Notes on Elementary Linear Analysis

Erik Bédos

Institute of Mathematics, University of Oslo

Version 1.1

May 23, 2019

Acknowledgements

The first draft of these notes were written during the spring of 2018 and tested simultaneously as a part of the curriculum for the course MAT3400/4400 at the University of Oslo. Hearthy thanks to Adam Sørensen, who was in charge of the lectures, for his patience and for his comments.

I have tested myself the notes during the spring of 2019 and used this opportunity to perform some minor revisions. I am grateful to Ulrik Enstad and to some of the students following the course for pointing out some misprints in the text. If you find others, or if you have some suggestions, I'll be happy if you send me a message.

E.B.

Contents

CHAPTER 1

Preliminaries

In this chapter we fix some notation and give a review of some of the concepts and results that we will need. These are usually covered in undergraduate courses in real analysis, and the reader may consult the book of T. Lindstrøm, *Spaces: an introduction to real analysis* (AMS 2017), or any other standard book in real analysis, for details and proofs.

1.1 Normed spaces

Throughout these notes $\mathbb{F}$ will denote either $\mathbb{R}$ (the real numbers) or $\mathbb{C}$ (the complex numbers). If X, Y are sets, we let $X \times Y$ denote their Cartesian product, i.e.,

$$X \times Y = \Big\{(x, y) : x \in X, y \in Y\Big\}.$$

A metric space (X, d) is called *complete* when every Cauchy sequence in (X, d) is convergent.

Definition 1.1.1. A *normed space* $(X, \|\cdot\|)$ *over* $\mathbb{F}$ is a vector space X over $\mathbb{F}$ which is equipped with a norm $\|\cdot\|$. We recall that X is then a metric space with respect to the metric given by $d(x, y) = \|x - y\|$ for $x, y \in X$. We will only consider normed spaces over $\mathbb{F}$ in these notes, and we will often just write X to denote such a normed space, assuming tacitly that some norm on X is given.

When $x \in X$ and $r > 0$, we let $B_r^X(x)$ denote the closed ball in X with center in x and radius r, that is,

$$B_r^X(x) := \Big\{y \in X : \|x - y\| \leq r\Big\}.$$

When there is no danger of confusion, we just write $B_r(x)$ instead of $B_r^X(x)$. We also set

$$X_1 := B_1^X(0), \text{ i.e., } X_1 = \Big\{x \in X : \|x\| \leq 1\Big\}.$$

Definition 1.1.2. If $(X, \|\cdot\|)$ is a normed space, and $\|\cdot\|'$ is also a norm on X, we say that $\|\cdot\|$ and $\|\cdot\|'$ are *equivalent* when there exist positive real numbers K and L such that

$$\|x\|' \leq K\,\|x\| \quad \text{and} \quad \|x\| \leq L\,\|x\|' \quad \text{for all } x \in X.$$

When $\|\cdot\|$ and $\|\cdot\|'$ are equivalent, it is clear that a sequence $\{x_n\}_{n=1}^{\infty}$ in X converges to $x \in X$ w.r.t. $\|\cdot\|$ if and only if it converges to $x \in X$ w.r.t. $\|\cdot\|'$. The following proposition implies that for many purposes the choice of a norm in a finite-dimensional space can be made arbitrarily.

Proposition 1.1.3. *If X is a finite-dimensional vector space over $\mathbb{F}$, then all norms on X are equivalent.*

Definition 1.1.4. Assume $\{x_n\}_{n=1}^{\infty}$ is a sequence in a normed space $(X, \|\cdot\|)$. We say that the series $\sum_{n=1}^{\infty} x_n$ is *convergent in X* if there is some $x \in X$ such that $\|x - \sum_{n=1}^{N} x_n\| \to 0$ as $N \to \infty$, in which case we say that $\sum_{n=1}^{\infty} x_n$ converges to x (w.r.t. $\|\cdot\|$), and also write $x = \sum_{n=1}^{\infty} x_n$.

Definition 1.1.5. When a normed space $(X, \|\cdot\|)$ is complete with respect to the associated metric given by

$$d(x, y) = \|x - y\|$$

for all $x, y \in X$, we say that X is a *Banach space* (over $\mathbb{F}$).

To check that a normed space is a Banach space, the following result is often useful:

Theorem 1.1.6. *Let $(X, \|\cdot\|)$ be a normed space. Then X is a Banach space if and only if every absolutely convergent series in X is convergent in X, that is, if and only if the following condition holds :*

Whenever $\sum_{n=1}^{\infty} x_n$ is a series in X such that $\sum_{n=1}^{\infty} \|x_n\| < \infty$, then $\sum_{n=1}^{\infty} x_n$ is convergent in X.

Remark 1.1.7. It is good to know that if X is a normed space, then we can always form its *completion*; this means that whenever needed, we can assume that X sits as a dense subspace of a Banach space $\widetilde{X}$ where the norm of $\widetilde{X}$ extends the norm on X. An elegant way to construct $\widetilde{X}$ (as an application of the so-called Hahn-Banach theorem) is covered in more advanced courses on linear analysis.

1.2 Inner product spaces

Definition 1.2.1. An *inner product space over* $\mathbb{F}$ is a vector space X over $\mathbb{F}$ which is equipped with an inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$. This means that for $x, y, z \in X$ and $\lambda \in \mathbb{F}$ we have:

i) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,

ii) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,

iii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$,

iv) $\langle x, x \rangle \geq 0$,

v) $\langle x, x \rangle = 0$ if and only if $x = 0$.

Remark 1.2.2. a) Properties i) and ii) say that the inner product is linear in the first variable.

b) When $\mathbb{F} = \mathbb{R}$, property iii) says that the inner product is symmetric, i.e., $\langle y, x \rangle = \langle x, y \rangle$; combining i) and ii) with iii), we then get that the inner product is also linear in the second variable.

c) When $\mathbb{F} = \mathbb{C}$, we get that the inner product is *conjugate-linear* in the second variable; this means that we have

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad \text{and} \quad \langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle.$$

Some authors prefer to use inner products that are linear in the second variable and conjugate-linear in the first variable. This is common in textbooks related to physics or mathematical physics. As one can go from one type to the other by setting $\langle x, y \rangle' := \langle y, x \rangle$, it is mainly a matter of taste which convention one chooses to use.

In the sequel, by an inner product space, we will always mean an inner product space over $\mathbb{F}$. An inequality of fundamental importance is:

Theorem 1.2.3 (*The Cauchy-Schwarz inequality*)**.** *Let X be an inner product space. For $x \in X$ set $\|x\| := \langle x, x \rangle^{1/2}$. Then we have*

$$|\langle x, y \rangle| \leq \|x\| \, \|y\| \tag{1.2.1}$$

for all $x, y \in X$, with equality if and only if x and y are linearly dependent.

If X is an inner product space, then using the Cauchy-Schwarz inequality, one deduces that $\|x\| = \langle x, x \rangle^{1/2}$ gives a norm on X. Thus, X is then a normed space, and its norm is easily seen to satisfy the *parallellogram law*, that is, for all $x, y \in X$ we have

$$\|x + y\|^2 + \|x - y\|^2 = 2 \, \|x\|^2 + 2 \, \|y\|^2. \tag{1.2.2}$$

Definition 1.2.4. Let X be an inner product space. If $x, y \in X$, then x and y are said to be *orthogonal* (to each other) when $\langle x, y\rangle = 0$. A subset $S \subset X$ is called *orthogonal* if x and y are orthogonal for all $x, y \in S$ such that $x \neq y$. Moreover, S is called *orthonormal* if S is orthogonal and $\|x\| = 1$ for all $x \in S$.

Proposition 1.2.5 (*Pythagoras*)**.** *Assume* $\{x_1, \ldots, x_n\}$ *is a finite* orthogonal *subset of an inner product space* X*. Then we have*

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

Proposition 1.2.6. *Assume* $S = \{u_1, \ldots, u_n\}$ *is a finite* orthonormal *subset of an inner product space* X*. Then* S *is linearly independent. Moreover, if* $u \in Span\{u_1, \ldots, u_n\}$*, i.e., if* u *is a linear combination of the vectors in* S*, then*

$$u = \sum_{j=1}^{n} \langle u, u_j\rangle\, u_j \quad \textit{and} \quad \|u\|^2 = \sum_{j=1}^{n} |\langle u, u_j\rangle|^2.$$

Proposition 1.2.7 (*Bessel's inequality*)**.** *Assume* $S = \{u_j : j \in J\}$ *is a countable* orthonormal *subset of an inner product space* X*. Then for any* $x \in X$ *we have*

$$\sum_{j\in J} |\langle x, u_j\rangle|^2 \,\leq\, \|x\|^2.$$

Definition 1.2.8. An inner product space X (over $\mathbb{F}$) is called an *Hilbert space* (*over* $\mathbb{F}$) when X is complete with respect to the norm associated with its inner product.

Remark 1.2.9. Assume X is an inner product space. Considering X as a normed space, we may form its completion $\widetilde{X}$ (cf. Remark 1.1.7), and extend the inner product on X to an inner product on $\widetilde{X}$ as follows: if $y, y' \in \widetilde{X}$, then we can pick sequences $\{x_n\}_{n=1}^{\infty}, \{x'_n\}_{n=1}^{\infty}$ in X converging respectively to y and y'; after checking that $\{\langle x_n, x'_n\rangle\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{F}$, hence is convergent, we may set

$$\langle y, y'\rangle := \lim_{n\to\infty} \langle x_n, x'_n\rangle.$$

It is then a somewhat tedious exercise to verify that this gives a well-defined inner product on $\widetilde{X}$ which extends the one on X. This means that whenever needed, we may assume that X sits as a dense subspace of a Hilbert space $\widetilde{X}$ (called the *completion* of X) where the inner product on $\widetilde{X}$ extends the inner product on X.

1.3 Linear operators

Definition 1.3.1. Assume that X and Y are both vectors spaces over $\mathbb{F}$. Then a map $T : X \to Y$ is called a *linear operator* if we have

$$T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T(x_1) + \lambda_2 T(x_2)$$

for all $\lambda_1, \lambda_2 \in \mathbb{F}$ and all $x_1, x_2 \in X$.

We denote by $\mathcal{L}(X, Y)$ the set of all linear operators from X to Y. One readily checks that $\mathcal{L}(X, Y)$ is a vector space over $\mathbb{F}$ with respect to the operations defined by

$$(S + T)(x) = S(x) + T(x), \quad (\lambda T)(x) = \lambda T(x)$$

for $S, T \in \mathcal{L}(X, Y)$, $\lambda \in \mathbb{F}$ and $x \in X$. We also set $\mathcal{L}(X) := \mathcal{L}(X, X)$. We let $I_X \in \mathcal{L}(X)$ denote the *identity map from X into itself*, that is, $I_X(x) = x$ for all $x \in X$. We just I instead of I_X if no confusion is possible.

Definition 1.3.2. Assume that X and Y are both normed spaces over $\mathbb{F}$. Then a linear operator $T : X \to Y$ is called *bounded* if there exists some real number $M > 0$ such that

$$\|T(x)\| \leq M \|x\| \quad \forall x \in X,$$

or, equivalently, such that $\|T(x)\| \leq M$ for all $x \in X_1$.

Proposition 1.3.3. *Assume that X and Y are both normed spaces over $\mathbb{F}$ and let $T \in \mathcal{L}(X, Y)$. Then the following conditions are equivalent:*

(*a*) *T is bounded.*

(*b*) *T is uniformly continuous on X.*

(*c*) *T is continuous on X.*

(*d*) *T is continuous at $x = 0$.*

We denote the set of all bounded linear operators from X to Y by $\mathcal{B}(X, Y)$. We follow tradition here and use the qualifying adjective "bounded", although we could equally well have used "continuous" instead. One readily checks that $\mathcal{B}(X, Y)$ is a subspace of $\mathcal{L}(X, Y)$. We also set $\mathcal{B}(X) = \mathcal{B}(X, X)$.

Proposition 1.3.4. *Assume that X and Y are both normed spaces over $\mathbb{F}$. For $T \in \mathcal{B}(X,Y)$, set*

$$\|T\| := \sup\Big\{\|T(x)\| : x \in X_1\Big\} < \infty.$$

Then the map $T \to \|T\|$ is a norm on $\mathcal{B}(X,Y)$, called the operator norm*. Moreover, we have*

$$\|T\| = \sup\Big\{\|T(x)\| : x \in X, \|x\| = 1\Big\} \quad (\textit{when } X \neq \{0\}),$$

and

$$\|T(x)\| \leq \|T\|\,\|x\| \quad \forall\, x \in X.$$

Theorem 1.3.5. *Assume that X is a normed space over $\mathbb{F}$, while Y is a Banach space. Then $\mathcal{B}(X,Y)$ is Banach space. In particular, $\mathcal{B}(X)$ is a Banach space whenever X is a Banach space.*

An immediate consequence of this theorem is that $\mathcal{B}(X,\mathbb{F})$ is a Banach space whenever X is normed space over $\mathbb{F}$. Elements of $\mathcal{L}(X,\mathbb{F})$ are called *linear functionals*. Thus $\mathcal{B}(X,\mathbb{F})$ consists of the *bounded linear functionals on X*; it is usually called the *dual space of X* and denoted by X^* in many books, or by $X^\sharp$ in others.

Definition 1.3.6. A map $T : X \to Y$ between two vector spaces over $\mathbb{F}$ is called a (*vector space*) *isomorphism* if $T \in \mathcal{L}(X,Y)$ and T is bijective (that is, T is both one-to-one and onto). It is then easy to check that the inverse map of T, $T^{-1} : Y \to X$, is linear, i.e., $T^{-1} \in \mathcal{L}(Y,X)$.

Definition 1.3.7. Assume that X and Y are normed spaces over $\mathbb{F}$. A map $T : X \to Y$ is called an *isomorphism of normed spaces* if T is a (vector space) isomorphism such that both T and T^{-1} are bounded.

Definition 1.3.8. Assume that X is a normed space and $T \in \mathcal{B}(X)$. Then we say that T is *invertible in $\mathcal{B}(X)$* if T is an isomorphism of normed spaces. In other words, an operator $T \in \mathcal{B}(X)$ is invertible in $\mathcal{B}(X)$ if T is bijective and $T^{-1} \in \mathcal{B}(X)$.

Proposition 1.3.9. *Let X, Y, Z be normed spaces over $\mathbb{F}$, and let $T \in \mathcal{B}(X,Y)$, $S \in \mathcal{B}(Y,Z)$. Set $ST := S \circ T : X \to Z$. Then $ST \in \mathcal{B}(X,Z)$ and*

$$\|ST\| \leq \|S\|\,\|T\|.$$

Corollary 1.3.10. *Assume that X is a normed space and $S \in \mathcal{B}(X)$. For each $n \in \mathbb{N}$, let $S^n := S \cdots S$ denote the product of S with itself n times, then $S^n \in \mathcal{B}(X)$ and $\|S^n\| \leq \|S\|^n$. Note that by setting $S^0 = I_X$, this formula also holds when $n = 0$.*

Theorem 1.3.11. *Assume that X is a Banach space and $S \in \mathcal{B}(X)$ is such that $\|S\| < 1$. Then $I - S$ is invertible in $\mathcal{B}(X)$ and*

$$(I - S)^{-1} = \sum_{n=0}^{\infty} S^n \quad (\textit{convergence w.r.t. operator norm}).$$

Moreover, $\|(I - S)^{-1}\| \leq \frac{1}{1 - \|S\|}$.

CHAPTER 2

On L^p-spaces

An important class of Banach spaces over $\mathbb{F}$ associated with measure spaces are the so-called L^p-spaces, where $1 \leq p \leq \infty$. We will assume that $\mathbb{F} = \mathbb{C}$, and just mention that the case where $\mathbb{F} = \mathbb{R}$ may be handled in a similar way. Our presentation is somewhat more detailed than the one given in section 7.7 and 7.9 of Lindstrøm's book.

2.1 The case $1 \leq p < \infty$

Let $(X, \mathcal{A}, \mu)$ be a measure space and set

$$\mathcal{M} = \mathcal{M}(X, \mathcal{A}) := \Big\{ f : X \to \mathbb{C} : f \text{ is } \mathcal{A}\text{-measurable} \Big\},$$

which we know is a vector space (with its natural operations). We will be interested in subspaces of $\mathcal{M}$ associated with any $p \in [1, \infty]$. We first consider the case $1 \leq p < \infty$. For each $f \in \mathcal{M}$ we note that the function $|f|^p$ is non-negative and belongs to $\mathcal{M}$ (since the function $z \to |z|^p$ is continuous on $\mathbb{C}$), so we can set

$$\|f\|_p := \Big(\int_X |f|^p \, d\mu \Big)^{1/p} \in [0, \infty]$$

(using the convention that $\infty^{1/p} = \infty$). Moreover, we set

$$\mathcal{L}^p(X, \mathcal{A}, \mu) := \{ f \in \mathcal{M} : \|f\|_p < \infty \}.$$

We will just write $\mathcal{L}^p$ when there is no danger of confusion, and note that some authors write $\mathcal{L}^p(\mu)$. It is then clear that $\mathcal{L}^1$ consists of all the complex functions on X which are integrable (w.r.t. μ). When $\mathcal{A} = \mathcal{P}(X)$ and μ is the counting measure on $\mathcal{A}$, it is common to write $\ell^p(X)$ instead of $\mathcal{L}^p(X, \mathcal{A}, \mu)$.

It is not difficult to see that $\mathcal{L}^p$ is a subspace of $\mathcal{M}$. For example, closedness under addition follows readily from the inequality $|z+w|^p \leq 2^p(|z|^p + |w|^p)$, which is easily seen to hold for all $z, w \in \mathbb{C}$. On the other hand, it is not true in general that $\|\cdot\|_p$ is a norm on $\mathcal{L}^p$. The reason is that for $f \in \mathcal{L}^p$, we have

$$\|f\|_p = 0 \Leftrightarrow \int_X |f|^p \, d\mu = 0 \Leftrightarrow |f|^p = 0 \;\; \mu\text{-a.e.} \Leftrightarrow f = 0 \;\; \mu\text{-a.e.}$$

As we will soon see, $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p$ in the following sense:

Definition 2.1.1. A *seminorm* on a vector space V (over $\mathbb{F}$) is a function $v \to \|v\|$ from V into $[0, \infty)$ satisfying $\|\lambda v\| = |\lambda| \, \|v\|$ and the *triangle inequality* $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$ and $\lambda \in \mathbb{F}$.

We note that a seminorm $\|\cdot\|$ is a norm if it also satisfies that $\|v\| = 0$ only if $v = 0$. Using the triangle inequality for $|\cdot|$ on $\mathbb{C}$, one readily deduces that $\|\cdot\|_1$ gives a seminorm on $\mathcal{L}^1$. To handle the case $p > 1$ we will need:

Theorem 2.1.2 (*Hölder's inequality*)**.** *Assume $p \in (1, \infty)$ and let $q \in (1, \infty)$ denote p's* conjugate exponent *given by $q = \frac{p}{p-1}$, so that $\frac{1}{p} + \frac{1}{q} = 1$.*

Let $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$. Then $fg \in \mathcal{L}^1$ and

$$\|fg\|_1 = \int_X |fg| \, d\mu \;\leq\; \|f\|_p \, \|g\|_q \, . \tag{2.1.1}$$

Proof. We first note that if a, b are nonnegative real numbers, then we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \, . \tag{2.1.2}$$

A geometric way to prove this inequality (called *Young's inequality*) is to observe that $\frac{a^p}{p}$ is the area given $\int_0^a x^{p-1} \, dx$, while $\frac{b^q}{q}$ is the area given by $\int_0^b y^{q-1} \, dy$. As $q - 1 = 1/(p-1)$, we have $y = x^{p-1} \Leftrightarrow x = y^{q-1}$ when $x, y \geq 0$. By considering the graph of $y = x^{p-1}$ and the rectangle $[0, a] \times [0, b]$ in the xy-plane, one realizes that (2.1.2) must be true.

Next, we note that we may assume that $\|f\|_p = \|g\|_q = 1$. Indeed, assume that (2.1.1) holds whenever $\|f\|_p = \|g\|_q = 1$, and consider $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then both sides of (2.1.1) are equal to zero. On the other hand, if $\|f\|_p$ and $\|g\|_q$ are both nonzero, then we may use that (2.1.1) holds for the functions $f/\|f\|_p$ and $g/\|g\|_q$, and deduce that it holds in the general case.

Hence, assume that $\|f\|_p = \|g\|_q = 1$. Then, using (2.1.2) with $a = |f(x)|$ and $b = |g(x)|$ for each $x \in X$, and linearity of the integral, we get

$$\begin{aligned}
\int_X |fg|\, d\mu &= \int_X |f(x)|\, |g(x)|\, d\mu(x) \\
&\leq \frac{1}{p} \int_X |f(x)|^p\, d\mu(x) + \frac{1}{q} \int_X |g(x)|^q\, d\mu(x) \\
&= \frac{1}{p} \|f\|_p^p + \frac{1}{q} \|g\|_q^q \\
&= \frac{1}{p} + \frac{1}{q} = 1 \\
&= \|f\|_p \, \|g\|_q \,,
\end{aligned}$$

as desired. ■

Corollary 2.1.3. *Let $p \in [1, \infty)$. Then $\| \cdot \|_p$ is a seminorm on $\mathcal{L}^p$. In particular, for all $f, g \in \mathcal{L}^p$, we have*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad \text{(Minkowski's inequality)} \tag{2.1.3}$$

Proof. As already mentioned, the case $p = 1$ is straightforward. So assume $p \in (1, \infty)$. The reader should have no problem to see that we have $\|\lambda f\|_p = |\lambda| \, \|f\|_p$ for all $\lambda \in \mathbb{C}$ and all $f \in \mathcal{L}^p$. Next, let $f, g \in \mathcal{L}^p$, and let q be p's conjugate exponent. As $(p-1)q = p$ and $p/q = p - 1$, we have

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

Since $f + g \in \mathcal{L}^p$, this shows that $|f+g|^{p-1} \in \mathcal{L}^q$; moreover, using Hölder's inequality (at the 4th step), we get

$$\begin{aligned}
\|f+g\|_p^p &= \int_X |f+g|^p\, d\mu = \int_X |f+g|\, |f+g|^{p-1}\, d\mu \\
&\leq \int_X |f|\, |f+g|^{p-1}\, d\mu + \int_X |g|\, |f+g|^{p-1}\, d\mu \\
&\leq \|f\|_p \, \| \, |f+g|^{p-1} \|_q + \|g\|_p \, \| \, |f+g|^{p-1} \|_q \\
&= (\|f\|_p + \|g\|_p) \, \| |f+g|^{p-1} \|_q \\
&= (\|f\|_p + \|g\|_p) \, \|f+g\|_p^{p-1} \,,
\end{aligned}$$

and Minkowski's inequality clearly follows. ■

Let $\{f_n\}$ be a sequence in $\mathcal{L}^p$ and $f \in \mathcal{L}^p$. We note that it may happen that $f_n \to f$ pointwise on X while $\|f_n - f\|_p \not\to 0$ as $n \to \infty$. For example one may let $X = \mathbb{R}$, $\mathcal{A} = \mathcal{B}_\mathbb{R}$, $\mu =$ Lebesgue measure on $\mathcal{B}_\mathbb{R}$, and consider the sequence given by $f_n = \chi_{[n,n+1]}$ for each $n \in \mathbb{N}$: it converges pointwise to 0 on $\mathbb{R}$ as $n \to \infty$, and satisfies $\|f_n\|_p = 1$ for all $n \in \mathbb{N}$.

The following $\mathcal{L}^p$-version of Lebesgue's Dominated Convergence Theorem gives conditions ensuring that a pointwise limit is also convergent w.r.t. $\|\cdot\|_p$.

Proposition 2.1.4. *Let $p \in [1, \infty)$ and $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^p$. Assume that there exist some $g \in \mathcal{L}^p$ such that $|f_n| \le g$ μ-a.e. for all $n \in \mathbb{N}$, and some $f \in \mathcal{M}$ such that $f_n \to f$ pointwise μ-a.e. on X.*

Then $f \in \mathcal{L}^p$ and $\|f_n - f\|_p \to 0$ as $n \to \infty$.

Proof. The assumptions imply that $|f_n|^p \le g^p$ μ-a.e. for all $n \in \mathbb{N}$ and that $|f_n|^p \to |f|^p$ pointwise μ-a.e. on X. It follows that we $|f|^p \le g^p$ μ-a.e., so

$$\int_X |f|^p \, d\mu \le \int_X g^p \, d\mu < \infty \,,$$

hence $f \in \mathcal{L}^p$. Further, we get

$$|f_n - f|^p \le \Big(|f_n| + |f|\Big)^p \le (2\,g)^p = 2^p\, g^p \quad \mu\text{-a.e.},$$

and $|f_n - f|^p \to 0$ pointwise μ-a.e. on X. Since $2^p\, g^p \in \mathcal{L}^1$, we can apply Lebesgue's Dominated Convergence Theorem and get

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

which gives that $\|f_n - f\|_p \to 0$ as $n \to \infty$, as desired. ∎

Let $p \in [1, \infty)$. It follows from Corollary 2.1.3 that we obtain a normed space L^p by identifying functions in $\mathcal{L}^p$ that agree μ-a.e. To achieve this in a formal way, we first define a relation $\sim$ on $\mathcal{L}^p$ by setting

$$f \sim g \;\Leftrightarrow\; f = g \quad \mu\text{-a.e.}$$

for $f, g \in \mathcal{L}^p$. In other words, $f \sim g \;\Leftrightarrow\; \|f - g\|_p = 0$. It is almost immediate that $\sim$ is an equivalence relation, and we will denote the equivalence class of $f \in \mathcal{L}^p$ by $[f]$, that is, we set

$$[f] := \Big\{ g \in \mathcal{L}^p : f \sim g \Big\}.$$

It is then a routine matter to check that

$$L^p = L^p(X, \mathcal{A}, \mu) := \Big\{ [f] : f \in \mathcal{L}^p \Big\}$$

becomes a normed space w.r.t.

$$[f] + [g] := [f + g]\,, \quad \lambda\,[f] := [\lambda\, f]\,, \quad \|\,[f]\,\|_p := \|f\|_p$$

where $f, g \in \mathcal{L}^p$ and $\lambda \in \mathbb{C}$. (The reader may consult Exercise 2.1 for a more general statement.) Moreover, we have:

Theorem 2.1.5. *Let $p \in [1, \infty)$. Then $(L^p, \|\cdot\|_p)$ is a Banach space.*

Proof. Let $\{[f_n]\}_{n\in\mathbb{N}} \subseteq L^p$ be such that $\sum_{n=1}^{\infty} \|[f_n]\|_p < \infty$, i.e., such that $S := \sum_{n=1}^{\infty} \|f_n\|_p < \infty$. According to Theorem 1.1.6 we have to show that the series $\sum_{n=1}^{\infty}[f_n]$ is convergent in L^p. It suffices to show that there exists some $F \in \mathcal{L}^p$ such that $\lim_{N\to\infty} \|\sum_{n=1}^{N} f_n - F\|_p = 0$, because this will give that

$$\lim_{N\to\infty} \|\sum_{n=1}^{N} [f_n] - [F]\|_p = \lim_{N\to\infty} \|\Big[\sum_{n=1}^{N} f_n - F\Big]\|_p = \lim_{N\to\infty} \|\sum_{n=1}^{N} f_n - F\|_p = 0\,,$$

thus showing that $\sum_{n=1}^{\infty}[f_n]$ converges to $[F]$ in L^p.

For each $N \in \mathbb{N}$, set $g_N := \sum_{n=1}^{N} |f_n|$. Also, let $g : X \to [0, \infty]$ be given by

$$g(x) := \sum_{n=1}^{\infty} |f_n(x)| \quad \text{for all } x \in X.$$

Clearly, the sequence $\{g_N^p\}$ of $\mathcal{A}$-measurable nonnegative functions is nondecreasing, and it converges pointwise to the $\mathcal{A}$-measurable function g^p on X. Further, using Minkowski's inequality, we get

$$\|g_N\|_p \;\le\; \sum_{n=1}^{N} \|\,|f_n|\,\|_p = \sum_{n=1}^{N} \|f_n\|_p \;\le\; S$$

for all $N \in \mathbb{N}$. Hence, using the Monotone Convergence Theorem, we get

$$\int_X g^p \, d\mu = \lim_{N\to\infty} \int_X g_N^p \, d\mu = \lim_{N\to\infty} \|g_N\|_p^p \;\le\; S^p \;<\; \infty\,.$$

Since $g^p \ge 0$, it follows from [L; Exercise 7.5.6] that g^p is finite μ-a.e., hence that g is finite μ-a.e. This means that the series $\sum_{n=1}^{\infty} f_n(x)$ is absolutely convergent for every x belonging to some $E \in \mathcal{A}$ such that $\mu(E^c) = 0$. We may therefore define $F \in \mathcal{M}$ by

$$F(x) = \begin{cases} \sum_{n=1}^{\infty} f_n(x) & \text{if } x \in E, \\ 0 & \text{if } x \in E^c. \end{cases}$$

With $F_N := \sum_{n=1}^{N} f_n$ we then have $|F_N| \le g_N \le g \in \mathcal{L}^p$ for every $N \in \mathbb{N}$, and $F_N \to F$ pointwise μ-a.e. on X as $N \to \infty$. Proposition 2.1.4 gives now that $F \in \mathcal{L}^p$ and $\lim_{N\to\infty} \|F_N - F\|_p = 0$, as we wanted to show. ∎

2.2 The case $p = \infty$

We now consider the case $p = \infty$. Let $\mathcal{F}(X)$ denote the vector space consisting of all complex functions on X (with its natural operations). By an *algebra of complex functions on* X, we will mean a subspace of $\mathcal{F}(X)$ which is also closed under pointwise multiplication. For example, $\mathcal{M} = \mathcal{M}(X, \mathcal{A})$ is an algebra of complex functions on X. Another natural algebra is the one consisting of those functions in $\mathcal{M}$ which are *bounded*. We will actually be interested in a slightly larger algebra.

Definition 2.2.1. A function $f \in \mathcal{M}$ is said to be *essentially bounded* (w.r.t. μ) if there exists some real number $M > 0$ such that

$$|f| \leq M \quad \mu\text{-a.e.},$$

in which case we set $\|f\|_\infty := \inf \Big\{ M > 0 : |f| \leq M \ \mu\text{-a.e.} \Big\}$.

Example 2.2.2. a) Asume $g \in \mathcal{M}$ is bounded and set $\|g\|_u := \sup_{x \in X} |g(x)|$. Then g is essentially bounded (w.r.t. μ), and we have

$$\|g\|_\infty \leq \|g\|_u \, .$$

Indeed, we have $\mu\Big(\{x \in X : |g(x)| > \|g\|_u\}\Big) = \mu(\varnothing) = 0$. This gives that $|g| \leq \|g\|_u$ μ-a.e., and both assertions follow readily.

We note that it may happen that $\|g\|_\infty < \|g\|_u$. For example, consider the Borel function g on $X = \mathbb{R}$ given by $g = \chi_{\{0\}}$; letting μ be the Lebsgue measure on $\mathcal{B}_\mathbb{R}$, we get

$$\|g\|_\infty = 0 < 1 = \|g\|_u \, .$$

b) Consider $X = [0, \infty)$, $\mathcal{A} =$ the Borel subsets of X and $\mu =$ the Lebesgue measure on $\mathcal{A}$. Let $f \in \mathcal{M}$ be given by

$$f(x) = e^{ix} + \sum_{n=1}^{\infty} n \, \chi_{\{2n\pi\}}(x), \quad x \geq 0 \, .$$

Then $f(2k\pi) = k+1$ for every $k \in \mathbb{N}$, so f is unbounded. On the other hand, f is essentially bounded (w.r.t. μ), with $\|f\|_\infty = 1$, since $\mu\Big(|f|^{-1}((M, \infty))\Big)$ is equal to 0 if $M \geq 1$ and to ∞ if $0 < M < 1$.

The following useful observation may seem obvious, but it requires a proof.

Lemma 2.2.3. *Let $f \in \mathcal{M}$ be essentially bounded (w.r.t. μ). Then we have*

$$|f| \leq \|f\|_\infty \quad \mu\text{-a.e.} \tag{2.2.1}$$

Proof. Set $B := \{x \in X : |f(x)| > \|f\|_\infty\} \in \mathcal{A}$ and assume (for contradiction) that $\mu(B) > 0$. For each $n \in \mathbb{N}$, set

$$B_n := \left\{x \in X : |f(x)| > \|f\|_\infty + \frac{1}{n}\right\} \in \mathcal{A}\,.$$

Clearly, $B_n \subseteq B_{n+1}$ for every n, and $B = \bigcup_{n=1}^\infty B_n$, so we have

$$\lim_{n\to\infty} \mu(B_n) = \mu(B) > 0\,.$$

Hence there must exist at least one $N \in \mathbb{N}$ such that $\mu(B_N) > 0$. Now, by definition of $\|f\|_\infty$, we can find $M > 0$ such that $\|f\|_\infty \leq M < \|f\|_\infty + \frac{1}{N}$ and $|f| \leq M$ μ-a.e. But this implies that $|f| \leq \|f\|_\infty + \frac{1}{N}$ μ-a.e., i.e., $\mu(B_N) = 0$, and we have reached a contradiction. ■

Using Lemma 2.2.3, it is straightforward to verify that the set $\mathcal{L}^\infty = \mathcal{L}^\infty(X, \mathcal{A}, \mu)$ consisting of all functions in $\mathcal{M}$ that are essentially bounded (w.r.t. μ) is an algebra of complex functions on X (cf. Exercise 2.9). Another application is the following Hölder-type inequality:

Proposition 2.2.4. *Let $q \in [1, \infty)$, $f \in \mathcal{L}^\infty$ and $g \in \mathcal{L}^q$. Then $fg \in \mathcal{L}^q$ and*

$$\|fg\|_q \leq \|f\|_\infty \|g\|_q\,.$$

Proof. Using Lemma 2.2.3 we get that $|fg|^q = |f|^q\,|g|^q \leq \|f\|_\infty^q\,|g|^q$ μ-a.e. It follows that

$$\int_X |fg|^q\,d\mu \leq \|f\|_\infty^q \int_X |g|^q\,d\mu < \infty\,.$$

Hence $fg \in \mathcal{L}^q$. Moreover, taking the q-th root, we obtain the desired inequality. ■

Convergence in $\mathcal{L}^\infty$ with respect to $\|\cdot\|_\infty$ is closely related to uniform convergence:

Proposition 2.2.5. *Let $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}^\infty$ and $f \in \mathcal{L}^\infty$. Then we have that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$ if and only if there exists some $E \in \mathcal{A}$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E.*

Proof. Assume $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. For each $n \in \mathbb{N}$, set

$$F_n := \{x \in X : |f_n(x) - f(x)| > \|f_n - f\|_\infty\} \in \mathcal{A}.$$

Since $f_n - f \in \mathcal{L}^\infty$, we have $\mu(F_n) = 0$ for each n. Hence, $F := \cup_{n\in\mathbb{N}} F_n \in \mathcal{A}$ and $\mu(F) = 0$.

Set now $E := F^c \in \mathcal{A}$. Then $\mu(E^c) = 0$ and

$$E = \{x \in X : |f_n(x) - f(x)| \leq \|f_n - f\|_\infty \text{ for all } n \in \mathbb{N}\}.$$

It is then obvious that $f_n \to f$ uniformly on E. The proof of the reverse implication goes along the same lines, and we leave it as an exercise (cf. Exercise 2.11). ■

As with the $\mathcal{L}^p$-spaces for $1 \leq p < \infty$, an annoying fact is that in general $\|\cdot\|_\infty$ is only a seminorm on $\mathcal{L}^\infty$. To get a norm we have to identify functions that agree μ-a.e. Thus, for each $f \in \mathcal{L}^\infty$ we set $[f] = \{g \in \mathcal{L}^\infty : g = f \ \mu\text{-a.e.}\}$. Then

$$L^\infty = L^\infty(X, \mathcal{A}, \mu) := \Big\{ [f] : f \in \mathcal{L}^\infty \Big\}$$

becomes a vector space w.r.t. the operations given by $[f] + [g] := [f + g]$, $\lambda[f] := [\lambda f]$ (where $f, g \in \mathcal{L}^\infty$ and $\lambda \in \mathbb{C}$), and $\|[f]\|_\infty := \|f\|_\infty$ gives a norm on L^∞ (cf. Exercise 2.1).

Theorem 2.2.6. *$(L^\infty, \|\cdot\|_\infty)$ is a Banach space.*

Proof. We have to show that L^∞ is complete w.r.t. the metric associated with $\|\cdot\|_\infty$.

Let $\{[f_n]\}_{n\in\mathbb{N}}$ be a Cauchy sequence in L^∞. So each f_n belongs to $\mathcal{L}^\infty$ and for any given $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that

$$m, n \geq N \ \Rightarrow \ \|\,[f_m] - [f_n]\,\|_\infty \, < \, \varepsilon\,,$$

that is,

$$m, n \geq N \ \Rightarrow \ \|\, f_m - f_n \,\|_\infty \, < \, \varepsilon\,. \tag{2.2.2}$$

For each $m, n \in \mathbb{N}$, set

$$F_{m,n} := \Big\{x \in X : \ |f_m(x) - f_n(x)| > \|f_m - f_n\|_\infty \Big\}.$$

Then $F_{m,n} \in \mathcal{A}$ and $\mu(F_{m,n}) = 0$ for all $m, n \in \mathbb{N}$ (because $f_m - f_n \in \mathcal{L}^\infty$).

Next, set $F := \bigcup_{m,n\in\mathbb{N}} F_{m,n} \in \mathcal{A}$ and $E := F^c \in \mathcal{A}$.

Note that $\mu(E^c) = \mu(F) = 0$ (since $0 \leq \mu(F) \leq \sum_{m,n\in\mathbb{N}} \mu(F_{m,n}) = 0$). Moreover,

$$E = \bigcap_{m,n\in\mathbb{N}} (F_{m,n})^c = \bigcap_{m,n\in\mathbb{N}} \Big\{x \in X : \; |f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty\Big\}$$
$$= \Big\{x \in X : |f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty \text{ for all } m, n \in \mathbb{N}\Big\} .$$

Let now $\varepsilon > 0$ be given, and choose $N \in \mathbb{N}$ such that (2.2.2) holds.

Then for all $x \in E$ and all $m, n \geq N$, we have

$$|f_m(x) - f_n(x)| \;\leq\; \| \, f_n - f_m \, \|_\infty \;<\; \varepsilon \, . \tag{2.2.3}$$

It follows that $\{f_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{C}$ for each $x \in E$. Since $\mathbb{C}$ is complete, this implies that $\{f_n(x)\}_{n\in\mathbb{N}}$ is convergent for each $x \in E$, hence that $\lim_{n\to\infty} f_n(x) = g(x)$ for some $g(x) \in \mathbb{C}$ for each $x \in E$. Thereby we obtain a function $g : E \to \mathbb{C}$, which is $\mathcal{A}_E$-measurable since g is the pointwise limit of the restriction of the f_n's to E. (Here, $\mathcal{A}_E$ denotes the σ-algebra of all sets in $\mathcal{A}$ which are contained in E).

We can now extend g to an $\mathcal{A}$-measurable function $f : X \to \mathbb{C}$ by setting $f(x) = g(x)$ if $x \in E$, and $f(x) = 0$ otherwise.

Again, let $\varepsilon > 0$ be given and choose N as above. Then, for all $x \in E$ and all $m \in \mathbb{N}$ such that $m \geq N$, we get from (2.2.3) that

$$|f_m(x) - f(x)| = |f_m(x) - g(x)| = \lim_{n\to\infty} \, |f_m(x) - f_n(x)| \;\leq\; \varepsilon \, .$$

This implies that $\{f_m\}_{m\in\mathbb{N}}$ converges uniformly to f on E.

Moreover, set $D := E \cap \{x \in X : |f_N(x)| \;\leq\; \|f_N\|_\infty\} \in \mathcal{A}$. Then we have

$$|f(x)| = |f(x) - f_N(x) + f_N(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \varepsilon + \|f_N\|_\infty$$

for all $x \in D$. As

$$0 \leq \mu(D^c) \leq \mu(F) + \mu\Big(\{x \in X : |f_N(x)| > \|f_N\|_\infty\}\Big) = 0 \, ,$$

we have $\mu(D^c) = 0$, so

$$|f| \;\leq \varepsilon + \|f_N\|_\infty \;\; \mu\text{-a.e.}$$

This shows that $f \in \mathcal{L}^\infty$. Using Proposition 2.2.5, we can now conclude that $\|f_m - f\|_\infty \to 0$ as $m \to \infty$. Thus

$$\| \, [f_m] - [f] \, \|_\infty = \|f_m - f\|_\infty \;\to\; 0 \;\text{ as } m \to \infty \, .$$

This means that $\{[f_m]\}_{j\in\mathbb{N}}$ converges to $[f]$ in L^∞. We have thereby shown that every Cauchy sequence in L^∞ is convergent and the proof is finished.

■

2.3 Exercises

Exercise 2.1. Let V be a vector space (over $\mathbb{F}$) and let $\|\cdot\|$ denote a seminorm on V. Define a relation $\sim$ on V by setting

$$v \sim w \Leftrightarrow \|v - w\| = 0$$

for $v, w \in V$.

a) Check that $\sim$ is an equivalence relation.

Denote the equivalence class of $v \in V$ by $[v]$, that is,

$$[v] := \Big\{ w \in V : v \sim w \Big\},$$

and set $\widetilde{V} := \Big\{ [v] : v \in V \Big\}$. Moreover, for $v, w \in V$, and $\lambda \in \mathbb{F}$, set

$$[v] + [w] := [v + w]\,, \quad \lambda\,[v] := [\lambda\, v]\,, \quad \|\,[v]\,\| := \|v\|\,.$$

b) Show that these operations on $\widetilde{V}$ are well-defined, that is, show that if $v, v', w, w' \in V$ are such that $v' \sim v, w' \sim w$, and $\lambda \in \mathbb{C}$, then $(v' + w') \sim (v + w)$, $\lambda v' \sim \lambda v$ and $\|v'\| = \|v\|$.

c) Verify that $(\widetilde{V}, \|\cdot\|)$ is a normed space. (Check at least three of the axioms.)

In the following exercises, unless otherwise specified, $(X, \mathcal{A}, \mu)$ denotes a measure space and $\mathcal{M}$ denotes the space of $\mathcal{A}$-measurable complex functions on X.

Exercise 2.2. Assume that $X = [1, \infty)$, $\mathcal{A} =$ the Borel subsets of X and μ is the Lebesgue measure on $\mathcal{A}$. Let $f \in \mathcal{M}$ be given by

$$f(x) = \frac{1}{x} \quad \text{for all } x \geq 1,$$

and let $1 \leq p < \infty$. Show that $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ if and only if $p > 1$, and compute $\|f\|_p$ in this case.

Exercise 2.3. Assume that $X = \mathbb{R}$, $\mathcal{A} =$ the Borel subsets of X and μ is the Lebesgue measure on $\mathcal{A}$. Let $f \in \mathcal{M}$ be given by

$$f(x) = e^{-x^2} \quad \text{for all } x \in \mathbb{R}.$$

Show that $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ for all $p \in [1, \infty)$ and compute $\|f\|_p$. (You are allowed to use that $\lim_{N\to\infty} \int_{-N}^{N} e^{-t^2}\, dt = \sqrt{\pi}$ without proof.)

Exercise 2.4. Assume that $X = (0,1]$, $\mathcal{A} =$ the Borel subsets of X and μ is the Lebesgue measure on $\mathcal{A}$. Let $f \in \mathcal{M}$ be given by

$$f(x) = \frac{1}{\sqrt{x}} \quad \text{for all } x \in (0,1],$$

and let $1 \leq p < \infty$.

a) Show that $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ if and only if $p < 2$, and compute $\|f\|_p$ in this case.

b) Let ν be the measure on $\mathcal{A}$ given by

$$\nu(A) = \int_A x \, d\mu(x) \qquad \text{for all } A \in \mathcal{A}\,.$$

Show that $f \in \mathcal{L}^p(X, \mathcal{A}, \nu)$ if and only if $p < 4$, and compute $\|f\|_p$ in this case.

Exercise 2.5. Assume that $X = [1, \infty)$, $\mathcal{A} =$ the Borel subsets of X and μ is the Lebesgue measure on $\mathcal{A}$. For each $n \in \mathbb{N}$, define $f_n \in \mathcal{M}$ by

$$f_n(x) = \frac{n}{n\,x^{1/3} + 1} \qquad \text{for all } x \geq 1.$$

a) Show that $f_n \in \mathcal{L}^p$ for all $n \in \mathbb{N}$ whenever $3 < p < \infty$.

b) Assume that $3 < p < \infty$. Decide whether the sequence $\{[f_n]\}_{n\in\mathbb{N}}$ is convergent in L^p and find its limit if it converges.

Exercise 2.6. Let $p \in [1, \infty)$. Let $\mathcal{E}$ denote the space of simple functions in $\mathcal{M}$ and $\mathcal{E}^0$ denote the subspace of $\mathcal{E}$ spanned by $\{\chi_A : A \in \mathcal{A},\, \mu(A) < \infty\}$.

a) Show that $\mathcal{E}^0 = \mathcal{E} \cap \mathcal{L}^p$.

b) Let $f \in \mathcal{L}^p$. Show that there exists a sequence $\{g_n\}$ in $\mathcal{E}^0$ such that $\|f - g_n\|_p \to 0$ as $n \to \infty$. Deduce that the space

$$[\mathcal{E}^0] := \Big\{[g] : g \in \mathcal{E}^0\Big\}$$

is dense in L^p with respect to $\|\cdot\|_p$.

Exercise 2.7. Let $a, b \in \mathbb{R}$, $a < b$, $\mathcal{A}$ denote the Lebesgue measurable subsets of $X = [a, b]$ and μ denote the Lebesgue measure on $\mathcal{A}$. Finally, let $C([a, b])$ denote the space of all continuous complex functions on $[a, b]$. Let $p \in [1, \infty)$.

a) Let $A \in \mathcal{A}$ and $\delta > 0$. Show that there exists some $k \in C([a, b])$ such that $\|\chi_A - k\|_p < \delta$.

b) Use a) and Exercise 2.6 to show that the space $\Big\{[f] : f \in C([a, b])\Big\}$ is dense in $L^p([a, b], \mathcal{A}, \mu)$ with respect to $\|\cdot\|_p$.

Exercise 2.8. Assume that $X = \mathbb{R}$, $\mathcal{A} =$ the Lebesgue measurable subsets of $\mathbb{R}$ and μ is the Lebesgue measure on $\mathcal{A}$. Say that a function $f : \mathbb{R} \to \mathbb{C}$ has *compact support* if $f = 0$ outside some closed, bounded interval. Let $C_c(\mathbb{R})$ denote the space of all continuous complex functions on $\mathbb{R}$ which have compact support. Let $p \in [1, \infty)$.

Show that the space $\{[f] : f \in C_c(\mathbb{R})\}$ is dense in L^p with respect to $\|\cdot\|_p$.

Exercise 2.9. Check that $\|\cdot\|_\infty$ is a seminorm on $\mathcal{L}^\infty$ (so that $\|\cdot\|_\infty$ gives a norm on L^∞). Check also that $\mathcal{L}^\infty$ is an algebra of functions on X and that we have $\|fg\|_\infty \leq \|f\|_\infty \, \|g\|_\infty$ for all $f, g \in \mathcal{L}^\infty$.

Exercise 2.10. Let $f \in \mathcal{M}$. Show that $f \in \mathcal{L}^\infty$ if and only if there exists a bounded function $g \in \mathcal{M}$ such that $f = g$ μ-a.e., in which case we have

$$\|f\|_\infty = \inf\{ \|g\|_u : g \in \mathcal{M} \text{ is bounded and } g = f \ \mu\text{-a.e.}\}.$$

Exercise 2.11. Finish the proof of Proposition 2.2.5.

Exercise 2.12. Let $1 \leq p \leq r < \infty$ and X be a nonempty set. Show that

$$\ell^p(X) \subseteq \ell^r(X) \subseteq \ell^\infty(X).$$

Exercise 2.13. Let $p \in [1, \infty)$ and assume that $(X, \mathcal{A}, \mu)$ is *finite*, that is, $\mu(X) < \infty$.

a) Show that $\mathcal{L}^\infty \subseteq \mathcal{L}^p$.

b) Consider $1 \leq p \leq r < \infty$ and let $f \in \mathcal{L}^r$. Show that $f \in \mathcal{L}^p$ and

$$\|f\|_p \leq \mu(X)^{\frac{1}{p} - \frac{1}{r}} \, \|f\|_r \, .$$

Hint: Use Hölder's inequality in a suitable way.

Note that this shows that $\mathcal{L}^r \subseteq \mathcal{L}^p$. In particular, we have $\mathcal{L}^\infty \subseteq \mathcal{L}^2 \subseteq \mathcal{L}^1$.

c) Consider the Lebesgue measure on the Borel subsets of $\mathbb{R}$. Give an example of a function which is in $\mathcal{L}^2$, but not in $\mathcal{L}^1$ Give also an example of a function which is in $\mathcal{L}^\infty$, but not in $\mathcal{L}^2$.

Exercise 2.14. Let $\mathcal{E}$ denote the space of simple functions in $\mathcal{M}$ and let $f \in \mathcal{L}^\infty$. Show that there exists a sequence $\{h_n\}$ in $\mathcal{E}$ such that $\|f - h_n\|_\infty \to 0$ as $n \to \infty$. Deduce that the space $[\mathcal{E}] := \{[h] : h \in \mathcal{E}\}$ is dense in L^∞ with respect to $\|\cdot\|_\infty$.

CHAPTER 3

More on normed spaces and linear operators

3.1 Aspects of finite dimensionality

Unless otherwise specified, we always assume that the space $\mathbb{F}^n$, $n \in \mathbb{N}$, is equipped with the Euclidean norm $\|\cdot\|_2$ given by

$$\|x\|_2 = \Big(|x_1|^2 + \cdots + |x_n|^2\Big)^{1/2} \quad \text{for } x = (x_1, \ldots, x_n) \in \mathbb{F}^n,$$

and with the metric induced by this norm. As we recalled in Section 1.1, all norms on a finite-dimensional vector space are equivalent. The usual way to prove this is to consider first $\mathbb{F}^n$ and show that any other norm on $\mathbb{F}^n$ is equivalent to $\|\cdot\|_2$. A crucial fact in the proof is that a subset of $\mathbb{F}^n$ is compact (w.r.t. the metric associated with $\|\cdot\|_2$) if and only if it is closed and bounded. It will be useful for us to know that this property, sometimes called the *Heine-Borel property*, holds in any finite-dimensional normed space. We will need the following lemma.

Lemma 3.1.1. *Let X and Y be finite-dimensional normed spaces. Assume that X and Y are isomorphic as vector spaces and let $T \in \mathcal{L}(X,Y)$ be an isomorphism. Then T is an isomorphism of normed spaces.*

Proof. We have to show that T and T^{-1} are bounded. To avoid confusion, we let $\|\cdot\|$ and $\|\cdot\|'$ denote the respective norms on X and Y. For $x \in X$ set

$$\|x\|_T := \|T(x)\|'\,.$$

Clearly, the map $x \to \|x\|_T$ is a seminorm on X; in fact, it is a norm since

$$\|x\|_T = 0 \;\Leftrightarrow\; \|T(x)\|' = 0 \;\Leftrightarrow\; T(x) = 0 \;\Leftrightarrow\; x = 0,$$

the last equivalence being a consequence of the injectivity of T. Since X is finite-dimensional, $\|\cdot\|_T$ is equivalent to $\|\cdot\|$. In particular, this means that there exists some $C > 0$ such that

$$\|T(x)\|' = \|x\|_T \leq C\,\|x\| \quad \text{for all } x \in X,$$

which shows that T is bounded. Similarly, by considering the norm on Y given by $\|y\|_{T^{-1}} := \|T^{-1}(y)\|$ for $y \in Y$, one deduces that T^{-1} is also bounded. ■

Proposition 3.1.2. *Let X be a finite-dimensional normed space. Then a subset K of X is compact (w.r.t. the metric induced by the given norm) if and only if K is closed and bounded.*

Proof. Since a compact subset of a metric space is always closed and bounded, we only have to show the reverse implication. So let $K \subseteq X$ be closed and bounded. We must show that K is compact. If $X = \{0\}$, this is obviously true, so we may assume that $m := \dim(X) \geq 1$. Let then $T : X \to \mathbb{F}^m$ denote the coordinate map w.r.t. some basis for X. Lemma 3.1.1 gives that T is an isomorphism of normed spaces. Set $K' := T(K) \subseteq \mathbb{F}^m$. Then K' is bounded (since T is bounded). Moreover, K' is closed. Indeed, as $K' = (T^{-1})^{-1}(K)$, this follows from the continuity of T^{-1}. By the Heine-Borel property of $\mathbb{F}^m$, we can conclude that K' is compact. As $K = T^{-1}(K')$ and T^{-1} is continuous, this implies that K is compact, as desired. ■

Since the unit ball X_1 of a normed space is closed and bounded we get:

Corollary 3.1.3. *The unit ball X_1 of a finite-dimensional normed space X is compact.*

We note that if X is an *infinite-dimensional* normed space, then X_1 is *not* compact. (See Exercises 3.1 and 3.2.) In particular, this implies that an infinite-dimensional normed space never has the Heine-Borel property.

Another property which is automatically satisfied for a finite-dimensional normed space is completeness:

Proposition 3.1.4. *Let X be a finite-dimensional normed space. Then X is a Banach space.*

Proof. We may clearly assume that $X \neq \{0\}$. To show that X is complete, we let $\{x_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in X and have to prove that it is convergent. As in the proof of Proposition 3.1.2, we can pick an isomorphism of normed spaces $T : X \to \mathbb{F}^m$, where $m = \dim(X)$. For each $n \in \mathbb{N}$, set $y_n := T(x_n)$. Since $\|y_n - y_k\|_2 = \|T(x_n - x_k)\|_2 \leq \|T\|\,\|x_n - x_k\|$ for all

$k, n \in \mathbb{N}$, we see that $\{y_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{F}^m$. Since $\mathbb{F}^m$ is complete, there exists $y \in \mathbb{F}^m$ such that $\|y_n - y\|_2 \to 0$ as $n \to \infty$. Set $x := T^{-1}(y) \in X$. Then we get

$$\|x_n - x\| = \|T^{-1}(y_n - y)\| \ \leq\ \|T^{-1}\|\,\|y_n - y\|_2 \ \to 0 \quad \text{as } n \to \infty.$$

Thus, $\{x_n\}_{n\in\mathbb{N}}$ is convergent, as desired. ■

Corollary 3.1.5. *Assume M is a finite-dimensional subspace of a normed space X. Then M is closed in X.*

Proof. Assume $\{x_n\}_{n\in\mathbb{N}} \subseteq M$ converges to $x \in X$. We have to show that $x \in M$. As M is complete by Proposition 3.1.4, and $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in M, it follows that $\{x_n\}_{n\in\mathbb{N}}$ converges to some $y \in M$. Thus we get that $x = \lim_{n\to\infty} x_n = y \in M$. ■

Finite dimensionality has also some impact on linear operators.

Example 3.1.6. Let $m, n \in \mathbb{N}$ and let $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Then T is bounded. Indeed, let $A = [a_{i,j}]$ denote the standard matrix of T. Then we have $T(x) = (F_1(x), \ldots, F_n(x))$, where $F_i(x) := \sum_{j=1}^n a_{i,j}\, x_j$ for each $i = 1, \ldots, m$ and $x = (x_1, \ldots, x_n) \in \mathbb{F}^n$. Since each component F_i is clearly a continuous function from $\mathbb{F}^n$ to $\mathbb{F}$, we get that T is continuous, and therefore bounded.

More generally, we have:

Proposition 3.1.7. *Let X and Y be normed spaces and let $T \in \mathcal{L}(X, Y)$. Assume that X is finite-dimensional. Then T is bounded.*

Proof. By replacing Y with $T(X)$ if necessary, we may assume that Y is finite-dimensional. Moreover, we may also assume that both X and Y are different from $\{0\}$. Set $n = \dim(X), m = \dim(Y)$, and let $C : X \to \mathbb{F}^n$, $D : Y \to \mathbb{F}^m$ be isomomorphims, which are then necessarily isomorphisms of normed spaces by Lemma 3.1.1. The composition $T' := D \circ T \circ C^{-1}$ is then a linear map from $\mathbb{F}^n$ to $\mathbb{F}^m$, hence it is bounded by the previous example. It follows that $T = D^{-1} \circ T' \circ C$, being the composition of bounded maps, is bounded. ■

Note that the above result is not true in general if we instead assume that Y is finite-dimensional, even in the case where $Y = \mathbb{F}$: a linear functional $T : X \to \mathbb{F}$ may be unbounded when X is an infinite-dimensional normed space. For an example, see Exercise 3.3.

Definition 3.1.8. A linear operator $T : X \to Y$ between two vector spaces X and Y is said to have *finite-rank* if the range of T is finite-dimensional, i.e., if $\dim(T(X)) < \infty$.

It is obvious that a linear functional on a normed space has always finite-rank. As such a linear functional can be unbounded, we get that a finite-rank linear operator T between normed spaces is not necessarily bounded; in fact, it can be shown that such an operator T is bounded if and only if $\ker(T)$ is closed. Bounded finite-rank operators have the following interesting property:

Proposition 3.1.9. *Let X and Y be normed spaces over $\mathbb{F}$, and assume that $T \in \mathcal{B}(X, Y)$ has finite-rank. Then, for any given bounded sequence $\{x_n\}_{n\in\mathbb{N}}$ in X, we have that the sequence $\{T(x_n)\}_{n\in\mathbb{N}}$ has a convergent subsequence in Y.*

Proof. Assume $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ satisfies $\|x_n\| \leq M$ for all $n \in \mathbb{N}$ for some $M > 0$. Then we have

$$\|T(x_n)\| \;\leq\; \|T\|\,\|x_n\| \;\leq\; \|T\|\,M$$

for all $n \in \mathbb{N}$. Now, the ball $B := \{y \in Y : \|y\| \leq \|T\|\,M\}$ is closed in Y. Considering $T(X)$ as a normed space w.r.t. to the norm it inherits from Y, we get that the set $K := T(X) \cap B$ is a closed and bounded subset of $T(X)$. Since $T(X)$ is finite-dimensional (by assumption), it follows from Proposition 3.1.2 that K is compact in $T(X)$. As $\{T(x_n)\}_{n\in\mathbb{N}}$ is a sequence in K, we can therefore conclude that it has a convergent subsequence. ■

An operator $T \in \mathcal{L}(X, Y)$ satisfying the property described in the conclusion of Proposition 3.1.9 is said to be *compact*. We will have a closer look at this important class of operators in Chapter 5.

3.2 Direct sums and projections

We first discuss the concepts of direct sums and projections in a purely linear algebraic setting. Let X be a vector space over $\mathbb{F}$, and let M_1 and M_2 be subspaces of X. We define the *sum of* M_1 *and* M_2 as the subset of X given by

$$M_1 + M_2 := \Big\{x_1 + x_2 : x_1 \in M_1,\ x_2 \in M_2\Big\}.$$

It is straightforward to verify that it the least subspace of X containing both M_1 and M_2.

Definition 3.2.1. We will say that X is the (*internal*) *algebraic direct sum of* M_1 *and* M_2, and write $X = M_1 \dot{+} M_2$, when

$$X = M_1 + M_2 \quad \text{and} \quad M_1 \cap M_2 = \{0\}\,.$$

Obviously, we have $X = M_1 \dot{+} M_2$ if and only if $X = M_2 \dot{+} M_1$.

We first make a simple, but fundamental, observation:

Lemma 3.2.2. *The following two conditions are equivalent*:

(*i*) $X = M_1 \dot{+} M_2$ *;*

(*ii*) *every* $x \in X$ *can be written* in a unique way *as* $x = x_1 + x_2$ *with* $x_1 \in M_1$ *and* $x_2 \in M_2$ *.*

Proof. Assume (*i*) holds and let $x \in X$. Then we have $x = x_1 + x_2$ for some $x_1 \in M_1, x_2 \in M_2$. If we also have $x = x_1' + x_2'$ for some $x_1' \in M_1, x_2' \in M_2$, then we get

$$x_1 - x_1' = x_2' - x_2 \ \in \ M_1 \cap M_2\,.$$

Since $M_1 \cap M_2 = \{0\}$, this implies that $x_1' = x_1$ and $x_2' = x_2$. Thus (*ii*) holds.

Conversely, assume (*ii*) holds. It then obvious that $X = M_1 + M_2$. Consider $y \in M_1 \cap M_2$. Then we have $y = y + 0$ with $y \in M_1$, $0 \in M_2$, and $y = 0 + y$ with $0 \in M_1$, $y \in M_2$. By uniqueness, we get $y = 0$. Thus, $M_1 \cap M_2 = \{0\}$, so (*i*) holds. ■

Remark 3.2.3. If V_1 and V_2 are vector spaces over $\mathbb{F}$, then one may form their direct product $V_1 \times V_2$, which is often called the (external) algebraic direct sum of V_1 and V_2. (This concept is presumably well-known; the definition is recalled in Exercise 3.5). In the case of an (internal) algebraic direct sum $X = M_1 \dot{+} M_2$, it can easily be verified that X is isomorphic to $M_1 \times M_2$.

Example 3.2.4. a) Let X be the space of all $n \times n$ matrices over $\mathbb{F}$, and let M_1 (resp. M_2) denote the subspace of X consisting of all upper (resp. lower) triangular matrices in X. Then it is obvious that we have $X = M_1 + M_2$; but X is *not* the algebraic direct sum of M_1 and M_2, since $M_1 \cap M_2$ consists of all the diagonal matrices.

b) Let X be the space of all $n \times n$ matrices over $\mathbb{R}$, and let M_1 (resp. M_2) denote the subspace of symmetric (resp. skew-symmetric) matrices in X. (We recall that $A \in X$ is called skew-symmetric when $A^t = -A$.) Then we have $X = M_1 \dot{+} M_2$. Indeed, if $A \in X$, then $A = A_1 + A_2$, where

$$A_1 := \frac{1}{2}(A + A^t) \in M_1 \quad \text{and} \quad A_2 := \frac{1}{2}(A - A^t) \in M_2 \,.$$

Moreover, if $A \in M_1 \cap M_2$, then we have $A = A^t = -A$, so $A = 0$.

There is a tight connection between projection operators and directs sums.

Definition 3.2.5. Let X be a vector space. An operator $P \in \mathcal{L}(X)$ is called a *projection* when P is an idempotent map, that is, when it satisfies $P^2 = P$.

One readily checks that $P \in \mathcal{L}(X)$ is a projection if and only if $I - P$ is a projection. We leave it as an exercise to check the following:

Proposition 3.2.6. *Assume $X = M_1 \dot{+} M_2$, and define $P_1, P_2 : X \to X$ by*

$$P_1(x) := x_1 \,, \quad P_2(x) := x_2 \,,$$

whenever $x = x_1 + x_2$ with $x_1 \in M_1$ and $x_2 \in M_2$.

Then P_1, P_2 are projections in $\mathcal{L}(X)$ such that

$$P_1 + P_2 = I, \quad P_1 P_2 = P_2 P_1 = 0,$$

$$P_1(X) = M_1 = \ker(P_2) \quad \textit{and} \quad P_2(X) = M_2 = \ker(P_1).$$

The map P_1 is called the projection (from X) on M_1 along M_2, *while the map P_2 is called* the projection (from X) on M_2 along M_1.

Example 3.2.7. Consider $X = \mathbb{R}^2$. The most familiar direct sum decomposition of $\mathbb{R}^2$ is of course

$$\mathbb{R}^2 = M_1 \dot{+} M_2$$

where $M_1 = \{(s,0) : s \in \mathbb{R}\}$ and $M_2 = \{(0,t) : t \in \mathbb{R}\}$, in which case P_1 and P_2 are the usual coordinate maps, i.e.,

$$P_1((s,t)) = (s,0) \quad \text{and} \quad P_2((s,t)) = (0,t)\,.$$

However, there are infinitely ways of writing $\mathbb{R}^2$ as a direct sum, even if we fix M_1 to be the first axis: indeed, we can then let M_2 be any line through the origin which is different from M_1. For example, if we choose $M_2 = \{(t,t) : t \in \mathbb{R}\}$, then $M_1 \cap M_2 = \{(0,0)\}$, and for any $(u,v) \in \mathbb{R}^2$ we have

$$(u,v) = (u-v,0) + (v,v)\,, \quad \text{with } (u-v,0) \in M_1 \text{ and } (v,v) \in M_2.$$

Thus, in this case, we get that the projection maps $P_1, P_2 : \mathbb{R}^2 \to \mathbb{R}^2$ are given by $P_1((u,v)) = (u-v,0)$ and $P_2((u,v)) = (v,v)$ for all $(u,v) \in \mathbb{R}^2$.

A converse to Proposition 3.2.6 is the following:

Proposition 3.2.8. *Assume $P \in \mathcal{L}(X)$ is a projection. Then we have*

$$X = P(X) \dotplus \ker(P).$$

Moreover, we have $P(X) = \ker(I-P)$, $\ker(P) = (I-P)(X)$, and P is the projection from X on $P(X)$ along $\ker(P)$.

Proof. Let $x \in X$. Note that

$$x = P(x) + (x - P(x))\,. \tag{3.2.1}$$

Since

$$P\Big(x - P(x)\Big) = P(x) - P^2(x) = 0\,,$$

that is, $(x - P(x)) \in \ker(P)$, this shows that $X = P(X) + \ker(P)$.

Next, assume that $x \in P(X) \cap \ker(P)$. Thus we have $x = P(y)$ for some $y \in X$ and $P(x) = 0$. This gives that

$$x = P(y) = P^2(y) = P(P(y)) = P(x) = 0\,.$$

Hence, $P(X) \cap \ker(P) = \{0\}$, so $X = P(X) \dotplus \ker(P)$.

If we now set $M_1 := P(X)$ and $M_2 := \ker(P)$, then, using the notation from Proposition 3.2.6, we get from equation (3.2.1) that $P_1 = P$ and $P_2 = I - P$, so the last assertions follow readily from this proposition. ∎

Remark 3.2.9. If M_1 is a subspace of a vector space X, then it can be shown that M_1 can be algebraically complemented, i.e., that there exists a subspace M_2 of X such that $X = M_1 \dot{+} M_2$. (In fact, if $\{0\} \neq M_1 \neq X$, then there exist infinitely many such subspaces, which are all isomorphic to each other.) When X is infinite-dimensional, the proof requires the axiom of choice, in the form of Zorn's lemma, as explained in more advanced textbooks.

We now turn our attention to normed spaces.

Definition 3.2.10. Assume that X is a normed space over $\mathbb{F}$, and let M_1 and M_2 be subspaces of X. We will say that X is the (internal) *direct sum* of M_1 and M_2, and write

$$X = M_1 \oplus M_2 \,,$$

when $X = M_1 \dot{+} M_2$ *and* both M_1 and M_2 are closed in X.

Proposition 3.2.11. *Let X be a normed space and assume $P \in \mathcal{L}(X)$ is a projection which is bounded (so $P \in \mathcal{B}(X)$). Then we have*

$$X = P(X) \oplus \ker(P) \,.$$

Proof. We know from Proposition 3.2.8 that $X = P(X) \dot{+} \ker(P)$, so it remains only to check that $P(X)$ and $\ker(P)$ are closed in X. Since $\ker(P) = P^{-1}(\{0\})$ and P is continuous, $\ker(P)$ is closed. Moreover, since $P(X) = \ker(I - P)$ and $I - P$ is continuous, we also get that $P(X)$ is closed. ∎

Example 3.2.12. Let V_1, V_2 be normed spaces over $\mathbb{F}$. As is readily verified (if not already known), the direct product $V := V_1 \times V_2$ becomes a normed space with respect to the norm given by

$$\|(v_1, v_2)\| := \|v_1\| + \|v_2\| \,.$$

Moreover, with $\widetilde{V_1} := \{(v_1, 0) : v_1 \in V_1\}$ and $\widetilde{V_2} := \{(0, v_2) : v_2 \in V_2\}$, we have $V = \widetilde{V_1} \dot{+} \widetilde{V_2}$ (cf. Exercise 3.5). Let $P_1 \in \mathcal{L}(V)$ denote the projection from V on $\widetilde{V_1}$ along $\widetilde{V_2}$. Then

$$\|P_1((v_1, v_2))\| = \|(v_1, 0)\| = \|v_1\| + \|0\| = \|v_1\| \leq \|(v_1, v_2)\|$$

for all $(v_1, v_2) \in V$, so P_1 is bounded. Thus, Proposition 3.2.11 gives that $V = P_1(V) \oplus \ker(P_1)$. As $P_1(V) = \widetilde{V_1}$ and $\ker(P_1) = \widetilde{V_2}$, we get that

$$V_1 \times V_2 = \widetilde{V_1} \,\oplus\, \widetilde{V_2} \,.$$

Remark 3.2.13. Somewhat surprisingly, if X is a normed space and $X = M_1 \oplus M_2$ for some closed subspaces M_1, M_2, then it may happen that the projection P_1 from X on M_1 along M_2 is unbounded (cf. Exercise 3.7), in which case the projection P_2 on M_2 along M_1 is also unbounded (since $P_1 + P_2 = I$). However this peculiarity does not arise if X is a Banach space, but the proof of this fact is beyond the scope of these notes. (One may for example invoke the so-called closed graph theorem, proven in more advanced courses).

Remark 3.2.14. It is common to say that a closed subspace M of a normed space X can be *complemented* when there exists a closed subspace N of X such that $X = M \oplus N$. It is not true that a closed subspace can always be complemented, even if X is a Banach space; for example, it is known that the closed subspace

$$c_0(\mathbb{N}) = \{f \in \ell^\infty(\mathbb{N}) : \lim_{n\to\infty} f(n) = 0\}$$

can not be complemented in $\ell^\infty(\mathbb{N})$ (with uniform norm), but we don't have yet the tools necessary to prove this. Proposition 3.2.11 tells us that if a closed subspace M of a normed space X is the range of a projection P in $\mathcal{B}(X)$, then M can be complemented. The previous remark implies that the converse holds when X is a Banach space. It is also known that a finite dimensional subspace of a normed space can always be complemented. We will see in the next chapter that any closed subspace of a Hilbert space can be complemented (by its orthogonal complement).

Direct sums and projections are useful in connection with the study of linear operators.

Proposition 3.2.15. *Assume X is a vector space such that $X = M_1 \dotplus M_2$ for some subspaces M_1, M_2. To each $S_1 \in \mathcal{L}(M_1)$ and $S_2 \in \mathcal{L}(M_2)$, we may associate an operator $S = S_1 \dotplus S_2 \in \mathcal{L}(X)$ given by*

$$(S_1 \dotplus S_2)(x) := S_1(x_1) + S_2(x_2)$$

for $x = x_1 + x_2 \in X$ with $x_1 \in M_1$ and $x_2 \in M_2$.

If P_1 (resp. P_2) denote the projection from X on M_1 along M_2 (resp. on M_2 along M_1), then $S = S_1 \dotplus S_2$ commutes with each P_j, that is, we have $SP_j = P_jS$ for $j = 1, 2$.

Moreover, if X is a normed space, M_1 and M_2 are closed in X, and P_1 is bounded (or, equivalently, P_2 is bounded), then $S_1 \dotplus S_2$ is bounded if and only if S_1 and S_2 are bounded.

(Note that if X is a Banach space, then P_1 and P_2 are automatically bounded, as mentioned in Remark 3.2.13).

Proof. The reader should have no difficulty to provide the necessary details, so we leave this as an exercise. ■

Definition 3.2.16. Let notation be as in Proposition 3.2.15. When an operator $S \in \mathcal{L}(X)$ can be written as $S = S_1 \dot{+} S_2$ for some $S_1 \in \mathcal{L}(M_1)$ and $S_2 \in \mathcal{L}(M_2)$, then we say that S is *decomposable* w.r.t. $X = M_1 \dot{+} M_2$.

When an operator is decomposable w.r.t. a direct sum decomposition, we may study it by studying each of its components. It is therefore of interest to know when this happens. The following notion will be useful.

Definition 3.2.17. Let X be a vector space and $T \in \mathcal{L}(X)$. A subset M of X is said to be *invariant under* T when $T(M) \subseteq M$.

Example 3.2.18. Let notation be as in Proposition 3.2.15, and set $S := S_1 \dot{+} S_2 \in \mathcal{L}(X)$. Then M_1 and M_2 are both invariant under S. Indeed, if $x_1 \in M_1$, then $S(x_1) = S_1(x_1) \in M_1$. Similarly, $S(x_2) \in M_2$ for all $x_2 \in M_2$.

Example 3.2.19. Assume X is a vector space over $\mathbb{F}$ and $T \in \mathcal{L}(X)$. The range of T is then a subspace of X which is invariant under T: indeed, with $M = T(X)$, we have $T(M) \subseteq T(X) = M$.

Moreover, for $\lambda \in \mathbb{F}$, set

$$E_\lambda^T := \ker(T - \lambda I)\,.$$

Then E_λ^T is also a subspace of X, which is invariant under T: indeed, for every $x \in E_\lambda^T$, we have $T(x) = \lambda x \in E_\lambda^T$. Of course, when $E_\lambda^T \neq \{0\}$, then λ is an eigenvalue of T, and E_λ^T is the associated eigenspace.

Proposition 3.2.20. *Assume X is a vector space over $\mathbb{F}$ such that $X = M_1 \dot{+} M_2$ for some subspaces M_1, M_2. Let P_1 (resp. P_2) denote the projection on M_1 along M_2 (resp. on M_2 along M_1) and consider $S \in \mathcal{L}(X)$. Then the following conditions are equivalent:*

(*a*) *S is decomposable w.r.t. $X = M_1 \dot{+} M_2$;*

(*b*) *both M_1 and M_2 are invariant under S ;*

(*c*) *S commutes with P_1 ;*

(*d*) *S commutes with P_2 .*

Proof. If (a) holds, then it follows from Proposition 3.2.15 that (c) and (d) hold. Since $P_2 = I - P_1$, it is elementary that (c) is equivalent to (d).

Assume that (c) holds. Let $x_1 \in M_1$. Then we have

$$S(x_1) = S(P_1(x_1)) = P_1(S(x_1)) \in P_1(X) = M_1 .$$

Thus, M_1 is invariant under S. Moreover, as (d) also holds, we get in a similar way that M_2 is invariant under S. Hence, (b) holds.

Finally, assume (b) holds. Then, for $j = 1, 2$, we may define $S_j \in \mathcal{L}(M_j)$ by

$$S_j(x_j) := S(x_j) \quad \text{for all } x_j \in M_j .$$

Let $x \in X$. Then $x = x_1 + x_2$ for $x_1 \in M_1$ and $x_2 \in M_2$, so we get

$$S(x) = S(x_1 + x_2) = S(x_1) + S(x_2) = S_1(x_1) + S_2(x_2) = (S_1 \dot{+} S_2)(x) .$$

This shows that $S = S_1 \dot{+} S_2$, hence that (a) holds. ■

Remark 3.2.21. Assume that X is a vector space over $\mathbb{F}$ and $T \in \mathcal{L}(X)$ has an eigenvalue $\lambda \in \mathbb{F}$. (For example, if X is finite dimensional and $\mathbb{F} = \mathbb{C}$, then every $T \in \mathcal{L}(X)$ has an eigenvalue). A natural question is then whether the eigenspace $M_1 = E_\lambda^T$, which is invariant under T, can be complemented in X by some subspace M_2 which is also invariant under T. This may not be the case (see Exercise 3.9), but if it happens, then we have $T = \lambda I_{M_1} \dot{+} T_2$ where $T_2 = T_{|M_2} \in \mathcal{L}(M_2)$, and we can focus on T_2. Moreover, in good cases, one can proceed further in an inductive way. This is basically the main idea used in the proof of the spectral theorem for symmetric real matrices. The same idea can also be used for compact self-adjoint operators on Hilbert spaces.

Finally, we mention for completeness that one can also consider direct sums decompositions of a vector space with more than two summands.

Let X is a vector space over $\mathbb{F}$, and assume that $M_1, M_2, \ldots, M_n$ are subspaces of X. Then X is said to be the (internal) *algebraic direct sum of* $M_1, M_2, \ldots, M_n$ if $X = M_1 + M_2 + \cdots + M_n$ and the following independence condition holds: if $x_1 \in M_1, x_2 \in M_2, \ldots, x_n \in M_n$ and

$$x_1 + x_2 + \cdots + x_n = 0 ,$$

then $x_1 = x_2 = \cdots = x_n = 0$. We leave it as an easy exercise to check that these two conditions are equivalent to requiring that every $x \in X$ can be written *in a unique way* as $x = x_1 + \cdots + x_n$ with $x_1 \in M_1, \ldots, x_n \in M_n$.

3.3 Extension by density and continuity

This short section is devoted to a very useful principle in linear analysis, often called *the principle of extension by density and continuity.* We will need the following elementary lemma, which is probably well-known.

Lemma 3.3.1. *Assume that X and Y are metric spaces and f, g are continuous maps from X to Y which agree on a dense subset X_0 of X. Then $f = g$.*

Proof. Let $x \in X$. Since X_0 is dense in X, there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ in X_0 which converges to x. By continuity of f and g, we get

$$f(x) = \lim_n f(x_n) = \lim_n g(x_n) = g(x)\,.$$

■

Theorem 3.3.2. *Assume that X is a normed space and Y is a Banach space (both over $\mathbb{F}$). Assume also that X_0 is a dense subspace of X, while Y_0 is a subspace of Y. Let $T_0 \in \mathcal{B}(X_0, Y_0)$. Then T_0 extends in a unique way to an operator $T \in \mathcal{B}(X, Y)$. It satisfies that $\|T\| = \|T_0\|$.*

Proof. Let $x \in X$. Since X_0 is dense in X, there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ in X_0 such that $\|x - x_n\| \to 0$ as $n \to \infty$. In particular, $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in X_0. We claim that $\{T_0(x_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. Indeed, let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ such that

$$\|x_m - x_n\| < \varepsilon/\|T_0\| \quad \text{for all } m, n \geq N\,.$$

Then, for all $m, n \in N$, we get

$$\|T_0(x_m) - T_0(x_n)\| = \|T_0(x_m - x_n)\| = \|T_0\|\,\|x_m - x_n\| < \varepsilon\,,$$

as desired.

Since Y is complete, we can conclude that there exists some $y \in Y$ such that $\lim_n T_0(x_n) = y$. Note that y only depends on x. Indeed, assume $\{x_n'\}_{n\in\mathbb{N}}$ is another sequence in X_0 converging to x. Then the sequence $x_1, x_1', x_2, x_2', \ldots, x_n, x_n', \ldots$ in X_0 also converges to x, so, arguing as above, we get that there exists some $z \in Y$ such that the sequence $T_0(x_1), T_0(x_1'), T_0(x_2), T_0(x_2'), \ldots, T_0(x_n), T_0(x_n'), \ldots$ converges to z. This implies that

$$\lim_n T_0(x_n') = z = \lim_n T_0(x_n) = y\,.$$

Hence it makes sense to define $T(x) := y$. Doing this for every $x \in X$, we get a map $T : X \to Y$, and it is easy to check that T is linear, so we leave this as an exercise.

Next, we show that T is bounded. Let $x \in X$ and pick $\{x_n\}_{n\in\mathbb{N}}$ in X_0 converging to x. As $T(x) = \lim_n T_0(x_n)$ and $\|T_0(x_n)\| \leq \|T_0\| \, \|x_n\|$ for all $n \in \mathbb{N}$, we get

$$\|T(x)\| = \lim_n \|T_0(x_n)\| \leq \|T_0\| \lim_n \|x_n\| = \|T_0\| \, \|x\| \, .$$

It follows that $T \in \mathcal{B}(X, Y)$ with $\|T\| \leq \|T_0\|$.

Further, T is an extension of T_0. Indeed, let $x \in X_0$. Then set $x_n := x$ for all $n \in \mathbb{N}$. Since $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in X_0 converging to x, we get that

$$T(x) = \lim_n T_0(x_n) = T_0(x) \, .$$

The uniqueness of T as an extension of T_0 is immediate from Lemma 3.3.1. Finally, we have

$$\begin{aligned} \|T_0\| &= \sup\{\|T_0(x)\| : x \in X_0, \|x\| \leq 1\} \\ &= \sup\{\|T(x)\| : x \in X_0, \|x\| \leq 1\} \\ &\leq \sup\{\|T(x)\| : x \in X, \|x\| \leq 1\} = \|T\| \leq \|T_0\| \, . \end{aligned}$$

Thus, $\|T\| = \|T_0\|$, as desired. ■

Remark 3.3.3. The conclusion of Theorem 3.3.2 is not necessarily true if Y is a normed space which is not complete (cf. Exercise 3.15).

An interesting special case of Theorem 3.3.2 is when T_0 is an isometry. We recall that a linear map between normed spaces is an *isometry* when it is norm-preserving. A linear isometry is clearly bounded.

Corollary 3.3.4. *Assume that X is a normed space, Y is a Banach space, X_0 is a dense subspace of X, Y_0 is a subspace of Y, and $U_0 \in \mathcal{L}(X_0, Y_0)$ is an isometry. Then the unique extension of U_0 to an operator U in $\mathcal{B}(X, Y)$ is also an isometry.*

Proof. Theorem 3.3.2 guarantees that U_0 extends in a unique way to $U \in \mathcal{B}(X, Y)$. Let $x \in X$ and pick $\{x_n\}_{n\in\mathbb{N}}$ in X_0 converging to x. We then have $U(x) = \lim_n U_0(x_n)$, so we get

$$\|U(x)\| = \lim_n \|U_0(x_n)\| = \lim_n \|x_n\| = \|x\| \, .$$

■

Using Corollary 3.3.4, it can be shown that the completion of a (non-complete) normed space is unique up to isometric isomorphism (cf. Exercise 3.17). We also record an important particular case of Theorem 3.3.2.

Corollary 3.3.5. *Assume that X is a Banach space and X_0 is a dense subspace of X. Then every $T_0 \in \mathcal{B}(X_0)$ extends in a unique way to an operator $T \in \mathcal{B}(X)$, which satisfies that $\|T\| = \|T_0\|$.*

Example 3.3.6. Let $a, b \in \mathbb{R}$, $a < b$, and equip the space $C([a, b])$ of all continuous complex functions on $[a, b]$ with the norm $\|f\|_2 = (\int_a^b |f(s)|^2\, ds)^{1/2}$. Considering the square $[a, b] \times [a, b]$ as a metric space w.r.t. the Euclidean metric inherited from $\mathbb{R}^2$, let $K : [a, b] \times [a, b] \to \mathbb{C}$ be a continuous function. One can then associate to K an *integral operator* T_K on $C([a, b])$ as follows.

Let $f \in C([a, b])$. Since the function $t \to K(s, t)\, f(t)$ is continuous on $[a, b]$ for each $s \in [a, b]$, we may define a function $T_K(f) : [a, b] \to \mathbb{C}$ by

$$[T_K(f)](s) = \int_a^b K(s, t)\, f(t)\, dt \quad \text{for all s} \in [a, b]\,.$$

We leave it as an exercise to verify, using basic knowledge from elementary analysis, that $T_K(f)$ is continuous on $[a, b]$ and satisfies

$$\|T_K(f)\|_2 \leq \Big(\int_a^b \int_a^b |K(s, t)|^2\, dsdt\Big)^{1/2} \|f\|_2\,.$$

As the map $f \to T_K(f)$ is then clearly linear, it follows that T_K is a bounded linear operator from $C([a, b])$ into itself.

Let now $L^2([a, b])$ denote the L^2-space associated with the measure space $([a, b], \mathcal{A}, \mu)$, where μ is the Lebesgue measure on the σ-algebra $\mathcal{A}$ of all Lebesgue measurable subsets of $[a, b]$.

As we may identify $C([a, b])$ with a dense closed subspace of $L^2([a, b])$ (cf. Exercise 2.7), we get from Corollary 3.3.5 that T_K has a unique extension to a bounded operator on $L^2([a, b])$, also denoted by T_K. The function K is usually called the *kernel* of the integral operator T_K. We will come back to such integral operators later.

We note that more generally, one can define integral operators associated with kernels K which are L^2-functions on $[a, b] \times [a, b]$ (with respect to the Lebesgue product measure), but this requires a thorough knowledge of integration theory on product spaces.

3.4 Exercises

Exercise 3.1. Let H be a Hilbert space which is infinite-dimensional (as a vector space). Argue first that there exists an orthonormal sequence $\{x_n\}_{n\in\mathbb{N}}$ in H. Then use this sequence to show that the unit ball H_1 is not compact.

Exercise 3.2. Let X be a normed space. Let M denote a finite-dimensional subspace of X, and assume $M \neq X$.

a) Let $x \in X \setminus M$. Show that $d := \inf_{m\in M} \|x - m\| > 0$.

b) Show that there exists $y \in X$ such that $\|y\| = 1$ and

$$\frac{1}{2} \leq \|y - m\| \quad \text{for all } m \in M.$$

c) Assume that X is infinite-dimensional (as a vector space). Show that the unit ball X_1 is not compact.

(*Hint* : Use b) to construct inductively a sequence $\{y_n\}_{n\in\mathbb{N}}$ in X_1 such that $1/2 \leq \|y_n - y_k\|$ for all $1 \leq k < n$.)

Exercise 3.3. Let X be the subspace of $\ell^\infty(\mathbb{N})$ given by

$$X = \{f : \mathbb{N} \to \mathbb{C} : f(n) = 0 \text{ for all but finitely many } n\}.$$

a) Show that X is infinite-dimensional.

b) Consider X as a normed space w.r.t. $\|f\|_u = \sup_{n\in\mathbb{N}} |f(n)|$ and let $L : X \to \mathbb{C}$ be defined by

$$L(f) = \sum_{n=1}^{\infty} f(n)$$

for all $f \in X$. Clearly, $L \in \mathcal{L}(X, \mathbb{C})$. Show that L is unbounded. Check also that $\ker(L)$ is not closed in X.

Exercise 3.4. Let $\mathcal{P}_\mathbb{R}$ denote the real vector space consisting of all polynomials in one real variable with real coefficients. For $p \in \mathcal{P}_\mathbb{R}$, set

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

a) Explain why $p \to \|p\|$ gives a well-defined norm on $\mathcal{P}_\mathbb{R}$.

b) Define a linear operator $D : \mathcal{P}_\mathbb{R} \to \mathcal{P}_\mathbb{R}$ by

$$D(p) = p' \;\; \text{(the derivative of } p).$$

Show that D is unbounded. Conclude that $\mathcal{P}_\mathbb{R}$ is infinite-dimensional.

Exercise 3.5. Let V_1, V_2 be vector spaces over $\mathbb{F}$. We recall that their (external) algebraic direct sum (also called their algebraic direct product) is the vector space

$$V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1,\ v_2 \in V_2\},$$

with operations given by

$$(v_1, v_2) + (v_1', v_2') := (v_1 + v_1', v_2 + v_2')\,,$$
$$\lambda\,(v_1, v_2) := (\lambda v_1, \lambda v_2)$$

for $v_1, v_1' \in V_1, v_2, v_2' \in V_2$ and $\lambda \in \mathbb{F}$.

a) Set $\widetilde{V}_1 = \{(v_1, 0) : v_1 \in V_1\}$ and $\widetilde{V}_2 = \{(0, v_2) : v_2 \in V_2\}$. Check that $\widetilde{V}_i$ is a subspace of $V_1 \times V_2$ which is isomorphic to V_i for $i = 1, 2$, and that

$$V_1 \times V_2 = \widetilde{V}_1 \dot{+} \widetilde{V}_2\,.$$

b) Assume X is a vector space, M_1 and M_2 are subspaces of X and $X = M_1 \dot{+} M_2$. Show that X is isomorphic to $M_1 \times M_2$.

Exercise 3.6. Let V_1, V_2 be normed spaces over $\mathbb{F}$ and set $V := V_1 \times V_2$. For $p \in [1, \infty)$ and $(v_1, v_2) \in V$, set

$$\|(v_1, v_2)\|_p := \Big(\|v_1\|^p + \|v_2\|^p\Big)^{1/p}.$$

Set also $\|(v_1, v_2)\|_\infty := \max\{\|v_1\|, \|v_2\|\}$.

a) Check that $\|\cdot\|_p$ gives a norm, called the p-norm on V, for each $p \in [1, \infty]$. Then show that all these p-norms on V are equivalent.

b) Set $\widetilde{V}_1 := \{(v_1, 0) : v_1 \in V_1\}$ and $\widetilde{V}_2 := \{(0, v_2) : v_2 \in V_2\}$, so $V = \widetilde{V}_1 \dot{+} \widetilde{V}_2$ (cf. Exercise 3.5). Let $P_1 \in \mathcal{L}(V)$ denote the projection from V on $\widetilde{V}_1$ along $\widetilde{V}_2$, and consider the normed space $(V, \|\cdot\|_p)$ for some $p \in [1, \infty]$. Show that P_1 is bounded.

Exercise 3.7. Let X be the subspace of $\ell^1(\mathbb{N})$ given by

$$X = \{f : \mathbb{N} \to \mathbb{C} : f(n) = 0 \text{ for all but finitely many } n\}$$

and consider X as a normed space w.r.t. the 1-norm $\|f\| := \sum_{n\in\mathbb{N}} |f(n)|$. Let M_1 be the subspace of X given by

$$M_1 = \{f \in X : f(2n) = n\, f(2n-1) \text{ for all } n \in \mathbb{N}\}\,,$$

and let M_2 be the subspace of X given by

$$M_2 = \{f \in X : f(2n-1) = 0 \text{ for all } n \in \mathbb{N}\}\,.$$

Show that $X = M_1 \oplus M_2$, and that the projection from X on M_1 along M_2 is unbounded.

Exercise 3.8. Prove Proposition 3.2.15.

Exercise 3.9. Set $X = \mathbb{C}^2$ and let $\{e_1, e_2\}$ denote the standard basis of X. Let $T \in \mathcal{L}(X)$ be the linear operator satisfying $T(e_1) = e_1$ and $T(e_2) = i\,e_1 + e_2$. Clearly, 1 is an eigenvalue of T. Set $M_1 = E_1^T$, so M_1 is a subspace of X which is invariant under T.

Show that there is no subspace M_2 of X which is invariant under T and satisfies that $X = M_1 \dot{+} M_2$.

Exercise 3.10. Let X be a vector space over $\mathbb{F}$.

a) Assume $X = M_1 \dot{+} M_2$ for some subspaces M_1 and M_2 of X, and let P_1, P_2 denote the associated projection maps. Define $S \in \mathcal{L}(X)$ by

$$S(x) := P_1(x) - P_2(x) = 2\,P_1(x) - x\,.$$

Check that $S^2 = I$. Moreover, check that

$$M_1 = \ker(I - S) = \{x \in X : S(x) = x\},$$

$$M_2 = \ker(I + S) = \{x \in X : S(x) = -x\}.$$

The map S is called the *symmetry through* M_1 *along* M_2.

b) Assume $S \in \mathcal{L}(X)$ satisfies $S^2 = I$. Show that $(I+S)(X) = \ker(I-S)$ and $(I - S)(X) = \ker(I + S)$. Moreover, show that

$$X = \ker(I - S) \dot{+} \ker(I + S)$$

and that S is the symmetry through $\ker(I - S)$ along $\ker(I + S)$. Finally, check that S is decomposable with respect to this direct sum decomposition.

c) Assume now that X is a normed space and that $S \in \mathcal{B}(X)$ satisfies $S^2 = I$. Deduce that $X = \ker(I - S) \oplus \ker(I + S)$.

d) Let $a > 0$ and consider the space $X = C([-a, a])$ with the uniform norm. Define $S : X \to X$ by

$$[S(f)](t) = f(-t) \quad \text{for all } f \in X \text{ and } t \in [-a, a]\,.$$

Check that S is bounded and $S^2 = I$. Deduce that $X = X_{\mathsf{even}} \oplus X_{\mathsf{odd}}$, where

$$X_{\mathsf{even}} := \{g \in X : g(-t) = g(t) \text{ for all } t \in [-a, a]\} \quad \text{and}$$

$$X_{\mathsf{odd}} := \{h \in X : h(-t) = -h(t) \text{ for all } t \in [-a, a]\}.$$

Exercise 3.11. Let $X = \mathbb{R}^3$ and consider X as an inner product space w.r.t. the Euclidean inner product. Let $R \in \mathcal{L}(X)$, $R \neq I_X$. Assume that the standard matrix U of R is orthogonal (i.e., $U^tU = I$) and has determinant equal to 1.

a) Show that 1 is an eigenvalue of R, and that the associated eigenspace $M := E_1^R$ is 1-dimensional.

b) Let $N = M^\perp$ denote the orthogonal complement of M. As should be well-known, we have $X = M \dot{+} N$.

Show that R is decomposable w.r.t. $X = M \dot{+} N$, so we may write $R = I_M \dot{+} R'$ with $R' \in \mathcal{L}(N)$.

c) Let $\mathcal{B}'$ be an orthonormal basis for N. Show that the matrix of R' w.r.t. $\mathcal{B}'$ is a 2×2 rotation matrix.

d) Describe how R acts in geometrical terms.

Exercise 3.12. Let X be a vector space over $\mathbb{F}$ and let M be a subspace of X. Define a relation $\sim_M$ on X by $x \sim_M y$ if and only if $y - x \in M$.

a) Check that $\sim_M$ is an equivalence relation on X.

The equivalence class of $x \in X$ w.r.t. $\sim_M$ is the set $\{x + m : m \in M\}$, which we will denote $x + M$. The set consisting of all these equivalence classes is called the *quotient space* (of X by M), and is denoted by X/M.

b) Check that X/M becomes a vector space over $\mathbb{F}$ with respect to the operations given by

$$(x + M) + (x' + M) := (x + x') + M\,, \quad \lambda\,(x + M) := (\lambda x) + M$$

for all $x, x' \in X$ and $\lambda \in \mathbb{F}$. You should first argue that these operations are well-defined.

The map $Q : X \to X/M$ given by $Q(x) = x + M$ is called the *quotient map*. It is evident that Q is linear.

c) Assume now that $X = M \dot{+} N$ for some subspaces M and N of X. Show that X/M is isomorphic to N. (Similarly, X/N is isomorphic to M).

Hint : Consider the map $\pi : N \to X/M$ given by $\pi := Q_{|N} : N \to X/M$, i.e.,

$$\pi(y) := y + M \quad \text{for all } y \in N\,, \tag{3.4.1}$$

and show that π is an isomorphism.

Exercise 3.13. Assume that X is a normed space and M is a closed subspace of X. For each element $x+M$ in the quotient space X/M (defined in the previous exercise), set

$$\|x+M\| := \inf_{m\in M} \|x+m\| \ (= \inf_{m\in M} \|x-m\|)\,.$$

a) Show that the map $x+M \to \|x+M\|$ gives a norm on X/M, called the *quotient norm*. Then check that the quotient map $Q : X \to X/M$ is contractive, i.e., $\|Q(x)\| \leq \|x\|$ for all $x \in X$. In particular, we have $Q \in \mathcal{B}(X, X/M)$.

b) Let now N be a subspace of X such that $X = M \dot{+} N$.
Let $\pi : N \to X/M$ be defined by (3.4.1), and let $P' : X \to X$ denote the projection from X on N along M. Consider X/M as a normed space w.r.t. the quotient norm, and $M \times N$ as a normed space w.r.t. any choice of p-norm, cf. Exercise 3.6. Show that the following assertions are equivalent:

(i) N is closed in X (so $X = M \oplus N$) and $\pi : N \to X/M$ is an isomorphism of normed spaces;

(ii) The map $(m,n) \to m+n$ from $M \times N$ to X is an isomorphism of normed spaces;

(iii) P' is bounded.

Exercise 3.14. Consider $X = \mathbb{R}^2$. Find three subspaces M_1, M_2, M_3 of X such that

- $X = M_1 + M_2 + M_3$;
- $M_1 \cap M_2 \cap M_3 = \{(0,0)\}$;
- X is *not* the algebraic direct sum of M_1, M_2 and M_3.

This illustrates why the definition of an algebraic direct sum of more than two subspaces must be formulated in a different way than the one you possibly had guessed.

Exercise 3.15. Let X be a Banach space having a dense subspace X_0 which is not complete. Consider the identity map $I_0 : X_0 \to X_0$. Show that I_0 does not have an extension to a bounded linear map $I_0 : X \to X_0$.

Exercise 3.16. Assume that X is a normed space and Y is a Banach space (both over $\mathbb{F}$), and let $\{T_k\}_{k\in\mathbb{N}}$ be a sequence in $\mathcal{B}(X,Y)$ which is *uniformly bounded* in the sense that $M := \sup_{k\in\mathbb{N}} \|T_k\| < \infty$.

Moreover, assume that there exists a dense subset S of X such that $\{T_k(x)\}_{k\in\mathbb{N}}$ converges in Y for every $x \in S$.

Show that there exists $T \in \mathcal{B}(X,Y)$ such that

$$T(x) = \lim_k T_k(x) \quad \text{for all } x \in X\,.$$

Exercise 3.17. Assume X_0 is a normed space and let (X, i) denote a *completion of* X_0, that is, X is a Banach space and $i : X_0 \to X$ is a linear isometry such that $i(X_0)$ is dense in X. (As mentioned in Remark 1.1.7, such a completion always exists.)

Show that (X, i) is unique up to isometric isomorphism, meaning that the following holds: if (X', i') is another completion of X_0, then there exists an isometric isomorphism $U : X \to X'$ such that $i' = U \circ i$.

Exercise 3.18. a) Provide the details missing in Example 3.3.6.

We outline below how one may define more directly integral operators on $L^2([a,b])$. Let μ denote the Lebesgue measure on the Lebesgue measurable subsets of $[a,b]$ and let K be a continuous complex function on $[a,b]\times[a,b]$. For each $s \in [a,b]$, let $k_s : [a,b] \to \mathbb{C}$ denote the continuous function defined by $k_s(t) := K(s,t)$ for all $t \in [a,b]$.

b) Let $f \in \mathcal{L}^2([a,b])$ and $s \in [a,b]$. Show the function $k_s\, f$ is Lebesgue integrable on $[a,b]$ and satisfies

$$\int_{[a,b]} k_s\, f\; d\mu\,| \leq \|k_s\|_2\, \|f\|_2\,.$$

c) Let $f \in \mathcal{L}^2([a,b])$ and define $g : [a,b] \to \mathbb{K}$ by

$$g(s) = \int_{[a,b]} k_s\, f\; d\mu = \int_{[a,b]} K(s,t)\, f(t)\; d\mu(t) \quad \text{for each } s \in [a,b]\,.$$

Show that g is continuous and check that

$$\|g\|_2 \;\leq\; M\,\|f\|_2\,, \quad \text{where } M := \Big(\int_a^b\int_a^b |K(s,t)|^2\; dsdt\Big)^{1/2}.$$

Deduce that we obtain a linear map $T_K^0 : \mathcal{L}^2([a,b]) \to \mathcal{L}^2([a,b])$ by setting

$$\Big(T_K^0(f)\Big)(s) := \int_{[a,b]} k_s\, f\, d\mu \quad \text{for each } f \in \mathcal{L}^2([a,b]) \text{ and all } s \in [a,b],$$

which satisfies that $\|T_K^0(f)\|_2 \;\leq\; M\,\|f\|_2$ for all $f \in \mathcal{L}^2([a,b])$.

d) Check that the operator $T_K : L^2([a,b]) \to L^2([a,b])$ defined by

$$T_K([f]) = [T_K^0(f)] \quad \text{for all } [f] \in L^2([a,b])$$

is well-defined, linear and bounded, with $\|T_K\| \leq M$.

CHAPTER 4

More on Hilbert spaces

By a Hilbert space we always mean a Hilbert space over $\mathbb{F}$, unless otherwise specified.

4.1 Geometry in Hilbert spaces

In courses in elementary linear algebra, one learns that if M is finite-dimensional subspace of an inner product space H, then every vector in H can be written in a unique way as the sum of a vector in M and a vector in the orthogonal complement $M^\perp$. Since M and $M^\perp$ are both closed subspaces of H, this means that $H = M \oplus M^\perp$. The projection of H on M along $M^\perp$ is then called the orthogonal projection of H on M. As we are going to establish, such a decomposition of H also holds when H is a Hilbert space and M is closed subspace of H, not necessarily finite-dimensional.

We recall first that if (X, d) is a metric space, $x \in X$ and A is a nonempty subset of X, then the *distance from x to A* is defined by

$$d(x, A) = \inf\{d(x, y) : y \in A\}\,.$$

If for example A is compact, then the function $y \to d(x, y)$, being continuous, will attain its minimum on A; hence, in this case, there exists some (not necessarily unique) $x_A \in A$ such that $d(x, A) = d(x, x_A)$. However, if A is only closed, then such an x_A may not exist (cf. Exercise 4.1).

If we now consider a Hilbert space H with the metric d_H associated to its norm, a vector $x \in H$ and a closed subspace M of H, then M is not compact and the result above does not apply. However, if M is finite-dimensional, then we know from previous courses that there exists a *unique* $x_M \in M$ which gives the best approximation to x in M, i.e., which satisfies that

$$\|x - x_M\| \;\leq\; \|x - y\| \quad \text{for all } y \in M,$$

that is, we have $d_H(x, x_M) = d_H(x, M)$. Moreover, we also know that x_M is given as the orthogonal projection of x on M. When M is not finite-dimensional, but still closed, we will reverse this prosess by showing first that there exists a unique best approximation x_M to x in M, and then use this to define the orthogonal projection of x on M.

We will actually prove a more general result, valid for any closed convex subset of H. We recall that a subset C of some vector space V (over $\mathbb{F}$) is called *convex* if C contains the line segment between any two elements of C, i.e., if we have $(1-t)x + ty \in C$ whenever $x, y \in C$ and $t \in [0,1]$.

Clearly, any subspace of a vector space is convex, as is any ball in a normed space. Using that the norm in a Hilbert space satisfies the parallellogram law, we will prove the following result, which the reader is advised to illustrate geometrically by looking at various examples in $\mathbb{R}^2$.

Theorem 4.1.1. *Let C be a nonempty closed convex subset of a Hilbert space H and let $x \in H$. Then there is a* unique *vector $x_C \in C$ such that $d_H(x, x_C) = d_H(x, C)$, that is, such that*

$$\|x - x_C\| \leq \|x - y\| \quad \textit{for all } y \in C.$$

The vector x_C is called the best approximation to x in C.

Proof. We first consider the case where $x = 0$. We then have to show that there is a unique vector $0_C \in C$ of minimal norm, i.e, which satisfies that

$$\|0_C\| = \inf\Big\{\|y\| : y \in C\Big\}.$$

Set $s := \inf\Big\{\|y\|^2 : y \in C\Big\}$. For each $n \in \mathbb{N}$ we can find $y_n \in C$ such that

$$s \leq \|y_n\|^2 < s + \frac{1}{2n}\,. \tag{4.1.1}$$

Then the sequence $\{y_n\}_{n\in\mathbb{N}}$ is Cauchy in H. Indeed, consider $m, n \in \mathbb{N}$. Then, using the parallellogram law and (4.1.1), we get that

$$\|y_n + y_m\|^2 + \|y_n - y_m\|^2 = 2\,\|y_n\|^2 + 2\,\|y_m\|^2 < 4\,s + \frac{1}{n} + \frac{1}{m}\,.$$

Now, since C is convex, we have $c := \frac{1}{2}\,y_n + \frac{1}{2}\,y_m \in C$. Hence,

$$\|y_n + y_m\|^2 = 4\,\|c\|^2 \geq 4\,s\,,$$

so we get

$$\|y_n - y_m\|^2 < 4\,s + \frac{1}{n} + \frac{1}{m} - \|y_n + y_m\|^2 \leq \frac{1}{n} + \frac{1}{m}\,.$$

Thus, given $\varepsilon > 0$, we can choose $N \in \mathbb{N}$ such that $N \geq (2\varepsilon^2)^{-1}$, and obtain that $\|y_n - y_m\| < \varepsilon$ for all $n, m \geq N$, as desired.

As H is complete, there exists $y_0 \in H$ such that $\lim_n y_n = y_0$. Since C is closed, $y_0 \in C$. Letting $n \to \infty$ in (4.1.1), we get that

$$\|y_0\| = \sqrt{s} = \inf\{\|y\| : y \in C\}.$$

If $y_0' \in C$ also satisfies that $\|y_0'\| = \inf\{\|y\| : y \in C\}$, then we can consider the sequence $\{z_n\}_{n\in\mathbb{N}}$ in C given by $z_n = y_0'$ if n is odd and $z_n = y_0$ if n is even. Since z_n satisfies (4.1.1) (with $y_n = z_n$) for each n, we can conclude as above that $\{z_n\}_{n\in\mathbb{N}}$ is convergent. This clearly implies that $y_0' = y_0$. Thus, y_0 is the unique vector in C satisfying $\|y_0\| = \inf\{\|y\| : y \in C\}$, and we can set $0_C := y_0$.

In the general case where $x \in H$, we note that the set

$$D := \{x - y : y \in C\}$$

is closed and convex. Using the first part, we get that there exists a unique vector $0_D \in D$ such that $\|0_D\| = \inf\{\|z\| : z \in D\} = d(x, C)$. Then $x_C := x - 0_D \in C$ has the desired properties. ■

One important application is when C is a closed subspace M of H.

Theorem 4.1.2. *Let M be a closed subspace of a Hilbert space H. Then we have*

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

The associated projection P_M of H on M along $M^\perp$ is given by

$$P_M(x) = x_M \quad \textit{for all } x \in H ,$$

where $x_M \in M$ is the best approximation to x in M (cf. Theorem 4.1.1). We call P_M the orthogonal projection of H on M *and write sometimes* Proj_M *instead of P_M. The linear map P_M is bounded, with $\|P_M\| = 1$ if $M \neq \{0\}$. Moreover, we have*

$$(M^\perp)^\perp = M \quad \textit{and} \quad P_{M^\perp} = I_H - P_M .$$

Proof. Let $x \in H$ and set $x^\perp := x - x_M$. We claim that $x^\perp$ belongs to $M^\perp$.

To show this, let $y \in M$ and $\varepsilon > 0$. Since $(x_M + \varepsilon\, y) \in M$, we get from Theorem 4.1.1 that

$$\begin{aligned}\|x^\perp\|^2 = \|x - x_M\|^2 \;\leq\; \|x - (x_M + \varepsilon\, y)\|^2 = \|x^\perp - \varepsilon\, y\|^2 \\ = \|x^\perp\|^2 - 2\,\varepsilon\, \mathrm{Re}(\langle x^\perp, y\rangle) + \varepsilon^2\, \|y\|^2 ,\end{aligned}$$

which gives that

$$2\,\mathrm{Re}(\langle x^\perp, y\rangle) \;\le\; \varepsilon\, \|y\|^2 .$$

As this holds for every $\varepsilon > 0$, we obtain that $\mathrm{Re}(\langle x^\perp, y\rangle) \le 0$. Applying this to $-y \in M$, we also get that $-\mathrm{Re}(\langle x^\perp, y\rangle) \le 0$, i.e., $\mathrm{Re}(\langle x^\perp, y\rangle) \ge 0$. Thus, it follows that $\mathrm{Re}(\langle x^\perp, y\rangle) = 0$. If $\mathbb{F} = \mathbb{R}$, this means that $\langle x^\perp, y\rangle = 0$. If $\mathbb{F} = \mathbb{C}$, we also have that $iy \in M$, and this gives that

$$\mathrm{Im}(\langle x^\perp, y\rangle) = \mathrm{Re}(-i\,\langle x^\perp, y\rangle) = \mathrm{Re}(\langle x^\perp, i\,y\rangle) = 0.$$

Thus, $\langle x^\perp, y\rangle = 0$ in this case too. As this holds for every $y \in M$, the claim is proven.

Since $x = x_M + x^\perp$, by definition of $x^\perp$, we get that

$$H = M \;+\; M^\perp .$$

Now, we also have that $M \cap M^\perp = \{0\}$ (because if $y \in M \cap M^\perp$, then $\langle y, y\rangle = 0$, so $y = 0$), while M and $M^\perp$ are both closed in H. Thus, $H = M \oplus M^\perp$, as we wanted to show.

The projection map $P_M : H \to H$ on M along $M^\perp$ is then clearly given by $P_M(x) = x_M$ for $x \in H$. Using Pythagoras' identity, we get that

$$\|P_M(x)\|^2 = \|x_M\|^2 \le \|x_M\|^2 + \|x^\perp\|^2 = \|x_M + x^\perp\|^2 = \|x\|^2$$

for all $x \in H$, showing that $\|P_M\| \le 1$. Since $P_M(y) = y$ whenever $y \in M$, we have that $\|P_M(y)\| = 1$ if $y \in M$ and $\|y\| = 1$. It follows that $\|P_M\| = 1$ if $M \neq \{0\}$, as asserted.

Consider now $y \in M$. Then for all $z \in M^\perp$, we have $\langle y, z\rangle = 0$. This implies that $y \in (M^\perp)^\perp$. Hence we have $M \subseteq (M^\perp)^\perp$.

To show the reverse inclusion, that is $(M^\perp)^\perp \subseteq M$, we first observe that by applying the first part of the theorem to $M^\perp$, we get that

$$H = M^\perp \oplus (M^\perp)^\perp .$$

Now let $x \in (M^\perp)^\perp$, and set again $x^\perp := x - x_M$. Since $x^\perp \in M^\perp$ and $x_M \in M \subseteq (M^\perp)^\perp$, we can write

$$\begin{aligned} x &= x^\perp + x_M\,, \quad \text{where } x^\perp \in M^\perp \text{ and } x_M \in (M^\perp)^\perp, \quad \text{and} \\ x &= \;0 \;+\; x\,, \quad\;\; \text{where } 0 \in M^\perp \text{ and } \;x \in (M^\perp)^\perp. \end{aligned}$$

By the uniqueness of decomposition in a direct sum, we get that $x = x_M$, so $x \in M$. Thus, we have shown that $(M^\perp)^\perp \subseteq M$, and we can conclude that $(M^\perp)^\perp = M$.

Finally, for $x \in H$, we have

$$x = (x - x_M) + x_M \,,$$

where $(x - x_M) \in M^\perp$ and $x_M \in M = (M^\perp)^\perp$. This gives that

$$P_{M^\perp}(x) = x - x_M = (I_H - P_M)(x) \,.$$

Hence, $P_{M^\perp} = I_H - P_M$.

■

Remark 4.1.3. Assume that M is finite-dimensional subspace of a Hilbert space H and that $\mathcal{B} = \{u_1, \ldots, u_n\}$ is an orthonormal basis for M. Then we know that the orthogonal projection P_M of H on M is given by

$$P_M(x) = \sum_{j=1}^{n} \langle x, u_j \rangle \, u_j \quad \text{for all } x \in H \,.$$

A similar formula holds when M is only assumed to be a closed subspace of H, as we will see in the next section after having discussed orthonormal bases in Hilbert spaces.

Corollary 4.1.4. *Let M be closed subspace of a Hilbert space H. Then $M = H$ if and only if $M^\perp = \{0\}$.*

In connection with the next corollary, we recall that if S is a nonempty subset of a vector space V, then Span (S) denote the subspace of V consisting of all possible finite linear combinations of vectors in S.

Corollary 4.1.5. *Let S denote a nonempty subset of a Hilbert space H. Then Span (S) is dense in H if and only if $S^\perp = \{0\}$.*

Proof. Set $M := \overline{\text{Span}\,(S)}$, which is a closed subspace of H. Then Span (S) is dense in H if and only if $M = H$. As one easily verifies that $S^\perp = M^\perp$ (cf. Exercise 4.3), the result follows from Corollary 4.1.4. ■

A nonempty subset S of a normed space X is sometimes called *total in* X when Span (S) is dense in X. So the corollary above says that S is total in H if and only if $S^\perp = \{0\}$.

Example 4.1.6. Let $(X, \mathcal{A}, \mu)$ be a measure space and set $L^2 := L^2(X, \mathcal{A}, \mu)$. We can organize L^2 as a Hilbert space (over $\mathbb{C}$) as follows.

Let $f, g \in \mathcal{L}^2$. Then $\overline{g}$ is measurable (since $\overline{g} = \text{Re}(g) - i\,\text{Im}(g)$) and $\int_X |\overline{g}|^2 \,\mathrm{d}\mu = \int_X |g|^2 \,\mathrm{d}\mu = \|g\|_2^2 < \infty$, so $\overline{g} \in \mathcal{L}^2$. Hence, $f\,\overline{g} \in \mathcal{L}^1$, and we can set

$$\Big\langle [f], [g] \Big\rangle := \int_X f\,\overline{g} \,\mathrm{d}\mu \,.$$

We leave it as an exercise to check that this gives a well-defined inner product on L^2. As the associated norm obviously coincides with $\|\cdot\|_2$, L^2 is complete w.r.t. this norm and we can conclude that L^2 is a Hilbert space.

Now, let $E \in \mathcal{A}$ and set $F := E^c \in \mathcal{A}$. If $g : X \to \mathbb{C}$ is measurable, let us say that g *lives essentially on* E when $\mu\big(\{x \in F : g(x) \neq 0\}\big) = 0$. Then let M_E be the subset of L^2 given by

$$M_E := \Big\{ [g] : g \in \mathcal{L}^2 \text{ and } g \text{ lives essentially on } E \Big\}.$$

Similarly, we can define M_F. We claim that

$$M_F = (M_E)^\perp \quad\text{and}\quad M_E = (M_F)^\perp . \tag{4.1.2}$$

To prove this, assume first that $[g] \in M_E$ and $[h] \in M_F$. Then one easily sees that $g = g\,\mathbf{1}_E$ μ-a.e. and $h = h\,\mathbf{1}_F$ μ-a.e., so, as $E \cap F = \varnothing$, we get

$$\Big\langle [g], [h] \Big\rangle = \int_X g\,\mathbf{1}_E\,\overline{h\,\mathbf{1}_F}\,\mathrm{d}\mu = \int_X g\,\overline{h}\,\mathbf{1}_{E\cap F}\,\mathrm{d}\mu = 0\,.$$

Since this is true for all $[g] \in M_E$, this implies that $[h] \in (M_E)^\perp$. As this holds for all $[h] \in M_F$, we get that $M_F \subseteq (M_E)^\perp$.

To show the reverse inclusion, let $[h] \in (M_E)^\perp$. Then we have

$$\int_X g\,\overline{h}\,\mathrm{d}\mu = 0 \quad \text{whenever } [g] \in M_E\,.$$

In particular, since $[h\,\mathbf{1}_E] \in M_E$, we get

$$\int_X |h|^2\,\mathbf{1}_E\,\mathrm{d}\mu = \int_X (h\,\mathbf{1}_E)\,\overline{h}\,\mathrm{d}\mu = 0\,.$$

Since $|h|^2\,\mathbf{1}_E$ is nonnegative on X, this implies that

$$\mu\Big(\{x \in X : |h(x)|^2\,\mathbf{1}_E(x) \neq 0\}\Big) = 0\,.$$

As $\{x \in E : h(x) \neq 0\} = \{x \in X : |h(x)|^2\,\mathbf{1}_E(x) \neq 0\}$, we get that $\mu\Big(\{x \in E : h(x) \neq 0\}\Big) = 0$, hence that h lives essentially on F. Thus, $[h] \in M_F$. This shows that $(M_E)^\perp \subseteq M_F$.

Altogether, we have proved that $M_F = (M_E)^\perp$. Interchanging E and F, we get that $M_E = (M_F)^\perp$, and the proof of (4.1.2) is finished.

Since the orthogonal complement of any subset is a closed subspace, we can conclude that M_E and M_F are closed subspaces of L^2. Theorem 4.1.2 now gives that

$$L^2 = M_E \oplus (M_E)^\perp = M_E \oplus M_F\,.$$

We note that the fact that $L^2 = M_E \dot{+} M_F$ is a simple consequence of the equation

$$[f] = [f\,\mathbf{1}_E] + [f\,\mathbf{1}_F]\,,\text{ where } [f\,\mathbf{1}_E] \in M_E\,,\ [f\,\mathbf{1}_F] \in M_F,$$

which holds for all $[f] \in L^2$. From this equation, we now see that the orthogonal projection of L^2 on M_E (resp. M_F) is given by

$$P_{M_E}([f]) = [f\,\mathbf{1}_E] \qquad (\text{resp. } P_{M_F}([f]) = [f\,\mathbf{1}_F]\,).$$

4.2 Orthonormal bases in Hilbert spaces

The notion of an orthonormal basis for a finite-dimensional inner product space, which is well-known from elementary linear algebra, have a natural generalization to Hilbert spaces.

Definition 4.2.1. A nonempty subset $\mathcal{B}$ of a Hilbert space H is called an *orthonormal basis* for H when $\mathcal{B}$ is orthonormal and $\text{Span}\,(\mathcal{B})$ is dense in H.

Suppose a Hilbert space H is finite-dimensional (and nonzero). Then an orthonormal set $\mathcal{B}$ in H has to be finite, so $\text{Span}\,(\mathcal{B})$, being finite-dimensional, is closed in H; hence, $\text{Span}\,(\mathcal{B})$ is dense in H if and only if $\text{Span}\,(\mathcal{B}) = H$. Thus we see that Definition 4.2.1 agrees with the usual one when H is finite-dimensional. We also mention that some authors like to define the empty set to be an orthonormal basis for the trivial Hilbert space $H = \{0\}$.

Our first example is of great importance in Fourier analysis.

Example 4.2.2. Let $H = L^2([-\pi,\pi],\mathcal{A},\mu)$, where $\mathcal{A}$ denotes the σ-algebra of all Lebesgue measurable subsets of $[-\pi,\pi]$, and μ is the *normalized* Lebesgue measure on $\mathcal{A}$, that is,

$$\mu(A) := \frac{1}{2\,\pi}\,\lambda(A) \quad \text{for all } A \in \mathcal{A}\,,$$

where λ denotes the Lebesgue measure on $\mathbb{R}$. In particular, we have $\mu([-\pi,\pi]) = 1$. For each $n \in \mathbb{Z}$, let $e_n : [-\pi,\pi] \to \mathbb{C}$ denote the continuous function given by

$$e_n(t) := e^{\,int} \quad \text{for all } t \in [-\pi,\pi]\,.$$

As is probably well-known (and easy to check), the set

$$\mathcal{B} := \{[e_n] : n \in \mathbb{Z}\}$$

is an orthonormal subset of H. Moreover, $\mathrm{Span}\,(\mathcal{B})$ is dense in H.

To show this, let $\mathcal{T}$ denote the space of all (complex) trigonometrical polynomials, i.e., $\mathcal{T} := \mathrm{Span}\,(\{e_n : n \in \mathbb{Z}\})$. Clearly, we have

$$\mathrm{Span}\,(\mathcal{B}) = \{[h] : h \in \mathcal{T}\}\,.$$

Further, let $C_{\mathsf{per}}([-\pi,\pi]) = \{k \in C([-\pi,\pi]) : k(-\pi) = k(\pi)\}$. We will use the fact (shown for example in Lindstrøm's book) that $\mathcal{T}$ is dense in $C_{\mathsf{per}}([-\pi,\pi])$ w.r.t. the uniform norm $\|\cdot\|_u$.
Let $[f] \in H$ and $\varepsilon > 0$. Using Exercise 2.7 we can find $g \in C([-\pi,\pi])$ such that

$$\|\,[f]-[g]\,\|_2 < \varepsilon/3\,. \tag{4.2.1}$$

Moreover, it is easy to see that we can pick $k \in C_{\mathsf{per}}([-\pi,\pi])$ such that

$$\|\,[g]-[k]\,\|_2 = \|\,g-k\,\|_2 < \varepsilon/3\,. \tag{4.2.2}$$

Now, as mentioned above, we can find $h \in \mathcal{T}$ such that $\|k-h\|_u < \varepsilon/3$. Since

$$\begin{aligned}\|\,[k]-[h]\,\|_2^2 &= \int_{[-\pi,\pi]} |k-h|^2 \,\mathrm{d}\mu \\ &\le \|k-h\|_u^2 \int_{[-\pi,\pi]} \mathrm{d}\mu \\ &= \|k-h\|_u^2 \;\mu([-\pi,\pi]) \\ &= \|k-h\|_u^2\,,\end{aligned}$$

we get

$$\|\,[k]-[h]\,\|_2 \le \|k-h\|_u < \varepsilon/3\,. \tag{4.2.3}$$

Using the triangle inequality, (4.2.1), (4.2.2) and (4.2.3), we obtain that

$$\begin{aligned}\|\,[f]-[h]\,\|_2 &= \|\,[f]-[g]+[g]-[k]+[k]-[h]\,\|_2 \\ &\le \|\,[f]-[g]\,\|_2 + \|\,[g]-[k]\,\|_2 + \|\,[k]-[h]\,\|_2 \\ &< \varepsilon/3+\varepsilon/3+\varepsilon/3 = \varepsilon\,.\end{aligned}$$

This shows that $[f] \in \overline{\mathrm{Span}\,(\mathcal{B})}$. Hence, $\overline{\mathrm{Span}\,(\mathcal{B})} = H$, as asserted.

We can now conclude that $\mathcal{B} = \{[e_n] : n \in \mathbb{Z}\}$ is an orthonormal basis for H.

More generally, one may consider the L^2-space associated to an interval $[a,b]$ and the normalized Lebesgue measure $\mu := \frac{1}{b-a}\,\lambda$. Then, letting e'_n be defined for each $n \in \mathbb{Z}$ by

$$e'_n(t) = e^{int\,2\pi/(b-a)} \quad \text{for all } t \in [a,b],$$

one may argue in a similar way as above, and conclude that $\mathcal{B}' = \{e'_n : n \in \mathbb{Z}\}$ is an orthonormal basis for this L^2-space.

An immediate consequence of Corollary 4.1.5 is the following useful characterization of orthonormal bases:

Proposition 4.2.3. *Assume that $\mathcal{B}$ is an orthonormal subset of a Hilbert space H. Then $\mathcal{B}$ is an orthonormal basis for H if and only if $\mathcal{B}^\perp = \{0\}$.*

Example 4.2.4. Let X be a nonempty set. Then $\ell^2(X)$ has a natural orthonormal basis $\mathcal{E}$ which is the analogue of the standard basis $\{e_1, \ldots, e_n\}$ for $\mathbb{F}^n$ (which may be identified with $\ell^2(\{1, \ldots, n\})$).

Indeed, for each $x \in X$, let $e_x \in \ell^2(X)$ be defined by $e_x = \mathbf{1}_{\{x\}}$, and set

$$\mathcal{E} := \{e_x : x \in X\}\,.$$

Then $\mathcal{E}$ is clearly orthonormal. Moreover, let $f \in \ell^2(X)$, $f \in \mathcal{E}^\perp$. Thus, for each $x \in X$, we have $\langle f, e_x \rangle = 0$. As

$$\langle f, e_x \rangle = \sum_{y \in X} f(y)\overline{e_x(y)} = \sum_{y \in \{x\}} f(y) = f(x)\,,$$

we get that $f(x) = 0$ for all $x \in X$, i.e., $f = 0$. This shows that $\mathcal{E}^\perp = \{0\}$, and Proposition 4.2.3 gives that $\mathcal{E}$ is an orthonormal basis for $\ell^2(X)$.

It will be shown in more advanced courses that every Hilbert space (which is non-zero) has an orthonormal basis. The proof is nonconstructive as it relies on Zorn's lemma, i.e., on the axiom of choice. We will take this fact as granted. Of course, in applications, it is better to have at hand a concrete orthonormal basis whenever possible.

Example 4.2.5. The *Gram-Schmidt orthonormalization prosess*, of great importance in the finite-dimensional case, can be generalized to cover the following situation:

Let H be a Hilbert space, $H \neq \{0\}$. Let $\{x_j\}_{j \in \mathbb{N}}$ be a sequence of vectors in $H \setminus \{0\}$ and set $S := \{x_j : j \in \mathbb{N}\}$. Assume that $\mathrm{Span}\,(S)$ is dense in H.

We remark that such a sequence exists whenever H is finite-dimensional (since repetitions are allowed in a sequence). More generally, it exists whenever H is *separable*, i.e., whenever H contains a countable dense subset, cf. Exercise 4.9.

For each $n \in \mathbb{N}$, set $M_n := \mathrm{Span}\,(\{x_1, \ldots, x_n\})$. We note that for each n we have $M_n \subseteq M_{n+1}$. Moreover, $\mathrm{Span}\,(S) = \bigcup_{n \in \mathbb{N}} M_n$.

Proceeding inductively, we can construct an orthonormal basis $\mathcal{B}_n$ for each M_n as follows:

i) We set $\mathcal{B}_1 := \left\{\frac{1}{\|x_1\|}\, x_1\right\}$. Clearly, $\mathcal{B}_1$ is an orthonormal basis for M_1.

ii) Let $n \in \mathbb{N}$ and assume that we have constructed an orthonormal basis $\mathcal{B}_n$ for M_n.

If $x_{n+1} \in M_n$, then set $\mathcal{B}_{n+1} := \mathcal{B}_n$. Otherwise, set

$$y_{n+1} := x_{n+1} - \mathrm{Proj}_{M_n}(x_{n+1}) \quad \text{and} \quad \mathcal{B}_{n+1} := \mathcal{B}_n \cup \left\{\frac{1}{\|y_{n+1}\|}\, y_{n+1}\right\}.$$

It follows readily that $\mathcal{B}_{n+1}$ is an orthonormal basis for M_{n+1}.

Set now $\mathcal{B} := \bigcup_{n\in\mathbb{N}} \mathcal{B}_n$. Then $\mathcal{B}$ is orthonormal, and $\mathrm{Span}\,(\mathcal{B}) = \mathrm{Span}\,(S)$, so

$$\overline{\mathrm{Span}\,(\mathcal{B})} = \overline{\mathrm{Span}\,(S)} = H\,.$$

Hence, $\mathcal{B}$ is an orthonormal basis for H.

We observe that since each $\mathcal{B}_n$ is finite, $\mathcal{B}$ is countable. Conversely, if H has a countable orthonormal basis, then it can be shown that H is separable (cf. Exercise 4.9).

When H is a nontrivial finite-dimensional inner product space, and $\mathcal{B} = \{u_1, \ldots, u_n\}$ is an orthonormal basis for H, we know that every $x \in H$ has a Fourier expansion w.r.t. $\mathcal{B}$, i.e., we have

$$x = \sum_{j=1}^{n} \langle x, u_j\rangle\, u_j\,.$$

As we will soon see, a similar expansion also holds in any infinite dimensional Hilbert space.

We will use the following notation. If $j \to t_j$ is a function from a nonempty set J into $[0, \infty)$, then we set

$$\sum_{j\,\in\, J} t_j := \sup\left\{\sum_{j\,\in\, F} t_j : F \subseteq J,\ F \text{ is finite and nonempty}\right\} \in [0, \infty]\,.$$

Equivalently, $\sum_{j\,\in\, J} t_j$ is the integral of the nonnegative function $j \to t_j$ w.r.t. the counting measure on $\mathcal{P}(J)$ (= the σ-algebra of all subsets of J).

We first note that Bessel's inequality holds for any orthonormal set:

Lemma 4.2.6. *Assume that $\mathcal{B}$ is an orthonormal set in an inner product space H, and let $x \in H$. Then*

$$\sum_{u\,\in\,\mathcal{B}} \left|\langle x, u\rangle\right|^2 \,\le\, \|x\|^2\,,$$

and the set $\mathcal{B}_x := \left\{u \in \mathcal{B} : \langle x, u\rangle \neq 0\right\}$ is countable.

Proof. Let F be a nonempty finite subset of $\mathcal{B}$. As F is orthonormal, Bessel's inequality for F gives that

$$\sum_{u \in F} \left|\langle x, u\rangle\right|^2 \leq \|x\|^2 .$$

Thus we get that

$$\sup\Big\{ \sum_{u \in F} \left|\langle x, u\rangle\right|^2 : F \subseteq \mathcal{B},\ F \text{ is finite and nonempty}\Big\} \leq \|x\|^2,$$

which proves the first assertion.

Further, this implies that the set $\mathcal{B}_{x,n} := \{u \in \mathcal{B} : |\langle x, u\rangle|^2 \geq 1/n\}$ is finite for every $n \in \mathbb{N}$. Hence, $\mathcal{B}_x = \bigcup_{n\in\mathbb{N}} \mathcal{B}_{x,n}$ is countable.

■

The next lemma will be useful at several occasions.

Lemma 4.2.7. *Assume $\{u_j : j \in \mathbb{N}\}$ is a countably infinite orthonormal set of distinct vectors in a Hilbert space H and let $\{c_j\}_{j\in\mathbb{N}}$ be any sequence in $\mathbb{F}$ satisfying that*

$$\sum_{j=1}^{\infty} |c_j|^2 < \infty .$$

Then the series $\sum_{j=1}^{\infty} c_j\, u_j$ converges to some $y \in H$, and we have that

$$\langle y, u_k\rangle = c_k \ \text{ for every } k \in \mathbb{N}.$$

Proof. This result is essentially shown in Lindstrøm's book, but we sketch the argument for the ease of the reader. For each $n \in \mathbb{N}$, set $y_n = \sum_{j=1}^{n} c_j\, u_j$. Then, for any $m > n$, Pythagoras' identity gives that

$$\|y_m - y_n\|^2 = \sum_{j=n+1}^{m} \|c_j\, u_j\|^2 = \sum_{j=n+1}^{m} |c_j|^2.$$

Using the assumption, the sum above can be made as small as we want by choosing m and n large enough. Thus the sequence $\{y_n\}_{n\in\mathbb{N}}$ is Cauchy in H, so it converges to some $y \in H$, i.e., we have

$$y = \sum_{j=1}^{\infty} c_j\, u_j .$$

For each $k \in \mathbb{N}$, continuity and linearity of the inner product in the first variable gives then that

$$\langle y, u_k\rangle = \sum_{j=1}^{\infty} c_j\, \langle u_j, u_k\rangle = c_k .$$

■

Theorem 4.2.8. *Let H be a Hilbert space, $H \neq \{0\}$, and let $\mathcal{B}$ be an orthonormal subset of H. Then the following conditions are equivalent:*

(a) $\mathcal{B}$ is an orthonormal basis for H.

(b) Every $x \in H \setminus \{0\}$ has a Fourier expansion

$$x = \sum_{u \in \mathcal{B}_x} \langle x, u \rangle \, u \tag{4.2.4}$$

where $\mathcal{B}_x = \{u \in \mathcal{B} : \langle x, u \rangle \neq 0\}$ is countable (cf. Lemma 4.2.6) and nonempty.

By (4.2.4) we mean that if $\mathcal{B}_x$ is not finite, and $\mathcal{B}_x = \{u_j : j \in \mathbb{N}\}$ is any enumeration of the distinct elements of $\mathcal{B}_x$, then we have

$$\lim_{n\to\infty} \Big\| x - \sum_{j=1}^{n} \langle x, u_j \rangle \, u_j \Big\| = 0 \,, \text{ i.e.,} \quad x = \sum_{j=1}^{\infty} \langle x, u_j \rangle \, u_j \,.$$

(c) For every $x \in H$ we have $\|x\|^2 = \sum_{u \in \mathcal{B}} \big|\langle x, u \rangle\big|^2$.

The formula in (c) is called Parseval's identity.

Proof. $(a) \Rightarrow (b)$: Assume that $\mathcal{B}$ is an orthonormal basis for H and let $x \in H \setminus \{0\}$.

We first observe that $\mathcal{B}_x \neq \varnothing$. Indeed, suppose that $\mathcal{B}_x = \varnothing$. This means that $x \in \mathcal{B}^\perp$. But $\mathcal{B}^\perp = \{0\}$ by Proposition 4.2.3, so $x = 0$, a contradiction.

We now consider the case where $\mathcal{B}_x$ is countably infinite. (The case where $\mathcal{B}_x$ is finite is much easier and left to the reader). Let $\{u_j : j \in \mathbb{N}\}$ be an enumeration of the distinct elements of $\mathcal{B}_x$. Since $\mathcal{B}_x$ is orthonormal, Bessel's inequality gives that

$$\sum_{j=1}^{\infty} \big|\langle x, u_j \rangle|^2 \, \leq \, \|x\|^2.$$

Applying Lemma 4.2.7 with $c_j = \langle x, u_j \rangle$ for every $j \in \mathbb{N}$, we get that the series $\sum_{j=1}^{\infty} \langle x, u_j \rangle \, u_j$ converges to some $y \in H$, which satisfies that

$$\langle y, u_k \rangle = c_k = \langle x, u_k \rangle \quad \text{for every } k \in \mathbb{N}.$$

Moreover, if $u \in \mathcal{B} \setminus \mathcal{B}_x$, we get that

$$\langle y, u \rangle = \sum_{j=1}^{\infty} \langle x, u_j \rangle \, \langle u_j, u \rangle = 0 = \langle x, u \rangle \,.$$

It follows that $x - y \in \mathcal{B}^\perp = \{0\}$, hence that $x = y$. This shows that the assertion in (b) holds in this case.

$(b) \Rightarrow (c)$: Assume (b) holds, and let $x \in H \setminus \{0\}$. Again we consider the more difficult case where $\mathcal{B}_x$ is countably infinite, so $\mathcal{B}_x = \{u_j : j \in \mathbb{N}\}$ as above. By continuity of the norm and Pythagoras' identity, we get

$$\|x\|^2 = \sum_{j=1}^{\infty} \left|\langle x, u_j\rangle\right|^2 .$$

Hence, given $\varepsilon > 0$, we can find $n \in \mathbb{N}$ such that $\|x\|^2 - \sum_{j=1}^{n} \left|\langle x, u_j\rangle\right|^2 < \varepsilon$, giving

$$\|x\|^2 - \varepsilon \; < \; \sum_{j=1}^{n} \left|\langle x, u_j\rangle\right|^2 \; \leq \; \sum_{u \in \mathcal{B}} \left|\langle x, u\rangle\right|^2 .$$

Since this holds for every $\varepsilon > 0$, we get that $\|x\|^2 \leq \sum_{u \in \mathcal{B}} \left|\langle x, u\rangle\right|^2$. Combining this inequality with Lemma 4.2.6, we see that (c) holds.

$(c) \Rightarrow (a)$: Assume $\|x\|^2 = \sum_{u \in \mathcal{B}} \left|\langle x, u\rangle\right|^2$ for every $x \in H$. If $x \in \mathcal{B}^{\perp}$, i.e., $\langle x, u\rangle = 0$ for every $u \in \mathcal{B}$, then we get $\|x\|^2 = 0$, so $x = 0$. Hence, $\mathcal{B}^{\perp} = \{0\}$, and Proposition 4.2.3 gives that (a) holds.

■

Remark 4.2.9. The Fourier expansion of x in Theorem 4.2.8 (b) can be written in the form

$$x = \sum_{u \in \mathcal{B}} \langle x, u\rangle \, u \tag{4.2.5}$$

if one takes care of giving a meaning to convergence of generalized sums in normed spaces. We discuss this in Exercise 4.12. In these notes, we will sometimes use (4.2.5) as a short form of the Fourier expansion of x given by (4.2.4).

Example 4.2.10. Let M be a closed subspace of a Hilbert space H and assume that we have found an orthonormal basis $\mathcal{C}$ for M. Then we can use it to compute the orthogonal projection P_M of H on M:

Let $x \in H$. If $x \in M^{\perp}$, then $P_M(x) = 0$, so we can assume $x \in H \setminus M^{\perp}$. Since $\mathcal{C}$ is orthonormal in H, we know that $\mathcal{C}_x := \{v \in \mathcal{C} : \langle x, v\rangle \neq 0\}$ is countable. Set $x_M := P_M(x) \in M$ and $x^{\perp} := x - x_M \in M^{\perp}$, and note that $x_M \neq 0$. Now, for each $v \in \mathcal{C}$, we have

$$\langle x, v\rangle = \langle x_M, v\rangle + \langle x^{\perp}, v\rangle = \langle x_M, v\rangle .$$

Hence, $\mathcal{C}_x = \mathcal{C}_{x_M}$. Moreover, applying Theorem 4.2.8 to M, $x_M \in M \setminus \{0\}$ and $\mathcal{C}$, we get that

$$x_M = \sum_{v \in \mathcal{C}_{x_M}} \langle x_M, v\rangle \, v .$$

Using our previous observations, this formula can be rewritten as

$$P_M(x) = \sum_{v \in \mathcal{C}_x} \langle x, v \rangle \, v \,,$$

generalizing the usual formula for $P_M(x)$ when M is finite-dimensional.

A straightforward consequence of Theorem 4.2.8 is the following:

Corollary 4.2.11. *Assume a Hilbert space H contains a countably infinite orthonormal subset $\mathcal{B}$, enumerated as $\mathcal{B} = \{v_k : k \in \mathbb{N}\}$. Then $\mathcal{B}$ is an orthonormal basis for H if and only if*

$$x = \sum_{k=1}^{\infty} \langle x, v_k \rangle \, v_k$$

for all $x \in H$, if and only if

$$\|x\|^2 = \sum_{k=1}^{\infty} \Big| \langle x, v_k \rangle \Big|^2$$

for all $x \in H$.

Example 4.2.12. Let $\mathcal{B} = \{[e_n] : n \in \mathbb{Z}\}$ denote the orthonormal basis for $H = L^2([-\pi, \pi], \mathcal{A}, \lambda/2\pi)$ described in Example 4.2.2. For $[f] \in H$ and $n \in \mathbb{Z}$ it is common to set

$$\widehat{[f]}\,(n) := \langle\, [f], [e_n]\, \rangle = \frac{1}{2\pi} \int_{[-\pi,\pi]} f(t) e^{-int} d\lambda(t) \,,$$

which is called the *Fourier coefficient* of $[f]$ at n.

In fact, it is usual to write f instead of $[f]$, having in mind that one identifies functions which agree μ-a.e. Hence, the Fourier coefficient of f at n is denoted by $\widehat{f}(n)$, and the Fourier expansion of f w.r.t. $\mathcal{B}$ is then written as

$$f = \sum_{n \in \mathbb{Z}} \widehat{f}(n) \, e_n \,,$$

meaning that

$$f = \lim_{m \to \infty} \sum_{n=-m}^{m} \widehat{f}(n) \, e_n \quad (\text{w.r.t. } \| \cdot \|_2).$$

This follows from Corollary 4.2.11 by enumerating $\mathcal{B}$ as e_0, e_{-1}, e_1, e_{-2}, e_2, etc. Similarly, we have

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} \Big| \widehat{f}(n) \Big|^2 .$$

4.3 Adjoint operators

Let X be a normed space (over $\mathbb{F}$). We recall that the dual space X^* consists of the bounded linear functionals on X (with values in $\mathbb{F}$), and that X^* is a Banach space w.r.t. the norm $\|\varphi\| = \sup\{|\varphi(x)| : x \in X_1\}$. The goal of functional analysis is to gain new insight by exploiting the interplay between a space and its dual. This is particularly successful when X is a Hilbert space because the dual space may then be identified in a natural way with the space itself.

Theorem 4.3.1. *Let H be a Hilbert space (over $\mathbb{F}$). For each $y \in H$, define $\varphi_y : H \to \mathbb{F}$ by*

$$\varphi_y(x) := \langle x, y \rangle \quad \text{for all } x \in H\,.$$

Then $\varphi_y \in H^*$ for all $y \in H$.

Moreover, the map $y \to \varphi_y$ is a bijection from H onto H^*, which is isometric, and conjugate-linear in the sense that

$$\varphi_{\lambda_1 y_1 + \lambda_2 y_2} = \overline{\lambda_1}\,\varphi_{y_1} + \overline{\lambda_2}\,\varphi_{y_2}$$

for all $\lambda_1, \lambda_2 \in \mathbb{F}$ and all $y_1, y_2 \in H$.

Proof. Let $y \in H$. Then the map φ_y is clearly linear. Moreover, for all $x \in H$, we have

$$|\varphi_y(x)| = |\langle x, y\rangle| \;\le\; \|x\|\,\|y\|\,.$$

Hence, φ_y is bounded, with $\|\varphi_y\| \le \|y\|$. If $y \neq 0$, then

$$\Big|\,\varphi_y\Big(\frac{1}{\|y\|}\,y\Big)\,\Big| = \frac{1}{\|y\|}\,\langle y, y\rangle = \|y\|,$$

so $\|\varphi_y\| \;\ge\; \|y\|$. Thus, $\|\varphi_y\| = \|y\|$.

This shows that the map $y \to \varphi_y$ is an isometry from H into H^*. In particular, it is injective. To show that it is surjective, let $\varphi \in H^*$. If $\varphi = 0$, then we have $\varphi = \varphi_0$. So assume $\varphi \neq 0$ and set $M := \ker\varphi$. Then M is a closed subspace of H such that $M \neq H$. By Corollary 4.1.4, $M^\perp \neq \{0\}$, so we can pick $z \in M^\perp$ such that $\|z\| = 1$, and set

$$y := \overline{\varphi(z)}\, z \in H\,.$$

We claim that $\varphi = \varphi_y$. Indeed, let $x \in H$ and set $m := \varphi(x)\,z - \varphi(z)\,x \in H$. Then we have

$$\varphi(m) = \varphi(x)\,\varphi(z) - \varphi(z)\,\varphi(x) = 0\,,$$

so $m \in M$. As $z \in M^\perp$, we get $\langle m, z\rangle = 0$, i.e.,

$$\Big\langle \varphi(x)\, z, z \Big\rangle = \Big\langle \varphi(z)\, x, z \Big\rangle .$$

Hence

$$\varphi(x) = \varphi(x)\, \|z\|^2 = \langle \varphi(x)\, z, z\rangle = \langle \varphi(z)\, x, z\rangle = \varphi(z)\, \langle x, z\rangle$$
$$= \langle x, \overline{\varphi(z)}\, z\rangle = \langle x, y\rangle = \varphi_y(x)\,,$$

which shows the claim above, hence that the map $y \to \varphi_y$ is surjective.

Altogether, we have shown that this map is an isometric bijection from H onto H^*, as desired.

The final assertion is an obvious consequence of the conjugate-linearity of the inner product in the second variable. ∎

This theorem, which is one among a diversity of results being called *the Riesz representation theorem*, has several useful consequences that will be covered in later courses. Our main application here will be to use it to associate an adjoint operator to every bounded operator on a Hilbert space. Some people like to think of the adjoint as a kind of twin (or as a kind of shadow), which happens to coincide with the original operator in many cases of interest.

Theorem 4.3.2. *Let H be a Hilbert space (over $\mathbb{F}$). For each $T \in \mathcal{B}(H)$, there is a unique operator $T^* \in \mathcal{B}(H)$, called the* adjoint of T, *satisfying*

$$\Big\langle T(x), y \Big\rangle = \Big\langle x, T^*(y) \Big\rangle \tag{4.3.1}$$

for all $x, y \in H$.

The $*$-operation on $\mathcal{B}(H)$, $T \to T^*$, enjoys the following properties:
For all $S, T \in \mathcal{B}(H)$ and all $\alpha, \beta \in \mathbb{F}$, we have

- *i)* $(\alpha\, S + \beta\, T)^* = \overline{\alpha}\, S^* + \overline{\beta}\, T^*$; *ii)* $(ST)^* = T^* S^*$; *iii)* $(T^*)^* = T$;
- *iv)* $\|T^*\| = \|T\|$; *v)* $\|T^*T\| = \|T\|^2$.

Remark 4.3.3. If H and K are Hilbert spaces (over the same $\mathbb{F}$), then one may associate to each $T \in \mathcal{B}(H, K)$ a unique adjoint operator $T^* \in \mathcal{B}(K, H)$ satisfying (4.3.1) for all $x \in H$ and all $y \in K$, and enjoying similar properties. We leave this as an exercise.

Proof of Theorem 4.3.2. Let $T \in \mathcal{B}(H)$ and consider $y \in H$. Using the linearity of T and the linearity of the inner product in the first variable, we get that the map $\varphi : H \to \mathbb{F}$ defined by

$$\varphi(x) := \Big\langle T(x), y \Big\rangle \quad \text{for all } x \in H\,,$$

is a linear functional on H. Moreover, as we have

$$\left|\left\langle T(x), y\right\rangle\right| \leq \|T(x)\|\,\|y\| \leq \|T\|\,\|x\|\,\|y\|$$

for all $x \in H$, φ is bounded with $\|\varphi\| \leq \|T\|\,\|y\|$. Hence, $\varphi \in H^*$, and Theorem 4.3.1 gives that there exists a unique vector in H, that we denote by $T^*(y)$, such that $\varphi = \varphi_{T^*(y)}$, i.e., such that

$$\left\langle T(x), y\right\rangle = \left\langle x, T^*(y)\right\rangle \tag{4.3.2}$$

for all $x \in H$. This theorem also gives that

$$\|T^*(y)\| = \|\varphi_{T^*(y)}\| = \|\varphi\| \leq \|T\|\,\|y\|\,. \tag{4.3.3}$$

As what we have done above holds for every $y \in H$, we obtain a map $T^* : H \to H$ which sends each $y \in H$ to $T^*(y) \in H$. In view of (4.3.2), it is clear that (4.3.1) holds for all $x, y \in H$.

To show that T^* is linear, let $y, y' \in H$ and $\alpha \in \mathbb{F}$. Then, for all $x \in H$, we have

$$\begin{aligned}\left\langle x, T^*(\alpha\, y + y')\right\rangle &= \left\langle T(x), \alpha\, y + y'\right\rangle \\ &= \overline{\alpha}\left\langle T(x), y\right\rangle + \left\langle T(x), y'\right\rangle \\ &= \overline{\alpha}\left\langle x, T^*(y)\right\rangle + \left\langle x, T^*(y')\right\rangle \\ &= \left\langle x, \alpha\, T^*(y) + T^*(y')\right\rangle.\end{aligned}$$

This implies that $T^*(\alpha\, y + y') = \alpha\, T^*(y) + T^*(y')$, as desired.

Next, from (4.3.3), we see that T^* is bounded with $\|T^*\| \leq \|T\|$. To show the asserted uniqueness property of T^*, assume that $S \in \mathcal{B}(H)$ satisfies the same property as T^*, i.e.,

$$\left\langle T(x), y\right\rangle = \left\langle x, S(y)\right\rangle \quad \text{for all } x, y \in H\,.$$

Let $y \in H$. Then, for all $x \in H$, we get

$$\left\langle x, S(y)\right\rangle = \left\langle T(x), y\right\rangle = \left\langle x, T^*(y)\right\rangle.$$

This implies that $S(y) = T^*(y)$. Thus, $S = T^*$.

We leave the proof of properties $i)$ and $ii)$ as an exercise. To show the other properties, let $T \in \mathcal{B}(H)$. Then, for each $y \in H$, using equation (4.3.1) for T^* instead of T, we get that, for all $x \in H$, we have

$$\begin{aligned}\left\langle x, (T^*)^*(y)\right\rangle &= \left\langle T^*(x), y\right\rangle = \overline{\left\langle y, T^*(x)\right\rangle} \\ &= \overline{\left\langle T(y), x\right\rangle} = \left\langle x, T(y)\right\rangle\end{aligned}$$

This implies that $(T^*)^*(y) = T(y)$. Thus, $(T^*)^* = T$, i.e., *iii*) holds

Now, we have seen that $\|T^*\| \leq \|T\|$ holds for *all* $T \in \mathcal{B}(H)$. Thus we get

$$\|T\| = \|(T^*)^*\| \leq \|T^*\| \leq \|T\|.$$

Hence $\|T^*\| = \|T\|$, i.e., *iv*) holds.

Further, using *iv*), we get $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$. On the other hand, for every $x \in H$, we have

$$\begin{aligned}\|T(x)\|^2 &= \Big\langle T(x), T(x)\Big\rangle = \Big\langle x, T^*(T(x))\Big\rangle \\ &= \Big|\Big\langle x, (T^*T)(x)\Big\rangle\Big| \leq \|x\| \, \|(T^*T)(x)\| \\ &\leq \|T^*T\| \, \|x\|^2 .\end{aligned}$$

This implies that $\|T\|^2 \leq \|T^*T\|$. Hence we get $\|T\|^2 = \|T^*T\|$, i.e., *v*) holds. ■

Example 4.3.4. Consider $H = \mathbb{F}^n$ for some $n \in \mathbb{N}$ with its usual inner product, and $T \in \mathcal{B}(H)$. Let A denote the standard matrix of T. Then the standard matrix of T^* is $A^* := \overline{A}^t$.

Here, $\overline{A}$ denotes the matrix obtained by conjugating every coefficient of A, while B^t denotes the transpose of a matrix B. Of course, if $\mathbb{F} = \mathbb{R}$, then we just get $A^* = A^t$.

Alternatively, we can formulate our assertion above by saying that if $T_A \in \mathcal{B}(H)$ denotes the operator given by multiplication with a matrix $A \in M_n(\mathbb{F})$, then we have

$$(T_A)^* = T_{A^*} .$$

To prove this, let $x, y \in H$. Recall that $\langle x, y\rangle = x^t\,\overline{y}$. So we get

$$\langle T_A(x), y\rangle = (Ax)^t\,\overline{y} = x^t A^t\,\overline{y} = x^t\,\overline{A^*y} = \langle x, T_{A^*}(y)\rangle.$$

Since this holds for all $x, y \in H$, this implies that $(T_A)^* = T_{A^*}$, as asserted.

More generally, if H is a nontrivial finite-dimensional Hilbert space, $\mathcal{B}$ is an orthonormal basis for H, and $[T]_{\mathcal{B}}$ is the matrix of T relative to $\mathcal{B}$, then we have

$$[T^*]_{\mathcal{B}} = \Big([T]_{\mathcal{B}}\Big)^* .$$

The verification of this fact is an easy exercise. (One may argue in a similar way as in the next example).

Example 4.3.5. Assume a Hilbert space H has a countably infinite orthonormal basis, enumerated as $\mathcal{B} = \{u_j\}_{j\in\mathbb{N}}$. Let $T \in \mathcal{B}(H)$. For each $(j,k) \in \mathbb{N}\times\mathbb{N}$, set

$$A(j,k) := \Big\langle T(u_k), u_j\Big\rangle \in \mathbb{F}.$$

We may think of the map A sending each (j,k) to $A(j,k)$ as the (infinite) matrix of T (w.r.t. $\mathcal{B}$) since, for each $k \in \mathbb{N}$, we have

$$T(u_k) = \sum_{j=1}^{\infty} \Big\langle T(u_k), u_j \Big\rangle \, u_j = \sum_{j=1}^{\infty} A(j,k)\, u_j \,. \tag{4.3.4}$$

Now, as every $x \in H$ has a Fourier expansion w.r.t. $\mathcal{B}$, it is clear that T is uniquely determined as a bounded operator on H by its values on $\mathcal{B}$. Thus we see from (4.3.4) that T is uniquely determined by its matrix A.

As $T^* \in \mathcal{B}(H)$ and

$$\Big\langle T^*(u_k), u_j \Big\rangle = \Big\langle u_k, T(u_j) \Big\rangle = \overline{\Big\langle T(u_j), u_k \Big\rangle} = \overline{A(k,j)},$$

we can conclude that the matrix of T^* w.r.t. $\mathcal{B}$ is A^*, where

$$A^*(j,k) := \overline{A(k,j)}\,,$$

so that, for all $k \in \mathbb{N}$, we have

$$T^*(u_k) = \sum_{j=1}^{\infty} A^*(j,k)\, u_j = \sum_{j=1}^{\infty} \overline{A(k,j)}\, u_j \,.$$

From (4.3.4) and Parseval's identity, we also get that

$$\sum_{j=1}^{\infty} |A(j,k)|^2 = \|T(u_k)\|^2 \le \|T\|^2 < \infty$$

for each $k \in \mathbb{N}$, so the ℓ^2-norms of the column vectors of A are uniformly bounded. However, such a condition on the column vectors of an infinite matrix A is not sufficient in general to ensure that A is the matrix of some operator in $\mathcal{B}(H)$. There are some known conditions guaranteeing this, but we will only look at two cases below where one can argue directly.

a) Let $\{\lambda_j\}_{j\in\mathbb{N}}$ be a bounded sequence in $\mathbb{F}$, so that

$$M := \sup\{|\lambda_j| : j \in \mathbb{N}\} < \infty.$$

In other words, the function $j \to \lambda_j$ belongs to $\ell^\infty(\mathbb{N}, \mathbb{F})$.

It is not difficult to see that there exists an operator $D \in \mathcal{B}(H)$ satisfying that

$$D(u_k) = \lambda_k\, u_k \quad \text{for each } k \in \mathbb{N}. \tag{4.3.5}$$

Indeed, consider $x \in H$. Then Parseval's identity gives that

$$\sum_{j=1}^{\infty} \Big|\lambda_j \langle x, u_j \rangle\Big|^2 \;\le\; M^2 \sum_{j=1}^{\infty} \Big|\langle x, u_j \rangle\Big|^2 = M^2\, \|x\|^2 \;<\; \infty.$$

Hence, Lemma 4.2.7 gives that the vector

$$D(x) := \sum_{j=1}^{\infty} \lambda_j \langle x, u_j \rangle u_j$$

satisfies that $\langle D(x), u_j \rangle = \lambda_j \langle x, u_j \rangle$ for each $j \in \mathbb{N}$. Thus, using Parseval's identity again, we get that

$$\|D(x)\|^2 = \sum_{j=1}^{\infty} \Big| \lambda_j \langle x, u_j \rangle \Big|^2 \leq M^2 \|x\|^2 .$$

It follows now readily that the map $x \to D(x)$ gives an operator $D \in \mathcal{B}(H)$ such that $\|D\| \leq M$ and satisfying (4.3.5). Since $\|D\| \geq \|D(u_k)\| = |\lambda_k|$ for all $k \in \mathbb{N}$, we also have that $\|D\| \geq M$. Hence, $\|D\| = M$.

It is now obvious that the matrix of D (w.r.t. $\mathcal{B}$) is the diagonal (infinite) matrix Λ defined for each $(j, k) \in \mathbb{N}$ by

$$\Lambda(j,k) = \begin{cases} \lambda_j & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

The operator D is often called *the diagonal operator associated to* $\{\lambda_j\}_{j \in \mathbb{N}}$ *(w.r.t.* $\mathcal{B}$*)*.

From our discussion in the first part, we get that the matrix of D^* is Λ^*. Thus we have $D^*(u_k) = \overline{\lambda_k}\, u_k$ for all $k \in \mathbb{N}$, so D^* is the diagonal operator associated to $\{\, \overline{\lambda_j}\, \}_{j \in \mathbb{N}}$ (w.r.t. $\mathcal{B}$).

b) We may also easily argue that there exists an operator $S \in \mathcal{B}(H)$ satisfying that

$$S(u_k) = u_{k+1} \quad \text{for all } k \in \mathbb{N}. \tag{4.3.6}$$

Indeed, since $\sum_{n=2}^{\infty} |\langle x, u_{n-1} \rangle|^2 = \sum_{j=1}^{\infty} |\langle x, u_j \rangle|^2 = \|x\|^2 < \infty$ for all $x \in H$, we may use Lemma 4.2.7 to define a map $S : H \to H$ by

$$S(x) = \sum_{n=2}^{\infty} \langle x, u_{n-1} \rangle u_n$$

which is then a linear isometry satisfying that $S(u_{n-1}) = u_n$ for all $n \geq 2$, i.e., such that (4.3.6) holds. The map S is called *the right shift operator on* H*(w.r.t.* $\mathcal{B}$*)*. The matrix of S (w.r.t. $\mathcal{B}$) is the (infinite) matrix σ given by

$$\sigma(j,k) = \begin{cases} 1 & \text{if } j = k+1, \\ 0 & \text{otherwise.} \end{cases}$$

for each $(j,k) \in \mathbb{N}$. Thus, the matrix of S^*(w.r.t. $\mathcal{B}$) is the matrix σ^* given by

$$\sigma^*(j,k) = \overline{\sigma(k,j)} = \begin{cases} 1 & \text{if } k = j+1, \\ 0 & \text{otherwise.} \end{cases}$$

for all $j,k \in \mathbb{N}$, so we get that

$$S^*(u_k) = \sum_{j=1}^{\infty} \sigma^*(j,k)\, u_j = \begin{cases} 0 & \text{if } k = 1, \\ u_{k-1} & \text{if } k \geq 2. \end{cases}$$

The operator S^* is called *the left shift operator on* H *(w.r.t.* $\mathcal{B}$*)*. We note that S^* is not isometric, in fact not even injective, since $S^*(u_1) = 0$.

Example 4.3.6. Let $(X, \mathcal{A}, \mu)$ be a measure space. Set $\mathcal{L}^\infty := \mathcal{L}^\infty(X, \mathcal{A}, \mu)$ and $H := L^2(X, \mathcal{A}, \mu)$. For each $f \in \mathcal{L}^\infty$, we may define a "multiplication" operator $M_f \in \mathcal{B}(H)$ by

$$M_f\big([g]\big) = [fg] \quad \text{for all } [g] \in H\,.$$

Indeed, this follows readily from Proposition 2.2.4 (with $q = 2$). Now, for all $[g], [h] \in H$, we have

$$\Big\langle M_f([g]), [h] \Big\rangle = \int_X fg\,\overline{h}\;\mathrm{d}\mu = \int_X g\,\overline{\overline{f}\,h}\;\mathrm{d}\mu = \Big\langle [g], M_{\overline{f}}([h]) \Big\rangle\,.$$

This implies that $(M_f)^* = M_{\overline{f}}$.

Example 4.3.7. Let $K : [a,b] \times [a,b] \to \mathbb{C}$ be a continuous function and let T_K denote the associated integral operator on $H = L^2([a,b])$, cf. Example 3.3.6 and Exercise 3.18. Then we leave it as an exercise to check that $(T_K)^* = T_{K^*}$, where $K^*(s,t) := \overline{K(t,s)}$ for all $s,t \in [a,b]$.

As an illustration that the adjoint operator contains valuable information about the original operator, we include a proposition showing the connection between the fundamental subspaces associated to these operators.

Proposition 4.3.8. *Let* H *be a Hilbert space (over* $\mathbb{F}$*) and let* $T \in \mathcal{B}(H)$*. Then we have:*

(*a*) $\ker(T) = T^*(H)^\perp$ *and* $\ker(T^*) = T(H)^\perp$.

(*b*) $\overline{T(H)} = \ker(T^*)^\perp$ *and* $\overline{T^*(H)} = \ker(T)^\perp$.

Proof. Both equalities in (a) are immediate consequences of (4.3.1). Using Exercise 4.3 with $N = T(H)$, we then get $\overline{T(H)} = (T(H)^\perp)^\perp = \ker(T^*)^\perp$. The second equality in (b) is shown similarly (or by replacing T with T^* in the first one). ■

Corollary 4.3.9. *Let H be a Hilbert space (over $\mathbb{F}$) and let $T \in \mathcal{B}(H)$. Then the image of T is dense in H if and only if T^* is injective (i.e., is one-to one).*

Proof. Using Proposition 4.3.8 and Corollary 4.1.4, we get

$$\overline{T(H)} = H \Leftrightarrow \ker(T^*)^\perp = H \Leftrightarrow \ker(T^*) = \{0\}\,.$$

As T^* is linear, we also have $\ker(T^*) = \{0\} \Leftrightarrow T^*$ is injective. ■

As another illustration, we also mention:

Proposition 4.3.10. *Let H be a Hilbert space (over $\mathbb{F}$) and let $T \in \mathcal{B}(H)$. Then T is invertible in $\mathcal{B}(H)$ if and only if T^* is invertible in $\mathcal{B}(H)$, in which case we have $(T^*)^{-1} = (T^{-1})^*$.*

Proof. Left to the reader as Exercise 4.17. ■

4.4 Self-adjoint operators

In this section, we introduce one of the most important classes of bounded operators on a Hilbert space and discuss some of their properties.

Definition 4.4.1. Let H be a Hilbert space (over $\mathbb{F}$). An operator $T \in \mathcal{B}(H)$ is called *self-adjoint* when $T^* = T$, that is, we have

$$\Big\langle T(x), y \Big\rangle = \Big\langle x, T(y) \Big\rangle \quad \text{for all } x, y \in H\,.$$

If $\mathbb{F} = \mathbb{C}$, a self-adjoint operator in $\mathcal{B}(H)$ is also called *Hermitian*, while it is often called *symmetric* if $\mathbb{F} = \mathbb{R}$.

We note that if $T, T' \in \mathcal{B}(H)$ are self-adjoint, and $\lambda \in \mathbb{R}$, then it is obvious that $\lambda T + T'$ is also self-adjoint.

Example 4.4.2. Let $A = [a_{j,k}] \in M_n(\mathbb{F})$ and let $T_A \in \mathcal{B}(\mathbb{F}^n)$ denote the operator given by multiplication with A (cf. Example 4.3.4). Then T_A is self-adjoint if and only if $A^* = A$, i.e., $\overline{a_{k,j}} = a_{j,k}$ for all $j, k \in \{1, \ldots, n\}$. In particular, when $\mathbb{F} = \mathbb{R}$, T_A is self-adjoint if and only if A is symmetric.

Example 4.4.3. Assume H is a Hilbert space with a countably infinite orthonormal basis $\mathcal{B}$. Let D denote a diagonal operator associated to a bounded sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ in $\mathbb{F}$ (w.r.t. $\mathcal{B}$), as in Example 4.3.5.

Then D is self-adjoint if and only if $\overline{\lambda_j} = \lambda_j$ for all $j \in \mathbb{N}$, i.e., $\lambda_j \in \mathbb{R}$ for all $j \in \mathbb{N}$. In particular, D is always self-adjoint when $\mathbb{F} = \mathbb{R}$.

Let S denote the right shift operator on H (w.r.t. $\mathcal{B}$). Then S^* is the left shift operator and it is obvious that $S^* \neq S$. So S is not self-adjoint.

Example 4.4.4. Let $(X, \mathcal{A}, \mu)$ be a measure space and set $H := L^2(X, \mathcal{A}, \mu)$. If $f \in \mathcal{L}^\infty$, then the multiplication operator $M_f \in \mathcal{B}(H)$ defined in Example 4.3.6 is self-adjoint if and only if $M_{\overline{f}} = M_f$.

Thus, M_f is self-adjoint whenever f is real-valued (μ-a.e.). It can be shown that the converse statement holds whenever $(X, \mathcal{A}, \mu)$ satisfies the mild assumption that it is semifinite (cf. Exercise 4.22).

Example 4.4.5. Let $K : [a, b] \times [a, b] \to \mathbb{C}$ be a continuous function and let T_K denote the associated integral operator on $H = L^2([a, b])$, cf. Example 4.3.7. Then T_K is self-adjoint if and only if $T_{K^*} = T_K$ (where $K^*(s, t) = \overline{K(t, s)}$). Hence it is clear that T_K is self-adjoint whenever K is real-valued. We leave it as an exercise to check that the converse statement also holds.

Example 4.4.6. Let M be a closed subspace of a Hilbert space H and let P_M denote the ortogonal projection of H on M. Then P_M is self-adjoint.

Indeed, let $x, y \in H$. As $P_M(x) \in M$ and $y - P_M(y) \in M^\perp$, we have

$$\Big\langle P_M(x), y - P_M(y) \Big\rangle = 0\,.$$

Similarly, we have $\Big\langle x - P_M(x), P_M(y) \Big\rangle = 0$. Hence we get

$$\begin{aligned}
\Big\langle P_M(x), y \Big\rangle &= \Big\langle P_M(x), P_M(y) + y - P_M(y) \Big\rangle \\
&= \Big\langle P_M(x), P_M(y) \Big\rangle + \Big\langle P_M(x), y - P_M(y) \Big\rangle \\
&= \Big\langle P_M(x), P_M(y) \Big\rangle \\
&= \Big\langle P_M(x), P_M(y) \Big\rangle + \Big\langle x - P_M(x), P_M(y) \Big\rangle \\
&= \Big\langle P_M(x) + x - P_M(x), P_M(y) \Big\rangle \\
&= \Big\langle x, P_M(y) \Big\rangle
\end{aligned}$$

It is easy to create self-adjoint operators.

Proposition 4.4.7. *Let H be a Hilbert space (over $\mathbb{F}$) and $T \in \mathcal{B}(H)$.*

Then $T + T^*$, T^*T and TT^* are all self-adjoint. Moreover, if $\mathbb{F} = \mathbb{C}$, then $-i\,(T - T^*)$ is also self-adjoint.

Proof. The reader should have no difficulty to verify these assertions by using the properties of the $*$-operation on $\mathcal{B}(H)$ listed in Theorem 4.3.2. ■

A noteworthy consequence is that bounded self-adjoint operators on a *complex* Hilbert space have a canonical decomposition similar to the one enjoyed by complex numbers.

Corollary 4.4.8. *Let H be a Hilbert space over $\mathbb{C}$ and let $T \in \mathcal{B}(H)$. Set*

$$\mathrm{Re}(T) := \frac{1}{2}\Big(T + T^*\Big)\,, \quad \mathrm{Im}(T) := \frac{1}{2\,i}\Big(T - T^*\Big)\,.$$

Then $\mathrm{Re}(T)$ *and* $\mathrm{Im}(T)$ *are both self-adjoint, and we have*

$$T = \mathrm{Re}(T) + i\ \mathrm{Im}(T)\,.$$

Proof. The first assertion follows readily from Proposition 4.4.7. The second one is elementary. ■

Consider a bounded operator T on a Hilbert space $H \neq \{0\}$. The *numerical range* of T is defined as the subset of $\mathbb{F}$ given by

$$W_T := \Big\{ \Big\langle T(x), x \Big\rangle : x \in H, \|x\| = 1 \Big\}.$$

Some properties of T are reflected in the geometric properties of W_T and of its closure, see Exercise 4.31 for some facts illustrating this. We will mainly be interested in the *numerical radius* of T, given by

$$N_T := \sup\{ |\lambda| : \lambda \in W_T\} = \sup\{ |\langle T(x), x\rangle| :\ x \in H, \|x\| = 1\}\,.$$

We note that the Cauchy-Schwarz inequality implies that $N_T \leq \|T\|$.

As we are going to prove below, a remarkable fact is that N_T agrees with $\|T\|$ when T is self-adjoint. We observe first that if T is self-adjoint, then $W_T \subseteq \mathbb{R}$. Indeed, if $T^* = T$, then for every $x \in H$, we have

$$\langle T(x), x\rangle = \langle x, T(x)\rangle = \overline{\langle T(x), x\rangle}\,,$$

and the claim clearly follows.

Theorem 4.4.9. *Let H be a Hilbert space, $H \neq \{0\}$, and let $T \in \mathcal{B}(H)$ be self-adjoint. Then we have*

$$\|T\| = N_T\,.$$

Proof. It suffices to prove that $\|T\| \le N_T$, hence that

$$\|T(x)\| \le N_T \quad \text{for all } x \in H_1. \tag{4.4.1}$$

We first note that if $v \in H$, then $|\langle T(v), v\rangle| \le N_T \|v\|^2$.

Indeed, if $v = 0$, the claim is trivial. Otherwise, if $v \neq 0$ and $u := \frac{1}{\|v\|} v$, so $v = \|v\| u$, then

$$\left|\Big\langle T(v), v\Big\rangle\right| = \|v\|^2 \left|\Big\langle T(u), u\Big\rangle\right| \le N_T \|v\|^2 .$$

Let now $x \in H_1$. If $T(x) = 0$, then the inequality in (4.4.1) is trivially satisfied, so we can assume that $T(x) \neq 0$ and set $y := \frac{1}{\|T(x)\|} T(x) \in H_1$. Then we have

$$\|T(x)\| = \frac{1}{\|T(x)\|} \Big\langle T(x), T(x)\Big\rangle = \Big\langle T(x), y\Big\rangle. \tag{4.4.2}$$

Similarly, $\|T(x)\| = \Big\langle y, T(x)\Big\rangle$. As T is self-adjoint, we get

$$\|T(x)\| = \Big\langle T(y), x\Big\rangle. \tag{4.4.3}$$

Combining (4.4.2) and (4.4.3), and using our previous observations, as well as the parallellogram law and the fact that $\|x\| \le 1$, $\|y\| = 1$, we get

$$\begin{aligned}
\|T(x)\| &= \frac{1}{2}\left(\Big\langle T(x), y\Big\rangle + \Big\langle T(y), x\Big\rangle\right)\\
&= \frac{1}{4}\left(\Big\langle T(x+y), x+y\Big\rangle - \Big\langle T(x-y), x-y\Big\rangle\right)\\
&\le \frac{1}{4} N_T \left(\|x+y\|^2 + \|x-y\|^2\right)\\
&= \frac{1}{2} N_T \left(\|x\|^2 + \|y\|^2\right)\\
&\le N_T .
\end{aligned}$$

This shows that (4.4.1) is satisfied, as desired. ∎

Having in mind the spectral theorem for symmetric real matrices, it is legitimate to wonder whether it could be true that every self-adjoint operator $T \in \mathcal{B}(H)$ is *diagonalizable* in the sense that there exists an orthonormal basis for H consisting of eigenvectors for T. One quickly realizes that this can not be the case, as a self-adjoint operator may not have any eigenvalue at all!

Example 4.4.10. Let $H = L^2([0,1])$ (with usual Lebesgue measure) and let $T = M_f$ be the self-adjoint operator in $\mathcal{B}(H)$ given by multiplication with the bounded continuous function $f(t) = t$ on $[0,1]$, cf. Example 4.4.4. Then the reader should have no trouble in checking that T has no eigenvalues.

We will see in the next chapter that every *compact* self-adjoint operator can be diagonalized in the sense mentioned above. Theorem 4.4.9 will help us to make the first step in proving this, by showing that a compact self-adjoint operator T has at least one an eigenvalue, namely $\|T\|$ or $-\|T\|$.

4.5 Unitary operators

In this section, we look at another important class of operators on Hilbert spaces. As a warm-up, we first characterize the linear operators which are isometric. We recall that if H is a Hilbert space, then a map $T : H \to H$ is said to *preserve the inner product* when it satisfies

$$\Big\langle T(x), T(y) \Big\rangle = \langle x, y \rangle \quad \text{for all } x, y \in H.$$

Proposition 4.5.1. *Let $H \neq \{0\}$ be a Hilbert space (over $\mathbb{F}$) and let $S : H \to H$. Then the following conditions are equivalent:*

(*i*) $S \in \mathcal{B}(H)$ *and* $S^*S = I_H$;

(*ii*) S *is linear and preserves the inner product*;

(*iii*) S *is a linear isometry.*

Proof. (*i*) $\Rightarrow$ (*ii*): Assume $S \in \mathcal{B}(H)$ satisfies $S^*S = I_H$. Then S is linear and for all $x, y \in H$, we have

$$\Big\langle S(x), S(y) \Big\rangle = \Big\langle x, (S^*S)(y) \Big\rangle = \langle x, y \rangle \, ,$$

so (*ii*) holds.
(*ii*) $\Rightarrow$ (*iii*): Any map preserving the inner product is isometric, so this is evident.
(*iii*) $\Rightarrow$ (*i*): Assume S is a linear isometry. Then $S \in \mathcal{B}(H)$ and $T := S^*S - I_H \in \mathcal{B}(H)$ is self-adjoint. Then for any $x \in H$, we have

$$\Big\langle T(x), x \Big\rangle = \Big\langle (S^*S - I)(x), x \Big\rangle = \Big\langle S(x), S(x) \Big\rangle - \langle x, x \rangle = \|S(x)\|^2 - \|x\|^2 = 0$$

Thus, $W_T = \{0\}$, so, using Theorem 4.4.9, we get that $\|T\| = N_T = 0$. Hence, $T = 0$, i.e., $S^*S = I_H$, so (*i*) holds. ∎

Example 4.5.2. Assume H is finite-dimensional and $S : H \to H$ is a linear isometry, so $S^*S = I_H$, cf. Proposition 4.5.1. As S is injective, it is also surjective (since $\dim(S(H)) = \dim(H) - \dim(\ker(S)) = \dim(H)$, so $S(H) = H$). Thus, S is bijective, so it has an inverse S^{-1} (which is also a linear isometry). Since $S^*S = I_H$, we get that $S^{-1} = S^*$. In particular, we also have $SS^* = I_H$.

Remark 4.5.3. When H is infinite-dimensional, then a linear isometry S is not necessarily surjective. A typical example is the right shift operator S considered in Example 4.3.5, whose range does not contain the first basis vector; in this case, we have $S^*S = I_H$, while $SS^* \neq I_H$ (cf. Exercise 4.20).

Definition 4.5.4. Let H be a Hilbert space (over $\mathbb{F}$). An operator $U \in \mathcal{B}(H)$ is called *unitary* when it satisfies

$$U^*U = UU^* = I_H .$$

Thus, $U \in \mathcal{B}(H)$ is unitary if and only if U is bijective and $U^{-1} = U^*$.

When $\mathbb{F} = \mathbb{R}$, some authors use the word *orthogonal* instead of unitary.

Proposition 4.5.5. *Let H be a Hilbert space (over $\mathbb{F}$) and let $U : H \to H$. Then the following conditions are equivalent:*

(*i*) *$U \in \mathcal{B}(H)$ and U is unitary;*

(*ii*) *U is bijective, linear and preserves the inner product;*

(*iii*) *U is a surjective linear isometry.*

Proof. $(i) \Rightarrow (ii)$: If $U \in \mathcal{B}(H)$ is unitary, then U is bijective and linear, and Proposition 4.5.1 gives that it preserves the inner product. Hence, (ii) holds.
$(ii) \Rightarrow (iii)$: This implication is evident.
$(iii) \Rightarrow (i)$: Suppose U is a surjective linear isometry. As a linear isometry is injective, U is bijective. Moreover, Proposition 4.5.1 gives that $U^*U = I_H$. So we get that $U^{-1} = U^*$, i.e., U is unitary, and (i) holds. ■

Example 4.5.6. Assume H has a countably infinite orthonormal basis $\mathcal{B}$ and D is the diagonal operator associated to a bounded sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ in $\mathbb{F}$ (w.r.t. $\mathcal{B}$).

Then it is straightforward to check that D is unitary if and only if $\overline{\lambda_j}\lambda_j = 1$, i.e., $|\lambda_j| = 1$, for all $j \in \mathbb{N}$.

Example 4.5.7. Let $(X, \mathcal{A}, \mu)$ be a measure space and set $H := L^2(X, \mathcal{A}, \mu)$. For $f \in \mathcal{L}^\infty$, consider the multiplication operator $M_f \in \mathcal{B}(H)$. Then we clearly have

$$(M_f)^* M_f = M_{|f|^2} = M_f (M_f)^*,$$

so we see that M_f is unitary whenever $|f| = 1$ μ-a.e. The converse holds if μ is semifinite, cf. Exercise 4.29.

Example 4.5.8. Let $H = \ell^2(\mathbb{Z})$. We may then define *the bilateral forward shift operator* $U : H \to H$ by

$$[U(\xi)](j) = \xi(j-1) \quad \text{for all } \xi \in H \text{ and all } j \in \mathbb{Z}.$$

Indeed, since the counting measure on $\mathbb{Z}$ is obviously translation-invariant, we have $\sum_{j\in\mathbb{Z}} |\xi(j-1)|^2 = \sum_{j\in\mathbb{Z}} |\xi(j)|^2 < \infty$, so we see that $U(\xi) \in H$ and $\|U(\xi)\|_2 = \|\xi\|_2$. Thus U is isometric.

Clearly, U is also linear. Moreover, it is surjective: if $\eta \in H$, then we have $U(\xi) = \eta$, where ξ is defined by $\xi(j) := \eta(j+1)$ for every $j \in \mathbb{Z}$; one may here argue as above to see that $\xi \in H$.

We may now conclude from Proposition 4.5.5 that U is unitary. Its adjoint $U^* = U^{-1}$ is called *the bilateral backward shift operator* (on $H = \ell^2(\mathbb{Z})$). We note that if $\mathcal{B} = \{e_n\}_{n\in\mathbb{Z}}$ denotes the canonical basis of $H = \ell^2(\mathbb{Z})$ as in Example 4.2.4, then we have

$$U(e_n) = e_{n+1} \quad \text{and} \quad U^*(e_n) = e_{n-1} \quad \text{for all } n \in \mathbb{Z}.$$

Let now H, K be Hilbert spaces (over $\mathbb{F}$). A bijective, linear map U from H onto K which preserves the inner product is often called an *isomorphism of Hilbert spaces*. As in Proposition 4.5.5, one shows that it is equivalent to require that U is a surjective linear isometry, or that $U \in \mathcal{B}(H, K)$ is unitary in the sense that we have $U^*U = I_H$ and $UU^* = I_K$. (Here, $U^* \in \mathcal{B}(K, H)$ denotes the adjoint of U, cf. Remark 4.3.3). We will say therefore say that H and K are *isomorphic as Hilbert spaces* when such a map $U : H \to K$ exists.

Theorem 4.5.9. *Let $H \neq \{0\}$ be a Hilbert space over $\mathbb{C}$, and let $\mathcal{B}$ be an orthonormal basis of H. Then H and $\ell^2(\mathcal{B})$ are isomorphic as Hilbert spaces.*

Proof. Let $x \in H$ and define $\widehat{x} : \mathcal{B} \to \mathbb{C}$ by

$$\widehat{x}(u) := \langle x, u \rangle \quad \text{for all } u \in \mathcal{B}.$$

Then Parseval's identity says that $\sum_{u\in\mathcal{B}} |\widehat{x}(u)|^2 = \|x\|^2$. In particular, we have $\widehat{x} \in \ell^2(\mathcal{B})$ and $\|\widehat{x}\| = \|x\|$. Thus we can define an isometric map $U : H \to \ell^2(\mathcal{B})$ by

$$U(x) = \widehat{x} \quad \text{for all } x \in H.$$

It is elementary to check that U is linear. Moreover, U is surjective.

Indeed, let $\xi \in \ell^2(\mathcal{B})$. As $\sum_{u\in\mathcal{B}} |\xi(u)|^2 < \infty$, the set

$$\mathcal{B}_\xi := \{u \in \mathcal{B} : \xi(u) \neq 0\}$$

must be countable. Let $\{u_j\}_{j\in N}$ be an enumeration of $\mathcal{B}_\xi$, where $N = \{1, \ldots, n\}$ for some $n \in \mathbb{N}$ or $N = \mathbb{N}$. If $N = \mathbb{N}$, we have $\sum_{j=1}^{\infty} |\xi(u_j)|^2 < \infty$, and this implies readily that the sequence $\left\{ \sum_{j=1}^{k} \xi(u_j)\, u_j \right\}_{k\in\mathbb{N}}$ is Cauchy, hence convergent in H. Thus we may define $x \in H$ by $x := \sum_{j\in N} \xi(u_j)\, u_j$, and we then have

$$\widehat{x}(u) = \langle x, u\rangle = \sum_{j\in N} \xi(u_j)\langle\, u_j, u\rangle = \begin{cases} \xi(u_k) & \text{if } u = u_k \text{ for some } k \in N, \\ 0 & \text{if } u \in \mathcal{B} \setminus \mathcal{B}_\xi\, , \end{cases}$$

i.e., $\widehat{x}(u) = \xi(u)$ for all $u \in \mathcal{B}$. Hence, $U(x) = \xi$, showing that U is surjective.

We can now conclude that U is an isomorphism of Hilbert spaces from H to $\ell^2(\mathcal{B})$, as we wanted to show.

■

Remark 4.5.10. Theorem 4.5.9 is also true when $H \neq \{0\}$ is a Hilbert space over $\mathbb{R}$, but one has then to replace $\ell^2(\mathcal{B})$ with the *real* ℓ^2-space

$$\ell^2_{\mathbb{R}}(\mathcal{B}) := \Big\{\xi : \mathcal{B} \to \mathbb{R} : \sum_{u\in\mathcal{B}} |\xi(u)|^2 < \infty, \Big\}$$

considered as a Hilbert space over $\mathbb{R}$.

Remark 4.5.11. If $H \neq \{0\}$ is a Hilbert space over $\mathbb{C}$, and $\mathcal{B}, \mathcal{B}'$ are both orthonormal bases of H, then we get from Theorem 4.5.9 that $\ell^2(\mathcal{B})$ and $\ell^2(\mathcal{B}')$ are isomorphic as Hilbert spaces. It can be shown that this implies that (and in fact is equivalent to) $\mathcal{B}$ and $\mathcal{B}'$ having the same cardinality, meaning that there exists a bijection between $\mathcal{B}$ and $\mathcal{B}'$. (A similar statement holds if $H \neq \{0\}$ is a Hilbert space over $\mathbb{R}$).

4.6 Exercises

In the exercises of this chapter, H always denotes a Hilbert space over $\mathbb{F}$, unless otherwise stated.

Exercise 4.1. Consider $X := \ell^\infty(\mathbb{N})$ as a metric space w.r.t. $d(f,g) = \|f - g\|_u$. Let A be the subset of X given by

$$A := \left\{a^{(N)} : N \in \mathbb{N}\right\},$$

where $a^{(N)}(n) = 1$ if $1 \leq n \leq N$ and $a^{(N)}(n) = 0$ if $n > N$.

a) Show that A is closed in X.

b) Let $x \in X$ be given by $x(n) = 1 + 1/n$ for all $n \in \mathbb{N}$.
Show that $d(x, A) = 1$ and that $1 < d(x, a^{(N)})$ for all $N \in \mathbb{N}$.

Exercise 4.2. Let $c \in H$, $r > 0$ and set $B := B_r(c) = \{y \in H : \|y - c\| \leq r\}$. Check that B is closed and convex, and give a formula for x_B when $x \in H \setminus B$.

Exercise 4.3. Let S denote a nonempty subset of H and set
$M := \overline{\text{Span}\,(S)}$.

Verify that $S^\perp = M^\perp$. Then deduce that $M = (S^\perp)^\perp$. Deduce also that if N is a subspace of H, then $\overline{N} = (N^\perp)^\perp$.

Exercise 4.4. Let M be a closed subspace of H, and let $x \in H$.

Show that $P_M(x) = y_0$ for some $y_0 \in M$ if and only if $x - y_0 \in M^\perp$. Show also that $P_M(x)$ is the unique vector y_0 in M such that $x - y_0 \in M^\perp$.

Exercise 4.5. Assume $P \in \mathcal{B}(H)$ is a projection satisfying $\|P\| = 1$. Show that P is the orthogonal projection of H on $M := P(H)$.

Hint: Recall that $H = P(H) \oplus \ker P$ (cf. Proposition 3.2.11).

Exercise 4.6. Consider $H = L^2([a,b], \mathcal{A}, \mu)$, where $\mathcal{A}$ denotes the σ-algebra of all Lebesgue measurable subsets of $[a,b]$, and μ is the usual Lebesgue measure on $\mathcal{A}$. Set

$$M := \left\{[g] \in H : g \in \mathcal{L}^2, \int_{[a,b]} g \, \mathrm{d}\mu = 0\right\}.$$

Check that M is a closed subspace of H. Then, given $[f] \in H$, find an expression for the best approximation of $[f]$ in M.

Exercise 4.7. Let H_1, H_2 be Hilbert spaces over $\mathbb{F}$ and consider $H := H_1 \times H_2$ as a vector space over $\mathbb{F}$. For $(x_1, x_2), (y_1, y_2) \in H$, set

$$\Big\langle (x_1, x_2), (y_1, y_2) \Big\rangle := \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle .$$

a) Check that this gives an inner product on H such that H is a Hilbert space. Check also that the associated norm on H corresponds to the norm $\| \cdot \|_2$ arising from the norms on H_1 and H_2.

b) Set $\widetilde{H}_1 := \{(x_1, 0) : x_1 \in H_1\}$ and $\widetilde{H}_2 := \{(0, x_2) : x_2 \in H_2\}$, so $H = \widetilde{H}_1 \dot{+} \widetilde{H}_2$ (cf. Exercise 3.5).

Check that $(\widetilde{H}_1)^\perp = \widetilde{H}_2$ and $(\widetilde{H}_2)^\perp = \widetilde{H}_1$. Deduce that the projection of H on $\widetilde{H}_1$ along $\widetilde{H}_2$ is the orthogonal projection of H on $\widetilde{H}_1$.

Exercise 4.8. Let $(X, \mathcal{A}, \mu)$ be a measure space.

a) Show that

$$\Big\langle [f], [g] \Big\rangle := \int_X f \, \overline{g} \, \mathrm{d}\mu$$

gives a well-defined inner product on $L^2 = L^2(X, \mathcal{A}, \mu)$ (cf. Example 4.1.6).

b) Let $E \in \mathcal{A}$. Set $\mathcal{A}_E = \{A \cap E : A \in \mathcal{A}\}$ and $\mu_E = \mu_{|\mathcal{A}_E}$. We recall that $(E, \mathcal{A}_E, \mu_E)$ is a measure space.

Show that there exists an isometric isomorphism from $L^2(E, \mathcal{A}_E, \mu_E)$ onto the space M_E defined in Example 4.1.6.

Exercise 4.9. Let $H \neq \{0\}$. Show that the following conditions are equivalent:

(a) H is separable;

(b) There is a sequence satisfying the assumptions in Example 4.2.5;

(c) H has a countable orthonormal basis.

Note that Example 4.2.5 shows that (b) $\Rightarrow$ (c). So it suffices to show that (a) $\Rightarrow$ (b), and (c) $\Rightarrow$ (a).

Exercise 4.10. Let H_1 and H_2 be Hilbert spaces over $\mathbb{F}$, and let H be the (external) direct product of H_1 and H_2, as defined in Exercise 4.7. Assume $\mathcal{B}_1$ and $\mathcal{B}_2$ are orthonormal bases for H_1 and H_2, respectively.

Find an orthonormal basis $\mathcal{B}$ for H in terms of $\mathcal{B}_1$ and $\mathcal{B}_2$.

Exercise 4.11. Let $H = L^2([-1,1], \mathcal{A}, \mu)$, where $\mathcal{A}$ denote the Lebesgue-measurable subsets of $[-1,1]$ and μ is the restriction of the usual Lebesgue measure to $\mathcal{A}$.

For each $n \in \{0\} \cup \mathbb{N}$, let $p_{n+1} : [-1,1] \to \mathbb{C}$ be defined by $p_{n+1}(t) = t^n$, and set $S := \Big\{[p_{n+1}] : n \in \{0\} \cup \mathbb{N}\Big\} \subseteq H$.

a) Show that $\operatorname{Span}(S)$ is dense in H.

b) Apply the Gram-Schmidt orthonormalization process to S to obtain an orthonormal basis $\mathcal{B} = \Big\{[q_{n+1}] : n \in \{0\} \cup \mathbb{N}\Big\}$ for H, where each q_{n+1} is the polynomial on $[-1,1]$ given by

$$q_{n+1}(t) = \frac{\sqrt{n + \frac{1}{2}}}{2^n \, n!} \, \frac{d^n}{dt^n}\Big((t^2 - 1)^n\Big) .$$

(These polynomials are called *the normalized Legendre polynomials.*)

Exercise 4.12. The concept of generalized sums can be used to provide an alternative way of describing Fourier expansions in Hilbert spaces.

Let X be a normed space, J be a nonempty set, $\{x_j\}_{j\in J}$ be a family of elements of X, and $x \in X$. Then one says that the generalized sum $\sum_{j\in J} x_j$ converges to x *when the following holds: given $\varepsilon > 0$, there exists a finite subset $F_0 \subseteq J$ such that for all finite subsets F of J containing F_0, we have*

$$\Big\| \, x - \sum_{j\in F} x_j \, \Big\| < \varepsilon ,$$

in which case we write

$$x = \sum_{j\in J} x_j .$$

Consider a Hilbert space H and $x \in H$.

a) Show that we have

$$x = \sum_{u\in\mathcal{B}} \langle x, u\rangle \, u .$$

b) Show also that if M is a closed subspace of H and $\mathcal{C}$ is an orthonormal basis for M, then we have

$$P_M(x) = \sum_{v\in\mathcal{C}} \langle x, v\rangle \, v .$$

Exercise 4.13. In the context of Fourier analysis described in Example 4.2.12 (see also Example 4.2.2), the formula

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} \left|\widehat{f}(n)\right|^2$$

is called *Parseval's identity.* (The more general equality obtained in Theorem 4.2.8 c) is also often called Parseval's identity.)

a) Set $f(t) = t$ for all $t \in [-\pi, \pi]$. Compute the Fourier coefficients of f.

b) Use a) and Parseval's identity to show that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

c) Set $g(t) = e^t$ for all $t \in [-\pi, \pi]$. Use Parseval's identity to obtain a formula for the sum of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + 1}.$$

Exercise 4.14. Let H be the L^2-space on $[-\pi, \pi]$ w.r.t. to the normalized Lebesgue measure μ, as in Example 4.2.2. Set

$$H_{\mathsf{even}} := \{[f] \in H : f \text{ is even}\} \quad \text{and} \quad H_{\mathsf{odd}} := \{[f] \in H : f \text{ is odd}\}.$$

We recall that $f : [-\pi, \pi] \to \mathbb{C}$ is called even if $f(-t) = f(t)$ for all t, while it is called odd if $f(-t) = -f(t)$ for all t.

a) Show that H_{even} is a closed subspace of H and that $(H_{\mathsf{even}})^{\perp} = H_{\mathsf{odd}}$. Describe the orthogonal projection P of H on H_{even}.
Hint: It might be helpful to consider the map $[f] \to [\widetilde{f}\,]$, where $\widetilde{f}(t) := f(-t)$.

b) Find an orthonormal basis for H_{even} and one for H_{odd}.

Exercise 4.15. Let $T \in \mathcal{B}(H)$. Assume H_0 is a dense subspace of H which is invariant under T, and let $T_0 \in \mathcal{B}(H_0)$ denote the restriction of T to H_0. Further, assume there exists some $S_0 \in \mathcal{B}(H_0)$ such that

$$\left\langle T_0(x), y \right\rangle = \left\langle x, S_0(y) \right\rangle \quad \text{for all } x, y \in H_0.$$

Show that $T^* = S$, where $S \in \mathcal{B}(H)$ is the unique extension of S_0 provided by Theorem 3.3.2.

Exercise 4.16. Show that the formula for $(T_K)^*$ in Example 4.3.7 is correct.

Hint : Consider $H_0 = [g] : g \in C([a, b])\}$ and use Exercise 4.15.

Exercise 4.17. Prove Proposition 4.3.10.

Exercise 4.18. Let $v, w \in H$ and consider the linear operator $T_{v,w} : H \to H$ defined by

$$T_{v,w}(x) := \langle x, v\rangle\, w \qquad \text{for all } x \in H\,.$$

Note that $T_{v,w}$ has rank one if $v, w \in H \setminus \{0\}$.

a) Show that $T_{v,w}$ is bounded with norm $\|T_{v,w}\| = \|v\|\,\|w\|$. Then show that $(T_{v,w})^* = T_{w,v}$.

b) Show that every $T \in \mathcal{B}(H)$ which has rank one is of the form $T = T_{v,w}$ for some $v, w \in H \setminus \{0\}$.

c) Assume $T \in \mathcal{B}(H)$ is a finite-rank operator, $T \neq 0$. Show that T may be written as a finite sum of rank one operators in $\mathcal{B}(H)$.

Hint : Start by picking an orthonormal basis for $T(H)$.

d) Show that if $T \in \mathcal{B}(H)$ is a finite-rank operator, then so is T^*.

Exercise 4.19. Let $T \in \mathcal{B}(H)$ and let M be a closed subspace of H. Show that

M is invariant under T if and only if $M^\perp$ is invariant under T^*.

Exercise 4.20. Let $T \in \mathcal{B}(H)$.

a) Show that $\ker(T) = \ker(T^*T)$ and $\overline{T^*(H)} = \overline{(T^*T)(H)}$.

b) Assume T is *normal*, i.e., satisfies $T^*T = TT^*$. Show that

$$\ker(T^*) = \ker(T) \text{ and } \overline{T^*(H)} = \overline{T(H)}\,.$$

c) Assume T is normal and has an eigenvalue λ. Show that $\overline{\lambda}$ is an eigenvalue of T^*, and that $E^{T^*}_{\overline{\lambda}} = E^T_\lambda$.

d) Assume H has a countably infinite orthonormal basis $\mathcal{B} = \{u_j\}_{j\in\mathbb{N}}$ and let $S \in \mathcal{B}(H)$ be the right shift operator (w.r.t. $\mathcal{B}$). Set $T = S^*$.

Show that T is not normal by checking that $TT^* = S^*S = I_H$, while $T^*T = SS^* = P$, where P is the orthogonal projection of H on $\{u_1\}^\perp$.

Check also that 0 is an eigenvalue for T, while 0 is not an eigenvalue of $T^* = S$. (This shows that c) does not necessarily hold when T is not normal.)

Finally, if you are in the right mood, show that S has no eigenvalues, while every λ satisfying $|\lambda| < 1$ is an eigenvalue of T.

Exercise 4.21. Let H and K be Hilbert spaces over $\mathbb{F}$, and let $T \in \mathcal{B}(H, K)$.

a) Show that there exists a unique operator $T^* \in \mathcal{B}(K, H)$ (called the adjoint of T) satisfying that

$$\left\langle T(x), y \right\rangle = \left\langle x, T^*(y) \right\rangle \quad \text{for all } x \in H \text{ and all } y \in K.$$

b) Let $T' \in \mathcal{B}(H, K)$ and $\alpha, \beta \in \mathbb{F}$. Let also L be a Hilbert space over $\mathbb{F}$, and let $S \in \mathcal{B}(K, L)$, so that $ST \in \mathcal{B}(H, L)$. Show that the following properties hold:

- *i)* $(\alpha T + \beta T')^* = \overline{\alpha}\, T^* + \overline{\beta}\, T'^*$; *ii)* $(ST)^* = T^* S^*$; *iii)* $(T^*)^* = T$;
- *iv)* $\|T^*\| = \|T\|$; *v)* $\|T^*T\| = \|T\|^2$.

Exercise 4.22. Let $(X, \mathcal{A}, \mu)$ be a measure space. One says that $(X, \mathcal{A}, \mu)$ is *semifinite* when the following condition holds: if $E \in \mathcal{A}$ and $\mu(E) = \infty$, then there exists $F \subseteq E$, $F \in \mathcal{A}$ such that $0 < \mu(F) < \infty$.

a) Show that $(X, \mathcal{A}, \mu)$ is semifinite whenever it is σ-finite.

Assume from now on that $(X, \mathcal{A}, \mu)$ *is semifinite.* Set $H := L^2(X, \mathcal{A}, \mu)$. Let $f \in \mathcal{L}^\infty$ and consider the multiplication operator $M_f \in \mathcal{B}(H)$ defined in Example 4.4.4.

b) Show that $\|M_f\| = \|f\|_\infty$.

c) Show that if M_f is self-adjoint, then f is real-valued μ-a.e. (As observed in Example 4.3.6, the converse is true without any restriction on $(X, \mathcal{A}, \mu)$.)

Exercise 4.23. Assume $P \in \mathcal{B}(H)$ is a self-adjoint projection, i.e., it satisfies that $P^* = P = P^2$. Show that P is the orthogonal projection of H on $M := P(H)$ (which is closed subspace of H, cf. Proposition 3.2.11).

Exercise 4.24. Let $H \neq \{0\}$.

a) Assume $T \in \mathcal{B}(H)$ is self-adjoint. Deduce from Theorem 4.4.9 that $T = 0$ if and only if $\left\langle T(x), x \right\rangle = 0$ for all $x \in H$.

b) Suppose $\mathbb{F} = \mathbb{R}$. Give an example with $H = \mathbb{R}^2$ showing that the equivalence in a) may fail when T is not self-adjoint.

c) Assume $\mathbb{F} = \mathbb{C}$ and let $T \in \mathcal{B}(H)$. Show that $T = 0$ if and only if $\left\langle T(x), x \right\rangle = 0$ for all $x \in H$.

Exercise 4.25. Show that the set $\mathcal{B}(H)_{\mathsf{sa}} := \{T \in \mathcal{B}(H) : T^* = T\}$ is closed in $\mathcal{B}(H)$.

Exercise 4.26. Let $H \neq \{0\}$. If $T \in \mathcal{B}(H)$ is self-adjoint, we have seen that $W_T \subseteq \mathbb{R}$; of course, if $\mathbb{F} = \mathbb{R}$, this gives no information on T as this inclusion is then true for any T in $\mathcal{B}(H)$. We assume therefore in this exercise that $H \neq \{0\}$ *is a Hilbert space over* $\mathbb{C}$.

Let $T \in \mathcal{B}(H)$. Then show that the following assertions are equivalent:

(i) T is self-adjoint;

(ii) $W_T \subseteq \mathbb{R}$;

(iii) $\big\langle T(x), x \big\rangle \in \mathbb{R}$ for all $x \in H$.

Exercise 4.27. A self-adjoint operator T in $\mathcal{B}(H)$ is called *positive* when

$$\big\langle T(x), x \big\rangle \geq 0 \quad \text{for all } x \in H\,, \tag{4.6.1}$$

in which case we write $T \geq 0$.

(We note that if $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{B}(H)$ satisfies (4.6.1), then T is automatically self-adjoint, as follows from the previous exercise.)

a) Let $S \in \mathcal{B}(H)$, and let $R \in \mathcal{B}(H)$ be self-adjoint.
Check that $S^*S \geq 0$ and $R^2 \geq 0$. Then show that

$$\|S\| \leq 1 \Leftrightarrow (I_H - S^*S) \geq 0\,.$$

b) Let M be a closed subspace of H. Check that $P_M \geq 0$.

c) Assume that $T, T' \in \mathcal{B}(H)$ are positive and $\lambda \in [0, \infty)$.
Check that $T + T'$ and λT are also positive.

d) Show that the set of positive operators in $\mathcal{B}(H)$ is closed in $\mathcal{B}(H)$.

Exercise 4.28. Let $H = L^2([0,1])$ (with usual Lebesgue measure) and let $T = M_f$ be the self-adjoint operator in $\mathcal{B}(H)$ given by multiplication with the function $f(t) = t$ on $[0,1]$, cf. Example 4.4.4. Show that T has no (complex) eigenvalues.

Exercise 4.29. Let $(X, \mathcal{A}, \mu)$ be a semifinite measure space (cf. Exercise 4.22), and let $f \in \mathcal{L}^\infty$. Suppose that the multiplication operator M_f on $H = L^2(X, \mathcal{A}, \mu)$ is unitary. Then show that $|f| = 1$ μ-a.e.
(As observed in Example 4.5.7, the converse statement is true without any restriction on $(X, \mathcal{A}, \mu)$.)

Exercise 4.30. Assume $H \neq \{0\}$ is separable (cf. Exercise 4.9) and infinite-dimensional. Let then $\mathcal{B}$ be an orthonormal basis for H indexed by $\mathbb{Z}$, say $\mathcal{B} = \{v_k\}_{k\in\mathbb{Z}}$. One may then define *the bilateral shift operator* $V : H \to H$ (w.r.t. $\mathcal{B}$) by

$$V(x) = \sum_{k\in\mathbb{Z}} \langle x, v_k\rangle\, v_{k+1} \quad \text{for all } x \in H, \quad \text{i.e., by}$$

$$V(x) = \lim_{n\to\infty} \sum_{k=-n}^{n} \langle x, v_k\rangle\, v_{k+1} \quad \text{for all } x \in H.$$

a) Show that V is a unitary operator in $\mathcal{B}(H)$.

b) Describe V as a multiplication operator when $H = L^2([-\pi, \pi])$ (with normalized Lebesgue measure μ) and $v_k(t) = e^{ikt}$ for every $k \in \mathbb{Z}$.

c) Assume $\mathbb{F} = \mathbb{C}$. Let $U : H \to \ell^2(\mathbb{Z})$ denote the isomorphism of Hilbert spaces defined in the proof of Theorem 4.5.9. Show that UVU^* is the bilateral forward shift operator on $\ell^2(\mathbb{Z})$.

Exercise 4.31. Let T be a bounded operator on a Hilbert space $H \neq \{0\}$. Check that the following properties of W_T and N_T hold:

(*a*) $W_{T^*} = \left\{\overline{\lambda} : \lambda \in W_T\right\}$; hence, $N_{T^*} = N_T$.

(*b*) W_T contains all the possible eigenvalues of T.

(*c*) $W_{\alpha T + \beta I_H} = \alpha\, W_T + \beta \quad$ for all $\alpha, \beta \in \mathbb{F}$.

(*d*) $W_{UTU^*} = W_T$, hence $N_{UTU^*} = N_T$, for every unitary $U \in \mathcal{B}(H)$.

(*e*) W_T is a compact subset of $\mathbb{F}$ when H is finite-dimensional.

It can also be shown that W_T is a convex subset of $\mathbb{F}$. This result is called the Toeplitz-Hausdorff Theorem, but the proof is beyond the scope of these notes.

CHAPTER 5

On compact operators

5.1 Introduction to compact operators between normed spaces

We had a very brief encounter with compact operators at the end of Section 3.1. For the ease of the reader, we recall their definition. In this section, X and Y will denote normed spaces, both over $\mathbb{F}$, unless otherwise specified.

Definition 5.1.1. An operator $T \in \mathcal{L}(X, Y)$ is called *compact* if the sequence $\{T(x_n)\}_{n\in\mathbb{N}}$ has a convergent subsequence in Y whenever $\{x_n\}_{n\in\mathbb{N}}$ is a bounded sequence in X.

We set $\mathcal{K}(X, Y) := \{T \in \mathcal{L}(X, Y) : T \text{ is compact}\}$.

To appreciate this definition, the concept of relative compactness for subsets of a metric space will be helpful.

A subset A of a metric space (Z, d) is called *relatively compact in* Z if its closure $\overline{A}$ is a compact subset of Z. (Some authors say *precompact* instead of relatively compact.) Equivalently, and this may be taken as the definition for our purposes, a subset A of Z is relatively compact in Z if and only if every sequence in A has a subsequence which converges in Z. In comparison, we recall that A is compact if and only if every sequence in A has a subsequence which converges in A.

We also remark that $A \subseteq Z$ is bounded whenever A is relatively compact in Z: indeed, if A is not bounded, then we can pick (any) $z \in Z$ and find a sequence $\{a_n\}_{n\in\mathbb{N}}$ in A such that $d(a_n, z) > n$ for all $n \in \mathbb{N}$; it is then rather easy to see that $\{a_n\}_{n\in\mathbb{N}}$ has no convergent subsequence in Z, so A is not relatively compact.

Proposition 5.1.2. *Let $T \in \mathcal{L}(X, Y)$. Then T is compact if and only if $T(B)$ is relatively compact in Y whenever B is a bounded subset of X.*

Proof. Assume first that T is compact and let $B \subseteq X$ be bounded. We want to show that $T(B)$ is relatively compact in Y. So let $\{y_n\}_{n\in\mathbb{N}}$ be a sequence in $T(B)$. For each $n \in \mathbb{N}$ we may then write $y_n = T(x_n)$ for some $x_n \in B$. As the sequence $\{x_n\}_{n\in\mathbb{N}}$ lies in B, it is bounded. Hence, by compactness of T, $\{y_n\}_{n\in\mathbb{N}} = \{T(x_n\}_{n\in\mathbb{N}}$ has a convergence subsequence in Y. Thus, $T(B)$ is relatively compact, as desired.

Conversely, assume that T maps bounded subsets of X into relatively compact subsets of Y. We want to show that T is compact. So let $\{x_n\}_{n\in\mathbb{N}}$ be a bounded sequence in X. Set $B := \{x_n : n \in \mathbb{N}\}$. Then B is a bounded subset of X, so $T(B) = \{T(x_n) : n \in \mathbb{N}\}$ is relatively compact in Y. As $\{T(x_n)\}_{n\in\mathbb{N}}$ is a sequence in $T(B)$, we can conclude that it has a convergent subsequence in Y. Thus, T is compact, as desired. ■

Corollary 5.1.3. *Assume $T \in \mathcal{L}(X,Y)$ is compact. Then T is bounded. Thus, $\mathcal{K}(X,Y) \subseteq \mathcal{B}(X,Y)$.*

Proof. Set $B := X_1$. Since B is a bounded subset of X, we get from Proposition 5.1.2 that $T(B)$ is relatively compact subset of Y. This implies that $T(B)$ is bounded. Hence we can find $M > 0$ such that $\|T(x)\| \leq M$ for all $x \in X_1$, and it follows that T is bounded with $\|T\| \leq M$. ■

As we have seen previously in Section 3.1, cf. Proposition 3.1.9, an important class of compact operators consists of the finite-rank operators in $\mathcal{B}(X,Y)$.

Example 5.1.4. Consider $X = C([0,1],\mathbb{R})$ with the uniform norm $\|\cdot\|_u$. For $g \in X$, define $T(g) : [0,1] \to \mathbb{R}$ by

$$[T(g)](s) = \int_0^1 \sin(s-t)\, g(t)\ \mathrm{d}t \quad \text{for all } s \in [0,1]\,.$$

Since $\sin(s-t) = \sin(s)\cos(t) - \cos(s)\sin(t)$, we have that

$$[T(g)](s) = \Big(\int_0^1 \cos(t)\, g(t)\ \mathrm{d}t\Big)\ \sin(s) - \Big(\int_0^1 \sin(t)\, g(t)\ \mathrm{d}t\Big)\ \cos(s)$$

for all $s \in [0,1]$. It follows that $T(g) \in X$. Moreover, the map $T : X \to X$ sending g to $T(g)$ is clearly linear. As $T(X)$ is 2-dimensional, T has finite-rank. Further, since

$$\Big|[T(g)](s)\Big| \ \leq\ \int_0^1 |\sin(s-t)g(t)|\ \mathrm{d}t \ \leq\ \int_0^1 |g(t)|\ \mathrm{d}t \ \leq\ \|g\|_u$$

for all $s \in [0,1]$, we get that $\|T(g)\|_u \leq \|g\|_u$ for all $g \in X$. Hence, T is bounded. We can therefore conclude that T is compact.

More generally, using the Arzelà-Ascoli Theorem (cf. Lindstrøm's book), it can be shown that if a function $K : [a,b] \times [c,d] \to \mathbb{R}$ is continuous, then the associated integral operator $T : C([c,d],\mathbb{R}) \to C([a,b],\mathbb{R})$, defined by

$$[T(g)](s) = \int_c^d K(s,t)\, g(t)\ \mathrm{d}t \quad \text{for all } s \in [a,b]\,,$$

is compact.

Theorem 5.1.5. *$\mathcal{K}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$. Moreover, if Y is a Banach space, then $\mathcal{K}(X,Y)$ is closed in $\mathcal{B}(X,Y)$, and it follows that $\mathcal{K}(X,Y)$ is a Banach space.*

Proof. We leave the proof of the first assertion as an exercise. So assume that Y is Banach space, and let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{K}(X,Y)$ converging to some $T \in \mathcal{B}(X,Y)$. To show that $\mathcal{K}(X,Y)$ is closed in $\mathcal{B}(X,Y)$, we have to show that T is compact.

So let $\{x_n\}_{n\in\mathbb{N}}$ be a bounded sequence in X. Choose $M > 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$.

- Since T_1 is compact, there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ such that $T_1(x_{n_k}) \to y_1$ as $k \to \infty$ for some $y_1 \in Y$.

  We set $x_{1,k} := x_{n_k}$ for each $k \in \mathbb{N}$. We then have $T_1(x_{1,n}) \to y_1$ as $n \to \infty$.

- Similarly, since $\{x_{1,n}\}_{n\in\mathbb{N}}$ is bounded and T_2 is compact, we can find a sequence $\{x_{2,n}\}_{n\in\mathbb{N}}$, which is a subsequence of $\{x_{1,n}\}_{n\in\mathbb{N}}$, and therefore of $\{x_n\}_{n\in\mathbb{N}}$, such that $T_2(x_{2,n}) \to y_2$ as $n \to \infty$ for some $y_2 \in Y$.

- Proceeding inductively, for each $m \in \mathbb{N}$, $m \geq 2$, we can find a sequence $\{x_{m,n}\}_{n\in\mathbb{N}}$, which is a subsequence of $\{x_{m-1,n}\}_{n\in\mathbb{N}}$, and therefore of $\{x_n\}_{n\in\mathbb{N}}$, such that $T_m(x_{m,n}) \to y_m$ as $n \to \infty$ for some $y_m \in Y$.

We now set $x'_k := x_{k,k} \in X$ for each $k \in \mathbb{N}$, and claim that

$$\{T(x'_k)\}_{k\in\mathbb{N}} \text{ is a Cauchy sequence in } Y. \tag{5.1.1}$$

Since Y is complete, we will then be able to conclude that $\{T(x'_k)\}_{k\in\mathbb{N}}$ has a convergent subsequence. As this subsequence will then be a subsequence of the sequence $\{T(x_n)\}_{n\in\mathbb{N}}$, we will thereby have shown that T is compact.

To establish (5.1.1), we first observe that for any $k, l, m \in \mathbb{N}$, we have

$$\begin{aligned} \|T(x'_l) - T(x'_k)\| &\leq \|(T - T_m)(x'_l) + T_m(x'_l) - T_m(x'_k) + (T_m - T)(x'_k)\| \\ &\leq \|(T - T_m)(x'_l)\| + \|T_m(x'_l) - T_m(x'_k)\| + \|(T_m - T)(x'_k)\| \\ &\leq \|T - T_m\|\, \|x'_l\| + \|T_m(x'_l) - T_m(x'_k)\| + \|T_m - T\|\, \|x'_k\| \\ &\leq \|T_m(x'_l) - T_m(x'_k)\| + 2M\, \|T - T_m\|\,. \end{aligned}$$

Let then $\varepsilon > 0$ and choose $m \in \mathbb{N}$ such that $\|T - T_m\| < \varepsilon/3M$. By construction, for each $k \geq m$, we have that $T_m(x'_k) = T_m(x_{k,k})$ is an element of the sequence $\{T_m(x_{m,n})\}_{n\in\mathbb{N}}$, which is convergent to y_m. It follows that the sequence $\{T_m(x'_k)\}_{k\in\mathbb{N}}$ is convergent, hence that it is Cauchy. So we can pick $N \in \mathbb{N}$ such that $\|T_m(x'_l) - T_m(x'_k)\| < \varepsilon/3$ for all $k, l \geq N$. This gives that

$$\|T(x'_l) - T(x'_k)\| \leq \|T_m(x'_l) - T_m(x'_k)\| + 2M\,\|T - T_m\| < \varepsilon/3 + 2M\,(\varepsilon/3M) = \varepsilon$$

for all $k, l \geq N$. Hence we have shown that the claim (5.1.1) is true.

Finally, as Y is a Banach space, we know that $\mathcal{B}(X, Y)$ is a Banach space too, and this implies that $\mathcal{K}(X, Y)$, being closed in $\mathcal{B}(X, Y)$, is also a Banach space. ∎

An immediate consequence is the following:

Corollary 5.1.6. *Assume Y is a Banach space and set*

$$\mathcal{F}(X, Y) := \{T \in \mathcal{B}(X, Y) : T \text{ has finite-rank}\}.$$

Then we have

$$\overline{\mathcal{F}(X, Y)} \subseteq \mathcal{K}(X, Y).$$

Example 5.1.7. Let $1 \leq p < \infty$ and set $X := \ell^p(\mathbb{N})$, which we know is a Banach space w.r.t. $\|\cdot\|_p$. For each $\lambda \in \ell^\infty(\mathbb{N})$, we may consider the multiplication operator $M_\lambda \in \mathcal{B}(X)$ given by

$$[M_\lambda(x)](n) = \lambda(n)\,x(n)$$

for all $x \in X$ and all $n \in \mathbb{N}$. One readily checks that $\|M_\lambda\| = \|\lambda\|_\infty$.

Now, assume that $\lambda \in c_0(\mathbb{N})$, i.e., $\lim_{n\to\infty} \lambda(n) = 0$. Then M_λ is compact. Indeed, for each $k \in \mathbb{N}$, let $\lambda^{(k)} \in \ell^\infty(\mathbb{N})$ be defined by

$$\lambda^{(k)}(n) = \begin{cases} \lambda(n) & \text{if } 1 \leq k \leq n, \\ 0 & \text{otherwise,} \end{cases}$$

for every $n \in \mathbb{N}$. Then it is clear that each $M_{\lambda^{(k)}}$ has finite-rank; moreover,

$$\|M_\lambda - M_{\lambda^{(k)}}\| = \|\lambda - \lambda^{(k)}\|_\infty \to 0 \quad \text{as } k \to \infty.$$

Thus $M_\lambda \in \overline{\mathcal{F}(X, X)} \subseteq \mathcal{K}(X, X)$.

We set $\mathcal{K}(X) = \mathcal{K}(X, X)$, so that $\mathcal{K}(X)$ is a subspace of $\mathcal{B}(X)$. If X is finite-dimensional, then every operator in $\mathcal{B}(X)$ has finite-rank, so $\mathcal{K}(X) = \mathcal{B}(X)$. On the other hand, if X is infinite-dimensional, then $\mathcal{K}(X) \neq \mathcal{B}(X)$, the reason being that the identity operator I_X is not compact in this case: indeed, if X is infinite-dimensional, then $I_X(X_1) = X_1$ is closed, but not compact, (cf. Exercise 3.2).

We also mention (cf. Exercise 5.1) that $\mathcal{K}(X)$ is a *two-sided ideal in* $\mathcal{B}(X)$, meaning that we have

$$ST \in \mathcal{K}(X) \text{ if } S \in \mathcal{B}(X) \text{ and } T \in \mathcal{K}(X), \text{ or if } S \in \mathcal{K}(X) \text{ and } T \in \mathcal{B}(X).$$

This property implies that no operator in $\mathcal{K}(X)$ can have a bounded inverse when X is infinite-dimensional (for if $T \in \mathcal{K}(X)$ has an inverse $T^{-1} \in \mathcal{B}(X)$, then we must have that $I_X = T^{-1}T \in \mathcal{K}(X)$, so $\dim(X) < \infty$).

We end this section with an interesting result concerning the possible eigenvalues of a compact operator.

Theorem 5.1.8. *Let $T \in \mathcal{K}(X)$. Then the following facts hold:*

(a) *Let $\delta > 0$. Then $\{\lambda \in \mathbb{F} : \lambda$ is an eigenvalue of T and $|\lambda| > \delta\}$ is a finite set.*

(b) *If $\lambda \in \mathbb{F}$ is a non-zero eigenvalue of T, then the associated eigenspace is finite-dimensional.*

(c) *The set of eigenvalues of T (which may be empty) is countable and bounded. If this set is countably infinite and $\{\lambda_k : k \in \mathbb{N}\}$ is an enumeration of it, then* $\lim_{k\to\infty} \lambda_k = 0$.

As we will be mostly interested in compact self-adjoint operators acting on Hilbert spaces in this course, for which much more can be said (cf. Theorem 5.3.4), we skip the proof of thi more general theorem.

5.2 On compact operators on Hilbert spaces

In view of Corollary 5.1.6, it is natural to wonder whether any compact operator from a normed space to a Banach space may be approximated in operator norm by bounded finite-rank operators. This problem was open until 1973, when a counterexample was exhibited by P. Enflo. Happily, the situation is as nice as possible when the target space is a Hilbert space.

Theorem 5.2.1. *Let X be a normed space and H be a Hilbert space (both over $\mathbb{F}$). Then we have*

$$\overline{\mathcal{F}(X, H)} = \mathcal{K}(X, H)\,.$$

Proof. By Corollary 5.1.6, we only have to show that $\mathcal{K}(X,H) \subseteq \overline{\mathcal{F}(X,H)}$. So let $T \in \mathcal{K}(X,H)$, and let $\varepsilon > 0$. We need to prove that there exists $S \in \mathcal{F}(X,H)$ such that $\|T - S\| \leq \varepsilon$. Clearly, we can assume $T \neq 0$.

Set $A := \overline{T(X_1)}$. Since X_1 is bounded and T is compact, the set A is compact in H. As H is a metric space, this implies that A is totally bounded (cf. Proposition 3.5.12 in Lindstrøm's book). Hence we can cover A with some open balls $B_1, \ldots, B_n$ of radius $\varepsilon/4$, having respective centers $a_1, \ldots, a_n \in A$. For each $j = 1, \ldots, n$, we can then find $x_j \in X_1$ such that $\|a_j - T(x_j)\| < \varepsilon/4$.

Set now $F := \mathrm{Span}\left(\{T(x_1), \ldots, T(x_n)\}\right)$, which is a finite dimensional subspace of H, and let P_F denote the orthogonal projection of H on F. Since the range of $P_F T$ is contained in F, $P_F T$ has finite-rank, so $P_F T \in \mathcal{F}(X,H)$. We claim that

$$\|T - P_F T\| \leq \varepsilon\,.$$

Indeed, let $x \in X_1$. Then $T(x) \in A$, so $T(x) \in B_j$ for some $j \in \{1, \ldots, n\}$. Hence,

$$\|T(x) - T(x_j)\| \;\leq\; \|T(x) - a_j\| + \|a_j - T(x_j)\| \;<\; \varepsilon/4 + \varepsilon/4 = \varepsilon/2\,.$$

Further, since $T(x_j) \in F$, we have that $P_F(T(x_j)) = T(x_j)$. Thus, using also that $\|P_F\| = 1$, we obtain that

$$\begin{aligned}
\|(T - P_F T)(x)\| &= \|T(x) - T(x_j) + (P_F T)(x_j) - (P_F T)(x)\| \\
&\leq \|T(x) - T(x_j)\| + \|P_F\big(T(x_j) - T(x)\big)\| \\
&\leq \|T(x) - T(x_j)\| + \|P_F\|\|T(x_j) - T(x)\| \\
&= 2\,\|T(x) - T(x_j)\| \\
&< 2 \cdot \varepsilon/2 = \varepsilon\,.
\end{aligned}$$

As this holds for every $x \in X_1$, the claim follows. Hence, setting $S := P_F T$, we are done. ∎

Remark 5.2.2. Let X be a normed space and H be a Hilbert space, and let $T \in \mathcal{K}(X,H)$. Then it can be shown that $\overline{T(X)}$ is separable. We leave this as an exercise.

Theorem 5.2.1 immediately gives:

Corollary 5.2.3. *Let H be a Hilbert space. Set $\mathcal{K}(H) := \mathcal{K}(H,H)$ and $\mathcal{F}(H) := \mathcal{F}(H,H)$. Then we have*

$$\overline{\mathcal{F}(H)} = \mathcal{K}(H)\,.$$

An application of this result is the following:

Corollary 5.2.4. *Let H be a Hilbert space and let $T \in \mathcal{K}(H)$. Then $T^* \in \mathcal{K}(H)$. In other words, $\mathcal{K}(H)$ is closed under the adjoint operation.*

Proof. Using Corollary 5.2.3, we can find a sequence $\{T_n\}$ in $\mathcal{F}(H)$ such that $\|T - T_n\| \to 0$ as $n \to \infty$. Now, it is not difficult to see that $\mathcal{F}(H)$ is closed under the adjoint operation (cf. Exercise 4.18). Hence, $\{T_n^*\}$ is a sequence in $\mathcal{F}(H)$, and we have

$$\|T^* - T_n^*\| = \|(T - T_n)^*\| = \|T - T_n\| \to 0 \quad \text{as } n \to \infty.$$

Thus, $T^* \in \overline{\mathcal{F}(H)} = \mathcal{K}(H)$. ■

We recall from the previous section that if H is finite-dimensional, then $\mathcal{K}(H) = \mathcal{B}(H) = \mathcal{F}(H)$, while $\mathcal{K}(H) \neq \mathcal{B}(H)$ if H is infinite-dimensional. An elementary argument showing that I_H is not compact can be given in this case: letting then $\{u_j\}_{j\in\mathbb{N}}$ be any orthonormal sequence in H, we have $\|u_j - u_k\| = \sqrt{2}$ for all $j, k \in \mathbb{N}$, and it follows that the sequence $\{I_H(u_j)\}_{j\in\mathbb{N}} = \{u_j\}_{j\in\mathbb{N}}$ does not have any convergent subsequence.

Let H be an infinite-dimensional Hilbert space H. An interesting class of compact operators on H containing $\mathcal{F}(H)$ consists of the so-called Hilbert-Schmidt operators. For simplicity, we only consider the case where H is separable. We note that every orthonormal basis for H is then countable: indeed, assume (for contradiction) that H had an uncountable orthonormal basis $\mathcal{B}$. Then, as $\|u - u'\| = \sqrt{2}$ for all distinct $u, u' \in \mathcal{B}$, we see that any dense subset of H would have to be uncountable, contradicting the separability of H.

Lemma 5.2.5. *Assume H is a separable, infinite-dimensional Hilbert space (over $\mathbb{F}$). Let $\mathcal{B} = \{u_j\}_{j\in\mathbb{N}}$ and $\mathcal{C} = \{v_j\}_{j\in\mathbb{N}}$ be orthonormal bases for H, and let $T \in \mathcal{B}(H)$. Then we have*

$$\sum_{j=1}^{\infty} \|T(u_j)\|^2 = \sum_{j=1}^{\infty} \|T(v_j)\|^2 .$$

Proof. Using Parseval's identity (two times), we get

$$\begin{aligned}
\sum_{j=1}^{\infty} \|T(u_j)\|^2 &= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \left|\left\langle T(u_j), v_k\right\rangle\right|^2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \left|\left\langle u_j, T^*(v_k)\right\rangle\right|^2 \\
&= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \left|\left\langle T^*(v_k), u_j\right\rangle\right|^2 = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \left|\left\langle T^*(v_k), u_j\right\rangle\right|^2 \\
&= \sum_{k=1}^{\infty} \|T^*(v_k)\|^2 .
\end{aligned}$$

Note that the change of order of summation at the second but last step above is allowed since we are dealing with sums of non-negative numbers. Applying what we have done to the case where $\mathcal{B} = \mathcal{C}$, i.e., $u_j = v_j$ for every $j \in \mathbb{N}$, we get that

$$\sum_{j=1}^{\infty} \|T(v_j)\|^2 = \sum_{k=1}^{\infty} \|T^*(v_k)\|^2 .$$

Thus we obtain that

$$\sum_{j=1}^{\infty} \|T(u_j)\|^2 = \sum_{k=1}^{\infty} \|T^*(v_k)\|^2 = \sum_{j=1}^{\infty} \|T(v_j)\|^2 ,$$

as desired. ■

Remark 5.2.6. An analogous result is true when H is finite-dimensional and $\mathcal{B}, \mathcal{C}$ are orthonormal bases for H.

Definition 5.2.7. Let H be a separable, infinite-dimensional Hilbert space (over $\mathbb{F}$). An operator $T \in \mathcal{B}(H)$ is called an *Hilbert-Schmidt operator* when we have

$$\sum_{j=1}^{\infty} \|T(u_j)\|^2 < \infty$$

for some orthonormal basis $\mathcal{B} = \{u_j\}_{j\in\mathbb{N}}$ of H, in which case we set

$$\|T\|_2 := \Big(\sum_{j=1}^{\infty} \|T(u_j)\|^2\Big)^{1/2} .$$

Lemma 5.2.5 shows that the definition of T being a Hilbert-Schmidt operator, and the value of $\|T\|_2$, do not depend on the choice of orthonormal basis for H. We also set

$$\mathcal{HS}(H) := \{T \in \mathcal{B}(H) : T \text{ is a Hilbert-Schmidt operator}\} .$$

Proposition 5.2.8. *Let H be a separable, infinite-dimensional Hilbert space (over $\mathbb{F}$).*

Then $\mathcal{HS}(H)$ is a subspace of $\mathcal{K}(H)$, which contains $\mathcal{F}(H)$ and is closed under the adjoint operation.

Moreover, the map $T \to \|T\|_2$ is a norm on $\mathcal{HS}(H)$, which satisfies

$$\|T\| \leq \|T\|_2$$

for every $T \in \mathcal{HS}(H)$.

Proof. We first note that it is evident from the proof of Lemma 5.2.5 that $T^* \in \mathcal{HS}(H)$ whenever $T \in \mathcal{HS}(H)$.

Let $\mathcal{B} = \{u_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for H, and let $T, T' \in \mathcal{HS}(H)$. Define $\xi, \xi' \in \ell^2(\mathbb{N})$ by

$$\xi(j) := \|T(u_j)\| \quad \text{and} \quad \xi'(j) := \|T'(u_j)\| \quad \text{for each } j \in \mathbb{N},$$

so that $\|\xi\|_2 = \|T\|_2$ and $\|\xi'\|_2 = \|T'\|_2$. Using the triangle inequality, first in H, and then in $\ell^2(\mathbb{N})$, we get

$$\begin{aligned}\sum_{j=1}^{\infty} \|(T+T')(u_j)\|^2 \;&\le\; \sum_{j=1}^{\infty} \Big(\|T(u_j)\| + \|T'(u_j)\|\Big)^2 = \|\xi + \xi'\|_2^2 \\ &\le\; (\|\xi\|_2 + \|\xi'\|_2)^2 = (\|T\|_2 + \|T'\|_2)^2 \;<\; \infty\,.\end{aligned}$$

This shows that $T + T' \in \mathcal{HS}(H)$ and

$$\|T + T'\|_2 \le \|T\|_2 + \|T'\|_2\,.$$

Moreover, one easily checks that $\lambda T \in \mathcal{HS}(H)$ and $\|\lambda T\|_2 = |\lambda|\,\|T\|_2$ for every $\lambda \in \mathbb{F}$. If $\|T\|_2 = 0$, then we get that $\|T(u_j)\| = 0$ for every $j \in \mathbb{N}$, and this clearly implies that $T = 0$.

Hence, we have shown so far that $\mathcal{HS}(H)$ is a subspace of $\mathcal{B}(H)$ which is closed under the adjoint operation, and that $\|\cdot\|_2$ is a norm on $\mathcal{HS}(H)$.

To show that $\|T\| \le \|T\|_2$, let $x \in H \setminus \{0\}$. Set $v_1 = \frac{1}{\|x\|}x$ and let $\{v_j\}_{j\ge 2}$ be an orthonormal basis for $\{x\}^{\perp}$. Then $\{v_j\}_{j\in\mathbb{N}}$ is an orthonormal basis for H, so we get

$$\|T(x)\|^2 = \|x\|^2\,\|T(v_1)\|^2 \;\le\; \|x\|^2 \sum_{j=1}^{\infty} \|T(v_j)\|^2 = \|T\|_2^2\,\|x\|^2\,.$$

Thus, $\|T\| \le \|T\|_2$.

Next, we show that $T \in \mathcal{K}(H)$. For each $n \in \mathbb{N}$, let P_n denote the orthogonal projection of H on $\operatorname{Span}(\{u_1, \dots u_n\})$ and set $T_n := TP_n$. Then we have

$$\sum_{j=1}^{\infty} \|T_n(u_j)\|^2 = \sum_{j=1}^{n} \|T(u_j)\|^2 < \infty\,,$$

so $T_n \in \mathcal{HS}(H)$ for each $n \in \mathbb{N}$. Hence,

$$\|T - T_n\| \;\le\; \|T - T_n\|_2 = \Big(\sum_{j=n+1}^{\infty} \|T(u_j)\|^2\Big)^{1/2} \to 0 \quad \text{as } n \to \infty\,.$$

Since $T_n \in \mathcal{F}(H)$ for each n, Theorem 5.1.6 gives that $T \in \mathcal{K}(H)$. Hence, $\mathcal{HS}(H) \subseteq \mathcal{K}(H)$.

It only remains to show that $\mathcal{F}(H) \subseteq \mathcal{HS}(H)$, but we leave this as an exercise. ■

Remark 5.2.9. For additional properties of $\mathcal{HS}(H)$, see Exercise 5.6.

Remark 5.2.10. If $H \neq \{0\}$ is finite-dimensional and $\mathcal{B} = \{u_j\}_{j=1}^n$ is an orthonormal basis for H, then we get a norm on $\mathcal{B}(H)$ by setting

$$\|T\|_2 := \Big(\sum_{j=1}^{n} \|T(u_j)\|^2\Big)^{1/2}$$

(which does not depend on the choice of orthonormal basis for H).

Letting $A = [a_{i,j}]$ denotes the matrix of T w.r.t. $\mathcal{B}$, one readily checks that

$$\|T\|_2 = \Big(\sum_{i,j=1}^{n} |a_{i,j}|^2\Big)^{1/2},$$

i.e., $\|T\|_2$ coincides with the so-called Fröbenius-norm of A.

Example 5.2.11. Set $H = L^2([a,b], \mathcal{A}, \mu)$, where $\mathcal{A}$ denotes the Lebesgue measurable subsets of a closed interval $[a,b]$ and μ is the Lebesgue measure on $\mathcal{A}$. Let $K : [a,b] \times [a,b] \to \mathbb{C}$ be a continuous function and let $T_K \in \mathcal{B}(H)$ denote the associated integral operator on H, which is the extension of the integral operator $T_K : C([a,b]) \to C([a,b])$ given by

$$[T_K(f)](s) = \int_a^b K(s,t)\ \mathrm{d}t \quad \text{for } f \in C([a,b]) \text{ and } s \in [a,b].$$

cf. Example 3.3.6 and Exercise 3.18. Then T_K is a Hilbert-Schmidt operator on H (so T_K is compact by Proposition 5.2.8).

To show this, we start by picking an orthonormal basis $\mathcal{B} = \{[u_j]\}_{j\in\mathbb{N}}$ for H, where each u_j is a continuous functions on $[a,b]$. (One may for example construct $\mathcal{B}$ by applying the Gram-Schmidt orthonormalization process to the monomials $\{t^{j-1} : j \in \mathbb{N}\}$). We note that $\overline{\mathcal{B}} := \{[\,\overline{u_j}\,]\}_{j\in\mathbb{N}}$ is then also an orthonormal basis for H.

Let now $s \in [a,b]$ and let $k_s \in C([a,b])$ be given by $k_s(t) = K(s,t)$ for all $t \in [a,b]$. Note that for each $j \in \mathbb{N}$, we have

$$[T_K(u_j)](s) = \int_a^b K(s,t)\, u_j(t)\ \mathrm{d}t = \int_{[a,b]} k_s(t)\, \overline{\overline{u_j}(t)}\ \mathrm{d}\mu(t) = \Big\langle [k_s], [\overline{u_j}] \Big\rangle .$$

Moreover, Parseval's identity gives that

$$\|\, [k_s]\, \|_2 = \Big(\sum_{j=1}^{\infty} \Big|\Big\langle [k_s], [\overline{u_j}] \Big\rangle\Big|^2\Big)^{1/2} .$$

Thus, we obtain that

$$\sum_{j=1}^{\infty} \Big| [T_K(u_j)](s) \Big|^2 = \sum_{j=1}^{\infty} \Big| \Big\langle [k_s], [\overline{u_j}] \Big\rangle \Big|^2 = \| [k_s] \|_2^2 .$$

Now, using this and the Monotone Convergence Theorem, we get

$$\begin{aligned} \sum_{j=1}^{\infty} \| T_K([u_j]) \|_2^2 &= \sum_{j=1}^{\infty} \int_{[a,b]} \Big| [T_K(u_j)](s) \Big|^2 \, \mathrm{d}\mu(s) \\ &= \int_{[a,b]} \Big(\sum_{j=1}^{\infty} \Big| [T_K(u_j)](s) \Big|^2 \Big) \, \mathrm{d}\mu(s) \\ &= \int_{[a,b]} \| [k_s] \|_2^2 \, \mathrm{d}\mu(s) \\ &= \int_{[a,b]} \Big(\int_{[a,b]} |k_s(t)|^2 \, \mathrm{d}\mu(t) \Big) \, \mathrm{d}\mu(s) \\ &= \int_a^b \int_a^b |K(s,t)|^2 \, \mathrm{d}t \, \mathrm{d}s < \infty, \end{aligned}$$

which shows that $T_K \in \mathcal{HS}(H)$ with $\|T_K\|_2 \leq \int_a^b \int_a^b |K(s,t)|^2 \, \mathrm{d}s \, \mathrm{d}t$.

In the previous example, one may allow the kernel K to be discontinuous and still obtain an Hilbert-Schmidt operator T_K, as long as K is square-integrable w.r.t. to the product Lebesgue measure on $[a,b] \times [a,b]$. However, this requires a better knowledge of measure theory than the one we preassume in these notes.

5.3 The spectral theorem for a compact self-adjoint operator

Throughout this section we let H denote a Hilbert space (over $\mathbb{F}$) different from $\{0\}$. Our main goal is to generalize the spectral theorem for symmetric real matrices known from linear algebra, and prove that every compact self-adjoint compact operator T on H is diagonalizable in the sense that there exists an orthonormal basis for H consisting of eigenvectors of T.

We begin with a series of lemmas.

Lemma 5.3.1. *Assume $T \in \mathcal{K}(H)$ has a nonzero eigenvalue $\lambda \in \mathbb{F}$. Then the associated eigenspace $E_\lambda := \ker(T - \lambda I)$ is finite-dimensional.*

Proof. Assume for contraction that E_λ is infinite-dimensional. We may then find a sequence $\{v_n\}_{n\in\mathbb{N}}$ of unit vectors in E_λ which are pairwise orthogonal. By compactness of T, $\{T(v_n)\}_{n\in\mathbb{N}}$ has a convergent subsequence. So we may as well assume that $\{T(v_n)\}_{n\in\mathbb{N}}$ is convergent, hence that it is a Cauchy sequence. However, we have that

$$\|T(v_n) - T(v_m)\|^2 = \|\lambda\, v_n - \lambda\, v_m\|^2 = 2\,|\lambda|^2 \neq 0$$

for all $m, n \in \mathbb{N}$. So $\{T(v_n)\}_{n\in\mathbb{N}}$ is not a Cauchy sequence, giving a contradiction. ■

Lemma 5.3.2. *Let $T \in \mathcal{B}(H)$ be self-adjoint, and assume T has an eigenvalue $\lambda \in \mathbb{F}$. Then $\lambda \in \mathbb{R}$.*

Moreover, if λ' is an eigenvalue of T distinct from λ, then $E_\lambda \perp E_{\lambda'}$, i.e., $\langle x, y\rangle = 0$ whenever $x \in E_\lambda$ and $y \in E_{\lambda'}$.

Proof. Let $x \in E_\lambda$. If $\|x\| = 1$, then we have

$$\lambda = \lambda\,\langle x, x\rangle = \langle \lambda\, x, x\rangle = \langle T(x), x\rangle \in W_T \subseteq \mathbb{R}\,,$$

so $\lambda \in \mathbb{R}$. Moreover, assume that λ' is an eigenvalue of T distinct from λ, and let $y \in E_{\lambda'}$. Then we have that $\lambda' \in \mathbb{R}$, so

$$\lambda\,\langle x, y\rangle = \langle T(x), y\rangle = \langle x, T(y)\rangle = \lambda'\,\langle x, y\rangle\,.$$

Since $\lambda' \neq \lambda$, we get that $\langle x, y\rangle = 0$. ■

Lemma 5.3.3. *Let $T \in \mathcal{K}(H)$ be self-adjoint. Then T has an eigenvalue $\lambda \in \mathbb{R}$ such that $|\lambda| = \|T\|$.*

Proof. If $T = 0$, then the assertion is trivial. So assume that $T \neq 0$. Using Theorem 4.4.9, we can find a sequence $\{x_n\}_{n\in\mathbb{N}}$ of unit vectors in H such that $|\langle T(x_n), x_n\rangle| \to \|T\|$ as $n \to \infty$. Since $\langle T(x_n), x_n\rangle \in \mathbb{R}$ for every n, we can assume (by passing to a subsequence and relabelling) that

$$\langle T(x_n), x_n\rangle \to \lambda \text{ as } n \to \infty, \text{ where } \lambda = \pm\|T\|. \tag{5.3.1}$$

Moreover, since T is compact, we can also assume (by passing again to a subsequence and relabelling) that $T(x_n) \to y$ as $n \to \infty$ for some $y \in H$. Note that the Cauchy-Schwarz inequality gives that

$$\Big|\langle T(x_n), x_n\rangle\Big| \leq \|T(x_n)\| \quad \text{for every } n \in \mathbb{N},$$

so, letting $n \to \infty$, we get that $\|y\| \geq |\lambda| > 0$, so $y \neq 0$.

Now, using that T is self-adjoint, λ is real, $\|x_n\| = 1$, and (5.3.1), we get

$$\begin{aligned}
\|T(x_n) - \lambda x_n\|^2 &= \Big\langle T(x_n) - \lambda x_n, T(x_n) - \lambda x_n\Big\rangle \\
&= \|T(x_n)\|^2 - 2\lambda\,\langle T(x_n), x_n\rangle + \lambda^2\|x_n\|^2 \\
&\leq \|T\|^2 - 2\lambda\,\langle T(x_n), x_n\rangle + \lambda^2 \\
&= 2\lambda\Big(\lambda - \langle T(x_n), x_n\rangle\Big) \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}$$

Thus, $\|T(x_n) - \lambda x_n\| \to 0$ as $n \to \infty$, and this gives that

$$\|y - \lambda x_n\| \leq \|y - T(x_n)\| + \|T(x_n) - \lambda x_n\| \to 0 \text{ as } n \to \infty.$$

Hence,

$$T(y) = \lim_{n\to\infty} T(\lambda x_n) = \lambda \lim_{n\to\infty} T(x_n) = \lambda\, y\,.$$

Since $y \neq 0$, λ is an eigenvalue of T, as we wanted to show. ∎

We are now ready for the *spectral theorem for a compact self-adjoint operator* T. Intuitively, we could hope to be able to construct an orthonormal basis of eigenvectors for T by using Lemma 5.3.3 repeatedly as follows. Start by picking a unit eigenvector v_0 of T associated to the eigenvalue $\lambda_0 = \pm\|T\|$. Next, consider the restriction T_1 of T to $\{v_0\}^\perp$, and pick a unit eigenvector v_1 of T_1 associated to the eigenvalue $\lambda_1 = \pm\|T_1\|$. Then continue this process inductively. There are several technicalities involved in working out the details of this approach. We will follow a more pedestrian route, which also provides more information about T.

Theorem 5.3.4. *Let $T \in \mathcal{K}(H)$ be self-adjoint. Then there exists an orthonormal basis $\mathcal{E}$ for H which consists of eigenvectors of T.*

More precisely, the following facts hold when $T \neq 0$:

(a) *The set L consisting of all nonzero eigenvalues of T is a nonempty, countable subset of the interval $\big[-\|T\|, \|T\|\big]$, containing $\|T\|$ or $-\|T\|$.*

(b) *If L is countably infinite, and $\{\lambda_k : k \in \mathbb{N}\}$ is an enumeration of L, then we have $\lim_{k\to\infty} \lambda_k = 0$.*

(c) *The eigenspace $E_\lambda := \ker(T - \lambda I)$ is finite-dimensional for each $\lambda \in L$.*

(d) *For each $\lambda \in L$, let $\mathcal{E}_\lambda$ be an orthonormal basis for E_λ, and set*

$$\mathcal{E}' := \bigcup_{\lambda \in L} \mathcal{E}_\lambda \,.$$

Then $\mathcal{E}'$ is an orthonormal basis for $\overline{T(H)} = \ker(T)^\perp$, which is countable.

(e) *If $\ker(T) = \{0\}$, set $\mathcal{E}_0 := \varnothing$; otherwise, let $\mathcal{E}_0$ be an orthonormal basis for $\ker(T)$. Then $\mathcal{E} := \mathcal{E}_0 \cup \mathcal{E}'$ is an orthonormal basis for H which consists of eigenvectors of T.*

(f) *Let P_λ denote the orthogonal projection of H on E_λ for each $\lambda \in L$. Then $P_\lambda P_{\lambda'} = 0$ whenever $\lambda \neq \lambda'$ belong to L. Moreover, T has a* spectral decomposition

$$T = \sum_{\lambda \in L} \lambda\, P_\lambda \quad (\textit{w.r.t. operator norm}), \tag{5.3.2}$$

meaning that

- $T = \sum_{\lambda \in L} \lambda\, P_\lambda$ *if L is finite*;
- $\lim_{n\to\infty} \|T - \sum_{k=1}^n \lambda_k\, P_{\lambda_k}\| = 0$ *if L is countably infinite and $\{\lambda_k : k \in \mathbb{N}\}$ is an enumeration of L, as in* (b).

Proof. We can clearly assume that $T \neq 0$.

(a): The set L is a subset of $\mathbb{R}$ by Lemma 5.3.2, which contains $\|T\|$ or $-\|T\|$ by Lemma 5.3.3. If $\lambda \in L$, and v is an associated eigenvector in H_1, we have

$$|\lambda| = |\langle \lambda v, v\rangle| = |\langle T(v), v\rangle| \leq \|T\| \,.$$

Thus, $L \subseteq [-\|T\|, \|T\|]$

To show that L is countable, let $\varepsilon > 0$ and consider the subset of L given by $L_\varepsilon := \{\lambda \in L : |\lambda| \geq \varepsilon\}$. Then L_ε is finite.

Indeed, assume L_ε is nonempty. Then for each $\lambda \in L$, we can pick $v_\lambda \in H_1$ such that $T(v_\lambda) = \lambda v_\lambda$; for $\lambda, \lambda' \in L_\varepsilon, \lambda \neq \lambda'$, we then have $\lambda v_\lambda \perp \lambda' v_{\lambda'}$ by Lemma 5.3.2, so we get

$$\|T(v_\lambda) - T(v_{\lambda'})\|^2 = \|\lambda v_\lambda - \lambda' v_{\lambda'}\|^2 = |\lambda|^2 + |\lambda'|^2 \geq 2\varepsilon^2 .$$

Hence, if L_ε was infinite, we could find a sequence in H_1 which T maps into a sequence with no convergent subsequence, contradicting the compactness of T. Thus, L_ε is finite.

Now, since $L = \bigcup_{n\in\mathbb{N}} L_{1/n}$, it follows that L is countable.

(b): Assume L is countably infinite and $\{\lambda_k : k \in \mathbb{N}\}$ is an enumeration of L. Let $\varepsilon > 0$ be given. Then, as in (a), we get that the set $K := \{k \in \mathbb{N} : |\lambda_k| \geq \varepsilon\}$ is finite. So there exists $N \in \mathbb{N}$ such that $K \subseteq \{1, \ldots, N\}$. For every $k \geq N + 1$, we then have $|\lambda_k| < \varepsilon$. This shows that $\lim_{k\to\infty} \lambda_k = 0$.

(c): This is a consequence of Lemma 5.3.1.

(d): We first remark that since T is self-adjoint, we have

$$\overline{T(H)} = \overline{T^*(H)} = (\ker T)^\perp.$$

Next, it follows from Lemma 5.3.2 that $\mathcal{E}_\lambda \perp \mathcal{E}_{\lambda'}$ whenever $\lambda \neq \lambda'$ belong to L. So it is clear that $\mathcal{E}'$ is an orthonormal set in H, which is countable since each $\mathcal{E}_\lambda$ is finite and L is countable. Hence, $\mathcal{E}'$ is a countable orthonormal basis for $M := \overline{\mathrm{Span}\,(\mathcal{E}')}$, and it remains only to show that $M = \ker(T)^\perp$, i.e., that $M^\perp = \ker(T)$.

- $\ker(T) \subseteq M^\perp$: Assume $y \in \ker(T)$. Then for each $\lambda \in L$ and $v \in \mathcal{E}_\lambda$, we have

  $$\lambda \langle v, y\rangle = \langle T(v), y\rangle = \langle v, T(y)\rangle = \langle v, 0\rangle = 0 .$$

  Since $\lambda \neq 0$, this shows that $y \in (\mathcal{E}')^\perp = M^\perp$.

- $M^\perp \subseteq \ker(T)$: It is easy to check that M is invariant under T. Hence, $M^\perp$ is invariant under $T^* = T$ (cf. Exercise 4.19). We may therefore consider the restriction S of T to $M^\perp$. Then $S \in \mathcal{K}(M^\perp)$: if not, then there would exist a bounded sequence in $M^\perp$, hence in H, which is mapped by S, hence by T, to a sequence with no convergent subsequence, contradicting the compactness of T. Moreover, S is self-adjoint (this is an easy exercise).

  Now, assume that $S \neq 0$. Then Lemma 5.3.3 gives that S has an nonzero eigenvalue μ. This implies that μ is a nonzero eigenvalue of T, i.e., $\mu \in L$. But if $v \in M^\perp$ is an eigenvector for S associated with μ, we then have that $v \in E_\mu \subseteq M$, so $v \in M \cap M^\perp = \{0\}$, contradicting that $v \neq 0$ (since v is an eigenvector).

This means that S has to be 0. Thus we get $T(y) = S(y) = 0$ for all $y \in M^\perp$, as desired.

(e): If $\ker(T) = \{0\}$, then we get from (d) that $\mathcal{E} = \mathcal{E}'$ is an orthonormal basis for $\ker(T)^\perp = \{0\}^\perp = H$. If $\ker(T) \neq \{0\}$, then we have $\mathcal{E}_0 \subseteq \ker(T)$ and $\mathcal{E}' \subseteq \ker(T)^\perp$, so it is clear that $\mathcal{E}$ is an orthonormal set. Moreover, we have that

$$H = \overline{\operatorname{Span}(\mathcal{E})}.$$

Indeed, let $x \in H$. Then we may write

$$x = x_M + x_{M^\perp},$$

where $x_M \in M = \overline{\operatorname{Span}(\mathcal{E}')}$ and $x_{M^\perp} \in M^\perp = \ker(T) = \overline{\operatorname{Span}(\mathcal{E}_0)}$. So we may choose $\{x_n\}_{n\in\mathbb{N}} \subseteq \operatorname{Span}(\mathcal{E}')$ and $\{y_n\}_{n\in\mathbb{N}} \subseteq \operatorname{Span}(\mathcal{E}_0)$ such that $\lim_{n\to\infty} x_n = x_M$ and $\lim_{n\to\infty} y_n = x_{M^\perp}$. This gives that

$$\lim_{n\to\infty}(x_n + y_n) = x_M + x_{M^\perp} = x.$$

Hence, $x \in \overline{\operatorname{Span}(\mathcal{E})}$. This shows that $\mathcal{E}$ is an orthonormal basis for H.

(f): The first assertion follows readily from the fact that $E_\lambda \perp E_{\lambda'}$ whenever $\lambda \neq \lambda'$, cf. Lemma 5.3.2. Next, we consider the case where L is countably infinite and $\{\lambda_k : k \in \mathbb{N}\}$ is an enumeration of L, leaving the easier case where L is finite to the reader.

For each $k \in \mathbb{N}$, set $n_k := \dim(E_{\lambda_k}) < \infty$, and let $\{v_{k,1}, \ldots, v_{k,n_k}\}$ be an enumeration of $\mathcal{E}'_{\lambda_k}$. Then we have

$$\mathcal{E}' = \bigcup_{k\in\mathbb{N}} \mathcal{E}'_{\lambda_k} = \{v_{k,l} : k \in \mathbb{N}, 1 \leq l \leq n_k\}.$$

Consider $x \in H$. Since $T(x) \in \overline{T(H)}$ and $\mathcal{E}'$ is an orthonormal basis for $\overline{T(H)}$, we get from Corollary 4.2.11 that

$$\begin{aligned} T(x) &= \lim_{m\to\infty} \sum_{k=1}^{m} \sum_{l=1}^{n_k} \langle T(x), v_{k,l}\rangle\, v_{k,l} = \lim_{m\to\infty} \sum_{k=1}^{m} \sum_{l=1}^{n_k} \langle x, T(v_{k,l})\rangle\, v_{k,l} \\ &= \lim_{m\to\infty} \sum_{k=1}^{m} \lambda_k \Big(\sum_{l=1}^{n_k} \langle x, v_{k,l}\rangle\, v_{k,l}\Big) = \sum_{k=1}^{\infty} \lambda_k\, P_{\lambda_k}(x). \end{aligned}$$

Let now $\varepsilon > 0$. We have to show that there exists $N \in \mathbb{N}$ such that $\|T - \sum_{k=1}^{n} \lambda_k P_{\lambda_k}\| \leq \varepsilon$ for all $n \geq N$.

Using (b), we can choose $N \in \mathbb{N}$ such that $|\lambda_k| < \varepsilon$ for all $k > N$. Then for all $n \geq N$ and all $x \in H$, using continuity of the norm in H and Pythagoras' identity, we get

$$\Big\|\Big(T - \sum_{k=1}^{n} \lambda_k \, P_{\lambda_k}\Big)(x)\Big\|^2 = \Big\| \sum_{k=n+1}^{\infty} \lambda_k \, P_{\lambda_k}(x)\Big\|^2 = \sum_{k=n+1}^{\infty} |\lambda_k|^2 \, \|P_{\lambda_k}(x)\|^2$$
$$\leq \varepsilon^2 \sum_{k=n+1}^{\infty} \|P_{\lambda_k}(x)\|^2 \leq \varepsilon^2 \, \|x\|^2$$

and the assertion follows.

■

Remark 5.3.5. Let us say that an operator $T \in \mathcal{B}(H)$ is *diagonalizable* if there exists an orthonormal basis for H whose elements are eigenvectors for T. The spectral theorem says that T is diagonalizable if T is compact and self-adjoint. A more precise statement is as follows. We recall that $T \in \mathcal{B}(H)$ is called *normal* if T^* commutes with T.

Assume that $T \in \mathcal{K}(H)$. If $\mathbb{F} = \mathbb{R}$, then T is diagonalizable if and only if T is self-adjoint. On the other hand, if $\mathbb{F} = \mathbb{C}$, then T is diagonalizable if and only if T is normal.

We leave the proof to the reader (cf. Exercises 5.13 and 5.14).

As a corollary of the spectral theorem, an analogue of the singular value decomposition for matrices may be obtained for compact operators.

Indeed, let $S \in \mathcal{K}(H)$, $S \neq 0$. Then $T := S^*S$ is self-adjoint and compact, and $T \neq 0$ (as $\|T\| = \|S^*S\| = \|S\|^2 \neq 0$). Hence, the spectral theorem gives that we may find a countable orthonormal basis $\{v_j\}_{j \in N}$ for $\overline{T(H)} = \ker(T)^\perp = \ker(S^*S)^\perp = \ker(S)^\perp$ consisting of eigenvectors for T. For each $j \in N$, let μ_j denote the eigenvalue of T associated with v_j. Note that

$$\mu_j = \Big\langle \mu_j \, v_j, v_j \Big\rangle = \Big\langle T(v_j), v_j \Big\rangle = \Big\langle S(v_j), S(v_j) \Big\rangle = \|S(v_j)\|^2 \, \geq 0$$

for every $j \in N$. Since each μ_j is nonzero, we get that all μ_j's are positive. For each $j \in N$, set

$$\sigma_j := \sqrt{\mu_j} \, > \, 0 \quad \text{and} \quad u_j := \frac{1}{\sigma_j} S(v_j) \, .$$

The σ_j's are called the *singular values of* S. For all $j, k \in N$ we have

$$\begin{aligned}\left\langle u_j, u_k \right\rangle &= \frac{1}{\sigma_j \sigma_k} \left\langle S(v_j), S(v_k) \right\rangle = \frac{1}{\sigma_j \sigma_k} \left\langle T(v_j), v_k \right\rangle \\ &= \frac{\mu_j}{\sigma_j \sigma_k} \left\langle v_j, v_k \right\rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise,} \end{cases}\end{aligned}$$

so $\{u_j : j \in N\}$ is an orthonormal set in the range of S. Further, we have the following decomposition of S:

$$S(x) = \sum_{j \in N} \sigma_j \left\langle x, v_j \right\rangle u_j \quad \text{for all } x \in H\,. \tag{5.3.3}$$

Indeed, let $x \in H$ and set $M := \overline{T(H)}$, so $M^\perp = \ker(S)$.
With $z := x - P_M(x) \in M^\perp$, we get that

$$x \;=\; P_M(x) + z \;=\; \sum_{j \in N} \langle x, v_j \rangle\, v_j \;+\; z\,,$$

so

$$S(x) \;=\; \sum_{j \in N} \langle x, v_j \rangle\, S(v_j) \;+\; S(z) \;=\; \sum_{j \in N} \sigma_j \left\langle x, v_j \right\rangle u_j,$$

as asserted in (5.3.3).

It readily follows that $\{u_j : j \in N\}$ is an orthonormal basis for $\overline{S(H)}$. Finally we remark that the spectral theorem also gives that $\sigma_j = \sqrt{\mu_j} \to 0$ as $j \to \infty$ when N is countably infinite, and that

$$\|S\| = \|T\|^{1/2} = \max\{\, \mu_j : j \in N\}^{1/2} = \max\{\, \sigma_j : j \in N\}\,.$$

5.4 Application: The Fredholm Alternative

A useful application of linear algebra, and one of its original motivation, is the study of systems of linear equations, i.e., of equations of the type $A\,x = b$, where $A \in M_{m\times n}(\mathbb{F})$, $b \in \mathbb{F}^m$ and the (unknown) vector x belongs to $\mathbb{F}^n$. More generally, one may consider equations of the form

$$T(v) = w \tag{5.4.1}$$

where V, W are vector spaces (over $\mathbb{F}$), $T \in \mathcal{L}(V, W)$, $w \in W$ and the (unknown) vector x belongs to V. Whether such an equation is consistent, i.e., has some solution(s), relies on whether w lies in the range of T, in which case it follows readily that the solution set of (5.4.1) is given by

$$v_0 + \ker(T) := \left\{ v_0 + u \mid u \in \ker(T) \right\} \tag{5.4.2}$$

where $v_0 \in V$ is any vector satisfying (5.4.1), i.e., such that $T(v_0) = w$.

In the rest of this section, we consider the case where $V = W = H$ is a Hilbert space ($\neq \{0\}$), and $T \in \mathcal{B}(H)$. We can then exploit the relationship between the fundamental subspaces of T and T^*, cf. Proposition 4.3.8.

For example, using that $\overline{T(H)} = \ker(T^*)^\perp$, we get that if T has closed range (i.e., $T(H)$ is closed), then (5.4.1) will be consistent if and only if w is orthogonal to $\ker(T^*)$.

In particular, if T has closed range and $\ker(T^*) = \{0\}$ (i.e., T^* is one-to-one), then T must be surjective, hence (5.4.1) is consistent for all $w \in H$. Similarly, if T^* has closed range and $\ker(T) = \{0\}$, then it follows that T^* is surjective, so the equation $T^*(v') = w'$ is consistent for all $w' \in H$.

On the other hand, if it happens that T is surjective, then we get that $\ker(T^*) = \{0\}$, hence that the equation $T^*(v') = w'$ will have either no solution or a unique solution. Similarly, if T^* is surjective, then $\ker(T^*) = \{0\}$, and (5.4.1) will have either no solution or a unique solution.

A problem is that many bounded operators do not have a closed range. Moreover, in general, it may be a difficult task to decide whether the range of some given $T \in \mathcal{B}(H)$ is closed or not. However, we note that if $T \in \mathcal{B}(H)$ has finite-rank, then it has closed range (as $T(H)$ is finite-dimensional). In the case where H is finite-dimensional, much more can be said. The following terminology will be useful.

Definition 5.4.1. An operator $F \in \mathcal{B}(H)$ is said to satisfy the *Fredholm alternative* if one of the following two (mutually exclusive) situations occurs:

(a) $\ker(F) = \ker(F^*) = \{0\}$, and the equations $F(v) = w$, $F^*(v') = w'$ have both a unique solution for all $w, w' \in H$;

(b) $1 \leq \dim(\ker(F)) = \dim(\ker(F^*)) < \infty$, the equation $F(v) = w$ is consistent if and only if $w \in \ker(F^*)^\perp$, and the equation $F^*(v') = w'$ is consistent if and only if $w' \in \ker(F)^\perp$.

Example 5.4.2. Assume that H is finite-dimensional and $F \in \mathcal{B}(H)$, i.e., $F \in \mathcal{L}(H)$. Then F satisfies the Fredholm alternative.

The crux is that we have $\dim(\ker(F^*)) = \dim(\ker(F))$. To show this, we use the formula

$$\dim(M) + \dim(M^\perp) = \dim(H),$$

which is easily verified for any subspace M of H, and the dimension formula for F. We get that

$$\dim(\ker(F^*)) = \dim(F(H)^\perp) = \dim(H) - \dim(F(H)) = \dim(\ker(F)).$$

Combining this fact with our previous observations in this section, it is straightforward to deduce that either (a) or (b) in Definition 5.4.1 holds.

An important class of bounded operators satisfying the Fredholm alternative consists of operators of the form $F = T - \mu I$, where T is a compact operator on H and $\mu \in \mathbb{F} \setminus \{0\}$. In the special case where $T = T_K$ is an integral operator, cf. Example 5.2.11, an equation of the form $(T_K - \mu I)(f) = g$, i.e., $T_K(f) - \mu f = g$, is often called a *Fredholm integral equation of the second kind.*[1]

Consider $T \in \mathcal{K}(H)$ and $\mu \in \mathbb{F} \setminus \{0\}$. Then it can be shown that the following facts hold:

(i) $T - \mu I$ has closed range;

(ii) $\dim(\ker(T - \mu I)) = \dim(\ker((T - \mu I)^*)) < \infty$.

Since T^* is compact, we also get that $T^* - \overline{\mu} I = (T - \mu I)^*$ has closed range. Using these properties, and the general principles outlined before, one readily arrives at the conclusion that $F = T - \mu I$ satisfies the Fredholm alternative, as asserted above. We don't have time in this course to prove that (i) and (ii) hold. Instead, we will illustrate how the spectral theorem for compact self-adjoint operators can be applied to give a direct proof of the following:

Theorem 5.4.3. *Assume $T \in \mathcal{K}(H)$ is self-adjoint and $\mu \in \mathbb{F} \setminus \{0\}$. Then $F = T - \mu I$ satisfies the Fredholm alternative.*

Proof. Assume first that μ is not an eigenvalue of T, i.e., $\ker(T - \mu I) = \{0\}$. Then the spectral theorem implies that the equation $(T - \mu I)(x) = y$ has a unique solution for all $y \in H$. (You are asked to check this in Exercise 5.9.) Thus, $F = T - \mu I$ is surjective, and this implies that $\ker(F^*) = \ker(T - \overline{\mu} I) = \{0\}$, i.e., $\overline{\mu}$ is not an eigenvalue of T. Arguing as above, we get that the equation $(T - \overline{\mu} I)(x') = y'$, i.e., $(T - \mu I)^*(x') = y'$ has a unique solution for all $y' \in H$. This shows that (a) in Definition 5.4.1 holds in this case.

Next, assume that μ is an eigenvalue of T, i.e., $\ker(T - \mu I) \neq \{0\}$. Then $\mu \in \mathbb{R}$, so $F^* = F$. Moreover, as $\mu \neq 0$, we have that $T \neq 0$, and the spectral theorem tells us that $1 \leq \dim(\ker(F)) = \dim(\ker(T - \mu I)) < \infty$. Hence, to show that (b) in Definition 5.4.1 holds, it remains only to prove that the

[1] Such equations, and Fredholm integral equations of the first kind (i.e., equations of the form $T_K(f) = g$), were studied by I. Fredholm at the beginning of the 20th century. They arise in some practical problems in signal theory and in physics.

equation $F(x) = y$ is consistent if and only if $y \in \ker(F)^{\perp}$. This means that we have to prove that the equation

$$T(x) - \mu x = y \tag{5.4.3}$$

is consistent if and only if $\langle y, z\rangle = 0$ for all $z \in E_\mu := \ker(T - \mu I)$.

To prove this, let $\mathcal{E}' = \{u_j\}_{j\in J}$ be an enumeration of the orthonormal basis for $\overline{T(H)}$ obtained in the spectral theorem for T, and let $\mu_j \in \mathbb{R} \setminus \{0\}$ denote the eigenvalue of T corresponding to each u_j.

Since H is the direct sum of $\ker(T)$ and $\ker(T)^{\perp} = \overline{T(H)}$, we may write $y \in H$ as

$$y = y_0 + \sum_{j\in J} \langle y, u_j\rangle\, u_j,$$

where y_0 denote the orthogonal projection of y onto $\ker(T)$. Likewise, we may assume that the (unknown) vector x in equation (5.4.3) is written as

$$x = x_0 + \sum_{j\in J} c_j\, u_j,$$

where $x_0 \in \ker(T)$ and $\{c_j\}_{j\in J} \in \ell^2(J)$ are to be determined, if possible. Plugging this into equation (5.4.3), we get the equivalent equation

$$-\mu x_0 + \sum_{j\in J} (\mu_j - \mu)\, c_j\, u_j = y_0 + \sum_{j\in J} \langle y, u_j\rangle\, u_j.$$

Clearly, we can set $x_0 := (-1/\mu)\, y_0$, and equation (5.4.3) is then consistent if and only if the sequence $\{c_j\}_{j\in J} \in \ell^2(J)$ can be chosen so that

$$(\mu_j - \mu)\, c_j = \langle y, u_j\rangle \quad \text{for all } j \in J. \tag{5.4.4}$$

Now, as μ is a nonzero eigenvalue of T, we have that $\mu = \mu_k$ for some $k \in J$. Let $u_{j_1}, \ldots, u_{j_n}$ denote the vectors in $\mathcal{E}'$ giving an orthonormal basis for $E_\mu = E_{\mu_k}$. If $j \notin \{j_1, \ldots, j_n\}$, we have $\mu_j \neq \mu$, so

$$c_j := \frac{1}{\mu_j - \mu}\, \langle y, u_j\rangle$$

will satisfy (5.4.4) for every such j.

On the other hand, if $j \in \{j_1, \ldots, j_n\}$, we have $\mu_j - \mu = 0$. Hence, (5.4.4) will be satisfied for $j = j_1, \ldots, j_n$ if and only if we have $\langle y, u_j\rangle = 0$ for $j = j_1, \ldots, j_n$, i.e., if and only if $\langle y, z\rangle = 0$ for all $z \in E_\mu$. Moreover, when this condition holds, we can choose $c_{j_1}, \ldots, c_{j_n}$ freely and, regardless of this choice, the constructed sequence $\{c_j\}_{j\in J}$ is easily seen to belong to $\ell^2(J)$ (exercise: check this!), meaning that the associated vector x gives a solution to (5.4.3). Thus, we have proved the desired equivalence. ■

5.5 Exercises

Exercise 5.1. Let X, Y, Z denote normed spaces over $\mathbb{F}$. Consider $\lambda \in \mathbb{F}$, $T, T' \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$, so $ST \in \mathcal{B}(X, Z)$.

a) Show that $\lambda T + T' \in \mathcal{K}(X, Y)$ if $T, T' \in \mathcal{K}(X, Y)$.

b) Show that $ST \in \mathcal{K}(X, Z)$ if $T \in \mathcal{K}(X, Y)$.

c) Show that $ST \in \mathcal{K}(X, Z)$ if $S \in \mathcal{K}(Y, Z)$.

d) Set $\mathcal{K}(X) = \mathcal{K}(X, X)$. Deduce that
$ST \in \mathcal{K}(X)$ if $S \in \mathcal{B}(X)$ and $T \in \mathcal{K}(X)$, or if $S \in \mathcal{K}(X)$ and $T \in \mathcal{B}(X)$.

Exercise 5.2. Let $X = \ell^p(\mathbb{N})$, $\lambda \in \ell^\infty(\mathbb{N})$, and $M_\lambda \in \mathcal{B}(X)$ be the associated multiplication operator, cf. Example 5.1.7.

Show that $\lambda \in c_0(\mathbb{N})$ if M_λ is compact.

(It therefore follows that M_λ is compact if and only if $\lambda \in c_0(\mathbb{N})$.)

Exercise 5.3. Let X be a normed space, H be a Hilbert space, and let $T \in \mathcal{K}(X, H)$. Show that $\overline{T(X)}$ is separable.

Exercise 5.4. Let H be an infinite-dimensional Hilbert space and let $T \in \mathcal{K}(H)$. Show that $\left\langle T(u_n), u_n \right\rangle \to 0$ as $n \to \infty$ whenever $\{u_n\}_{n\in\mathbb{N}}$ is an orthonormal sequence in H.

Exercise 5.5. Let H be a Hilbert space and let $P \in \mathcal{B}(H)$ be a projection (i.e. $P^2 = P$). Show that P has finite-rank if (and only if) P is compact.

Exercise 5.6. Let H be a separable Hilbert space, $H \neq \{0\}$.

a) Show that $\mathcal{F}(H) \subseteq \mathcal{HS}(H)$, and that $\mathcal{F}(H)$ is dense in $\mathcal{HS}(H)$ w.r.t. $\|\cdot\|_2$.

b) Assume that $T \in \mathcal{HS}(H)$ and $S \in \mathcal{B}(H)$. Show that both ST and TS belong to $\mathcal{HS}(H)$, and that we have

$$\|ST\|_2 \leq \|S\| \, \|T\|_2 \,, \quad \|TS\| \leq \|T\|_2 \, \|S\| \,.$$

c) Let $\mathcal{B} = \{u_j\}_{j\in J}$ be an orthonormal basis for H, where $J = \{1, \ldots, n\}$ if $\dim(H) = n < \infty$, while $J = \mathbb{N}$ otherwise.

For $T, T' \in \mathcal{HS}(H)$, set

$$\left\langle T, T' \right\rangle_2 := \sum_{j\in J} \left\langle T(u_j), T'(u_j) \right\rangle .$$

Show that this gives a well-defined inner product on $\mathcal{HS}(H)$, and check that the associated norm is the Hilbert-Schmidt norm $\|\cdot\|_2$.

d) Show that $\mathcal{HS}(H)$ is complete w.r.t. $\|\cdot\|_2$, so that $\mathcal{HS}(H)$ is a Hilbert space w.r.t. the inner product from c).

Exercise 5.7. Let $H = L^2(\mathbb{R}, \mathcal{A}, \mu)$ where $\mathcal{A}$ denote all Lebesgue measurable subsets of $\mathbb{R}$ and μ is the Lebesgue measure. For which $f \in \mathcal{L}^\infty$ is the multiplication operator $M_f \in \mathcal{B}(H)$ compact ?

Exercise 5.8. Let H be a Hilbert space, $T \in \mathcal{K}(H)$ and $\lambda \in \mathbb{F}, \lambda \neq 0$. Assume that there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ of unit vectors in H such that $\|T(x_n) - \lambda\, x_n\| \to 0$ as $n \to \infty$. Show that λ is an eigenvalue of T.

Exercise 5.9. Let H be a Hilbert space, and let $T \in \mathcal{K}(H)$ be self-adjoint. Assume $\mu \in \mathbb{F}, \mu \neq 0$ is *not* an eigenvalue of T, i.e. $T - \mu I_H$ is injective.

Let $y \in H$, let $\mathcal{E}' = \{u_j\}_{j\in J}$ be an enumeration of the orthonormal basis for $M = \overline{T(H)}$ obtained in the spectral theorem for T, and let $\mu_j \neq 0$ denote the eigenvalue of T corresponding to u_j.

a) Show that the series

$$\sum_{j\in J} \frac{\langle y, u_j\rangle}{\mu_j - \mu}\, u_j$$

converges to some $h \in H$.

b) Set $z := y - P_M(y)$ and $x := h - \frac{1}{\mu}\, z$. Show that $(T - \mu I_H)(x) = y$.

c) Deduce that $T - \mu I_H$ is surjective (hence that it is bijective).

Exercise 5.10. Consider $H = L^2([-\pi, \pi])$ (with respect to the normalized Lebesgue measure). Let $g \in C([-\pi, \pi])$ be periodic, i.e. satisfies that $g(-\pi) = g(\pi)$, and extend g to a periodic function $\tilde{g}$ on $\mathbb{R}$ with period 2π. Define $G : [-\pi, \pi] \times [-\pi, \pi] \to \mathbb{C}$ by $G(s, t) = \tilde{g}(s - t)$.

a) Check that G is continuous, so that the associated integral operator T_G belongs to $\mathcal{HS}(H)$ (hence is compact).

c) Decide when T_G is self-adjoint.

b) Let $k \in \mathbb{Z}$ and recall that $e_k(t) = e^{ikt}$ for all $t \in [-\pi, \pi]$. Check that e_k is an eigenvector for the operator T_G. Deduce that T_G is diagonalizable (with respect to $\{e_k\}_{k\in\mathbb{Z}}$).

c) Show that $\|T_G\|_2 = \|g\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |g(t)|^2 \,\mathrm{d}t\right)^{1/2}$.

Exercise 5.11. Consider $H = L^2([0,1])$ (with respect to Lebesgue measure) and the integral operator $T_K \in \mathcal{B}(H)$ associated with $K(s,t) := \min(s,t)$ for all (s,t) in $[0,1] \times [0,1]$, cf. Example 5.2.11.

a) Explain why T_K is self-adjoint and compact. Then check that the set $\mathcal{U} := \{[u_n] : n \in \mathbb{N}\}$, where

$$u_n(t) := \sqrt{2}\, \sin\Big((n - \frac{1}{2})\pi\, t\Big) \quad \text{for all } t \in [0,1],\ n \in \mathbb{N}\,,$$

is an orthonormal set of eigenvectors for T_K.

b) It can be shown that $\mathcal{U}$ is an orthonormal basis for H. Is it possible to deduce this from a) and the spectral theorem for T_K ?

Exercise 5.12. Let $S,\, T \in \mathcal{B}(H)$.

a) Assume there exists an orthonormal basis for H whose elements are eigenvectors for both S and T. Check that S commutes with T.

b) Assume S and T are compact and self-adjoint, and that S commutes with T. Show that there exists an orthonormal basis for H whose elements are eigenvectors for both S and T.

Hint: Start by considering an eigenvalue λ of T and study how S acts on the corresponding eigenspace E^T_λ.

Exercise 5.13. Assume H is a Hilbert space over $\mathbb{R}$, and let $T \in \mathcal{B}(H)$.

a) Assume that T is diagonalizable (as defined in Remark 5.3.5). Check that T is self-adjoint.

b) Let T be compact. Deduce that T is diagonalizable if and only if T is self-adjoint.

Exercise 5.14. Assume H is a Hilbert space over $\mathbb{C}$, and let $T \in \mathcal{B}(H)$.

a) Assume that T is diagonalizable (as defined in Remark 5.3.5). Check that T is normal.

b) Show that T is normal if and only if $\mathrm{Re}(T)$ and $\mathrm{Im}(T)$ commutes with each other.

c) Let T be compact. Show that T is diagonalizable if and only if T is normal.

Hint: The implication ($\Rightarrow$) follows from a). For ($\Leftarrow$), use b) and Exercise 5.12 b).

Exercise 5.15. Let H be a separable Hilbert space with a countably infinite orthonormal basis $\mathcal{B} = \{v_j\}_{j\in\mathbb{N}}$. Let $\{\mu_j\}_{j\in\mathbb{N}}$ be a bounded sequence in $\mathbb{F}$ and let $D \in \mathcal{B}(H)$ denote the associated diagonal operator (w.r.t. $\mathcal{B}$).

a) Show that D is compact if and only if $\lim_{j\to\infty} \mu_j = 0$.

(*Note*: If you have looked at Example 5.1.7 and solved Exercise 5.2, this should not be difficult).

b) Show that D is Hilbert-Schmidt if and only if $\{\mu_j\}_{j\in\mathbb{N}} \in \ell^2(\mathbb{N})$, in which case we have $\|D\|_2 = \left(\sum_{j=1}^{\infty} |\mu_j|^2\right)^{1/2}$.

Exercise 5.16. Let H be a separable Hilbert space of infinite dimension and let $T \in \mathcal{K}(H)$ be selfadjoint, $T \neq 0$. Assume that you have found an orthonormal basis $\mathcal{B} = \{v_j\}_{j\in\mathbb{N}}$ for H consisting of eigenvectors for T, and let $\mu_j \in \mathbb{R}$ denote the eigenvalue of T corresponding to each v_j.

a) Show that the sequence $\{\mu_j\}_{j\in\mathbb{N}}$ is bounded, hence that T is the diagonal operator (w.r.t. $\mathcal{B}$) associated with this sequence. Deduce from the previous exercise that $\lim_{j\to\infty} \mu_j = 0$.

b) As in the spectral theorem, set

$$L := \{\lambda \in \mathbb{R} \mid \lambda \text{ is a nonzero eigenvalue of } T\}.$$

Set also

$$\widetilde{L} := \{\lambda \in \mathbb{R} \mid \lambda \text{ is an eigenvalue of } T\},$$

so $L = \widetilde{L} \setminus \{0\}$. Show the following assertions:

(i) $\widetilde{L} = \{\mu_j \mid j \in \mathbb{N}\}$ and $L = \{\mu_j \mid j \in \mathbb{N}, \mu_j \neq 0\}$.

(ii) If $\lambda \in L$ and $N_\lambda := \{j \in \mathbb{N} \mid \mu_j = \lambda\}$, then N_λ is a finite subset of $\mathbb{N}$ and $\{v_j \mid j \in N_\lambda\}$ is an o.n.b. for E_λ.

(iii) If $\mu_j \neq 0$ for all $j \in \mathbb{N}$, then $\ker(T) = \{0\}$.

(iv) If $N_0 := \{j \in \mathbb{N} \mid \mu_j = 0\}$ is nonempty, then $\{v_j \mid j \in N_0\}$ is an o.n.b. for $\ker(T)$.

Exercise 5.17. Let $H = L^2([0, 1])$ (with usual Lebesgue measure) and let $T = M_f$ be the self-adjoint operator in $\mathcal{B}(H)$ given by multiplication with the function $f(t) = t$ on $[0, 1]$, cf. Example 4.4.4.

Show that $T(H)$ is not closed, i.e., that T does not have closed range. Show also that T is not compact.

Exercise 5.18. Let $H = \ell^2(\mathbb{N})$, let $\lambda \in \ell^\infty(\mathbb{N})$ be given by $\lambda(n) = \frac{1}{n}$ for all $n \in \mathbb{N}$, and let $T = M_\lambda \in \mathcal{B}(H)$ denote the associated multiplication operator. Note that T is compact, as follows from Example 5.1.7.

Show that $\overline{T(H)} = H$ and $T(H) \neq H$, so T does not have closed range.

Exercise 5.19. Let H be a Hilbert space and $T \in \mathcal{B}(H)$. Let us say that T is *bounded from below* if there exists some $\alpha > 0$ such that $\alpha \|x\| \leq \|T(x)\|$ for all $x \in H$. For example, T is bounded from below when T is an isometry.

Show that if T is bounded from below, then T has closed range.

Exercise 5.20. Finish the proof of Theorem 5.9 by checking that the sequence $\{c_j\}_{j \in J}$ constructed in the final paragraph (under the assumption that y is orthogonal to E_μ) belongs to $\ell^2(J)$.