Functional Analysis

Course notes for MATH 421/510: Real Analysis II
Term 2: January-April 2025
The University of British Columbia.

Mathav Murugan

version of April 14, 2025

Preface

We will cover the following topics in this course:

1. Point set topology.
2. Normed vector spaces.
3. L^p spaces.
4. Hilbert spaces.
5. Compact operators and their spectrum.

This notes will be updated throughout the term and made available at https://personal.math.ubc.ca/~mathav/teaching/notes/421notes.pdf. Please check the date on the first page to determine if you have the latest version. If you find any mistakes, typos, or have any other feedback, please let me know. Most of this material can be found in Folland's book [Fol].

There are various exercises throughout these notes. Try them all! Some of the exercises will be part of assignments (available on Canvas) and results mentioned in some of the exercises will be used in lectures. The notes will include all the topics covered in class. However the notes are rather terse as many oral discussions that gave further explanations or put results into a broader context are omitted. Additionally, diagrams drawn on the blackboard in class will not be reproduced in the notes.

Contents

1 Point set topology **3**

1.1 Continuous maps 11
1.2 Weak and product topologies 14
1.3 Urysohn's lemma and Tietze extension theorem 17
1.4 Nets 19
1.5 Compact sets 21
1.6 Nets, subnets, and compactness 24

2 Normed vector spaces **28**

2.1 Bounded linear maps and linear functionals 31
2.2 Dual space 34
2.3 Adjoint operator 41
2.4 L^p spaces 43
2.5 Baire category theorem and applications 52
2.6 Weak and weak* topologies 56
2.7 Metrizability of weak topology 60

3 Hilbert spaces **63**

3.1 Projection onto a closed convex set 66
3.2 Dual of a Hilbert space 68
3.3 Orthonormal basis 70

4 Compact operators **72**

4.1 Riesz theory of compact operators 76
4.2 Spectral decomposition of self-adjoint compact operators 80
4.3 The adjoint operator revisited 85

5 Appendix: Integration and measure **88**

1 Point set topology

The notions of continuity, limits and convergence are central to analysis. Usually, we first learn this in the setting of metric spaces (for example, ϵ-δ definition of continuity or ϵ-N definition of limits of a sequence). There are good reasons to go beyond metric spaces to study these notions such as

(1) Many useful modes of convergence do not have a metric associated with them (for example, notion of pointwise convergence of functions).

(2) Even for metric spaces, the properties of continuity, limits and convergence do not depend on the specific choice of metric but rather on the *topology* induced by the metric; that is, the collection of open sets associated with a metric.

The basic idea is to *define* a family of open sets.

Definition 1.1 (Topology)**.** Let X be a non-empty set. A **topology** $\mathcal{T}$ on X is a family of subsets of X such that

(i) $\varnothing, X \in \mathcal{T}$.

(ii) (closed under arbitrary unions) If $\{U_\alpha : \alpha \in A\}$ is a collection such that $U_\alpha \in \mathcal{T}$ for all $\alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{T}$.

(iii) (closed under finite intersections) For any $n \in \mathbb{N}$ and $U_1, \ldots, U_n \in \mathcal{T}$, we have $\bigcup_{i=1}^{n} U_i \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ is called a *topological space*.

Example 1.2. (1) $\{\varnothing, X\}$ is a topology (called the *trivial topology*). $\mathcal{P}(X)$ (power set of X) is a topology (called the *discrete topology*).

(2) In a metric space, open sets with respect to the metric form a topology (Exercise). Recall that a *metric space* (X, d) is a set X along with a non-negative function $d : X \times X \to [0, \infty)$ (called the *metric* or *distance function*) such that

(i) $d(x, y) \geqslant 0$ for all $x, y \in X$ and $d(x, y) = 0$ if and only if $x = y$.

(ii) (symmetry) $d(x, y) = d(y, x)$ for all $x, y \in X$.

(iii) (triangle inequality) $d(x, y) \leqslant d(x, z) + d(z, y)$ for all $x, y, z \in X$.

A set U in a metric space (X, d) is said to be *open* if for any $x \in U$, there exists $r > 0$ (r can depend on x) such that $B(x, r) \subset U$, where $B(x, r) = \{y \in X : d(x, y) < r\}$ is the *open ball* with center x and radius r. Then the collection of open sets $\mathcal{T} = \{U : U \subset X \text{ and } U \text{ is open}\}$ is a topology on X and is called the *topology induced by the metric d*.

(3) Let X be an infinite set. Then

$$\mathcal{T} = \{A \subset X : A = \varnothing \text{ or } A^c \text{ is finite}\}$$

is a topology and is called the *co-finite topology*.

Notation: $A \subset B$ allows for the possibility of equality; that is $A \subset B$ is same as $A \subseteq B$.

There are various constructions of new topological spaces from old ones. For example, a topology on X induces a topology on any subset $Y \subset X$.

Definition 1.3 (Relative topology)**.** Let $(X, \mathcal{T})$ be a topological space and let $Y \subset X$. Then

$$\mathcal{T}_Y := \{Y \cap U : U \in \mathcal{T}\}$$

is a topology on X and is called the *relative topology* or *subspace topology* on Y.

Exercise 1.4. Verify that all the topologies in Example 1.2 and Definition 1.3 satisfy properties (i),(ii),(iii) in Definition 1.1.

Question 1.5. Let (X, d_X) be a metric space and let $A \subset X$ be a non-empty subset. The function $d_A : A \times A \to \mathbb{R}$ defined as $d_A(x, y) = d_X(x, y)$ for all $x, y \in A$ is a metric on A (called the *restricted metric* on A). Let $\mathcal{T}_X$ denote the topology on X induced by d_X. Let us consider two topologies on A.

1. The topology on A induced by the restricted metric d_A.

2. The relative topology on A in the topological space $(X, \mathcal{T}_X)$.

What is the relation between the above two topologies on A? Are they the same?

Unless stated otherwise, we always assume that $(X, \mathcal{T})$ is a topological space for the rest of this section.

Definition 1.6 (Open/closed sets)**.** Let $(X, \mathcal{T})$ be a topological space. Elements of $\mathcal{T}$ are called *open sets*. We say $A \subset X$ is *closed* if A^c is open; that is, $A^c \in \mathcal{T}$.

Let $A \subset X$. We define the *interior of* A (denoted by A°) and the *closure of* A (denoted by $\overline{A}$) as

$$A^\circ = \text{interior of } A = \bigcup_{\substack{V \text{ is open,} \\ V \subset A}} V,$$

$$\overline{A} = \text{closure of } A = \bigcap_{\substack{F \text{ is closed,} \\ A \subset F}} F.$$

Note that A° is an open set and $\overline{A}$ is a closed set. Furthermore, A° is the largest open subset of A and $\overline{A}$ is the smallest closed set that contains A.
The set $\overline{A} \backslash A^\circ$ is called the *boundary of* A and is denoted as ∂A.
If $\overline{A} = X$, we say that A is *dense* in X.
If $(\overline{A})^\circ = \varnothing$, we say A is *nowhere dense* in X.

Since closed sets are complement of open sets, each of properties in Definition 1.1 can be rephrased in terms of closed sets.

Exercise 1.7. Let $(X, \mathcal{T})$ be a topological space.

(i) $\varnothing, X$ are closed sets.

(ii) If $(F_\alpha)_{\alpha\in A}$ is a collection of closed sets, then $\bigcap_{\alpha\in A} F_\alpha$ is closed.

(iii) If $n \in \mathbb{N}$ and $F_1, \ldots, F_n$ are closed sets, then $\bigcup_{j=1}^{n} F_j$ is closed.

We collect some basic properties of closure and interior of sets.

Lemma 1.8. *Let $(X, \mathcal{T})$ be a topological space and let $A, A_1, A_2 \subset X$.*

(1) $A^\circ \subset A \subset \overline{A}$.

(2) A° is open and $\overline{A}$ is closed.

(3) A is open if and only if $A = A^\circ$.

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

(5) $(A^c)^\circ = (\overline{A})^c$ and $(A^\circ)^c = \overline{A^c}$.

(6) $(A^\circ)^\circ = A^\circ$ and $\overline{(\overline{A})} = \overline{A}$.

(7) $A_1 \subset A_2$ implies $A_1^\circ \subset A_2^\circ$ and $\overline{A_1} \subset \overline{A_2}$.

(8) $\overline{A_1 \cup A_2} = \overline{A_1} \cup \overline{A_2}$.

(9) $(A_1 \cap A_2)^\circ = A_1^\circ \cap A_2^\circ$.

Proof. (1),(2) are immediate from the definition.
(3) If A is open, then A is one of V in the definition of A° and hence $A^\circ \supset A$. Hence $A^\circ = A$. The converse follows from (2).
(4) is similar to (3).
(5) Since $A \subset \overline{A}$, we have $(\overline{A})^c \subset A^c$. Since $(\overline{A})^c$ is open (by (2)), we have $(\overline{A})^c \subset (A^c)^\circ$. Let $B = (A^c)^\circ$, so B is open (by (2)). Then $B \subset A^c$ (by (1)), so $A \subset B^c$. Since B^c is closed, we have $\overline{A} \subset B^c$. So $(A^c)^\circ = B \subset (\overline{A})^c$. Hence we conclude $(A^c)^\circ = (\overline{A})^c$.

Replacing A with A^c in $(A^c)^\circ = (\overline{A})^c$, we obtain $A^\circ = \overline{A^c}^{\,c}$, which in turn implies $(A^\circ)^c = \overline{A^c}$.
(6),(7) are easy (you do this!).
(8) $\overline{A_1} \cup \overline{A_2}$ is closed (by (2)) and contains $A_1 \cup A_2$ (by (1)) and hence $\overline{A_1 \cup A_2} \subset \overline{A_1} \cup \overline{A_2}$. Conversely, $A_1 \subset A_1 \cup A_2$ and $A_2 \subset A_1 \cup A_2$ implies (using (7)) that $\overline{A_1} \subset \overline{A_1 \cup A_2}$ and $\overline{A_2} \subset \overline{A_1 \cup A_2}$. This in turn implies $\overline{A_1} \cup \overline{A_2} \subset \overline{A_1 \cup A_2}$.
(9) is similar to (8). □

Exercise 1.9. Let A and B be nowhere dense subsets of a topological space. Then show that $A \cup B$ is also nowhere dense.

The notion of closure is related to accumulation points.

Definition 1.10 (Neighborhood and accumulation points)**.** Let $x \in X$ and $A \subset X$. We say that A is a *neighborhood of* x if there exists an open set U such that $x \in U$ and $U \subset A$. We say that x is an *accumulation point of* A if $A \cap (U \backslash \{x\}) \neq \varnothing$ for any neighborhood U of x. By $\text{acc}(A)$, we denote the set of accumulation points of A.

Note that if U is a neighborhood of x, then there exists an open set V with $x \in V \subset U$. So V is also a neighborhood of x. Therefore, x is an accumulation point of A if and only if $A \cap (V \backslash \{x\}) \neq \varnothing$ for all open neighborhood of x.

Proposition 1.11. *(1)* $\overline{A} = A \cup \text{acc}(A)$.

(2) A *is closed if and only if* $\text{acc}(A) \subset A$.

Proof. (1) "$\supset$": $A \subset \overline{A}$, so we need to show that $\text{acc}(A) \subset \overline{A}$. Equivalently, it suffices to show that $(\overline{A})^c \subset (\text{acc}(A))^c$. So let $x \in (\overline{A})^c$. Hence $V = (\overline{A})^c$ is an open neighborhood of x. Since $A \subset \overline{A}$, we have $V \cap A = \varnothing$, and hence $x \notin \text{acc}(A)$. In other words, $x \in (\text{acc}(A))^c$.
"$\subset$": We need to show $\overline{A} \subset A \cup \text{acc}(A)$ or equivalently, $A^c \cap (\text{acc}(A))^c \subset (\overline{A})^c$. To this end, let $x \in A^c \cap (\text{acc}(A))^c$, so $x \notin A, x \notin \text{acc}(A)$. Therefore, there exists an open neighborhood V of x such that $A \cap (V \backslash \{x\}) = \varnothing$. Since $x \notin A$, we have $A \cap V = \varnothing$; that is, $A \subset V^c$. so $\overline{A} \subset V^c$ (since V^c is closed) and hence $V \subset (\overline{A})^c$. Since $x \in V$, we have $x \in (\overline{A})^c$ which concludes the proof.

(2) $\Longrightarrow$: A is closed implies $A = \overline{A}$ and hence $\text{acc}(A) \subset \overline{A} = A$ (by (1)).
$\Longleftarrow$:$\text{acc}(A) \subset A$ implies $\overline{A} = A \cup \text{acc}(A) = A$ (by (1)). Hence A is closed.

□

We introduce a terminology to compare two topologies on a space.

Definition 1.12. Let $\mathcal{T}_1, \mathcal{T}_2$ be two topologies on X. If $\mathcal{T}_1 \subset \mathcal{T}_2$, we say that $\mathcal{T}_1$ is *weaker* (or *coarser*) than $\mathcal{T}_2$. If $\mathcal{T}_1 \subset \mathcal{T}_2$, we say that $\mathcal{T}_2$ is *stronger* (or *finer*) than $\mathcal{T}_1$.

The following property is readily verified from the definition of a topology.

Exercise 1.13. If $\{\mathcal{T}_i : i \in I\}$ is a collection of topologies on X, then $\bigcap_{i \in I} \mathcal{T}_i$ is a topology on X.

The following is a common way to define a topology (we will use this method to define weak topology and product topology in §1.2).

Definition 1.14 (Topology generated by a collection of sets)**.** Let X be a set and $\mathcal{E} \subset \mathcal{P}(X)$ be a collection of subsets of X. Then

$$\mathcal{T}(\mathcal{E}) = \bigcap \{\mathcal{T} : \mathcal{E} \subset \mathcal{T}, \mathcal{T} \text{ is a topology on } X \}$$

is a topology (see Exercise 1.13) and is called the *topology generated by* $\mathcal{E}$.

In other words, the topology $\mathcal{T}(\mathcal{E})$ is the *smallest* topology that contains $\mathcal{E}$ in that it has the *fewest open sets* among all topologies that contain $\mathcal{E}$.

The following exercise provides a more concrete description of the topology generated by $\mathcal{E}$.

Exercise 1.15. Let $\mathcal{E} \subset \mathcal{P}(X)$ denote a collection of subsets of X. Let $\mathcal{E}'$ denote the set of finite intersections of elements of $\mathcal{E}$; that is

$$\mathcal{E}' = \{\bigcap_{j=1}^{n} E_i : n \in \mathbb{N}, E_i \in \mathcal{E} \text{ for all } i = 1, \dots, n\}.$$

Show that

$$\mathcal{T}(\mathcal{E}) = \{\emptyset, X\} \cup \{\bigcup_{\alpha \in A} U_\alpha : A \text{ is an arbitrary set and } \{U_\alpha : \alpha \in A\} \text{ is a collection of sets in } \mathcal{E}'\}. \tag{1.1}$$

In other words, $\mathcal{T}(\mathcal{E})$ contains the empty set, X, and arbitrary unions of finite intersections of elements in $\mathcal{E}$. **Hint**: For the inclusion '$\supset$', use Definition 1.1. For the inclusion '$\subset$', show that the right hand side of (1.1) is a topology containing $\mathcal{E}$.

It is a convenient property that the usual topology on $\mathbb{R}$ is generated by unbounded open intervals as described in the following exercise.

Exercise 1.16. Let d be the usual metric on $\mathbb{R}$; that is $d(x, y) = |x - y|$ for $x, y \in \mathbb{R}$. Let $\mathcal{T}$ denote the topology on $\mathbb{R}$ induced by d (see Remark 1.2-(2)). Let $\mathcal{E}$ denote the collection

$$\mathcal{E} = \{(-\infty, t) : t \in \mathbb{R}\} \cup \{(t, \infty) : t \in \mathbb{R}\}.$$

Show that $\mathcal{T} = \mathcal{T}(\mathcal{E})$.

It is often useful to describe topology not by giving *all* open sets but just some collection of open sets. This is similar to how open balls are used to describe all open sets for metric spaces.

Definition 1.17 (Base and neighborhood base)**.** (1) Let $(X, \mathcal{T})$ be a topological space and $x \in X$. A *neighborhood base for* $\mathcal{T}$ *at* x is a family $\mathcal{N}_x$ of subsets of X such that

(a) every $V \in \mathcal{N}_x$ is a neighborhood of x.

(b) if $U \in \mathcal{T}$ and $x \in U$, then there exists $V \in \mathcal{N}_x$ such that $V \subset U$.

(2) A *base* $\mathcal{B}$ for $\mathcal{T}$ is a family $\mathcal{B} \subset \mathcal{T}$ such that $\mathcal{B}$ contains a neighborhood base for $\mathcal{T}$ at x for all $x \in X$.

The following example illustrates the above definition.

Example 1.18. Let (X, d) be a metric space and let $\mathcal{T}$ be the topology induced by d on X (recall from Example 1.2-(2)).

(i) Let $x \in X$. Then $\mathcal{N}_x = \{B(x, r) : r > 0\}$ is a neighborhood base at x.

(ii) For any $x \in X$, $\{B(x, n^{-1}) : n \in \mathbb{N}\}$ is a neighborhood base at x.

(iii) $\mathcal{B} = \{B(y, r) : y \in X, r > 0\}$ is a base for $\mathcal{T}$.

Any collection of neighborhood base uniquely determines the topology as outlined in the following exercise.

Exercise 1.19. (a) Suppose that X is a set and for each $x \in X$, we are given a collection $\mathcal{N}_x$ of subsets of X satisfying:

(i) for all $V \in \mathcal{N}_x$, we have $x \in V$.

(ii) if $V_1, V_2 \in \mathcal{N}_x$, there exists $V_3 \in \mathcal{N}_x$ such that $V_3 \subset V_1 \cap V_2$.

(iii) for each $x \in X$, $\mathcal{N}_x \neq \emptyset$.

(iv) For each $U \in \mathcal{N}(x)$, there exists $V \subset X$ such that $x \in V \subset U$ and such that for every $y \in V$, there exists $W \in \mathcal{N}_y$ such that $W \subset V$.

Then show that there is a *unique* topology $\mathcal{T}$ on X such that $\mathcal{N}_x$ is a neighborhood base at x, for all $x \in X$.

(b) Conversely, if $(X, \mathcal{T})$ is a topological space and $\mathcal{N}_x$ is a neighborhood base at x for each $x \in X$, then show that the collection $\mathcal{N}_x$ of subsets of X satisfy the properties (i), (ii), (iii), (iv) above.

Here is a description of all open sets in terms of sets in the base.

Proposition 1.20. *Let $(X, \mathcal{T})$ is a topological space and $\mathcal{E} \subset \mathcal{T}$. Then $\mathcal{E}$ is a base for $\mathcal{T}$ if and only if every $U \in \mathcal{T}$ is a union of sets in $\mathcal{E}$.*

Proof. ($\Longrightarrow$): Let $\mathcal{E}$ be a base for $\mathcal{T}$ and $U \in \mathcal{T}$. If $x \in U$, there exists $V_x \in \mathcal{E}$ such that $x \in V_x \subset U$. Therefore $U = \bigcup_{x \in U} V_x$.
($\Longleftarrow$): For $x \in X$, we set $\mathcal{E}_x = \{V_x \in \mathcal{E} : x \in V\} \subset \mathcal{E}$. Let us check that $\mathcal{E}_x$ is a neighborhood base for $\mathcal{T}$ at x. Note that property (a) in Definition 1.17-(1) is true by the definition of $\mathcal{E}_x$. For (b), note that if $x \in U \in \mathcal{T}$, then $U = \bigcup_{\beta \in B} E_\beta$ for $E_\beta \in E$ for all $\beta \in B$ (since every $U \in \mathcal{T}$ is a union of sets in $\mathcal{E}$). So there exists $\beta_0 \in B$ such that $x \in E_{\beta_0}$. Thus $E_{\beta_0} \in \mathcal{E}_x$ and $E_{\beta_0} \subset U$. $\square$

Not every collection of subsets can be a base for a topology. The following proposition describes a necessary and sufficient condition for a collection of subsets to be a base for a topology.

Proposition 1.21. *Let $\mathcal{E} \subset \mathcal{P}(X)$. $\mathcal{E}$ is a base for a topology on X if and only if the following conditions are satisfied:*

(a) For all $x \in X$, there exists $V \in \mathcal{E}$ with $x \in V$.

(b) For all $U, V \in \mathcal{E}$ and for all $x \in U \cap V$, there exists $W \in \mathcal{E}$ such that $x \in W \subset U \cap V$.

Proof. ($\Longrightarrow$) (a) follows from the fact that $\mathcal{E}$ contains a neighborhood basis at x. For (b), note that $x \in U \cap V \in \mathcal{T}$, so there exists $W \in \mathcal{E}$ such that $x \in W \subset U \cap V$.
($\Longleftarrow$): Let $\mathcal{T} = \{U \subset X : \text{for each } x \in U, \text{there exists } V \in \mathcal{E} \text{ such that } x \in V \subset U\}$ (by Definition 1.17, this is the correct way to define the topology from the base). Let us verify that $\mathcal{T}$ is a topology on X. Note that $\varnothing, X \in \mathcal{T}$ (by (a)). If $\{U_\alpha : \alpha \in A\}$ is a collection of sets in $\mathcal{T}$, then $U = \bigcup_{\alpha \in A} U_\alpha$ also satisfies the definition condition for $\mathcal{T}$ and hence $U \in \mathcal{T}$. To show closure under finite intersections, it is enough to show that $U_1, U_2 \in \mathcal{T}$ implies $U_1 \cap U_2 \in \mathcal{T}$ (why?). To this end, let $U_1, U_2 \in \mathcal{T}$ and $x \in U_1 \cap U_2$. By the definition of $\mathcal{T}$, there exists $V_1, V_2 \in \mathcal{E}$ such that $x \in V_1 \subset U_1$ and $x \in V_2 \subset U_2$. By (b), there exists $W \in \mathcal{E}$ such that $x \in W \subset V_1 \cap V_2 \subset U_1 \cap U_2$. Therefore by the definition of $\mathcal{T}$, we have that $U_1 \cap U_2 \in \mathcal{T}$. Hence $\mathcal{T}$ is a topology on X and $\mathcal{E}$ is a base of $\mathcal{T}$ (by the definition of $\mathcal{T}$). □

The class of topological spaces are far too general. Topological spaces that occur in applications often satisfy additional conditions. We focus on two types of such conditions called *axioms of countability* and *axioms of separability*. We start with countability axioms for a topological space $(X, \mathcal{T})$ which imposes conditions on existence of countable base or neighborhood base (recall Definition 1.17).

Definition 1.22 (Countability axioms)**.** We say that $(X, \mathcal{T})$ is *first countable* if for each $x \in X$, there is a countable neighborhood base for $\mathcal{T}$ at x.
We say that $(X, \mathcal{T})$ is *second countable* if there is a countable base $\mathcal{B}$ for $\mathcal{T}$.

Obviously, every second countable space is first countable. Any metric space (X, d) is first countable (this follows from Example 1.18-(ii)).

Exercise 1.23. Find a first countable topological space that is not second countable.

First countable topological spaces can be understood using convergence properties of sequences.

Definition 1.24 (Convergence of a sequence)**.** Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X and let $x \in X$. We say that x_n converges to x as $n \to \infty$, if for any neighborhood U of x (or, equivalently, any *open* neighborhood of x), there exists $N \in \mathbb{N}$ such that $x_n \in U$ for all $n \geqslant N$. We denote this by $x_n \to \infty$ or $x_n \xrightarrow{n\to\infty} x$.

Question: Let $X = [0,1], \mathcal{T} = \{0, X\}$ and $x_n = \frac{1}{n}$ for all $n \in \mathbb{N}$. Which $y \in X$ does x_n converge to in the topological space $(X, \mathcal{T})$? **Hint**: Use the above definition.

The following exercise is meant to check that the notion of convergence in Definition 1.24 is a generalization of the usual definition in metric spaces.

Exercise 1.25. Let (X, d) be a metric space and let $\mathcal{T}$ denote the topology induced by the metric d. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X and let $x \in x$. Then the following are equivalent:

(a) $x_n \to x$ in the sense of Definition 1.24 on the topological space $(X, \mathcal{T})$.

(b) For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geqslant N$.

(c) $\lim_{n\to\infty} d(x_n, x) = 0$.

The following proposition describes the closure of a set in a first countable topological space using sequences. It is pleasing to know that the characterization of closure using limits of sequences for metric spaces also works for first countable topological spaces.

Proposition 1.26. *Let $(X, \mathcal{T})$ be a first countable space and $A \subset X$. Then $x \in \overline{A}$ if and only if there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in A such that $x_n \to x$.*

Proof. ($\Longleftarrow$): Suppose $x \notin \overline{A}$ and $(x_n)_{n\in\mathbb{N}}$. Then $V = (\overline{A})^c$ is an open neighborhood of x. Therefore $x_n \notin V$ for all $n \in \mathbb{N}$ (since $x_n \in A$ for all $n \in \mathbb{N}$ and $V \cap A = \varnothing$). Hence $(x_n)_{n\in\mathbb{N}}$ does not converge to x as $n \to \infty$. (Note that we don't need first countability for this implication. We only need it for the proof of the converse).
($\Longrightarrow$): Let $x \in \overline{A}$ and let $\{U_j : j \in \mathbb{N}\}$ be a countable neighborhood base for $\mathcal{T}$ at x. Set $V_n = \bigcap_{j=1}^n U_j$ for all $n \in \mathbb{N}$ (since finite intersections of neighborhoods of x is a neighborhood of x; why?). Then $x \in V_n$ and V_n is open for all $n \in \mathbb{N}$. Therefore $\{V_n : n \in \mathbb{N}\}$ is also a neighborhood base for $\mathcal{T}$ at x. Since $x \in \overline{A} = A \cup \operatorname{acc}(A)$ (see Proposition 1.11-(1)), we have $V_n \cap A \neq \varnothing$ for all $n \in \mathbb{N}$. Let $x_n \in V_n \cap A$ for all $n \in \mathbb{N}$.

We claim that (x_n) is the desired sequence that converges to x. To this end, let G be a open neighborhood of x. Since $\{V_n : n \in \mathbb{N}\}$ is a base, there exists $m \in \mathbb{N}$ such that $V_m \subset G$. Then $x_n \in V_n \subset V_m \subset G$ for all $n \geqslant m$. Hence $x_n \to x$. □

We say that $(X, \mathcal{T})$ is *separable* if there is a countable dense set (that is, there is a countable set $A \subset X$ such that $\overline{A} = X$).

Exercise 1.27. (1) If $(X, \mathcal{T})$ is second countable, then X is separable. (**Hint**: Pick one element from each set in a countable base to form a countable set and show that this set is dense).

(2) Suppose that $(X, \mathcal{T})$ is a topological space induced by a metric $d : X \times X \to [0, \infty)$ such that X is separable. Then show that $(X, \mathcal{T})$ is second countable. (**Hint**: Use a countable dense set to define a countable base).

Next, let us introduce the axioms of separation (usually denoted by T_0, T_1, T_2, T_3, T_4). The letter 'T' is due to the German term *Trennungsaxiom* which means *separation axiom*.

Definition 1.28 (Separation axioms)**.** (1) We say that $(X, \mathcal{T})$ is T_1 if for all $x, y \in X$ with $x \neq y$, there exists an open set $U_y \in \mathcal{T}$ such that $y \in U_y$ and $x \notin U_y$ (by symmetry, there exists $U_x \in \mathcal{T}$ such that $x \in U_x$ and $y \notin U_x$).

(2) We say that $(X, \mathcal{T})$ is T_2 (or *Hausdorff*) if for all $x, y \in X$ with $x \neq y$, there exist open sets U_x, U_y with $U_x \cap U_y = \varnothing, x \in U_x, y \in U_y$.

(3) We say that $(X, \mathcal{T})$ is T_3 (or *regular*) if $(X, \mathcal{T})$ is T_1 and if F is closed and $x \in F^c$, there exist disjoint open sets U, V with $F \subset U$ and $x \in V$.

(4) We say that $(X, \mathcal{T})$ is T_4 (or *normal*) if $(X, \mathcal{T})$ is T_1 and if F_1, F_2 are disjoint closed sets, there exist disjoint open sets U_1, U_2 with $F_1 \subset U_1$ and $F_2 \subset U_2$.

Singletons are closed in T_1 spaces as shown below.

Lemma 1.29. *$(X, \mathcal{T})$ is T_1 if and only if for all $x \in X$, $\{x\}$ is closed.*

Proof. $\Longrightarrow$: For $y \neq x$, let U_y be the open set with $y \in U_y, x \notin U_y$. Then $\bigcup_{y \in X \setminus \{x\}} U_y = X \setminus \{x\}$ is open, so $\{x\}$ is closed.
$\Longleftarrow$: If $x, y \in X$ with $x \neq y$, then $U_x = \{y\}^c, U_y = \{x\}^c$ satisfies the desired properties. □

Clearly every T_2 space is T_1. By Lemma 1.29, it follows that every T_4 space is T_3 and every T_3 space is T_2.
Sequences in Hausdorff (or T_2) spaces can have at most one limit.

Lemma 1.30. *Let $(X, \mathcal{T})$ be T_2 and $x, y \in X$ with $y \neq x$. If $x_n \to x$, then $x_n \not\to y$.*

Proof. Choose disjoint open sets U_x, U_y such that $x \in U_x, y \in U_y$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $x_n \in U_x$ for all $n \in \mathbb{N}$. Thus $x_n \notin U_y$ for all $n \geqslant N$ (since $U_x \cap U_y = \varnothing$). Therefore x_n does not converge to y. □

1.1 Continuous maps

We define the notion of continuity of a function between two topological spaces.

Definition 1.31. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$ be a function. We say that f is *continuous* if $f^{-1}(V)$ is open in X for all open sets V in Y; that is, $f^{-1}(V) \in \mathcal{T}_X$ for all $V \in \mathcal{T}_Y$.

Let $C(X, Y)$ denote the set to continuous maps from X to Y. If $Y = \mathbb{R}$ or $Y = \mathbb{C}$ with the usual topology induced by the Euclidean metric, then we abbreviate $C(X, Y)$ as $C(X)$.

Since closed sets are the complement of open sets and $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ for any $A \subset Y$, we have the following alternate criterion for continuity: a function $f : X \to Y$ is continuous if and only if $f^{-1}(A)$ is closed in X for any closed set A in Y.

The following the definition of continuity at a point.

Definition 1.32. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. Let $x \in X$ and $f : X \to Y$ be a function. We say that f is *continuous at* x if for any neighborhood V of $f(x)$ there exists a neighborhood U for x such that $f(U) \subset V$ (or equivalently, $f^{-1}(V)$ is a neighborhood of x).

Remark 1.33. If $f : X \to Y$ and $g : Y \to Z$ are continuous maps between topological spaces, then $h = g \circ f : X \to Z$ is continuous. **Proof**: Note that if V is an open subset of Z, then $h^{-1}(V) = f^{-1}(g^{-1}(V))$. Thus $g^{-1}(V)$ is open in Y (by continuity of g) and hence $f^{-1}(g^{-1}(V))$ is open in X by the continuity of f.

Exercise 1.34. Let $(X, \mathcal{T})$ be a topological space and let $f : X \to Y$ be a surjective (or onto) function. Then show that

$$\mathcal{T}_Y := \{U \subset Y | f^{-1}(U) \in \mathcal{T}\}$$

is a topology on Y (this is called the *quotient topology*). Furthermore, show that $\mathcal{T}_Y$ is the strongest (or finest) topology on Y such that f is a continuous function.

If X and Y are metric spaces, then Definitions 1.31 and 1.32 for the corresponding induces topologies are equivalent to the usual ϵ-δ definition. This is the content of the following exercise (you have likely encountered the equivalence between (b) and (c) below in an earlier course).

Exercise 1.35. Let $(X, d_X), (Y, d_Y)$ be metric spaces and let $\mathcal{T}_X, \mathcal{T}_Y$ denote the corresponding topologies induced by the metrics d_X, d_Y respectively. Let $f : X \to Y$ be a function and let $x \in X$. Then the following are equivalent:

(a) f is a continuous at x between the topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ in the sense of Definition 1.32.

(b) For any $\epsilon > 0$, there exists $\delta > 0$ such that if $x' \in X$ satisfies $d_X(x', x) < \delta$, then $d_Y(f(x'), f(x)) < \epsilon$.

(c) For any sequence $(x_n)_{n \in \mathbb{N}}$ such that $x_n \to x$, we have $f(x_n) \to f(x)$.

We confirm the familiar relationship between Definitions 1.31 and 1.32.

Proposition 1.36. *Let X, Y be topological spaces. Then $f : X \to Y$ is continuous if and only if $f : X \to Y$ is continuous at all $x \in X$.*

Proof. $\Longrightarrow$: Let $x \in V$ and V be a neighborhood of $f(x)$. Then there exists V' open in Y with $f(x) \in V' \subset V$. By the continuity of f, $f^{-1}(V')$ is open with $x \in f^{-1}(V')$. So $f^{-1}(V')$ is a neighborhood of x.
$\Longleftarrow$: Let $V \subset Y$ be open. If $x \in f^{-1}(V)$, then $f(x) \in V$, so V is a neighborhood of $f(x)$. By the continuity of f at x, $f^{-1}(V)$ is a neighborhood of x. So there exists an open set U_x such that $x \in U_x \subset f^{-1}(V)$. Therefore $x \in (f^{-1}(V))^\circ$. Hence $f^{-1}(V) \subset (f^{-1}(V))^\circ$ which implies that $f^{-1}(V)$ is open. □

Proposition 1.37. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces such that $\mathcal{T}_Y = \mathcal{T}(\mathcal{E})$ where $\mathcal{E} \subset \mathcal{P}(Y)$. Then $f : X \to Y$ is continuous if and only if $f^{-1}(E) \in \mathcal{T}_X$ for all $E \in \mathcal{E}$.*

Proof. Since $\mathcal{E} \subset \mathcal{T}_Y$, the 'only if' part follows immediately from the continuity of f.

For the converse, we use (1.1) along with $f^{-1}(\bigcap_{i=1}^n E_i) = \bigcap_{i=1}^n f^{-1}(E_i)$, $f^{-1}(\bigcup_{\alpha\in A} E_\alpha) = \bigcup_{\alpha\in A} f^{-1}(E_\alpha)$ and Definition 1.1. □

Definition 1.38 (Homeomorphism)**.** Let $f : X \to Y$ be a bijection (one-to-one and onto) between topological spaces such that f and $f^{-1} : Y \to X$ are continuous. Then we say that f is a *homeomorphism* between X and Y.

We say that two topological spaces X and Y are *homoemorphic* if there exists a homeomorphism f between X and Y.

A *topological property of a space* is one which is preserved under homeomorphisms.

Example 1.39. (1) $\mathbb{R}$ and $(0, 1)$ are homeomorphic. To see this, note that $x \mapsto \tan x$ is a homeomorphism between $(-\pi/2, \pi/2)$ to $\mathbb{R}$. Therefore $f : (0, 1) \to \mathbb{R}$ defined by $f(x) := \tan\left(\pi\left(x - \frac{1}{2}\right)\right)$ is one-to-one, onto and continuous such that f^{-1} is also continuous.

(2) The relation of being homeomorphic forms an equivalence relation among topological spaces. That is, every topological space is homeomorphic to itself (reflexive). If X and Y are homeomorphic, then Y and X are homeomorphic (symmetric). If X and Y are homeomorphic and Y and Z are homeomorphic, then X and Z are homeomorphic.

(3) $(0, 1)$ and $[0, 1]$ are not homeomorphic (with respect to the topologies induced by the usual metric). This is not clear now but we will see why by finding a suitable topological property which $[0, 1]$ has but $(0, 1)$ does not.

(4) The property of being Hausdorff (or T_2) is preserved under homeomorphism and hence is a topological property. This is also true for other axioms of separation and axioms of countability.

The following exercise outlines an argument that metric spaces are normal.

Exercise 1.40. Let (X, d) be a metric space and let $\mathcal{T}$ denote the topology induced by the metric d.

(a) If $F \subset X$ is closed, then show that the function $d_F : X \to \mathbb{R}$ defined by

$$d_F(x) = \inf\{d(x, y) : y \in F\}$$

is a continuous function such that $d_F(x) = 0$ if and only if $x \in F$. **Hint**: Show that $|d_F(x_1) - d_F(x_2)| \leqslant d(x_1, x_2)$ for all $x_1, x_2 \in X$ and use this to prove continuity.

(b) If F_1, F_2 are disjoint closed sets, show that the function $g : X \to \mathbb{R}$ defined by

$$g(x) = d_{F_1}(x) - d_{F_2}(x), \quad \text{for all } x \in X,$$

where $d_{F_1}, d_{F_2} : X \to \mathbb{R}$ is as given in (a) satisfies the following properties: g is continuous, $g(x) > 0$ for all $x \in F_2$ and $g(x) < 0$ for all $x \in F_1$.

(c) Using the function in (b), show that $(X, \mathcal{T})$ is T_4. **Hint**: Consider $g^{-1}(U)$ for suitable open sets U.

1.2 Weak and product topologies

We already saw one way to construct new topological spaces from old; namely, induced topology (see Definition 1.3). We will see two more constructions: weak topology and product topology. Weak topology is defined so as to *make* a collection of maps continuous.

Definition 1.41 (weak topology)**.** Let X be a set and $(Y_\alpha, \mathcal{T}_\alpha), \alpha \in A$ be a collection of topological spaces. Let $f_\alpha : X \to Y_\alpha, \alpha \in A$ be a collection of functions (usually, $Y_\alpha = \mathbb{R}$ or $\mathbb{C}$). Let $\mathcal{T}$ be the weakest (or coarsest) topology on X which makes all the functions f_α continuous. Then $\mathcal{T}$ is called the *weak topology on X generated by* $\{f_\alpha : \alpha \in A\}$.

Equivalently, $\mathcal{T}$ is the topology generated by $\{f_\alpha^{-1}(U_\alpha) : \alpha \in A, U_\alpha \in \mathcal{T}_\alpha\}$. A base for $\mathcal{T}$ is given by (check this using Proposition 1.21!)

$$\left\{ \bigcap_{i=1}^{n} f_{\alpha_i}^{-1}(U_{\alpha_i}) \middle| \alpha_i \in A, U_{\alpha_i} \in \mathcal{T}_{\alpha_i} \text{ for all } i = 1, \ldots, n \right\}.$$

In the following exercise, we see that the relative topology can be viewed as a special case of weak topology.

Exercise 1.42. Let $(X, \mathcal{T})$ be a topological space and let $Y \subset X$ be a non-empty subset of X. Let $\iota : Y \to X$ denote the inclusion map $\iota(y) = y$ for all $y \in Y$. Then show that the weak topology on Y generated by $\{\iota\}$ is the relative topology on Y (recall Definition 1.3).

The following exercise outlines an useful condition for continuity of functions whose target space is equipped with weak topology (cf. Proposition 1.37).

Exercise 1.43. Let X be a set and $(Y_\alpha, \mathcal{T}_\alpha), \alpha \in A$ be a collection of topological spaces. Let $f_\alpha : X \to Y_\alpha, \alpha \in A$ be a collection of functions and let X be equipped with the weak topology on X generated by $\{f_\alpha : \alpha \in A\}$. Let Z be another topological space and let $g : Z \to X$ be a function (not necessarily continuous). Then the following are equivalent:

(a) $g : Z \to X$ is continuous.

(b) $f_\alpha \circ g : Z \to Y_\alpha$ is continuous for all $\alpha \in A$.

Next, we would like to define a topology on product of topological spaces. Let $X_\alpha, \alpha \in A$ be a collection of sets. The product space (as a set) $\prod_{\alpha\in A} X_\alpha$ is defined as

$$\prod_{\alpha} X_\alpha = \{f : A \to \bigcup_{\alpha} X_\alpha : f(\alpha) \in X_\alpha \text{ for all } \alpha \in A\}. \tag{1.2}$$

This is consistent with the previous definition of products you may have seen. For example, elements of $\mathbb{R}\times\mathbb{R}$ are usually denotes as ordered pairs (x, y), which is equivalent to representation as a function $f : \{1, 2\} \to \mathbb{R}$ which corresponds to the ordered pair $(f(1), f(2))$. If $X_\alpha = Y$ for all $\alpha \in A$, we abbreviate $\prod_\alpha X_\alpha$ as Y^A. Similarly, if $X_\alpha = Y$ for all $\alpha \in A$ and $A = \{1, 2, \ldots, n\}$, we denote $\prod_\alpha X_\alpha$ as Y^n, which can also be though of as ordered n-tuples of elements of Y.

For each $\alpha_1 \in A$, we define projection maps (or coordinate maps) $\pi_{\alpha_1} : \prod_{\alpha\in A} X_\alpha \to X_{\alpha_1}$ by

$$\pi_{\alpha_1}(f) = f(\alpha_1), \quad \text{for all } f \in \textstyle\prod_{\alpha\in A} X_\alpha. \tag{1.3}$$

Question: If $X_\alpha \neq \varnothing$ for all $\alpha \in A$, then is $\prod_{\alpha\in A} X_\alpha \neq \varnothing$? Intuitively, we can choose $f : A \to \bigcup_{\alpha\in A} X_\alpha$ as $f(\alpha) = x_\alpha$ where x_α is an arbitrary element of X_α for each $\alpha \in A$. This gives $f \in \prod_{\alpha\in A} X_\alpha$ and hence $\prod_{\alpha\in A} X_\alpha \neq \varnothing$. However, this 'proof' does not follow from the usual axioms of set theory and this property has to be assumed as an additional axiom called the *axiom of choice*: if $X_\alpha \neq \varnothing$ for all $\alpha \in A$, then $\prod_{\alpha\in A} X_\alpha \neq \varnothing$. We will always assume that axiom of choice holds.

Now, that we have defined the product space $\prod_{\alpha\in A} X_\alpha$ as a set, we need to define a topology on it.

Definition 1.44 (Product topology)**.** Let $(X_\alpha, \mathcal{T}_\alpha), \alpha \in A$ be a family of topological spaces. Then the product topology $\mathcal{T}$ on $X := \prod_{\alpha\in A} X_\alpha$ is the weak topology on X generated by the projection maps $\{\pi_\alpha : X \to X_\alpha | \alpha \in A\}$ (see (1.2)). So a base for $\mathcal{T}$ is given by

$$\mathcal{B} = \left\{ \bigcap_{i=1}^{n} \pi_{\alpha_i}^{-1}(U_{\alpha_i}) \middle| \alpha_i \in A, U_{\alpha_i} \in \mathcal{T}_{\alpha_i} \text{ for all } i = 1, \ldots, n \right\}. \tag{1.4}$$

Warning: It is *not true in general* that if $U_\alpha \in \mathcal{T}_\alpha$ for all $\alpha \in A$, then $\prod_{\alpha\in A} U_\alpha \in \mathcal{T}$. That is, product of open sets need not be open in the product topology. However it is true if A is finite.

Product topology preserves some topological properties of the components.

Proposition 1.45. *If each $X_\alpha, \alpha \in A$ is a Hausdorff topological space, then $\prod_{\alpha\in A} X_\alpha$ equipped with the product topology $\mathcal{T}$ is Hausdorff.*

Proof. Let $f, g : A \to \bigcup_\alpha X_\alpha$ belong to $X = \prod_{\alpha\in A} X_\alpha$ such that $f \neq g$. Then there exists $\alpha_0 \in A$ such that $f(\alpha_0) \neq g(\alpha_0)$. Since $(X_{\alpha_0}, \mathcal{T}_{\alpha_0})$ is Hausdorff, there exist $U_{f(\alpha_0)}, U_{g(\alpha_0)} \in \mathcal{T}_\alpha$ such that $U_{f(\alpha_0)} \cap U_{g(\alpha_0)} = \varnothing, f(\alpha_0) \in U_{f(\alpha_0)}, g(\alpha_0) \in U_{g(\alpha_0)}$. Therefore $\pi_{\alpha_0}^{-1}(U_{f(\alpha_0)}), \pi_{\alpha_0}^{-1}(U_{g(\alpha_0)}) \in \mathcal{T}$ and $f \in \pi_{\alpha_0}^{-1}(U_{f(\alpha_0)}), g \in \pi_{\alpha_0}^{-1}(U_{g(\alpha_0)})$ with $\pi_{\alpha_0}^{-1}(U_{f(\alpha_0)}) \cap \pi_{\alpha_0}^{-1}(U_{g(\alpha_0)}) = \varnothing$. Therefore $(X, \mathcal{T})$ is Hausdorff. $\square$

We say that a sequence (f_n) of functions in X^A *converges pointwise* to $f : X \to A$, if for each $\alpha \in A$, $(f_n(\alpha))$ converges to $f(\alpha)$ in $(X, \mathcal{T})$ (see Definition 1.24). The product topology is the topology that corresponds to pointwise convergence of functions as we see below (see also Exercise 1.60).

Proposition 1.46. *Let $(X, \mathcal{T})$ be a topological space and let A be a set. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in X^A and let $f : A \to X$. Then the following are equivalent:*

(a) $f_n \to f$ in X^A with respect to product topology.

(b) f_n converges pointwise to f.

Proof. (a) $\implies$ (b): Fix $\alpha \in A$. Let U be a open neighborhood of $f(\alpha) \in A$. Then $V = \pi_\alpha^{-1}(U)$ is an open subset of f in X^A. Since $f_n \to f$ in X^A, there exists $N \in \mathbb{N}$ such that $f_n \in V$ for all $n \geqslant N$. Therefore $f_n(\alpha) \in U$ for all $n \geqslant N$; that is, $f_n(\alpha) \to f(\alpha)$ in $(X, \mathcal{T})$.
(b) $\implies$ (a): Let W be an open neighborhood of f in X^A. Since a base for product topology is given by (1.4), there exist $\alpha_1, \ldots, \alpha_k \in A$ and $U_{\alpha_i} \in \mathcal{T}$ for all $1 \leqslant i \leqslant k$ such that $f \in V$ and $V \subset W$, where

$$V = \{g \in X^A : g(\alpha_i) \in U_{\alpha_i} \quad \text{for all } 1 \leqslant i \leqslant k\}.$$

Since $f \in V$, $f(\alpha_i) \in U_{\alpha_i}$ for all $1 \leqslant i \leqslant k$. Since $f_n(\alpha_i) \to f(\alpha_i)$ as $n \to \infty$ for each $1 \leqslant i \leqslant k$, there exists $N_1, \ldots, N_n \in \mathbb{N}$ such that $f_n(\alpha_i) \in U_{\alpha_i}$ for all $n \geqslant N_i$ and for all $1 \leqslant i \leqslant k$. Therefore by letting $N = \max_{1\leqslant i\leqslant k} N_i$, we have $f_n \in V \subset W$ for all $n \geqslant N$. So $f_n \to f$ in X^A. ☐

The argument in the proof of Proposition 1.46 also works if all X_α are different spaces for different values of $\alpha \in A$.

The following exercise clarifies an useful relation between weak and subspace topologies.

Exercise 1.47. Let X be a set and $(Y_\alpha, \mathcal{T}_\alpha), \alpha \in A$ be a collection of topological spaces. Let $f_\alpha : X \to Y_\alpha, \alpha \in A$ be a collection of functions. Let $\mathcal{T}$ denote the weak topology on X generated by $\{f_\alpha : \alpha \in A\}$. Let $W \subset X$ be non-empty. We have two possible ways to define a topology on W.

(1) Let $\mathcal{S}$ denote the subpace (or induced) topology on W of the weak topology $\mathcal{T}$ generated by $\{f_\alpha : \alpha \in A\}$.

(2) Let $\tilde{\mathcal{S}}$ denote weak topology on W generated by $\{g_\alpha : \alpha \in A\}$, where $g_\alpha := f_\alpha\big|_W$ is the restriction of f_α to W for all $\alpha \in A$.

Show that

$$\mathcal{S} = \tilde{\mathcal{S}}.$$

In words, the above equality can be paraphrased as follows: *weak topology on subspace is equal to the subspace topology of weak topology.*

1.3 Urysohn's lemma and Tietze extension theorem

The next few results say that normal (or T_4) topological spaces have lots of continuous functions. We begin with a technical lemma that is key to construction of continuous functions.

Lemma 1.48. *Let X be a normal topological space and let A, B be disjoint closed sets. Let $\mathbb{D} = \{k2^{-n} : k, n \in \mathbb{N}, 1 \leqslant k \leqslant 2^n - 1\} \subset (0,1)$. Then there exists a family $\{U_\rho : \rho \in \mathbb{D}\}$ of open sets such that for all $0 < r < s < 1$ with $r, s \in \mathbb{D}$, we have*

$$A \subset U_r \subset \overline{U_r} \subset U_s \subset B^c.$$

Proof. As A, B are closed, there exists open U, V with $A \subset U, B \subset V, U \cap V = \varnothing$. Therefore

$$A \subset U \subset \overline{U} \subset V^c \subset B^c.$$

Set $U_{1/2} = U$.

Now, we can proceed by induction by the same argument as above. For, this purpose, set $U_0 = \overline{U_0} = A$ and $U_1 = B^c$. Suppose $U_{k2^{-n}} \subset \overline{U_{k2^{-n}}} \subset U_{(k+1)2^{-n}}$ is given for all $0 \leqslant k \leqslant 2^n - 1$, where $U_{(k+1)2^{-n}}$ is open. Then $\overline{U_{k2^{-n}}}$ and $U^c_{(k+1)2^{-n}}$ are disjoint closed sets. Therefore by the argument above, there exists an open set U such that

$$\overline{U_{k2^{-n}}} \subset U \subset \overline{U} \subset U_{(k+1)2^{-n}}.$$

We set $U_{(2k+1)2^{-(n+1)}} = U$. This completes the induction step. □

The family of open sets in Lemma 1.48 is useful to construct continuous functions. (Compare the proof below with the construction in Exercise 1.40-(b)).

Theorem 1.49 (Urysohn's lemma)**.** *Let X be a normal topological space and let A, B be disjoint closed sets. Then there exists a bounded, continuous function $f : X \to [0,1]$ such that $f \equiv 0$ on A and $f \equiv 1$ on B.*

Proof. Let $\{U_\rho : \rho \in \mathbb{D}\}$ be the family of open sets as given in Lemma 1.48. Set $U_0 = A, U_1 = X$. Define $f : X \to \mathbb{R}$ as

$$f(x) := \inf\{\rho \in \mathbb{D} \cup \{0,1\} : x \in U_\rho\}. \tag{1.5}$$

We claim that f satisfies the desired properties. Clearly $0 \leqslant f(x) \leqslant 1$ for all $x \in X$, $f \equiv 0$ on A and $f \equiv 1$ on B^c.

It remains to show that f is continuous. So it is enough to prove that $f^{-1}((-\infty, t))$ and $f^{-1}((t, \infty))$ are open for all $t \in \mathbb{R}$ (Do you see why? If not, review Proposition 1.37).

Let us first consider $f^{-1}((-\infty, t))$ for $t \in \mathbb{R}$. If $t > 1$ (resp. $t \leqslant 0$), $f^{-1}((-\infty, t))$ is X (resp. $\varnothing$) and hence open. It suffices to assume $t \in (0,1]$. Note that $f(x) < t$ if and only if there exists $\rho < t, \rho \in \mathbb{D}$ such that $x \in U_\rho$. Therefore

$$f^{-1}((-\infty, t)) = \{x : f(x) < t\} = \bigcup_{\substack{\rho < t, \\ \rho \in \mathbb{D}}} U_\rho,$$

which is open (being an union of open sets).

Next, we need to show that $f^{-1}((t,\infty))$ is open for all $t \in \mathbb{R}$. If $t < 0$ (resp. $t \geqslant 1$), then $f^{-1}((t,\infty))$ is X (resp. $\varnothing$). So it suffices to assume $t \in [0,1)$. Note that $f(x) > t$ if and only if $x \notin U_\rho$ for some $\rho > t, \rho \in \mathbb{D}$ which in turn holds if and only if $x \notin \overline{U_q}$ for some $q > t, q \in \mathbb{D}$ (since $\overline{U_q} \subset U_\rho$ for all $t < q < \rho$). Therefore

$$f^{-1}((t,\infty)) = \bigcup_{\substack{\rho > t,\\ \rho\in\mathbb{D}}} U_\rho^c = \bigcup_{\substack{q > t,\\ q\in\mathbb{D}}} (\overline{U_q})^c$$

which is open (being an union of open sets). □

Notation: Let $B(X,\mathbb{R})$ denote the set of *bounded functions* $f : X \to \mathbb{R}$.
Let $BC(X,\mathbb{R})$ denote the set of *bounded and continuous functions* $f : X \to \mathbb{R}$.
Let $d_\infty(f,g) = \sup_{x\in X} |f(x) - g(x)|$ for all $f, g \in B(X,\mathbb{R})$.

Lemma 1.50. *(i) $(B(X,\mathbb{R}), d_\infty)$ is a complete metric space.*

(ii) $BC(X,\mathbb{R})$ is a closed subspace of $(B(X,\mathbb{R}), d_\infty)$.

Proof. (i) Let f_n be a Cauchy sequence in $B(X,\mathbb{R})$. Note that $|f_n(x) - f_m(x)| \leqslant d_\infty(f_n, f_m)$ for all $m, n \in \mathbb{N}$, for each $x \in X$. So for any $x \in X$, $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, and converges to a limit, say $f(x) \in \mathbb{R}$. Since (f_n) is Cauchy, for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(x) - f_m(x)| < \epsilon$ for all $m, n \geqslant N$ and all $x \in X$. Let $m \to \infty$, to obtain $\sup_{x\in X} |f_n(x) - f(x)| \leqslant \epsilon$ for all $n \in \mathbb{N}$ with $n \geqslant N$. Therefore $\lim_{n\to\infty} d_\infty(f_n, f) = 0$, $\sup_{x\in X} |f(x)| \leqslant \sup_{x\in X} |f_N(x)| + \epsilon < \infty$ and hence $f \in B(X,\mathbb{R})$.

(ii) Let $f_n \in BC(X,\mathbb{R})$, $f_n \to f$ in $(B(X,\mathbb{R}), d_\infty)$; that is $\lim_{n\to\infty} d_\infty(f_n, f) = 0$.

By (i), it suffices show that $f \in C(X,\mathbb{R})$. To this end, let $x \in X, \epsilon > 0$. Let $N \in \mathbb{N}$ be such that $d_\infty(f_n, f) < \epsilon/3$ for all $n \geqslant N$. Using the continuity of f_N at x, choose a neighborhood U of x such that $|f_N(y) - f_N(x)| < \epsilon/3$ for all $y \in U$. Therefore, for all $y \in U$, we have

$$\begin{aligned} |f(y) - f(x)| &\leqslant |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| \\ &\leqslant 2d_\infty(f_N, f) + |f_N(y) - f_N(x)| < \epsilon. \end{aligned}$$

Therefore, f is continuous at x.

□

By Lemma 1.50, the space $(BC(X,\mathbb{R}), d_\infty)$ is a complete metric space.

Our next goal is to prove Tietze extension theorem.

Theorem 1.51 (Tietze extension theorem)**.** *Let X be a normal topological space and let $A \subset X$ be closed and $f \in C(A,[a,b])$. There exists $F \in C(X,[a,b])$ such that $F\big|_A = f$; that is $F(x) = f(x)$ for all $x \in A$.*

Remark 1.52. The assumption that A is closed is necessary. For example, if $X = [0,1], A = (0,1], f(x) = \sin(1/x)$ for all $x \in A$, then there is no continuous extension F of f to X.

The proof of Tietze extension theorem relies on repeated use of the following lemma.

Lemma 1.53. *Let X be a normal topological space and let $A \subset X$ be closed and $h \in C(A,[0,\lambda])$, where $\lambda > 0$. Then there exists $g \in C(X,[0,\lambda/3])$ such that $g \leqslant h \leqslant g + \frac{2}{3}\lambda$ on A.*

Proof. Set $F_1 = h^{-1}([0,\lambda/3])$ and $F_2 = h^{-1}([2\lambda/3,\lambda])$. By the continuity of h, F_1 and F_2 are closed subsets of A (with respect to the relative topology). Since A is closed, F_1 and F_2 are closed in X (verify this!). By Urysohn's lemma (Theorem 1.49), there exists $\tilde{g} \in C(X,[0,1])$ such that $g \equiv 0$ on F_1 and $g \equiv 1$ on F_2. Set $g = \frac{\lambda}{3}\tilde{g} \in C(X,[0,\lambda/3])$. Then we have

$$\begin{aligned} g(x) = 0, \quad & 0 \leqslant h(x) \leqslant \frac{\lambda}{3}, \quad && \text{for all } x \in F_1, \\ g(x) = \frac{\lambda}{3}, \quad & \frac{2}{3}\lambda \leqslant h(x) \leqslant \lambda, \quad && \text{for all } x \in F_2, \\ 0 \leqslant g(x) \leqslant \frac{\lambda}{3}, \quad & \frac{\lambda}{3} \leqslant h(x) \leqslant \frac{2\lambda}{3}, \quad && \text{for all } x \in A\backslash(F_1 \cup F_2). \end{aligned}$$

Combining the three cases, we obtain $g \leqslant h \leqslant g + \frac{2}{3}\lambda$ on A. □

Proof of Theorem 1.51. By replacing f with $(f-a)/(b-a)$ if necessary, we may assume $[a,b] = [0,1]$.

Use Lemma 1.53 with $h = f, \lambda = 1$, to obtain g_1 such that $g_1 \in C(X,[0,1/3])$, $g_1 \leqslant f \leqslant \frac{2}{3} + g_1$ on A; that is, $0 \leqslant f - g_1 \leqslant \frac{2}{3}$ on A.

Use Lemma 1.53 again with $h = f - g_1, \lambda = 2/3$ to obtain $g_2 \in C(X,[0,2/3^2])$ with $0 \leqslant f - g_1 - g_2 \leqslant (2/3)^2$ on A.

Continuing, we obtain $g_1, g_2, \ldots, g_n, \ldots$ such that $0 \leqslant g_n \leqslant \frac{2^{n-1}}{3^n}, 0 \leqslant f - \sum_{k=1}^n g_k \leqslant \left(\frac{2}{3}\right)^n$ on A. Therefore $F_n := \sum_{k=1}^n g_k$ converges in $(BC(X,\mathbb{R}), d_\infty)$ to a continuous function F (by Lemma 1.50) as $n \to \infty$ such that $F = f$ on A. □

1.4 Nets

Nets are generalizations of sequences. Sequences are indexed by $\mathbb{N}$ while nets are indexed by a (possibly uncountable) directed set.

Definition 1.54 (Directed set)**.** A directed set A is a set with a relation $\leqslant$ such that

(i) (reflexive) $\alpha \leqslant \alpha$ for all $\alpha \in A$.

(ii) (transitive) $\alpha \leqslant \beta$ and $\beta \leqslant \gamma$ implies $\alpha \leqslant \gamma$ for all $\alpha, \beta, \gamma \in A$.

(iii) (existence of arbitrarily large elements) For any $\alpha, \beta \in A$, there exists $\gamma \in A$ such that $\alpha \leqslant \gamma$ and $\beta \leqslant \gamma$.

Notation: We will denote $\alpha \leqslant \beta$ also as $\beta \geqslant \alpha$.

Example 1.55. (1) $(\mathbb{N}, \leqslant)$ is a directed set, where $\leqslant$ has the usual meaning; that is $a \leqslant b$ if and only if $b - a$ is non-negative.

(2) $\mathbb{Z}^d$ with relation $\leqslant$ defined by $(x_1, \ldots, x_d) \leqslant (y_1, \ldots, y_d)$ if $x_i \leqslant y_i$ for all $i = 1, \ldots, d$.

(3) This is the most important example for us. Let $(X, \mathcal{T})$ be a topological space and let $x \in X$. Let $\mathcal{N}_x$ denote the set of neighborhoods of x. We define $U \leqslant V$ for $U, V \in \mathcal{N}_x$ if $V \subset U$ (order by *reverse inclusion*).

(4) Let $(A, \leqslant_A), (B, \leqslant_B)$ be two directed sets. Then the product $A \times B$ equipped with the relation $\leqslant$ defined by $(a_1, b_1) \leqslant (a_2, b_2)$ if and only if $a_1 \leqslant_A a_2$ and $b_1 \leqslant_B b_2$ is a directed set.

Exercise 1.56. Verify that each of the directed sets in Example 1.55 satisfies the properties in Definition 1.54.

Since nets generalize sequences (due to Example 1.55-(1)), we need to define a suitable notion of convergence (or limits) of nets. Cluster points are generalization of subsequential limits.

Definition 1.57. A net in X is a function $\alpha : A \to X$; often denoted as $\alpha \mapsto x_\alpha$ or $(x_\alpha)_{\alpha \in A}$, where $(A, \leqslant)$ is a directed set.

Let $(X, \mathcal{T})$ be a topological space and let $(x_\alpha)_{\alpha \in A}$ be a net.

1. We say that $(x_\alpha)_{\alpha \in A}$ *converges to* $x \in X$ if for every neighborhood U of x, there exists $\alpha_0 \in A$ such that $x_\beta \in U$ for all $\beta \geqslant \alpha_0$. We denote this by $x_\alpha \to x$.

2. We say that $x \in X$ is a *cluster point* of $(x_\alpha)_{\alpha \in A}$ if for every neighborhood U of x and for any $\alpha \in A$, there exists $\beta \geqslant \alpha$ such that $x_\beta \in U$.

Let us recall that points in the closure of a set can be characterized as sequential limits for first countable topological spaces (see Proposition 1.26). However, if the topological space is not first countable, not every point in the closure is a sequential limit as given in Proposition 1.26 (you will encounter such an example in Assignment). The following extension of Proposition 1.26 is one of the motivations behind studying nets.

Proposition 1.58. *Let $(X, \mathcal{T})$ be a topological space, $E \subset X$, and $x \in X$.*

(a) $x \in \operatorname{acc}(E)$ if and only if there exists a net in $E \backslash \{x\}$ which converges to x.

(b) $x \in \overline{E}$ if and only if there exists a net in E which converges to x.

Proof. We only prove (a) as (b) is very similar.
$\Longrightarrow$: Let $(\mathcal{N}_x, \leqslant)$ denote the directed set in Example 1.55-(3). For $U \in \mathcal{N}_x$, since x is an accumulation point of E, there exists $x_U \in (U\backslash\{x\}) \cap E \neq \varnothing$. So $(x_U)_{U \in \mathcal{N}_x}$ is a net in $E\backslash\{x\}$. For any $U \in \mathcal{N}_x$, and any $V \in \mathcal{N}_x$ such that $U \leqslant V$, we have $x_V \in V \subset U$, so $x_U \to x$.
$\Longleftarrow$: Let $x_\alpha \to x$, where $(x_\alpha)_{\alpha \in A}$ is a net in $E\backslash\{x\}$. Given any $U \in \mathcal{N}_x$, by the convergence of (x_α) to x, there exists $\alpha \in A$ such that $x_\alpha \in U \cap (E\backslash\{x\})$. So $U \cap (E\backslash\{x\}) \neq \varnothing$. Since U is an arbitrary neighborhood of x, $x \in \mathrm{acc}(E)$. □

Recall that for a map $f : (X, d_X) \to (Y, d_Y)$ between metric spaces, f is continuous at $x \in X$ if and only if for any sequence $(x_n)_{n \in \mathbb{N}}$ converging to x, we have that the sequence $(f(x_n))_{n \in \mathbb{N}}$ converges to $f(x)$ (see Exercise 1.35). The following is an extension of that statement to arbitrary topological spaces.

Proposition 1.59. *Let $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ be topological spaces, $x \in X$, and $f : X \to Y$ be a function. Then f is continuous at x if and only if for every net (x_α) that converges to x, we have that $(f(x_\alpha))$ converges to $f(x)$.*

Proof. $\Longrightarrow$: (this implication is shown exactly as in the sequence case) Let $x_\alpha \to x$. Let V be a neighborhood of $f(x)$. By the continuity of f, $f^{-1}(V)$ is a neighborhood of x. As $x_\alpha \to x$, there exists α_0 such that $x_\alpha \in f^{-1}(V)$ for all $\alpha \geqslant \alpha_0$. So $f(x_\alpha) \in V$ for all $\alpha \geqslant \alpha_0$ and hence $f(x_\alpha) \to f(x)$.
$\Longleftarrow$: (Contrapositive) Suppose f is not continuous at x. There is a neighborhood V of $f(x)$ such that $f^{-1}(V)$ is not a neighborhood of x; that is $x \notin (f^{-1}(V))^\circ$. By Lemma 1.8-(5), $x \in ((f^{-1}(V))^\circ)^c = \overline{f^{-1}(V)^c} = \overline{f^{-1}(V^c)}$. By Proposition 1.58, there exists a net $(x_\alpha)_{\alpha \in A}$ in $f^{-1}(V^c)$ such that $x_\alpha \to x$. Hence $f(x_\alpha) \in V^c$ for all $\alpha \in A$, so $f(x_\alpha) \not\to f(x)$ (since V is a neighborhood of $f(x)$). □

The following exercise is an application of Proposition 1.59. It provides an alternate description of limit of nets in weak topology.

Exercise 1.60. Let X be a set and $(Y_i, \mathcal{T}_i), i \in I$ be a collection of topological spaces. Let $f_i : X \to Y_i, i \in I$ be a collection of functions. Let X be equipped with the weak topology $\mathcal{T}$ generated by $\{f_i : i \in I\}$. Let $(x_\alpha)_{\alpha \in A}$ be a net in X and let $x \in X$. Then show that the following are equivalent.

(a) $(x_\alpha)_{\alpha \in A}$ converges to x in $(X, \mathcal{T})$.

(b) $(f_i(x_\alpha))_{\alpha \in A}$ converges to $f_i(x)$ in $(Y_i, \mathcal{T}_i)$ for each $i \in I$.

1.5 Compact sets

Compactness is an important topological property.

Definition 1.61. We say that $(X, \mathcal{T})$ is a compact topological space if whenever $(U_i)_{i\in I}$ is an open cover of X (that is, $U_i \in \mathcal{T}$ for all $i \in I$ and $\bigcup_{i\in I} U_i = X$), then there exists a finite subcover (that is, there exists a finite subset $\{i_1, \ldots, i_k\}$ of I such that $\bigcup_{j=1}^k U_{i_j} = X$).

We say a subset $A \subset X$ is compact if it is compact in the relative topology.

Here is an equivalent description of compactness of a subset $A \subset X$. If (V_α) is an open cover of A in X (that is $V_\alpha \in \mathcal{T}$ for all α and $\bigcup_\alpha V_\alpha \supset A$), there exists ia finite subcover of A (that is $\alpha_1, \ldots, \alpha_k$ such that $\bigcup_{i=1}^k V_{\alpha_i} \supset A$). To see the equivalence, note that $U_\alpha = A \cap V_\alpha$ are open in the relative topology.

Several familiar results for compact subsets of a metric space are also true more generally for topological spaces. For instance, closed subset of a compact space is compact.

Proposition 1.62. *Let $(X, \mathcal{T})$ be a compact topological space and let $F \subset X$ be closed. Then F is compact.*

Proof. Let $F \subset \bigcup_\alpha W_\alpha$, where $W_\alpha \in \mathcal{T}$ for all α. Then $X \subset F^c \cup \bigcup_\alpha W_\alpha$, so by compactness of X, there exists $\alpha_1, \ldots, \alpha_n$ such that

$$X \subset F^c \cup \bigcup_{i=1}^n W_{\alpha_i}.$$

Then $F \subset \bigcup_{i=1}^n W_{\alpha_i}$. $\square$

Compactness of a set can be characterized using finite intersection property.

Definition 1.63 (Finite intersection property)**.** Let $(F_\alpha)_{\alpha\in A}$ be a collection of sets. We say that $(F_\alpha)_{\alpha\in A}$ has the *finite intersection property* (abbreviated as FIP) if for any $n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in A$, we have $\bigcap_{i=1}^n F_{\alpha_i} \neq \varnothing$.

Proposition 1.64. *X is compact if and only if whenever $(F_\alpha)_{\alpha\in A}$ is a collection of closed sets satisfying the finite intersection property, then $\bigcap_{\alpha\in A} F_\alpha \neq \varnothing$.*

Proof. $\Longrightarrow$: Let X be compact and let $(F_\alpha)_{\alpha\in A}$ be a collection of closed sets satisfying the FIP. Assume to the contrary that $\bigcap_{\alpha\in A} F_\alpha = \varnothing$. Then $(F_\alpha^c)_{\alpha\in A}$ is an open cover of X, which given as finite cover $\bigcup_{i=1}^n F_{\alpha_i}^c = X$. Therefore $\bigcap_{i=1}^n F_{\alpha_i} = \varnothing$ which contradicts the FIP.
$\Longleftarrow$: Let (U_α) be an open cover of X and let $F_\alpha = U_\alpha^c$. Then since $\bigcap_\alpha F_\alpha = (\bigcup_\alpha U_\alpha)^c = \varnothing$, so (F_α) does not admit FIP. Therefore, there exists $\alpha_1, \ldots, \alpha_n$ such that $\bigcap_{i=1}^n F_{\alpha_i} = \varnothing$. Therefore $X = \bigcup_{i=1}^n U_{\alpha_i}$. $\square$

Any compact subset of a Hausdorff topological space is closed.

Proposition 1.65. *Let $(X, \mathcal{T})$ be a Hausdorff topological space. If $F \subset X$ is compact, and $x \notin F$, there exists disjoint open sets $U, V \in \mathcal{T}$ such that $x \in U, F \subset V$. In particular, F is closed.*

Proof. Let $y \in F$, there are disjoint open sets U_y, V_y such that $x \in U_y, y \in V_y$. Then $F \subset \bigcup_{y \in F} V_y$, so $(V_y)_{y \in F}$ is an open cover of F. By compactness of F, there exists $y_1, \ldots, y_n \in F$ such that $F \subset \bigcup_{i=1}^n V_{y_i} = V$. Then if $U = \bigcap_{i=1}^n U_{y_i}$, we have $x \in U, U \in \mathcal{T}$ and U, V are disjoint.

Finally, if $x \in F^c$, there exists U open such that $x \in U \subset F^c$, so F^c is a neighborhood of each of its points and hence open (see Lemma 1.8-(3)). □

Compact Hausdorff spaces are normal.

Proposition 1.66. *If $(X, \mathcal{T})$ is a compact, Hausdorff (T_2) space, then it is normal (T_4).*

Proof. Let E, F be disjoint, closed sets; so they are compact (by Proposition 1.62). By Proposition 1.65 and compactness of F, if $x \in E$ there exist disjoint open sets $\widetilde{U_x}, \widetilde{V_x}$ such that $x \in \widetilde{U_x}, F \subset \widetilde{V_x}$.

Let us repeat the argument from the proof of Proposition 1.65. Since $(\widetilde{U_x})_{x \in E}$ is an open cover of E, (by compactness of E) there exists $x_1, \ldots, x_n \in E$ such that $E \subset \bigcup_{i=1}^n \widetilde{U_{x_i}} = U \in \mathcal{T}$. Then $F \subset \bigcap_{i=1}^n \widetilde{V_{x_i}} = V \in \mathcal{T}$ with $U \cap V = \varnothing$. □

Continuous image of a compact set is compact.

Proposition 1.67. *If X is compact and $f : X \to Y$ is continuous, then $f(X)$ is compact.*

Proof. Let (V_α) be an open cover of $f(X)$. Then by the continuity of f, $(f^{-1}(V_\alpha))$ is an open cover of X, so there exists a finite subcover $(f^{-1}(V_{\alpha_i}))_{1 \leqslant i \leqslant m}$ of X. Therefore $(V_{\alpha_i})_{1 \leqslant i \leqslant m}$ is a finite subcover of $f(X)$. □

The following exercise is an important application of Proposition 1.67.

Exercise 1.68. Show that compactness is a topological property of a space (cf. Definition 1.38). Using this verify the claim in Example 1.39-(3).

Another application of Proposition 1.67 is relevant for optimization problems.

Corollary 1.69. *If X is compact and $f : X \to \mathbb{R}$ is continuous, then $f(X)$ is compact. In particular, f attains its supremum and infimum, and $C(X, \mathbb{R}) = BC(X, \mathbb{R})$.*

Proof. Since $f(X)$ is compact subset of $\mathbb{R}$, it is closed and bounded (by Heine-Borel theorem) and hence $C(X, \mathbb{R}) = BC(X, \mathbb{R})$. Since $\sup_{x \in X} f(x), \inf_{x \in X} f(x) \in \overline{f(X)} = f(X)$, f attains its supremum and infimum. □

Proposition 1.70. *Let $f : X \to Y$ be a continuous bijection, where X is compact and Y is Hausdorff. Then f is a homeomorphism.*

Proof. Let $g = f^{-1} : Y \to X$, then $g^{-1}(E) = f(E)$ for $E \subset X$.

If E is closed in X, then E is compact (by Proposition 1.62), so $f(E)$ is closed (by Proposition 1.65). Hence $g^{-1}(E)$ is closed in Y, whenever E is closed in X. So g is continuous, and hence f is a homeomorphism. □

1.6 Nets, subnets, and compactness

A metric space X is compact if and only if every sequence in X has a convergent subsequence. More precisely, we have the following characterization of compact subsets of metric space.

Theorem 1.71. *Let (X, d) be a metric space and let $A \subset X$. Then the following are equivalent:*

(a) A is compact.

(b) Every sequence in A has a subsequence that converges to a limit in A.

(c) $(A, d|_{A\times A})$ is a complete metric space and is totally bounded *(that is, for any $\epsilon > 0$, A is covered by finitely many balls of radii ϵ).*

The equivalence between (a) and (b) in Theorem 1.71 can be extended to arbitrary topological spaces if we replaces sequences with nets. The following is an important characterization of compactness in terms of existence of cluster points for nets.

Theorem 1.72. *Let X be a topological space. Then the following are equivalent:*

(a) X is compact.

(b) Every net in X has a cluster point.

Proof. (a) $\Longrightarrow$ (b): Let X be compact and let (x_α) be a net in X. Set $E_\alpha = \{x_\beta : \beta \geqslant \alpha\}$ and $F_\alpha = \overline{E_\alpha}$. For any finite collection, $\alpha_1, \ldots, \alpha_m$, there exists γ such that $\alpha_i \leqslant \gamma$ for all $i = 1, \ldots, m$ (by using (iii) in Definition 1.54 repeatedly). So $x_\gamma \in E_{\alpha_i}$ for all $1 \leqslant i \leqslant m$ and hence $\bigcap_{i=1}^m F_{\alpha_i} \supset \bigcap_{i=1}^m E_{\alpha_i} \neq \varnothing$, so (F_α) satisfies FIP. Hence $\bigcap_\alpha F_\alpha \neq \varnothing$ (by Proposition 1.64). Let $x \in \bigcap_\alpha F_\alpha$ and U be an open neighborhood of x. Then $x \in F_\alpha$ for all α and hence $U \cap E_\alpha \neq \varnothing$ for all α. So for all α, there exists $\beta \geqslant \alpha$ such that $x_\beta \in U$. In other words, x is cluster point of (x_α).
(b) $\Longrightarrow$ (a): We will show the contrapositive. Let us assume that X is not compact. We need to show that there exists a net with no cluster point.

Choose an open cover $(V_\alpha)_{\alpha\in A}$ of X with no finite subcover. Now let $(\mathcal{B}, \leqslant)$ be the directed set of finite subsets of A ordered by inclusion (that is, $B_1 \leqslant B_2$ if and only if $B_1 \subset B_2$ for all $B_1, B_2 \in \mathcal{B}$). Since $(V_\alpha)_{\alpha\in A}$ has no finite subcover, for any $B \in \mathcal{B}$, $(\bigcup_{\alpha\in\mathcal{B}} V_\alpha)^c \neq \varnothing$. Let us choose $x_B \in (\bigcup_{\alpha\in\mathcal{B}} V_\alpha)^c$ for all $B \in \mathcal{B}$, so that $(x_B)_{B\in\mathcal{B}}$ is a net.

Suppose $y \in X$, so that there exists α_0 such that $y \in V_{\alpha_0}$ (since $(V_\alpha)_{\alpha\in A}$ covers X) and let $B_0 = \{\alpha_0\} \in \mathcal{B}$. If $B_1 \geqslant B_0$, then $x_{B_1} \notin V_{\alpha_0}$, so y is not a cluster point of $(x_B)_{B\in\mathcal{B}}$. Since $y \in X$ is arbitrary, there are no cluster points for the net $(x_B)_{B\in\mathcal{B}}$. □

Cluster points are generalizations of sub-sequential limits. In order to describe this, we need to first come up with a notion of subnet (analogue of subsequence). There are two ways to think about a subsequence:

(1) by choosing a subset of a sequence and renumbering the elements, or

(2) by composing the sequence considered as a function $\mathbb{N} \to X$, with a strictly increasing function from $\mathbb{N}$ to $\mathbb{N}$.

Perhaps a first guess would be to consider some subset of the directed set that forms the index of a net. The actual definition is a modification of the second way mentioned above.

Definition 1.73. A subnet of a net $(x_\alpha)_{\alpha\in A}$ is a pair

(i) $(y_\beta)_{\beta\in B}$ which is a net; and

(ii) a map $\beta \mapsto \alpha_\beta$ from $B \to A$ such that

   (a) for all $\alpha_0 \in A$, there exists $\beta_0 \in B$ such that $\alpha_\beta \geqslant \alpha_0$ for all $\beta \geqslant \beta_0$.

   (b) $y_\beta = x_{\alpha_\beta}$ for all $\beta \in B$.

The following lemma clarifies the relationship between cluster points and convergence along subnet.

Lemma 1.74. *Let $(x_\alpha)_{\alpha\in A}$ be a net in X. Then $x \in X$ is a cluster point of $(x_\alpha)_{\alpha\in A}$ if and only if there is a subnet $(y_\beta)_{\beta\in B}$ of $(x_\alpha)_{\alpha\in A}$ which converges to x.*

Proof. $\Longleftarrow$: Let $y_\beta \to x$ and U be a neighborhood of x. We need to show that for all $\alpha \in A$ there exists $\alpha' \in A$ such that $x_{\alpha'} \in U$ and $\alpha' \geqslant \alpha$.

As $y_\beta \to x$, there exists $\beta_1 \in B$ such that $y_\beta \in U$ for all $\beta \geqslant \beta_1$. Now let $\alpha \in A$. Then there exists β_0 such that $\alpha_\beta \geqslant \alpha$ for all $\beta \geqslant \beta_0$ (recall Definition 1.73-(ii)-(a)). Choose β_2 such that $\beta_2 \geqslant \beta_1, \beta_2 \geqslant \beta_0$ (recall Definition 1.54-(iii)). Then $\alpha' = \alpha_{\beta_2} \geqslant \alpha$ (as $\beta_2 \geqslant \beta_0$) and $x_{\alpha_{\beta_2}} = y_{\beta_2} \in U$ as $\beta_2 \geqslant \beta_1$.
$\Longrightarrow$ We need to construct a subnet that converges to x. Let $\mathcal{N}_x$ denote the set of neighborhoods of x and let $B = \mathcal{N}_x \times A$. We define a relation $\leqslant$ on B: say $(U, \alpha) \leqslant (U', \alpha')$ if and only if $U' \subset U$ and $\alpha \leqslant \alpha'$ (this is the relation defined in Example 1.55-(4)). It is easy to see that $(B, \leqslant)$ is a directed set (you check this!).

For $(U, \gamma) \in B$, let $\alpha = \alpha_{(U,\gamma)}$ be such that

(1) $\alpha \geqslant \gamma$,

(2) $x_\alpha \in U$ (such a point exists since x is a cluster point of (x_α)).

Now set, $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}}$ for all $(U, \gamma) \in B$. We claim that $(y_{(U,\gamma)})_{(U,\gamma)\in B}$ along with the map $(U, \gamma) \mapsto \alpha_{(U,\gamma)}$ is the desired subnet. First, let us check why it is a subnet.

- It is a not on B since B is a directed set.
- $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}}$ for all $(U, \gamma) \in B$.
- If $\alpha_0 \in A$, let $U_0 \in \mathcal{N}_x$ be arbitrary. Set $\beta_0 = (U_0, \alpha_0)$. If $\beta = (V, \gamma) \geqslant (U_0, \alpha_0)$, then $\alpha_\beta = \alpha_{(V,\gamma)} \geqslant \gamma \geqslant \alpha_0$.

So it is a subnet.

Let U be a neighborhood of x and let $\alpha_0 \in A$. If $(V,\gamma) \geqslant (U,\alpha_0)$, then $y_{(V,\gamma)} = x_{\alpha_{(V,\gamma)}} \in V \subset U$. Therefore $(y_\beta)_{\beta\in B}$ converges to x. □

Combining Lemma 1.74 and Theorem 1.72, we obtain the following

Theorem 1.75. *Let X be a topological space. Then the following are equivalent.*

(i) X is compact.

(ii) Every net in X has a cluster point.

(iii) Every net in X has a convergent subnet.

Next, we will study compactness properties of product spaces using Theorem 1.75.

Lemma 1.76. *Let X, Y be topological spaces such that Y is compact. Let $(z_\alpha)_{\alpha\in A}$ be a net in $X \times Y$ and let $x \in X$ be a cluster point of $(\pi_X(z_\alpha))_{\alpha\in A}$. Then there exists $y \in Y$ such that (x, y) is a cluster point of $(z_\alpha)_{\alpha\in A}$.*

Proof. By Lemma 1.74, there exists a subnet $(u_\beta)_{\beta\in B}$ of $(\pi_X(z_\alpha))_{\alpha\in A}$ such that $u_\beta \to x$. Set $x_\alpha = \pi_X(z_\alpha), y_\alpha = \pi_Y(z_\alpha)$ for all $\alpha \in A$, so $z_\alpha = (x_\alpha, y_\alpha)$. Note that $u_\beta = x_{\alpha_\beta}$ for all $\beta \in B$. Set $v_\beta = y_{\alpha_\beta}$ for $\beta \in B$. Since $(v_\beta)_{\beta\in B}$ is a net in a compact space Y, there is a cluster point $y \in Y$ for $(v_\beta)_{\beta\in B}$ (by Theorem 1.72).

Next, we show that $z = (x, y)$ is a cluster point of $(z_\alpha)_{\alpha\in A}$. Let U be a neighborhood of z; then (using the base in (1.4)) there exists open sets V_x and V_y in X, Y such that $x \in V_x, y \in V_y$ and $(x, y) \in V_x \times V_y \subset U$.

Let $\alpha_0 \in A$. By Definition 1.73-(ii)(a), there exists $\beta_0 \in B$ such that $\beta \geqslant \beta_0$ implies $\alpha_\beta \geqslant \alpha_0$. As $u_\beta \to x$, there exists $\beta_1 \in B$ such that $u_\beta \in V_x$ for all $\beta \geqslant \beta_1$. Let $\beta_2 \in B$ be such that $\beta_2 \geqslant \beta_0, \beta_2 \geqslant \beta_1$ (cf. Definition 1.54-(iii)). Since y is a cluster point of $(v_\beta)_{\beta\in B}$, there exists $\beta_3 \geqslant \beta_2$ such that $v_{\beta_3} \in V_y$. Since $\beta_3 \geqslant \beta_2 \geqslant \beta_1$, we have $u_{\beta_3} \in V_x$. Therefore $z_{\alpha_{\beta_3}} = (u_{\beta_3}, v_{\beta_3}) \in V_x \times V_y \subset U$ and $\alpha_{\beta_3} \geqslant \alpha_0$, since $\beta_3 \geqslant \beta_2 \geqslant \beta_0$. Since $U \in \mathcal{N}_z$ and $\alpha_0 \in A$ are arbitrary, z is a cluster point of $(z_\alpha)_{\alpha\in A}$. □

The following is a consequence of Lemma 1.76.

Corollary 1.77. *Let X, Y be compact topological spaces, then $X \times Y$ is compact.*

Proof. Let $(z_\alpha)_{\alpha\in A}$ be a net in $X \times Y$. Then by the compactness of X, there exists a cluster point $x \in X$ for the net $(\pi_X(z_\alpha))_{\alpha\in A}$. By Lemma 1.76, there exists is a cluster point of $(z_\alpha)_{\alpha\in A}$. □

We would like to extend the above result to arbitrary product of compact spaces. The basic idea is to extend the compactness to one coordinate at a time using Lemma 1.76. This works without much difficulty in the case of finite products but for infinite products we need an useful tool called *Zorn's lemma*.

Definition 1.78 (Zorn's lemma)**.** (a) We say that a relation $\leqslant$ on a set P is called a *partial order* if it is

- reflexive ($x \leqslant x$ for all $x \in P$),
- transitive ($x \leqslant y, y \leqslant z$ implies $x \leqslant z$ for all $x, y, z \in P$),
- antisymmetric (if $x \leqslant y$ and $y \leqslant x$ for $x, y \in P$, then $x = y$).

(b) We say that a subset $Q \subset P$ of a partially ordered set is *totally ordered* (or *linearly ordered*), if for all $x, y \in Q$, either $x \leqslant y$ or $y \leqslant x$ (or both).

(c) Let $Q \subset P$ be a subset of a partially ordered set P and let $c \in P$. We say that c *is an upper bound for* Q if $a \leqslant c$ for all $a \in Q$.

(d) We say that $m \in P$ is a maximal element of P if there is no element $x \in P$ such that $m \leqslant x$ other than $x = m$. (Note that a maximal element of P need not be an upper bound for P).

(e) **Zorn's lemma**: If P is a non-empty, partially ordered set and every totally ordered subset of P has an upper bound, then P has a maximal element.

Zorn's lemma is useful to show various existence results as illustrated in the exercise below (see also Exercise 2.60-(i)).

Exercise 1.79. Show that the Zorn's lemma implies the axiom of choice: if $X_\alpha \neq \varnothing$ for all $\alpha \in A$, then $\prod_{\alpha \in A} X_\alpha \neq \varnothing$ (this is the axiom of choice. In fact, the converse is also true as axiom of choice implies Zorn's lemma).

We will prove several important results in this course using Zorn's lemma such as Tychonoff's theorem (Theorem 1.80) and Hahn-Banach theorem (Theorem 2.26).

Compactness of a space can be viewed as existence of cluster points for nets (see Theorem 1.72) and hence Zorn's lemma is useful as we illustrate below.

Theorem 1.80 (Tychonoff's theorem)**.** *Let $X_i, i \in I$ be a collection of compact topological spaces. Then $X = \prod_{i \in I} X_i$ is compact.*

Proof. Let $(x_\alpha)_{\alpha \in A}$ be a net in $X = \prod_{i \in I} X_i$. For $J \subset I$, we define the projection map $\pi_J : X \to \prod_{i \in J} X_i$ as $\pi_J(x) = x\big|_J$, where $x\big|_J : J \to \bigcup_{i \in J} X_i$ denote the restriction of $x : I \to \bigcup_{i \in I} X_i$ to J for all $x \in X$. Let

$$P = \left\{ (J, p) : J \subset I, p \in \prod_{i \in J} X_i, p \text{ is a cluster point of } (\pi_J(x_\alpha))_{\alpha \in A} \right\}.$$

We define a partial order on P as $(J_1, p_1) \leqslant (J_2, p_2)$ if $J_1 \subset J_2$ and p_1 is the restriction of p_2 on J_1 (or p_2 is an extension of p_1).

Let us verify that $(P, \leqslant)$ satisfies the assumptions of Zorn's lemma.

(i) P is non-empty since if $J = \{i\}$ for some $i \in I$, then $(\pi_J(x_\alpha))_{\alpha \in A}$ has a cluster point by the compactness of X_i and Theorem 1.72.

(ii) Next, let us check that every totally ordered subset has an upper bound. Let $Q = \{(J_\beta, p_\beta) : \beta \in B\}$ be a totally ordered subset of P. Then define $J = \bigcup_{\beta \in B} J_\beta$ and $p \in \prod_{i \in J} X_i$ by

$$p(i) = p_\beta(i), \quad \text{for all } i \in J_\beta \text{ and for all } \beta \in B.$$

We need to check that the above function p is well-defined as the same i can be in different J_β's. Suppose $i \in J_{\beta_1} \cap J_{\beta_2}$ for some $\beta_1, \beta_2 \in B$, then either $(J_{\beta_1}, p_{\beta_1}) \leqslant (J_{\beta_2}, p_{\beta_2})$ or $(J_{\beta_2}, p_{\beta_2}) \leqslant (J_{\beta_1}, p_{\beta_1})$ (since Q is totally ordered). In either case, since p_{β_1} and p_{β_2} agree on $J_{\beta_1} \cap J_{\beta_2}$, we have $p_{\beta_1}(i) = p_{\beta_2}(i)$ and hence $p \in \prod_{i \in J} X_i$ is well-defined.

We need to show that p is a cluster point of $(\pi_J(x_\alpha))_{\alpha \in A}$ in $X_J := \prod_{i \in J} X_i$. Let U be an open neighborhood of p in X_J and let $\alpha_0 \in A$. So $\prod_{i \in J} V_i \subset U$, where $U_i = X_i$ for all but finitely many $i \in J$ and U_i open in X_i for all $i \in J$ (see (1.4)). Let $K = \{i \in J : V_i \neq X_i\} = \{i_1, \ldots, i_m\}$. Then there exists $\beta \in B$ such that $K \subset J_\beta$ (since Q is totally ordered). Since p_β is a cluster point in $\prod_{i \in J_\beta} X_i$, there exists $\alpha \geqslant \alpha_0$ such that $\pi_{J_\beta}(x_\alpha) \in \prod_{i \in J_\beta} V_i$. So $x_\alpha(i) \in V_i$ for all $i \in K$ and hence $\pi_J(x_\alpha) \in U$. Hence p is a cluster point of $(\pi_J(x_\alpha))_{\alpha \in A}$; that is, $(J, p) \in P$. Also, it is clear that $(J_\beta, p_\beta) \leqslant (J, p)$ for all $\beta \in B$. So (J, p) is an upper bound for Q.

By Zorn's lemma, there exists a maximal element $(\tilde{J}, \tilde{p}) \in P$ of P. If $\tilde{J} \neq I$, let $i \in I \backslash \tilde{J}$ and consider $\prod_{j \in \tilde{J}} X_j \times X_i$. By the conclusion of Lemma 1.76, there exists $p' \in \prod_{j \in \tilde{J} \cup \{i\}} X_j$ such that $(J \cup \{i\}, p') \in P$ with $(\tilde{J}, \tilde{p}) \leqslant (\tilde{J} \cup \{i\}, p')$ which contradicts the maximality of $(\tilde{J}, \tilde{p})$. So $\tilde{J} = I$. The compactness of X follows from Theorem 1.72. ☐

2 Normed vector spaces

Throughout §2, $K = \mathbb{R}$ or $\mathbb{C}$ and let X be a vector space over K.
Notation: $0 \in X$ is the zero vector. A *subspace* Y of X is a subset $Y \subset X$ that is also a vector space over K. If Y_1, Y_2 are two subspaces of X, then

$$Y_1 + Y_2 = \{y_1 + y_2 : y_1 \in Y_1, y_2 \in Y_2\}.$$

$Y_1 + Y_2$ is also a subspace of X (check this).

Definition 2.1 (Norm)**.** A *norm* $\|\cdot\|$ is a function from X to $[0, \infty) = \mathbb{R}_+$ (denoted by $x \mapsto \|x\|$) such that

(i) $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|\lambda x\| = |\lambda| \, \|x\|$ for all $\lambda \in K, x \in X$.

(iii) $\|x+y\| \leqslant \|x\| + \|y\|$ for all $x, y \in X$.

If $\|\cdot\|$ just satisfies (ii) and (iii), we say that is a *seminorm*. A vector space equipped with a norm is called a *normed vector space*.

A norm induces a metric and hence a topology (by Example 1.2-(2)). We call this topology the *norm topology* on X.

Exercise 2.2. If $\|\cdot\|$ is a norm on a vector space X, then

$$d(x,y) = \|x-y\|, \quad \text{for all } x, y \in X \tag{2.1}$$

defines a metric. **Hint**: The properties (i)-(iii) in Definition 2.1 correspond to the analogous properties in Example 1.2-(2).

Exercise 2.3. Let $(X, \|\cdot\|)$ be a normed vector space. Then show that the norm $\|\cdot\| : X \to \mathbb{R}$ is a continuous function, where X is equipped with the norm topology. (Hint: Show that $|\|x\| - \|y\|| \leqslant \|x-y\|$ for all $x, y \in X$.)

Since the norm induces a metric, we can speak of metric properties like Cauchy sequences, completeness, boundedness, on a normed vector space. For example, a sequence $(x_n)_{n\in\mathbb{N}}$ in a normed vector space is Cauchy if for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|x_n - x_m\| < \epsilon$ for all $n, m \geqslant N$.

Definition 2.4. A normed vector space is complete if every Cauchy sequence converges. A *Banach space* is a complete normed vector space.

Example 2.5. Let X be a topological space, let $\|\cdot\|_{\text{sup}}$ denote the supremum norm on $B(X,\mathbb{R})$ defined as $\|f\|_{\text{sup}} = \sup_{x\in X} |f(x)|$ for all $f \in B(X,\mathbb{R})$. Then by Lemma 1.50, $(B(X,\mathbb{R}), \|\cdot\|_{\text{sup}})$ and $(BC(X,\mathbb{R}), \|\cdot\|_{\text{sup}})$ are Banach spaces.

Not all normed vector spaces are complete, however any normed vector spaces that is not complete must necessarily be infinite dimensional (you will see why in an assignment). Here is an example of a normed vector space that is not complete.

Example 2.6. Let $X = \{f \in C([0,2],\mathbb{R}) : f(0) = 0\}$. Clearly X is a vector space as linear combinations of continuous functions that vanish at 0 are continuous functions that vanish at 0. We define the norm on X as

$$\|f\| = \int_0^2 |f(x)|\, dx, \quad \text{for all } f \in X.$$

Consider the sequence of functions $(f_n)_{n\in bN}$ in X defined by (notation: $a \wedge b = \min(a,b), a \vee b = \max(a,b)$)

$$f_n(x) = x^n \wedge 1, \quad \text{for all } x \in [0,1].$$

Note that

$$\|f_n - f_m\| = \int_0^1 |x^n - x^m|\, dx \leqslant \frac{1}{n+1} + \frac{1}{m+1}, \quad \text{for all } m, n \in \mathbb{N},$$

and hence $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. Let $g : [0,2] \to \mathbb{R}$ be defined as $g(x) = 0$ if $x < 1$ and $g(x) = 1$ if $x \geqslant 1$. Note that

$$\int_0^2 |f_n(x) - g(x)|\, dx = \int_0^1 x^n\, dx = \frac{1}{n+1}.$$

This suggests that g should be the limit but it does not belong to X, as it is not continuous at 1. We claim that (f_n) does not converge in X. Suppose to the contrary that $f_n \to f$ for some f in X, then for any $n \in \mathbb{N}$

$$\int_0^2 |f(x) - g(x)|\, dx \leqslant \int_0^2 |f(x) - f_n(x)|\, dx + \int_0^2 |f_n(x) - g(x)|\, dx = \int_0^2 |f(x) - f_n(x)|\, dx + \frac{1}{n+1}.$$

Letting $n \to \infty$ and using $\lim_{n\to\infty} \|f_n - f\| \to 0$, we obtain

$$\int_0^2 |f(x) - g(x)|\, dx = 0.$$

Therefore $f(x) = g(x)$ for almost every (with respect to Lebesgue measure) $x \in [0,2]$. This contradicts the continuity of f at $x = 1$ (why?).

Definition 2.7 (Equivalent norms)**.** We say that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are *equivalent* if there exists $C \in (0,\infty)$ such that

$$C^{-1}\|x\|_1 \leqslant \|x\|_2 \leqslant C\|x\|_1, \quad \text{for all } x \in X.$$

The terminology *equivalent* is justified by the following exercise.

Exercise 2.8. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be equivalent norms on X. Then show that the corresponding norm topologies are the same. Furthermore, show that $(X, \|\cdot\|_1)$ is complete if and only if $(X, \|\cdot\|_2)$ is complete.

Exercise 2.9. Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed vector spaces.

(a) Show that $\|(x,y)\|_{X\times Y} = \max(\|x\|_X, \|y\|_Y)$ defines a norm the vector space on $X \times Y$.

(b) Show that the norm topology induced by the norm in (a) coincides with the product of the norm topolgies for X, Y.

(c) Show that the norm in (a) is equivalent to the norms $\|(x,y)\| \mapsto \|x\|_X + \|y\|_Y$ and $\|(x,y)\| \mapsto \left(\|x\|_X^2 + \|y\|_Y^2\right)^{1/2}$.

(d) Show that $(X \times Y, \|\cdot\|_{X\times Y})$ is a Banach space if and only if both $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces.

Definition 2.10. Let (x_n) be a sequence in a normed vector space $(X, \|\cdot\|)$. We say that the series $\sum_{n=1}^\infty x_n$ *converges* (in X) if there exists $x \in X$ such that $\lim_{N\to\infty} \left\| x - \sum_{n=1}^N x_n \right\| = 0$. We say that the series $\sum_{n=1}^\infty x_n$ is *absolutely convergent* if $\sum_{n=1}^\infty \|x_n\| < \infty$.

Theorem 2.11. *A normed vector space $(X, \|\cdot\|)$ is complete if and only if every absolutely convergent series in X converges.*

Proof. $\implies$: (This part is same as the usual argument in $\mathbb{R}$) Let $\sum_{i=1}^{\infty} \|x_i\| < \infty$ and $\epsilon > 0$. Let $y_n = \sum_{i=1}^{n} x_i$ for all $n \in \mathbb{N}$. Since $\sum_{i=1}^{\infty} \|x_i\| < \infty$, there exists $N \in \mathbb{N}$ such that $\sum_{i=N}^{\infty} \|x_i\| < \infty$. Therefore for any $m, n \neq N$, we have

$$\|y_m - y_n\| \leqslant \sum_{i=m\wedge n}^{m\vee n} \|x_i\| \leqslant \sum_{i=N}^{\infty} \|x_i\| < \epsilon.$$

Therefore $(y_n)_{n\in\mathbb{N}}$ is Cauchy.
$\impliedby$: Let $(y_n)_{n\in\mathbb{N}}$ be Cauchy (in X). Since (y_n) is Cauchy, there exists a subsequence $(y_{n_i})_{i\in\mathbb{N}}$ such that $n_i < n_{i+1}$ for all $i \in \mathbb{N}$ and $\|y_m - y_{n_i}\| < 2^{-i}$ for all $m \geqslant n_i$ and all $i \in \mathbb{N}$. Let $x_i = y_{n_i} - y_{n_{i-1}}$ for all $i \geqslant 2$ and $x_1 = y_{n_1}$, so that the partial sums of $\sum_{i=1}^{\infty} x_i$ coincides with the subsequence $(y_{n_i})_{i\in\mathbb{N}}$. Since $\|x_i\| = \left\|y_{n_i} - y_{n_{i-1}}\right\| \leqslant 2^{1-i}$ for all $i \geqslant 2$, the series $\sum_{i=1}^{\infty} x_i$ is absolutely convergent and hence the subsequence $(y_{n_i})_{i\in\mathbb{N}}$ converges to say $y \in X$. Since $(y_n)_{n\in\mathbb{N}}$ is Cauchy, $y_n \to y$ as well. □

2.1 Bounded linear maps and linear functionals

Let $T : X \to Y$ be a linear map (also known as operator or linear operator) between normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$; that is, $T(x_1 + x_2) = T(x_1) + T(x_2)$ and $T(\lambda x) = \lambda T(x)$ for all $\lambda \in K, x, x_1, x_2 \in X$.

Definition 2.12 (Bounded linear map)**.** We say that a linear map $T : X \to Y$ is *bounded* if there exists $C \in (0, \infty)$ such that $\|T(x)\|_Y \leqslant C \|x\|_X$ for all $x \in X$.

It turns out that boundedness of a linear map is same as continuity (with respect to norm topology).

Proposition 2.13. *Let $T : X \to Y$ be a linear map between normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$. Then the following are equivalent:*

(a) T is continuous.

(b) T is continuous at 0.

(c) T is bounded.

Proof. (a) $\implies$ (b) is trivial.
(b) $\implies$ (c): Since T is continuous at 0, by Exercise 1.35, for every $\epsilon > 0$ there exists $\delta > 0$ such that $\|x\|_X < \delta$ implies $\|T(x)\|_Y < \epsilon$. Let $z \in X$ be arbitrary. If $z \neq 0$, then $w = \delta(2\|z\|_X)^{-1} z \in X$ satisfies $\|w\|_X = \delta/2 < \delta$ and hence

$$\frac{\delta}{2\|z\|_X} \|T(z)\|_Y = \|T(w)\|_Y < \epsilon.$$

Therefore

$$\|T(z)\|_Y \leqslant \frac{2\epsilon}{\delta} \|z\|_X \quad \text{for all } z \in X\backslash\{0\}.$$

The above estimate also holds for $z = 0$ since $T(z) = 0$ by linearity of T.
(c) $\implies$ (a): If $\|T(x)\|_Y \leqslant C\|x\|_X$ for all $x \in X$, then $\|T(y) - T(x)\|_Y = \|T(y-x)\|_Y \leqslant C\|y-x\|_X$. For any $\epsilon > 0$, the choice $\delta = \epsilon/C$ ensures the ϵ-δ definition of continuity at any $x \in X$. □

Definition 2.14. Given, normed vector spaces $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$, let $\mathcal{L}(X,Y)$ denote the set of all bounded linear maps $T : X \to Y$. Note that $\mathcal{L}(X,Y)$ is a vector space as $a_1T_1 + a_2T_2 \in \mathcal{L}(X,Y)$ for any $a_1, a_2 \in K, T_1, T_2 \in \mathcal{L}(X,Y)$, where $(a_1T_1 + a_2T_2)(x) = a_1T_1(x) + a_2T_2(x)$ for all $x \in X$.

For $T \in \mathcal{L}(X,Y)$, the *operator norm of* T is defined as

$$\|T\| := \sup\{\|T(x)\|_Y : \|x\|_X \leqslant 1\}.$$

The operator norm is also equal to (check that they are equal!)

$$\|T\| = \sup\{\|T(x)\|_Y : \|x\|_X \leqslant 1\} = \sup_{\substack{x \in X, \\ x \neq 0}} \frac{\|T(x)\|_Y}{\|x\|_X} = \sup\{\|T(x)\|_Y : \|x\|_X = 1\}. \tag{2.2}$$

Exercise 2.15. Check the claims made in the above definition. That is $\mathcal{L}(X,Y)$ is a vector space and that the operator norm defines a norm on $\mathcal{L}(X,Y)$. Show that $\|T(x)\|_Y \leqslant \|T\| \|x\|_X$ for all $x \in X$.

Proposition 2.16. *Let $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ be normed vector spaces. If Y is complete, then so is $\mathcal{L}(X,Y)$.*

Proof. Let $(T_n)_{n \in \mathbb{N}}$ be Cauchy in $\mathcal{L}(X,Y)$. Fix $x \in X$. Since

$$\|T_n(x) - T_m(x)\|_Y = \|(T_n - T_m)(x)\|_Y \leqslant \|T_n - T_m\| \|x\|_X,$$

$(T_n(x))_{n \in \mathbb{N}}$ is a Cauchy sequence in Y, and hence it converges to say $y \in Y$. Define $T : X \to Y$ as $T(x) = y$, where $y \in Y$ is as above. Then it is easy to check

(i) T is linear.

(ii) $\|T\| < \infty$; that is, $T \in \mathcal{L}(X,Y)$.

(iii) $\|T - T_n\| \to 0$ as $n \to \infty$. In particular, $\|T\| = \lim_{n\to\infty} \|T_n\|$. □

Exercise 2.17. Check (i),(ii),(iii) in the proof above.

Lemma 2.18. *Let X, Y, Z be normed vector spaces and let $T \in \mathcal{L}(X,Y), S \in \mathcal{L}(Y,Z)$, then $ST \in \mathcal{L}(X,Z)$ and $\|ST\| \leqslant \|S\| \|T\|$, where ST is the composition of S and T ($ST(x) = S(T(x))$ for all $x \in X$).*

Proof. For any $x \in X$, we have

$$\|ST(x)\|_Z \leqslant \|S(T(x))\|_Z \leqslant \|S\| \, \|T(x)\|_Y \leqslant \|S\| \, \|T\| \, \|x\|_X .$$

Therefore $ST \in \mathcal{L}(X, Z)$ and $\|ST\| \leqslant \|S\| \, \|T\|$. ☐

Two normed vector spaces are the 'same' if there is a norm-preserving bijection between them. This notion is called isometric isomorphism.

Definition 2.19. We say that a linear map $f : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ between two normed linear space is an *isometry* if $\|f(x)\|_Y = \|x\|_X$ for all $x \in X$. We say that f is an *isometric isomorphism*, if it is a bijective (or equivalently, surjective) isometry. We say that two normed vector spaces X and Y are *isometrically isomorphic*, if there exists an isometric isomorphism $f : X \to Y$ between them.

Any two norms in finite dimensional vector space are equivalent. So by Exercise 2.8, there is the norm topology is uniquely determined. The following exercise outlines a proof of this fact.

Exercise 2.20. Let $(X, \|\cdot\|)$ be a finite dimensional vector space over K. Then there is a finite basis $(e_i)_{1 \leqslant i \leqslant n}$ for the vector space X. Therefore, the linear map $\phi : K^n \to X$ defined by

$$\phi(a_1, \ldots, a_n) = \sum_{i=1}^{n} a_i e_i, \quad \text{for all } (a_1, \ldots, a_n) \in K^n,$$

is a bijection. Let K^n be equipped with the norm $\|(a_1, \ldots, a_n)\|_1 = \sum_{i=1}^{n} |a_i|$ for all $i = 1, \ldots, n$.

(a) Show that ϕ is a bounded linear map from $(K^n, \|\cdot\|_1)$ to $(X, \|\cdot\|)$.

(b) Show that the unit sphere in K^n, $S = \{(a_1, \ldots, a_n) \in K^n : \sum_{i=1}^{n} |a_i| = 1\}$ is a compact subset of K^n.

(c) Show that $\phi^{-1} : (K^n, \|\cdot\|_1) \to (X, \|\cdot\|)$ is also bounded. (Hint: Use previous parts and Exercise 2.3)

(d) Conclude that any two norms on X are equivalent.

(e) Show that any finite dimensional normed vector space is a Banach space.

Since closed and bounded sets in K^n are compact, by Exercise 2.20, the closed unit ball $\{x \in X : \|x\| \leqslant 1\}$ is compact on any finite dimensional normed vector space. It turns out that the compactness of closed unit ball characterizes finite dimensional normed vector spaces. To prove this, we need Riesz's lemma.

Lemma 2.21 (Riesz's lemma)**.** *Let X be a normed linear space and M is a closed proper subspace. Then for any $\epsilon \in (0, 1)$, there exists $x_0 \in X \backslash M$ such that $\|x_0\| = 1$ and $\|x_0 - x\| \geqslant 1 - \epsilon$ for all $x \in M$.*

Proof. Let $y \in X \in M$ be arbitrary. Since M is closed $d = \inf_{x \in M} \|y - x\| > 0$ (by Exercise 1.40-(a)). Let $\delta > 0$. Then there exists $y_1 \in M$ such that $d \leqslant \|y - y_1\| \leqslant d + \delta$. Let

$$x_0 = \frac{y - y_1}{\|y - y_1\|},$$

so that $\|x_0\| = 1$ and for any $x \in X$,

$$\|x_0 - x\| = \frac{\|y - y_1 - x\,\|y - y_1\|\|}{\|y - y_1\|} \geqslant \|y - y_1\|^{-1} \inf_{z \in M} \|y - z\| \geqslant \frac{d}{d + \delta} = 1 - \frac{\delta}{d + \delta}.$$

So we obtain the desired conclusion by choosing $\delta = \frac{d\epsilon}{1-\epsilon}$. $\square$

Theorem 2.22. *A normed vector space is finite dimensional if and only if the closed unit ball is compact in the norm topology.*

Proof. As explained after Exercise 2.20, it suffices to show that the closed unit balls is not compact for an infinite dimensional normed vector space.

Let X be an infinite dimensional normed vector space. We will construct a sequence $(x_n)_{n \in \mathbb{N}}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\|x_n - x_m\| \geqslant \frac{1}{2}$ for all $n \neq m$. This implies the desired conclusion due to Theorem 1.72 (or Theorem 1.71) as this sequence cannot have a Cauchy (and hence convergent) subsequence.

We construct such a sequence by induction. Let x_1 be any vector such that $\|x_1\| = 1$. Suppose $x_1, \ldots, x_n$ have been chosen, then we construct $x_{n+1} \in X \backslash M_n$ by choosing $\epsilon = \frac{1}{2}$ in Lemma 2.21, where M_n is the subspace $\operatorname{span}\{x_1, \ldots, x_n\}$. Note that M_n is a closed subspace due to Exercise 2.20-(e) and is proper since X is infinite dimensional. $\square$

2.2 Dual space

Definition 2.23 (linear functionals, dual space)**.** Let X be a vector space over $K = \mathbb{R}$ or $\mathbb{C}$. A bounded linear map $f : X \to K$ is called a linear functional (that is, $f \in \mathcal{L}(X, K)$).

The set of all bounded linear functionals of a normed vector space X is called the dual space of X and is denoted by X^* (that is, $X^* = \mathcal{L}(X, K)$). The dual space is also a normed vector space equipped with the operator norm.

For $z \in \mathbb{C}$, we define

$$\operatorname{sgn}(z) = \begin{cases} \frac{\bar{z}}{|z|}, & \text{if } z \neq 0, \\ 0 & \text{if } z = 0, \end{cases} \tag{2.3}$$

so that $|\operatorname{sgn}(z)| \leqslant 1$ and $\operatorname{sgn}(z) z = |z|$ for all $z \in \mathbb{C}$.

Note that by Proposition 2.16, the dual space X^* is a Banach space (even if X is not). So far, we did not distinguish between the cases $K = \mathbb{R}$ and $K = \mathbb{C}$. If X is a normed vector space over $\mathbb{C}$, it is also a normed vector space of $\mathbb{R}$. So we can consider dual space either over $\mathbb{R}$ or over $\mathbb{C}$. Let us see how these dual spaces are related.

Proposition 2.24. *Let X be a vector space over $\mathbb{C}$.*

(a) Let $f : X \to \mathbb{C}$ be complex linear function on X. Then $u = \mathrm{Re}(f) : X \to \mathbb{R}$ is a real linear function and $f(x) = u(x) - iu(ix)$ for all $x \in X$.

(b) If $u : X \to \mathbb{R}$ is a real linear function then $f(x) = u(x) - iu(ix)$ for all $x \in X$ is a $\mathbb{C}$-linear function.

Moreover, if f and u are related as above, then their operator norms are equal.

Proof. (a) Let f be $\mathbb{C}$-linear. Then

$$u(x_1+x_2) = \mathrm{Re}(f)(x_1+x_2) = \mathrm{Re}(f(x_1)+f(x_2)) = u(x_1)+u(x_2), \quad \text{for any } x_1, x_2 \in X,$$

and

$$u(\lambda x) = \mathrm{Re}(f)(\lambda x) = \mathrm{Re}(\lambda f(x)) = \lambda \,\mathrm{Re}(f)(x) = \lambda u(x) \text{for any } x \in X, \lambda \in \mathbb{R}.$$

Therefore u is $\mathbb{R}$-linear. Similarly, $v = \mathrm{Im}(f) : X \to \mathbb{R}$ is also $\mathbb{R}$-linear and $f(x) = u(x) + iv(x)$ for all $x\ inX$. For any $x \in X$, by $\mathbb{C}$-linearity of f, we have

$$u(ix) + v(ix) = f(ix) = if(x) = iu(x) - v(x), \quad \text{for all } x \in X.$$

Therefore $v(x) = -u(ix)$ for all $x \in X$.

(b) The function f is clearly $\mathbb{R}$-linear. So we just need to check $f(ix) = if(x)$ for all $x \in X$. To this end, note that

$$f(ix) = u(ix) - iu(i^2x) = u(ix) + iu(x), \quad if(x) = i[u(x) - iu(ix)] = u(x) + u(ix),$$

for all $x \in X$.

Let us now consider the operator norms. Note that

$$\|u\| = \sup\{|u(x)| : \|x\| \leqslant 1\}, \quad \|f\| = \sup\{|f(x)| : \|x\| \leqslant 1\}.$$

Since $|u(x)| = |\mathrm{Re}(f)(x)| \leqslant |f(x)|$ for all $x \in X$, we have $\|u\| \leqslant \|f\|$.

Let $x \in X$. There exists $\alpha = \mathrm{sgn}(f(x)) \in \mathbb{C}$ with $|\alpha| \leqslant 1$ such that $|f(x)| = \alpha f(x)$. Therefore $|f(x)| = \alpha f(x) = f(\alpha x) = |u(\alpha x)| \leqslant \|u\| \|\alpha x\| \leqslant \|u\| \|x\|$ (since $|\alpha| \leqslant 1$). Therefore $\|f\| \leqslant \|u\|$. □

The following exercise gives a description of the dual space of a finite dimensional normed vector space.

Exercise 2.25. Let X be a finite dimensional normed vector space over K. If the dimension is n, let us choose a basis $e_1, \ldots, e_n \in X$. Then every $x \in x$ can be uniquely written as a linear combination $x = \sum_{i=1}^n a_i e_i$, where $a_i \in K$ for all $1 \leqslant i \leqslant n$. For each $1 \leqslant i \leqslant n$, $e_i^* : X \to K$ as

$$e_i^*\left(\sum_{j=1}^n a_j e_j\right) = a_i, \quad \text{for all } (a_1, \ldots, a_n) \in K^n.$$

(i) Show that $e_i^* \in X^*$ for each $i = 1, \dots, n$. (Hint: See Exercise 2.20)

(ii) Show that $\{e_i^* : 1 \leqslant i \leqslant n\}$ is a basis for the vector space X^*.

(iii) Conclude that the dimension of X^* is same as the dimension of X.

For infinite dimensional spaces, the Hahn-Banach theorem allows us to construct lots of bounded linear functionals.

Theorem 2.26 (Hahn-Banach extension theorem:real version). *Let X be a vector space over $\mathbb{R}$ and let $p : X \to \mathbb{R}$ be a function satisfying*[1]

$$p(\lambda x) = \lambda p(x), \quad \text{for all } x \in X, \lambda > 0, \tag{2.4}$$

$$p(x + y) \leqslant p(x) + p(y), \quad \text{for all } x, y \in X. \tag{2.5}$$

Let M be a subpace of X and $f : M \to \mathbb{R}$ be a linear function on M such that $f(x) \leqslant p(x)$ for all $x \in M$. Then there exists a linear function $F : X \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in M$ and $F(x) \leqslant p(x)$ for all $x \in X$.

Proof. Let P denote the set of all linear functions $g : D(g) \to \mathbb{R}$ that satisfies the following:

- The domain $D(g)$ is a subspace of X that contains M.
- $g(x) = f(x)$ for all $x \in M$.
- $g(x) \leqslant p(x)$ for all $x \in D(g)$.

We define a partial order $\leqslant$ on P: for any $g_1, g_2 \in P$, $g_1 \leqslant g_2$ if and only if $D(g_1) \subset D(g_2)$ and $g_1(x) = g_2(x)$ for all $x \in D(g_1)$.

It is clear that P is non-empty as $f : M \to X$ belongs to P. Let $Q \subset P$ be a totally ordered subset. Let $Q = \{h_i : i \in I\}$. We define $h : D(h) \to \mathbb{R}$ by

$$D(h) = \bigcup_{i \in I} D(h_i), \quad h(x) = h_i(x), \quad \text{if } x \in D(h_i), i \in I.$$

It is easy to that h is well-defined (by the same argument as in the proof of Theorem 1.80). We can therefore use Zorn's lemma to find a maximal element $F : D(F) \to \mathbb{R}$.

We claim that $D(F) = X$. Suppose to the contrary that $D(F) \neq X$. Let $x_0 \in X \backslash D(F)$. Set $D(G) = \operatorname{span}(D(F) \cup \{x_0\}) = \{x + tx_0 : x \in D(F), t \in \mathbb{R}\}$. We would like to define a linear function $G : D(G) \to \mathbb{R}$ such that $G(x) = F(x)$ for all $x \in D(F)$ and $G(x_0) = \alpha \in \mathbb{R}$. Then $G \in P$ if and only if

$$F(x) + t\alpha \leqslant p(x + tx_0) \quad \text{for all } x \in D(F), t \in \mathbb{R};$$

which is equivalent (divide by $|t|$ on both sides) to

$$F(x) + \alpha \leqslant p(x + x_0), \quad \text{and } F(x) - \alpha \leqslant p(x - x_0), \quad \text{for all } x \in D(F),$$

[1] a function satisfying (2.4) and (2.5) is called a sublinear functional.

which in turn is equivalent to

$$\sup_{x \in D(F)} (F(x) - p(x - x_0)) \leqslant \alpha \leqslant \inf_{y \in D(F)} (p(y + x_0) - F(y)) \,.$$

Such an α exists, since for all $x, y \in D(F)$

$$F(x) + F(y) = F(x + y) \leqslant p(x + y) \leqslant p(x + x_0) + p(y - x_0),$$

which in turn implies

$$\sup_{x \in D(F)} (F(x) - p(x - x_0)) \leqslant \inf_{y \in D(F)} (p(y + x_0) - F(y)) \,.$$

This implies $F \leqslant G$ and $F \neq G$, which contradicts the maximality of F. □

Theorem 2.27 (Hahn-Banach extension theorem: complex version)**.** *Let X be a vector space over $\mathbb{C}$ and let $p : X \to \mathbb{R}$ be seminorm. Let M be a subpace of X and $f : M \to \mathbb{C}$ be a linear function on M such that $|f(x)| \leqslant p(x)$ for all $x \in M$. Then there exists a linear function (over $\mathbb{C}$) $F : X \to \mathbb{C}$ such that $F(x) = f(x)$ for all $x \in M$ and $|F(x)| \leqslant p(x)$ for all $x \in X$.*

Proof. Let $u = \mathrm{Re}(f) : M \to \mathbb{R}$ be $\mathbb{R}$-linear functional as given in Proposition 2.24. By Theorem 2.26, there exists a linear function $U : X \to \mathbb{R}$ over $\mathbb{R}$ such that $U(x) = u(x)$ for all $x \in M$ and $U(x) \leqslant p(x)$ for all $x \in X$. Define $F : X \to \mathbb{C}$ as $F(x) = U(x) - iU(ix)$ for all $x \in X$. By Proposition 2.24, F is $\mathbb{C}$-linear function. For $x \in X$, choose $\alpha = \mathrm{sgn}(F(x)) \in \mathbb{C}$ with $|\alpha| \leqslant 1$ and $F(\alpha x) = \alpha F(x) = |F(x)|$ as given in the proof of Proposition 2.24. Therefore

$$|F(x)| = U(\alpha x) \leqslant p(\alpha x) = |\alpha| p(x) \leqslant p(x).$$

□

Corollary 2.28. *Let X be a normed vector space over K and let M be a subspace of X. Let $f : M \to K$ belong to $\mathcal{L}(M, K)$. Then there exists $F \in X^*$ such that $F(x) = f(x)$ for all $x \in M$ and $\|F\| = \|f\|$.*

Proof. Note that $|f(x)| \leqslant \|G\| \, \|x\|$ for all $x \in M$. By applying either Theorem 2.26 (if $K = \mathbb{R}$) or Theorem 2.27 (if $K = \mathbb{C}$), with $p(x) = \|f\| \, \|x\|$ for all $x \in X$, we obtain the existence of F with $\|F\| \leqslant \|f\|$. We obtain the desired conclusion since

$$\|F\| = \sup\{|F(x)| : \|x\| \leqslant 1, x \in X\} \geqslant \sup\{|f(x)| : \|x\| \leqslant 1, x \in M\} = \|f\| \,.$$

□

The operator norm of F in Corollary 2.28 is as small as possible as can be verified in the following easy exercise.

Exercise 2.29. Let X be a normed vector space over K and let M be a subspace of X. Let $F \in X^*$. Then show that the restriction $f := F\big|_M : M \to K$ (that is; $f(x) = F(x)$ for all $x \in M$) belongs to the dual space M^* (M is equipped with the restriction of the norm on X) and satisfies

$$\|f\| \leqslant \|F\| \tag{2.6}$$

The following corollary provides a large family of linear functionals.

Corollary 2.30. *Let X be a normed vector space over K.*

(i) For any $x \in X$ with $x \neq 0$, there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$.

(ii) For any $x \in X$, the evaluation map $E_x : X^* \to K$ defined by $E_x(f) = f(x)$ satisfies $E_x \in X^{**}$ and the map $x \mapsto E_x$ is a linear isometry from X to X^{**}.

Proof. (i) Let $X \ni x \neq 0$. Define the linear functional $h : \text{span}\{x\} \to K$ as $h(\lambda x) = \lambda \|x\|$ for all $\lambda \in K$. Note that $\|h\| = 1$. So by Corollary 2.28, there is $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$.

(ii) Note that E_x is linear over K, since

$$E_x(\lambda_1 f_1 + \lambda_2 f_2) = (\lambda_1 f_1 + \lambda_2 f_2)(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x) = \lambda_1 E_x(f_1) + \lambda_2 E_x(f_2)$$

for all $f_1, f_2 \in X^*$, and for all $\lambda_1, \lambda_2 \in K$. Since

$$|E_x(f)| = |f(x)| \leqslant \|x\| \|f\|, \text{for all } f \in X^*, x \in X,$$

we have $\|E_x\| \leqslant \|x\|$, so we have the desired conclusion if $x = 0$.

If $x \neq 0$, by considering f by (i), we have

$$\|x\| = |f(x)| = |E_x(f)| \leqslant \|E_x\| \|f\| = \|E_x\|.$$

Combining the estimates, we obtain $\|E_x\| = \|x\|$ for all $x \in X$. It is clear that $x \mapsto E_x$ is linear over K.

□

Definition 2.31 (Reflexive spaces)**.** Let $(X, \|\cdot\|)$ is a normed linear space. Let $J : X \to X^{**}$ denote the linear isometry in Corollary 2.30-(ii). We say that the normed vector space $(X, \|\cdot\|)$ is *reflexive* if the isometry J is surjective.

Note that reflexive normed vector spaces are necessarily Banach spaces due to Proposition 2.16 but not all Banach spaces are reflexive. Nevertheless, all finite dimensional normed vector spaces are reflexive.

Exercise 2.32. Show that if $(X, \|\cdot\|)$ is a finite dimensional normed vector space, then it is reflexive. (Hint: see Exercise 2.25).

Exercise 2.33. If X is an infinite dimensional normed vector space, show that X^* is also infinite dimensional. (Hint: see Exercise 2.25).

The Hahn-Banach theorem also has useful geometric consequences.

Definition 2.34. Let X be a normed vector space over $\mathbb{R}$.

(1) A *closed hyperplane* in X is a subset $H \subset X$ of the form

$$H = \{x \in X : f(x) = \alpha\},$$

where $f : X \to \mathbb{R}$ is a non-zero bounded linear functional and $\alpha \in \mathbb{R}$. We abbreviate H as $[f = \alpha]$ in this case.

(2) Let $A, B \subset X$. We say that the closed hyperplane $[f = \alpha]$ *separates A and B* if

$$f(x) \leqslant \alpha \quad \text{for all } x \in A, \quad \text{and} \quad f(x) \geqslant \alpha \quad \text{for all } x \in B.$$

(3) We say that the closed hyperplane $[f = \alpha]$ *strictly separates A and B* if there exists $\epsilon > 0$ such that

$$f(x) \leqslant \alpha - \epsilon \quad \text{for all } x \in A, \quad \text{and} \quad f(x) \geqslant \alpha + \epsilon \quad \text{for all } x \in B.$$

(4) We say that $C \subset X$ is *convex*, if

$$tx + (1-t)y \in C, \quad \text{for all } t \in [0,1] \text{ and for all } x, y \in X.$$

Exercise 2.35. Let $(X, \|\cdot\|)$ be a normed vector space, $x \in X$ and $r > 0$. Then the open ball

$$B_X(x,r) = \{y \in X : \|y - x\| < r\}, \tag{2.7}$$

and the closed ball

$$\overline{B_X}(x,r) = \{y \in X : \|y - x\| \leqslant r\} \tag{2.8}$$

are convex sets.

Convex sets also leads to sublinear functionals. The functional p in the following lemma is sometimes called the Minkowski functional associated with a convex set.

Lemma 2.36. *Let X be a normed vector space over $\mathbb{R}$. Let C be an open, convex subset of X such that $0 \in C$. Define $p : X \to [0, \infty]$ as*

$$p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in C\}.$$

The p is a sublinear functional (that is; satisfies (2.4) and (2.5)) and such that

$$C = \{x \in X : p(x) < 1\}. \tag{2.9}$$

Furthermore, there exists $M > 0$ such that

$$p(x) \leqslant M \|x\|, \quad \text{for all } x \in X. \tag{2.10}$$

Proof. It is clear that (2.4) holds.

Let $r > 0$ be such that $B_X(0, r) \subset C$ (as C is open and $0 \in C$). Then $p(x) \leqslant r^{-1}\|x\|$ for all $x \in X$ which implies (2.10).

Let $x \in C$, then $(1+\epsilon)x \in C$ for some $\epsilon > 0$. Therefore $p(x) \leqslant (1+\epsilon)^{-1} < 1$. Hence $C \subset \{x \in X : p(x) < 1\}$. Conversely, if $p(x) < 1$, then there exits $\alpha \in (0, 1)$ such that $\alpha^{-1}x \in C$. Since $0, \alpha^{-1}x \in C$ and C is convex, we have $x = \alpha(\alpha^{-1}x) + (1-\alpha)0 \in C$. This proves (2.9).

In order to prove (2.5), consider $x, y \in X$ and let $\epsilon > 0$. Using (2.9), we have

$$(p(x) + \epsilon)^{-1}x, (p(y) + \epsilon)^{-1}y \in C.$$

Choosing

$$t = \frac{p(x) + \epsilon}{p(x) + p(y) + 2\epsilon},$$

by the convexity of C, we have

$$t(p(x) + \epsilon)^{-1}x + (1-t)(p(y) + \epsilon)^{-1}y = (p(x) + p(y) + 2\epsilon)^{-1}(x + y) \in C$$

Therefore $p(x + y) \leqslant p(x) + p(y) + 2\epsilon$. Since $\epsilon > 0$ is arbitrary, we obtain (2.5). □

The following lemma shows that points outside an open convex set is separated by a closed hyperplane.

Lemma 2.37. *Let X be a normed vector space over $\mathbb{R}$. Let $C \subset X$ be a non-empty open convex set and let $x_0 \in X \backslash C$. Then there exists $f \in X^*$ such that $f(x) < f(x_0)$ for all $x \in C$. In particular, the closed hyperplane $[f = f(x_0)]$ separates $\{x_0\}$ and C.*

Proof. After a translation, we may assume that $0 \in C$. Consider the sublinear function $p : X \to [0, \infty)$ defined Lemma 2.36. Consider the subspace $\{\lambda x_0 : \lambda \in \mathbb{R}\}$ and the linear functional $g : G \to \mathbb{R}$ defined as $g(tx_0) = t$ for all $t \in \mathbb{R}$. Note that

$$g(x) \leqslant p(x) \quad \text{for all } x \in G.$$

To see this, if $x = tx_0$ for $t \leqslant 0$ then the above inequality is true since $p(x) \geqslant 0$. If $t > 0$, then since $x_0 \notin C$, $p(x) = tp(x_0) \geqslant t = g(x)$ (by Lemma 2.36). By Theorem 2.26, there exists $f \in X^*$ such that $f(x) \leqslant p(x)$ for all $x \in E$ and $f(x_0) = 1$. By Lemma 2.36, $f(x) < 1$ for all $x \in C$. □

Theorem 2.38 (Hahn-Banach separation theorem)**.** *Let X be a normed vector space over $\mathbb{R}$.*

(1) Let $A, B \subset X$ be two non-empty, disjoint, convex subsets. Assume that one of them is open. Then there exists a closed hyperplane that separates A and B.

(2) Let $A, B \subset X$ be two non-empty, disjoint, convex subsets. Assume that A is closed and B is compact. Then there exists a closed hyperplane that strictly separates A and B.

Proof. (1) Assume that A is open. Let $C = A - B = \{a - b : a \in A, b \in B\}$. It is easy to check (do this!) that C is convex, $0 \notin C$ and $C = \bigcup_{b \in B}(A - \{b\})$ is open. By Lemma 2.37, there exists $f \in X^*$ such that $f(z) < 0$ for all $z \in C$. So

$$f(x) < f(y) \quad \text{for all } x \in A \text{ and for all } y \in B.$$

Therefore, there exists $\alpha \in \mathbb{R}$ such that

$$\sup_{x \in A} f(x) \leqslant \alpha \leqslant \inf_{y \in B} f(y).$$

So $[f = \alpha]$ is the desired closed hyperplane that separates A and B.

(2) Let $C = A - B$. It is easy to check (do this!) that C is convex, and $0 \notin C$. We claim that C is closed. To show this claim, it is enough to prove that every accumulation point of C belongs to C (see Proposition 1.11). To this end, let $y \in \operatorname{acc}(A)$. Then there exists a sequence $(a_n - b_n)_{n \in \mathbb{N}}$ such that $a_n \in A$ and $b_n \in B$ for all $n \in \mathbb{N}$ and $\lim_{n \to \infty}(a_n - b_n) = y$. By the compactness of B and passing to a subsequence if necessary we may assume that $(b_n)_{n \in \mathbb{N}}$ converges to $b \in B$ (see Theorem 1.71(b)). Therefore $a_n \to \lim_{n \to \infty}(a_n - b_n) + \lim_{n \to \infty} b_n = y + b$. Since A is closed by Proposition 1.58(b), we have that $y + b \in A$. Therefore $y = (y + b) - b \in A - B = C$. This concludes the proof that C is closed.

So there exists $r > 0$ such that $B_X(0, r)$ and C are disjoint convex sets. By the previous part, there exists $f \in X^*$ such that $f \not\equiv 0$ such that

$$f(x - y) \leqslant f(rz), \quad \text{for all } x \in A, y \in B, z \in B_X(0, 1).$$

It follows that (since $\inf_{z \in B_X(0,1)} f(rz) = -r \|f\|$; why?)

$$f(x - y) \leqslant -r \|f\|, \quad \text{for all } x \in A, y \in B.$$

Therefore, letting $\epsilon = \frac{1}{2} r \|f\| > 0$, there exists $\alpha \in \mathbb{R}$ such that

$$\sup_{x \in A}(f(x) + \epsilon) \leqslant \alpha \leqslant \inf_{y \in B}(f(y) - \epsilon).$$

Hence the closed hyperplane $[f = \alpha]$ strictly separates A and B.

□

2.3 Adjoint operator

A linear operator between normed vector spaces induces another linear operator between the dual spaces called the *adjoint operator*.

Definition 2.39 (Adjoint operator)**.** Let X, Y be normed linear spaces over K and let $T \in \mathcal{L}(X, Y)$. We define the adjoint operator $T^* : Y^* \to X^*$ as

$$T^*(f) = f \circ T, \quad \text{for all } f \in Y^*.$$

Clearly T^* is linear over K. We claim that T^* is bounded. To see this note by Lemma 2.18, that

$$\|T^*(f)\| = \|f \circ T\| \leqslant \|f\| \, \|T\| \, . \tag{2.11}$$

This implies $T^* \in \mathcal{L}(Y^*, X^*)$ and $\|T^*\| \leqslant \|T\|$. The fact that $\|T^*\| = \|T\|$ is a consequence of Hahn-Banach theorem.

Proposition 2.40. *For any $T \in \mathcal{L}(X, Y)$, we have $T^* \in \mathcal{L}(Y, X)$ and $\|T^*\| = \|T\|$.*

Proof. By (2.11), it suffices to show $\|T^*\| \geqslant \|T\|$. Let $x_0 \in X$ be such that $\|x_0\| = 1$ and $T(x_0) \neq 0$. Then there exists $f_0 \in Y^*$ with $\|f_0\| = 1$ and $f_0(T(x_0)) = \|T(x_0)\|$ (by Corollary 2.30-(i)). Therefore,

$$\begin{aligned}\|T^*\| &= \sup\{\|T^*(f)\| : \|f\| = 1\} \geqslant \|f_0 \circ T\| \\ &= \sup\{|f(T(x))| : \|x\| = 1\} \geqslant |f_0(T(x_0))| = \|T(x_0)\| \, .\end{aligned}$$

Since $\|T\| = \sup\{|T(x)| : \|x\| = 1\}$, we obtain the inequality $\|T^*\| \geqslant \|T\|$. □

Here are a few properties of the adjoint operator that follow easily from the definitions.

Exercise 2.41. (a) Let X, Y, Z be normed linear spaces with $S \in \mathcal{L}(X, Y), T \in \mathcal{L}(Y, Z)$, then $(T \circ S)^* = S^* \circ T^* \in \mathcal{L}(Z^*, X^*)$.

(b) Let X, Y be normed linear spaces over K with $T_1, T_2 \in \mathcal{L}(X, Y)$ and $\lambda_1, \lambda_2 \in K$, then $(\lambda_1 T_1 + \lambda_2 T_2)^* = \lambda_1 T_1^* + \lambda_2 T_2^*$.

(c) Let X, Y be normed linear spaces and $T \in \mathcal{L}(X, Y)$, then $T^{**} \in \mathcal{L}(X^{**}, Y^{**})$ satisfies

$$T^{**} \circ J_X = J_Y \circ T \in \mathcal{L}(X, Y^{**}),$$

where $J_X : X \to X^{**}, J_Y : Y \to Y^{**}$ are the linear isometries described in Corollary 2.30-(ii).

(d) If $T \in \mathcal{L}(X, Y)$ is such that $T^{-1} \in \mathcal{L}(Y, X)$, then $(T^*)^{-1} = (T^{-1})^*$.

Exercise 2.42. Let X, Y be normed linear spaces and $T \in \mathcal{L}(X, Y)$. Then the following are equivalent.

(i) T^* is one-to-one.

(ii) The range of T is dense in Y.

The next exercise shows how adjoint operator can be viewed as a version of matrix transpose.

Exercise 2.43. Let X be a finite dimensional normed vector space of dimension $n \in \mathbb{N}$ over K and let $\{e_i : 1 \leqslant i \leqslant n\}$ be a basis for X. Let $\{e_i^* : 1 \leqslant i \leqslant n\}$ denote the base for X^* as defined in Exercise 2.25. Let Y be a finite dimensional normed vector space over K with dimension $m \in \mathbb{N}$ with basis $\{f_i : 1 \leqslant j \leqslant m\}$. Let $\{f_i^* : 1 \leqslant i \leqslant m\}$ denote the the base for Y^* as defined in Exercise 2.25.

(a) Show that the map $V_X : X \to X^*$ defined by

$$V_X\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n a_i e_i^*, \quad \text{for all } (a_1, \ldots, a_n) \in K^n$$

is an invertible bounded linear map with bounded inverse.

(b) If X has a non-zero vector, show that there are no invertible bounded linear maps $L_X : X \to X^*$ and $L_Y : Y \to Y^*$ such that

$$L_X = T^* \circ L_Y \circ T, \quad \text{for all } T \in \mathcal{L}(X, Y).$$

(c) For any $T \in \mathcal{L}(X, Y)$, let $M_T \in K^{m \times n}$ denote the $m \times n$ matrix over K associated to T with respect to bases $\{e_j : 1 \leqslant j \leqslant n\}$ and $\{f_i : 1 \leqslant i \leqslant m\}$; that is if M_{ij}^T denotes the element in the i-th row and j-th column of M_T, we have

$$T\left(\sum_{j=1}^n a_j e_j\right) = \sum_{i=1}^m \left(\sum_{j=1}^n M_{ij}^T a_j\right) f_i, \quad \text{for all } (a_1, \ldots, a_n) \in K^n.$$

Similarly, let M_{T*} denote the $n \times m$ matrix over K associated with the adjoint operator T^* with respect to bases $\{f_j : 1 \leqslant j \leqslant m\}$ and $\{e_i^* : 1 \leqslant i \leqslant n\}$; that is if $M_{ij}^{T^*}$ denotes the element in the i-th row and j-th column of M_{T*}, we have

$$T^*\left(\sum_{j=1}^m a_j f_j^*\right) = \sum_{i=1}^m \left(\sum_{j=1}^n M_{ij}^{T^*} a_j\right) e_i^*, \quad \text{for all } (a_1, \ldots, a_m) \in K^m.$$

Show that M_{T*} is the transpose of M_T.

2.4 L^p spaces

We describe an important example of Banach spaces. Throughout 2.4, we fix a measure space $(X, \mathcal{M}, \mu)$. We will use various results concerning measure and integration from the prerequisite (MATH 420) and we refer to the appendix in §5 for a brief review.

Definition 2.44 (Semifinite, σ-finite and finite measures)**.** We say that μ is *semifinite* if whenever $E \in \mathcal{M}$ with $\mu(E) = \infty$, there exists $F \in \mathcal{M}$ with $F \subset E$ and $0 < \mu(F) < \infty$.
We say that μ is *σ-finite*, if there exists a sequence $(E_j)_{j \in \mathbb{N}}$ of sets in $\mathcal{M}$ such that $X = \bigcup_{j \in \mathbb{N}} E_j$ and $\mu(E_j) < \infty$ for all $j \in \mathbb{N}$.
We say that μ is *finite*, if $\mu(X) < \infty$.

Exercise 2.45. (i) Show that every finite measure space is σ-finite and every σ-finite measure space is semifinite.

(ii) Give examples of σ-finite measure space that is not finite and semifinite measure space that is not σ-finite.

Definition 2.46 (L^p spaces)**.** If $f : X \to \mathbb{C}$ is a measurable function on X and $p \in (0, \infty)$, we define

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p},$$

and define

$$L^p(X, \mathcal{M}, \mu) = \{[f] : f : X \to \mathbb{C}, f \text{ is measurable and } \|f\|_p < \infty\},$$

where $[f]$ is the equivalence class of f corresponding to the relation that identifies functions that are equal almost everywhere.
If μ is the counting measure on $(A, \mathcal{P}(A))$, then we abbreviate $L^p(A, \mathcal{P}(A), \mu)$ as $\ell^p(A)$.

Although, elements of L^p spaces are equivalence classes of functions, it is customary to denote them as functions; for example, $f \in L^p(X, \mathcal{M}, \mu)$ instead of $[f] \in L^p(X, \mathcal{M}, \mu)$.

First let us observe that $L^p(X, \mathcal{M}, \mu)$ forms a vector space as

$$|f + g|^p \leqslant (2\max(|f|, |g|))^p \leqslant 2^p(|f|^p + |g|^p).$$

Next, we examine, whether $(L^p(X, \mathcal{M}, \mu), \|\cdot\|_p)$ is a normed linear space. The properties (i) and (ii) in Definition 2.1 are easy to verify (property (i) also explains the need for looking at equivalence class of functions) from the definition (do it!). For the triangle inequality, we need to restrict ourselves to $p \in [1, \infty)$ as can be from the following example.

Example 2.47. Let $p \in (0, \infty), A = \{0, 1\}$ and consider $f, g \in \ell^p(A)$ such that $f(0) = g(1) = 1$ and $g(0) = f(1) = 0$. Then $\|f + g\|_p = 2^{1/p}, \|f\|_p = \|g\|_p = 1$. So, in this case, the triangle inequality is usatisfied if and only if $p \in [1, \infty)$.

The next few results are preparation to obtain the triangle inequality for the case $p \geqslant 1$.

Lemma 2.48 (Young's inequality)**.** *If $a \geqslant 0, b \geqslant 0$, and $p, q \in (1, \infty)$ be such that $p^{-1} + q^{-1} = 1$, then*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q},$$

with equality if and only if $a^p = b^q$.

Proof. The result is clear if either $a = 0$ or $b = 0$.
Set $x = ab^{-q/p} = ab^{1-q}$. So the desired claim can be rewritten as (dividing by b^q on both sides)

$$x \leqslant \frac{x^p}{p} + \frac{1}{q}, \quad \text{for all } x > 0,$$

with equality if and only if $x = 1$. This claim can be verified by setting $f(x) = \frac{x^p}{p} + \frac{1}{q} - x$ and noting that satisfies $f'(x) < 0$ for all $x \in (0, 1)$ and $f'(x) > 0$ for all $x \in (1, \infty]$. Therefore f attains its minimum at $x = 1$, and hence $f(x) \geqslant f(1) = 0$ for all $x \in (0, \infty)$. □

Our next result is Hölder's inequality which is extremely useful.

Theorem 2.49 (Hölder's inequality). *Let $(X, \mathcal{M}, \mu)$ be a measure space and let $p, q \in (1, \infty)$ satisfy $p^{-1} + q^{-1} = 1$. If $f, g : X \to \mathbb{C}$ be measurable functions, then*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q .$$

In particular, if $f \in L^p(X, \mathcal{M}, \mu), g \in L^q(X, \mathcal{M}, \mu)$, then $fg \in L^1(X, \mathcal{M}, \mu)$ and the equality $\|fg\|_1 = \|f\|_p \|g\|_q$ holds in the case that both sides are finite if and only if there exists $\alpha, \beta \in \mathbb{C}$ not both zero such that $\alpha|f|^p(x) = \beta|g|^q(x)$ μ-almost every $x \in X$.

Proof. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $fg = 0$ μ-almost everywhere and hence the desired inequality holds.

So, we may assume $\|f\|_p \neq 0$ and $\|g\|_q \neq 0$. If either $\|f\|_p = \infty$ or $\|g\|_q = \infty$, then the inequality is trivial. So we consider the case $\|f\|_p , \|g\|_q \in (0, \infty)$. Let

$$F(x) = \|f\|_p^{-1} f(x), \quad G(x) = \|g\|_q^{-1} g(x), \quad \text{for all } x \in X.$$

Then, by Lemma 2.48,

$$|F(x)G(x)| \leqslant \frac{|F(x)|^p}{p} + \frac{|G(x)|^q}{q}.$$

Integrating both sides and using the linearity of integral, we obtain

$$\frac{1}{\|f\|_p \|g\|_q} \int |fg| \, d\mu \leqslant \frac{1}{p} + \frac{1}{q} = 1.$$

Therefore,

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q .$$

Let us consider the case when $f \in L^p(X, \mathcal{M}, \mu), g \in L^q(X, \mathcal{M}, \mu)$ such that $\|fg\|_1 = \|f\|_p \|g\|_q$. If $\|f\|_p = 0$, then we may choose $\beta = 0$ and $\alpha \neq 0$ and conclude $\alpha|f|^p(x) = \beta|g|^q(x)$ μ-almost every $x \in X$. The case $\|f\|_q = 0$ is similar as we can choose $\alpha = 0, \beta = 1$.

If $\|f\|_p , \|g\|_q \in (0, \infty)$, then the equality case of Lemma 2.48 in the argument above, we have equality if and only if $|F(x)|^p = |G(x)|^q$, μ-almost everywhere or equivalently,

$$\|f\|_p^{-p} |f(x)|^p = \|g\|_q^{-q} |g(x)|^q, \quad \text{for } \mu\text{-almost every } x \in X.$$

∎

The triangle inequality for L^p norm is known as Minkowski's inequality.

Theorem 2.50 (Minkowski's inequality). *If $1 \leqslant p < \infty$ and $f, g \in L^p$, then*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p .$$

Proof. The result follows from integrating the inequality $|f+g| \leqslant |f| + |g|$ if $p = 1$.

We consider $p \in (1, \infty)$ and let $q = p/(p-1) \in (1, \infty)$. Note that,

$$|f+g|^p = |f+g||f+g|^{p-1} \leqslant |f||f+g|^{p-1} + |g||f+g|^{p-1}.$$

Then by integrating the above estimate and using Hölder inequality (Theorem 2.49), we obtain (note that $(p-1)q = p$)

$$\begin{aligned}\int_X |f+g|^p \, d\mu &\leqslant \|f\|_p \left(\int_X |f+g|^{(p-1)q} \, d\mu \right)^{1/q} + \|g\|_p \left(\int_X |f+g|^{(p-1)q} \, d\mu \right)^{1/q} \\ &= (\|f\|_p + \|g\|_p) \left(\int_X |f+g|^p \, d\mu \right)^{1/q},\end{aligned}$$

which implies the desired inequality (as $1 - q^{-1} = p^{-1}$). □

Now that we have verifies that L^p-norm is indeed a norm, we next show the completeness.

Theorem 2.51. *Let $(X, \mathcal{M}, \mu)$ is a measure space and $p \in [1, \infty)$. Then $(L^p(X, \mathcal{M}, \mu), \|\cdot\|_p)$ is a Banach space.*

Proof. Due to Theorem 2.50, we have that $(L^p(X, \mathcal{M}, \mu), \|\cdot\|_p)$ is a normed vector space (see the discussion before Example 2.47). By Theorem 2.11, it suffices to show that every absolutely convergent series converges.

To this end, let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^p(X, \mathcal{M}, \mu)$ such that $S := \sum_{n=1}^{\infty} \|f_n\|_p < \infty$. Then $G_n = \sum_{k=1}^{n} |f_k|$ satisfies (by Theorem 2.50) $\|G_n\|_p = \sum_{k=1}^{n} \|f_k\|_p \leqslant S$. By montone convergence theorem $G = \lim_{n\to\infty} G_n = \sum_{k=1}^{\infty} |f_k|$ satisfies

$$\int_X G^p \, d\mu = \lim_{n\to\infty} \int_X |G_n|^p \, d\mu \leqslant S^p < \infty.$$

Therefore $G < \infty$ μ-almost everywhere, which implies that $\sum_{k=1}^{\infty} f_k(x)$ converges for μ-almost every $x \in X$. Let $F(x) = \limsup_{n\to\infty} \sum_{k=1}^{n} f_k(x)$. Note that $|F| \leqslant G$ and hence $F \in L^p(X, \mathcal{M}, \mu)$. By the triangle inequality, $|F - \sum_{k=1}^{n} f_k|^p \leqslant (2G)^p$ for all $n \in \mathbb{N}$. So by almost everywhere convergence of $|F - \sum_{k=1}^{n} f_k|$ to zero and the dominated convergence theorem, we have

$$\lim_{n\to\infty} \left\| F - \sum_{k=1}^{n} f_k \right\|_p = 0.$$

Therefore $\sum_{k=1}^{n} f_k$ converges to F in $L^p(X, \mathcal{M}, \mu)$. □

The family of L^p space can be extended to the case $p = \infty$ as we define below.

Definition 2.52 (L^∞ space)**.** Let $(X, \mathcal{M}, \mu)$ be a measure space If $f : X \to \mathbb{C}$ is a measurable function, then we define *μ-essential supremum* (or *essential supremum* or L^∞-norm) of f as

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x \in X} |f(x)| = \inf\{t \geqslant 0 : \mu(\{x \in X : |f(x)| > t\}) = 0\},$$

with the convention that $\inf \varnothing = +\infty$. Equivalently, $\|f\|_\infty$ is the smallest $t \geqslant 0$ such that $|f(x)| \leqslant t$ for μ-almost every $x \in X$.

$L^\infty = L^\infty(X, \mathcal{M}, \mu)$ is then defined as

$$L^\infty = \{[f] | f : X \to \mathbb{C} \text{ is measurable and } \|f\|_\infty < \infty\},$$

with the usual convention that $[f]$ denotes the equivalence class of functions that agree almost everywhere with f.

Exercise 2.53. (a) Show that $L^\infty(X, \mathcal{M}, \mu)$ equipped with the essential supremum norm $\|\cdot\|_\infty$ is a normed vector space.

(b) If $f \in L^\infty(X, \mathcal{M}, \mu), g \in L^1(X, \mathcal{M}, \mu)$, show that Hölder inequality extends to the case $p = \infty$ by proving

$$\|fg\|_1 \leqslant \|f\|_\infty \|g\|_1 .$$

Similar to the case $p \in [1, \infty)$, L^∞ is also a Banach space.

Theorem 2.54. *$L^\infty(X, \mathcal{M}, \mu)$ is a Banach space.*

Proof. By Exercise 2.53-(a), it suffices to verify the completeness of the normed vector space $(L^\infty(X, \mathcal{M}, \mu), \|\cdot\|_\infty)$. Let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in L^∞. For any $k \in \mathbb{N}$ there exists $N_k \in \mathbb{N}$ such that $\|f_m - f_n\|_\infty \leqslant k^{-1}$ for all $n, m \geqslant N_k$. Hence there exists $E_k \in \mathcal{M}$ with $\mu(E_k) = 0$ such that (why?)

$$|f_m(x) - f_n(x)| \leqslant k^{-1}, \quad \text{for all } x \in X \backslash E_k, \text{ and for all } m, n \geqslant N_k. \tag{2.12}$$

We then let $E = \bigcup_{k \in \mathbb{N}} E_k$, so that $E \in \mathcal{M}$, $\mu(E) = 0$, and for all $x \in X \backslash E$, $(f_n(x))_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{C}$. Therefore $f_n(x) \to f(x)$ for all $x \in X \backslash E$. By passing to limit $m \to \infty$ in (2.12), we obtain

$$|f(x) - f_n(x)| \leqslant k^{-1} \quad \text{for all } x \in X \backslash E, \text{ and all } n \geqslant N_k. \tag{2.13}$$

Therefore $f \in L^\infty(X, \mathcal{M}, \mu)$ and $\|f - f_n\|_\infty \leqslant k^{-1}$ for all $n \geqslant N_k$. So f is the limit of $(f_n)_{n \in \mathbb{N}}$ in L^∞. ☐

Our next subject is dual of L^p spaces. Let $p, q \in [1, \infty]$ be such that $p^{-1} + q^{-1} = 1$ (so that $p = 1$ implies $q = \infty$). Note that, we have Hölder's inequality, $\|fg\|_1 \leqslant \|f\|_p \|g\|_q$ for any $f \in L^p$ and $g \in L^q$. This suggest a natural construction of linear functionals in L^p as follows: for $g \in L^q$, we define $\phi_g : L^p(X, \mathcal{M}, \mu) \to \mathbb{C}$ as

$$\phi_g(f) = \int_X fg \, d\mu. \tag{2.14}$$

Clearly ϕ_g is linear (over $\mathbb{C}$) and by Hölder inequality (see Theorem 2.49 and Exercise 2.53-(b)), we have

$$|\phi_g(f)| \leqslant \|fg\|_1 \leqslant \|g\|_q \|f\|_p \quad \text{for all } f \in L^p.$$

Therefore $\phi_g \in (L^p)^*$ with the operator norm $\|\phi_g\|$ satisfying $\|\phi_g\| \leqslant \|g\|_q$. The upper bound on the operator norm of ϕ_g given by Hölder inequality is sharp as we verify in the proposition below.

Proposition 2.55. *(a) If $p \in (1, \infty]$ (so $q \in [1, \infty)$), then $\|\phi_g\| = \|g\|_q$.*

(b) If μ is semi-finite and $p = 1$ (so $q = \infty$), then $\|\phi_g\| = \|g\|_\infty$.

Proof. (a) If $g = 0$, then $\|\phi_g\| = 0$. So we may assume $\|g\|_q \neq 0$. It suffices to show $\|\phi_g\| \geqslant \|g\|_q$. To this end, consider the function

$$f = \begin{cases} |g|^{q-1}\operatorname{sgn}(g), & \text{if } q \in (1, \infty), \\ \operatorname{sgn}(g), & \text{if } q = 1, \end{cases}$$

so that

$$\|f\|_p = \begin{cases} \left(\int_X |g|^{(q-1)p}\, d\mu\right)^{1/p} = \left(\int_X |g|^q\, d\mu\right)^{1/p} = \|g\|_q^{q/p} < \infty, & \text{if } q \in (1, \infty), \\ 1, & \text{if } q = 1. \end{cases}$$

and

$$|\phi_g(f)| = \begin{cases} \int_X |g|^{q-1}\operatorname{sgn}(g) g\, d\mu = \|g\|_q^q, & \text{if } q \in (1, \infty), \\ \int_X \operatorname{sgn}(g) g\, d\mu = \|g\|_1, & \text{if } q = 1. \end{cases}$$

Therefore

$$|\phi_g(f)| = \|g\|_q^q = \|g\|_q \|g\|_q^{q/p} = \|g\|_q \|f\|_p.$$

(b) Again, we may assume $\|g\|_\infty \in (0, \infty)$. It is enough to show that for any $\epsilon > 0$, there exists $f \in L^1$ with $\|f\|_1 = 1$ such that $|\phi_g(f)| \geqslant \|g\|_\infty - \epsilon$.

To this end, let $\epsilon > 0$. Define $A = \{x : |g(x)| \geqslant \|g\|_\infty - \epsilon\}$. By definition of essential supremum, $\mu(A) > 0$. By the semifiniteness of the measure μ, there exists $B \in \mathcal{M}$ such that $B \subset A$ and $\mu(B) \in (0, \infty)$. Set

$$f = \frac{1}{\mu(B)} \chi_B \frac{\overline{g}}{|g|},$$

so that $\|f\|_1 = 1$ and

$$|\phi_g(f)| = \phi_g(f) = \frac{1}{\mu(B)} \int_B |g|\, d\mu \geqslant \frac{1}{\mu(B)} \int_B (\|g\|_\infty - \epsilon)\, d\mu = (\|g\|_\infty - \epsilon).$$

□

The following lemma shows that every function in L^p can be approximated by simple functions if $p \in [1, \infty)$.

Lemma 2.56. *Let $(X, \mathcal{M}, \mu)$ be a measure space and let $1 \leqslant p \leqslant \infty$.*

(a) If $p \in [1, \infty)$, then the set of simple functions g of the form $g = \sum_{i=1}^{n} a_i \chi_{E_i}$, where $n \in \mathbb{N}, a_i \in \mathbb{C}, E_i \in \mathcal{M}$ and $\mu(E_i) < \infty$ for all $i = 1, \ldots, n$ is dense in L^p.

(b) Simple functions are dense in L^∞.

(c) Let $1 \leqslant p < \infty$ (so $p \neq \infty$). Let $(E_j)_{j\in\mathbb{N}}$ be a sequence of pairwise disjoint, measurable sets such that $E = \bigcup_{j\in\mathbb{N}} E_j$ satisfies $\mu(E) < \infty$. Then

$$\lim_{n\to\infty} \left\| \chi_E - \sum_{j=1}^{n} \chi_{E_j} \right\|_p = 0.$$

Proof. (a) Let $f \in L^p$. Then by approximation using simple functions (see Theorem 5.1), there exists a sequence of measurable simple functions $(f_n)_{n\in\mathbb{N}}$ such that $0 \leqslant |f_n| \leqslant f$ for all $n \in \mathbb{N}$ and such that $f_n(x) \to f(x)$ for all $x \in X$. Note that, for all $n \in \mathbb{N}$

$$|f_n - f|^p \leqslant (|f_n| + |f|)^p \leqslant 2^p |f|^p \in L^1.$$

So by dominated convergence theorem,

$$\lim_{n\to\infty} \int_X |f_n - f|^p \, d\mu = \int_X \lim_{n\to\infty} |f_n - f|^p \, d\mu = 0.$$

Since $f \in L^p$ is arbitrary, simple functions of the are dense in L^p. Note that each simple function f_n above can be written in the desired form since $f_n \in L^p$.

(b) Let $f \in L^\infty$. Then there exists (see Theorem 5.1) a sequence of measurable simple functions $(f_n)_{n\in\mathbb{N}}$ such that $0 \leqslant |f_n| \leqslant f$ for all $n \in \mathbb{N}$ such that f_n converges uniformly to f in the set $\{x \in X : |f(x)| \leqslant \|f\|_\infty\}$. Therefore $f_n \to f$ in L^∞.

(c) Note that

$$\left\| \chi_E - \sum_{j=1}^{n} \chi_{E_j} \right\|_p^p = \left\| \sum_{j=n+1}^{\infty} \chi_{E_j} \right\|_p^p = \mu\left(\bigcup_{j=n+1}^{\infty} E_j \right) \xrightarrow{n\to\infty} 0.$$

□

We saw in Proposition 2.55, that functions in L^q can be used to construct operators in the dual space $(L^p)^*$. We show that in many cases, every operator in the dual space is of the form described in Proposition 2.55.

Theorem 2.57. *Let μ a σ-finite measure on a measurable space $(X, \mathcal{M})$ and $1 \leqslant p < \infty$. Let $q \in (1, \infty]$ be defined by $q^{-1} + p^{-1} = 1$. Then if $\phi \in (L^p)^*$, then there exists $g \in L^q$ such that $\phi(f) = \phi_g(f)$ for all $f \in L^p$, where ϕ_g is as defined in (2.14). Furthermore, $\|\phi_g\| = \|\phi\| = \|g\|_q$. In other words, $(L^p)^*$ is isometrically isomorphic to L^q.*

Proof. We first consider the case $\mu(X) < \infty$. Let $\phi \in (L^p)^*$. For $E \in \mathcal{M}$, define

$$\nu(E) = \phi(\chi_E).$$

Note that $\chi_E \in L^p$ since μ is a finite measure.

Let us prove that ν is a complex measure (see Defintion 5.5). By Lemma 2.56-(c) and the continuity of $\phi : L^p \to \mathbb{C}$, we have the following: for any sequence $(E_j)_{j\in\mathbb{N}}$ of pairwise disjoint, measurable sets such that $E = \bigcup_{j\in\mathbb{N}} E_j$, we have

$$\sum_{i=1}^{\infty} \nu(E_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \phi(\chi_{E_i}) = \lim_{n\to\infty} \phi\left(\sum_{i=1}^{n} \chi_{E_i}\right) = \phi(\chi_E) = \nu(E). \tag{2.15}$$

In order to conclude that ν is a complex measure, we need to verify that $\sum_{i=1}^{\infty} |\nu(E_i)| < \infty$. To this end, let $a_i \in \mathbb{C}$ be such that $|a_i| = 1, a_i\nu(E_i) = |\nu(E_i)|$ for all $i \in \mathbb{N}$. Therefore,

$$\sum_{i=1}^{\infty} |\nu(E_i)| = \lim_{n\to\infty} \sum_{i=1}^{n} a_i\nu(E_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \phi(a_i\chi_{E_i}).$$

By the same argument as in the proof of Lemma 2.56, $\sum_{i=1}^{n} a_i\chi_{E_i}$ converges in L^p to $h := \sum_{i=1}^{\infty} a_i\chi_{E_i}$ and $\|h\|_p = \mu(E)^{1/p}$. Hence by the continuity of ϕ, we have

$$\sum_{i=1}^{\infty} |\nu(E_i)| = \phi(h) \leqslant \|\phi\| \, \|h\|_p \leqslant \|\phi\| \, \mu(E)^{1/p} < \infty,$$

where $\|\phi\| \in [0, \infty)$ is the operator norm of ϕ. This along with (2.15) concludes the proof that ν is a complex measure.

Note that by the boundedness of ϕ, we have

$$|\nu(E)| \leqslant \|\phi\| \, \|\chi_E\|_p = \|\phi\| \, \mu(E)^{1/p}, \quad \text{for any } E \in \mathcal{M}.$$

Therefore $\nu \ll \mu$. By the Radon-Nikodym theorem (see Theorem 5.6), there exists an integrable function g such that

$$\phi(\chi_E) = \nu(E) = \int_E g \, d\mu = \int_X \chi_E g \, d\mu, \quad \text{for all } E \in \mathcal{M}. \tag{2.16}$$

By the linearity of ϕ and (2.16), we have

$$\phi(f) = \int_X fg \, d\mu, \tag{2.17}$$

for all simple functions f.

Next, we show that $g \in L^q$. By approximation by simple functions (Theorem 5.1), there exists a sequence of simple functions $(g_n)_{n\in\mathbb{N}}$ such that $g_n(x) \to g(x)$ for all $x \in X$ and $|g_n| \uparrow |g|$. If $g = 0$ almost everywhere, there is nothing to show. So, by passing to a subsequence if necessary, we may assume that g_n is not identically zero for all $n \in \mathbb{N}$. Let us first consider the case $q \in (1, \infty)$. Thus, for any $n \in \mathbb{N}$, the function

$$f_n = \frac{|g_n|^{q-1} \mathrm{sgn}(g)}{\|g_n\|_q^{q-1}}$$

satisfies

$$\int_X |f_n|^p \, d\mu = \|g_n\|_q^{p(1-q)} \int_X |g_n|^{p(q-1)} \, d\mu = \|g_n\|_q^{-q} \int_X |g_n|^q \, d\mu = 1$$

and

$$\int_X |f_n g_n| \, d\mu = \|g_n\|_q^{1-q} \int_X |g_n|^q \, d\mu = \|g_n\|_q .$$

By Fatou's lemma, we have

$$\begin{aligned} \|g\|_q &= \liminf_{n\to\infty} \|g_n\|_q = \liminf_{n\to\infty} \int_X |f_n g_n| \, d\mu \\ &\leqslant \liminf_{n\to\infty} \int_X |f_n g| \, d\mu \quad (\text{since } |g_n| \leqslant |g| \text{ for all } n) \\ &= \liminf_{n\to\infty} \int_X f_n g \, d\mu \quad (\text{since } |f_n g| = f_n g) \\ &= \liminf_{n\to\infty} \phi(f_n), \quad (\text{by } (2.17)) \\ &\leqslant \|\phi\| \quad (\text{since } \|f_n\|_p = 1 \text{ for all } n) \end{aligned}$$

This concludes the proof of that $g \in L^q$.

Next, we consider the case $q = \infty$. We claim that $\|g\|_\infty \leqslant \|\phi\|$ in this case as well. Suppose to the contrary, there exists $\epsilon > 0$ such that

$$A := \{x \in X : |g(x)| \geqslant \|\phi\| + \epsilon\},$$

satisfies $\mu(A) > 0$. Then

$$f := \mu(A)^{-1} \mathrm{sgn}(g) \chi_A$$

satisfies $\|f\|_p = \|f\|_1 = 1$ and

$$\phi(f) \overset{(2.17)}{=} \int_X f g \, d\mu = \mu(A)^{-1} \int_A |g| \, d\mu \geqslant \mu(A)^{-1} \int_A (\|\phi\| + \epsilon) \, d\mu = \|\phi\| + \epsilon,$$

which contradicts the inequality, $|\phi(f)| \leqslant \|\phi\| \, \|f\|_p = \|\phi\|$. Therefore, we have

$$\|g\|_q \leqslant \|\phi\| , \quad \text{for all } q \in (1, \infty]. \tag{2.18}$$

Since ϕ and ϕ_g agree on a dense set of L^p(due to (2.17) and Lemma 2.56-(a)), $\phi(f) = \phi_g(f)$ for all $f \in L^p$ due to the continuity of ϕ and ϕ_g (ϕ_g is continuous due to Proposition 2.55).

Next, let us consider the case that μ is σ-finite. There exists an increasing sequence of measurable sets $(E_n)_{n\in\mathbb{N}}$ such that $\mu(E_n) \in (0, \infty)$ for all $n \in \mathbb{N}$ and $X = \bigcup_{n=1}^\infty E_n$. For each $n \in \mathbb{N}$, $L^p(E_n)$ can be viewed as a subset of $L^p(X)$ consisting of functions that vanish μ-almost everywhere on $X \backslash E_n$. By the previous case of finite measure, we have that for each $n \in \mathbb{N}$, there exists $g_n \in L^q(E_n)$ such that

$$\phi(f) = \int_X f g_n \, d\mu, \quad \text{for all } f \in L^p(E_n). \tag{2.19}$$

Therefore for any $n \leqslant m$, we have

$$g_n = g_m \quad \mu\text{-almost everywhere on } E_n. \tag{2.20}$$

By (2.20), where exists a measurable function $g : X \to \mathbb{C}$ such that for all $n \in \mathbb{N}$, we have

$$g = g_n \quad \mu\text{-almost everywhere on } E_n. \tag{2.21}$$

By monotone convergence theorem and (2.18),

$$\|g\|_q = \lim_{n\to\infty} \|g_n\|_q \leqslant \lim_{n\to\infty} \left\| \phi|_{L^p(E_n)} \right\| \overset{(2.6)}{\leqslant} \|\phi\|,$$

where $\phi|_{L^p(E_n)}$ denotes the restriction of ϕ to $L^p(E_n)$. Moreover by the dominated convergence theorem, for any $f \in L^p$, we have $\lim_{n\to\infty} \|f\chi_{E_n} - f\|_p = 0$ and hence by the continuity of ϕ, we have

$$\phi(f) = \lim_{n\to\infty} \phi(f\chi_{E_n}) = \lim_{n\to\infty} \int_X f g_n \chi_{E_n} \, d\mu \overset{(2.21)}{=} \lim_{n\to\infty} \int_X f g \chi_{E_n} \, d\mu = \int_X f g \, d\mu,$$

where the last equality above follows from dominated convergence theorem, since $|fg|$ is integrable due to Hölder inequality.

The equality $\|\phi\| = \|g\|_q$ also follows from Proposition 2.55. □

Exercise 2.58. Explain why (2.20) follows from (2.19) in the above proof. Likewise, explain why g in the statement of Theorem 2.57 is uniquely determined (up to μ-almost everywhere equivalence).

2.5 Baire category theorem and applications

Recall from Definition 1.6 that $A \subset X$ is dense (respectively, nowhere dense) if $\overline{A} = X$ (resp., $(\overline{A})^\circ = \varnothing$).

Theorem 2.59 (Baire category theorem)**.** *Let (X, d) be a complete metric space.*

(1) If $(U_n)_{n\in\mathbb{N}}$ is a sequence of open dense sets, then $\bigcap_{i=1}^\infty U_i$ is dense in X.

(2) If $(E_n)_{n\in\mathbb{N}}$ is a sequence of nowhere dense sets, then $\bigcup_{i=1}^\infty E_i \neq X$.

Proof. It suffices to show (1) and (2) is an easy consequence of (1). If $\overline{(E_n)_{n\in\mathbb{N}}}$ satisfies the hypotheses of (2), then $U_n = (\overline{E_n})^c$ is U_n is open and dense, since $\overline{U_n} = \overline{(\overline{E_n})^c} = (\overline{E_n}^\circ)^c = X$ (due to Lemma 1.8-(5)). So (2) follows from (1).

Let W be a non-empty open subset of X. It suffices to show that $W \cap \bigcap_{i=1}^\infty U_i \neq \varnothing$. Since W is open and non-empty, pick $x_0 \in W, r_0 \in (0,1)$ such that $\overline{B(x_0, r_0)} \subset W$. Since $B(x_0, r_0) \cap U_1$ is open and U_1 is dense, there exists $x_1 \in B(x_0, r_0) \cap U_1$ and $0 < r_1 < 2^{-1}$ such that $\overline{B(x_1, r_1)} \subset B(x_0, r_0) \cap U_1$. By induction, for all $n \in \mathbb{Z}_+$, there exist $x_n \in X, 0 < r_n < 2^{-n}$ such that

$$\overline{B(x_{n+1}, r_{n+1})} \subset B(x_n, r_n) \cap U_{n+1}.$$

Therefore (x_n) is a Cauchy sequence in X and therefore converges to $x \in X$. Note that $x \in \overline{B(x_n, r_n)} \subset W \cap U_n$ for all $n \in \mathbb{N}$ and therefore $x \in W \cap \bigcap_{i=1}^\infty U_i$. □

Exercise 2.60. Let $(X, \|\cdot\|)$ be a normed vector space over K. Recall that an *algebraic basis* of X is a subset $(e_i)_{i\in I}$ such that every $x \in X$ can be written uniquely as finite linear combination of elements of $(e_i)_{i\in I}$; that is,

$$x = \sum_{i\in J} x_i e_i, \quad \text{with } J \subset I,\ J \text{ finite, } x_i \in K \text{ for all } i \in J.$$

The cardinality of a algebraic basis is called the *dimension* of the normed vector space.

(i) Prove using Zorn's lemma that there exists an algebraic basis $(e_i)_{i\in I}$ of X such that $\|e_i\| = 1$ for all $i \in I$.

(ii) If X is infinite dimensional, show that there is a linear functional $f : X \to K$ that is *not* continuous. (Hint: Use Proposition 2.13)

(iii) If X is an infinite dimensional Banach space, show that I is not countably infinite. (Hint: Use Baire category theorem = Theorem 2.59).

The following remarkable result improves pointwise estimates to global (or uniform) estimates.

Theorem 2.61 (Uniform boundedness principles or Banach–Steinhaus theorem)**.** *Let X, Y be Banach spaces and let $(T_i)_{i\in I}$ be a family (not necessarily countable) such that $T_i \in \mathcal{L}(X, Y)$ for all $i \in I$. Suppose that*

$$\sup_{i\in I} \|T_i(x)\|_Y < \infty, \quad \textit{for all } x \in X.$$

Then

$$\sup_{i\in I} \|T_i\| < \infty.$$

Proof. Let $X_n = \{x \in X : \sup_{i\in I} \|T_i(x)\|_Y \leqslant n\}$ for all $n \in \mathbb{N}$. Note that X_n is closed (why?) and by our assumption, we have

$$\bigcup_{n=1}^\infty X_n = X.$$

By Baire category theorem (Theorem 2.59-(2)), we have $X_m^\circ \neq \varnothing$ for some $m \in \mathbb{N}$. Therefore, there exist $x_0 \in X_m$ and $r > 0$ such that for all $z \in X$ with $\|z\|_X < 1$, we have $x_0 + rz \in X_m^\circ$. Hence

$$\|T_i(x_0 + rz)\|_Y \leqslant m, \quad \text{for all } i \in I \text{ and } z \in X \text{ with } \|z\|_X < 1.$$

Therefore

$$r\,\|T_i\| \leqslant m + \|T_i(x_0)\|_Y\,, \text{for all } i \in I.$$

and hence $\sup_{i\in I} \|T_i\| \leqslant \frac{2m}{r} < \infty$. □

The following two exercises are applications of the uniform boundedness principle.

Exercise 2.62. Let X, Y be Banach spaces and let $(T_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear operators in $\mathcal{L}(X, Y)$ such that for each $x \in X$, the sequence $(T_n(x))_{n\in\mathbb{N}}$ converges to a limit, say $T(x) \in Y$. Show that

(a) $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$.

(b) $T \in \mathcal{L}(X, Y)$.

(c) $\|T\| \leqslant \liminf_{n\to\infty} \|T_n\|$.

Exercise 2.63. Let X be a Banach space over K and let $B \subset X$. Show that the following are equivalent:

(a) B is bounded in X (that is, there exists $R > 0$ such that $\|x\| \leqslant R$ for all $x \in B$).

(b) For any $T \in X^*$, the set $\{T(x) : x \in B\}$ is bounded (in K).

For a normed space $(X, \|\cdot\|)$, we denote by $B(x, r)$ (or $B_X(x, r)$), the open ball of radius r centered at x; that is,

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

For $A, B \subset X$ and $\lambda \in K$, we use the notation

$$A + B = \{a + b : a \in A, b \in B\}, \quad \lambda A = \{\lambda a : a \in A\}.$$

If $A = \{a\}$, then $A + B$ is also denoted as $a + B$ or $B + a$.

Exercise 2.64. Let X be a normed vector space and let $A, B \subset X$.

(a) If either A or B is open, then $A + B$ is open.

(b) If A is closed and B is compact, then $A + B$ is closed.

Our next results are the open mapping principle and closed graph theorem.

Theorem 2.65 (Open mapping principle)**.** *Let X, Y be Banach space and let $T \in \mathcal{L}(X, Y)$ be surjective. Then there exists $c > 0$ such that*

$$T(B_X(0,1)) \supset B_Y(0,c).$$

Remark 2.66. The above theorem implies that image of open sets under T are open sets (such a map is called open map). To see this, let U be open and let $y_0 \in T(U)$. It suffices to show that $y_0 \in (T(U))^\circ$. Let $x_0 \in U$ be such that $T(x_0) = y_0$. Since U is open, there exists $r > 0$ such that $B_X(x_0, r) \subset U$. Therefore $T(U) \supset T(B_X(x_0, r)) = T(x_0 + B_X(0, r)) = T(x_0) + rT(B_X(0,1)) \supset y_0 + rB_Y(0,c) = B_Y(y_0, cr)$.

Proof of Theorem 2.65. Let $T \in \mathcal{L}(X, Y)$ be surjective. We split the proof into two steps.
Step 1: There exists $c > 0$ such that $\overline{T(B_X(0,1))} \supset B_Y(0, 2c)$.
First, we prove the above claim. Set $Y_n := n\overline{T(B_X(0,1))}$. Since T is surjective, we have $\bigcup_{n=1}^{\infty} Y_n = Y$. Therefore by Baire category theorem (Theorem 2.59-(2)), there exists $m \in \mathbb{N}$ such that $Y_m^\circ \neq \varnothing$. So there exists $y_0 \in \left(\overline{T(B_X(0,1))}\right)^\circ$ and hence there exists $c > 0$ such that

$$B_Y(y_0, 4c) \subset \overline{T(B_X(0,1))}.$$

By symmetry $-y_0 \in \overline{T(B_X(0,1))}$, and hence

$$B_Y(0, 4c) = -y_0 + B_Y(y_0, 4c) \subset \overline{T(B_X(0,1))} + \overline{T(B_X(0,1))}.$$

Since $\overline{T(B_X(0,1))}$ is convex, we have $\overline{T(B_X(0,1))} + \overline{T(B_X(0,1))} = 2\overline{T(B_X(0,1))}$, and hence

$$B_Y(0, 2c) \subset \overline{T(B_X(0,1))}. \tag{2.22}$$

Step 2: If $c > 0$ is as given in (2.22), then $T(B_X(0,1)) \supset B_Y(0,c)$.
Let $y \in Y$ with $\|y\|_Y < c$. We need to show that there exists $x \in X$ such that $\|x\|_X < 1$ and $T(x) = y$. By (2.22), we have

$$\text{for any } \epsilon > 0, \text{ there exists } z \in X \text{ with } \|z\|_X < \tfrac{1}{2} \text{ and } \|y - T(z)\|_Y < \epsilon. \tag{2.23}$$

Choose $\epsilon = c/2$ in (2.23), there exists $z_1 \in X$ with

$$\|z_1\|_X < \frac{1}{2}, \quad \text{and} \quad \|y - T(z_1)\| < \frac{c}{2}.$$

Repeating the same argument with y replaced by $y - T(z_1)$ and with $\epsilon = c/4$, there exists $z_2 \in X$ such that

$$\|z_2\|_X < \frac{1}{4}, \quad \text{and} \quad \|y - T(z_1) - T(z_2)\| < \frac{c}{4}.$$

By induction, we obtain a sequence $(z_n)_{n\in\mathbb{N}}$ in X such that

$$\|z_n\|_X < \frac{1}{2^n}, \quad \text{and} \quad \|y - T(z_1 + z_2 + \ldots + z_n)\| < \frac{c}{2^n}, \quad \text{for all } n \in \mathbb{N}.$$

It follows that $x_n = \sum_{i=1}^{n} z_i$ is a Cauchy sequence in X with $x_n \to x$ for some $x \in B_X(0,1)$ with $T(x) = y$. □

Corollary 2.67. *Let X, Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$ be a bijection. Then T is an isomorphism; that is, $T^{-1} \in \mathcal{L}(Y, X)$.*

Proof. If T is bijective, continuity of T^{-1} is equivalent to the property that $T(U)$ is open in Y whenever U is open in X. This follows from the open mapping theorem (Theorem 2.65) as explained in Remark 2.66. □

Definition 2.68. Let $T : X \to Y$ be a linear map between normed vector spaces X and Y. We define the *graph of* T to be

$$\Gamma(T) = \{(x, y) \in X \times Y : y = T(x)\},$$

which is a subspace of $X \times Y$. We say that a linear map $T : X \to Y$ is a *closed linear map*, if $\Gamma(T)$ is closed in $X \times Y$, where $X \times Y$ is equipped with the product norm $\|(x, y)\|_{X\times Y} = \max(\|x\|_X, \|y\|_Y)$.

We always endow $X \times Y$ with the product norm $\|(x, y)\|_{X\times Y} = \max(\|x\|_X, \|y\|_Y)$ which induces the product topology (see Exercise 2.9(a)). If T is continuous, then the graph $\Gamma(T)$ is a closed subspace of $X \times Y$ endowed with the product norm (Can you see why? If not, review Proposition 1.26). The converse is also true and this is called the closed graph theorem.

Theorem 2.69. *Let X and Y be Banach spaces and let $T : X \to Y$ be a closed linear map. Then T is bounded.*

Proof. Let $\pi_1 : \Gamma(T) \to X$, $\pi_2 : \Gamma(T) \to Y$ be the projections to X and Y respectively; that is, $\pi_1(x, Tx) = x, \pi_2(x, Tx) = Tx$ for all $(x, Tx) \in \Gamma(T)$. Note that $\pi_1 \in \mathcal{L}(\Gamma(T), X), \pi_2 \in \mathcal{L}(\Gamma(T), Y)$, since $\|x\|_X \leqslant \|(x, Tx)\|_{X\times Y}, \|Tx\|_Y \leqslant \|(x, Tx)\|_{X\times Y}$ for all $(x, Tx) \in \Gamma(T)$. Since $\Gamma(T)$ is a closed subset of a Banach space, it is a Banach space (see Exercise 2.9(d)) π_1 is a bijection, $\pi_1^{-1} : X \to \Gamma(T)$ is bounded by Corollary 2.67. Therefore by Lemma 2.18, $T = \pi_2 \circ \pi_1^{-1} : X \to Y$ is bounded. □

2.6 Weak and weak* topologies

So far the only topology on a normed vector space X is the topology induced by the norm. We introduce another important topology on X called the weak topology (recall Definition 1.41).

Definition 2.70. Let $(X, \|\cdot\|)$ be a normed vector space over K and let X^* denote the dual space. Then the weak topology $\mathcal{T}(X, X^*)$ is the coarsest topology on X such that the collections of functions $\{f : X \to K | f \in X^*\}$ are continuous; that is, $\mathcal{T}(X, X^*)$ is the weak topology on X generated by $\{f : f \in X^*\}$.

Remark 2.71. There are two properties of a topological space $(X, \mathcal{T})$ that are desirable:

(i) Lots of continuous functions $f : X \to Y$.

(ii) Lots of compact sets.

However, these is a conflict between these two properties as finer topologies has fewer compact sets and coarser topologies have fewer continuous functions[2]. The weak topology can be viewed as an attempt to reconcile the two conflicting desires by prescribing the coarsest topology with a such that a given family of functions is continuous.

By Proposition 2.13, the weak topology $\mathcal{T}(X, X^*)$ is coarser than the topology induced by the norm. As a result, the weak topology has more compact sets. Since compactness plays an important role in existence of limits (for example, in minimization problems) such topologies are useful.

Let X be a normed vector space and let X^* denotes its dual space. So far, we have seen two topologies on X^*:

(i) the norm topology induced by the operator norm on X^*.

(ii) the weak topology $\mathcal{T}(X^*, X^{**})$ on X^* as given in Definition 2.70.

Now we are going to define a third topology on X^* called the weak* topology (read as 'weak star'). Recall from Definition 2.31 and Corollary 2.30-(ii) that there is a natural isometric linear map from X to X^{**}, where the map $J : X \to X^{**}$ is defined by

$$(J(x))(f) = f(x), \quad \text{for all } f \in X^* \text{ and all } x \in X.$$

Definition 2.72. Let $(X, \|\cdot\|)$ be a normed vector space over K and let X^* denote the dual space. Then the weak* topology $\mathcal{T}(X^*, X)$ is the coarsest topology on X^* such that the collections of functions $\{J(x) : X^* \to K | x \in X\}$ are continuous; that is, $\mathcal{T}(X^*, X)$ is the weak topology on X generated by $\{J(x) : x \in X\}$.

Note that if X is reflexive (that is, $J(X) = X^{**}$), then the weak and weak* topologies on X^* coincide. In general, the weak* topology on X^* is coarser than weak topology on X^* which in turn is coarser than the topology induced by the operator norm on X^*.

Proposition 2.73. *Let X be a normed vector space over K and let $\mathcal{T}(X, X^*)$ denote the weak topology on X.*

(i) The weak topology on X is Hausdorff.

(ii) Let $x_0 \in X$. Then sets of the form

$$V(f_1, \ldots, f_n; \epsilon) := \{x \in X : |f_i(x) - f_i(x_0)| < \epsilon \quad \text{for all } i = 1, \ldots, n\}$$

obtained by varying $\epsilon \in (0, \infty), n \in \mathbb{N}$ and $f_1, \ldots, f_n \in X^*$ form a neighborhood base of x_0 for the weak topology.

[2] For example, *every* function from a space equipped with discrete topology is continuous but the only compact subsets are finite sets. On the other extreme, every subset of the trivial topology is compact but the only continuous functions are constants.

Proof. (i) Let $x_1, x_2 \in X$ be distinct. By Hahn-Banach theorem (Corollary 2.30), there exists $f \in X^*$ such that $f(x_1) \neq f(x_2)$. Since K is Hausdorff, there exist disjoint open sets U_1, U_2 in K such that $f(x_1) \in U_1$ and $f(x_2) \in U_2$. Hence $f^{-1}(U_1)$ and $f^{-1}(U_2)$ are disjoint open neighborhoods of x_1 and x_2 respectively in the weak topology $\mathcal{T}(X, X^*)$.

(ii) Clearly $x_0 \in V(f_1, \ldots, f_n; \epsilon)$. Also $V(f_1, \ldots, f_n; \epsilon)$ is open in weak topology since

$$V(f_1, \ldots, f_n; \epsilon) = \bigcap_{i=1}^{n} f_i^{-1}(U_i), \quad \text{where } U_i = \{a \in K : |a - f_i(x_0)| < \epsilon\},$$

and each U_i is open in K.

Now let U be any open set containing x_0 in the weak topology. Then by the base of weak topology described after Definition 1.41, there exist $f_1, f_2, \ldots, f_n \in X^*$ and open sets $V_1, \ldots, V_n$ in K such that $f_i(x_0) \in V_i$ for each $i = 1, \ldots, n$ and

$$\bigcap_{i=1}^{n} f_i^{-1}(V_i) \subset U.$$

Since each V_i is open, there exists $\epsilon > 0$ such that $\{a \in K : |a - f_i(x_0)| < \epsilon\} \subset V_i$ for each $i = 1, \ldots, n$. Hence $V(f_1, \ldots, f_n; \epsilon) \subset U$ which concludes the proof that sets of the form $V(f_1, \ldots, f_n; \epsilon)$ form a base of weak topology by varying $\epsilon \in (0, \infty), n \in \mathbb{N}$ and $f_1, \ldots, f_n \in X^*$.

□

Proposition 2.74. *Let X be a normed vector space over K. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X.*

(i) $x_n \to x$ in the weak topology if and only if $f(x_n) \to f(x)$ for all $f \in X^*$.

(ii) $x_n \to x$ in the norm topology implies that $x_n \to x$ in the weak topology.

(iii) If $x_n \to x$ in the weak topology, then $(\|x_n\|)_{n \in \mathbb{N}}$ is bounded and $\|x\| \leqslant \liminf_{n \to \infty} \|x_n\|$.

Proof. (i) This is a special case of Exercise 1.60.

(ii) This follows from (i), Proposition 1.59, and the fact that every $f \in X^*$ is continuous in the norm topology of X.

(iii) Consider the sequence of evaluation maps $E_{x_n} \in X^{**}$ defined in Corollary 2.30-(ii). Since for any $f \in X^*$, we have

$$\lim_{n \to \infty} E_{x_n}(f) = \lim_{n \to \infty} f(x_n) = f(x),$$

by the uniform boundedness principle (Theorem 2.61) and Corollary 2.30-(ii),

$$\sup_{n \in \mathbb{N}} \|E_{x_n}\| = \sup_{n \in \mathbb{N}} \|x_n\| < \infty.$$

Since $|f(x_n)| \leqslant \|f\| \|x_n\|$ for any $f \in X^*$, we have

$$|f(x)| = \lim_{n\to\infty} |f(x_n)| \leqslant \|f\| \liminf_{n\to\infty} \|x_n\|, \quad \text{for all } f \in X^*.$$

Therefore, the desired result follows from Corollary 2.30-(i).

□

Similar results with almost the same proofs also hold for weak* topology and is left as an exercise.

Exercise 2.75. Formulate and prove versions of Propositions 2.73 and 2.74 for the weak* topology on the dual space X^* of a normed vector space X.

Recall that the compactness of closed unit ball in the norm topology is a characterization of finite dimensional normed vector spaces (see Theorem 2.22). On the other hand, closed unit ball is compact in the weak* topology even on infinite dimensional spaces. The proof involves relating the weak* topology with product topology and using Tychonoff's theorem.

Theorem 2.76 (Banach-Alaoglu theorem)**.** *Let X be a normed vector space over K and let X^* denote the dual space of X. Let $\overline{B_{X^*}}(0,1)$ denote the closed unit ball in X^*. Then $\overline{B_{X^*}}(0,1)$ is compact with respect to the subspace topology of the weak* topology on X^*.*

Proof. Let X^* denote the dual space of X equipped with the weak* topology $\mathcal{T}(X^*, X)$. Let $Y = \prod_{x\in X} K = K^X$ denote the space of all functions from X to K equipped with the standard product topology. Let $\pi_x : Y \to K, x \in X$, denote the cannonical projection maps. Let $\Phi : X^* \to Y$ denote the cannonical injective map such that $\Phi(f) = (f(x))_{x\in X}$ for all $f \in X^*$. By Exercise 1.43, the map Φ is continuous since $\pi_x \circ \Phi : X^* \to K$ is the map $f \mapsto f(x)$ which in turn is continuous by the definition of weak* topology.

Let us verify that the inverse map $\Phi^{-1} : \Phi(X^*) \to X^*$ is also continuous if $\Phi(X^*)$ is equipped with the subspace topology inherited from Y. Again by Exercise 1.43, the map $\Phi^{-1} : \Phi(X^*) \to X^*$ is continuous since $J(x) \circ \Phi^{-1} = \pi_x\big|_{\Phi(X^*)} : \Phi(X^*) \to K$ for all $x \in X$, which is continuous by the definition of subspace and product topologies. Therefore, $\Phi : X^* \to \Phi(X^*)$ is a homeomorphism.

Let $\overline{B_K}(0,r) = \{a \in K : |a| \leqslant r\}$ denote the closed ball of radius r centered at 0 in K. Note that

$$\Phi(\overline{B_{X^*}}(0,1)) = \left(\bigcap_{x\in X} \{\omega \in Y : \pi_x(\omega) \in \overline{B_K}(0, \|x\|)\}\right) \cap \left(\bigcap_{x,y\in X, a,b\in K} \{\omega \in Y : a\pi_x(\omega) + b\pi_y(\omega) = \pi_{ax+by}(\omega)\}\right)$$

Note that $\left(\bigcap_{x\in X} \{\omega \in Y : \pi_x(\omega) \in \overline{B_K}(0, \|x\|)\}\right)$ is compact, since it is the product of compact sets $\prod_{x\in X} \overline{B_K}(0, \|x\|)$ by Tychnoff's theorem (see Theorem 1.80 and Exercise 1.47). For each $x, y \in X$ and $a, b \in K$, the function $\omega \mapsto \pi_{ax+by}(\omega) - a\pi_x(\omega) - b\pi_y(\omega)$ is a continuous function on Y and hence

$$\bigcap_{x,y\in X, a,b\in K} \{\omega \in Y : a\pi_x(\omega) + b\pi_y(\omega) = \pi_{ax+by}(\omega)\}$$

is a closed subset of Y. Since the closed subset of a compact space is compact (by Proposition 1.62), we have that $\Phi(\overline{B_{X^*}}(0,1))$ is a compact subset of Y and hence $\Phi(X^*)$. Since $\Phi : X^* \to \Phi(X^*)$ is a homeomorphism, this implies the desired conclusion. □

2.7 Metrizability of weak topology

Definition 2.77. We say that a topology $\mathcal{T}$ on a set X is metrizable, if there is a metric $d : X \times X \to [0, \infty)$ on X such that the topology induced by the metric coincides with $\mathcal{T}$.

By definition, the norm topology on a normed vector space is metrizable. The following theorem clarifies when the weak topology is metrizable.

Theorem 2.78. *Let X be a Banach space over K. Then the following are equivalent:*

(a) The weak topology $\mathcal{T}(X, X^*)$ on X is metrizable.

(b) X is a finite dimensional space.

The proof of Theorem 2.78 requires some preliminary results. First, we show that finite dimensional normed vector spaces have metrizable weak topology by showing that weak and norm topologies are the same.

Proposition 2.79. *Let X be a finite dimensional normed vector space over K. Then the weak topology $\mathcal{T}(X, X^*)$ is same as the norm topology on X. In particular, the weak topology is metrizable.*

Proof. Let $(e_i)_{1\leqslant i\leqslant n}$ be a basis of X such that $\|e_i\| = 1$ for all $i = 1, \ldots, n$. Let $(f_i)_{1\leqslant i\leqslant n}$ be the basis of X^* as defined in Exercise 2.25, so that $x = \sum_{j=1}^n f_j(x)e_j$ for all $x \in X$.

Since weak topology is coarser than the norm topology it suffices to show that every open set in norm topology is open in the weak topology. So it suffices to show that for any $x_0 \in X, r > 0$, the open ball $B(x_0, r) = \{x \in X : \|x - x_0\| < r\}$ is a neighborhood of x_0 in the weak topology. To this end, note that if

$$y \in V(f_1, \ldots, f_n; \epsilon) := \{x \in X : |f_i(x) - f_i(x_0)| < \epsilon \quad \text{for all } i = 1, \ldots, n\},$$

then by the triangle inequality

$$\|y - x_0\| = \left\|\sum_{i=1}^n (f_i(y) - f_i(y_0))e_i\right\| \leqslant \sum_{i=1}^n |f_i(y) - f_i(y_0)| < n\epsilon.$$

So

$$V(f_1, \ldots, f_n; r/n) \subset B(x_0, r)$$

for any $x_0 \in X, r > 0$, which concludes the proof by Proposition 2.73-(ii). □

If X is a normed vector space over $\mathbb{C}$, it can also be viewed as a normed vector space over $\mathbb{R}$. Then it has two dual spaces from these two different viewpoints, say $X^*_{\mathbb{R}}$ and $X^*_{\mathbb{C}}$. So there are two weak topologies on X that arise $\mathcal{T}(X, X^*_{\mathbb{R}})$ and $\mathcal{T}(X, X^*_{\mathbb{C}})$ by viewing X as a vector space over $\mathbb{R}$ and $\mathbb{C}$ respectively. We claim that these two weak topologies are the same. To see this, recall relationship between $X^*_{\mathbb{R}}$ and $X^*_{\mathbb{C}}$ in Proposition 2.24

$$X^*_{\mathbb{R}} = \{\mathrm{Re}(f) | f \in X^*_{\mathbb{C}}\}, \quad X^*_{\mathbb{C}} = \{F : X \to \mathbb{C} | F(x) = u(x) - iu(ix) \text{ for all } x \in X, \text{ where } u \in X^*_{\mathbb{R}}\}.$$

Let $f \in X^*_{\mathbb{C}}$. Then $\mathrm{Re}(f), \mathrm{Im}(f) \in X^*_{\mathbb{R}}$. This shows that $\mathcal{T}(X, X^*_{\mathbb{R}}) \supset \mathcal{T}(X, X^*_{\mathbb{C}})$ as every $f \in X^*_{\mathbb{C}}$ is continuous with respect to $\mathcal{T}(X, X^*_{\mathbb{R}})$ as both real and imaginary parts are continuous. Conversely, if $u \in X^*_{\mathbb{R}}$, there exists $f \in X^*_{\mathbb{C}}$ such that $u = \mathrm{Re}(f)$. So u is continuous with respect to $\mathcal{T}(X, X^*_{\mathbb{C}})$ and hence $\mathcal{T}(X, X^*_{\mathbb{R}}) \subset \mathcal{T}(X, X^*_{\mathbb{C}})$. This concludes the proof that $\mathcal{T}(X, X^*_{\mathbb{R}}) = \mathcal{T}(X, X^*_{\mathbb{C}})$.

So for the purposes of proving Theorem 2.78, we may assume without any loss of generality that X is a vector space over $\mathbb{R}$ for the remainder of §2.7.

Let us recall the definition of the notion of kernel and range of a linear operator.

Definition 2.80. Let $T \in \mathcal{L}(X, Y)$. Then the kernel (or nullspace) of T denoted by $\mathcal{N}(T)$ is defined as

$$\mathcal{N}(T) = \{x \in X : T(x) = 0\}.$$

The range of T denoted by $\mathcal{R}(T)$ is defined as

$$\mathcal{R}(T) = \{T(x) : x \in X\}.$$

Note that $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are subspaces of X and Y respectively and $\mathcal{N}(T)$ is closed in X (check these elementary facts). The following algebraic lemma is a consequence of Hahn-Banach separation theorem.

Lemma 2.81. *Let X be a normed vector space over $\mathbb{R}$ and let $\phi_1, \phi_2, \ldots, \phi_n \in X^*$ be linear functionals. Let $\phi : X \to \mathbb{R}$ be a linear functional such that*

$$\bigcap_{i=1}^{n} \mathcal{N}(\phi_i) \subset \mathcal{N}(\phi). \tag{2.24}$$

Then there exists $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$\phi(x) = \sum_{i=1}^{n} \lambda_i \phi_i(x), \quad \textit{for all } x \in X.$$

Proof. Define $T : X \to \mathbb{R}^{n+1}$, where

$$T(x) = (\phi(x), \phi_1(x), \ldots, \phi_n(x)), \quad \text{for all } x \in X,$$

where $\mathbb{R}^{n+1}$ is a normed linear space equipped with the norm given in Exercise 2.20. Then by (2.24), the point $y_0 = (1, 0, \ldots, 0)$ does not belong to the range $\mathcal{R}(T)$ of T.

By Hahn-Banach separation theorem (Theorem 2.38-(2)), there exists $G \in (\mathbb{R}^{n+1})^*$ and $\alpha \in \mathbb{R}$ such that

$$G(y_0) < \alpha < G(T(x)), \quad \text{for all } x \in X.$$

By Exercise 2.25, there exists $\lambda, \lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $G((z_0, z_1, \ldots, z_n)) = \lambda z_0 + \sum_{i=1}^n \lambda_i z_i$ for all $(z_0, z_1, \ldots, z_n) \in \mathbb{R}^{n+1}$. Therefore

$$\lambda < \alpha < \lambda\phi(x) + \sum_{i=1}^{n} \lambda_i \phi_i(x), \quad \text{for all } x \in X. \tag{2.25}$$

Substituting $x = 0$ in the above gives $\lambda < \alpha < 0$. If $\lambda\phi(x) + \sum_{i=1}^n \lambda_i\phi_i(x) \neq 0$ for some $x \in X$, then replacing x with tx in (2.25) leads to a contradiction as either $t \to \infty$ or $t \to -\infty$. Hence $\lambda \neq 0$ and $\lambda\phi(x) + \sum_{i=1}^n \lambda_i\phi_i(x) = 0$ for all $x \in X$. □

Proposition 2.82. *Let X be a normed vector space over $\mathbb{R}$ such that the weak topology is metrizable, then X is finite dimensional.*

Proof. As the proof is long, we break it into three steps:

1. X^* admits a basis that is either finite or countably infinite.
2. X^* is finite dimensional.
3. X is finite dimensional.

Step 1: Let $d : X \times X \to [0, \infty)$ be a metric on X such that the metric space (X, d) induces the weak topology $\mathcal{T}(X, X^*)$. Recall from Example 1.18-(ii) that

$$\{B_k : k \in \mathbb{N}\}, \quad \text{where } B_k = \{x : d(x, 0) < k^{-1}\}$$

is a neighborhood base of 0 in the weak topology. By Proposition 2.73-(ii), for all $k \in \mathbb{N}$ there exists a finite set $F_k \subset X^*$ and $\epsilon_k \in (0, \infty)$ such that

$$\{x \in X : |f(x)| < \epsilon_k \quad \text{for all } f \in F_k\} \subset B_k. \tag{2.26}$$

Let $F = \bigcup_{k=1}^\infty F_k$. Since each F_k is finite, F is either finite or countably infinite. We claim that $\operatorname{span}(F) = X^*$; that is, every $g \in X^*$ is a finite linear combination of elements of F. Let $g \in X^*$ be arbitrary. Since $\{x \in X : |g(x)| < 1\}$ is a neighborhood of zero (by Proposition 2.73-(ii)), and $\{B_k : k \in \mathbb{N}\}$ is a neighborhood base, by (2.26), there exists $m \in \mathbb{N}$ such that

$$\{x \in X : |f(x)| < \epsilon_m \quad \text{for all } f \in F_m\} \subset B_m \subset \{x \in X : |g(x)| < 1\}.$$

So if $x \in \bigcap_{f \in F_m} \mathcal{N}(f)$, then $= |t||g(x)| = |g(tx)| < 1$ for all $t \in \mathbb{R}$, which in turn implies $g(x) = 0$. Therefore

$$\bigcap_{f \in F_m} \mathcal{N}(f) \subset \mathcal{N}(g).$$

By Lemma 2.81, $g \in \text{span}(F_m) \subset \text{span}(F)$ and hence

$$X^* = \text{span}(F).$$

Since F is either finite or countably infinite, by choosing a maximal (by inclusion) linearly independent subset of F using Zorn's lemma, we obtain a basis that is either finite or countably infinite.
Step 2: From step 1, we need to rule out the possibility that X^* has a countably infinite basis. This is an easy consequence of Exercise 2.60-(iii) and the fact that X^* is a Banach space (by Proposition 2.16). Therefore X^* is finite dimensional space.
Step 3: By step 2 and Exercise 2.25-(iii), X^{**} is finite dimensional. Since there is a one-to-one linear map $J : X \to X^{**}$ (by Corollary 2.30-(ii)), we conclude that X is finite dimensional. □

We are now ready to prove Theorem 2.78.

Proof of Theorem 2.78. (b) $\Longrightarrow$ (a) follows from Proposition 2.79.
(a) $\Longrightarrow$ (b) follows from Proposition 2.82 and the discussion after Proposition 2.79 which reduces the analysis to the case $K = \mathbb{R}$. □

A slight modification of the proof of Theorem 2.78 also shows a similar result for weak* topologies as stated in the exercise below.

Exercise 2.83. Let X be a Banach space over K. Then the following are equivalent:

(a) The weak* topology $\mathcal{T}(X^*, X)$ on X^* is metrizable.

(b) X is a finite dimensional space.

Hint: Imitate the proof of Theorem 2.78. It might help to solve Exercise 2.75 first.

3 Hilbert spaces

Hilbert spaces can be viewed as generalizations of Banach space where the norm is replaced with an inner product. The notion of inner product can be viewed as a refined version of norm.

Definition 3.1 (inner product)**.** Let $\mathcal{H}$ be a vector space over $\mathbb{C}$. An *inner product* on $\mathcal{H}$ is a function $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that:

(i) For any $x, y, z \in \mathcal{H}$ and $a, b \in \mathbb{C}$, we have

$$\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle.$$

(ii) For any $x, y \in \mathcal{H}$,

$$\langle y, x \rangle = \overline{\langle x, y \rangle}.$$

(iii) $\langle x, x\rangle \in [0, \infty)$ for all $x \in \mathcal{H}$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

A complex vector space $\mathcal{H}$ equipped with an inner product is called a *pre-Hilbert space*. If $\mathcal{H}$ is a pre-Hilbert space with inner product $\langle \cdot, \cdot\rangle$, we define

$$\|x\| := \sqrt{\langle x, x\rangle}, \quad \text{for all } x \in \mathcal{H}. \tag{3.1}$$

Remark 3.2. Note that properties (i) and (ii) in Definition 3.1 implies that, for all $x, y, z \in \mathcal{H}$ and $a, b \in \mathbb{C}$, we have

$$\langle x, ay + bz\rangle = \bar{a}\langle x, y\rangle + \bar{b}\langle x, z\rangle.$$

The notation in (3.1) suggests that every inner product defines a norm. In order to verify the triangle inequality for the 'norm' defined in (3.1), we need the Schwarz inequality.

Theorem 3.3 (Schwarz inequality)**.** *Let $\mathcal{H}$ be a pre-Hilbert space with inner product $\langle \cdot, \cdot\rangle$. Then for all $x, y \in \mathcal{H}$, we have*

$$|\langle x, y\rangle| \leqslant \|x\| \, \|y\| \, ,$$

with equality if and only if x and y are linearly dependent.

Proof. If $\langle x, y\rangle = 0$, then the result holds trivially.

So we may assume that $\langle x, y\rangle \neq 0$ (and hence $\|x\| \neq 0, \|y\| \neq 0$; why?) Let $\alpha = \operatorname{sgn}(\langle y, x\rangle)$, so that if $z = \alpha y$, we have $\langle x, z\rangle = \overline{\langle \alpha y, x\rangle} = \bar{\alpha}\langle x, y\rangle = \operatorname{sgn}(x, y)\langle x, y\rangle = |\langle x, y\rangle|$. For any $t \in \mathbb{R}$, we have

$$0 \leqslant \langle x - tz, x - tz\rangle = \langle z, z\rangle t^2 - 2|\langle x, y\rangle| t + \langle x, x\rangle.$$

Since the quadratic function $t \mapsto \langle z, z\rangle t^2 - 2|\langle x, y\rangle| t + \langle x, x\rangle$ is non-negative and it achieves its minimum at $t_0 = \|y\|^{-2} |\langle x, y\rangle|$ and hence

$$0 \leqslant \langle x - t_0 z, x - t_0 z\rangle = \|x\|^2 - \|y\|^{-2} |\langle x, z\rangle|^2.$$

Note that this implies the desired inequality, with equality if and only if (by Definition 3.1-(iii))

$$x = t_0 z = t_0 \alpha y$$

which happens if and only if x and y are linearly independent. ☐

Proposition 3.4. *The function $x \mapsto \|x\|$ defined in (3.1) is a norm on $\mathcal{H}$.*

Proof. Note that $\|x\| \geqslant 0$ with equality if and only if $x = 0$ (by Definition 3.1-(iii)). By Definition 3.1-(i),(ii), we have $\|\lambda x\| = |\lambda| \, \|x\|$ for all $x \in \mathcal{H}, \lambda \in \mathbb{C}$. It remains to verify the triangle inequality. To this end, note that

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2 \operatorname{Re}\langle x, y\rangle. \tag{3.2}$$

Therefore by Theorem 3.3, for all $x, y \in \mathcal{H}$, we have

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2 \operatorname{Re}\langle x, y\rangle \leqslant \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \\ &\leqslant \|x\|^2 + \|y\|^2 + 2 \|x\| \, \|y\| = (\|x\| + \|y\|)^2 \, . \end{aligned}$$

☐

Given the norm induced by the inner product one can recover the inner product using the following formula (called the polarization identity).

Exercise 3.5. Let $\mathcal{H}$ be an inner product space equipped with inner product $\langle\cdot,\cdot\rangle$ and the norm given by (3.1). Then we have

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3}\left(i^k \left\|x + i^k y\right\|^2\right), \quad \text{for all } x, y \in \mathcal{H}.$$

By Proposition 3.4, every pre-Hilbert space is a normed vector space, and hence is a metric space (recall Exercise 2.2). The definition of Hilbert space is similar to that of Banach space.

Definition 3.6 (Hilbert space)**.** A pre-Hilbert spce what is complete with respect to the norm $\|x\| = \sqrt{\langle x, x\rangle}$ is called a *Hilbert space*.

Example 3.7. Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $L^2(X, \mathcal{M}, \mu)$ be equipped with the inner product (Exercise: check the properties of inner product)

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu.$$

Then the corresponding norm is the L^2-norm in Definition 2.46 and hence the above inner product is called the L^2-inner product. By Theorem 2.51, this pre-Hilbert space a Hilbert space.

Lemma 3.8 (Parallelogram law)**.** *For all $x, y \in \mathcal{H}$,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

Proof. This follows from (3.2) as

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\langle x, y\rangle, \quad \|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\operatorname{Re}\langle x, y\rangle.$$

□

Remark 3.9. We can also define *real* Hilbert spaces for vector spaces of $\mathbb{R}$. In this case, the *real inner product* on a vector space $\mathcal{H}_{\mathbb{R}}$ over $\mathbb{R}$ is a function $\langle\cdot,\cdot\rangle : \mathcal{H}_{\mathbb{R}} \times \mathcal{H}_{\mathbb{R}} \to \mathbb{R}$ such that:

(i) For any $x, y, z \in \mathcal{H}$ and $a, b \in \mathbb{C}$, we have

$$\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle.$$

(ii) For any $x, y \in \mathcal{H}$,

$$\langle y, x\rangle = \langle x, y\rangle.$$

(iii) $\langle x, x\rangle \in [0, \infty)$ for all $x \in \mathcal{H}$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

The proof that a real inner product defines a norm $\|x\| = \sqrt{\langle x, x\rangle}$ is exactly the same as in the complex case. In fact, all results in §3 works for real vector spaces.

3.1 Projection onto a closed convex set

Theorem 3.10 (Projection onto a closed convex set)**.** *Let K be a nonempty closed convex set of a Hilbert space $\mathcal{H}$. Then for every $x \in \mathcal{H}$, there exists a unique $u \in K$ such that*

$$\|x-u\| = \min_{v\in K} \|x-v\| . \tag{3.3}$$

Moreover, the distance minimizing property (3.3) is equivalent to

$$u \in K, \quad \textit{and } \operatorname{Re}\left(\langle x-u, v-u\rangle\right) \leqslant 0, \quad \textit{for all } v \in K. \tag{3.4}$$

Proof. Existence: Let $(v_n)_{n\in\mathbb{N}}$ be a minimizing sequence for $\inf_{v\in K}\|x-v\|$; that is, $v_n \in K$ for all $n \in \mathbb{N}$ and

$$d_n := \|x-v_n\| \to d := \inf_{v\in K} \|x-v\| .$$

We claim that (v_n) is a Cauchy sequence. By parallelogram law (Lemma 3.8) applied to $\frac{x-v_n}{2}$ and $\frac{x-v_m}{2}$, we have

$$\left\|x - \frac{v_n+v_m}{2}\right\|^2 + \frac{1}{4}\|v_n-v_m\|^2 = \frac{1}{2}\left(d_n^2+d_m^2\right)$$

Since $\frac{v_n+v_m}{2} \in K$ (by the convexity of K) and thus $\left\|x-\frac{v_n+v_m}{2}\right\| \geqslant d$. It follows that

$$\frac{1}{4}\|v_n-v_m\|^2 \leqslant \frac{1}{2}\left(d_n^2+d_m^2\right) - d^2, \quad \text{and} \quad \lim_{m,n\to\infty}\|v_n-v_m\| = 0.$$

Therefore (v_n) converges to some $u \in K$ (since K is closed) with $d = \|x-u\|$. This completes the proof fo existence of $u \in K$ that satisfies (3.3).

Equivalence: Before we show uniqueness, we show the equivalence between (3.3) and (3.4). Assume that $u \in K$ satisfies (3.3) and let $w \in K$. By the convexity of K we have

$$v = (1-t)u + tw \in K, \quad \text{for all } t \in [0,1],$$

and hence

$$\|x-u\| \leqslant \|x-((1-t)u+tw)\| = \|x-u-t(w-u)\| .$$

Therefore by (3.2),

$$\|x-u\|^2 \leqslant \|x-u\|^2 - 2t\operatorname{Re}\left(\langle x-u, v-u\rangle\right) + t^2\|w-u\|^2 .$$

As $t \downarrow 0$, we obtain (3.4).

Conversely, assume that $u \in K$ satsifies (3.4), then for any $v \in K$, we have by (3.2)

$$\|u-x\|^2 - \|v-x\|^2 = 2\operatorname{Re}\langle x-u, v-u\rangle - \|v-u\|^2 \overset{(3.4)}{\leqslant} 0,$$

which implies (3.3).

Uniqueness: Assume that $u_1, u_2 \in K$ satisfy (3.4). Therefore

$$\operatorname{Re}\left(\langle x-u_1, v-u_1\rangle\right) \leqslant 0, \quad \text{for all } v \in K. \tag{3.5}$$

$$\text{Re}\left(\langle x - u_2, v - u_2\rangle\right) \leqslant 0, \quad \text{for all } v \in K. \tag{3.6}$$

By setting $v = u_2$ in (3.5) and $v = u_1$ in (3.6), and adding the inequalities, we obtain

$$0 \geqslant \text{Re}\left(\langle x - u_1, u_2 - u_1\rangle + \langle x - u_2, u_1 - u_2\rangle\right) = \text{Re}(\|u_1 - u_2\|^2) = \|u_1 - u_2\|^2 .$$

Therefore $u_1 = u_2$. □

Notation: For $x \in \mathcal{H}$ and a closed convex subset $K \subset \mathcal{H}$, by $P_K(x) \in K$, we denote the unique $u = P_K(x) \in K$ that satisfies (3.3) in Theorem 3.10. We call $P_K(x)$ the projection of x onto K.

For a Banach space, the existence and uniqueness described in Theorem 3.10 can fail even for the case $x = 0$ (equivalent to elements of minimal norm) as outlined in the exercise below.

Exercise 3.11. (a) (failure of existence) Consider the Banach space $X = C([0,1])$ equipped with the supremum(or uniform) norm. Then show that the set

$$M = \left\{ f \in X : \int_0^{\frac{1}{2}} f(x)\, dx - \int_{\frac{1}{2}}^1 f(x)\, dx = 1 \right\}$$

is a nonempty closed convex set of X with *no* element of minimal notm.

(b) (failure of uniqueness) Consider the Banach space $Y = L^1(\mathbb{R}, \mathcal{B}, m)$, where m is the Lebesgue measure and $\mathcal{B}$ is the Borel-σ-field. Show that the set $K = \{f \in X : \int_{\mathbb{R}} f(x)\, dx = 1\}$ is a non-empty closed convex subset of Y with infinitely many elements of minimal norm.

The projection onto K maps cannot increase distances as shown below.

Proposition 3.12. *Let K be a nonempty closed convex set of a Hilbert space $\mathcal{H}$. Then*

$$\|P_K(x_1) - P_K(x_2)\| \leqslant \|x_1 - x_2\| , \quad \textit{for all } x_1, x_2 \in \mathcal{H}.$$

Proof. Let $x_1, x_2 \in \mathcal{H}$ and $u_1 = P_K(x_1), u_2 = P_K(x_2) \in K$. Then by (3.4) in Theorem 3.10, we have

$$\text{Re}\left(\langle x_1 - u_1, v - u_1\rangle\right) \leqslant 0, \quad \text{for all } v \in K. \tag{3.7}$$

$$\text{Re}\left(\langle x_2 - u_2, v - u_2\rangle\right) \leqslant 0, \quad \text{for all } v \in K. \tag{3.8}$$

By setting $v = u_2$ in (3.5) and $v = u_1$ in (3.6), and adding the inequalities, we obtain

$$\begin{aligned} 0 &\geqslant \text{Re}\left(\langle x_1 - u_1, u_2 - u_1\rangle + \langle x_2 - u_2, u_1 - u_2\rangle\right) \\ &= \text{Re}(\|u_1 - u_2\|^2) - \text{Re}\left(\langle x_1 - x_2, u_1 - u_2\rangle\right) = \|u_1 - u_2\|^2 - \text{Re}\left(\langle x_1 - x_2, u_1 - u_2\rangle\right). \end{aligned}$$

Therefore by Schwarz inequality (Theorem 3.3), we have

$$\|u_1 - u_2\|^2 \leqslant \text{Re}\left(\langle x_1 - x_2, u_1 - u_2\rangle\right) \leqslant |\langle x_1 - x_2, u_1 - u_2\rangle| \leqslant \|u_1 - u_2\| \, \|x_1 - x_2\| .$$

This implies the desired estimate. □

Corollary 3.13. *Assume that $M \subset \mathcal{H}$ is a closed subspace of a Hilbert space $\mathcal{H}$. Let $x \in \mathcal{H}$. Then $u = P_M(x)$ is characterized by*

$$u \in M, \quad \text{and} \quad \langle x - u, v \rangle = 0, \quad \text{for all } v \in M. \tag{3.9}$$

Furthermore, $P_M : \mathcal{H} \to M$ is a bounded linear operator and is called the orthogonal projection onto M.

Proof. Suppose $u \in M$ satisfies

$$\|x - u\| = \min_{v \in M} \|x - v\|.$$

Then by (3.4), we have

$$\operatorname{Re}(\langle x - u, w - u \rangle) \leqslant 0, \quad \text{for all } w \in M.$$

For any $v \in M, \alpha \in \mathbb{C}$, we have $w = \alpha v + u \in M$ and hence

$$\operatorname{Re}(\langle x - u, \alpha v \rangle) \leqslant 0, \quad \text{for all } v \in M \text{ and } \alpha \in \mathbb{C}.$$

By choosing $\alpha \in \mathbb{C}$ such that $\operatorname{Re}(\langle x - u, \alpha v \rangle) = \operatorname{Re}(\overline{\alpha}\langle x - u, v \rangle) = |\langle x - v, v \rangle|$ in the above estimate, we obtain (3.9).

Conversely, if $u \in M$ satisfies (3.9), then it satisfies

$$\langle x - u, w - u \rangle = 0, \quad \text{for all } w \in M.$$

and hence implies (3.4). Therefore by Theorem 3.10, we have $u = P_M(x)$.

Note that for all $a, b \in \mathbb{C}$, $x, y \in \mathcal{H}$ and $v \in M$, by (3.9), we have

$$\langle ax + by - (aP_M(x) + bP_M(y)), v \rangle = a\langle x - P_M(x), v \rangle + b\langle y - P_M(y), v \rangle \overset{(3.9)}{=} 0.$$

By the characterization in (3.9), we conclude that

$$P_M(ax + by) = aP_M(x) + bP_M(y), \quad \text{for all } x, y \in \mathcal{H} \text{ and } a, b \in \mathbb{C}.$$

Hence P_M is linear (over $\mathbb{C}$). Furthermore, by Proposition 3.12, we have $\|P_M(x)\| \leqslant \|x\|$ for all $x \in \mathcal{H}$. So P_M is bounded. □

3.2 Dual of a Hilbert space

Next, we describe the dual space of a Hilbert space. For any $y \in \mathcal{H}$, the map $\phi_y : \mathcal{H} \to \mathbb{C}$ defined by

$$\phi_y(x) = \langle x, y \rangle \tag{3.10}$$

is clearly linear (over $\mathbb{C}$) by Definition 3.1-(i). Note that, by Schwarz inequality,

$$|\phi_y(x)| = |\langle x, y \rangle| \leqslant \|y\| \|x\|.$$

Therefore $\phi_y \in \mathcal{H}^*$ and the operator norm $\|\phi\|$ satisfies $\|\phi\| \leqslant \|y\|$. The fact that $\|\phi\| = \|y\|$ follows from $\|y\|^2 = |\phi_y(y)| \leqslant \|\phi_y\| \|y\|$. The map $y \mapsto \phi_y$ is *conjugate linear* from $\mathcal{H}$ to $\mathcal{H}^*$, as

$$\phi_{a_1y_1+a_2y_2} = \overline{a_1}\phi_{y_1} + \overline{a_2}\phi_{y_2}, \quad \text{for all } a_1, a_2 \in \mathbb{C} \text{ and } y_1, y_2 \in \mathcal{H}.$$

Hence $y \mapsto \phi_y$ defines a conjugate linear isometry from $\mathcal{H}$ to $\mathcal{H}^*$. The surjectivity of this map is called the Riesz–Fréchet representation theorem.

Theorem 3.14 (Riesz–Fréchet representation theorem). *For any $\phi \in \mathcal{H}^*$, there exists a unique $y \in \mathcal{H}$ such that $\phi(x) = \langle x, y\rangle$ for all $x \in \mathcal{H}$.*

Proof. Existence: If $\phi \equiv 0$, then we may choose $y = 0$. Otherwise, consider the kernel of ϕ, that is, $M = \phi^{-1}(\{0\})$. Then M is a closed subspace (due to continuity and linearity of ϕ) and $M \neq \mathcal{H}$. Choose $x_1 \in \mathcal{H}\backslash M$ and let $u_1 = P_M(x_1)$. Then by Corollary 3.13, $x_2 = \|x_1 - u_1\|^{-1}(x_1 - u_1)$ satisfies

$$x_2 \notin M, \quad \|x_2\| = 1, \quad \text{and} \quad \langle v, x_2\rangle = 0, \quad \text{for all } v \in M. \tag{3.11}$$

Since $\phi(x_2) \neq 0$, for any $x \in \mathcal{H}$ we have

$$x - \frac{\phi(x)}{\phi(x_2)}x_2 \in M,$$

and hence by (3.11), we have

$$0 = \langle x - \frac{\phi(x)}{\phi(x_2)}x_2, x_2\rangle = \langle x, x_2\rangle - \frac{\phi(x)}{\phi(x_2)}\langle x_2, x_2\rangle = \langle x, x_2\rangle - \frac{\phi(x)}{\phi(x_2)}, \quad \text{for all } x \in \mathcal{H}.$$

Therefore, choosing $y = \overline{\phi(x_2)}x_2$, we have

$$\phi(x) = \phi(x_2)\langle x, x_2\rangle = \langle x, \overline{\phi(x_2)}x_2\rangle = \langle x, y\rangle, \quad \text{for all } x \in \mathcal{H}.$$

Uniqueness: If there exist $y_1, y_2 \in \mathcal{H}$ such that $\phi(x) = \langle x, y_1\rangle = \langle x, y_2\rangle$ for all $x \in \mathcal{H}$, by choosing $x = y_1 - y_2$, we have $\|y_1 - y_2\|^2 = 0$ and hence $y_1 = y_2$. □

Note that the Riesz-Fréchet representation theorem defines a bijection $\mathfrak{C} : \mathcal{H} \to \mathcal{H}^*$ between a Hilbert space and its dual defined by

$$(\mathfrak{C}(y))(x) = \langle x, y\rangle \quad \text{for all } x, y \in \mathcal{H}.$$

It is not linear over $\mathbb{C}$, but rather *conjugate linear*, that is,

$$(\mathfrak{C}(x))(a_1y_1 + a_2y_2) = \overline{a_1}(\mathfrak{C}(x))(y_1) + \overline{a_2}(\mathfrak{C}(x))(y_2), \quad \text{for all } a_1, a_2 \in \mathbb{C}, x, y_1, y_2 \in \mathcal{H}.$$

Note that, if $T : \mathcal{H} \to \mathcal{H}$ is a bounded linear operator on a Hilbert space, then the adjoint $T^* : \mathcal{H}^* \to \mathcal{H}^*$ is a bounded linear operator on its dual (recall Definition 2.39). By the Riesz-Fréchet representation theorem, we can define can view the adjoint operator as an operator in $\mathcal{L}(\mathcal{H},\mathcal{H})$ be conjugating with the map $\mathfrak{C} : \mathcal{H} \to \mathcal{H}^*$; that is, we define the *Hilbert space adjoint* as the operator $T^\dagger := \mathfrak{C}^{-1} \circ T^* \circ \mathfrak{C} : \mathcal{H} \to \mathcal{H}$. So we have

$$\langle x, T^\dagger(y)\rangle = \langle x, (\mathfrak{C}^{-1} \circ T^* \circ \mathfrak{C})(y)\rangle = (T^*(\mathfrak{C}(y)))(x) = (\mathfrak{C}(y))(T(x)) = \langle T(x), y\rangle,$$

for all $x, y \in \mathcal{H}$. It is easy to verify that $T^\dagger$ is linear and bounded. By Proposition 2.40 and Theorem 3.14, we have

$$\left\|T^\dagger\right\| = \|T\|.$$

That is, the Hilbert space adjoint $T^\dagger : \mathcal{H}^* \to \mathcal{H}^*$ of $T \in \mathcal{L}(\mathcal{H},\mathcal{H})$ is the unique linear operator in $\mathcal{L}(\mathcal{H},\mathcal{H})$ characterized by

$$\langle T(x), y\rangle = \langle x, T^\dagger(y)\rangle, \quad \text{for all } x, y \in \mathcal{H}. \tag{3.12}$$

We say that a bounded operator $\mathcal{T} \in \mathcal{L}(\mathcal{H},\mathcal{H})$ is *self-adjoint* if $T = T^\dagger$.

3.3 Orthonormal basis

Definition 3.15. Let $\mathcal{H}$ be a Hilbert space. We say a subset $\{u_\alpha : \alpha \in A\} \subset \mathcal{H}$ is *orthonormal* if $\langle u_\alpha, u_\alpha\rangle = 1$ for all $\alpha \in A$ and

$$\langle u_\alpha, u_\beta\rangle = 0, \quad \text{for all } \alpha, \beta \in A \text{ such that } \alpha \neq \beta.$$

Proposition 3.16 (Bessel's inequality)**.** *If $\{u_\alpha : \alpha \in A\}$ is an orthonormal subset of $\mathcal{H}$ and $x \in \mathcal{H}$, then*

$$\sum_{\alpha \in A} |\langle x, u_\alpha\rangle|^2 \leqslant \|x\|^2.$$

In particular, $\{\alpha \in A : \langle x, u_\alpha\rangle \neq 0\}$ is countable.

Proof. It suffices to show that $\sum_{\alpha \in F} |\langle x, u_\alpha\rangle|^2 \leqslant \|x\|^2$ for any finite subset $F \subset A$. Note that

$$\begin{aligned} 0 \leqslant \left\| x - \sum_{\alpha \in F} \langle x, u_\alpha\rangle u_\alpha \right\|^2 &= \|x\|^2 + \sum_{\alpha \in F} \|\langle x, u_\alpha\rangle u_\alpha\|^2 - \sum_{\alpha \in F} 2\,\mathrm{Re}\langle x, \langle x, u_\alpha\rangle u_\alpha\rangle \\ = \|x\|^2 + \sum_{\alpha \in F} |\langle x, u_\alpha\rangle|^2 - 2\sum_{\alpha \in F} |\langle x, u_\alpha\rangle|^2 &= \|x\|^2 - \sum_{\alpha \in F} |\langle x, u_\alpha\rangle|^2. \end{aligned}$$

By Bessel's inequality, $\{\alpha \in A : |\langle x, u_\alpha| > n^{-1}\}$ is finite for each $n \in \mathbb{N}$. ☐

An useful consequence of Schwarz inequality is the continuity of inner product in both variables.

Lemma 3.17. *If $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof.

$$\begin{aligned}|\langle x_n y_n\rangle - \langle x,y\rangle| &\leqslant |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\ &\leqslant |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \leqslant \|x_n - x\| \, \|y_n\| + \|x\| \, \|y_n - y\| \,.\end{aligned}$$

Since $\|y_n\| \to \|y\|$, we obtain the result. □

We describe properties of a maximal (see Exercise 3.20) orthonormal set.

Theorem 3.18. *If $\{u_\alpha : \alpha \in A\}$ is an orthonormal subset of $\mathcal{H}$, the following properties are equivalent:*

(a) (Completeness) If $x \in \mathcal{H}$ satisfies $\langle x, u_\alpha\rangle = 0$ for all $\alpha \in A$, then $x = 0$.

(b) (Parseval's identity) For each $x \in \mathcal{H}$, we have $\|x\|^2 = \sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2$.

(c) For each $x \in \mathcal{H}$, $x = \sum_{\alpha\in A}\langle x, u_\alpha\rangle u_\alpha$, where the sum has only countably many non-zero terms and converges in the norm topology regardless of how the terms are ordered.

Proof. (b) implies (a) follows from the non-degeneracy of norm.
(a) $\Longrightarrow$ (c): Let $x \in \mathcal{H}$. Let $\alpha_1, \alpha_2, \ldots$ be an enumeration (finite or infinite sequence) of all α's such that $\langle x, u_\alpha\rangle \neq 0$. Since the sum converges if it is finite, we assume that we have an infinite sum. By Bessel's inequality, we have $\sum_j \left|\langle x, u_{\alpha_j}\rangle\right|$ converges, and hence

$$\left\|\sum_{j=m}^{n} \langle x, u_{\alpha_j}\rangle u_{\alpha_j}\right\|^2 = \sum_{j=m}^{n} \left|\langle x, u_{\alpha_j}\rangle\right|^2 \to 0, \qquad \text{as } m, n \to \infty.$$

By the completeness of $\mathcal{H}$, the series $\sum_j \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$ converges. If $y = x - \sum_j \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$, then $\langle y, u_\alpha\rangle = 0$ for all $\alpha \in A$ (by the continuity of inner product; see Lemma 3.17). So by the completeness of the orthonormal set, we conclude that $y = 0$, or equivalently $x = \sum_j \langle x, u_{\alpha_j}\rangle u_{\alpha_j}$.
(c) $\Longrightarrow$ (b): With α_j's as above, the calculation in the proof of Bessel's inequality implies that

$$\|x\|^2 - \sum_{j=1}^{n} \left|\langle x, u_{\alpha_j}\right|^2 = \left\|x - \sum_{j=1}^{n} \langle x, u_{\alpha_j}\rangle u_{\alpha_j}\right\|^2 \xrightarrow{n\to\infty} 0.$$

□

Definition 3.19. If an orthonormal subset of a Hilbert space satisfies any of the equivalent conditions of Theorem 3.18, we say that it is a *orthonormal basis*.

Exercise 3.20. Let $\mathcal{H}$ be a Hilbert space.

(a) Consider the collection of all orthonormal subsets partially ordered by inclusion. Show using Zorn's lemma that there is a maximal orthonormal subset.

(b) Show that if $\{u_\alpha \in A\}$ is a maximal orthonormal subset of $\mathcal{H}$, then it is complete (in the sense of Theorem 3.18-(a)).

(c) Conclude that every Hilbert space has an orthonormal basis.

The following is a notion is when two Hilbert spaces can be considered the 'same' (or isomorphism).

Definition 3.21. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ respectively. We say that a map $U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary, if it is linear, invertible and preserves inner products

$$\langle U(x), U(y) \rangle_2 = \langle x, y \rangle_1, \quad \text{for all } x, y \in \mathcal{H}_1.$$

Using a orthonormal basis every Hilbert space can be viewed as an L^2 space.

Proposition 3.22. *Let $\{e_\alpha : \alpha \in A\}$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. The map $U : \mathcal{H} \to \ell^2(A)$ defined by $(U(x))(\alpha) = \langle x, e_\alpha \rangle$ is a unitary map.*

Proof. Linearity of U follows from the linearity of inner product in the first argument.

Completeness of orthonormal basis implies that U is one-to-one. If $f \in \ell^2(A)$, then it is easy to check (similar to proof of (a) implies (c) Theorem 3.18), that $\sum_{\alpha \in A} f(\alpha) e_\alpha$ has atmost countably many non-zero terms and converges in $\mathcal{H}$ in the norm topology such that the limit does not depend on the order of terms. Furthermore $U(\sum_{\alpha \in A} f(\alpha) e_\alpha) = f$. This proves that U is surjective.

By Parseval's theorem,

$$\|U(x)\|_{\ell^2(A)} = \|x\|, \quad \text{for all } x \in \mathcal{H}.$$

This along with the polarization identity (Exercise 3.5) implies that U preserves inner product. ☐

4 Compact operators

We introduce the notion of compact operators. Throughout §4, we assume that $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach spaces over the field K, where K is $\mathbb{R}$ or $\mathbb{C}$. So by Proposition 2.16, the vector space of bounded linear maps $\mathcal{L}(X, Y)$ from X to Y equipped with the operator norm is a Banach space. Compact operators are a special class of bounded linear operators.

Definition 4.1. A bounded linear operator $T \in \mathcal{L}(X, Y)$ is said to be *compact* is image of every bounded set under T has compact closure.

There are various equivalent definitions of compact operators.

Exercise 4.2. Show that the following are equivalent for an operator $T \in \mathcal{L}(X, Y)$.

(a) The image of every bounded set under T has compact closure.

(b) The image of the unit ball (centered at origin) in X under T has compact closure.

(c) The image of the unit ball (centered at origin) is *totally bounded* (that is, for any $\epsilon > 0$, the image of the unit ball is covered by finitely many balls of radii ϵ).

(d) For any bounded sequence $(x_n)_{n\in\mathbb{N}}$ in X, the sequence $(T(x_n))_{n\in\mathbb{N}}$ has a convergent subsequence in Y.

Hint: recall Theorem 1.71.

Compact operators composed with bounded operators lead to compact operators as explained in the exercise below.

Exercise 4.3. Let X, Y, Z be Banach spaces over K and let $T \in \mathcal{L}(X,Y), S \in \mathcal{L}(Y,Z)$. Then $S \circ T \in \mathcal{K}(X,Z)$ if either $T \in \mathcal{K}(X,Y)$ or $S \in \mathcal{K}(Y,Z)$. (Hint: Use Exercise 4.2)

The set of compact operators in $\mathcal{L}(X,Y)$ is denoted by $\mathcal{K}(X,Y)$.

Proposition 4.4. *The set of compact operators $\mathcal{K}(X,Y)$ forms a closed subspace of $\mathcal{L}(X,Y)$.*

Proof. The fact that $\mathcal{K}(X,Y)$ is a subspace follows from the characterization in Exercise 4.2-(c). For any $a_1, a_2 \in K, T_1, T_2 \in \mathcal{K}(X,Y)$ and any bounded sequence $(x_n)_{n\in\mathbb{N}}$ in X, by the compactness of T_1 and T_2, there is a *common* subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $(T_1(x_{n_k}))_{k\in\mathbb{N}}$ and $(T_2(x_{n_k}))_{k\in\mathbb{N}}$ converge. Therefore $((a_1T_1 + a_2T_2)(x_{n_k}))_{k\in\mathbb{N}}$ converges and hence $a_1T_1 + a_2T_2 \in \mathcal{K}(X,Y)$.

We again use the characterization in Exercise 4.2-(c) to show that $\mathcal{K}(X,Y)$ is closed. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{K}(X,Y)$ such that it converges to $T \in \mathcal{L}(X,Y)$; that is $\lim_{n\to\infty} \|T - T_n\| = 0$. Let $(x_n)_{n\in\mathbb{N}}$ be any bounded sequence in X and let $B \in (0,\infty)$ be such that $\|x_n\| \leqslant B$ for all $n \in \mathbb{N}$. By the compactness of T_1, there exists a subsequence $(x_{n_{j,1}})_{j\in\mathbb{N}}$ such that $(T_1(x_{n_{j,1}}))_{j\in\mathbb{N}}$ converges. By the compactness of T_2, this subsequence in turn has a further subsequence subsequence $(x_{n_{j,2}})_{j\in\mathbb{N}}$ such that $(T_1(x_{n_{j,2}}))_{j\in\mathbb{N}}$ converges. Repeating this procedure, and choosing the diagonal subsequence

$$y_k = x_{n_{k,k}}, \quad \text{for all } k \in \mathbb{N},$$

we have that the sequence $(T_n(y_k))_{k\in\mathbb{N}}$ converges for each $n \in \mathbb{N}$. Hence

$$\begin{aligned}\|T(y_k) - T(y_l)\| &\leqslant \|T(y_k) - T_n(y_k)\| + \|T_n(y_k) - T_n(y_l)\| + \|T_n(y_k) - T_n(y_l)\| \\ &\leqslant \|T - T_n\| \|y_k\| + \|T_n(y_k) - T_n(y_l)\| + \|T - T_n\| \|y_l\| \\ &\leqslant 2\|T - T_n\| B + \|T_n(y_k) - T_n(y_l)\| .\end{aligned}$$

For any $\epsilon > 0$, there exists $n \in \mathbb{N}$ such that $2\|T - T_n\| B < \epsilon/2$ (since $\lim_{m\to\infty} \|T - T_m\| = 0$). Since $(T_n(y_m)))_{m\in\mathbb{N}}$ converges, there exists $N \in \mathbb{N}$ such that $\|T_n(y_k) - T_n(y_l)\| < \epsilon/2$ for all $k, l \geqslant N$. ☐

Recall the definition of kernel $\mathcal{N}(T)$ and range $\mathcal{R}(T)$ of a linear operator T from Definition 2.80.

Example 4.5. Let $T \in \mathcal{L}(X,Y)$ be a *finite rank* operator; that is, the range $\mathcal{R}(T) = \{T(x) : x \in X\}$ is finite dimensional. Then $T \in \mathcal{K}(X,Y)$.
Proof. If (x_n) is a bounded sequence in X, then $(T(x_n))_{n\in\mathbb{N}}$ is a bounded sequence in the range $\mathcal{R}(T)$. Since closed balls centered at origin is compact (see Exercise 2.20), $(T(x_n))_{n\in\mathbb{N}}$ has a bounded subsequence due to Theorem 1.75.

As a result of Proposition 4.4 and Example 4.5, any limit of finite rank bounded operators is compact.

Theorem 4.6 (Schauder's theorem)**.** *Let X, Y be Banach spaces over K and let $T \in \mathcal{K}(X,Y)$ be a compact operator. Then $T^* \in \mathcal{K}(Y^*, X^*)$.*

Proof. By Exercise 4.2, it suffices to show that the image $T^*(B_{Y^*})$ of the unit ball $B_{Y^*} = \{g \in Y^* : \|g\| < 1\}$ is totally bounded. To this end, let $\epsilon > 0$. By Exercise 4.2, the image $T(B_X)$ of the unit ball $B_X := \{x \in X : \|x\| < 1\}$ is totally bounded. Hence there exists $x_1, x_2, \ldots, x_n \in B_X$ such that

$$\min_{1\leqslant j\leqslant n} \|T(x) - T(x_j)\| < \frac{\epsilon}{3}, \quad \text{for all } x \in B_X. \tag{4.1}$$

Define $A : Y^* \to K^n$ as

$$A(g) = (g(T(x_1)), \ldots, g(T(x_n))), \quad \text{for all } g \in Y^*.$$

Here we equip K^n with the norm $\|(a_1, \ldots, a_n)\| = \sum_{i=1}^n |a_i|$ for all $(a_1, \ldots, a_n) \in K^n$. Clearly, A is linear. Note that $A \in \mathcal{L}(Y^*, K^n)$ as

$$\|A(g)\| = \sum_{i=1}^n |g(T(x_i))| \leqslant \sum_{i=1}^n \|g\| \|T(x_i)\| \leqslant \sum_{i=1}^n \|g\| \|T\| = n \|T\| \|g\|, \quad \text{for all } g \in Y^*.$$

By Exercise 4.5, A is a compact operator. Therefore by Exercise 4.2, there exist $g_1, \ldots, g_m \in B_{Y^*}$ such that

$$\min_{1\leqslant k\leqslant m} \|A(g) - A(g_k)\| < \frac{\epsilon}{3}, \quad \text{for all } g \in B_{Y^*}. \tag{4.2}$$

We claim that balls of radii ϵ centered at $T^*(g_1), \ldots, T^*(g_m)$ cover $T^*(B_{Y^*})$ or equivalently,

$$\min_{1\leqslant k\leqslant m} \|T^*(g) - T^*(g_k)\| < \epsilon, \quad \text{for all } g \in B_{Y^*}. \tag{4.3}$$

In order to prove (4.3), let $x \in B_X, g \in B_{Y^*}$ be arbitrary. By (4.2), there exists k such that $1 \leqslant k \leqslant m$ and $\|A(g) - A(g_k)\| < \frac{\epsilon}{3}$. By (4.1), there exists j such that $1 \leqslant j \leqslant n$ and $\|T(x) - T(x_j)\| < \frac{\epsilon}{3}$. Combining these we obtain

$$|T^*(g)(x) - T^*(g_k)(x)| = |g(T(x)) - g_k(T(x))|$$

$$\leqslant |g(T(x)) - g(T(x_j))| + |g(T(x_j)) - g_k(T(x_j))| + |g_k(T(x_j)) - g_k(T(x))|$$
$$\leqslant \|g\| \, \|T(x) - T(x_j)\| + |g(T(x_j)) - g_k(T(x_j))| + \|g_k\| \, \|T(x) - T(x_j)\|$$
$$\leqslant 2\,\|T(x) + T(x_j)\| + \|A(g) - A(g_k)\| < \epsilon.$$

Since the above estimate holds for all $x \in B_X, g \in B_{Y^*}$, we have

$$\min_{1 \leqslant k \leqslant m} \|T^*(g) - T^*(g_k)\| < \epsilon \quad \text{for all } g \in B_{Y^*}.$$

This implies $T^*(B_{Y^*})$ is totally bounded, or equivalently, $T^* \in \mathcal{K}(Y^*, X^*)$. □

Schauder's theorem can be used to prove its converse.

Exercise 4.7. Let X, Y be Banach spaces over K and let $T \in \mathcal{L}(X, Y)$ be such that $T^* \in \mathcal{K}(Y^*, X^*)$ is a compact operator. Then $T \in \mathcal{K}(X, Y)$. (Hint: Schauder's theorem implies that $T^{**} \in \mathcal{K}(X^{**}, Y^{**})$).

We would like to analyze the spectrum of a compact operator whose definition we introduce below.

Definition 4.8 (Spectrum and its classification)**.** Let X be a Banach space over K and let $T \in \mathcal{L}(X, X)$. Let I denote the identity map in $\mathcal{L}(X, X)$. The *resolvent set* of T, denoted by $\rho(T)$, is defined as

$$\rho(T) : \{\lambda \in K : (T - \lambda I) \text{ is a bijection from } X \text{ onto } X\}.$$

The *spectrum of* T, denoted by $\sigma(T)$, is the complement of the resolvent set, that is,

$$\sigma(T) = K \backslash \rho(T).$$

A number $\lambda \in K$ is said to be an *eigenvalue of* T if $\mathcal{N}(T - \lambda I) \neq \{0\}$. The set of all eigenvalues is called the *point spectrum* of T, denoted by $\sigma_p(T)$. If $\lambda \in \sigma(T) \backslash \sigma_p(T)$, then $T - \lambda I$ is injective but not surjective. In this case, we further subdivide the spectrum into two cases depending on whether or not the range $\mathcal{R}(T - \lambda I)$ is dense. If $\lambda \in \sigma(T) \backslash \sigma_p(T)$ and $\mathcal{R}(T - \lambda I)$ is dense in X, then we say that λ belongs to the *continuous spectrum* (denoted by $\sigma_c(T)$). If $\lambda \in \sigma(T) \backslash \sigma_p(T)$ and $\mathcal{R}(T - \lambda I)$ is not dense in X, then we say that λ belongs to the *residual spectrum* (denoted by $\sigma_r(T)$). This classification expresses the spectrum $\sigma(T)$ as a disjoint union $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

Remark 4.9. (a) Note that if $\lambda \in \rho(T)$, then $(T - \lambda I)^{-1} \in \mathcal{L}(X, X)$ due to open mapping principle (Corollary 2.67).

(b) If X is finite dimensional, then $\sigma(T) = \sigma_p(T)$. To see this, note that if $\lambda \notin \sigma_p(T)$, then $T - \lambda I$ is one-to-one and hence surjective (due to the rank-nullity theorem). So $\lambda \in \rho(T)$, or equivalently $\lambda \notin \sigma(T)$. However, on infinite dimensional spaces it is possible that $\sigma_p(T) \subsetneq \sigma(T)$ (see assignment for such an example).

(c) If $T \in \mathcal{K}(X,X)$ and $0 \in \rho(T)$, then X is finite dimensional. By (a), $T^{-1} \in \mathcal{L}(X,X)$ and hence by Exercise 4.3, $I = T \circ T^{-1} \in \mathcal{K}(X,X)$. So by Theorem 2.22, X is finite dimensional.

(d) Equivalently, if X is an infinite dimensional Banach space and $T \in \mathcal{K}(X,X)$, then $0 \in \sigma(T)$.

By Remark 4.9-(b), the spectrum can be considered as a generalization of the set of eigenvalues that we know from linear algebra. The following exercise outlines an argument that the spectrum of a bounded operator is a compact subset.

Exercise 4.10. Let X be a Banach space over K.

(i) If $T \in \mathcal{L}(X,X)$ and $\|I - T\| < 1$, then show that T is invertible and $T^{-1} \in \mathcal{L}(X,X)$. (Hint: show that $T^{-1} = \sum_{n=0}^{\infty}(I-T)^n$ and that the series converges is in the Banach space $\mathcal{L}(X,X)$).

(ii) If $S, T \in \mathcal{L}(X,X)$ is such that S is invertible with a bounded inverse $S^{-1} \in \mathcal{L}(X,X)$ and $\|S - T\| < \|S^{-1}\|^{-1}$, then show that T is invertible with a bounded inverse. (Hint: Note that (ii) is a generalization of (i)).

(iii) If $T \in \mathcal{L}(X,X)$ and $\lambda \in K$ is such that $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$. (Hint: Note that $T - \lambda I$ is a bijection if and only if $I - \lambda^{-1}T$ is a bijection and use (i)).

(iv) If $T \in \mathcal{L}(X,X)$, then show that the resolvent set $\rho(T)$ is open in K. (Hint: Use (ii) and Remark 4.9-(a))

(v) Conclude that the spectrum $\sigma(T)$ for any $T \in \mathcal{L}(X,X)$ is a compact subset of K.

4.1 Riesz theory of compact operators

Throughout §4.1, let X be a Banach space over K and let $T \in \mathcal{K}(X,X)$.

Lemma 4.11. *If $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in X and $((I-T)(x_n))_{n\in\mathbb{N}}$ converges, then $(x_n)_{n\in\mathbb{N}}$ has a convergent subsequence.*

Proof. Since T is compact, there is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $(T(x_{n_k}))_{k\in\mathbb{N}}$ converges, to say y. Then if $w = \lim_{n\to\infty}(I-T)(x_n)$, we have $\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty}(I - T)(x_{n_k}) + T(x_{n_k}) = w + y$. □

Definition 4.12. An operator $S \in \mathcal{L}(Y,Z)$ between normed linear spaces Y, Z is said to be *bounded below* if there exists $c \in (0,\infty)$ such that $\|S(y)\|_Z \geqslant c\,\|y\|_Y$ for all $y \in Y$.

Lemma 4.13. *If $I - T$ is one-to-one, then it is bounded below.*

Proof. Suppose to the contrary that $I - T$ is not bounded below. Then there is a sequence $(x_n)_{n\in\mathbb{N}}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \|(I-T)(x_n)\| = 0$. Therefore, by Lemma 4.11, (x_n) has a subsequence that converges, to say $x \in X$. By continuity of norm (see Exercise 2.3) and $I - T$, we have $\|x\| = 1$ and $(I - T)(x) = 0$, which contradicts the assumption that $I - T$ is one-to-one. □

Proposition 4.14. *Let X be a Banach space over K and let $T \in \mathcal{K}(X, X)$. Then $\mathcal{N}(I-T)$ is finite dimensional.*

Proof. Let $Y = \mathcal{N}(I - T)$. Note that $T\big|_Y = I\big|_Y$, so the closed unit ball in Y is compact but the compactness of T. Therefore Y is finite dimensional by Theorem 2.22. □

The following is an algebraic notion associated with a subspace of a vector space.

Definition 4.15 (Complement of a subspace)**.** Let X be a vector space over K and let Y be a subspace of X. Then a subspace M of X is said to be a *complement of Y in X* if the following hold:

(i) $Y \cap M = \{0\}$

(ii) $Y + M = X$; that is, for all $x \in X$, there exists $y \in Y, m \in M$ such that $x = y + m$.

If (i) and (ii) hold, we denote this by $Y \oplus M = X$. Note that (i) implies that the decomposition in $x = y + m$ in (ii) is unique.

Every finite dimensional subspace has a closed complement. This is a consequence of Hahn-Banach theorem.

Lemma 4.16. *Let F be a finite dimensional subspace of a normed vector space Y (over K). Then there is a closed subspace $M \subset Y$ such that $Y = F \oplus M$.*

Proof. Let $e_1, \ldots, e_n$ be a basis of F. For each $1 \leqslant i \leqslant n$, the linear functional $f_i : F \to K$ defined by

$$f_i\left(\sum_{j=1}^{n} a_j e_j\right) = a_i, \quad \text{for all } (a_1, \ldots, a_n) \in K^n$$

is a bounded linear functional (by Exercise 2.25). So for each $1 \leqslant i \leqslant n$, by the Hahn-Banach extension theorem, there is a bounded extension $F_i : X \to K$ such that $f_i = F_i\big|_F$. Define an operator $P_F : Y \to Y$ such that

$$P_F(y) = \sum_{i=1}^{n} F_j(y) e_j.$$

Note that, $P_F \in \mathcal{L}(Y, Y)$, $P_F(x) = x$ for all $x \in F$, $\mathcal{R}(P_F) = F$ and $P_F \circ P_F = P_F$ (you check this!).

Now let $M = \mathcal{N}(P_F)$. Since P_F is continuous M is closed and $M \cap F = \{0\}$. If $y \in Y$, then $y = P_F(y) + (y - P_F(y))$, where $P_F(y) \in F$ and $y - P_F(y) \in M$ (since $P_F \circ (I - P_F) = 0$). □

Proposition 4.17. *Let X be a Banach space over K and let $T \in \mathcal{K}(X, X)$. Then $\mathcal{R}(I-T)$ is closed.*

Proof. Since the kernel $\mathcal{N}(I-T)$ is finite dimensional (by Proposition 4.14), it has a closed complementary subspace M by Lemma 4.16 such that $X = \mathcal{N}(I-T) \oplus M$. So $\mathcal{R}(I-T) = \mathcal{R}((I-T)\big|_M)$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in M such that $(I-T)(x_n)$ converges, to say $w \in X$. We need to show that $w \in \mathcal{R}(I-T)$. By the argument in Lemma 4.13, $(I-T)\big|_M$ is bounded below. So $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence, and converges, to say $x \in M$. Therefore $w = (I-T)(x) \in \mathcal{R}(I-T)$. □

Proposition 4.18. *Let X be a Banach space over K and let $T \in \mathcal{K}(X, X)$. Then the following are equivalent:*

(a) $I - T$ is onto.

(b) $I - T$ is one-to-one.

Proof. (a) $\Longrightarrow$ (b): Let $X_0 = \{0\}$ and for $n \in \mathbb{N}$, set $X_n = \mathcal{N}((I-T)^n)$, where $(I-T)^n$ is the n-fold composition of $I-T$. Clearly, $\{0\} = X_0 \subset X_1 \subset X_2 \subset \cdot$. Note that each X_n is a closed subspace by Lemma 2.18.

Assume to the contrary that if $I-T$ is not one-to-one; that is, there exists $x_1 \in X_1 \backslash X_0$. Then since $I-T$ is onto, there exists $x_2 \in X$ such that $(I-T)(x_2) = x_1$. By induction, we obtain $x_n \in X$ such that $(I-T)(x_n) = x_{n-1}$ for all $n \geqslant 2$. Therefore, $(I-T)^n(x_n) = (I-T)(x_1) = 0$ and $(I-T)^{n-1}(x_n) = x_1$ for all $n \geqslant 2$. So $x_n \in X_n \backslash X_{n-1}$ for all $n \in \mathbb{N}$.

By Lemma 2.21, for each $n \in \mathbb{N}$, there exists $y_n \in X_n$ with $\|y_n\| = 1$ and $\|y_n - x\| \geqslant \frac{1}{2}$ for all $x \in X_{n-1}$. For any $n > m$, we have

$$T(y_n) - T(y_m) = y_n - (y_m - (I-T)(y_m) + (I-T)(y_n)) = y_n - x, \quad \text{where } x \in X_{n-1}.$$

So $\|T(y_n) - T(y_m)\| \geqslant \frac{1}{2}$ for any $m \neq n$, which contradicts the compactness of T.
(b) $\Longrightarrow$ (a): Suppose $I-T$ is one-to-one. Then by Lemma 4.13, $I-T$ is bounded below. So if M is a closed subspace of X, then so is $(I-T)(M) := \{(I-T)x : x \in M\}$ (why?).

Suppose to the contrary that $I-T$ is not onto. Let $Y_0 = X, Y_1 = (I-T)(Y_0), Y_2 = (I-T)(Y_1), \ldots$. Since $(I-T)$ is one-to-one and not onto, we have that Y_{m+1} is a proper closed subspace of Y_m for all m. By Lemma 2.21, there exists $z_n \in Y_n$ such that $\|z_n\| = 1$ and $\|z_n - x\| \geqslant \frac{1}{2}$ for all $x \in Y_{n+1}$ for all $n \in \mathbb{N}$. For any $n > m$, we have

$$T(z_m) - T(z_n) = z_m - (z_n + (I-T)(z_m) - (I-T)(z_n)) = y_n - x, \quad \text{where } x \in X_{m+1}.$$

So $\|T(z_n) - T(z_m)\| \geqslant \frac{1}{2}$ for any $m \neq n$, which contradicts the compactness of T. □

The following theorem describes important properties of the spectrum of a compact operator.

Theorem 4.19. *Let X be a Banach space over K and let $T \in \mathcal{K}(X, X)$ be a compact operator. Then we have the following.*

(i) $\sigma(T)\backslash\{0\} \subset \sigma_p(T)$*; that is, all non-zero points in the spectrum are eigenvalues. Furthermore, if* $\lambda \in \sigma(T)\backslash\{0\}$*, then* $\dim(\mathcal{N}(T-\lambda I)) < \infty$*; that is, each non-zero eigenvalue has finite multiplicity.*

(ii) If $(\lambda_n)_{n\in\mathbb{N}}$ *is a sequence of distinct points in* $\sigma(T)$ *such that* $\lim_{n\to\infty}\lambda_n = \lambda$*, then* $\lambda = 0$*. In other words, the only possible cluster point in the spectrum is zero, or equivalently, all non-zero points in the spectrum are isolated.*

Proof. (i) Let $\lambda \in \sigma(T)\backslash\{0\}$. Then $I - \lambda^{-1}T$ is not invertible. So $I - \lambda^{-1}T$ is either not one-to-one or not onto. By Proposition 4.18, we have that $I - \lambda^{-1}T$ is not one-to-one in both cases. Therefore $\lambda \in \sigma_p(T)$. Note that by Proposition 4.14, we have

$$\dim(\mathcal{N}(T-\lambda I)) = \dim(\mathcal{N}(I-\lambda^{-1}T)) < \infty.$$

(ii) By (a) and removing one point in the sequence if necessary, we assume that $\lambda_n \neq 0$ for all $n \in \mathbb{N}$ so that $\lambda_n \in \sigma_p(T)$ for all $n \in \mathbb{N}$. So for each $n \in \mathbb{N}$, there exists 'eigenvectors' $x_n \in X\backslash\{0\}$ such that $T(x_n) = \lambda_n x_n$.

We claim that $\{x_i : i \in \mathbb{N}\}$ is a linearly independent set. We prove this by induction on n, by verifying that $\{x_i : 1 \leqslant i \leqslant n\}$ is linearly independent for each $n \in \mathbb{N}$. Clearly, this is true for $n = 1$ as $x_1 \neq 0$. Suppose the induction hypothesis that $\{x_i : 1 \leqslant i \leqslant n\}$ is linearly independent. Assume to the contrary that $\{x_i : 1 \leqslant i \leqslant n+1\}$ is not linearly independent. Then there are scalars $a_1, \ldots, a_n \in K$ such that $x_{n+1} = \sum_{i=1}^n a_i x_i$. Hence

$$\sum_{i=1}^n \lambda_{n+1} a_i x_i = \lambda_{n+1} x_{n+1} = T(x_{n+1}) = \sum_{i=1}^n a_i T(x_i) = \sum_{i=1}^n a_i \lambda_i x_i.$$

So $a_i(\lambda_{n+1} - \lambda_i) = 0$ for all $1 \leqslant i \leqslant n$, which in turn implies $a_i = 0$ for all $1 \leqslant i \leqslant n$, which contradicts $x_{n+1} \neq 0$

Define $X_n = \operatorname{span}\{x_i : 1 \leqslant i \leqslant n\}$ for all $n \in \mathbb{N}$ and $X_0 = \{0\}$. By the linear independence of $\{x_i : i \in \mathbb{N}\}$ and Exercise 2.20-(e), X_n is a proper, closed subspace of X_{n+1} for all $n \in \mathbb{N}$. So by Riesz's lemma (Lemma 2.21), for each $n \in \mathbb{N}$, there exists $y_n \in X_n$ such that $\|y_n\| = 1$ and $\|y_n - x\| \geqslant \frac{1}{2}$ for all $x \in X_{n-1}$. By writing any vector X_n in terms of the basis $\{x_i : 1 \leqslant i \leqslant n\}$, we note that

$$(T - \lambda_n I)(x) \in X_{n-1}, \quad \text{for all } x \in X_n \text{ and } n \in \mathbb{N}.$$

Therefore for any $n > m$, we have

$$T(\lambda_n^{-1} y_n) - T(\lambda_m^{-1} y_m) = (T - \lambda_n I)(\lambda_n^{-1} y_n) - (T - \lambda_m I)(\lambda_m^{-1} y_m) + y_n - y_m = y_n - x,$$

where $x \in X_{n-1}$. So for any $n > m$, we have

$$\left\|T(\lambda_n^{-1} y_n) - T(\lambda_m^{-1} y_m)\right\| \geqslant \frac{1}{2}. \tag{4.4}$$

If $\lambda_n \to \lambda$ and $\lambda \neq 0$, then (4.4) would contradict the compactness of T as $(\lambda_n^{-1} y_n)_{n\in\mathbb{N}}$ is a bounded sequence such that $(T(\lambda_n^{-1} y_n))_{n\in\mathbb{N}}$ does not have a convergent subsequence. Hence $\lambda = 0$.

□

Exercise 4.20. Let $1 \leqslant p \leqslant \infty$ and let $X = \ell^p(\mathbb{N})$. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a bounded sequence of complex numbers and $T \in \mathcal{L}(X, X)$ is defined by

$$T((x_1, x_2, \ldots)) = (\lambda_1 x_1, \lambda_2 x_2, \ldots).$$

Prove that $T \in \mathcal{K}(X, X)$ if and only if $\lim_{n\to\infty} \lambda_n = 0$.

Exercise 4.21. Let X, Y be Banach spaces and let $T \in \mathcal{K}(X, Y)$. Assume that the range $\mathcal{R}(T)$ is a closed subspace of Y.

(a) Show that T is a finite rank operator (cf. Example 4.5). Hint: Use the open mapping principle.

(b) If in addition the nullspace $\mathcal{N}(T)$ is finite-dimensional, then X is a finite dimensional space.

4.2 Spectral decomposition of self-adjoint compact operators

Definition 4.22. Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$. We say that T is *self-adjoint* if $T^\dagger = T$.

The following exercise outlines some useful relations between the spectrum of an operator and its adjoint.

Exercise 4.23. Let $T \in \mathcal{L}(X, X)$ where X is a Banach space. Show the following.

1. If λ is in the residual spectrum of T, then show that λ is in the point spectrum of $T^* \in \mathcal{L}(X^*, X^*)$; that is $\sigma_r(T) \subset \sigma_p(T^*)$.

2. If λ is in the point spectrum of T^*, then show that λ is either in the point spectrum or the residual spectrum of T. In other words, $\sigma_p(T^*) \subset \sigma_p(T) \cup \sigma_r(T)$.

(c) Let $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ be operator on a Hilbert space $\mathcal{H}$. Let $T^* \in \mathcal{L}(\mathcal{H}^*, \mathcal{H}^*)$ and $T^\dagger \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ denote the adjoint and the Hilbert space adjoint respectively. Then show the following

$$\sigma_p(T^*) = \overline{\sigma_p(T^\dagger)}, \quad \sigma_c(T^*) = \overline{\sigma_c(T^\dagger)}, \quad \sigma_r(T^*) = \overline{\sigma_r(T^\dagger)}.$$

Here for $A \subset \mathbb{C}$, we set $\overline{A} = \{\overline{a} : a \in A\}$.

The following theorem lists some basic properties of the spectrum of a self-adjoint operator.

Theorem 4.24. *Let $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ be a self-adjoint operator on a Hilbert space $\mathcal{H}$.*

(a) $\langle T(u), u\rangle \in \mathbb{R}$ for all $u \in \mathcal{H}$.

(b) If $M := \sup_{\substack{u\in\mathcal{H},\\ \|u\|=1}} \langle T(u), u\rangle$ and $m := \inf_{\substack{u\in\mathcal{H},\\ \|u\|=1}} \langle T(u), u\rangle$, then we have

$$\{m, M\} \subset \sigma(T) \subset [m, M], \quad \|T\| = \max(|m|, |M|).$$

(c) If $\sigma(T) = \{0\}$, then $T = 0$.

Proof. (a) For any $u \in \mathcal{H}$, since $T^\dagger = T$ we have

$$\overline{\langle T(u), u\rangle} = \langle u, T(u)\rangle = \langle u, T^\dagger(u)\rangle \overset{(3.12)}{=} \langle T(u), u\rangle.$$

Therefore $\langle T(u), u\rangle \in \mathbb{R}$ for all $u \in \mathcal{H}$.

(b) **Step 1: $\sigma(T) \subset [m, M]$.** Let $\lambda \in \mathbb{C}$. Then by (a), we have

$$\mathrm{Re}(\langle (T-\lambda I)u, u\rangle) = \langle T(u), u\rangle - \mathrm{Re}(\lambda)\langle u, u\rangle, \quad \mathrm{Im}(\langle (T-\lambda I)u, u\rangle) = -\mathrm{Im}(\lambda)\langle u, u\rangle$$

Setting $c = \max(|\mathrm{Im}(\lambda)|, \mathrm{Re}(\lambda) - M, m - \mathrm{Re}(\lambda))$, by the above equality and Schwarz inequality, we have

$$\|(T-\lambda I)(u)\| \, \|u\| \geqslant |\langle (T-\lambda I)u, u\rangle| \geqslant c \, \|u\|^2 .$$

Note that $c > 0$ if $\lambda \notin [m, M]$ and hence $T - \lambda I$ is bounded below for all $\lambda \in \mathbb{C}\backslash[m, M]$. This implies that $T - \lambda I$ is injective and has a closed range. Therefore $\lambda \in \mathbb{C}\backslash[m, M]$ implies that $\lambda \in \rho(T) \cup \sigma_r(T)$, or equivalently,

$$\mathbb{C}\backslash[m, M] \subset \rho(T) \cup \sigma_r(T).$$

We claim that $\mathbb{C}\backslash[m, M] \subset \rho(T)$. Suppose to the contrary that $\lambda \in \sigma_r(T) \cap (\mathbb{C}\backslash[m, M])$, then by Exercise 4.23 we have $\lambda \in \sigma_p(T^*)$ which in turn implies $\overline{\lambda} \in \sigma_p(T^\dagger) = \sigma_p(T)$. Since $\lambda \in \mathbb{C}\backslash[m, M]$, we have $\overline{\lambda} \in \sigma_p(T) \cap (\mathbb{C}\backslash[m, M]) \subset \sigma_p(T) \cap (\rho(T) \cup \sigma_r(T)) = \varnothing$, a contradiction. Therefore $\mathbb{C}\backslash[m, M] \subset \rho(T)$ or equivalently,

$$\sigma(T) \subset [m, M].$$

Step 2: $m, M \subset \sigma(T)$. We define a function $a : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that

$$a(u, v) := \langle (Mu - T(u), v\rangle.$$

It is easy to obtain the following (inner product-like) properties:

- $a(u, u) \in [0, \infty)$ for all $u \in \mathcal{H}$.
- $a(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 a(u_1, v) + \lambda_2 a(u_2, v)$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$ and $u_1, u_2 \in \mathcal{H}$.

- For all $u, v \in \mathcal{H}$, $\overline{a(u,v)} = a(v,u)$. To see this note, that

$$\begin{aligned}\overline{a(u,v)} &= \overline{M\langle u,v\rangle - \langle T(u),v\rangle} = M\langle v,u\rangle - \langle v,T(u)\rangle \\ &= M\langle v,u\rangle - \langle v,T^\dagger(u)\rangle = M\langle v,u\rangle - \langle T(v),u\rangle = a(v,u).\end{aligned}$$

These properties imply the Scharwz inequality,

$$|a(u,v)| \leqslant \sqrt{a(u,u)a(v,v)}, \tag{4.5}$$

as the quadratic function $f(t) = a(u+tv, u+tv)$ for all $t \in \mathbb{R}$ does not have two distinct real roots. Since $M = \sup_{\substack{u\in\mathcal{H},\\ \|u\|=1}} \langle T(u), u\rangle$, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathcal{H}$ such that

$$\|u_n\| = 1 \quad \text{for all } n \in \mathbb{N} \text{ and,} \quad \lim_{n\to\infty} \langle Mu_n - T(u_n), u_n\rangle = 0. \tag{4.6}$$

By Riesz-Fréchet representation theorem (Theorem 3.14) and Hahn-Banach extension theorem (Corollary 2.30-(i)), for all $w \in \mathcal{H}$, we have

$$\|w\| = \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} |\langle w, v\rangle|. \tag{4.7}$$

By Schwarz inequality, for any $v \in \mathcal{H}$ with $\|v\| = 1$, we have

$$|a(v,v)| = |\langle Mv - T(v), v\rangle| \leqslant \|Mv - T(v)\| \, \|v\| \leqslant \|MI - T\| \leqslant M + \|T\|. \tag{4.8}$$

Hence for all $n \in \mathbb{N}$,

$$\|Mu_n - T(u_n)\| \overset{(4.7)}{=} \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} |a(u_n,v)| \overset{(4.5)}{\leqslant} \sqrt{a(u_n,u_n)} \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} a(v,v) \overset{(4.8)}{\leqslant} (M+\|T\|)\sqrt{a(u_n,u_n)}.$$

Hence by (4.6), we conclude,

$$\lim_{n\to\infty} \|Mu_n - T(u_n)\| = 0, \quad \text{with } \|u_n\| = 1 \text{ for all } n \in \mathbb{N}. \tag{4.9}$$

This implies that $MI - T$ does not have a bounded inverse (since if $(MI-T)^{-1} \in \mathcal{L}(\mathcal{H},\mathcal{H})$, we have $\|u\| \leqslant \|(MI-T)^{-1}\| \, \|Mu - T(u)\|$ for all $u \in \mathcal{H}$ which contradicts (4.9)). Therefore $M \in \sigma(T)$.

The proof that $m \in \sigma(T)$ is similar by considering the function $b(u,v) = \langle T(u) - mu, v\rangle$ (or by replacing T with $-T$).

Step 3: $\|T\| = \max(|m|, |M)|$. The desired lower bound on $\|T\|$ follows easily from Schwarz inequality as

$$\max(|m|,|M|) = \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} |\langle T(v), v\rangle| \leqslant \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} \|T(v)\| \, \|v\| = \sup_{\substack{v\in\mathcal{H},\\ \|v\|=1}} \|T(v)\| = \|T\|.$$

For each $u, v \in \mathcal{H}$, there exists $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ so that $|\langle T(u), v\rangle| = \langle T(u), \alpha v\rangle$ (for example, $\alpha = \frac{\langle T(u), \alpha v\rangle}{|\langle T(u), \alpha v\rangle|}$ if $\langle T(u), \alpha v\rangle \neq 0$ and 1 otherwise). This along with (4.7) implies that

$$\|T(u)\| = \sup_{\substack{v \in \mathcal{H}, \\ \|v\|=1}} \operatorname{Re}(\langle T(u), v\rangle)$$

and hence

$$\|T\| = \sup_{\substack{u,v \in \mathcal{H}, \\ \|u\|=\|v\|=1}} \operatorname{Re}(\langle T(u), v\rangle) \tag{4.10}$$

Since $T = T^\dagger$ using properties of inner product (why?), we obtain

$$\begin{aligned} \operatorname{Re}(\langle T(u), v\rangle) &= \frac{1}{2}\left(\langle T(u), v\rangle + \langle v, T(u)\rangle\right) = \frac{1}{2}\left(\langle T(u), v\rangle + \langle v, T^\dagger(u)\rangle\right) \\ &= \frac{1}{2}\left(\langle T(u), v\rangle + \langle T(v), u\rangle\right) \\ &= \frac{1}{4}\left(\langle T(u+v), u+v\rangle - \langle T(u-v), u-v\rangle\right), \quad \text{for all } u, v \in \mathcal{H}. \end{aligned}$$

Using the definition of M and m, for all $u, v \in \mathcal{H}$ with $\|u\| = \|v\| = 1$, we have

$$\begin{aligned} \operatorname{Re}(\langle T(u), v\rangle) &\leqslant \frac{1}{4}\left(M\langle u+v, u+v\rangle - m\langle u-v, u-v\rangle\right) \\ &\leqslant \frac{1}{4}\left(|M|\langle u+v, u+v\rangle + |m|\langle u-v, u-v\rangle\right) \\ &\leqslant \frac{1}{4}\max(|m|, |M|)\left(\langle u+v, u+v\rangle + \langle u-v, u-v\rangle\right) \\ &= \max(|m|, |M|)\frac{1}{2}\left(\|u\|^2 + \|v\|^2\right) = \max(|m|, |M|). \end{aligned}$$

Hence (4.10) implies the desired upper bound

$$\|T\| \leqslant \max(|m|, |M|).$$

(c) Since $\{0\} = \sigma(T) \supset \{m, M\}$, we have $M = m = 0$. Since $\|T\| = \max(|m|, |M|) = 0$, we conclude $T = 0$.

□

We can now combine the spectral theorems for compact operators (Theorem 4.19) and for self-adjoint operators (Theorem 4.24) to obtain a spectral theorem for self-adjoint compact operators. This can be viewed as a generalization of the result in linear algebra that any Hermitian matrix can be diagonalized.

Theorem 4.25. *Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{K}(\mathcal{H}, \mathcal{H})$ be a self-adjoint compact operator. Then there exists an orthonormal basis $\{e_i : i \in I\}$ of $\mathcal{H}$ such that for each $i \in I$, e_i is an eigenvector of T.*

Proof. For $\lambda \in \sigma_p(T)$, let $E_\lambda = \mathcal{N}(T - \lambda I)$ denote the corresponding eigenspace. We claim that if $\lambda, \mu \in \sigma_p(T)$ are distinct ($\mu \neq \lambda$), then E_λ and E_μ are mutually orthogonal; that is,

$$\langle u, v \rangle = 0, \quad \text{for all } u \in E_\lambda, v \in E_\mu. \tag{4.11}$$

To prove this, let λ, μ be two distinct eigenvalues. By Theorem 4.24-(b), we have $\lambda, \mu \in \mathbb{R}$. Hence for all $u \in E_\lambda, v \in E_\mu$, we have

$$\lambda \langle u, v \rangle = \langle T(u), v \rangle = \langle u, T^\dagger(v) \rangle = \langle u, T(v) \rangle = \overline{\mu} \langle u, v \rangle = \mu \langle u, v \rangle.$$

This implies $(\lambda - \mu)\langle u, v \rangle = 0$ and hence $\langle u, v \rangle = 0$ which prove (4.11)

Since for each $\lambda \in \sigma_p(T)$, E_λ is a closed subspace of a Hilbert space (being the kernel of a bounded operator), E_λ is a Hilbert space and hence admits an orthonormal basis B_λ of E_λ. By (4.11), we have that

$$B = \bigcup_{\lambda \in \sigma_p(T)} B_\lambda$$

is an orthonormal set in $\mathcal{H}$ such that each vector in B is an eigenvector. We claim that B is an orthonormal basis of $\mathcal{H}$. We verify the completeness of B. To this end set

$$M = \{v \in \mathcal{H} : \langle v, e \rangle = 0 \text{ for all } e \in B\}.$$

Note that M is a closed subspace of $\mathcal{H}$ (why?) and is a Hilbert space. We claim that $T(M) \subset M$. To see this note that for any $v \in M$, $e \in B_\lambda, \lambda \in \sigma_p(T)$, we have

$$\langle T(v), e \rangle = \langle v, T^\dagger(e) \rangle = \langle v, T(e) \rangle = \overline{\lambda} \langle v, e \rangle = 0.$$

This implies that the restriction $T\big|_M : M \to M$ can be viewed as a compact, self-adjoint operator on the Hilbert space M. We claim that

$$\sigma(T\big|_M) = \{0\}. \tag{4.12}$$

Suppose to the contrary, if $\lambda \in \sigma(T\big|_M) \backslash \{0\}$, then by Theorem 4.24-(i) we have $\lambda \in \sigma_p(T)$ and there exists $v \in M \cap E_\lambda$ with $v \neq 0$. This implies the existence of $v \neq 0$ with $v \in E_\lambda$ and $\langle v, e \rangle = 0$ for all $e \in B_\lambda$ which contradicts the completeness of orthonormal set B_λ. This completes the proof of (4.12). Now by Theorem 4.24-(c), we have $T(v) = 0$ for all $v \in M$. We claim that $M = \{0\}$. Suppose not. Then $0 \in \sigma_p(T)$ and there exists $v \in E_0 \cap M$ with $v \neq 0$. As above we obtain a contradiction to the completeness of orthonormal set B_0. Therefore $M = \{0\}$, or equivalently, B is a complete orthonormal set (orthonormal basis). □

Remark 4.26. Every compact, self-adjoint operator T on a Hilbert space can be viewed as a limit of finite rank operators.

To see this, we use Theorem 4.19 and Exercise 4.10-(iii), to conclude the set of eigenvalues $\sigma_p(T)$ is countable (finite or infinite). This is because the set $\{\lambda \in \sigma_p(T) : |\lambda| > n^{-1}\} \subset \{\lambda \in \mathbb{C} : n^{-1} \leqslant |\lambda| \leqslant \|T\|\}$ is finite (otherwise, we obtain a contradiction to Theorem 4.19-(ii) as $\{\lambda \in \mathbb{C} : n^{-1} \leqslant |\lambda| \leqslant \|T\|\}$ is compact).

Let $\{\lambda_n : n \in \mathbb{N}, 1 \leqslant n \leqslant N\}$ where $N \in \mathbb{N} \cup \infty$ denote an enumeration of $\sigma(T)\backslash\{0\} = \sigma_p(T)\backslash\{0\}$. By Theorem 4.25, T is a limit of $T_n = \sum_{k=1}^{n} \lambda_n P_{E_{\lambda_n}}$, where $P_{E_{\lambda_n}}$ is the projection map (see Corollary 3.13) to the (closed) eigenspace $E_{\lambda_n} = \mathcal{N}(T - \lambda_n I)$ (why?).

4.3 The adjoint operator revisited

We will show that several properties of an operator is shared by its adjoint operator. We already encountered one such result as we showed that compactness of a operator is equivalent to its adjoint (Theorem 4.6 and Exercise 4.7).

Theorem 4.27. *Let X, Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$. Then the following are equivalent.*

(a) T is a bijection.

(b) $T^* \in \mathcal{L}(Y^*, X^*)$ is a bijection.

Proof. (a) $\implies$ (b): Let $S = T^{-1} : Y \to X$. Then by Corollary 2.67, $S \in \mathcal{L}(Y, X)$ with $S \circ T = I_X$ and $T \circ S = I_Y$. By Exercise 2.41-(a), we have

$$S^* \circ T^* = (T \circ S)^* = I_Y^* = I_{Y^*}, \quad T^* \circ S^* = (S \circ T)^* = I_X^* = I_{X^*}.$$

Therefore T^* is a bijection.
(b) $\implies$ (a): Conversely, if T^* is a bijection, then by Theorem 2.65 (open mapping principle) there exists $c > 0$ such that

$$T^*(B_{Y^*}(0,1)) \supset T^*(B_{X^*}(0,c)).$$

Therefore for any $x \in X$, by Corollary 2.30-(i) we have

$$\|T(x)\| = \sup_{f \in B_{Y^*}(0,1)} |f(T(x))| = \sup_{g \in T^*(B_{Y^*}(0,1))} |g(x)| \geqslant \sup_{g \in B_{X^*}(0,c)} |g(x)| \geqslant c\,\|x\|\,, \quad \text{for all } x \in X.$$

This implies that T is injective as $T(x) = 0$ implies $c\,\|x\| \leqslant \|T(x)\| = 0$ which in turm implies $x = 0$.

It remains to show that T is onto. Next, we show that $\mathcal{R}(T)$ is closed. To see this, let $y \in \overline{\mathcal{R}(T)}$, then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\lim_{n\to\infty} T(x_n) = y$. Since $\|x_n - x_m\| \leqslant c^{-1}\,\|T(x_n) - T(x_n)\|$ for all $n, m \in \mathbb{N}$, we conclude that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X and hence converges to $x = \lim_{n\to\infty} x_n$. By the continuity of T, we have $y = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x) \in \mathcal{R}(T)$. Therefore $\mathcal{R}(T)$ is closed; that is $\mathcal{R}(T) = \overline{\mathcal{R}(T)}$. Since T^* is injective, by Exercise 2.42, $\overline{\mathcal{R}(T)} = Y$ and hence $\mathcal{R}(T) = \overline{\mathcal{R}(T)} = Y$. ☐

Corollary 4.28. *Let X be a Banach space and let $T \in \mathcal{L}(X, X)$. Then $\sigma(T) = \sigma(T^*)$.*

Proof. Note that $(T - \lambda I_X)^* = T^* - \lambda I_{X^*}$ (by Exercise 2.41-(b)). Therefore by Theorem 4.27, $\lambda \in \rho(T)$ if and only if $\lambda \in \rho(T^*)$. ☐

Definition 4.29. Let X be a normed vector space and let X^* denote the dual space. For a subspace M of X, we set

$$M^\perp := \{f \in X^* : f(x) = 0 \ \ \forall x \in M\}\,.$$

For a subspace N of X^*, we define

$$N^\perp := \{x \in X : f(x) = 0 \ \ \forall f \in N\}\,.$$

Exercise 4.30. Let X be a normed vector space and let X^* denote the dual space. Let M and N be subspaces X and X^* respectively. Show the following:

(a) $M^\perp$ and $N^\perp$ are closed subspaces of X^* and X respectively.

(b) $\overline{M} = (M^\perp)^\perp$.

(c) $(N^\perp)^\perp \supset \overline{N}$.

Lemma 4.31. *Let X and Y be Banach spaces over K. Let $T \in \mathcal{L}(X, Y)$ be such that T^* is injective and has a closed range. Then T is onto.*

Proof. We may assume that $K = \mathbb{R}$ without loss of generality by considering X and Y are vector spaces over $\mathbb{R}$. If $K = \mathbb{C}$, then the injectivity of T^* and the closed range property for the real case follows from Proposition 2.24.

Note that since $T^* \in \mathcal{L}(Y^*, X^*)$ is injective and closed range, T^* viewed as a bijection from the Banach space Y^* to the Banach space $T^*(Y^*)$. By open mapping principle (Corollary 2.67), there exists $c > 0$ such that

$$\|T^*(f)\| \geqslant c\,\|f\|, \quad \text{for all } f \in Y^*. \tag{4.13}$$

We claim that

$$\overline{T(B_X(0,1))} \supset B_Y(0,c), \tag{4.14}$$

where $B_X(x,r)$ (respectively, $B_Y(x,r)$) denotes an open ball with center x and radius r in X (respectively, Y). We prove (4.14) by contradiction. To this end, assume to the contrary that $y \in B_Y(0,c)\backslash\overline{T(B_X(0,1))}$ Since $\overline{T(B_X(0,1))}$ is closed, convex set (being a closure of a convex set; see HW 4, Question 1) and $f \notin \overline{T(B_X(0,1))}$ by Hahn-Banach separation theorem (Theorem 2.38-(2)) there exists $\alpha \in \mathbb{R}$ and $f \in Y^*$ such that

$$|(T^*(f))(x)| = |f(T(x))| \leqslant \alpha < f(y), \quad \text{for all } x \in B_X(0,1). \tag{4.15}$$

By (4.15) and $y \in B_Y(0,c)$, we have

$$\|f\| \geqslant \frac{|f(y)|}{\|y\|} > \frac{\alpha}{c}, \quad \|T^*(f)\| = \sup_{x \in B_X(0,1)} |(T^*(f))(x)| \leqslant \alpha. \tag{4.16}$$

By (4.13) and (4.16), we obtain the desired contradiction. This proves (4.14).

The same argument in the proof of Step 2 of Theorem 2.65 (open mapping principle) implies that

$$T(B_X(0,1)) \supset B_Y(0,c/2).$$

Hence S is surjective. □

Exercise 4.32 (Quotient space)**.** Let $(X, \|\cdot\|)$ be a normed vector space and let M be a proper, closed subspace of X. We define an *equivalence relation* $\sim$ on X as $x \sim y$ if and only if $x - y \in M$. We denote the equivalence class containing $x \in X$ by $x + M = \{x + y : y \in M\}$ and the collection of equivalence classes $\{x + M : x \in M\}$ by X/M.

1. Check that $\sim$ defines an equivalence relation on X and that X/M is a vector space equipped with the scalar multiplication and vector addition defined by

$$\alpha(x+M) = (\alpha x) + M, \quad (x+M) + (y+M) = (x+y) + M$$

for all $\alpha \in K$ and $x, y \in X$. Verify that the scalar multiplication and vector addition operations defined above are well-defined in the sense that they do not depend on the choice of representatives from the equivalence class.

2. Show that the *quotient norm* on X/M defined by

$$\|x + M\| := \inf_{y \in M} \|x + y\|$$

is a norm on the vector space X/M.

3. Let $\pi : X \to X/M$ denote the map defined by

$$\pi(x) = x + M, \quad \text{for all } x \in M.$$

Show that π is a bounded linear map whose operator norm is one. (Hint: Use Riesz's lemma).

4. Let $B_X = \{x \in X : \|x\| < 1\}$ and $B_{X/M} = \{x + M \in X/M : \|x + M\| < 1\}$ denote the open unit balls centered at zero in X and X/M respectively. Show that $\pi(B_X) = B_{X/M}$.

5. Show that a subset $U \subset X/M$ is open in X/M if and only if $\pi^{-1}(U)$ is open in X.

6. Show that if X is complete, then so is X/M equipped with the quotient norm.

Theorem 4.33 (Closed range theorem)**.** *Let X and Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$. Then $\mathcal{R}(T)$ is closed (in Y) if and only if $\mathcal{R}(T^*)$ is closed (in X^*).*

Proof. $\Longrightarrow$ Let $\mathcal{R}(T)$ be closed. Then the map $\widetilde{T} : X/\mathcal{N}(T) \to \mathcal{R}(T)$ defined by $\widetilde{T}(x + \mathcal{N}(T)) = T(x)$ is (well-defined) a bounded (since by Exercise 4.32-(4), we have $\left\|\widetilde{T}\right\| = \|T\|$), linear bijection between Banach spaces (see Exercise 4.32-(6) and recall that $\mathcal{R}(T)$ being a closed subspace of a Banach space is a Banach space). Therefore by Corollary 2.67, $\widetilde{T}$ is invertible and $\widetilde{T}^{-1} \in \mathcal{L}(\mathcal{R}(T), X/\mathcal{N}(T))$. Let $\pi \in \mathcal{L}(X, X/\mathcal{N}(T))$ denote the quotient map as defined in Exercise 4.32-(3) and let $\iota \in \mathcal{L}(\mathcal{R}(T), Y)$ denote the inclusion map, so that we have

$$T = \iota \circ \widetilde{T} \circ \pi.$$

By Exercise 2.41-(a), we have

$$T^* = \pi^* \circ \widetilde{T}^* \circ \iota^*.$$

By Hahn-Banach extension theorem (Corollary 2.28), $\mathcal{R}(\iota^*) = \mathcal{R}(T)^*$. Since $\widetilde{T}$ is invertible (that is, $\widetilde{T}^{-1} \in \mathcal{L}(\mathcal{R}(T), X/\mathcal{N}(T))$), by Exercise 2.41-(a), we have $\mathcal{R}(\widetilde{T}^* \circ \iota^*) = (X/\mathcal{N}(T))^*$. Therefore

$$\mathcal{R}(T^*) = \mathcal{R}(\pi^* \circ \widetilde{T}^* \circ \iota^*) = \mathcal{R}(\pi^*).$$

We claim that $\mathcal{R}(\pi^*) = \mathcal{N}(T)^\perp$. Note that if $f \in \mathcal{R}(\pi^*)$, then $f = g \circ \pi$ for some $g \in (X/\mathcal{N}(T))^*$ and hence $\mathcal{N}(f) \supset \mathcal{N}(\pi) = \mathcal{N}(T)$. Hence $f \in \mathcal{N}(T)^\perp$.

Conversely, if $f \in \mathcal{N}(T)^\perp$, then $g(x + \mathcal{N}(T)) = f(x)$ defines a well-defined, bounded linear functional such that $g \in (X/\mathcal{N}(T))^*$ that satisfies $\|g\| = \|f\|$ (by Exercise 4.32-(4)). So $f = g \circ \pi = \pi^*(g) \in \mathcal{R}(\pi^*)$. This proves

$$\mathcal{R}(T^*) = \mathcal{R}(\pi^*) = \mathcal{N}(T)^\perp.$$

Hence $\mathcal{R}(T^*)$ is closed by Exercise 4.30-(a).
$\Longleftarrow$: Conversely, let $\mathcal{R}(T^*)$ be closed. Let $Z = \overline{\mathcal{R}(T)}$ and let $S \in \mathcal{L}(X, Z)$ denote the map $S(x) = T(x)$ for all $x \in X$. By Hahn-Banach extension theorem (Corollary 2.28), $\mathcal{R}(S^*) = \mathcal{R}(T^*)$. Clearly, $S^* \in \mathcal{L}(Z^*, Y^*)$ is injective, since if $f \in Z^*$ satisfies $S^*(f) = 0$, then $f(T(x)) = f(S(x)) = 0$ for all $x \in X$ and hence $\mathcal{N}(f) \supset \mathcal{R}(T)$ which implies $\mathcal{N}(f) \supset \overline{\mathcal{R}(T)} = Z$ or equivalently $f = 0$ (see also Exercise 2.42). Since S^* is injective with closed range (as the range is $\mathcal{R}(T^*)$), S is onto by Lemma 4.31. Hence $\mathcal{R}(T) = \mathcal{R}(S) = \overline{\mathcal{R}(T)}$. □

5 Appendix: Integration and measure

This appendix is meant to be a reminder of some basic results concerning integration which were covered in the prerequisite (MATH 420). Let $(X, \mathcal{M})$ be a measurable space. If $E \in \mathcal{M}$, then the *indicator function of* E is the function $\chi_E : X \to \mathbb{R}$ is given by

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

A simple function $f : X \to \mathbb{C}$ is a measurable function with finite range; or equivalently, there exists $n \in \mathbb{N}$, $z_1, \ldots, z_n \in \mathbb{C}, E_1, \ldots, E_n \in \mathcal{M}$ such that $f(x) = \sum_{i=1}^n z_i \chi_{E_i}(x)$ for all $x \in X$. The following result concerning approximation of measurable functions by simple functions.

Theorem 5.1 (Approximation by simple functions)**.** *Let $(X, \mathcal{M})$ be a measurable space and let $f : X \to \mathbb{C}$ be a measurable function. Then there exists a sequence of simple functions $(\phi_n)_{n \in \mathbb{N}}$ such that $0 \leqslant |\phi_1| \leqslant |\phi_2| \leqslant \ldots \leqslant |f|$ and $\lim_{n\to\infty} \phi_n(x) = f(x)$ for all $x \in X$. Furthermore, the convergence is uniform on any set on which f is bounded.*

There are several results that concerning limits of functions and integrals as we recall below.

Theorem 5.2 (Monotone convergence theorem)**.** *Let* $(X, \mathcal{M}, \mu)$ *be a measure space. If* $(f_n)_{n\in\mathbb{N}}$ *is a sequence of non-negative, measurable functions such that* $0 \leqslant f_n(x) \leqslant f_{n+1}(x)$ *for all* $x \in X$ *and* $n \in \mathbb{N}$*. Then* $f = \lim_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} f_n$ *is measurable and satisfies*

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

Theorem 5.3 (Fatou's lemma)**.** *Let* $(X, \mathcal{M}, \mu)$ *be a measure space. If* $(f_n)_{n\in\mathbb{N}}$ *is a sequence of non-negative, measurable functions, then* $\liminf_{n\to\infty} f_n$ *is a measurable function and satisfies*

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leqslant \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

We recall the dominated convergence theorem. We say that a measurable function $f : X \to \mathbb{C}$ on a measure space $(X, \mathcal{M}, \mu)$ is *integrable*, if $\int_X |f| \, d\mu < \infty$.

Theorem 5.4. *Let* $(X, \mathcal{M}, \mu)$ *be a measure space. If* $(f_n)_{n\in\mathbb{N}}$ *be a sequence of integrable (complex-valued) measurable functions and such that*

(a) there exists $f : X \to \mathbb{C}$ *measurable such that* $f_n \to f$ μ*-almost everywhere;*

(b) there exists a non-negative, integrable and measurable function g *such that* $|f_n| \leqslant g$ μ*-almost everywhere for every* $n \in \mathbb{N}$*.*

Then f *is integrable and*

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

Definition 5.5 (Complex measure)**.** A complex measure on a measurable space $(X, \mathcal{M})$ is a map $\nu : \mathcal{M} \to \mathbb{C}$ such that

(a) $\nu(\varnothing) = 0$,

(b) if $(E_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{M}$, then

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n),$$

where the series above converges absolutely.

We recall the Lebesgue-Radon-Nikodym theorem for complex measures. This is a simple consequence of the corresponding result for signed measures as real and imaginary parts of a complex measure are signed measures.

Theorem 5.6 (Lebesgue-Radon-Nikodym theorem)**.** *Let* ν *be a complex measure and* μ *be a positive,* σ*-finite measure on a measurable space* $(X, \mathcal{M})$*. Then there exist a complex measure* λ *and an integrable function* $f : X \to \mathbb{C}$ *such that* $\lambda \perp \mu$ *and* $d\nu = d\lambda + f \, d\mu$*. Furthermore, the decomposition above is unique; that is, if there exist a complex measure* λ' *and an integrable function* $f' : X \to \mathbb{C}$ *such that* $\lambda' \perp \mu$ *and* $d\nu = d\lambda' + f' \, d\mu$*, then* $\lambda = \lambda'$ *and* $f = f'$ μ*-almost everywhere.*

References

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

[Fol] G. B. Folland, *Real analysis* Pure Appl. Math. (N. Y.) Wiley-Intersci. Publ. John Wiley & Sons, Inc., New York, 1999, xvi+386 pp.

[Lax] P. D. Lax, *Functional analysis*. Pure Appl. Math. (N. Y.) Wiley-Interscience [John Wiley & Sons], New York, 2002. xx+580 pp.

[RS] M. Reed, B. Simon, *Methods of modern mathematical physics. I. Functional analysis* Academic Press, New York-London, 1972, xvii+325 pp.