 2(\|x_j\|^2 + \|x_k\|^2) \\ &\xrightarrow[j,k\to\infty]{} 4r^2, \end{aligned}$$

implying $\|x_j - x_k\| \to_{j,k\to\infty} 0$. Thus $(x_k)_{k=1}^{\infty}$ is a Cauchy sequence, converging to some $z \in C$ with $\|z\| = r$ (recall that $\mathcal{H}$ is complete and $C \subset \mathcal{H}$ is closed). If $z' \in C$ satisfies $\|z'\| = d$ then the alternating sequence $(z, z', z, z', \cdots)$ would be a Cauchy sequence, by the reasoning above: hence $z = z'$. □

Exercise B.5.12 (Parallelogram identity). Show that the parallelogram identity (B.3):

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

holds for all $x, y \in \mathcal{H}$.

Lemma B.5.13. *Let M be a vector subspace in a Hilbert space $\mathcal{H}$. Let $\|z\| \leq \|z+u\|$ for every $u \in M$. Then $z \in M^{\perp}$.*

Proof. To get a contradiction, assume $\langle z, v\rangle \neq 0$ for some $v \in M$. Multiplying v by a scalar, we may assume that $\mathrm{Re}\langle z, v\rangle \neq 0$. If $r \in \mathbb{R}$ then

$$0 \leq \|z - rv\|^2 - \|z\|^2 = r^2\|v\|^2 - 2r\mathrm{Re}\langle z, v\rangle = r(r\|v\|^2 - 2\mathrm{Re}\langle z, v\rangle),$$

but this inequality fails when r is between 0 and $2\mathrm{Re}\langle z, v\rangle/\|v\|^2$. □

Definition B.5.14 (Orthogonal projection). Let M be a closed subspace of a Hilbert space $\mathcal{H}$. Then we may define $P_M : \mathcal{H} \to \mathcal{H}$ so that $P_M(x) \in M$ is the point in M closest to $x \in \mathcal{H}$. Mapping P_M is called the *orthogonal projection onto M*.

Proposition B.5.15. *Operator $P_M : \mathcal{H} \to \mathcal{H}$ defined above is linear, and $\|P_M\| = 1$ (unless $M = \{0\}$). Moreover, $P_{M^\perp} = I - P_M$.*

Proof. Let $x \in \mathcal{H}$, $P := P_M$ and $Q = I - P$. By Definition B.5.14, $P(x) \in M$ and

$$\|Q(x)\| \leq \|Q(x) + u\|$$

for every $u \in M$. This implies $Q(x) \in M^\perp$ by Lemma B.5.13. Let $x, y \in \mathcal{H}$ and $\lambda, \mu \in \mathbb{C}$. Since

$$\begin{aligned} \lambda x &= \lambda\left(P(x) + Q(x)\right), \\ \mu y &= \mu\left(P(y) + Q(y)\right), \\ \lambda x + \mu y &= P(\lambda x + \mu y) + Q(\lambda x + \mu y), \end{aligned}$$

we get

$$M \ni P(\lambda x + \mu y) - \lambda P(x) - \mu P(y) = \lambda Q(x) + \mu Q(y) - Q(\lambda x + \mu y) \in M^\perp.$$

This implies the linearity of P, because $M \cap M^\perp = \{0\}$. Finally,

$$\|x\|^2 = \|Px + Qx\|^2 = \|Px\|^2 + \|Qx\|^2 + 2\Re\langle Px, Qx\rangle = \|Px\|^2 + \|Qx\|^2;$$

in particular, $\|Px\| \leq \|x\|$. □

Remark B.5.16. We have proven that

$$\mathcal{H} = M \oplus M^\perp.$$

This means that $M, M^\perp$ are closed subspaces of the Hilbert space $\mathcal{H}$ such that $M \perp M^\perp$ and that $M + M^\perp = \mathcal{H}$.

Definition B.5.17 (Direct sum). Let $\{\mathcal{H}_j : j \in J\}$ be a family of pair-wise orthogonal closed subspaces of $\mathcal{H}$. If the span of $\bigcup_{j \in J} \mathcal{H}_j$ is dense in $\mathcal{H}$ then H is said to be a *direct sum* of $\{\mathcal{H}_j : j \in J\}$, denoted by

$$\mathcal{H} = \bigoplus_{j \in J} \mathcal{H}_j.$$

If $\mathcal{H}$ is a direct sum of $\{M_j\}_{j=1}^k$, we denote $\mathcal{H} = \bigoplus_{j=1}^k M_j$. Especially, $M_1 \oplus M_2 = \bigoplus_{j=1}^2 M_j$.

Remark B.5.18. If $\mathcal{H}$ be a Hilbert space, it is easy to see that $f = (x \mapsto \langle x, y\rangle) : \mathcal{H} \to \mathbb{C}$ is a linear functional, and $\|f\| = \|y\|$ due to the Cauchy–Schwarz inequality and to $f(y) = \|y\|^2$. Actually, there are no other kinds of bounded linear functionals on a Hilbert space:

Theorem B.5.19 (Riesz (Hilbert Space) Representation Theorem). *Let $f : \mathcal{H} \to \mathbb{C}$ be a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there exists a unique $y \in \mathcal{H}$ such that*

$$f(x) = \langle x, y \rangle$$

for every $x \in \mathcal{H}$. Moreover, $\|f\| = \|y\|$.

Sometimes this theorem is also called the Fréchet–Riesz (representation) theorem.

Proof. Assume the non-trivial case $f \neq 0$. Thus we may choose $u \in \mathrm{Ker}(f)^{\perp}$ for which $\|u\| = 1$. Pursuing for a suitable representative $y \in \mathcal{H}$, we notice that $f(u) = \langle u, \overline{f(u)}u \rangle$, inspiring an investigation:

$$\begin{aligned} \langle x, \overline{f(u)}u \rangle - f(x) &= \langle f(u)x, u \rangle - \langle f(x)u, u \rangle \\ &= \langle f(u)x - f(x)u, u \rangle \\ &= 0, \end{aligned}$$

since $f(u)x - f(x)u \in \mathrm{Ker}(f)$. Thus $f(x) = \langle x, \overline{f(u)}u \rangle$ for every $x \in \mathcal{H}$. Furthermore, if $f(x) = \langle x, y \rangle = \langle x, z \rangle$ for every $x \in \mathcal{H}$ then

$$0 = f(x) - f(x) = \langle x, y \rangle - \langle x, z \rangle = \langle x, y - z \rangle \overset{x=y-z}{=} \|y - z\|^2,$$

so that $y = z$. □

Definition B.5.20 (Adjoint operator). Let $\mathcal{H}$ be a Hilbert space, $z \in \mathcal{H}$ and $A \in \mathcal{L}(\mathcal{H})$. Then a bounded linear functional on $\mathcal{H}$ is defined by $x \mapsto \langle Ax, z \rangle$, so that by Theorem B.5.19 there exists a unique vector $A^*z \in \mathcal{H}$ satisfying

$$\langle Ax, z \rangle = \langle x, A^*z \rangle$$

for every $x \in \mathcal{H}$. This defines a mapping $A^* : \mathcal{H} \to \mathcal{H}$, which is called the *adjoint* of $A \in \mathcal{L}(\mathcal{H})$. If $A^* = A$ then A is called *self-adjoint*.

Exercise B.5.21. Let $\lambda \in \mathbb{C}$ and $A, B \in \mathcal{L}(\mathcal{H})$. Show that $(\lambda A)^* = \overline{\lambda} A^*$, $(A+B)^* = A^* + B^*$ and $(AB)^* = B^*A^*$.

Exercise B.5.22. Show that the adjoint operator $A^* : \mathcal{H} \to \mathcal{H}$ of $A \in \mathcal{L}(\mathcal{H})$ is linear and bounded. Moreover, show that $(A^*)^* = A$, $\|A^*A\| = \|A\|^2$ and $\|A^*\| = \|A\|$.

Lemma B.5.23. *Let $A^* = A \in \mathcal{L}(\mathcal{H})$. Then*

$$\|A\| = \sup_{x:\ \|x\| \leq 1} |\langle Ax, x \rangle| \,.$$

Proof. Let $r := \sup\{|\langle Ax, x \rangle| :\ x \in \mathcal{H},\ \|x\| \leq 1\}$. Then

$$r \overset{(B.2)}{\leq} \sup_{x:\ \|x\| \leq 1} \|Ax\|\ \|x\| \leq \|A\|.$$

Let us assume that $Ax \neq 0$ for $\|x\| = 1$, and let $y := Ax/\|Ax\|$. Since $A^* = A$, we have $\langle Ax, y\rangle = \langle x, Ay\rangle = \langle Ay, x\rangle \in \mathbb{R}$, so that

$$\begin{aligned}
\|Ax\| &= \langle Ax, y\rangle \\
&\overset{A^*=A}{=} \frac{1}{4}\left(\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle\right) \\
&\leq \frac{1}{4}\left(|\langle A(x+y), x+y\rangle| + |\langle A(x-y), x-y\rangle|\right) \\
&\leq \frac{r}{4}\left(\|x+y\|^2 + \|x-y\|^2\right) \\
&\overset{(B.3)}{=} \frac{r}{2}\left(\|x\|^2 + \|y\|^2\right) \\
&= r.
\end{aligned}$$

This concludes the proof. □

Lemma B.5.24. *Let $\mathcal{H} \neq \{0\}$. Let $A^* = A \in \mathcal{L}(\mathcal{H})$ be compact. Then there exists a non-zero $x \in \mathcal{H}$ such that $Ax = +\|A\|x$ or $Ax = -\|A\|x$.*

Proof. Assume the non-trivial case $\|A\| > 0$. By Lemma B.5.23, we may choose $\lambda \in \{\pm\|A\|\}$ to be an accumulation point of the set $\{\langle Ax, x\rangle : \ x \in \mathcal{H},\ \|x\| \leq 1\}$. For each $k \in \mathbb{Z}^+$, take $x_k \in \mathcal{H}$ such that $\|x_k\| \leq 1$ and $\langle Ax_k, x_k\rangle \to_{k\to\infty} \lambda$. Since A is compact, by Theorem A.13.4 it follows that the sequence $(Ax_k)_{k=1}^{\infty}$ has a convergent subsequence; omitting elements from the sequence, we may assume that $z := \lim_k Ax_k \in \mathcal{H}$ exists. Now

$$\begin{aligned}
0 &\leq \|Ax_k - \lambda x_k\|^2 \\
&= \|Ax_k\|^2 + \lambda^2\|x_k\|^2 - 2\lambda\langle Ax_k, x_k\rangle \\
&\leq \|A\|^2 + \lambda^2 - 2\lambda\langle Ax_k, x_k\rangle \\
&\xrightarrow[k\to\infty]{} 0,
\end{aligned}$$

implying that $\lim_k \lambda x_k$ exists and equals to $\lim_k Ax_k = z$. Finally, let $x := z/\lambda$, so that by continuity $Ax = \lim_k Ax_k = \lambda x$. □

Theorem B.5.25 (Diagonalisation of compact self-adjoint operators). *Let $\mathcal{H}$ be infinite-dimensional and $A^* = A \in \mathcal{L}(\mathcal{H})$ be compact. Then there exist $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}$ and an orthonormal set $\{x_k\}_{k=1}^{\infty} \subset \mathcal{H}$ such that $|\lambda_{k+1}| \leq |\lambda_k|$, $\lim_k \lambda_k = 0$ and*

$$Ax = \sum_{k=1}^{\infty} \lambda_k \langle x, x_k\rangle\ x_k$$

for every $x \in \mathcal{H}$.

Proof. By Lemma B.5.24, take $\lambda_1 \in \mathbb{R}$ and $x_1 \in \mathcal{H}$ such that $\|x_1\| = 1$, $Ax_1 = \lambda_1 x_1$ and $\|A_1\| = |\lambda_1|$. Then we proceed by induction as follows. Let $\mathcal{H}_k := \left(\{x_j\}_{j=1}^{k-1}\right)^{\perp}$. Then $A_k^* = A_k := A|_{\mathcal{H}_k} \in \mathcal{L}(\mathcal{H}_k)$ is compact as it is a finite-dimensional operator, so we may apply Lemma B.5.24 to choose $\lambda_k \in \mathbb{R}$ and $x_k \in \mathcal{H}_k$ such that $\|x_k\| = 1$, $Ax_k = \lambda_k x_k$ and $\|A_k\| := |\lambda_k|$. Since $\mathcal{H}$ is infinite-dimensional, we obtain an orthonormal family $\{x_k\}_{k=1}^{\infty} \subset \mathcal{H}$, and $Ax_k = \lambda_k x_k$ for each $k \in \mathbb{Z}^+$, where $|\lambda_{k+1}| \leq |\lambda_k|$.

Since A is compact, $(Ax_k)_{k=1}^{\infty}$ has a converging subsequence. Actually, $(A_k)_{k=1}^{\infty}$ itself must converge and $\lambda_k \to 0$, because

$$\|Ax_j - Ax_k\| = \|\lambda_j x_j - \lambda_k x_k\| = \sqrt{\lambda_j^2 + \lambda_k^2} \geq |\lambda_k|$$

for every $j, k \in \mathbb{Z}^+$. If $x \in \mathcal{H}$ then $z_k := x - \sum_{j=1}^{k-1} \langle x, x_k \rangle \; x_k \in \mathcal{H}_k$, and

$$\|Az_k\| = \|A_k z_k\| \leq \|A_k\| \|z_k\| = |\lambda_k| \; \|z_k\| \leq |\lambda_k| \; \|x\| \xrightarrow[k\to\infty]{} 0,$$

completing the proof. □

Corollary B.5.26 (**Hilbert–Schmidt Spectral Theorem).** *Let $A^* = A \in \mathcal{L}(\mathcal{H})$ be compact. Then $\sigma(A)$ is at most countable, and $\operatorname{Ker}(\lambda I - A)$ is finite-dimensional if $0 \neq \lambda \in \sigma(A)$. Moreover, $\sigma(A) \setminus \{0\}$ is discrete, and*

$$\mathcal{H} = \bigoplus_{\lambda \in \sigma(A)} \operatorname{Ker}(\lambda I - A).$$

Exercise B.5.27. Prove the Hilbert–Schmidt Spectral Theorem using Theorem B.5.25.

Definition B.5.28 (**Weak topology on a Hilbert space).** The *weak topology* of a Hilbert space $\mathcal{H}$ is the smallest topology for which mappings

$$(u \mapsto \langle u, v \rangle_{\mathcal{H}}) : \mathcal{H} \to \mathbb{C}$$

are continuous for all $v \in \mathcal{H}$.

Exercise B.5.29 (**Weak = weak* in Hilbert spaces).** Show that Hilbert spaces are reflexive. Prove that in a Hilbert space the weak topology is the same as the weak*-topology, introduced in Definition B.4.35.

As a consequence of Exercise B.5.29 and the Banach–Alaoglu Theorem B.4.36 we obtain:

Theorem B.5.30 (**Banach–Alaoglu Theorem for Hilbert spaces).** *Let $\mathcal{H}$ be a Hilbert space. Its closed unit ball*

$$\overline{\mathbb{B}} = \{v \in \mathcal{H} : \; \|v\|_{\mathcal{H}} \leq 1\}$$

is compact in the weak topology.

Exercise B.5.31. Let $\{e_\alpha\}_{\alpha \in I}$ be an orthonormal collection in $\mathcal{H}$ and let $x \in \mathcal{H}$. Show that

$$\sum_{\alpha \in I} |\langle x, e_\alpha \rangle|^2 \leq \|x\|^2. \tag{B.4}$$

(Hint: Pythagoras' theorem.) Consequently, deduce from Exercise B.4.3 that the set of α such that $\langle x, e_\alpha \rangle \neq 0$ is at most countable.

We finish with the following theorem which is of importance, because it allows one to decompose elements into "simpler ones", which is particularly important in applications.

Theorem B.5.32 (**Orthonormal sets in Hilbert space).** *Let $\{e_\alpha\}_{\alpha\in I}$ be an orthonormal set in the Hilbert space $\mathcal{H}$. Then the following conditions are equivalent:*

(i) *For every $x \in \mathcal{H}$ there are only countably many $\alpha \in I$ such that $\langle x, e_\alpha\rangle \neq 0$, and the equality*
$$x = \sum_{\alpha\in I} \langle x, e_\alpha\rangle e_\alpha$$
holds, where the series is converging in norm, independent of any ordering of its terms.

(ii) *If $\langle x, e_\alpha\rangle = 0$ for all $\alpha \in I$, then $x = 0$.*

(iii) *(Plancherel's identity) For every $x \in \mathcal{H}$ holds $\|x\|^2 = \sum_{\alpha\in I} |\langle x, e_\alpha\rangle|^2$.*

Proof. (i) ⇒ (iii). This follows by enumerating countably many e_α's with $\langle x, e_\alpha\rangle \neq 0$ by $\{e_j\}_{j=1}^\infty$, and using the identity
$$\|x\|^2 - \sum_{j=1}^n |\langle x, e_j\rangle|^2 = \|x - \sum_{j=1}^n \langle x, e_j\rangle e_j\|^2.$$
(iii) ⇒ (ii) is automatic. Finally, let us show (ii) ⇒ (i). It follows from the last part of Exercise B.5.31 that the collection of e_α with $\langle x, e_\alpha\rangle \neq 0$ is countable, so it can be enumerated by $\{e_j\}_{j=1}^\infty$. Now, the identity
$$\|\sum_{j=j_1}^{j_2} \langle x, e_j\rangle e_j\|^2 = \sum_{j=j_1}^{j_2} \|\langle x, e_j\rangle\|^2$$
and (B.4) imply that the right hand side → 0 as $j_1, j_2 \to \infty$. This means that the series $\sum_{j=1}^\infty \langle x, e_j\rangle e_j$ converges. Setting $y := x - \sum_{j=1}^\infty \langle x, e_j\rangle e_j$ we see that $\langle y, e_\alpha\rangle = 0$ for all $\alpha \in I$, which implies that $y = 0$. □

Exercise B.5.33. Verify the identities stated in the proof.

Definition B.5.34 (**Orthonormal basis).** An orthonormal set satisfying conditions of Theorem B.5.32 is called an *orthonormal basis* of the Hilbert space $\mathcal{H}$. Then we have the following properties

Theorem B.5.35 (**Every Hilbert space has an orthonormal basis).** *Every Hilbert space $\mathcal{H}$ has an orthonormal basis. An orthonormal basis is countable if and only if $\mathcal{H}$ is separable, in which case any other basis is also countable.*

Exercise B.5.36. Prove Theorem B.5.35: the first part follows from Zorn's lemma if we order orthonormal collections by inclusion, since the maximal element would satisfy property (ii) of Theorem B.5.32. The second part follows from the Gram–Schmidt process.

B.5.1 Trace class, Hilbert–Schmidt, and Schatten classes

Definition B.5.37 (Trace class operators). Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\{e_j \mid j \in J\}$. Let $A \in \mathcal{L}(\mathcal{H})$. Let us denote

$$\|A\|_{S_1} := \sum_{j \in J} |\langle A e_j, e_j \rangle_{\mathcal{H}}| \,;$$

this is the *trace norm* of A, and the *trace class* is the (Banach) space

$$S_1 = S_1(\mathcal{H}) := \{A \in \mathcal{L}(\mathcal{H}) : \ \|A\|_{S_1} < \infty\} \,.$$

The *trace* is the linear functional $\mathrm{Tr} : S_1(\mathcal{H}) \to \mathbb{C}$,

$$A \mapsto \sum_{j \in J} \langle A e_j, e_j \rangle_{\mathcal{H}}.$$

Exercise B.5.38. Verify that the definition of the trace is independent of the choice of the orthonormal basis for $\mathcal{H}$. Consequently, if $(a_{ij})_{i,j \in J}$ is the matrix representation of $A \in S_1$ with respect to the chosen basis, then $\mathrm{Tr}(A) = \sum_{j \in J} a_{jj}$.

Exercise B.5.39 (Properties of trace). Prove the following properties of the trace functional:

$$\begin{aligned}
\mathrm{Tr}(AB) &= \mathrm{Tr}(BA), \\
\mathrm{Tr}(A^*) &= \overline{\mathrm{Tr}(A)}, \\
\mathrm{Tr}(A^* A) &\geq 0, \\
\mathrm{Tr}(A \oplus B) &= \mathrm{Tr}(A) + \mathrm{Tr}(B), \\
\dim(\mathcal{H}) < \infty &\Rightarrow \begin{cases} \mathrm{Tr}(I_{\mathcal{H}}) = \dim(\mathcal{H}), \\ \mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\ \mathrm{Tr}(B). \end{cases}
\end{aligned}$$

Exercise B.5.40 (Trace on a finite-dimensional space). Show that the trace on a finite-dimensional vector space is independent of the choice of inner product. Thus, the trace of a square matrix is defined to be the sum of its diagonal elements; moreover, the trace is the sum of the eigenvalues (with multiplicities counted).

Exercise B.5.41. Let $\mathcal{H}$ be finite-dimensional. Let $f : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ be a linear functional satisfying

$$\begin{cases} f(AB) = f(BA), \\ f(A^* A) \geq 0, \\ f(I_{\mathcal{H}}) = \dim(\mathcal{H}) \end{cases}$$

for all $A, B \in \mathcal{L}(\mathcal{H})$. Show that $f = \mathrm{Tr}$.

Definition B.5.42 (Hilbert-Schmidt operators). The space of *Hilbert–Schmidt operators* is

$$S_2 = S_2(\mathcal{H}) := \{A \in \mathcal{L}(\mathcal{H}) : \ A^*A \in S_1(\mathcal{H})\} ,$$

and it can be endowed with a Hilbert space structure via the inner product

$$\langle A, B\rangle_{S_2} := \mathrm{Tr}(AB^*).$$

The Hilbert–Schmidt norm is then

$$\|A\|_{HS} = \|A\|_{S_2} := \langle A, A\rangle_{S_2}^{1/2}.$$

The case of the Hilbert–Schmidt norm on the finite-dimensional spaces will be discussed in more detail in Section 12.6.

Remark B.5.43. In general, there are inclusions $S_1 \subset S_2 \subset \mathcal{K} \subset S_\infty$, where $S_\infty := \mathcal{L}(\mathcal{H})$ and $\mathcal{K} \subset S_\infty$ is the subspace of compact linear operators. Moreover,

$$\|A\|_{S_\infty} \leq \|A\|_{S_2} \leq \|A\|_{S_1}$$

for all $A \in S_\infty$. One can show that the dual $\mathcal{K}' = \mathcal{L}(\mathcal{K}, \mathbb{C})$ is isometrically isomorphic to S_1, and that $(S_1)'$ is isometrically isomorphic to S_∞. In the latter case, it turns out that a bounded linear functional on S_1 is of the form $A \mapsto \mathrm{Tr}(AB)$ for some $B \in S_\infty$. These phenomena are related to properties of the sequence spaces $\ell^p = \ell^p(\mathbb{Z}^+)$. In analogy to the operator spaces, $\ell^1 \subset \ell^2 \subset c_0 \subset \ell^\infty$, where c_0 is the space of sequences converging to 0, playing the counterpart of space $\mathcal{K}$.

Remark B.5.44 (**Schatten classes**). Trace class operators S_1 and Hilbert–Schmidt operators S_2 turn out to be special cases of the *Schatten classes* S_p, $1 \leq p < \infty$. These classes can be introduced with the help of the singular values $\mu^2 \in \sigma(A^*A)$. To avoid the technicalities we assume that all the operators below are compact. Thus, for $A \in \mathcal{L}(\mathcal{H})$ we set

$$\|A\|_{S_p} := \left(\sum_{\mu^2 \in \sigma(A^*A)} \mu^p \right)^{1/p} .$$

We note that operators that satisfy $\|A\|_{S_p} < \infty$ must have at most countable spectrum $\sigma(A^*A)$ in view of Exercise B.4.3, but in our case this is automatically satisfied since we assumed that A is compact. Therefore, denoting the sequence of singular values $\mu_j^2 \in \sigma(A^*A)$, counted with multiplicities, we have

$$\|A\|_{S_p} = \|\{\mu_j\}_j\|_{\ell^p}.$$

The Schatten class S_p is then defined as the space

$$S_p = S_p(\mathcal{H}) := \left\{A \in \mathcal{L}(\mathcal{H}) : \ \|A\|_{S_p} < \infty\right\} .$$

With this norm, $S_p(\mathcal{H})$ is a Banach space, and $S_2(\mathcal{H})$ is a Hilbert space. In analogy to the trace class and Hilbert–Schmidt operators, once can show that actually $\|A\|_{S_p}^p = \mathrm{Tr}(|A^*A|^{p/2}) = \mathrm{Tr}(|A|^p)$ for a compact operator A.

Exercise B.5.45. Show that the Schatten classes S_1 and S_2 coincide with the previously defined trace class and Hilbert–Schmidt operators.

Exercise B.5.46 (Hölder's inequality for Schatten classes). Show that a Schatten class S_p is an ideal in $\mathcal{L}(\mathcal{H})$. Let $\mathcal{H}$ be separable. Show that if $1 \leq p \leq \infty$, $\frac{1}{p}+\frac{1}{q} = 1$, $A \in S_p$ and $B \in S_q$, then

$$\|AB\|_{S_1} \leq \|A\|_{S_p} \|B\|_{S_q}.$$

(Hint: approximate operators by matrices.)

Chapter C

Measure theory and integration

This chapter provides sufficient general information about measures and integration. The starting point is the concept of an outer measure, which "measures weights of subsets of a space". We should first consider how to sum such weights, which are either infinite or non-negative real numbers. For a finite set K, notation

$$\sum_{j \in K} a_j$$

abbreviates the usual sum of numbers $a_j \in [0, \infty]$ over the index set K. The conventions here are that $a + \infty = \infty$ for all $a \in [0, \infty]$, and that

$$\sum_{j \in \emptyset} a_j = 0.$$

Infinite summations are defined by limits as follows:

Definition C.0.47. The *sum* of numbers $a_j \in [0, \infty]$ over the index set J is

$$\sum_{j \in J} a_j := \sup \left\{ \sum_{j \in K} a_j : \quad K \subset J \text{ is finite} \right\}.$$

Exercise C.0.48. Let $0 < a_j < \infty$ for each $j \in J$. Suppose

$$\sum_{j \in J} a_j < \infty.$$

Show that J is at most countable.

The message of Exercise C.0.48 is that for positive numbers, only countable summations are interesting. In the measure theory, where summations are fundamental, such a "restriction to countability" will be encountered recurrently.

C.1 Measures and outer measures

Recall that for a set X, by $\mathcal{P}(X) := \{E \mid E \subset X\}$ we denote its *power set*, i.e. the family of all subsets of X. Let us write $E^c := X \setminus E = \{x \in X : x \notin E\}$ for the complement set, when the space X is implicitely known from the context.

C.1.1 Measuring sets

Definition C.1.1 (Outer measure). A mapping $\psi : \mathcal{P}(X) \to [0, \infty]$ is an *outer measure* on a set $X \neq \emptyset$ if

$$\begin{aligned} &\psi(\emptyset) = 0, \\ &E \subset F \quad \Rightarrow \quad \psi(E) \leq \psi(F), \\ &\psi\Big(\bigcup_{j=1}^{\infty} E_j\Big) \leq \sum_{j=1}^{\infty} \psi(E_j). \end{aligned}$$

for every $E, F \subset X$ and $\{E_j\}_{j=1}^{\infty} \subset \mathcal{P}(X)$.

Intuitively, an outer measure is weighing the subsets of a space.

Example. Define $\psi : \mathcal{P}(X) \to [0, \infty]$ by $\psi(\emptyset) = 0$ and $\psi(E) = 1$, when $\emptyset \neq E \subset X$. This is an outer measure.

Example. Let $\psi : \mathcal{P}(X) \to [0, \infty]$, where $\psi(E)$ is the number of the points in the set $E \subset X$. Such an outer measure is called a *counting measure* for obvious reasons.

At first sight, constructing meaningful non-trivial outer measures may appear difficult. However, there is an easy and useful method for generating outer measures out of simpler set functions, which we call the measurelets:

Definition C.1.2 (Measurelets). Let $\mathcal{A} \subset \mathcal{P}(X)$ cover X, i.e. $X = \bigcup \mathcal{A}$. We call a mapping $m : \mathcal{A} \to [0, \infty]$ a *measurelet* on X. Members of the family $\mathcal{A}$ are called the *elementary sets*. A measurelet $m : \mathcal{A} \to [0, \infty]$ on X *generates* a mapping $m^* : \mathcal{P}(X) \to [0, \infty]$ defined by

$$m^*(E) := \inf \left\{ \sum_{A \in \mathcal{B}} m(A) : \quad \mathcal{B} \subset \mathcal{A} \text{ is countable, } E \subset \bigcup \mathcal{B} \right\}.$$

Exercise C.1.3. Let

$$\mathcal{A} := \{\emptyset, \mathbb{R}^2\} \cup \{S \subset \mathbb{R}^2 : S \text{ a finite union of polygons}\}.$$

Let us define a measurelet $A : \mathcal{A} \to [0, \infty]$ by the following informal demands:

(1) $A(\mathit{rectangle}) = \mathit{base} \cdot \mathit{height}$.

(2) $A(S_1 \cup S_2) = A(S_1) + A(S_2)$, if the interiors of the sets S_1, S_2 are disjoint.

(3) The measurelet A does not change in translations nor rotations of sets.

Using these rules, calculate the measurelets of a parallelogram and a triangle.

Apparently, there are plenty of measurelets: almost anything goes. Especially, outer measures are measurelets.

Theorem C.1.4. *Let $m : \mathcal{A} \to [0, \infty]$ be a measurelet on a set X. Then $m^* : \mathcal{P}(X) \to [0, \infty]$ is an outer measure for which $m^*(A) \leq m(A)$ for every $A \in \mathcal{A}$.*

Proof. Clearly, $m^* : \mathcal{P}(X) \to [0, \infty]$ is well-defined, and $m^*(A) \leq m(A)$ for every $A \in \mathcal{A}$. We see that $m^*(\emptyset) = 0$, because $\sum_{A \in \emptyset} m(A) = 0$, $\emptyset \subset \mathcal{A}$ is countable, and $\emptyset \subset \bigcup \emptyset$. Next, if $E \subset F \subset X$ then $m^*(E) \leq m^*(F)$, because any cover $\{A_j\}_{j=1}^{\infty}$ of F is also a cover of E. Lastly, let $\{E_j\}_{j=1}^{\infty} \subset \mathcal{P}(X)$. Take $\varepsilon > 0$. For each $j \geq 1$, choose $\{A_{jk}\}_{k=1}^{\infty} \subset \mathcal{A}$ such that

$$E_j \subset \bigcup_{k=1}^{\infty} A_{jk} \quad \text{and} \quad m^*(E_j) + 2^{-j}\varepsilon \geq \sum_{k=1}^{\infty} m(A_{jk}).$$

Then $\{A_{jk}\}_{j,k=1}^{\infty} \subset \mathcal{A}$ is a countable cover of $\bigcup_{j=1}^{\infty} E_j \subset X$, and

$$\begin{aligned} m^*(\bigcup_{j=1}^{\infty} E_j) &\leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} m(A_{jk}) \\ &\leq \sum_{j=1}^{\infty} m^*(E_j) + \varepsilon. \end{aligned}$$

Thus $m^*(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} m^*(E_j)$; the proof is complete. □

Definition C.1.5 (Lebesgue's outer measure). On the Euclidean space $X = \mathbb{R}^n$, let us define the partial order $\leq$ by

$$a \leq b \quad \overset{\text{definition}}{\Longleftrightarrow} \quad \forall j \in \{1, \cdots, n\} : \ a_j \leq b_j.$$

When $a \leq b$, let the *n-interval* be

$$[a, b] := [a_1, b_1] \times \cdots \times [a_n, b_n] = \{x \in \mathbb{R}^n : \ a \leq x \leq b\}.$$

For $\mathcal{A} = \{[a, b] : \ a, b \in X, \ a \leq b\}$ let us define the *Lebesgue measurelet* $m : \mathcal{A} \to [0, \infty]$ by

$$m([a, b]) := volume([a, b]) = \prod_{j=1}^{n} |a_j - b_j|.$$

Then the generated outer measure $\lambda^* = \lambda^*_{\mathbb{R}^n} := m^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ is called the *Lebesgue outer measure* of $\mathbb{R}^n$.

Exercise C.1.6. Give an example of an outer measure that cannot be generated by a measurelet.

Definition C.1.7 (Outer measure measurability). Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an outer measure. A set $E \subset X$ is called *ψ-measurable* if

$$\psi(S) = \psi(E \cap S) + \psi(E^c \cap S)$$

for every $S \subset X$, where $E^c = X \backslash E$. The family of ψ-measurable sets is denoted by

$$\mathcal{M}(\psi) \subset \mathcal{P}(X).$$

Remark C.1.8. Notice that trivially

$$\psi(S) \leq \psi(E \cap S) + \psi(E^c \cap S)$$

by the properties of the outer measure. Intuitively, a measurable set E "sharply cuts" "rough" sets $S \subset X$ into two disjoint pieces, $E \cap S$ and $E^c \cap S$.

Remark C.1.9 (**Non-measurability**). The axiom of choice can be used to "construct" a subset $E \subset \mathbb{R}^n$ which is not Lebesgue measurable. We will discuss this topic in Section C.1.4.

Exercise C.1.10. Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an outer measure measure and $E \subset X$. Define $\psi_E : \mathcal{P}(X) \to [0,\infty]$ by $\psi_E(S) := \psi(E \cap S)$. Show that ψ_E is an outer measure for which $\mathcal{M}(\psi) \subset \mathcal{M}(\psi_E)$.

Lemma C.1.11. *Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an outer measure and $\psi(E) = 0$. Then $E \in \mathcal{M}(\psi)$.*

Proof. Let $S \subset X$. Then

$$\begin{aligned} \psi(S) &\leq \psi(E \cap S) + \psi(E^c \cap S) \\ &\leq \psi(E) + \psi(S) \\ &= \psi(S), \end{aligned}$$

so that $\psi(S) = \psi(E \cap S) + \psi(E^c \cap S)$; set E is ψ-measurable. □

Lemma C.1.12. *Let $E, F \in \mathcal{M}(\psi)$. Then $E^c, E \cap F, E \cup F \in \mathcal{M}(\psi)$.*

Proof. The definition of ψ-measurability is clearly complement symmetric, so that $E \in \mathcal{M}(\psi) \iff E^c \in \mathcal{M}(\psi)$. Next, it is sufficient to deal with $E \cup F$, since $E \cap F = (E^c \cup F^c)^c$. Take $S \subset X$. Then

$$\begin{aligned} \psi(S) &\leq \psi((E \cup F) \cap S) + \psi((E \cup F)^c \cap S) \\ &= \psi((E \cup F) \cap S) + \psi(E^c \cap F^c \cap S) \\ &\overset{E \in \mathcal{M}(\psi)}{=} \psi(E \cap S) + \psi(E^c \cap F \cap S) + \psi(E^c \cap F^c \cap S) \\ &\overset{F \in \mathcal{M}(\psi)}{=} \psi(E \cap S) + \psi(E^c \cap S) \\ &\overset{E \in \mathcal{M}(\psi)}{=} \psi(S). \end{aligned}$$

Hence $\psi(S) = \psi((E \cup F) \cap S) + \psi((E \cup F)^c \cap S)$, so that $E \cup F$ is ψ-measurable. □

Exercise C.1.13. Let $\psi : \mathcal{P}(X) \to [0, \infty]$ be an outer measure. Let $E \subset S \subset X$, $E \in \mathcal{M}(\psi)$ and $\psi(E) < \infty$. Show that $\psi(S \setminus E) = \psi(S) - \psi(E)$.

Definition C.1.14 (**σ-algebras**). A family $\mathcal{M} \subset \mathcal{P}(X)$ is called a *σ-algebra* on X (pronounced: *sigma-algebra*) if

1. $\bigcup \mathcal{E} \in \mathcal{M}$ for every countable collection $\mathcal{E} \subset \mathcal{M}$, and
2. $E^c \in \mathcal{M}$ for every $E \in \mathcal{M}$.

Remark C.1.15. Here, recall the conventions for the union and the intersection of the empty family: for $\mathcal{A} = \emptyset \subset \mathcal{P}(X)$, we naturally define $\bigcup \mathcal{A} = \emptyset$, but notice that $\bigcap \mathcal{A} = X$ (this is not as surprising as it might first seem). Thereby $\mathcal{M}$ is a σ-algebra on X if and only if

1. $\bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$ whenever $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$,
2. $E^c \in \mathcal{M}$ for every $E \in \mathcal{M}$, and
3. $\emptyset \in \mathcal{M}$.

Thus, a σ-algebra on X contains always at least subsets $\emptyset \subset X$ and $X \subset X$.

Proposition C.1.16. *Let $\mathcal{A} \subset \mathcal{P}(X)$. There exists the smallest σ-algebra $\Sigma(\mathcal{A})$ on X containing $\mathcal{A}$, called the* σ-algebra generated by $\mathcal{A}$.

A word of warning: there is no summation in this σ-algebra business here, even though we have used the capital sigma symbol Σ.

Exercise C.1.17. Prove Proposition C.1.16.

Definition C.1.18 (**Borel σ-algebra**). The *Borel σ-algebra* of a topological space (X, τ) is $\Sigma(\tau) \subset \mathcal{P}(X)$. The members of $\Sigma(\tau)$ are called *Borel sets*.

Definition C.1.19 (**Disjoint family**). A family $\mathcal{A} \subset \mathcal{P}(X)$ is called *disjoint* if $A \cap B = \emptyset$ for every $A, B \in \mathcal{A}$ for which $A \neq B$.

Remark C.1.20 (**Disjointisation**). In measure theory, the following "disjointisation" process comes often handy. Let $\mathcal{M}$ be a σ-algebra and $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$. Let $F_1 := E_1$ and

$$F_{k+1} := E_{k+1} \setminus \bigcup_{j=1}^{k} E_j.$$

Now $\{F_k\}_{k=1}^{\infty} \subset \mathcal{M}$ is a disjoint family satisfying $F_k \subset E_k$ and

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{k=1}^{\infty} F_k.$$

Proposition C.1.21. *Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an outer measure. Let $\{F_k\}_{k=1}^{\infty} \subset \mathcal{M}(\psi)$ be disjoint. Then $\bigcup_{k=1}^{\infty} F_k \in \mathcal{M}(\psi)$ and*

$$\psi(\bigcup_{k=1}^{\infty} F_k \cap S) = \sum_{k=1}^{\infty} \psi(F_k \cap S) \tag{C.1}$$

for every $S \subset X$, especially

$$\psi(\bigcup_{k=1}^{\infty} F_k) = \sum_{k=1}^{\infty} \psi(F_k).$$

Proof. Let $E := \bigcup_{k=1}^{\infty} F_k$. Take $S \subset X$. By Lemma C.1.12, $G_n := \bigcup_{k=1}^{n} F_k \in \mathcal{M}(\psi)$. Now

$$\begin{aligned}
\psi(S) &\leq \psi(E \cap S) + \psi(E^c \cap S) \\
&\leq \sum_{k=1}^{\infty} \psi(F_k \cap S) + \psi(E^c \cap S) \\
&= \lim_{n\to\infty} \left(\sum_{k=1}^{n} \psi(F_k \cap S) + \psi(E^c \cap S) \right) \\
&\overset{\{F_k\}_{k=1}^{n} \subset \mathcal{M}(\psi) \text{ disjoint}}{=} \lim_{n\to\infty} \left(\psi(G_n \cap S) + \psi(E^c \cap S) \right) \\
&\overset{E^c \subset G_n^c}{\leq} \lim_{n\to\infty} \left(\psi(G_n \cap S) + \psi(G_n^c \cap S) \right) \\
&\overset{G_n \in \mathcal{M}(\psi)}{=} \psi(S).
\end{aligned}$$

Hence $\psi(S) = \psi(E \cap S) + \psi(E^c \cap S)$, meaning that $E \in \mathcal{M}(\psi)$. Moreover, (C.1) follows from the above chain of (in)equalities. □

Corollary C.1.22. *Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an outer measure. For each $k \geq 1$, let $E_k \in \mathcal{M}(\psi)$ be such that $E_k \subset E_{k+1}$. Then*

$$\psi(\bigcup_{k=1}^{\infty} E_k) = \lim_{k\to\infty} \psi(E_k). \tag{C.2}$$

For each $k \geq 1$, let $F_k \in \mathcal{M}(\psi)$ such that $F_k \supset F_{k+1}$ and $\psi(F_1) < \infty$. Then

$$\psi(\bigcap_{k=1}^{\infty} F_k) = \lim_{k\to\infty} \psi(F_k). \tag{C.3}$$

Proof. Let us assume that $\psi(E_k) < \infty$ for every $k \geq 1$, for otherwise the first claim is trivial. Thereby

$$\begin{aligned}
\psi(\bigcup_{k=1}^{\infty} E_k) &= \psi(E_1 \cup \bigcup_{k=1}^{\infty}(E_{k+1} \setminus E_k)) \\
&\overset{\text{Prop. C.1.21}}{=} \psi(E_1) + \sum_{k=1}^{\infty} \psi(E_{k+1} \setminus E_k) \\
&\overset{\text{Exercise C.1.13}}{=} \psi(E_1) + \lim_{n\to\infty} \sum_{k=1}^{n} (\psi(E_{k+1}) - \psi(E_k)) \\
&= \lim_{n\to\infty} \psi(E_{n+1}).
\end{aligned}$$

Now

$$\begin{aligned}
\psi(F_1) &= \psi((\bigcap_{k=1}^{\infty} F_k) \cup \bigcup_{j=1}^{\infty}(F_1 \setminus F_j)) \\
&\overset{\text{Prop. C.1.21}}{=} \psi(\bigcap_{k=1}^{\infty} F_k) + \psi(\bigcup_{j=1}^{\infty}(F_1 \setminus F_j)) \\
&\overset{\text{(C.2)}}{=} \psi(\bigcap_{k=1}^{\infty} F_k) + \lim_{j\to\infty} \psi(F_1 \setminus F_j) \\
&\overset{\text{Exercise C.1.13}}{=} \psi(\bigcap_{k=1}^{\infty} F_k) + \lim_{j\to\infty} (\psi(F_1) - \psi(F_j)),
\end{aligned}$$

from which (C.3) follows, since $\psi(F_1) < \infty$. □

Exercise C.1.23. Give an example of an outer measure $\psi : \mathcal{P}(X) \to [0, \infty]$ and sets $E_k \subset X$ such that $E_k \subset E_{k+1}$ for all $k \in \mathbb{Z}^+$ and

$$\psi(\bigcup_{k=1}^{\infty} E_k) \neq \lim_{k\to\infty} \psi(E_k).$$

Exercise C.1.24. Give an example that shows the indispensability of the assumption $\psi(F_1) < \infty$ in Corollary C.1.22. For instance, find an outer measure $\varphi : \mathcal{P}(\mathbb{Z}) \to [0, \infty]$ and a family $\{F_k\}_{k=1}^{\infty} \subset \mathcal{M}(\varphi)$ for which

$$\varphi(\bigcap_{k=1}^{\infty} F_k) \neq \lim_{k\to\infty} \varphi(F_k),$$

even though $F_k \supset F_{k+1}$ for all k.

Theorem C.1.25. *Let $\psi : \mathcal{P}(X) \to [0, \infty]$ be an outer measure. Then the ψ-measurable sets form a σ-algebra $\mathcal{M}(\psi)$.*

Proof. $\emptyset \in \mathcal{M}(\psi)$ due to Lemma C.1.11. By Lemma C.1.12, we know that $\mathcal{M}(\psi)$ is closed under taking complements. We must prove that it is closed also under taking countable unions. Let $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}(\psi)$. Applying the disjointisation process of Remark C.1.20, we obtain a disjoint family $\{F_k\}_{k=1}^{\infty} \subset \mathcal{M}(\psi)$, for which $\bigcup_{j=1}^{\infty} E_j = \bigcup_{k=1}^{\infty} F_k$. Exploiting Proposition C.1.21, the proof is concluded. □

Definition C.1.26 (**Measures and measure spaces**). Let $\mathcal{M}$ be a σ-algebra on X. A mapping $\mu : \mathcal{M} \to [0, \infty]$ is called a *measure* on X if

$$\mu(\emptyset) = 0,$$
$$\mu\Big(\bigcup_{j=1}^{\infty} E_j\Big) = \sum_{j=1}^{\infty} \mu(E_j)$$

whenever $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ is a disjoint family. Then the triple $(X, \mathcal{M}, \mu)$ is called a *measure space*; such a measure space and the corresponding measure μ are called

- *finite*, if $\mu(X) < \infty$;
- *probability*, if μ is a finite measure with $\mu(X) = 1$;
- *complete*, if $F \in \mathcal{M}$ whenever there exists $E \in \mathcal{M}$ such that $F \subset E$ and $\mu(E) = 0$;
- *Borel*, if $\mathcal{M} = \Sigma(\tau)$, σ-algebra of the Borel sets in a topological space (X, τ). However, sometimes the Borel condition may mean $\Sigma(\tau) \subset \mathcal{M}$ (more on this later).

Theorem C.1.27. *Let $\psi : \mathcal{P}(X) \to [0, \infty]$ be an outer measure. Then the restriction $\psi|_{\mathcal{M}(\psi)} : \mathcal{M}(\psi) \to [0, \infty]$ is a complete measure.*

Proof. This follows by Proposition C.1.21 and Lemma C.1.11. □

Exercise C.1.28. Let $\mu_k : \mathcal{M} \to [0, \infty]$ be measures for which $\mu_k(E) \leq \mu_{k+1}(E)$ for every $E \in \mathcal{M}$ (and all $k \in \mathbb{Z}^+$). Show that $\mu : \mathcal{M} \to [0, \infty]$, where

$$\mu(E) := \lim_{k \to \infty} \mu_k(E).$$

Exercise C.1.29 (**Borel–Cantelli Lemma**). Let $(X, \mathcal{M}, \mu)$ be a measure space, $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ and

$$E := \left\{ x \in X \mid \{ j \in \mathbb{Z}^+ : \ x \in E_j \} \text{ is infinite} \right\}.$$

Prove that $\mu(E) = 0$ if

$$\sum_{j=1}^{\infty} \mu(E_j) < \infty.$$

This is the so-called *Borel–Cantelli Lemma*.

Remark C.1.30. By Theorem C.1.4, any measure μ generates the outer measure μ^*, whose restriction $\mu^*|_{\mathcal{M}(\mu^*)}$ is a complete measure, which generates an outer measure... Fortunately, this back-and-forth-process terminates, as we shall see in Theorems C.1.35 and C.1.36.

Lemma C.1.31. *Let $\mu : \mathcal{M} \to [0,\infty]$ be a measure on X. Then for every $S \subset X$ there exists $A \in \mathcal{M}$ such that*

$$S \subset A : \ \mu^*(S) = \mu(A).$$

Consequently,

$$\mu^*(S) = \min\left\{\mu(A) : \ S \subset A \in \mathcal{M}\right\}.$$

Remark C.1.32. An outer measure $\psi : \mathcal{P}(X) \to [0,\infty]$ is called $\mathcal{M}$*-regular* if $\mathcal{M} \subset \mathcal{M}(\psi)$ and

$$\forall S \subset X \ \exists A \in \mathcal{M} : S \subset A, \ \psi(S) = \psi(A);$$

according to Lemma C.1.31, the outer measure μ^* generated by a measure $\mu : \mathcal{M} \to [0,\infty]$ is $\mathcal{M}$-regular.

Proof. If $S \subset X$ then

$$\begin{aligned}
\mu^*(S) &= \inf\left\{\sum_{j=1}^{\infty}\mu(A_j) : \ S \subset \bigcup_{j=1}^{\infty} A_j, \ \{A_j\}_{j=1}^{\infty} \subset \mathcal{M}\right\} \\
&\geq \inf\left\{\mu(\bigcup_{j=1}^{\infty} A_j) : \ S \subset \bigcup_{j=1}^{\infty} A_j, \ \{A_j\}_{j=1}^{\infty} \subset \mathcal{M}\right\} \\
&= \inf\left\{\mu(A) : \ S \subset A, \ A \in \mathcal{M}\right\} \\
&\geq \mu^*(S).
\end{aligned}$$

Thus $\mu^*(S) = \inf\{\mu(A) : \ S \subset A \in \mathcal{M}\}$. For $\varepsilon > 0$, choose $A_\varepsilon \in \mathcal{M}$ such that $S \subset A_\varepsilon$ and $\mu^*(S) + \varepsilon \geq \mu(A_\varepsilon)$. Let $A_0 := \bigcap_{k=1}^{\infty} A_{1/k} \in \mathcal{M}$. Then $S \subset A_0$, and

$$\mu^*(S) \leq \mu(A_0) \leq \mu(A_\varepsilon) \leq \mu^*(S) + \varepsilon$$

implies $\mu^*(S) = \mu(A_0)$. □

Exercise C.1.33. Let $\psi : \mathcal{P}(X) \to [0,\infty]$ be an $\mathcal{M}$-regular outer measure and $E \in \mathcal{M}(\psi)$. Define $\psi_E : \mathcal{P}(X) \to [0,\infty]$ by $\psi_E(S) := \psi(E \cap S)$ as in Exercise C.1.10. Show that ψ_E is an $\mathcal{M}$-regular outer measure.

Exercise C.1.34. Let $(X, \mathcal{M}, \mu)$ is a measure space and $E_k \subset X$ such that $E_k \subset E_{k+1}$ for all $k \in \mathbb{Z}^+$. Show that

$$\mu^*(\bigcup_{k=1}^{\infty} E_k) = \lim_{k\to\infty} \mu^*(E_k).$$

Notice that this does not violate Exercise C.1.23.

Theorem C.1.35 (Carathéodory–Hahn extension). *Let $\mu : \mathcal{M} \to [0, \infty]$ be a measure. Then $\mathcal{M} \subset \mathcal{M}(\mu^*)$ and $\mu = \mu^*|_{\mathcal{M}}$.*

Proof. Let $E \in \mathcal{M}$. Then $\mu^*(E) = \mu(E)$, because trivially $\mu^*(E) \leq \mu(E)$ and because

$$\mu(E) \leq \mu(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \mu(E_j)$$

for any $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ covering E. To prove $\mathcal{M} \subset \mathcal{M}(\mu^*)$, we must show that

$$\mu^*(S) = \mu^*(E \cap S) + \mu^*(E^c \cap S)$$

for any $S \subset X$. This follows, because

$$\begin{aligned} & \mu^*(E \cap S) + \mu^*(E^c \cap S) \\ \geq\ & \mu^*(S) \\ \overset{\text{Lemma } C.1.31}{=}\ & \inf \{\mu(A) :\ S \subset A \in \mathcal{M}\} \\ =\ & \inf \{\mu(A \cap E) + \mu(A \cap E^c) :\ S \subset A \in \mathcal{M}\} \\ \geq\ & \mu^*(E \cap S) + \mu^*(E^c \cap S). \end{aligned}$$

This concludes the proof. □

Theorem C.1.36. *Let $\mu : \mathcal{M} \to [0, \infty]$ be a measure. Then $\mu^* = \left(\mu^*|_{\mathcal{M}(\mu^*)}\right)^*$.*

Proof. Let $\nu := \mu^*|_{\mathcal{M}(\mu^*)}$. We must show that $\nu^* = \mu^*$. Since $\mu = \mu^*|_{\mathcal{M}}$ and $\mathcal{M} \subset \mathcal{M}(\mu^*)$ by Theorem C.1.35, we see that μ is a restriction of ν, and thus the investigation of Definition C.1.2 yields $\nu^* \leq \mu^*$. Moreover,

$$\begin{aligned} \mu^*(S) \geq\ & \nu^*(S) \\ \overset{\text{Lemma } C.1.31}{=}\ & \inf \{\nu(A) :\ S \subset A \in \mathcal{M}(\mu^*)\} \\ \overset{\text{Lemma } C.1.31}{=}\ & \inf \{\mu(B) :\ S \subset A \in \mathcal{M}(\mu^*),\ A \subset B \in \mathcal{M}\} \\ \geq\ & \inf \{\mu(B) :\ S \subset B \in \mathcal{M}\} \\ \geq\ & \mu^*(S), \end{aligned}$$

so that $\mu^*(S) = \nu^*(S)$. □

Remark C.1.37. In the sequel, measures are often required to be complete. This restriction is not severe, as measures can always be completed, e.g. by the Carathéodory–Hahn extension, whose naturality is proclaimed by Theorems C.1.35 and C.1.36: if $(X, \mathcal{M}, \mu)$ is a measure space, $\mathcal{N} = \mathcal{M}(\mu^*)$ and $\nu = \mu^*|_{\mathcal{N}}$, then $(X, \mathcal{N}, \nu)$ is a complete measure space such that $\mathcal{M} \subset \mathcal{N}$ and $\mu = \nu|_{\mathcal{M}}$, with $\mu^* = \nu^*$. So, from this point onwards, we may assume that a measure $\mu : \mathcal{M} \to [0, \infty]$ is already Carathéodory–Hahn complete, i.e. that $\mathcal{M} = \mathcal{M}(\mu^*)$.

C.1.2 Borel regularity

Borel measure are particularly important providing a link with topology on the space. We will study such measures in this section.

Definition C.1.38 (**Borel regular outer measures**). Let (X,τ) be a topological space and $\Sigma(\tau)$ its Borel σ-algebra. An outer measure $\psi : \mathcal{P}(X) \to [0,\infty]$ is *Borel regular* if it is $\Sigma(\tau)$-regular.

Definition C.1.39 (**Metric outer measure**). An outer measure $\psi : \mathcal{P}(X) \to [0,\infty]$ on a metric space (X,d) is called a *metric outer measure* if it satisfies the following *Carathéodory condition*:

$$\operatorname{dist}(A,B) > 0 \quad \Rightarrow \quad \psi(A\cup B) = \psi(A) + \psi(B). \tag{C.4}$$

This condition characterises measurability of Borel sets of a metric space:

Theorem C.1.40. *Let τ_d be the metric topology of a metric space (X,d). An outer measure $\psi : \mathcal{P}(X) \to [0,\infty]$ is a metric outer measure if and only if $\tau_d \subset \mathcal{M}(\psi)$.*

Proof. The "if" part of the proof is left for the reader as Exercise C.1.41. Take $U \in \tau_d$. To show that $U \in \mathcal{M}(\psi)$, we need to prove $\psi(A\cup B) = \psi(A)+\psi(B)$ when $A \subset U$ and $B \subset U^c$. We may assume that $\psi(A), \psi(B) < \infty$. For each $k \in \mathbb{Z}^+$, let

$$A_k := \{x \in A \mid \operatorname{dist}(x, U^c) \geq 1/k\}.$$

Then $\operatorname{dist}(A_k, B) \geq 1/k$, enabling the application of the Carathéodory condition (C.4) in

$$\begin{aligned}\psi(A)+\psi(B) &\overset{\text{trivial}}{\geq} \psi(A\cup B) \\ &\overset{A\supseteq A_k}{\geq} \psi(A_k \cup B) \\ &\overset{(C.4)}{=} \psi(A_k)+\psi(B).\end{aligned}$$

Clearly

$$\psi(A_k) \leq \psi(A) \leq \psi(A_k) + \psi(A\setminus A_k),$$

so we have to show that $\psi(A\setminus A_k) \to 0$. Here $A = \bigcup_{k=1}^{\infty} A_k$, since U is open. Consequently

$$\begin{aligned}\psi(A\setminus A_k) &= \psi\Big(\bigcup_{l=k}^{\infty}(A_{l+1}\setminus A_l)\Big) \\ &\leq \sum_{l=k}^{\infty} \psi(A_{l+1}\setminus A_l) \\ &\overset{(C.4)}{=} \psi\Big(\bigcup_{m=1}^{\infty}(A_{k+2m+1}\setminus A_{k+2m})\Big) + \psi\Big(\bigcup_{m=1}^{\infty}(A_{k+2m}\setminus A_{k+2m-1})\Big) \\ &\leq 2\,\psi(A) \quad < \quad \infty.\end{aligned}$$

Thus $\psi(A \setminus A_k) \leq \sum_{l=k}^{\infty} \psi(A_{l+1} \setminus A_l) \xrightarrow[k\to\infty]{} 0.$ □

Exercise C.1.41. Let (X, d) be a metric space. Complete the proof of Theorem C.1.40 by showing that if $\Sigma(\tau_d) \subset \mathcal{M}(\psi)$ then ψ is a metric outer measure.

Theorem C.1.42 (Topological approximation of measurable sets). *Let (X, d) be a metric space and $\psi : \mathcal{P}(X) \to [0, \infty]$ be a Borel regular outer measure such that $\psi(X) < \infty$. Let $E \subset X$. Then the following statements are equivalent:*

1. *$E \in \mathcal{M}(\psi)$.*
2. *E can be ψ-approximated topologically: more precisely, for each $\varepsilon > 0$ there exist closed $F_\varepsilon \subset X$ and open $G_\varepsilon \subset X$ such that $G_\varepsilon \supset E \supset F_\varepsilon$ and $\psi(G_\varepsilon \setminus F_\varepsilon) < \varepsilon$.*

Proof. Let us assume the second condition. Let $E \subset X$ such that for each $\varepsilon > 0$ there exists a closed set $F_\varepsilon \subset X$ such that

$$F_\varepsilon \subset E \quad \text{and} \quad \psi(E \setminus F_\varepsilon) < \varepsilon.$$

If $F = \bigcup_{k=1}^{\infty} F_{1/k}$ then $E \supset F \in \Sigma(\tau_d) \subset \mathcal{M}(\psi)$, since we assume the measurability of the Borel sets. Moreover, $E \in \mathcal{M}(\psi)$, because

$$E = F \cup (E \setminus F),$$

where $E \setminus F \in \mathcal{M}(\psi)$ due to

$$0 \leq \psi(E \setminus F) \setminus \psi(E \setminus F_{1/k}) < \frac{1}{k} \xrightarrow[k\to\infty]{0} .$$

Thus the second condition of the theorem implies the first one. Notice that here we did not even need the assumption $\psi(X) < \infty$ nor the sets G_ε!

Conversely, we must show that ψ-measurable sets can be ψ-approximated topologically. This can be done by showing that

$$\mathcal{D} := \{A \in \mathcal{M}(\psi) \mid A \text{ can be } \psi\text{-approximated topologically}\}$$

is a σ-algebra containing τ_d; then the Borel regularity will imply $\mathcal{D} = \mathcal{M}(\psi)$. Trivially, $\emptyset \in \mathcal{D}$, and if $A \in \mathcal{D}$ then also $A^c \in \mathcal{D}$. Let $\{A_k\}_{k=1}^{\infty} \subset \mathcal{D}$; now $\mathcal{D}$ is a σ-algebra if $A := \bigcap_{k=1}^{\infty} A_k \in \mathcal{D}$. Clearly, $A \in \mathcal{M}(\psi)$, because each $A_k \in \mathcal{M}(\psi)$. By the topological ψ-approximation, for each $k \in \mathbb{Z}^+$ we can take closed $F_k \subset X$ and open $G_k \subset X$ such that

$$G_k \supset A_k \supset F_k \quad \text{and} \quad \begin{cases} \psi(G_k \setminus A_k) \leq 2^{-k}\varepsilon, \\ \psi(A_k \setminus F_k) \leq 2^{-k}\varepsilon. \end{cases}$$

Then the closed set $\bigcap_{k=1}^\infty F_k$ ψ-approximates the set $\bigcap_{k=1}^\infty A_k$ from inside:

$$\begin{aligned}\psi(\bigcap_{k=1}^\infty A_k \setminus \bigcap_{k=1}^\infty F_k) &\leq \psi(\bigcup_{k=1}^\infty (A_k \setminus F_k)) \\ &\leq \sum_{k=1}^\infty \psi(A_k \setminus F_k) \\ &\leq \varepsilon.\end{aligned}$$

On the other hand, for large enough $n \in \mathbb{Z}^+$, the open set $\bigcap_{k=1}^n G_k$ ψ-approximates the set $\bigcap_{k=1}^\infty A_k$ from outside:

$$\begin{aligned}\psi(\bigcap_{k=1}^n G_k \setminus \bigcap_{k=1}^\infty A_k) &\xrightarrow[n\to\infty]{\psi(X)<\infty} \psi(\bigcap_{k=1}^\infty G_k \setminus \bigcap_{k=1}^\infty A_k) \\ &\leq \quad \cdots \quad \leq \varepsilon.\end{aligned}$$

Thus we have seen that $\mathcal{D}$ is a σ-algebra, so in proving that $\tau_d \subset \mathcal{D}$, it suffices to show that $F \in \mathcal{D}$ when $F \subset X$ is closed. First, $F \in \mathcal{M}(\psi)$, because the Borel sets are measurable. Clearly, the closed set F ψ-approximates itself from inside, as $\psi(F \setminus F) = \psi(\emptyset) = 0$. Let $U_\varepsilon := \bigcup_{x\in F} B_\varepsilon(x)$, where $B_r(x) = \{y \in X : \ d(x,y) < r\}$ is an open ball. Thus $U_\varepsilon \subset X$ is open, $F \subset U_\varepsilon$, and

$$\psi(U_{1/k} \setminus F) \xrightarrow[k\to\infty]{\psi(X)<\infty} \psi(\bigcap_{k=1}^\infty (U_{1/k} \setminus F)) \overset{F \text{ closed}}{=} \psi(\emptyset) = 0.$$

Hence F can be ψ-approximated by open sets U_ε from outside.

Now we know that $\mathcal{D} \supset \Sigma(\tau_d)$ is a σ-algebra. Take $E \in \mathcal{M}(\psi)$. By the Borel regularity, there exist Borel sets $F, G \subset X$ such that

$$\begin{cases} G \supset E, & \psi(G) = \psi(E), \\ F^c \supset E^c, & \psi(F^c) = \psi(E^c). \end{cases}$$

By the topological ψ-approximation, take closed $F_\varepsilon \subset X$ and open $G_\varepsilon \subset X$ such that

$$\begin{aligned}&G_\varepsilon \supset G \supset E \supset F \supset F_\varepsilon, \\ &\psi(G_\varepsilon \setminus G) < \varepsilon, \quad \psi(F \setminus F_\varepsilon) < \varepsilon.\end{aligned}$$

Then

$$\begin{aligned}\psi(G_\varepsilon \setminus E) &\leq \psi(G_\varepsilon \setminus G) + \psi(G \setminus E) &< \varepsilon, \\ \psi(E \setminus F_\varepsilon) &\leq \psi(E \setminus F) + \psi(F \setminus F_\varepsilon) &< \varepsilon,\end{aligned}$$

completing the proof. □

Remark C.1.43. From the proof of Theorem C.1.42 we see that $E \in \mathcal{M}(\psi)$ if and only if $E = B \cup N$, where B is a Borel set and $\psi(N) = 0$.

Exercise C.1.44. In Theorem C.1.42 we assumed that $\psi(X) < \infty$. Prove analogous result assuming that $\psi(B_r(x)) < \infty$ whenever $0 < r < \infty$; prove also that Remark C.1.43 holds in this generalisation (hint: Exercises C.1.10 and C.1.33). Notice that this new assumption is satisfied by the Lebesgue outer measure $\lambda^*_{\mathbb{R}^n}$.

C.1.3 On Lebesgue measure

Recall the Lebesgue outer measure $\lambda_{\mathbb{R}^n} : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ from Definition C.1.5: first, we set the volume of an n-interval

$$I_{ab} := [a, b] = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$$

to be

$$m(I_{ab}) = volume(I_{ab}) = (b_1 - a_1) \cdots (b_n - a_n),$$

and define

$$\lambda^*_{\mathbb{R}^n}(E) = \inf \left\{ \sum_{k=1}^{\infty} m(I_k) : \ E \subset \bigcup_{j=1}^{\infty} I_k, \ I_k \subset \mathbb{R}^n \text{ an } n\text{−interval} \right\}.$$

Definition C.1.45 (Lebesgue measure). The *Lebesgue measure* $\lambda_{\mathbb{R}^n} : \mathcal{M} \to [0, \infty]$ is the restriction of the Lebesgue outer measure $\lambda^*_{\mathbb{R}^n}$ to the σ-algebra $\mathcal{M} = \mathcal{M}(\lambda^*_{\mathbb{R}^n})$.

For a measurable set $E \subset \mathbb{R}^n$, the number $\lambda_{\mathbb{R}^n}(E) \in [0, \infty]$ can be thought as an "n-dimensional volume". Next we try to justify this claim.

Proposition C.1.46. *For any n-interval $I_{ab} \subset \mathbb{R}^n$,*

$$\lambda^*_{\mathbb{R}^n}(I_{ab}) = volume(I_{ab}).$$

Proof. Trivially $\lambda^*_{\mathbb{R}^n}(I_{ab}) \leq volume(I_{ab})$ by the definition of the Lebesgue outer measure. Conversely, take $\varepsilon > 0$. Take a family $\{A_j\}_{j=1}^{\infty}$ of n-intervals such that $I_{ab} \subset \bigcup_{j=1}^{\infty} A_j$ and

$$\sum_{j=1}^{\infty} (A_j) < \lambda^*_{\mathbb{R}^n}(I_{ab}) + \varepsilon.$$

Take a family $\{B_j\}_{j=1}^{\infty}$ of n-intervals such that

$$A_j \subset \mathrm{int}(B_j) \quad \text{and} \quad volume(B_j) \leq volume(A_j) + 2^{-j}\varepsilon.$$

Then $\{\mathrm{int}(B_j)\}_{j=1}^{\infty}$ is an open cover of the compact set $I_{ab} \subset \mathbb{R}^n$, thus having a finite subcover $\{\mathrm{int}(B_{j_k})\}_{k=1}^{l}$. Take a family $\{K_i\}_{i=1}^{m}$ of n-intervals such that $I_{ab} = \bigcup_{i=1}^{m} K_i$, $\{\mathrm{int}(K_i)\}_{i=1}^{m}$ is disjoint and that for each i there exists j_k such that $K_i \subset \text{int}(B_{j_k})$; that is, the idea is to chop the n-interval I_{ab} into small enough n-intervals K_i. Then

$$\begin{aligned}
volume(I_{ab}) &= \sum_{i=1}^{m} volume(K_i) \\
&\leq \sum_{k=1}^{l} volume(B_{j_k}) \\
&\leq \sum_{j=1}^{\infty} volume(B_j) \\
&\leq \sum_{j=1}^{\infty} volume(A_j) + \varepsilon \\
&\leq \lambda^*_{\mathbb{R}^n}(I_{ab}) + 2\varepsilon.
\end{aligned}$$

Hence $volume(I_{ab}) \leq \lambda^*_{\mathbb{R}^n}(I_{ab})$. □

Exercise C.1.47. Let $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ $E \subset \mathbb{R}^n$,

$$\begin{aligned}
tE &:= \{ty \mid y \in E\} \quad \text{and} \\
x + E &:= \{x + y \mid y \in E\}.
\end{aligned}$$

Show that

$$\begin{aligned}
\lambda^*_{\mathbb{R}^n}(tE) &= |t|^n \ \lambda^*_{\mathbb{R}^n}(E), \\
\lambda^*_{\mathbb{R}^n}(x + E) &= \lambda^*_{\mathbb{R}^n}(E).
\end{aligned}$$

Moreover, show that $tE, x + E \in \mathcal{M}$ if $E \in \mathcal{M} = \mathcal{M}(\lambda^*_{\mathbb{R}^n})$.

Remark C.1.48 (**Translation and rotation invariance of Lebesgue measure**). The *translation invariance* of the Lebesgue (outer) measure refers to the invariance under mapping $E \mapsto x + E$ in Exercise C.1.47. The Lebesgue measure behaves well also under linear mappings: for a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^n$ and $E \subset \mathbb{R}^n$, let $AE := \{Ay \mid y \in E\}$. Then

$$\lambda^*_{\mathbb{R}^n}(AE) = |\det(A)| \ \lambda^*_{\mathbb{R}^n}(E),$$

where $\det(A) \in \mathbb{R}$ is the determinant of A. Moreover, $AE \in \mathcal{M}$ if $E \in \mathcal{M} = \mathcal{M}(\lambda^*_{\mathbb{R}^n})$. Especially, the invariance under the orthogonal mappings is called the *rotation invariance.*

Lemma C.1.49. *Let us define the half-space* $E = \{x \in \mathbb{R}^n \mid x_i \geq 0\}$*, where* $i \in \{1, \ldots, n\}$*. Then* $E \in \mathcal{M} = \mathcal{M}(\lambda^*_{\mathbb{R}^n})$*.*

Proof. Clearly, it is sufficient to deal with the case $i = 1$. Take $S \subset \mathbb{R}^n$. Let $\varepsilon > 0$. Take a family $\{A_j\}_{j=1}^n$ of n-intervals such that $S \subset \bigcup_{j=1}^\infty A_j$ and

$$\sum_{j=1}^n volume(A_j) \leq \lambda_{\mathbb{R}^n}^*(S) + \varepsilon.$$

Notice that $E \cap A_j$ and $\overline{E^c} \cap A_j$ are n-intervals, so that

$$\begin{aligned}
\lambda_{\mathbb{R}^n}^*(S) &\leq \lambda_{\mathbb{R}^n}^*(E \cap S) + \lambda_{\mathbb{R}^n}^*(E^c \cap S) \\
&\leq \lambda_{\mathbb{R}^n}^*(E \cap S) + \lambda_{\mathbb{R}^n}^*(\overline{E^c} \cap S) \\
&\leq \lambda_{\mathbb{R}^n}^*(\bigcup_{j=1}^\infty (E \cap A_j)) + \lambda_{\mathbb{R}^n}^*(\bigcup_{j=1}^\infty (\overline{E^c} \cap A_j)) \\
&\leq \sum_{j=1}^\infty \left(volume(E \cap A_j) + volume(\overline{E^c} \cap A_j) \right) \\
&= \sum_{j=1}^\infty volume(A_j) \\
&\leq \lambda_{\mathbb{R}^n}^*(S) + \varepsilon.
\end{aligned}$$

Thus $\lambda_{\mathbb{R}^n}^*(S) = \lambda_{\mathbb{R}^n}^*(E \cap S) + \lambda_{\mathbb{R}^n}^*(E^c \cap S)$. This proves the Lebesgue measurability of the half-space $E \subset \mathbb{R}^n$. □

Corollary C.1.50. *The closed n-interval $[a, b] \subset \mathbb{R}^n$ is Lebesgue measurable, and so is its interior.*

Proof. First,

$$[a, b] = \bigcap_{k=1}^n (\{x \in \mathbb{R}^n : \ a_k \leq x_k\} \cap \{x \in \mathbb{R}^n : \ x_k \leq b_k\}),$$

so it is measurable, as a finite intersection of measurable sets. Finally, if $c = (1, \cdots, 1) \in \mathbb{R}^n$ then the interior

$$\mathrm{int}([a, b]) = \bigcup_{k=1}^\infty [a + c/k, b - c/k].$$

Being a countable union of measurable sets, the interior is measurable. □

Definition C.1.51. For $x \in \mathbb{R}^n$ and $r > 0$, let the *open cube* be

$$\begin{aligned}
Q(x, r) &:= x + (-r, +r)^n \\
&= \{y \in \mathbb{R}^n \mid \forall i \in \{1, \cdots, n\} : \ |x_i - y_i| < r\}.
\end{aligned}$$

This is a Lebesgue measurable set, as it is the interior of the closed n-interval $\overline{Q(x, r)} = [a, b]$, where $a = x - (r, \cdots, r)$ and $b = x + (r, \cdots, r)$.

Corollary C.1.52 (Lebesgue outer measure is Borel regular). *Lebesgue outer measure $\lambda^* : \mathcal{P}(\mathbb{R}^n) \to [0,\infty]$ is Borel regular.*

Proof. Let $U \subset \mathbb{R}^n$ be open. It is easy to check that

$$Q(x, r/\sqrt{n}) \subset B_r(x) \subset Q(x,r).$$

Thus $x \in U$ if and only if $Q(x,r) \subset U$ for some $r > 0$. If $Q(x,r) \subset U$, take $z \in \mathbb{Q}^n \cap B_{r/2}(x)$; then $x \in Q(z, r/2) \subset Q(x,r) \subset U$. Thus

$$U = \bigcup \left\{ Q(z, 1/m) : \ z \in \mathbb{Q}^n, \ m \in \mathbb{Z}^+, Q(z,1/m) \subset U \right\},$$

which is measurable as a countable union of measurable sets. □

Remark C.1.53. It now turns out that Lebesgue measurable sets are nearly open or closed sets:

Theorem C.1.54 (Topological approximation of Lebesgue measurable sets). *Let $E \subset \mathbb{R}^n$. The following three conditions are equivalent:*

1. *$E \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$.*
2. *For every $\varepsilon > 0$ there exists an open set $U \subset \mathbb{R}^n$ such that $E \subset U$ and $\lambda^*_{\mathbb{R}^n}(U \setminus E) < \varepsilon$.*
3. *For every $\varepsilon > 0$ there exists a closed set $S \subset \mathbb{R}^n$ such that $S \subset E$ and $\lambda^*_{\mathbb{R}^n}(E \setminus S) < \varepsilon$.*

Proof. Let us show that the first condition implies the second one. Suppose $E \subset \mathbb{R}^n$ is Lebesgue measurable. Let $\varepsilon > 0$. For a moment, assume that

$$\lambda_{\mathbb{R}^n}(E) < \infty. \tag{C.5}$$

Take a family $\{A_j\}_{j=1}^\infty$ of n-intervals such $E \subset \bigcup_{j=1}^\infty A_j$ and

$$\sum_{j=1}^\infty volume(A_j) < \lambda_{\mathbb{R}^n}(E) + \varepsilon. \tag{C.6}$$

We may think that this is an ε-tight cover of E, and we may loosen it a bit by taking a family $\{B_j\}_{j=1}^\infty$ on n-intervals such that $A_j \subset \mathrm{int}(B_j)$ and

$$\lambda_{\mathbb{R}^n}(B_j) \leq \lambda_{\mathbb{R}^n}(A_j) + 2^{-j}\varepsilon. \tag{C.7}$$

Let $U := \bigcup_{j=1}^\infty \mathrm{int}(B_j)$. Then $U \subset \mathbb{R}^n$ is open, $E \subset U$ and

$$\begin{aligned}
\lambda_{\mathbb{R}^n}(U) &\leq \sum_{j=1}^\infty \lambda_{\mathbb{R}^n}(B_j) \\
&\overset{(C.7)}{\leq} \sum_{j=1}^\infty \lambda_{\mathbb{R}^n}(A_j) + \varepsilon \\
&\overset{(C.6)}{<} \lambda_{\mathbb{R}^n}(E) + 2\varepsilon.
\end{aligned}$$

From this we get (as E, U are measurable and $E \subset U$)

$$\lambda^*_{\mathbb{R}^n}(U \setminus E) = \lambda_{\mathbb{R}^n}(U \setminus E) \overset{(C.5)}{=} \lambda_{\mathbb{R}^n}(U) - \lambda_{\mathbb{R}^n}(E) < 2\varepsilon.$$

Thus the case of (C.5) is completely solved. Now let us forget the restriction (C.5), and let $E_k := E \cap B_d(0,k)$, where d is the Euclidean distance. Then

$$E_k \in \mathcal{M}(\lambda^*_{\mathbb{R}^n}), \quad \lambda_{\mathbb{R}^n}(E_k) < \infty, \quad E = \bigcup_{k=1}^{\infty} E_k.$$

By the earlier part of the proof, for each k there exists an open set $U_k \subset \mathbb{R}^n$ such that $E_k \subset U_k$ and

$$\lambda_{\mathbb{R}^n}(U_k \setminus E_k) < 2^{-k}\varepsilon.$$

Let $U := \bigcup_{k=1}^{\infty} U_k$. Then U is open, $E \subset U$ and

$$U \setminus E = \left(\bigcup_{k=1}^{\infty} U_k \right) \setminus E = \bigcup_{k=1}^{\infty} (U_k \setminus E) \subset \bigcup_{k=1}^{\infty} (U_k \setminus E_k),$$

implying

$$\lambda^*_{\mathbb{R}^n}(U \setminus E) \le \sum_{k=1}^{\infty} \lambda^*_{\mathbb{R}^n}(U_k \setminus E_k) < \sum_{k=1}^{\infty} 2^{-k}\varepsilon = \varepsilon.$$

Thus the first condition in Theorem C.1.54 implies the second one.

Let us now assume the second condition, about approximation by open sets from outside: thereby for each $k \in \mathbb{Z}^+$ there exists an open set $U_k \subset \mathbb{R}^n$ such that $E \subset U_k$ and

$$\lambda^*_{\mathbb{R}^n}(U_k \setminus E) < \frac{1}{k}.$$

Let $G := \bigcap_{k=1}^{\infty} U_k$. Then $E \subset G \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$, and $G \setminus E \subset U_j \setminus E$ for every $j \in \mathbb{Z}^+$. Hence

$$\lambda^*_{\mathbb{R}^n}(G \setminus E) \le \lambda^*_{\mathbb{R}^n}(U_j \setminus E) < \frac{1}{j} \xrightarrow[j \to \infty]{} 0,$$

so that $\lambda^*_{\mathbb{R}^n}(G \setminus E) = 0$. Thus $G \setminus E \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$ by Lemma C.1.11, so that $E = G \setminus (G \setminus E) \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$. This shows that the second condition implies the first one in Theorem C.1.54.

Let us now show that the first and the second conditions imply the third condition. Let $E \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$. Take $\varepsilon > 0$. Since $E^c \in \mathcal{M}(\lambda^*_{\mathbb{R}^n})$, there exists an open set $U \subset \mathbb{R}^n$ such that $E^c \subset U$ and

$$\lambda_{\mathbb{R}^n}(U \setminus E^c) < \varepsilon.$$

Now $S := U^c \subset \mathbb{R}^n$ is closed, $S \subset E$, and $E \setminus S = U \setminus E^c$. This establishes the third condition, about approximation by closed sets from inside.

The rest of the proof is left for the reader as an exercise. Naturally, the reasoning can be made similar to the case where the second condition implied the first one. $\square$

Exercise C.1.55. Complete the proof of Theorem C.1.54 by showing that the third condition implies the first one.

Remark C.1.56 (**Lebesgue is "almost" Borel**). From the proof of Theorem C.1.54 and from a solution to Exercise C.1.55, we may notice that a set $E \subset \mathbb{R}^n$ is Lebesgue measurable if and only if there exist Borel sets $F, G \subset \mathbb{R}^n$ such that $F \subset E \subset G$ and

$$\lambda_{\mathbb{R}^n}(G \setminus F) = 0.$$

Moreover, closer examination reveals that G can be taken as a countable intersection of open sets, and correspondingly F as a countably union of closed sets. In this sense, a Lebesgue measurable set is almost Borel (up to measure zero), and it looks nearly like open from outside, and nearly like closed from inside.

C.1.4 Lebesgue non-measurable sets

The Axiom of Choice (Axiom A.4.2) can be used to "construct" a Lebesgue non-measurable subset $S \subset \mathbb{R}^n$. Let $f : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}^n$ be a choice function. Let

$$S := \{f(x + \mathbb{Q}^n) \mid x \in \mathbb{R}^n\}.$$

Let us show that this set is non-measurable. Now $\lambda^*_{\mathbb{R}^n}(S) > 0$, because $\mathbb{R}^n = \mathbb{Q}^n + S$ is the union of a countable family $\{q + S \mid q \in \mathbb{Q}^n\}$, where $\lambda^*_{\mathbb{R}^n}(r+S) = \lambda^*_{\mathbb{R}^n}(S)$. Moreover, if $0 \neq q \in \mathbb{Q}^n$ then $S \cap (q + S) = \emptyset$. By the following result, this proves the non-measurability of S:

Proposition C.1.57. *Let $S \subset \mathbb{R}^n$ be Lebesgue measurable and $\lambda_{\mathbb{R}^n}(S) > 0$. Then there exists $\delta > 0$ such that $\lambda_{\mathbb{R}^n}(S \cap (x + S)) > 0$ whenever $\|x\|_{\mathbb{R}^n} < \delta$.*

Proof. Let $0 < \varepsilon < 1$. Since $\lambda(S) > 0$, there exists an n-interval $I = [a, b] \subset \mathbb{R}^n$ such that

$$\lambda(S \cap I) = (1 - \varepsilon)\ \lambda(I) > 0.$$

Let $E = S \cap I$. Then $\lambda(I \setminus E) = \lambda(I) - \lambda(E) = \varepsilon\ \lambda(I)$ due to the measurability of E. For any $x \in \mathbb{R}^n$,

$$I \cap (x + I) = (E \cup (x + E)) \cup (I \setminus E) \cup ((x + I) \setminus (x + E)),$$

so that

$$\begin{aligned} \lambda\left(I \cap (x + I)\right) &\leq \lambda\left(E \cap (x + E)\right) + \lambda\left(I \setminus E\right) + \lambda\left((x + I) \setminus (x + E)\right) \\ &= \lambda\left(E \cap (x + E)\right) + 2\varepsilon\ \lambda\left(I\right), \end{aligned}$$

where the last equality follows by the translation invariance of the Lebesgue measure. The reader easily verifies that $\lim_{x\to 0} \lambda(I + (x + I)) = \lambda(I)$. Thus the claim follows if we choose ε small enough. □

Exercise C.1.58. Let $I = [a,b] \subset \mathbb{R}^n$ be an n-interval. Show that

$$\lambda_{\mathbb{R}^n}(I \cap (x+I)) \xrightarrow[\|x\|_{\mathbb{R}^n} \to 0]{} \lambda_{\mathbb{R}^n}(I).$$

Actually, it can be shown that in the Zermelo–Fraenkel set theory without the Axiom of Choice, one cannot prove the existence of Lebesgue non-measurable sets: see [111]. In practice, we do not have to worry about non-Lebesgue-measurability much.

C.2 Measurable functions

In topology, continuous functions were essential; in measure theory, the nice functions are the measurable ones. Before going into details, let us sketch the common framework behind both continuity and measurability. Let us say that $f : X \to Y$ *induces* (or *pulls back*) from a family $\mathcal{B} \subset \mathcal{P}(Y)$ a new family $f^*(\mathcal{B}) \subset \mathcal{P}(X)$ defined by

$$f^*(\mathcal{B}) := \left\{ f^{-1}(B) \subset X \mid B \in \mathcal{B} \right\},$$

and $f : X \to Y$ *co-induces* (or *pushes forward*) from a family $\mathcal{A} \subset \mathcal{P}(X)$ a new family $f_*(\mathcal{A}) \subset \mathcal{P}(Y)$ defined by

$$f_*(\mathcal{A}) := \left\{ B \subset Y \mid f^{-1}(B) \in \mathcal{A} \right\}.$$

Here if $\mathcal{A}, \mathcal{B}$ are topologies (or respectively σ-algebras) then $f_*(\mathcal{A}), f^*(\mathcal{B})$ are also topologies (or respectively σ-algebras), since $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ preserves unions, intersections and complementations.

Exercise C.2.1. Let $\mathcal{A}, \mathcal{B}$ be σ-algebras. Check that $f_*(\mathcal{A}), f^*(\mathcal{B})$ are indeed σ-algebras.

C.2.1 Well-behaving functions

Definition C.2.2 (Measurability). Let $\mathcal{M}_X, \mathcal{M}_Y$ be σ-algebras on X and Y, respectively. A function $f : X \to Y$ is called $(\mathcal{M}_X, \mathcal{M}_Y)$-*measurable* if

$$f^{-1}(V) \in \mathcal{M}_X$$

for every $V \in \mathcal{M}_Y$; that is, if $f^*(\mathcal{M}_Y) \subset \mathcal{M}_X$.

Remark C.2.3. We see that the measurability behaves well in compositions provided that the involved σ-algebras naturally match: if

$$X \xrightarrow[(\mathcal{M},\mathcal{N})-\text{measurable}]{f} Y \xrightarrow[(\mathcal{N},\mathcal{O})-\text{measurable}]{g} Z$$

then $g \circ f : X \to Z$ is $(\mathcal{M}, \mathcal{O})$-measurable. For us, a most important case is $Y = Z = [-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$, for which the canonical σ-algebra will be the collection $\Sigma(\tau_\infty)$ of Borel sets, where $\tau_\infty \subset \mathcal{P}([-\infty, +\infty])$ is the smallest topology for which all the intervals $[a,b] \subset [-\infty, +\infty]$ are closed.

Definition C.2.4 (**Borel/Lebesgue mesurability).** Let $\mathcal{M}$ be a σ-algebra on X, and let τ_X be a topology of X. A function $f : X \to [-\infty, +\infty]$ is called

- $\mathcal{M}$*-measurable* if it is $(\mathcal{M}, \Sigma(\tau_\infty))$-measurable, and
- *Borel measurable* if it is $\Sigma(\tau_X)$-measurable.

A function $f : \mathbb{R}^n \to [-\infty, +\infty]$ is called *Lebesgue measurable* if it is $\mathcal{M}(\lambda^*_{\mathbb{R}^n})$-measurable.

Definition C.2.5. The *characteristic function* $\chi_E : X \to \mathbb{R}$ of a subset $E \subset X$ is defined by

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \in E^c. \end{cases}$$

Notice that χ_E is $\mathcal{M}$-measurable if and only if $E \in \mathcal{M}$.

Definition C.2.6. Let $a \in \mathbb{R}$ and $f, g : X \to [-\infty, +\infty]$. Denote

$$\begin{aligned} \{f > a\} &:= \{x \in X \mid f(x) > a\}, \\ \{f > g\} &:= \{x \in X \mid f(x) > g(x)\}. \end{aligned}$$

In the analogous manner one defines sets

$$\{f < a\},\ \{f \geq a\},\ \{f \leq a\},\ \{f = a\},\ \{f \neq a\},$$

$$\{f < g\},\ \{f \geq g\},\ \{f \leq g\},\ \{f = g\},\ \{f \neq g\},$$

and so on.

Theorem C.2.7. *Let $\mathcal{M}$ be a σ-algebra on X and $f : X \to [-\infty, +\infty]$. Then the following conditions are equivalent:*

1. *f is $\mathcal{M}$-measurable.*
2. *$\{f > a\}$ is measurable for each $a \in \mathbb{R}$.*
3. *$\{f \geq a\}$ is measurable for each $a \in \mathbb{R}$.*
4. *$\{f < a\}$ is measurable for each $a \in \mathbb{R}$.*
5. *$\{f \leq a\}$ is measurable for each $a \in \mathbb{R}$.*

Proof. If f is $\mathcal{M}$-measurable then $\{f > a\} = f^{-1}((a, +\infty]) \in \mathcal{M}$, because $(a, +\infty] \subset [-\infty, +\infty]$ is a Borel set.

Now suppose $\{f > a\} \in \mathcal{M}$ for every $a \in \mathbb{R}$: we have to show that f is $\mathcal{M}$-measurable. We notice that f is $(\mathcal{M}, \mathcal{D})$-measurable, where

$$\mathcal{D} := f_*(\mathcal{M}) = \left\{B \subset [-\infty, +\infty] \mid f^{-1}(B) \in \mathcal{M}\right\}.$$

Furthermore, f is $\mathcal{M}$-measurable, because $\Sigma(\tau_\infty) \subset \mathcal{D}$, because for every $[a, b] \subset [-\infty, +\infty]$ we have

$$\begin{aligned} f^{-1}([a,b]) &= \{f \geq a\} \cap \{f \leq b\} \\ &= \bigcap_{k=1}^{\infty} \{f > a - 1/k\} \cap \{f > b\}^c \quad \in \quad \mathcal{M}; \end{aligned}$$

recall that $\Sigma(\tau_\infty)$ is the smallest σ-algebra containing every interval. Thus f is $\mathcal{M}$-measurable. All the other claims have essentially similar proofs. □

Remark C.2.8. Let $f, g : X \to [-\infty, +\infty]$ be $\mathcal{M}$-measurable. By Theorem C.2.7, then $\{f > g\} \in \mathcal{M}$, because

$$\{f > g\} = \bigcup_{r \in \mathbb{Q}} (\{f > r\} \cap \{g < r\})\,;$$

notice that here the union is countable! Similarly, also

$$\{f \geq g\}, \{f < g\}, \{f \leq g\}, \{f = g\}, \{f \neq g\} \in \mathcal{M}.$$

Example. A continuous function $f : X \to [-\infty, +\infty]$ is Borel measurable, because $\{f \geq a\} \subset X$ is closed for each $a \in \mathbb{R}$. Therefore a continuous function $f : \mathbb{R}^n \to [-\infty, +\infty]$ is Lebesgue measurable, because Borel sets in $\mathbb{R}^n$ are Lebesgue measurable.

Theorem C.2.9. *Let $\lambda \in \mathbb{R}$ and $0 < p < \infty$. Let $f, g : X \to \mathbb{R}$ be $\mathcal{M}$-measurable. Then*

$$\lambda f, f + g, fg, |f|^p, \min(f, g), \max(f, g) : X \to \mathbb{R}$$

are $\mathcal{M}$-measurable. Moreover, if $0 \notin f(X) = \{f(x) : \ x \in X\}$ then $1/f$ is $\mathcal{M}$-measurable.

Proof. The reader may easily show that λf is $\mathcal{M}$-measurable. If $a \in \mathbb{R}$ then

$$\begin{aligned} \{f + g > a\} &= \bigcup_{q \in \mathbb{Q}:\ q > a} \{f + g > q\} \\ &= \bigcup_{r,s \in \mathbb{Q}:\ r+s > a} (\{f > r\} \cap \{g > s\}) \quad \in \quad \mathcal{M}, \end{aligned}$$

showing that $f + g$ is $\mathcal{M}$-measurable. If $a \geq 0$ then

$$\left\{f^2 > a\right\} = \left\{f > \sqrt{a}\right\} \cup \left\{f < -\sqrt{a}\right\},$$

so that f^2 is $\mathcal{M}$-measurable. Thereby also

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}$$

is $\mathcal{M}$-measurable. If $0 \notin f(X)$ then $1/f : X \to \mathbb{R}$ is $\mathcal{M}$-measurable, since it is a composition of

- the $\mathcal{M}$-measurable mapping $(x \mapsto f(x)) : X \to \mathbb{R} \setminus \{0\}$, and
- the continuous mapping $(t \mapsto 1/t) : \mathbb{R} \setminus \{0\} \to \mathbb{R}$.

The rest of the proof is left as Exercise C.2.10. □

Exercise C.2.10. Complete the proof of Theorem C.2.9.

Remark C.2.11. Notice that f^2 can be $\mathcal{M}$-measurable even if f is not: consider e.g. $f = \chi_E - 1/2$, where $E \notin \mathcal{M}$.

Definition C.2.12 (**μ-almost everywhere**). Let $(X, \mathcal{M}, \mu)$ be a complete measure space. We say that a property holds *μ-almost everywhere* (abbreviated *μ-a.e.*) if it holds in a set $N^c = X \setminus N$, where $N \in \mathcal{M}$ and $\mu(N) = 0$.

Theorem C.2.13. *Let $(X, \mathcal{M}, \mu)$ be complete and $f, g : X \to [-\infty, +\infty]$. Let f be $\mathcal{M}$-measurable and $f = g$ μ-a.e. Then g is $\mathcal{M}$-measurable.*

Proof. Let $N := \{f \neq g\} \in \mathcal{M}$. We have to show that $\{g > a\} \in \mathcal{M}$ for any $a \in \mathbb{R}$. Notice that

$$\begin{aligned} \{g > a\} &= (N \cap \{g > a\}) \cup (N^c \cap \{g > a\}) \\ &= (N \cap \{g > a\}) \cup (N^c \cap \{f > a\}). \end{aligned}$$

Clearly, $N^c \cap \{f > a\} \in \mathcal{M}$. Moreover, $N \cap \{g > a\} \in \mathcal{M}$, because μ is complete and $\mu^*(N \cap \{g > a\}) \leq \mu(N) = 0$. □

Definition C.2.14 (**Distinguishing functions?**). Let $(X, \mathcal{M}, \mu)$ be complete. Denote $f \sim_\mu g$, if $f = g$ μ-almost everywhere: we may identify those functions that μ "does not distinguish". Especially, if $f : X \to [-\infty, +\infty]$ such that $\mu(\{|f| = \infty\}) = 0$, we may identify f with $g : X \to \mathbb{R}$ defined by

$$g(x) := \begin{cases} f(x), & \text{when } f(x) \in \mathbb{R}, \\ 0, & \text{otherwise.} \end{cases}$$

C.2.2 Sequences of measurable functions

Theorem C.2.15. *Let $f_j : X \to [-\infty, +\infty]$ be $\mathcal{M}$-measurable for each $j \in \mathbb{Z}^+$. Then*

$$\sup_{j \in \mathbb{Z}^+} f_j, \quad \inf_{j \in \mathbb{Z}^+} f_j, \quad \limsup_{j \to \infty} f_j, \quad \liminf_{j \to \infty} f_j$$

are also $\mathcal{M}$-measurable.

Proof. First,

$$\left\{ \sup_{j \in \mathbb{Z}^+} f_j > a \right\} = \bigcup_{j=1}^{\infty} \{f_j > a\} \in \mathcal{M}.$$

Second, the case of the infimum is handled analogously. Third, these previous cases imply the results for lim sup and lim inf. □

Definition C.2.16 (Convergences). Let $f_j, f : X \to \mathbb{R}$, where $j \in \mathbb{Z}^+$. Let us define various convergences $f_j \to f$ in the following manner: We say that $f_j \to f$ *pointwise* (word "*pointwise*" often omitted) if

$$\forall x \in X : \; |f_j(x) - f(x)| \xrightarrow[j\to\infty]{} 0.$$

Saying that $f_j \to f$ *uniformly* means

$$\sup_{x\in X} : \; |f_j(x) - f(x)| \xrightarrow[j\to\infty]{} 0.$$

Let $(X, \mathcal{M}, \mu)$ be complete, $f_j : X \to \mathbb{R}$ be $\mathcal{M}$-measurable and $f : X \to [-\infty, +\infty]$. We say that $f_j \to f$ *μ-a.e.* if

$$f_j \to f \text{ pointwise } \mu\text{-a.e. on } X.$$

Saying that $f_j \to f$ *μ-almost uniformly* means that

$$\begin{cases} \forall \varepsilon > 0 \; \exists A_\varepsilon \in \mathcal{M} : \; (f_j - f)|_{A_\varepsilon} \xrightarrow[j\to\infty]{} 0 \text{ uniformly}, \\ \mu(A_\varepsilon^c) < \varepsilon. \end{cases}$$

Saying that $f_j \to f$ *in measure μ* means

$$\forall \varepsilon > 0 : \; \mu^* \left(\{ |f_j - f| \geq \varepsilon \} \right) \xrightarrow[j\to\infty]{} 0.$$

Exercise C.2.17. Let functions $f_j : X \to \mathbb{R}$ be $\mathcal{M}$-measurable for every $j \in \mathbb{Z}^+$. Show that $E \in \mathcal{M}$, where

$$E := \left\{ x \in X : \; \lim_{j\to\infty} f_j(x) \in \mathbb{R} \text{ exists} \right\}.$$

Exercise C.2.18. Let (X, τ) be a topological space, $f_j \in C(X)$ for each $j \in \mathbb{Z}^+$ and $f_j \to f$ uniformly. Show that $f : X \to \mathbb{R}$ is also continuous. This extends Theorem A.9.7.

Remark C.2.19. Let $(X, \mathcal{M}, \mu)$ be as above. By Theorems C.2.13 and C.2.15, if $f_j \to f$ μ-a.e. then $f : X \to [-\infty, +\infty]$ is $\mathcal{M}$-measurable. Moreover, if $f_j \to f$ in measure or $f_j \to f$ almost uniformly then f is $\mathcal{M}$-measurable, and $f(x) \in \mathbb{R}$ for μ-a.e. $x \in X$ (by Theorem C.2.24 and Exercise C.2.20, respectively).

Exercise C.2.20. Let $f_j \to f$ μ-almost uniformly.
a) Show that $f_j \to f$ in measure μ.
b) Show that $f_j \to f$ μ-almost everywhere.
These implications cannot be reversed: give examples.

Exercise C.2.21. For each $j \in \mathbb{Z}^+$, let $f_j : X \to \mathbb{R}$ be $\mathcal{M}$-measurable. Let $(f_j)_{j=1}^\infty$ be a *Cauchy sequence in measure* μ, that is

$$\forall \varepsilon > 0 : \quad \mu(\{|f_i - f_j| \geq \varepsilon\}) \xrightarrow[i,j\to\infty]{} 0.$$

Show that there exists $f : X \to [-\infty, +\infty]$ such that $f_j \to f$ in measure μ.

Exercise C.2.22. Let $f_j \to f$ μ-almost everywhere.
a) Show that $f_j \to f$ in measure μ, if $\mu(X) < \infty$.
b) Give an example where $\mu(X) = \infty$ and $f_j \not\to f$ in measure μ;
consequently, here also $f_j \not\to f$ μ-almost uniformly, by Exercise C.2.20.

For **finite** measure spaces the almost everywhere convergence implies the almost uniform convergence:

Theorem C.2.23 (Egorov: "finite pointwise is almost uniform"). *Let $(X, \mathcal{M}, \mu)$ be a complete finite measure space. Let $f_j \to f$ μ-almost everywhere. Then $f_j \to f$ almost uniformly.*

Proof. Take $\varepsilon > 0$. We want to find $A_\varepsilon \in \mathcal{M}$ such that $\mu(A_\varepsilon^c) < \varepsilon$ and $(f_j - f)|_{A_\varepsilon} \xrightarrow[j\to\infty]{} 0$ uniformly. Let

$$E := \{|f_j - f| \to 0\}.$$

Now $E \in \mathcal{M}$ and $\mu(E^c) = 0$, because $f_j \to f$ μ-almost everywhere. Moreover,

$$A_{jk} := \bigcap_{i=j}^\infty \left\{ |f_i - f| < \frac{1}{k} \right\} \in \mathcal{M}.$$

We may choose $j_k \in \mathbb{Z}^+$ such that

$$\mu(A_{j_k k}^c) < 2^{-k}\varepsilon, \tag{C.8}$$

because

$$\begin{aligned} \lim_{j\to\infty} \mu(A_{jk}^c) \quad &\overset{\mu(X)<\infty,\ A_{jk}^c \supset A_{(j+1)k}^c \in \mathcal{M}}{=} \quad \mu(\bigcap_{j=1}^\infty A_{jk}^c) \\ &\overset{E \subset \bigcup_{j=1}^\infty A_{jk}}{\leq} \quad \mu(E^c) \quad = \quad 0. \end{aligned}$$

Now $A_\varepsilon := \bigcap_{k=1}^\infty A_{j_k k} \in \mathcal{M}$ is the desired set:

$$\mu(A_\varepsilon^c) = \mu(\bigcup_{k=1}^\infty A_{j_k k}^c) \leq \sum_{k=1}^\infty \mu(A_{j_k k}^c) \overset{(C.8)}{<} \varepsilon,$$

and $(f_i - f)|_{A_\varepsilon} \to 0$ uniformly, because

$$A_\varepsilon = \bigcap_{k=1}^{\infty} \bigcap_{i=j_k}^{\infty} \left\{ |f_i - f| < \frac{1}{k} \right\},$$

so that $|f_i(x) - f(x)| < \dfrac{1}{k}$ for all $x \in A_\varepsilon$ whenever $i \geq j_k$. □

Theorem C.2.24. *Let $(X, \mathcal{M}, \mu)$ be complete and $f_j \to f$ in measure μ. Then there exists a subsequence $\{f_{j_k}\}_{k=1}^{\infty} \subset \{f_j\}_{j=1}^{\infty}$ such that $f_{j_k} \to f$ μ-almost everywhere.*

Proof. Since $f_j \to f$ in measure, for each $k \in \mathbb{Z}^+$ we may take $j_k \in \mathbb{Z}^+$ such that $j_1 = 1$, $j_{k+1} > j_k$ and

$$\mu^*(\left\{ |f_j - f| \geq \frac{1}{k} \right\}) < 2^k \tag{C.9}$$

whenever $j \geq j_k$. Let $N_k := \left\{ |f_{j_k} - f| \geq \frac{1}{k} \right\}$ and

$$N := \limsup_{k \to \infty} N_k = \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} N_k.$$

Then $\mu^*(N) = 0$ (thus $N \in \mathcal{M}$), because

$$\mu^*(N) \leq \mu^*(\bigcup_{k=j}^{\infty} N_k) \leq \sum_{k=j}^{\infty} \mu^*(N_k) \xrightarrow[j \to \infty]{(C.9)} 0.$$

Now

$$N^c = \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} \left\{ |f_{j_k} - f| < \frac{1}{k} \right\},$$

so that $|f_{j_k}(x) - f(x)| < \dfrac{1}{k}$ for all $x \in N^c$ whenever k is large enough. Thus $f_{j_k}(x) \xrightarrow[k \to \infty]{} f(x)$ for all $x \in N^c$. □

Corollary C.2.25. *If $f_j \to f$ in measure then f is measurable.* □

Exercise C.2.26. Draw a clear diagram about logical implications between different types of convergences $f_j \to f$.

C.2.3 Approximating measurable functions

Definition C.2.27 (Simple functions). A function $f : X \to \mathbb{R}$ is called *simple* if its range $f(X) = \{f(x) : x \in X\} \subset \mathbb{R}$ is finite. Its *normal form* is then

$$f = \sum_{a \in f(X)} a \, \chi_{f^{-1}(\{a\})}.$$

Definition C.2.28 (Positive and negative parts). The *positive and negative parts* of a function $f : X \to [-\infty, +\infty]$ are $f^+, f^- : X \to [0, \infty]$, respectively, where

$$\begin{aligned} f^+(x) &:= \max\{0, f(x)\}, \\ f^- &:= (-f)^+. \end{aligned}$$

Theorem C.2.29. *Let $f : X \to [-\infty, +\infty]$. Then there exist simple functions $f_j : X \to \mathbb{R}$ such that*

$$f_j \to f \text{ pointwise}.$$

Moreover, f_j can be chosen so that

$$\text{if } 0 \leq f \text{ then } 0 \leq f_j \leq f_{j+1} \leq f,$$

$$\text{if } f \text{ bounded then } f_j \to f \text{ uniformly},$$

$$\text{if } f \text{ measurable then } f_j \text{ measurable}.$$

Proof. Since $f = f^+ - f^-$, we may approximate f^+ and f^- separately. Thus assume $f \geq 0$. Define

$$f_i(x) := \begin{cases} \frac{k-1}{2^i}, & \text{when } \frac{k-1}{2^i} \leq f(x) < \frac{k}{2^i} \text{ and } 1 \leq k \leq 2^i i, \\ i, & \text{when } f(x) \geq i. \end{cases} \tag{C.10}$$

We leave the further details for the reader. □

Exercise C.2.30. Check that the functions f_i defined in (C.10) have the desired properties.

Theorem C.2.31 (Luzin: "measurable is almost continuous"). *Let (X, d) be a metric space, $(X, \mathcal{M}, \mu)$ be a complete finite measure space. Let $\tau_d \subset \mathcal{M}$ and $f : X \to \mathbb{R}$ be $\mathcal{M}$-measurable. Then for every $\varepsilon > 0$ there exists a closed set $F_\varepsilon \subset X$ such that $\mu(F_\varepsilon^c) < \varepsilon$ and $f|_{F_\varepsilon} : F_\varepsilon \to \mathbb{R}$ is continuous.*

Remark C.2.32. Notice that if $f : X \to \mathbb{R}$ is continuous and $E \subset X$ then $f|_E : E \to \mathbb{R}$ is continuous; this implication may not be reversible!

Proof. Since $f = f^+ - f^-$, it suffices to assume that $f \geq 0$. For each $i \in \mathbb{Z}^+$, $f(X)$ has a disjoint Borel cover $\{[(j-1)/i, j/i) : j \in \mathbb{Z}^+\}$, and

$$\left\{A_{ij} := f^{-1}([(j-1)/i, j/i))\right\}_{j=1}^\infty \subset \mathcal{M}$$

is a disjoint cover of X. Take closed sets $F_{ij} \subset X$ such that $F_{ij} \subset A_{ij}$ and

$$\mu(A_{ij} \setminus F_{ij}) < 2^{-(i+j)}\varepsilon.$$

Then

$$\begin{aligned}
\lim_{k\to\infty} \mu(X \setminus \bigcup_{j=1}^{k} F_{ij}) &\overset{\mu(X)<\infty}{=} \mu(X \setminus \bigcup_{j=1}^{\infty} F_{ij}) \\
&= \mu(\bigcup_{j=1}^{\infty} (A_{ij} \setminus F_{ij})) \\
&= \sum_{j=1}^{\infty} \mu(A_{ij} \setminus F_{ij}) \\
&< 2^{-i}\varepsilon.
\end{aligned}$$

Thereby let $B_i := \bigcup_{j=1}^{k_i} F_{ij}$, where k_i is so large that

$$\mu(X \setminus B_i) < 2^{-i}\varepsilon.$$

Now $F_\varepsilon := \bigcap_{i=1}^{\infty} B_i \subset X$ is closed, and

$$\mu(F_\varepsilon^c) = \mu(\bigcup_{i=1}^{\infty} B_i^c) \leq \sum_{i=1}^{\infty} \mu(B_i^c) < \varepsilon.$$

Let us define $g_i : B_i \to \mathbb{R}$ such that

$$g_i(x) := j/i, \quad \text{when} \quad x \in F_{ij}.$$

Then $g_i : B_i \to \mathbb{R}$ is continuous, because $g_i|_{F_{ij}}$ is constant and because the closed sets F_{ij} are disjoint. Next, $g_i|_{F_\varepsilon} \to f|_{F_\varepsilon}$ uniformly, because $|f(x) - g_i(x)| < 1/j$ whenever $x \in F_{ij} \subset A_{ij}$; the proof is complete, since continuity is preserved in uniform convergence. □

Exercise C.2.33. Let $(X, \mathcal{M}, \mu)$, where $X = [0,1] \subset \mathbb{R}$ and μ is the restriction of the Lebesgue measure $\lambda_\mathbb{R}$ to X. Consider the characteristic function $f = \chi_{\mathbb{Q}\cap X} : X \to \mathbb{R}$ in the light of Luzin's Theorem C.2.31.

Remark C.2.34 (**Littlewood's principles**). With a pinch of salt, measure theory may be crystallised in *Littlewood's principles*:

1. "A measurable set is almost open"
   (see Topological Approximation Theorem C.1.42).
2. "A measurable function is almost continuous"
   (see Luzin's Theorem C.2.31).
3. "Pointwise convergence is almost uniform convergence"
   (see Egorov's Theorem C.2.23).

C.3 Integration

In this section, let $(X, \mathcal{M}, \mu)$ be a complete measure space. The *μ-integral*

$$\int f \, \mathrm{d}\mu$$

of an $\mathcal{M}$-measurable function $f : X \to [-\infty, +\infty]$ is defined step-by-step:

1. first for a simple non-negative function;
2. then for a non-negative function;
3. finally, the general definition.

Definition C.3.1. Let $s : X \to [0, \infty)$ be a $\mathcal{M}$-measurable simple function. Its *integral* $\int s \, \mathrm{d}\mu \in [0, \infty]$ is defined as

$$\int s \, \mathrm{d}\mu = \int \sum_{a \in s(X)} a \, \chi_{\{s=a\}} \, \mathrm{d}\mu := \sum_{a \in s(X)} a \cdot \mu\left(\{s = a\}\right),$$

with the convention $0 \cdot \infty := 0$. Especially, $\int \chi_E \mathrm{d}\mu = \mu(E)$ for $E \in \mathcal{M}$.

Definition C.3.2. Let $f^+ : X \to [0, \infty]$ be a $\mathcal{M}$-measurable non-negative function. Its *integral* $\int f^+ \mathrm{d}\mu \in [0, \infty]$ is defined as

$$\int f^+ \mathrm{d}\mu := \sup \left\{ \int s \, \mathrm{d}\mu : \ 0 \leq s \leq f^+, \ s \text{ simple measurable} \right\}.$$

Definition C.3.3 (Integral). Let $f : X \to [-\infty, +\infty]$ be a $\mathcal{M}$-measurable function. Its *integral* $\int f \mathrm{d}\mu$ is defined as

$$\int f \, \mathrm{d}\mu := \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu$$

provided that $\int f^+ \mathrm{d}\mu < \infty$ *or* $\int f^- \mathrm{d}\mu < \infty$: we want to avoid a situation $\infty - \infty$ here. If $\int f^+ \mathrm{d}\mu < \infty$ *and* $\int f^- \mathrm{d}\mu < \infty$ then f is called *μ-integrable.* Let $f : X \to \mathbb{C}$ be $\mathcal{M}$-measurable such that $|f| : X \to \mathbb{R}$ is μ-integrable. The μ-integral of f is defined by

$$\int f \, \mathrm{d}\mu := \int \mathrm{Re} f \, \mathrm{d}\mu + \mathrm{i} \int \mathrm{Im} f \, \mathrm{d}\mu,$$

where $\mathrm{Re} f, \mathrm{Im} f : X \to \mathbb{R}$ are the real and imaginary parts of f, respectively. If we want to emphasize the variable in the integration, we may write

$$\int f \, \mathrm{d}\mu = \int f(x) \, \mathrm{d}\mu(x),$$

or even $\int f(x) \, \mathrm{d}x$, if the measure is clear from the context. We shall also use the abbreviation

$$\int_E f \, \mathrm{d}\mu := \int \chi_E f \, \mathrm{d}\mu,$$

where $E \in \mathcal{M}$; this is the *integral of f over $E \subset X$*. The *Lebesgue integral* is the integral with respect to the Lebesgue measure.

C.3.1 Integrating simple non-negative functions

It is simple to integrate simple functions. We leave the details as an exercise for the reader:

Exercise C.3.4. Let $r, s : X \to [0, \infty)$ be $\mathcal{M}$-measurable simple functions and $a \in [0, \infty)$. Show that

$$\int ar \, \mathrm{d}\mu = a \int r \, \mathrm{d}\mu \quad \text{and} \quad \int (r+s) \, \mathrm{d}\mu = \int r \, \mathrm{d}\mu + \int s \, \mathrm{d}\mu.$$

Moreover, if $r \leq s$, show that $\int r \, \mathrm{d}\mu \leq \int s \, \mathrm{d}\mu$.

C.3.2 Integrating non-negative functions

Let us now concentrate on integrating measurable non-negative functions. Recall that we are dealing with a complete measure space $(X, \mathcal{M}, \mu)$.

Exercise C.3.5. Let $S \in \mathcal{M}$ and

$$\mu_S(E) := \mu(E \cap S).$$

Show that $(X, \mathcal{M}, \mu_S)$ is a complete and that

$$\int f \, \mathrm{d}\mu_S = \int_S f \, \mathrm{d}\mu$$

for all $\mathcal{M}$-measurable $f \geq 0$.

As an easy consequence of Exercise C.3.4, for $\mathcal{M}$-measurable functions $f^+, g^+ : X \to [0, \infty]$ and $a \in \mathbb{R}^+$,

$$\begin{aligned} \int af^+ \, \mathrm{d}\mu \quad &= \quad a \int f^+ \, \mathrm{d}\mu, \\ \text{if} \quad f^+ \leq g^+ \quad \text{then} \quad &\int f^+ \, \mathrm{d}\mu \leq \int g^+ \, \mathrm{d}\mu. \end{aligned}$$

These observations will be used frequently. However, it is not evident whether

$$\int (f^+ + g^+) \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu + \int g^+ \, \mathrm{d}\mu.$$

This will soon be obtained as a consequence of the following fundamental result:

Theorem C.3.6 (**Monotone Convergence Theorem (Henri-Léon Lebesgue, Beppo Levi)**)**.** *For each $k \geq 1$, let $f_k : X \to [0, \infty]$ be $\mathcal{M}$-measurable such that $f_k \leq f_{k+1}$. Then*

$$\lim_{k \to \infty} \int f_k \, \mathrm{d}\mu = \int \lim_{k \to \infty} f_k \, \mathrm{d}\mu.$$

The case when the limit $f := \lim_{k\to\infty} f_k$ is integrable was proved by Henri-Léon Lebesgue, with the integrability assumption removed by Beppo Levi. In general, the convergence here is meant in $[0,\infty]$, i.e. the limit may be infinite.

Proof of Theorem C.3.6. The function $f := \lim\limits_{k\to\infty} f_k : X :\to [0,\infty]$ is measurable as a limit of measurable functions. Clearly, $f_k \leq f_{k+1} \leq f$, so the increasing sequence of integrals $\int f_k \mathrm{d}\mu \leq \int f \mathrm{d}\mu$ converges to the limit

$$\lim_{k\to\infty} \int f_k \, \mathrm{d}\mu \leq \int f \, \mathrm{d}\mu.$$

Let $0 < \varepsilon < 1$. Take a simple measurable function s such that $s \leq f$ and

$$\int s \, \mathrm{d}\mu \geq (1-\varepsilon) \int f \, \mathrm{d}\mu.$$

Let $E_k := \{f_k > (1-\varepsilon)s\}$. Since f_k and s are measurable, $E_k \in \mathcal{M}$. Furthermore,

$$\begin{aligned}
\int f_k \, \mathrm{d}\mu \quad & \geq \quad \int (1-\varepsilon)s \; \chi_{E_k} \, \mathrm{d}\mu \\
& = \quad \sum_{a\in s(X)} (1-\varepsilon)a \cdot \mu\left(E_k \cap \{s=a\}\right) \\
& \xrightarrow[k\to\infty]{} \quad (1-\varepsilon) \sum_{a\in s(X)} a \cdot \mu\left(\{s=a\}\right) \\
& = \quad (1-\varepsilon) \int s \, \mathrm{d}\mu \\
& \geq \quad (1-\varepsilon)^2 \int f \, \mathrm{d}\mu,
\end{aligned}$$

where the limit is due to $X = \bigcup_{k=1}^{\infty} E_k$, where $E_k \subset E_{k+1} \in \mathcal{M}$. Thus

$$\lim_{k\to\infty} \int f_k \, \mathrm{d}\mu \geq (1-\varepsilon)^2 \int f \, \mathrm{d}\mu.$$

Taking $\varepsilon \to 0$, the proof is complete. □

Corollary C.3.7. *Let $f, g : X \to [0,\infty]$ be $\mathcal{M}$-measurable. Then*

$$\int (f+g) \; \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu.$$

Proof. Take measurable simple functions $f_k, g_k : X \to [0,\infty)$ such that $f_k \leq f_{k+1}$ and $g_k \leq g_{k+1}$ for each $k \in \mathbb{Z}^+$, and $f_k \to f$ and $g_k \to g$ pointwise. Then $f_k + g_k : X \to [0,\infty)$ is measurable and simple, such that

$$f_k + g_k \leq f_{k+1} + g_{k+1} \xrightarrow[k\to\infty]{} f+g,$$

so that by the Monotone Convergence Theorem C.3.6,

$$\begin{aligned}\int (f+g)\ \mathrm{d}\mu &= \lim_{k\to\infty}\int (f_k+g_k)\ \mathrm{d}\mu \\ &\overset{\text{Exercise C.3.4}}{=} \lim_{k\to\infty}\left(\int f_k\ \mathrm{d}\mu + \int g_k\ \mathrm{d}\mu\right) \\ &= \int f\ \mathrm{d}\mu + \int g\ \mathrm{d}\mu,\end{aligned}$$

establishing the result. □

Corollary C.3.8. *Let $g_j : X \to [0,\infty]$ be $\mathcal{M}$-measurable for each $j \in \mathbb{Z}^+$. Then*

$$\int \sum_{j=1}^{\infty} g_j\ \mathrm{d}\mu = \sum_{j=1}^{\infty}\int g_j\ \mathrm{d}\mu.$$

Proof. For each $k \in \mathbb{Z}^+$, let us define functions $f_k, f : X \to [0,\infty]$ by

$$f_k := \sum_{j=1}^{k} g_j \quad \text{and} \quad f := \lim_{k\to\infty} f_k = \sum_{j=1}^{\infty} g_j.$$

These functions are measurable and $f_k \le f_{k+1} \le f$, so

$$\begin{aligned}\int \lim_{k\to\infty}\sum_{j=1}^{k} g_j\ \mathrm{d}\mu &\overset{\text{Monotone Convergence}}{=} \lim_{k\to\infty}\int \sum_{j=1}^{k} g_j\ \mathrm{d}\mu \\ &\overset{\text{Corollary C.3.7}}{=} \lim_{k\to\infty}\sum_{j=1}^{k}\int g_j\ \mathrm{d}\mu,\end{aligned}$$

completing the proof. □

Exercise C.3.9. Let $f \ge 0$ be $\mathcal{M}$-measurable and $\int f\ \mathrm{d}\mu < \infty$. Prove that

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall A \in \mathcal{M} : \quad \mu(A) < \delta \Rightarrow \int_A f\ \mathrm{d}\mu < \varepsilon.$$

Theorem C.3.10 (Fatou's Lemma). *Let $g_k : X \to [0,\infty]$ be $\mathcal{M}$-measurable for each $k \in \mathbb{Z}^+$. Then*

$$\int \liminf_{k\to\infty} g_k\ \mathrm{d}\mu \le \liminf_{k\to\infty}\int g_k\ \mathrm{d}\mu.$$

Proof. Notice that

$$\liminf_{k\to\infty} g_k = \sup_{k\ge 1}\inf_{j\ge k} g_j.$$

Define $f_k := \inf_{j \geq k} g_j$ for each $k \geq 1$. Now $f_k : X \to [0, \infty]$ is measurable and $f_k \leq f_{k+1}$, so that $\sup_{k \geq 1} f_k = \lim_{k \to \infty} f_k$, and

$$\begin{aligned}
\int \liminf_{k \to \infty} g_k \, \mathrm{d}\mu &= \int \sup_{k \geq 1} f_k \, \mathrm{d}\mu \\
&= \int \lim_{k \to \infty} f_k \, \mathrm{d}\mu \\
&\overset{\text{Monotone Convergence}}{=} \lim_{k \to \infty} \int f_k \, \mathrm{d}\mu \\
&= \liminf_{k \to \infty} \int f_k \, \mathrm{d}\mu \\
&\leq \liminf_{k \to \infty} \int g_k \, \mathrm{d}\mu.
\end{aligned}$$

The proof is complete. □

Exercise C.3.11. Sometimes $\int \liminf_{k \to \infty} g_k \, \mathrm{d}\mu < \liminf_{k \to \infty} \int g_k \, \mathrm{d}\mu$ happens in Fatou's Lemma C.3.10. Find an example.

Exercise C.3.12. Actually, the Monotone Convergence Theorem C.3.6 and Fatou's Lemma C.3.10 are logically equivalent: prove this.

Exercise C.3.13 (Reverse Fatou's lemma). Prove the following *reverse Fatou's lemma*. Let $g_k : X \to [0, \infty]$ be $\mathcal{M}$-measurable for each $k \in \mathbb{Z}^+$. Assume that $g_k \leq g$ for every k, where g is μ-integrable. Then

$$\int \limsup_{k \to \infty} g_k \, \mathrm{d}\mu \geq \limsup_{k \to \infty} \int g_k \, \mathrm{d}\mu.$$

C.3.3 Integration in general

Let $f : X \to [-\infty, +\infty]$ be an $\mathcal{M}$-measurable function. Recall that if

$$I^+ = \int f^+ \, \mathrm{d}\mu < \infty \qquad \textit{or} \qquad I^- = \int f^- \, \mathrm{d}\mu < \infty$$

then the μ-integral f is $\int f \mathrm{d}\mu = I^+ - I^-$. Moreover, if both I^+ *and* I^- are finite, f is called μ-integrable. We shall be interested mainly in μ-integrable functions.

Theorem C.3.14. *Let $a \in \mathbb{R}$ and $f : X \to [-\infty, +\infty]$ be μ-integrable. Then*

$$\int af \, \mathrm{d}\mu = a \int f \, \mathrm{d}\mu.$$

Moreover, if $g : X \to [-\infty, +\infty]$ is μ-integrable such that $f \leq g$, then

$$\int f \, \mathrm{d}\mu \leq \int g \, \mathrm{d}\mu.$$

Especially, $\left|\int f \, \mathrm{d}\mu\right| \leq \int |f| \, \mathrm{d}\mu.$

Exercise C.3.15. Prove Theorem C.3.14.

Exercise C.3.16. Let $E \in \mathcal{M}$ and $|f| \leq g$, where f is $\mathcal{M}$-measurable and g is μ-integrable. Show that f and $f\chi_E$ are μ-integrable.

Exercise C.3.17 (Chebyshev's inequality). Let $0 < a < \infty$, and let $f : X \to [-\infty, +\infty]$ be $\mathcal{M}$-measurable. Prove *Chebyshev's inequality*

$$\mu(\{|f| > a\}) \leq a^{-1} \int |f| \, \mathrm{d}\mu. \tag{C.11}$$

We continue with noticing the short-sightedness of integrals:

Lemma C.3.18. *Let $f, g : X \to [-\infty, +\infty]$ be μ-integrable. Then*

1. *Let $E \in \mathcal{M}$ such that $\mu(E) = 0$. Then $\int_E f \, \mathrm{d}\mu = 0$.*
2. *Let $f = g$ μ-almost everywhere. Then $\int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu$.*
3. *Let $\int |f| \, \mathrm{d}\mu = 0$. Then $f = 0$ μ-almost everywhere.*

Proof. First,

$$\begin{aligned} \int_E f^+ \, \mathrm{d}\mu &= \int f^+ \chi_E \, \mathrm{d}\mu \\ &= \sup \left\{ \int s \, \mathrm{d}\mu : \ s \leq f^+ \chi_E \text{ simple measurable} \right\} \\ &\overset{\mu(E)=0}{=} 0, \end{aligned}$$

proving the first result. Next, let us suppose $f = g$ μ-almost everywhere. Then

$$\begin{aligned} \int f^+ \, \mathrm{d}\mu &= \int \left(f^+ \chi_{\{f=g\}} + f^+ \chi_{\{f \neq g\}} \right) \, \mathrm{d}\mu \\ &\overset{\text{Corollary C.3.7}}{=} \int_{\{f=g\}} f^+ \, \mathrm{d}\mu + \int_{\{f \neq g\}} f^+ \, \mathrm{d}\mu \\ &\overset{\mu(\{f \neq g\})=0}{=} \int_{\{f=g\}} f^+ \, \mathrm{d}\mu, \end{aligned}$$

showing that $\int f^+ \mathrm{d}\mu = \int g^+ \mathrm{d}\mu$, establishing the second result. Finally,

$$\begin{aligned}
\mu(\{f \neq 0\}) &= \mu(\bigcup_{k=1}^{\infty} \{|f| > 1/k\}) \\
&\leq \sum_{k=1}^{\infty} \mu(\{|f| > 1/k\}) \\
&= \sum_{k=1}^{\infty} \int \chi_{\{|f|>1/k\}} \, \mathrm{d}\mu \\
&\leq \sum_{k=1}^{\infty} \int k|f| \, \mathrm{d}\mu \\
&= \sum_{k=1}^{\infty} k \int |f| \, \mathrm{d}\mu
\end{aligned}$$

so that if $\int |f| \mathrm{d}\mu = 0$, then $\mu(\{f \neq 0\}) = 0$. □

Proposition C.3.19. *Let $f : X \to [-\infty, +\infty]$ be μ-integrable. Then $f(x) \in \mathbb{R}$ for μ-almost every $x \in X$.*

Proof. First, $\{f^+ = \infty\} = \bigcap_{k=1}^{\infty} \{f^+ > k\} \in \mathcal{M}$, because f^+ is $\mathcal{M}$-measurable. Thereby

$$\begin{aligned}
\mu\left(\{f^+ = \infty\}\right) &= \frac{1}{k} \int k \cdot \chi_{\{f^+=\infty\}} \, \mathrm{d}\mu \\
&\leq \frac{1}{k} \int f^+ \, \mathrm{d}\mu \quad \xrightarrow[k\to\infty]{} \quad 0,
\end{aligned}$$

so that $\mu(\{f^+ = \infty\}) = 0$. Similarly, $\mu(\{f^- = \infty\}) = 0$. □

Remark C.3.20. By Lemma C.3.18 and Proposition C.3.19, when it comes to integration, we may identify a μ-integrable function $f : X \to [-\infty, +\infty]$ with the function $\tilde{f} : X \to \mathbb{R}$ defined by

$$\tilde{f}(x) = \begin{cases} f(x), & \text{when } f(x) \in \mathbb{R}, \\ 0, & \text{when } |f(x)| = \infty. \end{cases}$$

We shall do this identification without any further notice.

Theorem C.3.21 (Sum is integrable). *Let $f, g : X \to [-\infty, +\infty]$ be μ-integrable. Then $f + g$ is μ-integrable and*

$$\int (f + g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu.$$

Proof. For integrable $f, g : X \to \mathbb{R}$, the function $f + g : X \to \mathbb{R}$ is measurable. Notice that

$$f + g = \begin{cases} (f^+ - f^-) + (g^+ - g^-) \\ (f + g)^+ - (f + g)^-. \end{cases}$$

Since $(f + g)^+ \leq f^+ + g^+$, and $(f + g)^- \leq f^- + g^-$, the integrability of $f + g$ follows. Moreover, $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. By Corollary C.3.7,

$$\int (f + g)^+ \mathrm{d}\mu + \int f^- \mathrm{d}\mu + \int g^- \mathrm{d}\mu = \int (f + g)^- \mathrm{d}\mu + \int f^+ \mathrm{d}\mu + \int g^+ \mathrm{d}\mu,$$

implying

$$\begin{aligned} \int (f + g) \ \mathrm{d}\mu &= \int (f + g)^+ \ \mathrm{d}\mu - \int (f + g)^- \ \mathrm{d}\mu \\ &= \int f^+ \ \mathrm{d}\mu - \int f^- \ \mathrm{d}\mu + \int g^+ \ \mathrm{d}\mu - \int g^- \ \mathrm{d}\mu \\ &= \int f \ \mathrm{d}\mu + \int g \ \mathrm{d}\mu. \end{aligned}$$

The proof for the summation is thus complete. □

Theorem C.3.22 (Lebesgue's Dominated Convergence Theorem). *For each $k \geq 1$, let $f_k : X \to [-\infty, +\infty]$ be measurable and $f_k \xrightarrow[k\to\infty]{} f$ pointwise. Assume that $|f_k| \leq g$ for every $k \geq 1$, where g is μ-integrable. Then*

$$\begin{aligned} \int |f_k - f| \ \mathrm{d}\mu &\xrightarrow[k\to\infty]{} 0, \\ \int f_k \ \mathrm{d}\mu &\xrightarrow[k\to\infty]{} \int f \ \mathrm{d}\mu. \end{aligned}$$

Proof. The functions $f_k, f, |f_k - f|$ are μ-integrable, because they are measurable, g is μ-integrable, $|f_k|, |f| \leq g$ and $|f_k - f| \leq 2g$. For each $k \geq 1$, we define function $g_k := 2g - |f_k - f|$. Then the functions $g_k \geq 0$ satisfy the assumptions of Fatou's Lemma C.3.10, yielding

$$\begin{aligned} \int 2g \ \mathrm{d}\mu &= \int \liminf_{k\to\infty} g_k \ \mathrm{d}\mu \\ &\overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \int g_k \ \mathrm{d}\mu \\ &= \liminf_{k\to\infty} \left(\int 2g \ \mathrm{d}\mu - \int |f_k - f| \ \mathrm{d}\mu \right) \\ &= \int 2g \ \mathrm{d}\mu - \limsup_{k\to\infty} \int |f_k - f| \ \mathrm{d}\mu. \end{aligned}$$

Here we may cancel $\int 2g \, \mathrm{d}\mu \in \mathbb{R}$, getting

$$\limsup_{k\to\infty} \int |f_k - f| \, \mathrm{d}\mu \leq 0,$$

so that $\int |f_k - f| \, \mathrm{d}\mu \xrightarrow[k\to\infty]{} 0$. Finally,

$$\left| \int f_k \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \right| = \left| \int (f_k - f) \, \mathrm{d}\mu \right| \leq \int |f_k - f| \, \mathrm{d}\mu \xrightarrow[k\to\infty]{} 0,$$

which completes the proof. □

Remark C.3.23. It is easy to slightly generalise Lebesgue's Dominated Convergence Theorem C.3.22: the same conclusions hold even if we assume only that $f_k \to f$ *almost everywhere*, and that $|f_k| \leq g$ *almost everywhere*, where g is integrable. This is due to that integrals are not affected if we change values of functions in a set of measure zero.

Exercise C.3.24 (Indispensability of an integrable dominating function). Show that in Theorem C.3.22 it is indispensable to require the μ-integrability of a dominating function g. For this, consider $X = [0,1]$, μ the Lebesgue measure, and the sequence $(f_k)_{k=1}^{\infty}$ with $f_k(x) = k$ for $x \in (0, 1/k]$, and $f_k(x) = 0$ for $x \in (1/k, 1]$. Show that the function $h := \sup_k f_k \geq 0$ is not Lebesgue-integrable on $[0,1]$ (hence no dominating function here can be Lebesgue-integrable). Finally, show that the conclusion of Theorem C.3.22 fails for this sequence $(f_k)_{k=1}^{\infty}$.

Exercise C.3.25 (Fatou–Lebesgue Theorem). Prove the following **Fatou–Lebesgue Theorem**: *Let $(f_k)_{k=1}^{\infty}$ be a sequence of $\mathcal{M}$-measurable functions $f_k : X \to \mathbb{R}$ on a measure space $(X, \mathcal{M}, \mu)$. Assume that $|f_k| \leq g$ for every $k \geq 1$, where g is μ-integrable. Then $\liminf_{k\to\infty} f_k$ and $\limsup_{k\to\infty} f_k$ are μ-integrable and we have*

$$\int \liminf_{k\to\infty} f_k \, \mathrm{d}\mu \leq \liminf_{k\to\infty} \int f_k \, \mathrm{d}\mu \leq \limsup_{k\to\infty} \int f_k \, \mathrm{d}\mu \leq \int \limsup_{k\to\infty} f_k \, \mathrm{d}\mu \, .$$

Proposition C.3.26 (Riemann vs Lebesgue). *Let $f : \mathbb{R} \to \mathbb{R}$ be Riemann-integrable on the closed interval $[a,b] \subset \mathbb{R}$. Then $f\chi_{[a,b]}$ is Lebesgue-integrable and the Riemann- and Lebesgue-integrals coincide:*

$$\int_a^b f(x) \, \mathrm{d}x = \int_{[a,b]} f \, \mathrm{d}\lambda_{\mathbb{R}}.$$

Exercise C.3.27 (Riemann integration). Prove Proposition C.3.26. Recall the definition of the Riemann-integral: Let $g : [a,b] \to \mathbb{R}$ be bounded. A finite sequence $P_n = (x_0, \cdots, x_n)$ is called *partition of* $[a,b]$ if

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,$$

for which the *lower and upper Riemann sums* $L(g, P_n), U(g, P_n)$ are defined by

$$\begin{aligned} U(g, P_n) &= \sum_{k=1}^{n} \left(\sup_{x_{k-1} \leq x < x_k} g(x) \right) (x_k - x_{k-1}), \\ L(g, P_n) &= \sum_{k=1}^{n} \left(\inf_{x_{k-1} \leq x < x_k} g(x) \right) (x_k - x_{k-1}). \end{aligned}$$

Now $L(g) \leq U(g)$, where

$$\begin{cases} U(g) := \inf \left\{ U(g, P) : \ P \text{ is a partition of } [a, b] \right\}, \\ L(g) := \sup \left\{ L(g, P) : \ P \text{ is a partition of } [a, b] \right\}. \end{cases}$$

If $L(g) = U(g)$, we say that g is *Riemann-integrable* with Riemann-integral

$$\int_a^b g(x) \, \mathrm{d}x = L(g).$$

Exercise C.3.28. Prove the following *ϵ-criterion for Riemann integrability*: if for any $\varepsilon > 0$ there is a partition P of $[a, b]$ such that $U(g, P) - L(g, P) < \varepsilon$, then g is Riemann integrable over $[a, b]$.

Consequently, prove that if g is monotonic on $[a, b]$ or if g is continuous on $[a, b]$, then g is Riemann integrable over $[a, b]$.

C.4 Integral as a functional

C.4.1 Lebesgue spaces $L^p(\mu)$

In the sequel, $(X, \mathcal{M}, \mu)$ is a complete measure space. For instance, we may have $\mathcal{M} = \mathcal{M}(\mu^*)$.

Definition C.4.1 (**$L^p(\mu)$-norms**)**.** For $1 \leq p < \infty$, the $L^p(\mu)$-*norm* of an $\mathcal{M}$-measurable function $f : X \to [-\infty, +\infty]$ is

$$\|f\|_{L^p(\mu)} := \left(\int |f|^p \, \mathrm{d}\mu \right)^{1/p},$$

and let

$$\|f\|_{L^\infty(\mu)} := \inf \left\{ M \in [0, \infty] : \ |f| \leq M \ \mu\text{-a.e.} \right\},$$

Here L^p is read "L-p" or "Lebesgue-p". If μ is known from the context, notations $L^p = L^p(X) = L^p(\mu)$ are used: e.g. $L^p(\mathbb{R}^n) = L^p(\lambda_{\mathbb{R}^n})$.

Remark C.4.2. The quantities $\|f\|_{L^p(\mu)}$ are not the norms because the non-degeneracy fails: $\|f\|_{L^p(\mu)} = 0$ does only imply that $f = 0$ μ-a.e. In fact, clearly, $\|f\|_{L^p} \in [0, \infty]$, $\|f\|_{L^p} = 0$ if and only if $f = 0$ μ-almost everywhere, $\|\lambda f\|_{L^p} = |\lambda| \ \|f\|_{L^p}$, and $\|f\|_{L^1(\mu)} = \int |f| \, \mathrm{d}\mu$. Also, in Theorem C.4.5 we will see that $||f+g||_{L^p} \leq ||f||_{L^p} + ||g||_{L^p}$ so that the triangle inequality would be satisfied. Thus, we will modify the construction slightly (identifying functions equal μ-a.e.) in Definition C.4.6 to make them norms, so that the terminology "norm" will be justified.

Definition C.4.3 (**Lebesgue conjugate).** The *Lebesgue conjugate* of $p \in [1,\infty]$ is the number $p' \in [1,\infty]$ defined by

$$\frac{1}{p} + \frac{1}{p'} = 1$$

with the usual convention $1/\infty = 0$.

The converse to the following theorem (the converse of Hölder's inequality) will be shown in Theorem C.4.56.

Theorem C.4.4 (**Hölder's inequality).** *Let $1 \leq p \leq \infty$ and $q = p'$. Let $f, g : X \to [-\infty, +\infty]$ be $\mathcal{M}$-measurable. Then*

$$\|fg\|_{L^1} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Proof. For $p = 1$,

$$\begin{aligned} \|fg\|_{L^1} &= \int |f||g| \, \mathrm{d}\mu \\ &\leq \int |f| \, \mathrm{d}\mu \quad \|g\|_{L^\infty} \\ &= \|f\|_{L^1} \; \|g\|_{L^\infty}; \end{aligned}$$

the proof for $p = \infty$ is symmetric. Finally, let us assume that $1 < p < \infty$. We may assume the non-trivial case $0 < \|f\|_{L^p} < \infty$ and $0 < \|g\|_{L^q} < \infty$. Then

$$\|fg\|_{L^1} = \|f\|_{L^p} \; \|g\|_{L^q} \int ab \, \mathrm{d}\mu,$$

where $a = |f|/\|f\|_{L^p}$ and $b = |g|/\|g\|_{L^q}$. The concavity of the logarithm gives

$$\begin{aligned} \ln(ab) &= \ln(a^p)/p + \ln(b^q)/q \\ &\overset{1/q=1-1/p}{\leq} \ln\left(a^p/p + b^q/q\right), \end{aligned}$$

so $\displaystyle\int ab \, \mathrm{d}\mu \leq \int (a^p/p + b^q/q) \, \mathrm{d}\mu = 1/p + 1/q = 1.$ □

Theorem C.4.5 (**Minkowski's inequality).** *Let $1 \leq p \leq \infty$. Let $f, g : X \to [-\infty, +\infty]$ be $\mathcal{M}$-measurable. Then*

$$\|f+g\|_{L^p} \leq \|f\|_{L^p} + \|f\|_{L^p}. \tag{C.12}$$

Proof. First,

$$\begin{aligned} \|f+g\|_{L^1} &= \int |f+g| \ \mathrm{d}\mu \\ &\le \int |f| \ \mathrm{d}\mu + \int |g| \ \mathrm{d}\mu \\ &= \|f\|_{L^1} + \|g\|_{L^1}. \end{aligned}$$

Also $|f+g| \le |f| + |g|$, so that $|f+g| \le \|f\|_{L^\infty} + \|g\|_{L^\infty}$ almost everywhere, yielding

$$\|f+g\|_{L^\infty} \le \|f\|_{L^\infty} + \|g\|_{L^\infty}.$$

Finally, assume that $1 < p < \infty$. Then

$$\begin{aligned} \|f+g\|_{L^p}^p &= \int |f+g|^p \ \mathrm{d}\mu \\ &\le \int (|f|+|g|) \ |f+g|^{p-1} \ \mathrm{d}\mu \\ &= \left\|f \ |f+g|^{p-1}\right\|_{L^1} + \left\|g \ |f+g|^{p-1} \ \right\|_{L^1} \\ &\overset{\text{Hölder}}{\le} (\|f\|_{L^p} + \|g\|_{L^p}) \ \left\||f+g|^{p-1}\right\|_{L^q}, \end{aligned}$$

where

$$\begin{aligned} \left\||f+g|^{p-1}\right\|_{L^q} &= \left(\int |f+g|^{(p-1)q} \ \mathrm{d}\mu\right)^{1/q} \\ &= \left(\int |f+g|^p \ \mathrm{d}\mu\right)^{(p-1)/p} \\ &= \|f+g\|_{L^p}^{p-1}, \end{aligned}$$

concluding the proof. □

Definition C.4.6 (**$L^p(\mu)$-spaces**)**.** Let $1 \le p \le \infty$ and

$$V^p := \{f : \ \|f\|_{L^p} < \infty\}.$$

Noticing that $\|\lambda f\|_{L^p} = |\lambda| \ \|f\|_{L^p}$ for any scalar λ, and recalling Minkowski's inequality (C.12), we see that

$$(f \mapsto \|f\|_{L^p}) : V^p \to [0, \infty)$$

is a seminorm on the vector space V^p. Let us define an equivalence relation $\sim$ on V^p by

$$f \sim g \iff \|f-g\|_{L^p} = 0;$$

i.e. $f \sim g \iff f = g$ μ-almost everywhere. Let us denote the equivalence classes by

$$[f] := \{g \in V^p : \ f \sim g\}.$$

We obtain the quotient vector space

$$L^p(\mu) \quad := \quad V^p/\sim \quad = \quad \{[f] : \ f \in V^p\}$$

with the usual vector space operations. Moreover,

$$([f] \mapsto \|f\|_{L^p}) : L^p(\mu) \to [0, \infty)$$

is a norm on vector space $L^p(\mu)$. Customarily, $\ell^p(X) := L^p(\mu)$, where μ is the *counting measure*, i.e. $\mu(E)$ is the number of points in the set E.

Remark C.4.7. $f \in [f]$ is a function $X \to [-\infty, +\infty]$, but $[f] \in L^p(\mu)$ is not a function, but an equivalence class of functions. However in practice, to avoid cumbersome notation, one often identifies f and $[f]$, e.g. writing briefly $f \in L^p(\mu)$.

Definition C.4.8 (**Convergence in** $L^p(\mu)$). Let $f \in L^p$ and $\{f_j\}_{j=1}^\infty \subset L^p$. We say that $f_j \to f$ *in* L^p if

$$\|f_j - f\|_{L^p} \xrightarrow[j\to\infty]{} 0.$$

Theorem C.4.9. *$L^p(\mu)$ is a Banach space.*

Proof. The case $p = \infty$ is left as Exercise C.4.10; let us consider the case $1 \leq p < \infty$. We already know that $L^p(\mu)$ is a normed space. Given a Cauchy sequence $(f_j)_{j=1}^\infty$ in $L^p(\mu)$, we need a candidate f for the limit of this sequence. Now $(f_j)_{j=1}^\infty$ is a Cauchy sequence in measure μ, because

$$\begin{aligned}
\mu(\{|f_i - f_j| \geq \varepsilon\}) &= \mu(\{|f_i - f_j|^p \geq \varepsilon^p\}) \\
&\overset{\text{Chebyshev (C.11)}}{\leq} \varepsilon^{-p} \int |f_i - f_j|^p \,\mathrm{d}\mu \\
&= \varepsilon^{-p} \|f_i - f_j\|_{L^p}^p \\
&\xrightarrow[i,j\to\infty]{} 0.
\end{aligned}$$

Hence by Exercise C.2.21, $f_j \to f$ in measure μ for an $\mathcal{M}$-measurable function f. By Theorem C.2.24, $f_{j_k} \to f$ μ-almost everywhere for a subsequence $(f_{j_k})_{k=1}^\infty$ of $(f_j)_{j=1}^\infty$. Here $f \in L^p(\mu)$, because

$$\begin{aligned}
\|f\|_{L^p}^p &= \int |f|^p \,\mathrm{d}\mu \\
&= \int \liminf_{k\to\infty} |f_{j_k}|^p \,\mathrm{d}\mu \\
&\overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \int |f_{j_k}|^p \,\mathrm{d}\mu \\
&\leq \textit{constant} < \infty,
\end{aligned}$$

because Cauchy sequences in a normed space are bounded. Finally, $f_i \to f$ in $L^p(\mu)$, because

$$\begin{aligned} \|f_i - f\|_{L^p}^p &= \int |f_i - f|^p \, \mathrm{d}\mu \\ &= \int \liminf_{k\to\infty} |f_i - f_{j_k}|^p \, \mathrm{d}\mu \\ &\overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \|f_i - f_{j_k}\|_{L^p}^p \\ &\xrightarrow[i\to\infty]{(f_i)_{i=1}^\infty \text{ Cauchy}} 0. \end{aligned}$$

Thus $L^p(\mu)$ is a Banach space for $1 \leq p < \infty$. □

Exercise C.4.10. Complete the previous proof by showing that $L^\infty(\mu)$ is a Banach space.

Exercise C.4.11. Let $1 \leq p < \infty$ and $\|f_j - f\|_{L^p} \to 0$, where $f \in L^p$ and $\{f_j\}_{j=1}^\infty \subset L^p$. Show that

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall j \in \mathbb{Z}^+ : \quad \mu(E) < \delta \Longrightarrow \int_E |f_j|^p \, \mathrm{d}\mu < \varepsilon.$$

Why $p \neq \infty$ here?

Lemma C.4.12. *Let $g \in L^p(\mu)$, where $1 \leq p < \infty$. Show that*

$$\forall \varepsilon > 0 \ \exists E_g \in \mathcal{M} : \quad \mu(E_g) < \infty \quad \text{and} \quad \int_{E_g^c} |g|^p \, \mathrm{d}\mu < \varepsilon.$$

Exercise C.4.13. Prove Lemma C.4.12.

Theorem C.4.14 (**Vitali's Convergence Theorem**)**.** *Let $1 \leq p < \infty$. Let $f, f_j \in L^p(\mu)$ for each $j \in Z^+$. Then properties* (1,2,3) *imply* (0)*, and* (0) *implies properties* (2,3)*:*

(0) $f_j \to f$ *in L^p.*
(1) $f_j \to f$ *μ-almost everywhere.*
(2) $\forall \varepsilon > 0 \ \exists E \in \mathcal{M} \ \forall j \in \mathbb{Z}^+ : \ \mu(E) < \infty, \ \int_{E^c} |f_j|^p \, \mathrm{d}\mu < \varepsilon.$
(3) $\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall j \in \mathbb{Z}^+ \ \forall A \in \mathcal{M} : \ \mu(A) < \delta \ \Rightarrow \ \int_A |f_j|^p \, \mathrm{d}\mu < \varepsilon.$

Proof. First, let us show that $(1,2,3)$ implies (0). Take $\varepsilon > 0$. Take $\delta > 0$ as in (3). Take $E \in \mathcal{M}$ as in (2). Exploiting (1), Egorov's Theorem C.2.23 says that $(f_j - f)|_E \to 0$ μ-almost uniformly. Hence there exists $B \in \mathcal{M}$ such that

$$\begin{cases} B \subset E, \\ \mu(E \setminus B) < \delta, \\ (f_j - f)|_B \to 0 \quad \text{uniformly}. \end{cases} \tag{C.13}$$

We want to show that $\|f_j \to f\|_{L^p} \to 0$:

$$\begin{aligned} \|f_j - f\|_{L^p}^p &= \int |f_j - f|^p \, \mathrm{d}\mu \\ &= \int_B |f_j - f|^p \, \mathrm{d}\mu + \int_{B^c} |f_j - f|^p \, \mathrm{d}\mu, \end{aligned}$$

and here the integral over B tends to 0 as $j \to \infty$, by (C.13). What about the integral over B^c? Since $(t \mapsto t^p) : \mathbb{R}^+ \to \mathbb{R}$ is a convex function, we have $(a/2 + b/2)^p \leq a^p/2 + b^p/2$, so that

$$\begin{aligned} & \int_{B^c} |f_j - f|^p \, \mathrm{d}\mu \\ \leq \quad & \int_{B^c} 2^{p-1} \left(|f_j|^p + |f|^p \right) \, \mathrm{d}\mu \\ = \quad & 2^{p-1} \left(\int_{E^c} |f_j|^p \, \mathrm{d}\mu + \int_{E \setminus B} |f|^p \, \mathrm{d}\mu + \int_{B^c} \liminf_{j \to \infty} |f_j|^p \, \mathrm{d}\mu \right) \\ \overset{(2),\ (3),\ \text{Fatou}}{<} \quad & 2^{p-1} \left(\varepsilon + \varepsilon + 2\varepsilon \right); \end{aligned}$$

thus $\|f_j - f\|_{L^p} \to 0$: we have proven that (0) follows from $(1, 2, 3)$.

Implication $(0) \Rightarrow (3)$ is left as Exercise C.4.15.

Let us show that $(0) \Rightarrow (2)$. Let $f_j \to f$ in $L^p(\mu)$. Take $\varepsilon > 0$. Take $j_\varepsilon \in \mathbb{Z}^+$ such that $\|f_j - f\|_{L^p} < \varepsilon^{1/p}$ whenever $j > j_\varepsilon$. Take $E_f, E_{f_j} \in \mathcal{M}$ as in Lemma C.4.12. Let

$$E := E_f \cup \bigcup_{j=1}^{j_\varepsilon} E_{f_j}.$$

Then $E \in \mathcal{M}$ and $\mu(E) < \infty$. If $j \leq j_\varepsilon$ then

$$\int_{E^c} |f_j|^p \, \mathrm{d}\mu \leq \int_{E_{f_j}^c} |f_j|^p \, \mathrm{d}\mu < \varepsilon.$$

If $j > j_\varepsilon$ then

$$\begin{aligned} \int_{E^c} |f_j|^p \, \mathrm{d}\mu &\overset{\text{Minkowski}}{\leq} \left(\|\chi_{E^c}(f_j - f)\|_{L^p} + \|\chi_{E^c} f\|_{L^p} \right)^p \\ &\leq \left(\varepsilon^{1/p} + \varepsilon^{1/p} \right)^p, \end{aligned}$$

so that $\int_{E^c} |f_j|^p \, \mathrm{d}\mu \leq 2^p \varepsilon$ for every $j \in \mathbb{Z}^+$. We have shown that $(0) \Rightarrow (2)$.

Finally, let us prove that (0) $\Rightarrow$ (1). We have

$$\begin{aligned}\mu(\{|f_j - f| \geq \varepsilon\}) &= \mu(\{|f_j - f|^p \geq \varepsilon^p\}) \\ &\overset{\text{Chebyshev}}{\leq} \varepsilon^{-p} \int |f_j - f|^p \, \mathrm{d}\mu \\ &\xrightarrow[j\to\infty]{(0)} 0,\end{aligned}$$

so that $f_j \to f$ in measure μ. By Theorem C.2.24, there is a subsequence $(f_{j_k})_{k=1}^{\infty}$ such that $f_{j_k} \to f$ μ-almost everywhere. We have shown that (0) $\Rightarrow$ (1). □

Exercise C.4.15. Complete the proof of Vitali's Convergence Theorem C.4.14 by showing that (0) $\Rightarrow$ (3).

Exercise C.4.16. Let $1 \leq p \leq \infty$ and $f_j \to f$ μ-a.e., where $\{f_j\}_{j=1}^{\infty} \subset L^p$.
(a) Let $f_j \to g$ in L^p. Show that $f = g$ μ-a.e.
(b) Give an example where $f \in L^p$, but $f_j \not\to f$ in L^p.

Finally, we give without proof a very useful interpolation theorem. But first we introduce

Definition C.4.17 (Semifinite measures). A measure μ is called *semifinite* if for every $E \in \mathcal{M}$ with $\mu(E) = \infty$ there exists $F \in \mathcal{M}$ such that $F \subset E$ and $0 < \mu(F) < \infty$.

Theorem C.4.18 (M. Riesz–Thorin interpolation theorem). *Let μ, ν be semifinite measures and let $1 \leq p_0, p_1, q_0, q_1 \leq \infty$. For every $0 < t < 1$ define p_t and q_t by*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \qquad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

Assume that A is a linear operator such that

$$||Af||_{L^{q_0}(\nu)} \leq C_0 ||f||_{L^{p_0}(\mu)}, \quad ||Af||_{L^{q_1}(\nu)} \leq C_1 ||f||_{L^{p_1}(\mu)},$$

for all $f \in L^{p_0}(\mu)$ and $f \in L^{p_1}(\mu)$, respectively. Then for all $0 < t < 1$, the operator A extends to a bounded linear operator from $L^{p_t}(\mu)$ to $L^{q_t}(\nu)$ and we have

$$||Af||_{L^{q_t}(\nu)} \leq C_0^{1-t} C_1^t ||f||_{L^{p_t}(\mu)}$$

for all $f \in L^{p_t}(\mu)$.

C.4.2 Signed measures

Definition C.4.19 (Signed measures). Let $\mathcal{M}$ be a σ-algebra on X. A mapping $\nu : \mathcal{M} \to \mathbb{R}$ is called a *signed measure* on X if

$$\nu\Big(\bigcup_{j=1}^{\infty} E_j\Big) = \sum_{j=1}^{\infty} \nu(E_j)$$

for any disjoint countable family $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$.

Example. Let $\mu, \nu : \mathcal{M} \to [0, \infty]$ be *finite* measures on X, that is $\mu(X) < \infty$ and $\nu(X) < \infty$. Then

$$\mu - \nu : \mathcal{M} \to \mathbb{R}$$

is a signed measure. It will turn out that there are no other types of signed measures on X, see the Jordan decomposition result in Corollary C.4.26

Remark C.4.20. For simplicity and in view of the planned applications of this notion we restrict the exposition to what may be called finite signed measures. In principle, one can allow $\nu : \mathcal{M} \to [-\infty, +\infty]$ assuming that only one of infinities may be achieved. The statements and the proofs remain largely similar, so we may leave this case as an exercise for an interested reader. For example, only one of the measures in Theorem C.4.25 would be finite, etc.

Exercise C.4.21. Let $(X, \mathcal{M}, \mu)$ be a measure space and let $f : X \to [-\infty, +\infty]$ be μ-integrable. Define $\nu : \mathcal{M} \to \mathbb{R}$ by

$$\nu(E) := \int_E f \, \mathrm{d}\mu. \tag{C.14}$$

Show that ν is a signed measure. Moreover, prove that ν is a (finite) measure if and only if $f \geq 0$ μ-almost everywhere.

Definition C.4.22 (Variations of measures). Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure. Define mappings $\nu^+, \nu^-, |\nu| : \mathcal{M} \to [0, \infty]$ by

$$\begin{aligned} \nu^+(E) &:= \sup_{A \in \mathcal{M}:\ A \subset E} \nu(A), \\ \nu^- &:= (-\nu)^+, \\ |\nu| &:= \nu^+ + \nu^-. \end{aligned}$$

The mappings ν^+, ν^- are called the *positive and negative variations* (respectively) of ν, and the pair (ν^+, ν^-) is the *Jordan decomposition* of ν. The mapping $|\nu|$ is the *total variation* of ν.

Exercise C.4.23. Show that $\nu^+, \nu^-, |\nu| : \mathcal{M} \to [0, \infty]$ are measures.

Exercise C.4.24. Let $\nu(E) = \int_E f \, \mathrm{d}\mu$ as in (C.14). Show that

$$\nu^+(E) = \int_E f^+ \, \mathrm{d}\mu \quad \text{and} \quad \nu^-(E) = \int_E f^- \, \mathrm{d}\mu.$$

Hence here $\nu = \nu^+ - \nu^-$, but this happens even generally:

Theorem C.4.25. *Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure. Then the measures $\nu^+, \nu^- : \mathcal{M} \to [0, \infty]$ are finite.*

Proof. By Exercise C.4.23, ν^+ and ν^- are measures. Let us show that ν^+ (and similarly ν^-) is finite. To get a contradiction, assume that $\nu^+(X) = \infty$. Take $E_0 \in \mathcal{M}$ such that $\nu(E_0) \geq 0$. Take $A_0 \in \{E_0, X \setminus E_0\}$ such that $\nu^+(A_0) = \infty$. For $k \in \mathbb{Z}^+$, suppose $E_k, A_k \in \mathcal{M}$ have been chosen so that $\nu^+(A_k) = \infty$. Take $E_{k+1} \in \mathcal{M}$ such that

$$E_{k+1} \subset A_k \quad \text{and} \quad \nu(E_{k+1}) \geq 1 + \nu(E_k).$$

Take $A_{k+1} \in \{E_{k+1}, A_k \setminus E_{k+1}\}$ such that $\nu^+(A_{k+1}) = \infty$. Then

1. either $\quad \exists k_0\ \forall k \geq k_0 :\ A_{k+1} = E_{k+1}$
2. or $\quad \forall k_0\ \exists k \geq k_0 :\ A_{k+1} = A_k \setminus E_{k+1}$.

Here in the first case, $E \supset E_k \supset E_{k+1}$ for every $k \geq k_0$, and

$$\begin{aligned} \nu(E_{k_0}) &= \nu(\bigcap_{k=k_0}^{\infty} E_k) + \sum_{k=k_0}^{\infty} \nu(E_k \setminus E_{k+1}) \\ &= \nu(\bigcap_{k=k_0}^{\infty} E_k) + \sum_{k=k_0}^{\infty} (\nu(E_k) - \nu(E_{k+1})) \\ &= -\infty; \end{aligned}$$

of course, this is a contradiction, excluding the first case. In the second case, take a disjoint family $\{E_{k_j}\}_{j=1}^{\infty}$ where $k_{j+1} > k_j \in \mathbb{Z}^+$, so that

$$\nu(\bigcup_{j=1}^{\infty} E_{k_j}) = \sum_{j=1}^{\infty} \nu(E_{k_j}) = +\infty,$$

again a contradiction; therefore ν^+ and ν^- must be finite measures. □

Corollary C.4.26 (**Jordan Decomposition**). *Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure. Then*

$$\nu = \nu^+ - \nu^-.$$

Proof. Let $E \in \mathcal{M}$. For any $A \in \mathcal{M}$ we have

$$\begin{aligned} \nu(E) &= \nu(A \cap E) + \nu(A^c \cap E) \\ &\leq \nu^+(E) - (-\nu)(A^c \cap E), \end{aligned}$$

yielding $\nu(E) \leq \nu^+(E) - \nu^-(E)$. Similarly,

$$(-\nu)(E) \leq (-\nu)^+(E) - (-\nu)^-(E) = \nu^-(E) - \nu^+(E),$$

so that $\nu(E) \geq \nu^+(E) - \nu^-(E)$. □

Exercise C.4.27. Let $\mathcal{M}$ be a σ-algebra on X. Let $M(\mathcal{M})$ be the real vector space of all signed measures $\nu : \mathcal{M} \to \mathbb{R}$. For $\nu \in M(\mathcal{M})$, let $\|\nu\| = |\nu|(X)$; show that this gives a Banach space norm on $M(\mathcal{M})$.

Definition C.4.28 (Hahn decomposition). A pair (P, P^c) is called a *Hahn decomposition* of a signed measure $\nu : \mathcal{M} \to \mathbb{R}$ if $P \in \mathcal{M}$ and

$$\forall E \in \mathcal{M} : \quad \nu(P \cap E) \geq 0 \geq \nu(P^c \cap E).$$

Then P is called a *ν-positive set* and P^c is a *ν-negative set.*

Example. Let $\nu(E) = \int_E f \, \mathrm{d}\mu$ as in (C.14). Then (P, P^c) and (Q, Q^c) are Hahn decompositions of $\nu : \mathcal{M} \to \mathbb{R}$, where

$$P := \{f \geq 0\}, \qquad Q := \{f > 0\}.$$

Definition C.4.29 (Mutually singular measures). The measures $\mu, \lambda : \mathcal{M} \to [0, \infty]$ are *mutually singular*, denoted by $\mu \perp \lambda$, if there exists $P \in \mathcal{M}$ such that

$$\mu(P) = 0 = \lambda(P^c).$$

Here, the zero-measure condition $\mu(P) = 0$ can be interpreted so that the measure μ *does not see* the set $P \in \mathcal{M}$.

Theorem C.4.30 (Hahn Decomposition). *Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure. Then ν has a Hahn decomposition (P, P^c). More precisely,*

$$\begin{cases} \nu^+(E) = +\nu(P \cap E), \\ \nu^-(E) = -\nu(P^c \cap E) \end{cases}$$

for each $E \in \mathcal{M}$. Especially, $\nu^- \perp \nu^+$ such that $\nu^-(P) = 0 = \nu^+(P^c)$.

Proof. For each $k \in \mathbb{Z}^+$, take $A_k \in \mathcal{M}$ such that $\nu^+(X) - \nu(A_k) < 2^{-k}$. Then

$$P := \limsup_{k \to \infty} A_k = \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k \in \mathcal{M}.$$

Moreover,

$$\begin{aligned} \nu^-(P) \quad & \leq \quad \sum_{k=j}^{\infty} \nu^-(A_k) \\ & \overset{\text{Corollary C.4.26}}{=} \quad \sum_{k=j}^{\infty} \left(\nu^+(A_k) - \nu(A_k) \right) \\ & \leq \quad \sum_{k=j}^{\infty} 2^{-k} \quad = \quad 2^{1-j}, \end{aligned}$$

so that $\nu^-(P) = 0$. On the other hand,

$$\begin{aligned}
\nu^+(P^c) &= \nu^+(\liminf_{k\to\infty} A_k^c) \\
&= \lim_{j\to\infty} \nu^+(\bigcap_{k=j}^{\infty} A_k^c) \\
&\leq \lim_{j\to\infty} \nu^+(A_j^c) \\
&\leq \lim_{j\to\infty} 2^{-j} \quad = \quad 0,
\end{aligned}$$

so that $\nu^+(P^c) = 0$. Thereby

$$\begin{aligned}
\nu(P\cap E) &\overset{\text{Jordan}}{=} \nu^+(P\cap E) - \nu^-(P\cap E) \\
&= \nu^+(P\cap E) + \nu^+(P^c\cap E) \\
&= \nu^+(E),
\end{aligned}$$

and similarly $\nu(P^c \cap E) = -\nu^-(E)$. □

Exercise C.4.31. Let (P, P^c) and (Q, Q^c) be two Hahn decompositions of a signed measure ν. Show that $|\nu|(P \setminus Q) = 0$. The moral here is that all the Hahn decompositions are "essentially the same".

Exercise C.4.32. Let $\nu = \alpha - \beta$, where $\alpha, \beta : \mathcal{M} \to [0,\infty]$ are finite measures and $\alpha \perp \beta$. Show that

$$\alpha = \nu^+ \quad \text{and} \quad \beta = \nu^-.$$

In this respect the Jordan decomposition is the most natural decomposition of ν as a difference of two measures.

C.4.3 Derivatives of signed measures

In this section we study which signed measures $\nu : \mathcal{M} \to \mathbb{R}$ can be written in the integral form as in (C.15). The key property is the *absolute continuity of ν with respect to μ*, and the key result is the Radon–Nikodym Theorem C.4.38.

Definition C.4.33 (Radon–Nikodym derivative). Let $(X, \mathcal{M}, \mu)$ be a measure space and $f : X \to [-\infty, +\infty]$ be μ-integrable. Let a signed measure $\nu : \mathcal{M} \to \mathbb{R}$ be defined by

$$\nu(E) := \int_E f \, \mathrm{d}\mu. \tag{C.15}$$

Then $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} := f$ is the *Radon–Nikodym derivative* of ν with respect to μ.

Remark C.4.34. Actually, the Radon–Nikodym derivative $\mathrm{d}\nu/\mathrm{d}\mu = f$ is not an integrable function $X \to [-\infty, +\infty]$ but actually the equivalence class

$$\{g : X \to [-\infty, +\infty] \mid f \sim g\},$$

where $f \sim g \iff f = g$ μ-almost everywhere. The classical derivative of a function (a limit of a difference quotient) is connected to the Radon–Nikodym derivative in the case of the Lebesgue measure $\mu = \lambda_{\mathbb{R}}$, but this shall not be investigated here.

Definition C.4.35 (Absolutely continuous measures). A signed measure $\nu : \mathcal{M} \to \mathbb{R}$ is *absolutely continuous* with respect to a measure $\mu : \mathcal{M} \to [0, \infty]$, denoted by $\nu \ll \mu$, if

$$\forall E \subset \mathcal{M} : \quad \mu(E) = 0 \ \Rightarrow \ \nu(E) = 0.$$

Example. If $\nu(E) = \displaystyle\int_E f \, \mathrm{d}\mu$ as in (C.15) then $\nu \ll \mu$.

The following (ε, δ)-result justifies the term *absolute continuity* here:

Theorem C.4.36. *Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure and $\mu : \mathcal{M} \to [0, \infty]$ a measure. Then the following conditions are equivalent:*
(a) $\nu \ll \mu$.
(b) $\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall E \in \mathcal{M} : \ \mu(E) < \delta \ \Rightarrow \ |\nu(E)| < \varepsilon$.

Proof. The (ε, δ)-condition trivially implies $\nu \ll \mu$. On the other hand, let us show that $\nu \not\ll \mu$, when we assume

$$\exists \varepsilon > 0 \ \forall \delta > 0 \ \exists E_\delta \in \mathcal{M} : \quad \mu(E_\delta) < \delta \quad \text{and} \quad |\nu(E_\delta)| \geq \varepsilon.$$

Then

$$E := \limsup_{k \to \infty} E_{2^{-k}} = \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} E_{2^{-k}} \in \mathcal{M},$$

and $\mu(E) = 0$, because

$$\mu(E) \leq \mu(\bigcup_{k=j}^{\infty} E_{2^{-k}}) \leq \sum_{k=j}^{\infty} 2^{-k} = 2^{1-j}.$$

Now $|\nu|(E) > 0$, because

$$|\nu|(E) = \lim_{j \to \infty} |\nu|(\bigcup_{k=j}^{\infty} E_{2^{-k}}) \geq \varepsilon.$$

Hence $\nu^+(E) > 0$ or $\nu^-(E) > 0$, so that $|\nu(A)| > 0$ for some $A \subset E$, where $A \in \mathcal{M}$. Here $\mu(A) = 0$, so that $\nu \not\ll \mu$. □

Exercise C.4.37. Show that the following conditions are equivalent:

1. $\nu \ll \mu$.
2. $|\nu| \ll \mu$.
3. $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

Theorem C.4.38 (Radon–Nikodym). *Let $\mu : \mathcal{M} \to [0,\infty]$ be a finite measure and $\nu \ll \mu$. Then there exists a Radon–Nikodym derivative* $\mathrm{d}\nu/\mathrm{d}\mu$*, i.e.*

$$\nu(E) = \int_E \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \,\mathrm{d}\mu$$

for every $E \in \mathcal{M}$.

Exercise C.4.39 (σ-finite Radon–Nikodym). A measure space is called *σ-finite* if it is a countable union of sets of finite measure. Generalise the Radon–Nikodym Theorem to σ-finite measure spaces. For example, if μ is σ-finite, we can find a sequence $E_j \nearrow X$ with $\mu(E_j) < \infty$ and define

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} := \sup_j \frac{\mathrm{d}\nu|_{E_j}}{\mathrm{d}\mu|_{E_j}}.$$

Exercise C.4.40. Let $\nu = \lambda_{\mathbb{R}^n}$ be the Lebesgue measure, and let $(X, \mathcal{M}, \mu) = (\mathbb{R}^n, \mathcal{M}(\lambda^*_{\mathbb{R}^n}), \mu)$, where μ is the counting measure; this measure space is not σ-finite, but $\nu \ll \mu$. Show that ν cannot be of the form $\nu(E) = \int_E f \,\mathrm{d}\mu$. Thus there is no anologue to the Radon–Nikodym Theorem in this case.

Before proving the Radon–Nikodym Theorem C.4.38, let us deal with the essential special case of the result:

Lemma C.4.41. *Let $\mu, \nu : \mathcal{M} \to [0,\infty]$ be finite measures such that $\nu \le \mu$. Then there exists a Radon–Nikodym derivative* $\mathrm{d}\nu/\mathrm{d}\mu$*, i.e.*

$$\nu(E) = \int_E \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \,\mathrm{d}\mu$$

for every $E \in \mathcal{M}$. Moreover,

$$\int g^+ \,\mathrm{d}\nu = \int g^+ \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \,\mathrm{d}\mu \tag{C.16}$$

when $g^+ : X \to [0,\infty]$ is $\mathcal{M}$-measurable.

Proof. A *$\mathcal{M}$-partition* of a set X is a finite disjoint collection $\mathcal{P} \subset \mathcal{M}$, for which $X = \bigcup \mathcal{P}$. Let us define a partial order $\le$ on the family the $\mathcal{M}$-partitions by $\mathcal{P} \le \mathcal{Q}$ if and only if for every $Q \in \mathcal{Q}$ there exists $P \in \mathcal{P}$ such that $Q \subset P$. The *common refinement* of $\mathcal{M}$-partitions $\mathcal{P}, \mathcal{Q}$ is the $\mathcal{M}$-partition

$$\uparrow \{\mathcal{P}, \mathcal{Q}\} = \{P \cap Q : \ P \in \mathcal{P},\ Q \in \mathcal{Q}\}.$$

For a $\mathcal{M}$-partition $\mathcal{P}$, let us define $d_{\mathcal{P}} : \mathcal{M} \to \mathbb{R}$ by

$$d_{\mathcal{P}}(E) := \begin{cases} \frac{\nu(P)}{\mu(P)}, & \text{if } x \in P \in \mathcal{P} \text{ and } \mu(P) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then $0 \leq d_{\mathcal{P}} \leq 1$, $d_{\mathcal{P}}$ is simple and μ-integrable, and

$$d_{\mathcal{P}} = \sum_{P \in \mathcal{P}} \frac{\nu(P)}{\mu(P)} \chi_P.$$

The idea in the following is that the Radon–Nikodym derivative $\mathrm{d}\nu/\mathrm{d}\mu$ will be approximated by functions $d_{\mathcal{P}}$ is the $L^2(\mu)$-sense. If $\mathcal{P} \leq \mathcal{Q}$ and $E \in \mathcal{P}$ then

$$\nu(E) = \int_E d_{\mathcal{P}} \, \mathrm{d}\mu = \int_E d_{\mathcal{Q}} \, \mathrm{d}\mu, \tag{C.17}$$

because

$$\begin{aligned} \int_E d_{\mathcal{Q}} \, \mathrm{d}\mu &= \int_E \sum_{Q \in \mathcal{Q}} \frac{\nu(Q)}{\mu(Q)} \chi_Q \, \mathrm{d}\mu \\ &= \sum_{Q \in \mathcal{Q}} \frac{\nu(Q)}{\mu(Q)} \int \chi_{Q \cap E} \, \mathrm{d}\mu \\ &\overset{E \in \mathcal{P} \leq \mathcal{Q}}{=} \sum_{Q \in \mathcal{Q}:\ Q \subset E} \frac{\nu(Q)}{\mu(Q)} \mu(Q) \\ &\overset{E \in \mathcal{P} \leq \mathcal{Q}}{=} \nu(E). \end{aligned}$$

Moreover, here

$$\|d_{\mathcal{P}}\|^2_{L^2(\mu_E)} \leq \|d_{\mathcal{Q}}\|^2_{L^2(\mu_E)} = \|d_{\mathcal{P}}\|^2_{L^2(\mu_E)} + \|d_{\mathcal{Q}} - d_{\mathcal{P}}\|^2_{L^2(\mu_E)}, \tag{C.18}$$

because

$$\begin{aligned} \|d_{\mathcal{P}}\|^2_{L^2(\mu_E)} &\leq \|d_{\mathcal{P}}\|^2_{L^2(\mu_E)} + \|d_{\mathcal{Q}} - d_{\mathcal{P}}\|^2_{L^2(\mu_E)} \\ &= \int_E d_{\mathcal{P}}^2 \, \mathrm{d}\mu + \int_E (d_{\mathcal{Q}} - d_{\mathcal{P}})^2 \, \mathrm{d}\mu \\ &= \int_E d_{\mathcal{Q}}^2 \, \mathrm{d}\mu + 2 \int_E d_{\mathcal{P}} \, (d_{\mathcal{P}} - d_{\mathcal{Q}}) \, \mathrm{d}\mu \\ &\overset{E \in \mathcal{P}}{=} \|d_{\mathcal{Q}}\|^2_{L^2(\mu_E)} + 2 \int_E \frac{\nu(E)}{\mu(E)} \left(\frac{\nu(E)}{\mu(E)} - d_{\mathcal{Q}} \right) \mathrm{d}\mu \\ &\overset{(C.17)}{=} \|d_{\mathcal{Q}}\|^2_{L^2(\mu_E)} + 2 \frac{\nu(E)}{\mu(E)} \left(\frac{\nu(E)}{\mu(E)} - \nu(E) \right) \\ &= \|d_{\mathcal{Q}}\|^2_{L^2(\mu_E)}. \end{aligned}$$

Now

$$\begin{aligned} M &:= \sup\left\{ \|d_{\mathcal{P}}\|^2_{L^2(\mu)} \mid \mathcal{P} \text{ is an } \mathcal{M}\text{−partition} \right\} \\ &\overset{0\le d_{\mathcal{P}}\le 1}{\le} \mu(X) < \infty. \end{aligned}$$

Take a sequence of $\mathcal{M}$-partitions $\mathcal{P}_k$ such that

$$\|d_{\mathcal{P}_k}\|^2_{L^2(\mu)} \xrightarrow[k\to\infty]{} M.$$

We obtain an increasing sequence of partitions $\mathcal{Q}_k$ by common refinements:

$$\begin{cases} \mathcal{Q}_1 := \mathcal{P}_1, \\ \mathcal{Q}_{k+1} :=\uparrow \{\mathcal{P}_{k+1}, \mathcal{Q}_k\}. \end{cases}$$

Let us show that the sequence of functions $f_k := d_{\mathcal{Q}_k}$ converges to the Radon–Nikodym derivative $\mathrm{d}\nu/\mathrm{d}\mu$ in $L^2(\mu)$. First, $f_k \in L^2(\mu)$, because $\mu(X) < \infty$. Moreover, these functions form a Cauchy-sequence, because

$$\begin{aligned} \|f_j - f_k\|^2_{L^2(\mu)} &\overset{(C.18)}{=} \left| \|f_j\|^2_{L^2(\mu)} - \|f_k\|^2_{L^2(\mu)} \right| \\ &\xrightarrow[j,k\to\infty]{} 0, \end{aligned}$$

as $M \ge \|f_k\|^2_{L^2(\mu)} \ge \|d_{\mathcal{P}_k}\|^2_{L^2(\mu)} \to M$. Since $L^2(\mu)$ is a Banach space, there exists $f \in L^2(\mu)$ for which $\|f - f_k\|_{L^2(\mu)} \to 0$. Let us show that $f = \mathrm{d}\nu/\mathrm{d}\mu$. Take $E \in \mathcal{M}$. Let $d_k := d_{\mathcal{R}_k}$, where

$$\mathcal{R}_k :=\uparrow \{Q_k, \{E, X \setminus E\}\}.$$

Then

$$\begin{aligned} \nu(E) &\overset{(C.17)}{=} \int_E d_k \ \mathrm{d}\mu \\ &= \int_E (d_k - f_k) \ \mathrm{d}\mu + \int_E f_k \ \mathrm{d}\mu \\ &\xrightarrow[k\to\infty]{} \int_E f \ \mathrm{d}\mu, \end{aligned}$$

because $\int_E f_k \ \mathrm{d}\mu \to \int Ef \ \mathrm{d}\mu$ by the Monotone Convergence Theorem C.3.6 and by Vitali's Convergence Theorem C.4.14, and because

$$\begin{aligned} \left| \int_E (d_k - f_k) \ \mathrm{d}\mu \right| &\le \int_E |d_k - f_k| \ \mathrm{d}\mu \\ &\overset{\text{Hölder}}{\le} \left(\int_E |d_k - f_k|^2 \ \mathrm{d}\mu \right)^{1/2} \left(\int_E \mathrm{d}\mu \right)^{1/2} \\ &\le \|d_k - f_k\|^2_{L^2(\mu)} \ \mu(X)^{1/2} \\ &\overset{(C.18)}{\le} \left| \|d_k\|^2_{L^2(\mu)} - \|f_k\|^2_{L^2(\mu)} \right| \ \mu(X)^{1/2} \\ &\xrightarrow[k\to\infty]{} 0. \end{aligned}$$

Thus $f = \mathrm{d}\nu/\mathrm{d}\mu$. Finally, let $g^+ \geq 0$ be $\mathcal{M}$-measurable. Take simple $\mathcal{M}$-measurable functions s_k for which

$$0 \leq s_k(x) \leq s_{k+1}(x) \xrightarrow[k\to\infty]{} g^+(x).$$

Then

$$\begin{aligned}
\int g^+ \ \mathrm{d}\nu &= \int \lim_{k\to\infty} s_k \ \mathrm{d}\mu \\
&\overset{\text{Mon.conv.}}{=} \lim_{k\to\infty} \int s_k \ \mathrm{d}\mu \\
&\overset{(*)}{=} \lim_{k\to\infty} \int s_k \ \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \ \mathrm{d}\mu \\
&\overset{\text{Mon.conv.}}{=} \int \lim_{k\to\infty} s_k \ \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \ \mathrm{d}\mu \\
&= \int g^+ \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \ \mathrm{d}\mu,
\end{aligned}$$

where equality $(*)$ easily follows from $\int \chi_E \ \mathrm{d}\nu = \int \chi_E \ \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \ \mathrm{d}\mu$. □

Proof of the Radon–Nikodym Theorem C.4.38. Since $\nu = (\nu^+ - \nu^-) \ll \mu$, we have also $\nu^+, \nu^- \ll \mu$. If the Radon–Nikodym derivatives $\mathrm{d}\nu^+/\mathrm{d}\mu$ and $\mathrm{d}\nu^-/\mathrm{d}\mu$ exist, then by the linearity of the integral we have

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = \frac{\mathrm{d}\nu^+}{\mathrm{d}\mu} - \frac{\mathrm{d}\nu^-}{\mathrm{d}\mu}.$$

Thus we may assume that ν, μ are finite measures, where $\nu \ll \mu$. Then also $\mu + \nu : \mathcal{M} \to [0,\infty]$ is a finite measure. By Lemma C.4.41, the Radon–Nikodym derivatives $\mathrm{d}\mu/\mathrm{d}(\mu+\nu)$ and $\mathrm{d}\nu/\mathrm{d}(\mu+\nu)$ exist. Let

$$A := \left\{ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)} > 0 \right\};$$

let us show that $\mathrm{d}\nu/\mathrm{d}\mu = g^+$, where $g^+ : X \to [0,\infty]$ is defined by

$$g^+(x) := \begin{cases} \frac{\mathrm{d}\nu}{\mathrm{d}(\mu+\nu)} \Big/ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}, & \text{when } x \in A, \\ 0, & \text{when } x \in A^c. \end{cases}$$

Here $\mu(A^c) \overset{(C.16)}{=} \int_{A^c} \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}\, \mathrm{d}(\mu+\nu) = 0$, and if $E \in \mathcal{M}$ then

$$\begin{aligned}
\nu(E) &= \nu(A \cap E) + \nu(A^c \cap E) \\
&\overset{\mu(A^c\cap E)=0,\ \nu\ll\mu}{=} \nu(A \cap E) \\
&\overset{(C.16)}{=} \int_{A\cap E} \frac{\mathrm{d}\nu}{\mathrm{d}(\mu+\nu)}\, \mathrm{d}(\mu+\nu) \\
&= \int_{A\cap E} g^+ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}\, \mathrm{d}(\mu+\nu) \\
&\overset{(C.16)}{=} \int_{A\cap E} g^+\, \mathrm{d}\mu \\
&\overset{\mu(A^c)=0}{=} \int_E g^+\, \mathrm{d}\mu.
\end{aligned}$$

Thus $g^+ = \mathrm{d}\nu/\mathrm{d}\mu$, and the Radon–Nikodym Theorem C.4.38 is proven. □

Exercise C.4.42. Let $\lambda, \mu, \nu : \mathcal{M} \to [0, \infty]$ be σ-finite measures. Prove:
(a) If $\lambda \ll \mu$, $E \in \mathcal{M}$ and g is $\mathcal{M}$-measurable, then

$$\int_E g\, \mathrm{d}\lambda = \int_E g \frac{\mathrm{d}\lambda}{\mathrm{d}\mu}\, \mathrm{d}\mu.$$

(b) If $\lambda \ll \nu$ and $\mu \ll \nu$, then $\frac{\mathrm{d}(\lambda+\mu)}{\mathrm{d}\nu} = \frac{\mathrm{d}\lambda}{\mathrm{d}\nu} + \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$.
(c) If $\lambda \ll \mu$ and $\mu \ll \nu$, then $\frac{\mathrm{d}\lambda}{\mathrm{d}\nu} = \frac{\mathrm{d}\lambda}{\mathrm{d}\mu}\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$.
(d) If $\lambda \ll \mu$ and $\mu \ll \lambda$, then $\frac{\mathrm{d}\lambda}{\mathrm{d}\mu} = \left(\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\right)^{-1}$.

Definition C.4.43 (Lebesgue decomposition). Let $\mu, \nu : \mathcal{M} \to [0, \infty]$ be measures. A *Lebesgue decomposition* of ν with respect to μ is a pair (ν_0, ν_1) of measures $\nu_0, \nu_1 : \mathcal{M} \to [0, \infty]$ satisfying

$$\nu = \nu_0 + \nu_1, \quad \begin{cases} \nu_0 \perp \mu, \\ \nu_1 \ll \mu. \end{cases}$$

Theorem C.4.44 (Existence of Lebesgue decomposition). *Let $\mu, \nu : \mathcal{M} \to [0, \infty]$ be σ-finite measures. Then there exists a unique Lebesgue decomposition of ν with respect to μ.*

Proof. The Radon–Nikodym Theorem C.4.38 was formulated for a finite measure, but it can be easily generalised to σ-finite spaces: showing this was left as Exercise C.4.39. Let $A := \left\{ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)} > 0 \right\}$. Define measures $\nu_0, \nu_1 : \mathcal{M} \to [0, \infty]$ by

$$\begin{cases} \nu_0(E) := \nu(A^c \cap E), \\ \nu_1(E) := \nu(A \cap E). \end{cases}$$

Clearly $\nu = \nu_0 + \nu_1$, and $\nu_0 \perp \mu$ because

$$\begin{cases} \nu_0(A) = \nu(A^c \cap A) = \nu(\emptyset) = 0, \\ \mu(A^c) \overset{(C.16)}{=} \int_{A^c} \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}\ \mathrm{d}(\mu+\nu) = 0. \end{cases}$$

We will now prove that $\nu_1 \ll \mu$. Let $A_k := \left\{ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)} \geq 1/k \right\}$. Take $E \in \mathcal{M}$ such that $\mu(E) = 0$. Now $\nu_1(E) = 0$, because

$$\begin{aligned} \nu_1(E) &= \nu(A \cap E) \\ &\leq (\mu+\nu)(A \cap E) \\ &= \lim_{k\to\infty} (\mu+\nu)(A_k \cap E) \\ &\leq \int_{A_k \cap E} k\ \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}\ \mathrm{d}(\mu+\nu) \\ &\overset{\text{Radon-Nikodym}}{=} k\ \mu(A_k \cap E) \\ &\leq k\ \mu(E) \quad = \quad 0. \end{aligned}$$

Proving the uniqueness part is left as Exercise C.4.45. □

Exercise C.4.45. Show that the Lebesgue decomposition in Theorem C.4.44 is unique.

C.4.4 Integration as functional on function spaces

Assume $(X, \mathcal{M}, \mu)$ is a measure space, possibly with topology. On function spaces like $L^p(\mu)$ or $C(X) = C(X, \mathbb{R})$ (when X is e.g. a compact Hausdorff space), integration acts as a bounded linear functional by

$$f \mapsto \int fg\ \mathrm{d}\mu,$$

when g is a suitable weight function on X. It is natural to study necessary and sufficient conditions for g, and ask whether all the bounded linear functionals are of this form. The general functional analytic outline is as follows: Let V be a real Banach space, e.g. $L^p(\mu)$ or $C(X) = C(X, \mathbb{R})$. The dual of V is the Banach space

$$V' = \mathcal{L}(V, \mathbb{R}) := \{\phi : V \to \mathbb{R} \mid \phi \text{ bounded and linear}\},$$

endowed with the (operator) norm

$$\phi \mapsto \|\phi\| := \sup_{f \in V:\ \|f\|_V \leq 1} |\phi(f)|,$$

see Definition B.4.15 and Exercise B.4.16. Given a "concrete" space V, we would like to discover an intuitive representation of the dual.

C.4.5 Integration as functional on $L^p(\mu)$

Now we are going to find a concrete presentation for the dual of $V = L^p(\mu)$, where $(X, \mathcal{M}, \mu)$ is a measure space. We shall assume that $\mu(X) < \infty$, though often this technical assumption can be removed, since everything works for σ-finite measures just as well.

Lemma C.4.46. *Let μ be a finite measure. Let $1 \le p \le \infty$, and let $q = p'$ be its Lebesgue conjugate, i.e. $1/p + 1/q = 1$. Let $g \in L^q(\mu)$. Then $\phi_g \in L^p(\mu)'$, where*

$$\phi_g(f) := \int fg \, \mathrm{d}\mu,$$

and $\|\phi_g\| = \|g\|_{L^q}$.

Exercise C.4.47. Prove Lemma C.4.46.

Remark C.4.48. You may generalise Lemma C.4.46 as follows: the conclusion holds for a general measure μ if $1 < p \le \infty$, and for a σ-finite measure μ if $1 \le p \le \infty$.

Remark C.4.49. The next Theorem C.4.50 roughly says that the dual of L^p "is" L^q, under some technical assumptions. The result holds for a general measure μ if $1 < p < \infty$, and for a σ-finite measure if $1 \le p < \infty$.

Theorem C.4.50 (**Dual of $L^p(\mu)$**)**.** *Let μ be a finite measure. Let $1 \le p < \infty$, and let $q = p'$ be its Lebesgue conjugate. Then the mapping*

$$(g \mapsto \phi_g) : L^q(\mu) \to L^p(\mu)'$$

is an isometric isomorphism, i.e. $L^p(\mu)' \cong L^q(\mu)$.

Proof. By the previous Lemma C.4.46, it suffices to show that $\psi \in L^p(\mu)'$ is of the form $\psi = \phi_g$ for some $g \in L^q(\mu)$. Let us define $\nu : \mathcal{M} \to \mathbb{R}$ such that

$$\nu(E) := \psi(\chi_E),$$

where $\chi_E \in L^p(\mu)$ because $\mu(X) < \infty$. The idea in the proof is to show that $\mathrm{d}\nu/\mathrm{d}\mu \in L^q(\mu)$ and that

$$\psi(f) = \int f \, \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu. \tag{C.19}$$

The first step is to show that ν is a signed measure: Let $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ be a disjoint collection. Then $\nu(\bigcup_{j=1}^\infty E_j) = \sum_{j=1}^\infty \nu(E_j)$, because

$$\begin{aligned}
\left|\nu(\bigcup_{j=1}^{\infty} E_j) - \sum_{j=1}^{k} \nu(E_j)\right| &= \left|\psi(\chi_{\bigcup_{j=1}^\infty E_j}) - \sum_{j=1}^{k} \psi(E_j)\right| \\
&= \left|\psi(\sum_{j=k+1}^{\infty} \chi_{E_j})\right| \\
&\leq \|\psi\| \sum_{j=k+1}^{\infty} \|\chi_{E_j}\|_{L^p(\mu)} \\
&\xrightarrow[k\to\infty]{\mu(X)<\infty} 0,
\end{aligned}$$

where we used the linearity and boundedness of ψ, and the disjointness of $\{E_j\}_{j=1}^\infty$. Thus ν is a signed measure. Moreover $\nu \ll \mu$, because if $\mu(E) = 0$ then $\chi_E = 0$ μ-almost everywhere, implying

$$\nu(E) = \psi(\chi_E) = 0$$

as $\psi \in L^p(\mu)'$. Thus $\mathrm{d}\nu/\mathrm{d}\mu \in L^1(\mu)$ exists by the Radon–Nikodym Theorem C.4.38. We have to show that $\mathrm{d}\nu/\mathrm{d}\mu \in L^q(\mu)$ and that (C.19) holds for every $f \in L^p(\mu)$. At least

$$\psi(\chi_E) = \nu(E) \overset{\text{Radon−Nikodym}}{=} \int \chi_E \, \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu$$

for every $E \in \mathcal{M}$; by the linearity of ψ, this extends to

$$\psi(s) = \int s \, \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu \tag{C.20}$$

for every simple $\mathcal{M}$-measurable $s : X \to \mathbb{R}$. Next we show that $\mathrm{d}\nu/\mathrm{d}\mu \in L^q(\mu)$.

For a moment, let $p = 1$ (so that $q = \infty$). We shall soon see that $\|\mathrm{d}\nu/\mathrm{d}\mu\|_{L^q} \leq$

$\|\psi\|$. Take any $M > \|\psi\|$ and let $A_M := \left\{ \left| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right| > M \right\}$. Then

$$\begin{aligned} M\ \mu(A_M) &= \int_{A_M} M\ \mathrm{d}\mu \\ &\leq \int_{A_M} \left| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right|\ \mathrm{d}\mu \\ &= \chi_{A_M}\ \mathrm{sgn}(\frac{\mathrm{d}\nu}{\mathrm{d}\mu})\ \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\ \mathrm{d}\mu \\ &\overset{(C.20)}{=} \psi(\chi_{A_M}\ \mathrm{sgn}(\frac{\mathrm{d}\nu}{\mathrm{d}\mu})) \\ &\leq \|\psi\|\ \left\| \chi_{A_M}\ \mathrm{sgn}(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}) \right\|_{L^p(\mu)} \\ &\overset{p=1}{\leq} \|\psi\|\ \mu(A_M). \end{aligned}$$

Since $M > \|\psi\|$ and $\mu(A_M) < \infty$, we must have $\mu(A_M) = 0$, so $\|\mathrm{d}\nu/\mathrm{d}\mu\|_{L^\infty(\mu)} \leq \|\psi\|$.

Now let $1 < p < \infty$ (so that $\infty > q > 1$). Take simple $\mathcal{M}$-measurable functions $h_k : X \to \mathbb{R}$ such that

$$0 \leq h_k(x) \leq h_{k+1}(x) \xrightarrow[k\to\infty]{} \left| \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(x) \right|.$$

Then

$$\left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|^q_{L^q(\mu)} = \int \left| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right|\ \mathrm{d}\mu \overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \int h_k^q\ \mathrm{d}\mu,$$

so that $\mathrm{d}\nu/\mathrm{d}\mu \in L^q(\mu)$ follows if we show that $\|h_k\|_{L^q(\mu)} \leq \mathit{constant} < \infty$ for every $k \in \mathbb{Z}^+$:

$$\begin{aligned} \|h_k\|^q_{L^q(\mu)} &= \int h_k^q\ \mathrm{d}\mu \\ &\leq \int h_k^{q-1}\ \left| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right|\ \mathrm{d}\mu \\ &= \int h_k^{q-1}\ \mathrm{sgn}(\frac{\mathrm{d}\nu}{\mathrm{d}\mu})\ \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\ \mathrm{d}\mu \\ &\overset{(C.20)}{=} \psi(h_k^{q-1}\ \mathrm{sgn}(\frac{\mathrm{d}\nu}{\mathrm{d}\mu})) \\ &\leq \|\psi\|\ \left\| h_k^{q-1} \right\|_{L^p(\mu)} \\ &= \|\psi\|\ \|h_k\|_{L^{q/p}(\mu)}, \end{aligned}$$

because $p(q-1)=q$. Hence $\|h_k\|_{L^q(\mu)} = \|h_k\|_{L^q(\mu)}^{q(1-1/p)} \le \|\psi\|$.

Finally, we have to show that (C.19) holds for $f \in L^p(\mu)$. Take simple $\mathcal{M}$-measurable functions $f_k : X \to [-\infty, +\infty]$ such that $f_k \to f$ in $L^p(\mu)$. Then

$$\begin{aligned}
& \left|\psi(f) - \int f \,\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,\mathrm{d}\mu\right| \\
\overset{(C.20)}{=} \quad & \left|\psi(f-f_k) + \int (f_k - f)\,\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,\mathrm{d}\mu\right| \\
\le \quad & |\psi(f-f_k)| + \int |f_k - f| \left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right| \mathrm{d}\mu \\
\overset{\text{Hölder}}{\le} \quad & \|\psi\|\ \|f-f_k\|_{L^p(\mu)} + \|f_k - f\|_{L^p(\mu)} \left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_{L^q(\mu)} \\
\xrightarrow[k\to\infty]{} \quad & 0.
\end{aligned}$$

Thus the proof is complete. □

Exercise C.4.51. Generalise Theorem C.4.50 to the case, where μ is σ-finite (and $1 \le p < \infty$).

Exercise C.4.52. Generalise Theorem C.4.50 to the case, where μ is any measure and $1 < p < \infty$ (so that $1 < q < \infty$ also). (Hint: apply the result of Exercise C.4.51.)

Remark C.4.53. We have not dealt with the dual of $L^\infty(\mu)$. This case actually resembles the other L^p-cases, but is slightly different, see details e.g. in [150]. Often, however, $L^\infty(\mu)' \ncong L^1(\mu)$.

Exercise C.4.54. Let $X = [0,1]$ and $\mu = (\lambda_{\mathbb{R}})_X$. Show that there exists $\psi \in L^\infty(\mu)'$ which is not of the form $f \mapsto \int fg \,\mathrm{d}\mu$ for any $g \in L^1(\mu)$.
(Hint: Define a suitable bounded linear functional $f \mapsto \varphi(f)$ for continuous functions f, and extend it to ψ using the Hahn–Banach Theorem, see Theorem B.4.25).

Exercise C.4.55. Let $(X, \mathcal{M}, \mu)$ be a measure space, where X is uncountable, $\mathcal{M} = \{E \subset X :\ E \text{ or } E^c \text{ is countable}\}$ and μ is the counting measure. Show that there exists $\psi \in L^1(\mu)'$ which is not of the form $f \mapsto \int fg \,\mathrm{d}\mu$ for any $g \in L^\infty(\mu)$.
(Hint: You may use that there exists $S \in \mathcal{P}(X) \setminus \mathcal{M}$, which follows by using the Hausdorff Maximal Principle or other equivalents to the Axiom of Choice).

Theorem C.4.56 (**Converse of Hölder's inequality**)**.** *Let μ be a σ-finite measure, $1 \le p \le \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. Let S be the space of all simple functions that vanish outside a set of finite measure. Let g be $\mathcal{M}$-measurable such that $fg \in L^1(\mu)$ for all $f \in S$, and such that*

$$M_q(g) := \sup\left\{ \left|\int fg \,\mathrm{d}\mu\right| :\ f \in S,\ \|f\|_{L^p(\mu)} = 1 \right\}$$

is finite. Then $g \in L^q(\mu)$ and $M_q(g) = \|g\|_{L^q(\mu)}$.

Proof. From Hölder's inequality (Theorem C.4.4) we have the inequality $M_q(g) \leq ||g||_{L^q(\mu)}$. For the proof of $||g||_{L^q(\mu)} \leq M_q(g)$ we follow [35]. Assume first that $q < \infty$. Let $E_n \subset X$ be an increasing sequence of sets such that $0 < \mu(E_n) < \infty$ for all n, and such that $\bigcup_{n=1}^{\infty} E_n = X$. Let φ_n be a sequence of simple functions such that $\varphi_n \to g$ pointwise and $|\varphi_n| \leq g$, and let $g_n := \varphi_n \chi_{E_n}$, where χ_{E_n} is the characteristic function of the set E_n. Then $g_n \to g$ pointwise, $|g_n| \leq |g|$ and $g_n \in S$. Define $f_n := ||g_n||_{L^q(\mu)}^{1-q} |g_n|^{q-1} \frac{g}{|g|}$ when $g \neq 0$ and $f_n := 0$ when $g = 0$. The relation $\frac{1}{p} + \frac{1}{q} = 1$ implies $(q-1)p = q$, so that $||f_n||_{L^p(\mu)} = 1$, and by Fatou's lemma C.3.10 we have:

$$\begin{aligned} ||g||_{L^q(\mu)} &\leq \liminf ||g_n||_{L^q(\mu)} \\ &= \liminf \int |f_n g_n| \, \mathrm{d}\mu \\ &\leq \liminf \int |f_n g| \, \mathrm{d}\mu \\ &= \liminf \int f_n g \, \mathrm{d}\mu \leq M_q(g). \end{aligned}$$

The case $q = \infty$ is slightly different. Take $\epsilon > 0$ and denote $A := \{x \in X : |g(x)| \geq M_\infty(g) + \epsilon$. We need to show that $\mu(A) = 0$. If $\mu(A) > 0$, there exists some $B \subset A$ such that $0 < \mu(B) < \infty$. Let us define $f := \frac{1}{\mu(B)} \frac{g}{|g|} \chi_B$ when $g \neq 0$ and $f := 0$ when $g = 0$. Then $||f||_{L^1(\mu)} = 1$ and $\int fg \, \mathrm{d}\mu \geq M_\infty(g) + \epsilon$, a contradiction. □

C.4.6 Integration as functional on $C(X)$

Measure theory and topology have fundamental connections, as exemplified in this passage. For our purposes, it is enough to study compact Hausdorff spaces, though analogies hold for locally compact Hausdorff spaces. Let (X, τ) be a compact Hausdorff space and let $C(X) = C(X, \mathbb{R})$ denote the Banach space of continuous functions $f : X \to \mathbb{R}$, endowed with the supremum norm:

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

Appealing to the "geometry" of X, we are going to characterise the dual $C(X)' = \mathcal{L}(C(X), \mathbb{R})$.

Exercise C.4.57. Let (X, τ) be a compact Hausdorff space. Actually, $C(X)$ contains all the information about (X, τ): a set $S \subset X$ is closed if and only if $S = \{f = 0\}$ for some $f \in C(X)$. Prove this.

Remark C.4.58. Let (X, τ) be a topological space. Recall that the vector space of signed (Borel) measures

$$M(X) = M(\Sigma(\tau)) := \{\nu : \Sigma(\tau) \to \mathbb{R} \mid \nu \text{ is a signed measure}\}$$

is a Banach space with the norm

$$\|\nu\| := |\nu|(X).$$

Lemma C.4.59. *Let $\nu : \mathcal{M} \to \mathbb{R}$ be a signed measure on X, where $\tau \subset \mathcal{M}$. For $f \in C(X)$, let*

$$\begin{aligned} T_\nu(f) &:= \int f \ \mathrm{d}\nu \\ &= \int f \ \mathrm{d}\nu^+ - \int f \ \mathrm{d}\nu^-. \end{aligned}$$

Then $T_\nu \in C(X)'$ and $\|T_\nu\| \leq \|\nu\|$.

Proof. Each $f \in C(X)$ is $\mathcal{M}$-measurable, as $\tau \subset \mathcal{M}$. Furthermore, $\nu^+, \nu^- : \mathcal{M} \to [0, \infty]$ are finite measures. Consequently, $f \in C(X)$ is $\nu^\pm$-integrable, so $T_\nu(f) \in \mathbb{R}$ is well-defined, and

$$\begin{aligned} |T_\nu(f)| &\leq \int |f| \ \mathrm{d}\nu^+ + \int |f| \ \mathrm{d}\nu^- \\ &\leq \|f\| \ \left(\nu^+(X) + \nu^-(X)\right) \quad = \quad \|f\| \ \|\nu\|. \end{aligned}$$

The operator $T_\nu : C(X) \to \mathbb{R}$ is linear since integration is linear. □

Theorem C.4.60 (F. Riesz's Topological Representation Theorem). *Let (X, τ) be a compact Hausdorff space. Let $M(X)$ and $T_\nu \in C(X)'$ be as above. Then*

$$(\nu \mapsto T_\nu) : M(X) \to C(X)'$$

is an isometric isomorphism.

In other words, bounded linear functionals on $C(X)$ are exactly integrations with respect to signed measures, with the natural norms coinciding. We shall soon prove the Riesz Representation Theorem C.4.60 step-wise.

Definition C.4.61 (Positive functionals). Let (X, τ) be a compact Hausdorff space. A functional $T : C(X) \to \mathbb{R}$ is called *positive* if $T(f) \geq 0$ whenever $f \geq 0$.

Exercise C.4.62. Show that a positive linear functional $T \in C(X)'$ is bounded and that

$$\|T\| = T1,$$

where $1 \in C(X)$ is the constant function $x \mapsto 1$.

Lemma C.4.63. *Let $T \in C(X)'$, where (X, τ) is a compact Hausdorff space. Then there exist positive $T^+, T^- \in C(X)'$ such that*

$$\begin{aligned} T &= T^+ - T^-, \\ \|T\| &= \|T^+\| + \|T^-\|. \end{aligned}$$

Proof. For $f = f^+ - f^- \in C(X)$, let us define

$$T^+(f) := T^+(f^+) - T^+(f^-),$$

where

$$T^+(g^+) := \sup\left\{T(h^+) \mid h^+ \in C(X),\ 0 \leq h^+ \leq g^+\right\}.$$

Obviously, $0 = T(0) \leq T^+(g^+) \leq T(\|g^+\|1) = \|T\|\ \|g^+\|$. Thereby the functional $T^+ : C(X) \to \mathbb{R}$ is well-defined and positive. Let us show that T^+ is linear. If $0 < \lambda^+ \in \mathbb{R}$ then

$$\begin{aligned} T^+(\lambda^+ f^+) &= \sup\left\{Th \mid h \in C(X),\ 0 \leq h \leq \lambda^+ f^+\ \right\} \\ &= \sup\left\{T(\lambda^+ h) \mid h \in C(X),\ 0 \leq h \leq f^+\ \right\} \\ &\overset{T \text{ linear}}{=} \lambda^+\ T^+(f^+); \end{aligned}$$

from this we easily see that

$$T^+(\lambda f) = \lambda\ T^+(f)$$

for every $\lambda \in \mathbb{R}$ and $f \in C(X)$. Next,

$$T^+(f^+ + g^+) = T^+(f^+) + T^-(g^+)$$

whenever $0 \leq f^+, g^+ \in C(X)$, because

$$\text{if} \quad \begin{cases} 0 \leq h \leq f^+ + g^+, \\ 0 \leq h_1 \leq f^+, \\ 0 \leq h_2 \leq g^+ \end{cases} \qquad \text{then} \quad \begin{cases} 0 \leq h_1 + h_2 \leq f^+ + g^+, \\ 0 \leq \min(f^+, h) \leq f^+, \\ 0 \leq h - \min(f^+, h) \leq g^+. \end{cases}$$

Since

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+,$$

we get

$$T^+((f+g)^+) + T^+(f^-) + T^+(g^-) = T^+((f+g)^-) + T^+(f^+) + T^+(g^+),$$

so that

$$\begin{aligned} T^+(f+g) &= T^+((f+g)^+) - T^+((f+g)^-) \\ &= \left(T^+(f^+) - T^+(f^-)\right) + \left(T^+(g^+) - T^+(g^-)\right) \\ &= T^+(f) + T^+(g). \end{aligned}$$

Hence we have seen that $T^+ : C(X) \to \mathbb{R}$ is linear and positive, and that $\|T^+\| \leq \|T\|$. Next, let us define $T^- := T^+ - T \in C(X)'$. Then T^- is positive, because

$$T^-(f^+) = \sup\left\{Th - T(f^+) \mid h \in C(X):\ 0 \leq h \leq f^+\right\}.$$

Finally,

$$\begin{aligned}
\|T\| &= \|T^+ - T^-\| \\
&\leq \|T^+\| + \|T^-\| \\
&= T^+(1) + T^-(1) \\
&= 2\,T^+(1) - T(1) \\
&= \sup\{T(2h-1) \mid h \in C(X) : 0 \leq h \leq 1\} \\
&= \sup\{T(g) \mid g \in C(X) : -1 \leq g \leq 1\} \\
&= \|T\|,
\end{aligned}$$

so $\|T\| = \|T^+\| + \|T^-\|$. □

Remark C.4.64. Recall that the *support* $\mathrm{supp}(f) \subset X$ of a function $f \in C(X)$ is the closure of the set $\{f \neq 0\}$. Moreover, abbreviations

$$K \prec f, \quad f \prec U$$

mean that $0 \leq f \leq 1$, $K \subset X$ is compact such that $\chi_K \leq f$, and $U \subset X$ is open such that $\mathrm{supp}(f) \subset U$.

Theorem C.4.65. *Let $T^+ \in C(X)'$ be positive, where (X, τ) is a compact Hausdorff space. Then there exists a finite Borel measure $\mu : \Sigma(\tau) \to [0, \infty]$ such that*

$$Tf = \int f \, \mathrm{d}\mu$$

for every $f \in C(X)$.

Proof. Let us define a measurelet $m : \tau \to [0, \infty]$ such that

$$m(U) := \sup\{Tf \mid f \prec U\}.$$

Indeed, $m(\emptyset) = T(0) = 0$. Thus m generates an outer measure $m^* : \mathcal{P}(X) \to [0, \infty]$ by

$$m^*(E) = \inf\left\{\sum_{j=1}^{\infty} m(U_j) : \quad E \subset \bigcup_{j=1}^{\infty} U_j \in \tau, U_j \in \tau\right\}.$$

We have to show $\mu := m^*|_{\Sigma(\tau)}$ is the desired measure. First,

$$m^*(E) = \inf\{m(U) : \quad E \subset U \in \tau\} \tag{C.21}$$

follows, if we show that $m(\bigcup_{j=1}^{\infty} U_j) \leq \sum_{j=1}^{\infty} m(U_j)$. So let $f \prec \bigcup_{j=1}^{\infty} U_j$. Now $\mathrm{supp}(f) \subset X$ is compact, so $\mathrm{supp}(f) \subset \bigcup_{j=1}^{n} U_j$ for some $n \in \mathbb{Z}^+$. Let $\{g_j\}_{j=1}^{n}$ be a partition of unity for which

$$K \prec \sum_{j=1}^{n} g, \quad g_j \prec U_j.$$

Thereby

$$\begin{aligned}
Tf &= T(f\sum_{j=1}^{n} g_j) \\
&\overset{T \text{ linear}}{=} \sum_{j=1}^{n} T(fg_j) \\
&\overset{fg_j \prec U_j}{\leq} \sum_{j=1}^{n} m(U_j) \\
&\leq \sum_{j=1}^{\infty} m(U_j),
\end{aligned}$$

proving (C.21). Next, we show that $\tau \subset \mathcal{M}(m^*)$ by proving that $m^*(A \cup B) = m^*(A) + m^*(B)$ whenever $A \subset U \in \tau$ and $B \subset U^c$; let us assume the non-trivial case $m^*(A), m^*(B) < \infty$. Given $\varepsilon > 0$, there exists $V \in \tau$ such that $A \cup B \subset V$ and $m^*(A \cup B) + \varepsilon > m(V)$. Moreover, let

$$\begin{cases} f \prec U \cap V: & m(U \cap V) < Tf + \varepsilon, \\ g \prec \operatorname{supp}(f)^c \cap V: & m(\operatorname{supp}(f)^c \cap V) < Tg + \varepsilon. \end{cases}$$

We notice that $U \in \mathcal{M}(m^*)$, because

$$\begin{aligned}
m^*(A \cup B) + \varepsilon &> m(V) \\
&\overset{f+g \prec V}{\geq} T(f+g) \\
&\overset{T \text{ linear}}{=} Tf + Tg \\
&> m(U \cap V) + m(\operatorname{supp}(f)^c \cap V) - 2\varepsilon \\
&\geq m^*(A) + m^*(B) - 2\varepsilon \\
&\geq m^*(A \cup B) - 2\varepsilon.
\end{aligned}$$

Thus we can define the Borel measure $\mu := m^*|_{\Sigma(\tau)}$. Notice that $m(U) = \mu(U)$, $\mu(X) = T1 < \infty$ and that m^* is Borel-regular. If $\chi_E \leq g \leq \chi_F$, where $g \in C(X)$ and $E, F \in \Sigma(\tau)$, then

$$\int \chi_E \, \mathrm{d}\mu \leq \int g \, \mathrm{d}\mu \leq \int \chi_F \, \mathrm{d}\mu; \tag{C.22}$$

moreover,

$$\mu(E) \leq Tg \leq \mu(F), \tag{C.23}$$

because

$$\begin{array}{rcl} \mu(E) & \stackrel{E\subset\{g\geq 1\}}{\leq} & \mu(\{g\geq 1\}) \\ & \stackrel{\delta>0}{\leq} & \mu(\{g>1-\delta\}) \\ & \stackrel{\{g>1-\delta\}\in\tau}{=} & \sup\{Tf:\ f\prec\{g>1-\delta\}\} \\ & \stackrel{T\text{ positive},\ 0<\delta<1}{\leq} & T(g/(1-\delta)) \\ & \xrightarrow[\delta\to 0]{T\text{ linear}} & Tg \end{array}$$

implies $\mu(E)\leq Tg$, which implies

$$\mu(X\setminus F) \stackrel{\chi_{F^c}\leq 1-g}{\overset{}{T}} (1-g) = \mu(X) - Tg,$$

so $Tg\leq\mu(F)$. Let us show that $T(f^+) = \int f^+\ \mathrm{d}\mu$, when $0\leq f^+\in C(X)$. Take $\varepsilon>0$. Let us define $f_\varepsilon\in C(X)$ by

$$f_\varepsilon(x) := \min\left\{f^+(x),\varepsilon\right\}.$$

Then

$$f^+ = \sum_{k=0}^{\infty}\left(f_{(k+1)\varepsilon} - f_{k\varepsilon}\right), \tag{C.24}$$

where the sum is actually finite, as f^+ is bounded. Combining

$$\varepsilon\chi_{\{f^+\geq(k+1)\varepsilon\}} \leq f_{(k+1)\varepsilon} - f_{k\varepsilon} \leq \varepsilon\chi_{\{f^+\geq k\varepsilon\}}$$

with inequalities (C.22,C.23), we get

$$\begin{cases}\varepsilon\mu(\{f^+\geq(k+1)\varepsilon\}) \leq \int(f_{(k+1)\varepsilon}-f_{k\varepsilon})\ \mathrm{d}\mu \leq \varepsilon\mu(\{f^+\geq k\varepsilon\}),\\ \varepsilon\mu(\{f^+\geq(k+1)\varepsilon\}) \leq T(f_{(k+1)\varepsilon}-f_{k\varepsilon}) \leq \varepsilon\mu(\{f^+\geq k\varepsilon\}).\end{cases} \tag{C.25}$$

We obtain

$$\begin{array}{rcl} \left|\int f^+\ \mathrm{d}\mu - T(f^+)\right| & \stackrel{(C.24)}{\leq} & \sum_{k=0}^{\infty}\left|\int(f_{(k+1)\varepsilon}-f_{k\varepsilon})\ \mathrm{d}\mu - T(f_{(k+1)\varepsilon}-f_{k\varepsilon})\right| \\ & \stackrel{(C.25)}{\leq} & \varepsilon\sum_{k=0}^{\infty}\left(\mu(\{f^+\geq k\varepsilon\}) - \mu(\{f^+\geq(k+1)\varepsilon\})\right) \\ & = & \varepsilon\mu(X) \\ & \xrightarrow[\varepsilon\to 0]{\mu(X)<\infty} & 0, \end{array}$$

i.e. $T(f^+) = \int f^+\ \mathrm{d}\mu$. Consequently, $Tf = \int f\ \mathrm{d}\mu$ for every $f\in C(X)$. □

Proof of the Riesz Representation Theorem C.4.60. Let $T \in C(X)'$. Then $T = T^+ - T^-$ by Lemma C.4.63, and Theorem C.4.65 provides finite Borel measures $\mu, \lambda : \Sigma(\tau) \to [0, \infty]$ such that

$$\begin{cases} T^+(f) = \int f \, \mathrm{d}\mu, \\ T^-(f) = \int f \, \mathrm{d}\lambda \end{cases}$$

for all $f \in C(X)$. Thus $\mu - \lambda : \Sigma(\tau) \to \mathbb{R}$ is a signed measure for which

$$Tf = \int f \, \mathrm{d}(\mu - \lambda)$$

for every $f \in C(X)$. Moreover,

$$\begin{array}{rcl} \|T\| & = & \|T_{\mu-\lambda}\| \\ & \overset{\text{Lemma } C.4.59}{\leq} & \|\mu - \lambda\| \\ & \leq & \|\mu\| + \|\lambda\| \\ & = & \mu(X) + \lambda(X) \quad = \quad T^+1 + T^-1 \\ & = & \|T^+\| + \|T^-\| \\ & = & \|T\|, \end{array}$$

so $\|T\| = \|\mu - \nu\|$. □

Remark C.4.66. The Riesz Representation Theorem C.4.60 can be generalised. For instance, let (X, τ) be a locally compact Hausdorff space, which is also second countable (i.e. τ has a countable base). Endow the vector spaces

$$\begin{array}{rcl} C_0(X) & := & \{f \in C(X) \mid \{f \geq \varepsilon\} \subset X \text{ is compact for every } \varepsilon > 0\}, \\ M(X) & := & \{\nu : \Sigma(\tau) \to \mathbb{R} \mid \nu \text{ is a signed measure}\} \end{array}$$

with respective complete norms $f \mapsto \|f\| = \sup\{|f(x)| : x \in X\}$ and $\|\nu\| := |\nu|(X)$; the rough idea is that a function $f \in C_0(X)$ "*vanishes at infinity*". Then

$$(\nu \mapsto T_\nu) : M(X) \to C_0(X)'$$

is an isometric isomorphism, where

$$T_\nu f := \int f \, \mathrm{d}\nu = \int f \, \mathrm{d}\nu^+ - \int f \, \mathrm{d}\nu^-.$$

We shall not prove this claim here.

Exercise C.4.67. Let μ be a finite Borel measure on a compact metric space (X, d). Prove that $C(X)$ is dense in $L^p(\mu)$ (with the natural embedding) if $1 \leq p < \infty$. What could be a problem with $p = \infty$?

Exercise C.4.68. Let us define $(\tau_h f)(x) := f(x+h)$, when $f : \mathbb{R} \to [-\infty, \infty]$. Show that $\|\tau_h f - f\|_{L^p(\mathbb{R}^n)} \to_{h \to 0} 0$ if $f \in L^p(\mathbb{R}^n)$ and $1 \leq p < \infty$. Why $p \neq \infty$?

C.5 Product measure and integral

Let $(X, \mathcal{M}_\mu, \mu)$ and $(Y, \mathcal{M}_\nu, \nu)$ be measure spaces. We are going to study the possibility of changing the order of iterated integrals, i.e. whether

$$\int_X \int_Y f \; \mathrm{d}\nu \; \mathrm{d}\mu = \int_Y \int_X f \; \mathrm{d}\mu \; \mathrm{d}\nu$$

under some reasonable assumptions on μ, ν and $f : X \times Y \to [-\infty, +\infty]$. Of course, there are many technical issues involved: for instance,

$$\int_X \int_Y f \; \mathrm{d}\nu \; \mathrm{d}\mu = \int_X \left(x \mapsto \int_Y (y \mapsto f(x,y)) \; \mathrm{d}\nu \right) \; \mathrm{d}\nu,$$

where $y \mapsto f(x,y)$ is $\mathcal{M}_\nu$-measurable for μ-almost every $x \in X$, and so on. We start by defining the product measure $\mu \times \nu$ on $X \times Y$.

Definition C.5.1 (**Product measures).** Let $\mathcal{A} := \{A \times B : \; A \in \mathcal{M}_\mu, \; B \in \mathcal{M}_\nu\}$. Define a measurelet $m : \mathcal{A} \to [0, \infty]$ on $X \times Y$ by

$$m(A \times B) := \mu(A) \; \nu(B).$$

The outer measure $m^* : \mathcal{P}(X \times Y) \to [0, \infty]$ generated by the measurelet m is called the *product outer measure* of the measures μ and ν. The *product measure* of μ and ν is the restricted measure $\mu \times \nu := m^*|_{\mathcal{M}_{\mu\times\nu}}$, where $\mathcal{M}_{\mu\times\nu} := \mathcal{M}(m^*)$.

Remark C.5.2. Recall that m generates m^* by

$$m^*(S) = \inf \left\{ \sum_{j=1}^{\infty} m(A_j \times B_j) : \; S \subset \bigcup_{j=1}^{\infty} A_j \times B_j, \; \{A_j \times B_j\}_{j=1}^{\infty} \subset \mathcal{A} \right\}.$$

Actually, we can do here better: we may demand that $\{A_j \times B_j\}_{j=1}^{\infty} \subset \mathcal{A}$ is disjoint. Why is that? For any two families $\{A_j\}_{j=1}^{\infty} \subset \mathcal{M}_\mu$ and $\{B_j\}_{j=1}^{\infty} \subset \mathcal{M}_\nu$, let

$$\begin{cases} F_1 := A_1, \\ F_{k+1} := A_{k+1} \setminus \bigcup_{j=1}^{k} A_j, \end{cases} \qquad \begin{cases} G_1 := B_1, \\ G_{k+1} := B_{k+1} \setminus \bigcup_{j=1}^{k} B_j. \end{cases}$$

Clearly, $\{F_j \times G_k \mid j, k \in \mathbb{Z}^+\} \subset \mathcal{A}$ is a disjoint family, and it is easy to check that

$$E = \bigcup \left\{ F_j \times G_k \mid j, k \in \mathbb{Z}^+ : \; F_j \times G_k \subset E \right\},$$

where $E = \bigcup_{i=1}^{\infty} A_i$. Moreover,

$$\sum_{j=1}^{\infty} m(A_j \times B_j) \geq \sum_{j,k \leq n: \; F_j \times G_k \subset E} m(F_j \times G_k)$$

for each $n \in \mathbb{Z}^+$. Letting $n \to \infty$ yields the claim.

Proposition C.5.3. *Let $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$. Then $A \times B \in \mathcal{M}_{\mu\times\nu}$, and hence the $\mathcal{A}$-generated σ-algebra $\Sigma(\mathcal{A}) \subset \mathcal{M}_{\mu\times\nu}$. Moreover,*

$$(\mu \times \nu)(A \times B) = \mu(A)\ \nu(B).$$

Remark C.5.4. Of course $\int_X \int_Y \chi_{A\times B}\ \mathrm{d}\nu\ \mathrm{d}\mu = \int_Y \int_X \chi_{A\times B}\ \mathrm{d}\mu\ \mathrm{d}\nu = \mu(A)\nu(B)$ here, but this certainly does not prove the claims above.

Proof. To prove that $A \times B \in \mathcal{M}_{\mu\times\nu}$, it suffices to show that

$$m^*(S \cup T) \geq m^*(S) + m^*(T)$$

whenever $S \subset A \times B$ and $T \subset (A \times B)^c$ such that $m^*(S), m^*(T) < \infty$. Take $\varepsilon > 0$. Let $\{A_j \times B_j\}_{j=1}^\infty \subset \mathcal{A}$ be disjoint such that

$$\begin{aligned} S \cup T &\subset \bigcup_{j=1}^\infty A_j \times B_j, \\ \varepsilon + m^*(S \cup T) &> \sum_{j=1}^\infty m(A_j \times B_j). \end{aligned}$$

Let us define $S_j, T_j, U_j \subset X \times Y$ by

$$\begin{aligned} S_j &:= (A_j \times B_j) \cap (A \times B), \\ T_j &:= (A_j \times B_j) \cap (A \times (Y \setminus B)), \\ U_j &:= (A_j \times B_j) \cap ((X \setminus A) \times Y). \end{aligned}$$

Then $\{S_j, T_j, U_j\}_{j=1}^\infty \subset \mathcal{A}$ is disjoint, and $A_j \times B_j = S_j \cup T_j \cup U_j$. Moreover,

$$S \subset \bigcup_{j=1}^\infty S_j, \quad T \subset \bigcup_{j=1}^\infty (T_j \cup U_j),$$

so that

$$\begin{aligned} \varepsilon + m^*(S \cup T) &> \sum_{j=1}^\infty m(A_j \times B_j) \\ &= \sum_{j=1}^\infty m(S_j) + \sum_{j=1}^\infty (m(T_j) + m(U_j)) \\ &\geq m^*(S) + m^*(T). \end{aligned}$$

Thus we have shown that $A \times B \in \mathcal{M}_{\mu\times\nu}$. Finally, if $\{A_j \times B_j\}_{j=1}^{\infty} \subset \mathcal{A}$ is a cover of $A \times B$ then

$$\begin{aligned} m^*(A\times B) &\overset{\text{trivial}}{\leq} m(A\times B) \\ &= \mu(A)\ \nu(B) \\ &= \int_X \int_Y \chi_{A\times B}\ \mathrm{d}\nu\ \mathrm{d}\mu \\ &\leq \int_X \int_Y \sum_{j=1}^{\infty} \chi_{A_j\times B_j}\ \mathrm{d}\nu\ \mathrm{d}\mu \\ &\overset{\text{Mon.Conv.}}{=} \sum_{j=1}^{\infty} m(A_j\times B_j). \end{aligned}$$

Therefore $m^*(A \times B) = \mu(A)\ \nu(B)$. □

Exercise C.5.5. Show that for the Lebesgue measures, $\lambda_{\mathbb{R}^m} \times \lambda_{\mathbb{R}^n} = \lambda_{\mathbb{R}^{m+n}}$.

Definition C.5.6. For $x \in X$, the *x-slice* $S^x \subset Y$ of a set $S \subset X \times Y$ is

$$S^x := \{y \in Y \mid (x, y) \in S\}.$$

Remark C.5.7. Let

$$\mathcal{B} = \{R \in \Sigma(\mathcal{A}) :\ R^x \in \mathcal{M}_\nu \text{ for all } x \in X\}$$

Clearly $X \times Y \in \mathcal{A} \subset \mathcal{B}$. If $R \in \mathcal{B}$ then also $R^c = (X \times Y) \setminus R \in \mathcal{B}$, because $(R^c)^x = (R^x)^c$. Similarly, if $\{R_j\}_{j=1}^{\infty} \subset \mathcal{B}$ then $\bigcup_{j=1}^{\infty} R_j \in \mathcal{B}$, because $(\bigcup_{j=1}^{\infty} R_j)^x = \bigcup_{j=1}^{\infty} (R_j)^x$. Thus $\Sigma(\mathcal{A}) \subset \mathcal{B}$.

Lemma C.5.8. *The product outer measure m^* is $\Sigma(\mathcal{A})$-regular: for any $S \subset X \times Y$ there exists $R \in \Sigma(\mathcal{A})$ such that $S \subset R$ and $m^*(S) = m^*(R)$. Moreover, if μ, ν are finite then the x-slice $R^x \in \mathcal{M}_\nu$ for every $x \in X$, $x \mapsto \nu(R^x)$ is $\mathcal{M}_\mu$-measurable, and*

$$m^*(R) = \int_X \int_Y \chi_R\ \mathrm{d}\nu\ \mathrm{d}\mu.$$

Proof. For each $k \in \mathbb{Z}^+$, take a disjoint family $\{A_{kj} \times B_{kj}\}_{j=1}^{\infty} \subset \mathcal{A}$ such that

$$\begin{aligned} S &\subset \bigcup_{j=1}^{\infty} A_{kj} \times B_{kj}, \\ m^*(S) + \frac{1}{k} &\geq \sum_{j=1}^{\infty} m(A_{kj} \times B_{kj}). \end{aligned}$$

Let $R_n := \bigcap_{k=1}^{n} \bigcup_{j=1}^{\infty} (A_{kj} \times B_{kj})$ and $R := \bigcap_{n=1}^{\infty} R_n$. Then $S \subset R \in \Sigma(\mathcal{A})$. Moreover, we have $m^*(S) = m^*(R)$, because

$$\begin{aligned} m^*(S) &\leq m^*(R) \\ &\leq m^*(\bigcup_{j=1}^{\infty} A_{nj} \times B_{nj}) \\ &= \sum_{j=1}^{\infty} m(A_{nj} \times B_{nj}) \\ &\leq m^*(S) + \frac{1}{n}. \end{aligned}$$

The set R_n is the union of a disjoint family $\{C_{nj} \times D_{nj}\}_{j=1}^{\infty} \subset \mathcal{A}$, and

$$\chi_{R_n^x}(y) = \sum_{j=1}^{\infty} \chi_{C_{nj}}(x)\ \chi_{D_{nj}}(y).$$

Consequently, $\chi_{R_n^x} : Y \to \mathbb{R}$ is $\mathcal{M}_\nu$-measurable for all $x \in X$.

$$1 \leq \chi_{R_n^x}(y) \leq \chi_{R_{n+1}^x}(y) \xrightarrow[n\to\infty]{} \chi_{R^x}(y) \geq 0,$$

Lebesgue's Dominated Convergence Theorem C.3.22 yields

$$h_n(x) := \int_Y \chi_{R_n^x}\ \mathrm{d}\nu \xrightarrow[n\to\infty]{\nu(Y)<\infty} \int_Y \chi_{R^x}\ \mathrm{d}\nu =: h(x).$$

Then $h_n : X \to [0, \infty)$ is $\mathcal{M}_\mu$-measurable and

$$h_n(x) = \sum_{j=1}^{\infty} \nu(D_{nj})\ \chi_{C_{nj}}(x).$$

Moreover,

$$\infty > \nu(Y) \geq h_n(x) \geq h_{n+1}(x) \xrightarrow[n\to\infty]{} h(x) \geq 0,$$

so also $h : X \to [0, \infty)$ is $\mathcal{M}_\mu$-measurable and

$$\int_X \int_Y \chi_{R_n}\ \mathrm{d}\nu\ \mathrm{d}\mu = \int_X h_n\ \mathrm{d}\mu \xrightarrow[n\to\infty]{\mu(X)<\infty} \int_X h\ \mathrm{d}\mu = \int_X \int_Y \chi_R\ \mathrm{d}\nu\ \mathrm{d}\mu,$$

where Lebesgue's Dominated Convergence Theorem was used again. □

The following result will neatly justify calling m^* *the* product outer measure.

Corollary C.5.9. *Let μ, ν be finite. Then $(\mu \times \nu)^* = m^*$.*

Proof. If $E \subset X \times Y$ then

$$\begin{aligned} (\mu\times\nu)^*(E) &= \inf\{(\mu\times\nu)(S) \mid E \subset S \in \mathcal{M}_{\mu\times\nu}\} \\ &\overset{\text{Lemma C.5.8}}{=} \inf\{m^*(R) \mid E \subset R \in \Sigma(\mathcal{A})\} \\ &\geq m^*(E) \\ &\geq (\mu\times\nu)^*(E), \end{aligned}$$

that is $(\mu\times\nu)^*(E) = m^*(E)$. □

Exercise C.5.10. Generalise Corollary C.5.9 for σ-finite measures.

Exercise C.5.11. Let $\mu_0 = \mu^*|_{\mathcal{M}(\mu^*)}$ and $\nu_0 = \nu^*|_{\mathcal{M}(\nu^*)}$. Is $\mu\times\nu = \mu_0\times\nu_0$?

Proposition C.5.12. *Let μ,ν be finite and complete, $S \subset \mathcal{M}_{\mu\times\nu}$. Then the x-slice $S^x \in \mathcal{M}_\nu$ for μ-almost every $x \in X$, and $x \mapsto \nu(S^x)$ is $\mathcal{M}_\mu$-measurable. Moreover,*

$$(\mu\times\nu)(S) = \int_X \int_Y \chi_S \, \mathrm{d}\nu \, \mathrm{d}\mu.$$

Proof. By Lemma C.5.8, there exists a set $R \in \Sigma(\mathcal{A}) \subset \mathcal{M}_{\mu\times\nu}$ such that $S \subset R$ and

$$\begin{aligned} (\mu\times\nu)(S) &= m^*(S) = m^*(R) \\ &= \int_X \int_Y \chi_R \, \mathrm{d}\nu \, \mathrm{d}\mu. \end{aligned}$$

Thus the result follows, if we can show that $\nu(R^x) = \nu(S^x)$ for μ-almost every $x \in X$. Here $R^x = S^x \cup (R \setminus S)^x$. By Lemma C.5.8, there exists a set $Q \in \Sigma(\mathcal{A})$ such that $R \setminus S \subset Q$ and

$$m^*(R \setminus S) = \int_X \int_Y \chi_Q \, \mathrm{d}\nu \, \mathrm{d}\mu.$$

Now $m^*(R \setminus S) = m^*(R) - m^*(S) = 0$, because the measures are finite. Therefore $\nu(Q^x) = 0$ for μ-almost every $x \in X$. Because $R \setminus S \subset Q$ and ν is complete, we see that

$$(R \setminus S)^x \in \mathcal{M}_\nu, \quad \nu((R \setminus S)^x) = 0$$

for μ-almost every $x \in X$. This shows that $\nu(R^x) = \nu(S^x)$ for μ-almost every $x \in X$. □

Theorem C.5.13 (Fubini–Tonelli Theorem). *Let μ,ν be finite complete measures,* *and let $f \geq 0$ be $\mathcal{M}_{\mu\times\nu}$-measurable. Then*

$$\begin{aligned} x \mapsto f(x,y) \quad & \textit{is} \quad \mathcal{M}_\mu\textit{-measurable } \nu\textit{-a.e.}, \\ y \mapsto \int_X (x \mapsto f(x,y))\ \mathrm{d}\mu \quad & \textit{is} \quad \mathcal{M}_\nu\textit{-measurable, and} \\ \int_{X\times Y} f\ \mathrm{d}(\mu\times\nu) & = \int_Y \int_X f\ \mathrm{d}\mu\ \mathrm{d}\nu \\ & = \int_X \int_Y f\ \mathrm{d}\nu\ \mathrm{d}\mu. \end{aligned}$$

Proof. Take $\mathcal{M}_{\mu\times\nu}$-measurable simple functions $f_k : X \times Y \to [0,\infty)$ for which

$$f_k \leq f_{k+1} \xrightarrow[k\to\infty]{\text{pointwise}} f.$$

By Proposition C.5.12, take $N_k \in \mathcal{M}_\mu$ such that $\mu(N_k) = 0$ and

$$(y \mapsto f_k(x,y)) : Y \to [0,\infty)$$

is $\mathcal{M}_\nu$-measurable for all $x \in X \setminus N_k$. Then $N := \bigcup_{k=1}^\infty N_k \in \mathcal{M}_\mu$, $\mu(N) = 0$ and

$$(y \mapsto f(x,y)) : Y \to [0,\infty]$$

is $\mathcal{M}_\nu$-measurable for all $x \in X \setminus N$. Let us define $g_k, g : X \to [0,\infty]$ such that $g_k(x) := 0 =: g(x)$ if $x \in N$, and otherwise

$$g_k(x) := \int_Y f_k(x,y)\ \mathrm{d}\nu(y) \xrightarrow[k\to\infty]{\text{Mon.Conv.}} \int_Y f(x,y)\ \mathrm{d}\nu(y) =: g(x).$$

By Proposition C.5.12, $g_k : X \to [0,\infty]$ is $\mathcal{M}_\mu$-measurable, and clearly

$$g_k \leq g_{k+1} \xrightarrow[k\to\infty]{\text{pointwise}} g.$$

Thus $g : X \to [0,\infty]$ is $\mathcal{M}_\mu$-measurable, and

$$\begin{aligned} \int_X \int_Y f\ \mathrm{d}\nu\ \mathrm{d}\mu \quad & = \quad \int_X g\ \mathrm{d}\mu \\ & \overset{\text{Mon.Conv.}}{=} \lim_{k\to\infty} \int g_k\ \mathrm{d}\mu \\ & = \lim_{k\to\infty} \int_X \int_Y f_k\ \mathrm{d}\nu\ \mathrm{d}\mu \\ & \overset{\text{Prop. C.5.12}}{=} \lim_{k\to\infty} \int_{X\times Y} f_k\ \mathrm{d}(\mu\times\nu) \\ & \overset{\text{Mon.Conv.}}{=} \int_{X\times Y} f\ \mathrm{d}(\mu\times\nu). \end{aligned}$$

This concludes the proof. □

Corollary C.5.14 (Fubini theorem). *Let $f \in L^1(\mu \times \nu)$. Then*

$$\int_{X \times Y} f \, \mathrm{d}(\mu \times \nu) = \int_X \int_Y f \, \mathrm{d}\nu \, \mathrm{d}\mu = \int_Y \int_X f \, \mathrm{d}\mu \, \mathrm{d}\nu.$$

□

Theorem C.5.15 (σ-finite Fubini–Tonelli and Fubini theorems). *The results in Proposition C.5.12, the Fubini–Tonelli Theorem C.5.13 and its Corollary C.5.14 hold also for σ-finite complete measures μ, ν.*

Exercise C.5.16. Prove Theorem C.5.15 by taking an increasing sequence of measurable sets E_n of finite measure such that $\bigcup_n E_n = X \times Y$.

Exercise C.5.17. Let $X = [0,1] = Y$, $\mu := (\lambda_{\mathbb{R}})_X$ and $\nu := (E \mapsto \#E) : \mathcal{P}(Y) \to [0,\infty]$. Let $f = \chi_S$, where $S := \{(x,y) \in X \times Y : \ x = y\}$. Calculate
(a) $\int_X \int_Y f \, \mathrm{d}\nu \, \mathrm{d}\mu$,
(b) $\int_Y \int_X f \, \mathrm{d}\mu \, \mathrm{d}\nu$,
(c) $\int_{X \times Y} f \, \mathrm{d}(\mu \times \nu)$.

Remark C.5.18. In Exercise C.5.17, the measure ν is not σ-finite. Thus there is no contradiction with Theorem C.5.15.

Exercise C.5.19. Let $(X, \mathcal{M}, \mu)$ be a complete σ-finite measure space. Let $f : X \to [0,\infty]$ be $\mathcal{M}$-measurable. Show that

$$\int f \, \mathrm{d}\mu = \int_{[0,\infty)} \mu(\{f > t\}) \, \mathrm{d}\lambda_{\mathbb{R}}(t).$$

Corollary C.5.20 (Young's inequality). *Let μ, ν be σ-finite and $1 < p < \infty$. Assume that $K : X \times Y \to \mathbb{C}$ is a $\mathcal{M}_{\mu \times \nu}$-measurable function satisfying*

$$C_1 := \sup_{y \in Y} \int_X |K(x,y)| \, \mathrm{d}\mu(x),$$
$$C_2 := \sup_{x \in X} \int_Y |K(x,y)| \, \mathrm{d}\nu(y),$$

where $C_1, C_2 < \infty$. For any $u \in L^p(\nu)$ define $Au : X \to \mathbb{C}$ by

$$Au(x) = \int_Y K(x,y) \, u(y) \, \mathrm{d}\nu(y).$$

Then

$$\|Au\|_{L^p(\mu)} \le C_1^{1/p} \, C_2^{1/q} \, \|u\|_{L^p(\nu)},$$

where q is the conjugate exponent of p.

Remark C.5.21. Notice that this defines a unique bounded linear operator $A : L^p(\nu) \to L^p(\mu)$.

Remark C.5.22. It is clear that we can replace sup by the esssup in the definition of C_1, C_2, where the esssup would be taken with respect to ν and μ in C_1 and C_2, respectively:

$$\begin{aligned} C_1 &:= \nu - \text{esssup}_{y\in Y} \int_X |K(x,y)| \, \mathrm{d}\mu(x), \\ C_2 &:= \mu - \text{esssup}_{x\in X} \int_Y |K(x,y)| \, \mathrm{d}\nu(y), \end{aligned}$$

with the same proof.

Proof of Corollary C.5.20. First, $y \mapsto K(x,y)\ u(y)$ is $\mathcal{M}_\nu$-measurable, and

$$\begin{aligned} |Au(x)| &\leq \int_Y \left(|K(x,y)|^{1/p}|u(y)|\right)\left(|K(x,y)|^{1/q}\right) \, \mathrm{d}\nu(y) \\ &\overset{\text{Holder}}{\leq} \left(\int_Y |K(x,y)||u(y)|^p \, \mathrm{d}\nu(y)\right)^{1/p} \left(\int_Y |K(x,y)| \, \mathrm{d}\nu(y)\right)^{1/q} \\ &\leq \left(\int_Y |K(x,y)||u(y)|^p \, \mathrm{d}\nu(y)\right)^{1/p} C_2^{1/q}. \end{aligned}$$

Using this we get

$$\begin{aligned} \|Au\|^p_{L^p(\mu)} &= \int_X |Au(x)|^p \, \mathrm{d}\mu(x) \\ &\leq C_2^{p/q} \int_X \int_Y |K(x,y)||u(y)|^p \, \mathrm{d}\nu(y) \, \mathrm{d}\mu(x) \\ &\overset{\text{Fubini}}{=} C_2^{p/q} \int_Y |u(y)|^p \int_X |K(x,y)| \, \mathrm{d}\mu(x) \, \mathrm{d}\nu(y) \\ &\leq C_1 C_2^{p/q} \int_Y |u(y)|^p \, \mathrm{d}\nu(y) \\ &= \left(C_1^{1/p} C_2^{1/q}\right)^p \ \|u\|^p_{L^p(\mu)}, \end{aligned}$$

which gives the result. □

Theorem C.5.23 (Minkowski's inequality for integrals). *Let μ, ν be σ-finite and let $f : X \times Y \to \mathbb{C}$ be a $\mathcal{M}_{\mu\times\nu}$-measurable function. Let $1 \leq p < \infty$. Then*

$$\left\{\int \left(\int |f(x,y)| \, \mathrm{d}\nu(y)\right)^p \mathrm{d}\mu(x)\right\}^{1/p} \leq \int \left(\int |f(x,y)|^p \, \mathrm{d}\mu(x)\right)^{1/p} \mathrm{d}\nu(y).$$

Proof. If $p = 1$ the result follows from Theorem C.5.15 exchanging the order of the integration. For $1 < p < \infty$, taking $g \in L^q(\mu)$, $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\begin{aligned} & \int \left(\int |f(x,y)| \, \mathrm{d}\nu(y) \right) |g(x)| \, \mathrm{d}\mu(x) = \\ \overset{\text{Fubini}}{=} & \int \left(\int |f(x,y)||g(x)| \, \mathrm{d}\mu(x) \right) \, \mathrm{d}\nu(y) \\ \overset{\text{Holder}}{=} & \int \left(\int |f(x,y)|^p \, \mathrm{d}\mu(x) \right)^{1/p} \mathrm{d}\nu(y) ||g||_{L^q(\mu)}. \end{aligned}$$

Now the statement follows from the converse of Hölder's theorem (Theorem C.4.56). □

As a consequence, we obtain the second part of Minkowski's inequality for integrals:

Corollary C.5.24 (**Monotonicity of L^p-norm).** *Let μ, ν be σ-finite and let $f : X \times Y \to \mathbb{C}$ be a $\mathcal{M}_{\mu\times\nu}$-measurable function. Let $1 \le p \le \infty$. Assume that $f(\cdot, y) \in L^p(\mu)$ for ν-a.e. y, and assume that the function $y \mapsto ||f(\cdot, y)||_{L^p(\mu)}$ is in $L^1(\mu)$. Then $f(x, \cdot) \in L^1(\nu)$ for μ-a.e. x, the function $x \mapsto \int f(x, y) \, \mathrm{d}\nu(y)$ is in $L^p(\mu)$, and*

$$\| \int f(\cdot, y) \, \mathrm{d}\nu(y) \|_{L^p(\mu)} \le \int \|f(\cdot, y)\|_{L^p(\mu)} \, \mathrm{d}\nu(y).$$

Proof. For $p = \infty$ the statement follows from Theorem C.3.14. For $1 \le p < \infty$ it follows from Theorem C.5.23 and Fubini's Theorem C.5.15. □