Elemente der Funktionalanalysis

Delio Mugnolo

(delio.mugnolo@uni-ulm.de)

What is functional analysis?

One might say that *functional analysis* is the branch of mathematics that tries to capture fundamental properties of common analytical objects by setting them in a more abstract general framework. More specifically, functional analysis deals with infinite dimensional vector spaces and mappings acting between them. Occasionally, finite dimensional vector spaces are also considered. However, functional analytical theory of *finite* dimensional vector spaces usually goes under the name of *linear algebra*. A nice account of linear algebra based on a truly functional analytic approach is given in [**9**].

Functional analysis is essentially a science of the 20$^{\text{th}}$ century. One might argue that its birthday coincides with the development of the notion of "Hilbert space" around 1905 by Erhard Schmidt, a scholar of David Hilbert, and almost simultaneously by Maurice René Fréchet. (Hilbert spaces arose from Hilbert's investigations on concrete integral equations. In fact, a famous factoid of dubious origin reports that Hilbert once asked Richard Courant after a talk "Richard, what exactly is a Hilbert space?"). This opened the doors to the development of completely new methods, aiming at reformulating analytical assertions in an abstract way and then proving relevant results by techniques based, among others, on linear algebra and topology. A beautiful example for this approach is Stone's generalization (obtained in 1937) of the Weierstraß' classical result that each continuous function may be approximated by polynomials, see Theorem 4.32.

Very complete historical *excursi* can be found in [**20**]. However, it can be argued that the history of functional analysis has often been the history of individual scientists and their ideas. Short biographies of some of the most brilliant men and women of functional analysis appear in [**19**]. In the beautiful book [**15**], the early history of functional analysis is shown to be intertwined with the drama of Nazi occupation of Europe.

This manuscript has been developed for the course of introductory functional analysis held during the summer term 2009 at the University of Ulm. Due to the time limitations of the course only selected topics in functional analysis could be treated. I have decided to discuss almost exclusively Hilbert space theory.

As almost all human products, this manuscript is highly unlikely to be free of mistakes, in spite of the careful proof-reading of my assistent Manfred Sauter. Critics and corrections are welcome via e-mail to `delio.mugnolo@uni-ulm.de`.

Ulm, August 2, 2011.

Contents

CHAPTER 1

Metric, normed and Banach spaces

We assume the notion of vector space as well as some basic topological facts as known. In the following we will always consider a field $\mathbb{K}$ as being either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1. *Let X be a vector space over $\mathbb{K}$. A **distance** d on X is a mapping $d : X \times X \to \mathbb{R}_+$ such that for all $x, y, z \in X$*

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

*Then, (X, d) is called a **metric space**.*

Definition 1.2. *Let (X, d) be a metric space.*

*(1) A subset A of X is called **open** if for all $x \in A$ there exists $r > 0$ such that*

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

*It is called **closed** if $X \setminus A$ is open. An element $x \in A$ is called **interior point** of A if there exist $\delta > 0$ and $y \in X$ such that $y \in B_\delta(y) \subset A$.*

(2) We say that a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ **converges to** *$x \in X$, or that x* **is limit of** *$(x_n)_{n\in\mathbb{N}}$, if $\lim_{n\to\infty} d(x_n, x) = 0$. We say that $(x_n)_{n\in\mathbb{N}} \subset X$ is a* **Cauchy sequence** *if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m > N$.*

(3) X is called **complete** *if for each Cauchy sequence $(x_n)_{n\in\mathbb{N}} \subset X$ there exists $x \in X$ such that x is limit of $(x_n)_{n\in\mathbb{N}}$.*

(4) A subset Y of X is called **dense in** *X if each element of X is limit of a suitable sequence $(y_n)_{n\in\mathbb{N}}$ such that $y_n \in Y$ for all $n \in \mathbb{N}$, or equivalently if for all $x \in X$ and all $\epsilon > 0$ there is an $y \in Y$ such that $d(x, y) < \epsilon$.*

Metric spaces were introduced by Maurice René Fréchet in 1906.

Example 1.3. *The Euclidean distance d on the complex vector space $\mathbb{C}^n$ is defined by $d(x, y) := \sqrt{|x_1 - y_1|^2 + \ldots + |x_n - y_n|^2}$.*

Example 1.4. *Let X be a set. Then we can define the* **discrete distance** *by setting $d(x, y) = \delta_{xy}$, where δ is the Kronecker delta, i.e., $d(x, x) = 1$ and $d(x, y) = 0$ if $x \neq y$. In particular, any non-empty set is* metrisable*, i.e., there is always a metric space associated to it.*

The following fundamental result has been proved in 1899 by René Louis Baire.

Theorem 1.5. *Let (X,d) be a complete metric space and $(G_n)_{n\in\mathbb{N}}$ a sequence of nonempty open and dense subsets of X. Then also their intersection is dense in X.*

PROOF. Our aim is to show that for all $x\in X$ and all $\varepsilon>0$ there exists $y\in\bigcap_{n\in\mathbb{N}}G_n$ such that $d(x,y)<\varepsilon$.

Since G_1 is dense in X, we can find $y_1\in G_1\cap B_\varepsilon(x)$. Since both $G_1, B_\varepsilon(x)$ are open, also $G_1\cap B_\varepsilon(x)$ is open, i.e., there exists $\varepsilon_1>0$ such that $B_{\varepsilon_1}(y_1)\subset G_1\cap B_\varepsilon(x)$. Again because G_2 is dense in X, we can find $y_2\in G_2\cap B_{\tilde\varepsilon_1}(y_1)$, where $\tilde\varepsilon_1:=\min\{\frac{\varepsilon_1}{2},1\}$. Since both $G_1, B_{\tilde\varepsilon_1}(x)$ are open, also $G_1\cap B_{\tilde\varepsilon_1}(y_1)$ is open, i.e., there exists $\varepsilon_2>0$ such that $B_{\varepsilon_2}(y_2)\subset G_2\cap B_{\tilde\varepsilon_1}(y_1)$. Clearly, it is possible to extend this construction to obtain sequences $(y_n)_{n\in\mathbb{N}}$ of vectors and $(\varepsilon_n)_{n\in\mathbb{N}}$ of strictly positive numbers such that

$$B_{\varepsilon_{n+1}}(y_{n+1})\subset G_{n+1}\cap B_{\tilde\varepsilon_n}(y_n),\qquad n\in\mathbb{N},$$

where $\tilde\varepsilon_{n+1}:=\min\{\frac{\varepsilon_n}{2},\frac{1}{n}\}$. Take now $m,n\in\mathbb{N}$, $m>n$, and observe that

$$d(y_n,y_m)<\tilde\varepsilon_n<\frac{1}{n},$$

hence $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Because X is complete, $(y_n)_{n\in\mathbb{N}}$ converges to some $y\in X$. Since moreover $y_m\in G_{n+1}\cap B_{\tilde\varepsilon_n}(y_n)\subset B_{\tilde\varepsilon_n}(y_n)$ for all $m>n$, it follows that

$$y\in\overline{B_{\tilde\varepsilon_n}(y_n)}\subset B_{\varepsilon_n}(y_n)\subset G_n,\qquad n\in\mathbb{N},$$

and accordingly $y\in\bigcap_{n\in\mathbb{N}}G_n$. Moreover $y\in B_{\varepsilon_1}(y_1)\subset B_\varepsilon(x)$, i.e., $d(x,y)<\varepsilon$. This concludes the proof. □

Remark 1.6. *One can see that an equivalent formulation of Baire's theorem is the following:*

Let (X,d) be a complete metric space and $(F_n)_{n\in\mathbb{N}}$ a sequence of closed subsets of X such that $\bigcup_{n\in\mathbb{N}}F_n=X$. Then there exist $x\in X$, $\varepsilon>0$ and $n_0\in\mathbb{N}$, such that $B_\varepsilon(x)\subset A_{n_0}$, i.e., at least one set A_{n_0} has an interior point.

A subset U in X is called **nowhere dense** *if $\overline{U}$ has no interior point. Any subset of X that can be written as $\bigcup_{n\in\mathbb{N}}U_n$ for a sequence $(U_n)_{n\in\mathbb{N}}$ of nowhere dense subsets of X is said to be* **of first Baire category**. *A subset of X that is not of first category is said to be* **of second Baire category**. *Now, we may formulate Baire's theorem in yet another way by stating that*

Each nonempty complete metric space is of second Baire category.

One of the first applications of Baire's theorem is to the proof of the following fact – which was actually first shown by Vito Volterra:

No function from $\mathbb{R}$ to $\mathbb{R}$ can be continuous at each rational number and discontinuous at each irrational number.

Exercise 1.7. *1) Show that $\mathbb{Q}$ is of first Baire category in $\mathbb{R}$.*
2) Show that $\mathbb{R}\setminus\mathbb{Q}$ and $\mathbb{R}$ are of second Baire category in $\mathbb{R}$.
3) Show that $\mathbb{R}$ is of first Baire category in $\mathbb{C}$.

A special class of metric space is presented in the following.

Definition 1.8. *Let X be a vector space over a field $\mathbb{K}$. A* ***norm*** *$\|\cdot\|$ on X is a mapping $X\to\mathbb{R}_+$ such that for all $x,y\in X$ and all $\lambda\in\mathbb{K}$*

(1) $\|x\| = 0 \Leftrightarrow x = 0$,
(2) $\|\lambda x\| = |\lambda|\|x\|$, *and*
(3) $\|x + y\| \leq \|x\| + \|y\|$.

Then, $(X, \|\cdot\|)$ *is called a* ***normed vector space****. A complete normed vector space is called a* **Banach space**.

Example 1.9. *1) Consider the Euclidean distance* d *on the complex vector space* $\mathbb{C}^n$*. Then* $\|x\|_2 := d(0, x) := \sqrt{|x_1|^2 + \ldots + |x_n|^2}$ *defines a norm on* $\mathbb{C}^n$.

2) The vector space ℓ^1 *of all summable sequences becomes a normed vector space after introducing* $\|x\|_1 := \sum_{n\in\mathbb{N}} |x_n|$.

3) Both the vector space c_0 *of all sequences that converge to 0 and the vector space* ℓ^∞ *of all bounded sequences become a normed vector spaces after introducing* $\|x\|_\infty := \sup_{n\in\mathbb{N}} |x_n|$.

4) Let Ω *be a* σ*-finite measure space. The Lebesgue space* $L^p(\Omega)$ *of* p*-summable functions on* Ω *is a normed vector space for all* $p \in [1, \infty)$ *with respect to the norm defined by*

$$\|f\|_p := \left(\int_\Omega |f(x)|^p dx \right)^{\frac{1}{p}} .$$

Same holds for the Lebesgue space $L^\infty(\Omega)$ *of essentially bounded functions on* Ω*, normed by*

$$\|f\|_\infty := \operatorname*{ess\,sup}_{x\in\Omega} \|f(x)\|.$$

Exercise 1.10. *Prove that the norms* $\|\cdot\|_2$ *and* $\|\cdot\|_1$ *introduced in Exercise 1.9.(1)-(2), respectively, turn the respective vector spaces into Banach spaces. What about the norm introduced in Exercise 1.9.(3)?*

Remark 1.11. *We observe that* $L^p(\Omega)$ *is a Banach space for all* $p \in [1, \infty)$ *as well as for* $p = \infty$*, see e.g.* [**1**, § 1.3] *for a proof.*

Example 1.12. *Let* $(X_1, d_1), (X_2, d_2)$ *be metric spaces.*

(1) Prove that

$$d((x_1, y_1), (x_2, y_2)) := d_1(x_1, x_2) + d_2(y_1, y_2), \qquad x_1, x_2 \in X_1,\ y_1, y_2 \in X_2,$$

defines a metric on $X_1 \times X_2$.

(2) Show that if X_1, X_2 *are normed (resp., complete), then also* $X_1 \times X_2$ *is normed (resp., complete) with respect to the norm defined by*

$$\|(x_1, x_2)\|_X := \|x_1\|_{X_1} + \|x_2\|_{X_2}, \qquad x_1 \in X_1,\ x_2 \in X_2.$$

Definition 1.13. *Two norms* $\|\cdot\|$ *and* $|\!|\!|\cdot|\!|\!|$ *on* X *are called* **equivalent** *if there exists* $c > 0$ *such that*

$$c\|x\| \leq |\!|\!|x|\!|\!| \leq \frac{1}{c}\|x\| \qquad \text{for all } x \in X.$$

Remark 1.14. *Observe that in particular if a sequence converges with respect to a norm* $\|\cdot\|$*, then it converges with respect to any further norm that is equivalent to* $\|\cdot\|$*. Therefore, if a vector space is complete with respect to a norm* $\|\cdot\|$*, then it is complete also with respect to any further norm that is equivalent to* $\|\cdot\|$.

Exercise 1.15. *Consider the space $C([0,1])$ of all continuous functions on $[0,1]$ and define the* **sup-norm** *by*

$$\|f\|_\infty := \max_{x\in[0,1]} |f(x)| \quad \textit{and} \quad \|f\|_1 := \int_0^1 |f(x)|dx.$$

Show that

(1) $C([0,1])$ endowed with $\|\cdot\|_\infty$ becomes a Banach space;
(2) $\|\cdot\|_1$ defines a norm on $C([0,1])$, too;
(3) if a sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $C([0,1])$ converges with respect to $\|\cdot\|_\infty$, then it also converges with respect to $\|\cdot\|_1$;
(4) define the functions $f_k(t) := t^k$ for $k \in \mathbb{N}$, and $f(t) := 0$ ($t \in [0,1]$). Then the sequence $(f_k)_{k\in\mathbb{N}}$ converges to f with respect to $\|\cdot\|_1$;
(5) the norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are not equivalent (Hint: use (4));
(6) the norm $|||\cdot|||$ defined by

$$|||f||| := \|f\|_\infty + \|f\|_1$$

is equivalent to $\|\cdot\|_\infty$.

CHAPTER 2

Operators

A central notion in functional analysis is that of operator.

Definition 2.1. *Let X, Y be vector spaces. A mapping T from X to Y is usually called an* **operator***, and we usually write $Tx := T(x)$.*

We will devote our attention almost exclusively to **linear operators**, i.e., to operators that satisfy

$$T(x+y) = Tx + Ty \qquad \text{and} \qquad T(\lambda x) = \lambda Tx \qquad \text{for all } x, y \in H,\ \lambda \in \mathbb{K}.$$

Clearly, each $n \times m$-matrix is a linear operator mapping $\mathbb{K}^m$ to $\mathbb{K}^n$.

Definition 2.2. *Let X, Y be normed vector spaces. A linear operator T is called* **bounded***, and we write $T \in \mathcal{L}(X, Y)$, if*

$$\|T\|_{\mathcal{L}(X,Y)} := \sup_{\|x\|_X \le 1} \|Tx\|_Y < \infty. \tag{2.1}$$

The **null space** *and* **range** *of T are defined by*

$$\mathrm{Ker} T := \{x \in X : Tx = 0\}$$

and

$$\mathrm{Ran} T := \{Tx \in Y : x \in X\},$$

respectively.

We will mostly write $\|T\|$ instead of $\|T\|_{\mathcal{L}(X,Y)}$.

Remark 2.3. *Let X, Y be normed vector spaces and $T \in \mathcal{L}(X, Y)$. By definition, $\|Tx\|_Y \le \|T\|\|x\|_X$ for all $x \in X$. Let additionally Z be a normed vector space and $S \in \mathcal{L}(Y, Z)$. Then we obtain $\|STx\|_Z \le \|S\|\|Tx\|_Y \le \|S\|\|T\|\|x\|_X$, i.e., $ST \in \mathcal{L}(X, Z)$ and*

$$\|ST\| \le \|S\|\|T\|.$$

One says that the operator norm is **submultiplicative***.*

Remark 2.4. *Let X, Y be normed vector spaces. It is easy to see that a linear operator from X to Y is bounded if and only if*

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} < \infty. \tag{2.2}$$

Exercise 2.5. *Let X be a normed vector space and T be a bounded liner operator on X. Show that the* **square** *of T defined by*

$$T^2 : x \mapsto TTx, \qquad x \in X,$$

*and more generally the n^{th}***-power** *of T defined by*

$$T^n : x \mapsto \underbrace{T \dots T}_{n \text{ times}} x, \qquad x \in X,$$

define bounded linear operators on X for all $n \in \mathbb{N}$.

Example 2.6. *Define operators T, S acting on a space of sequences (to be made precise later) by*

$$T(x_n)_{n\in\mathbb{N}} := (0, x_1, x_2, \dots)$$

and

$$S(x_n)_{n\in\mathbb{N}} := (x_2, x_3, \dots).$$

It is clear that T, S are linear. They are called the **right** *and* **left shift***, respectively.*

One also sees that $STx = x$ for each sequence x. We can consider both T and S as operators acting on several sequence spaces, including $\ell^1, \ell^2, \ell^\infty$.

Example 2.7. *Let X_1, X_2 be normed spaces and consider the set $X := X_1 \times X_2$. It is easy to show that*

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_X := \|x\|_{X_1} + \|y\|_{X_2}, \qquad x \in X_1,\ y \in X_2,$$

defines a norm on X – with respect to which X is in fact a Banach space provided that X_1, X_2 are Banach spaces.

Now, consider four operators $A \in \mathcal{L}(X_1)$, $B \in \mathcal{L}(X_2, X_1)$, $C \in \mathcal{L}(X_1, X_2)$, and $D \in \mathcal{L}(X_2)$. Then, the mapping

$$X \ni \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} Ax + By \\ Cx + Dy \end{pmatrix}$$

defines a bounded linear operator on $X \times Y$ which is usually denoted by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

in analogy with the rule of composition between a matrix and a vector in linear algebra.

Example 2.8. *Erik Ivar Fredholm preformed a thoroug study of a class of so-called integral equations in several investigations at the turn of the 20^{th} century. With this purpose, he followed an abstract approach and introduced what is now called the* **Fredholm operator** *F_k by*

$$(Tf)(x) := \int_0^1 k(x,y) f(y) dy, \qquad f \in C([0,1]),\ x \in [0,1].$$

If $k \in C([0,1] \times [0,1])$, then T is linear and bounded and its norm can be estimated from above by $\|k\|_\infty := \max_{0 \le x,y \le 1} |k(x,y)|$. To see this, observe that for all $x \in [0,1]$

$$|(Tf)(x)| = \Big| \int_0^1 k(x,y)f(y)dy \Big| \le \|k\|_\infty \int_0^1 |f(y)|dy \le \|k\|_\infty \max_{0 \le x \le 1} |f(x)|.$$

Accordingly,

$$\|Tf\|_\infty = \max_{0 \le x \le 1} |(Tf)(x)| \le \|k\|_\infty \|f\|_\infty,$$

i.e., $\|T\| \le \|k\|_\infty$. The function k is often called **kernel** *of the operator T.*

The following example shows that a bounded linear operator – say, from X to Y – is always defined everywhere in X.

Example 2.9. *The operator T defined by*

$$(Tf)(x) := xf(x), \qquad f \in C([0,1]),\ x \in [0,1],$$

is linear and bounded on $C([0,1])$. It is called **position operator** *in the context of quantum mechanics.*

Not each linear operator is bounded. For example, the **momentum operator** *S defined by*

$$(Sf)(x) := f'(x), \qquad f \in C^1([0,1]),\ x \in [0,1],$$

also relevant in quantum mechanics, is linear but not bounded since continuous functions on a compact interval need not have bounded first derivative.

Example 2.10. *Let X, Y be normed vector spaces.*

(1) The operator mapping each $x \in X$ to $0 \in Y$, called the **zero operator***, is linear and bounded with norm 0. Its null space is clearly X and its range $\{0\}$.*

(2) The operator Id mapping each $x \in X$ to itself, called the **identity***, is linear and bounded on X. If $X \neq \{0\}$, then Id has norm 1. Its null space is $\{0\}$ and its range is X.*

(3) By Remark 2.3, if $T \in \mathcal{L}(X,Y)$ is invertible, then

$$1 = \|\,\mathrm{Id}\,\| = \|TT^{-1}\| \le \|T\|\|T^{-1}\|,$$

and therefore $\|T\|^{-1} \le \|T^{-1}\|$.

Exercise 2.11. *Let X be a Banach space. Show that a linear operator on X is bounded if and only if it is continuous if and only if it is Lipschitz continuous.*

Exercise 2.12. *Let Y be a normed vector space and $n \in \mathbb{N}$. Show that each linear operator from $\mathbb{K}^n$ to Y is bounded.*

Exercise 2.13. *Let X be a Banach space. Show that a linear operator on X is injective if and only if its null space is $\{0\}$.*

Remark 2.14. *Let X, Y be normed vector spaces and $T \in \mathcal{L}(X,Y)$. It is clear that $\mathrm{Ker}T$ and $\mathrm{Ran}T$ are vector spaces. Since T is continuous, $T^{-1}C$ is a closed subset of X for all closed subsets C of Y. In particular, $\mathrm{Ker}T = T^{-1}\{0\}$ is a closed subspace of X, while $\mathrm{Ran}T$ need not* *be closed. Can you find an example? A characterization of operators with closed range can be found in* [**2**, § II.7].

The set $\mathcal{L}(X,Y)$ of linear operators from X to Y has a vector space structure, after letting

$$(S+T)x := Sx + Tx \qquad \text{and} \qquad (\lambda T)x := \lambda Tx \qquad \text{for all } S,T \in \mathcal{L}(X,Y),\ \lambda \in \mathbb{K},\ x \in X.$$

In fact, $\mathcal{L}(X,Y)$ becomes a normed vector space when endowed with $\|\cdot\|$. More can be said.

Exercise 2.15. *Let X be a normed vector space and Y a Banach space. Show that $\mathcal{L}(X,Y)$ is complete, hence a Banach space.*

A most important consequence of Baire's theorem is the following, proved in 1927 by Stefan Banach and Hugo Steinhaus.

Theorem 2.16 (Uniform boundedness principle)**.** *Let X be a Banach space and Y be a normed vector space. Le $(T_j)_{j\in J}$ be a family of bounded linear operators from X to Y. If $\sup_{j\in J}\|T_j x\|_Y < \infty$ for all x in X, then $\sup_{j\in J}\|T_j\| < \infty$.*

Proof. Define a sequence of closed sets X_n by

$$X_n := \{x \in X : \sup_{j\in J}\|T_j x\|_Y \le n\}, \qquad n \in \mathbb{N},$$

so that $X = \bigcup_{n\in\mathbb{N}} X_n$. Since in particular X is a complete metric space, Baire's theorem – in the version presented in Remark 1.6 – applies. Thus, at least one X_{n_0} has an interior point, i.e., there exist $y \in X_{n_0}$ and $\delta > 0$ such that $B_\delta(y) \subset X_{n_0}$. Pick some $z \in X$ such that $\|z\|_X = \|(z+y) - y\|_X < \delta$. In this way, $y + z \in B_\delta(y) \subset X_{n_0}$ and hence for all $j \in J$, $\|T_j z\|_Y \le \|T_j(y+6z)\|_Y + \|T_j y\|_Y \le 2n_0$. Thus, $\|T_j\| \le \frac{2n_0}{\delta}$ for all $j \in J$, and the claim follows. ☐

A classical application of the uniform boundedness principle results in the proof of the following result, first obtained in 1929 by Stefan Banach. Usually, it goes under the name of **Open mapping theorem**.

Theorem 2.17. *If X and Y are Banach spaces and $T \in \mathcal{L}(X,Y)$ is surjective, then T is open, i.e., $T(U)$ is open in Y for all open subsets U of X.*

Proof. To begin with, set

$$Y_n := n\overline{T(B_1(0))} := \{ny \in Y : y \in \overline{T(B_1(0))}\}.$$

Because T is surjective one has $Y = \bigcup_{n\in\mathbb{N}} Y_n$. By Baire's theorem – again in the version presented in Remark 1.6 – at least one Y_{n_0} has an interior point, hence in particular $\overline{T(B_1(0))}$ has nonempty interior, i.e., there are $y \in Y$ and $\varepsilon > 0$ such that $B_{4\varepsilon}(y) \subset \overline{T(B_1(0))}$. In particular both y and (by similar reasons) $-y$ belong to $\overline{T(B_1(0))}$. Adding up we obtain

$$B_{4\varepsilon}(0) \subset \overline{T(B_1(0))} + \overline{T(B_1(0))} := \{y_1 + y_2 \in Y : y_1, y_2 \in \overline{T(B_1(0))}\}.$$

Since $\overline{T(B_1(0))}$ is convex (why?) we conclude that $\overline{T(B_1(0))} + \overline{T(B_1(0))} = 2\overline{T(B_1(0))}$, and this finally implies that

$$B_{2\varepsilon}(0) \subset \overline{T(B_1(0))} \qquad \text{for some } \epsilon > 0. \tag{2.3}$$

We will use (2.3) in order to prove that

$$B_{\varepsilon}(0) \subset T(B_1(0)) \qquad \text{for some } \epsilon > 0. \tag{2.4}$$

Let $z \in Y$ with $\|z\|_Y \leq \varepsilon$. It suffices to find $x \in X$ such that $\|x\| < 1$ and $Tx = z$. In fact, it follows by recurrence from (2.3) that for all $n \in \mathbb{N}$ there exists $x_n \in X$ such that

$$\|x_n\| < \frac{1}{2^n} \qquad \text{and} \qquad \Big\| z - T\sum_{k=1}^{n} x_k \Big\| < \frac{\varepsilon}{2^n}.$$

Accordingly, the series $\sum_{k\in\mathbb{N}} x_k$ converges by the Cauchy criterion. Let us denote $x := \sum_{k\in\mathbb{N}} x_k$. Then clearly $\|x\|_X < 1$ and $z = Tx$, by continuity of T.

We are finally in the position to prove the claim. Let $G \subset X$ be open and $y \in TG$. Pick $x \in X$ such that $x = Ty$. Then $B_\delta(x) \subset G$. By linearity there exists $\epsilon > 0$ such that $B_\epsilon(0) \subset TB_\delta(0)$. Then we conclude that

$$B_\epsilon(y) = y + B_\epsilon(0) \subset y + TB_\delta(0) = Tx + TB_\delta(0) = TB_\delta(x) \subset TG.$$

This concludes the proof. ☐

The following is usually called **bounded inverse theorem**.

Corollary 2.18. *If X and Y are Banach spaces and $T \in \mathcal{L}(X,Y)$ is bijective, then $T^{-1} \in \mathcal{L}(Y,X)$.*

Proof. Checking linearity of T^{-1} is an easy exercise. Let now $y \in Y$ such that $y = Tx$. Then (2.4) implies that if $\|Tx\|_Y \leq \epsilon$, then $\|x\|_X \leq 1$, and by linearity this shows that T^{-1} is bounded. ☐

Definition 2.19. *Let X, Y be normed spaces and T a bounded linear operator from X to Y. If T is bijective, then it is called an* **isomorphism**, *and in this case the spaces X, Y are called* **isomorphic**, *and we write $X \cong Y$. If T satisfies $\|Tx\|_Y = \|x\|_X$ for all $x \in X$, then T is called an* **isometry**.

Observe in particular that if a normed space X is isomorphic to a Banach space Y, then X is a Banach space, too.

Example 2.20. *The right shift is clearly an isometry with respect to all norms $\ell^1, \ell^2, \ell^\infty$, whereas the left shift is not – take e.g. the sequence $(1, 0, 0, 0, \ldots)$.*

Exercise 2.21. *Prove that each isometry is injective and in particular it has closed range.*

Theorem 2.22. *Let $(X, \|\cdot\|_X)$ be a normed space. If there is $n \in \mathbb{N}$ such that X is isomorphic to $\mathbb{R}^n$, then any further norm on X is equivalent to $\|\cdot\|_X$.*

Proof. Denote by $\|\cdot\|_2$ the norm associated with the Euclidean distance on $\mathbb{R}^n$. Let Φ be an isomorphism from $\mathbb{R}^n$ to X. Then in particular Φ, and hence the mapping $\phi : \mathbb{R}^n \ni x \mapsto \|\Phi x\|_X \in \mathbb{R}$ are continuous. Consider the unit sphere S^{n-1} of $\mathbb{R}^n$, i.e., the set of all vectors of $\mathbb{R}^n$ of unitary norm. Since S^{n-1} is compact, the restriction of ϕ to S^{n-1} has a (necessarily positive) minimum m. Let now $0 \neq x \in \mathbb{R}^n$ and observe that

$$\|\Phi\|\|x\|_2 \geq \Phi x\|_X = \|x\|_2 \left\|\Phi \frac{x}{\|x\|_2}\right\|_X \geq m\|x\|_2.$$

Let $|||\cdot|||_X$ be a further norm on X and denote by $|||\Phi|||$ the norm of Φ with respect to $|||\cdot|||_X$. As above we obtain

$$|||\Phi x|||x\|_2 \geq |||\Phi x|||_X \geq m\|x\|_2.$$

Denote by $\mathbf{m}$ the minimum of ϕ on the unit sphere *defined with respect to the norm* $|||\cdot|||_X$. Then for $y \in X$ and $x \in \mathbb{R}^n$ such that $\Phi x = y$ we have

$$|||y|||_X = |||\Phi x|||_X \leq |||\Phi|||\|x\|_2 \leq \frac{|||\Phi|||}{m}\|y\|_X.$$

One proves likewise that

$$\|y\|_X \leq \frac{\|\Phi\|}{\mathbf{m}}|||y|||_X.$$

This completes the proof. □

Exercise 2.23. *Is the right shift operator an isomorphism on either ℓ^1, ℓ^2, or ℓ^∞? Is it an isometry?*

Exercise 2.24. *Denote by c the vector space of converging sequences. Define an operator from c to c_0 by*

$$T(x_n)_{n\in\mathbb{N}} := (\lim_{n\to\infty} x_n, x_1 - \lim_{n\to\infty} x_n, x_2 - \lim_{n\to\infty} x_n, \ldots).$$

Show that T is an isomorphism.

Exercise 2.25. *Let Ω be a bounded open subset of $\mathbb{R}^n$. Consider an essentially bounded and measurable $q\Omega \to \mathbb{K}$ and define an operator M_q on $L^p(\Omega)$, called* **multiplication operator**, *by*

$$(M_q f)(x) := q(x)f(x) \qquad \text{for all } f \in L^p(\Omega) \text{ and a.e. } x \in \Omega.$$

Then M_q defines a bounded linear operator on $L^p(\Omega)$ due to the Hölder inequality. Show that $\|M_q\| = \|q\|_\infty$ and that M_q has a bounded inverse if and only if 0 is not in the **essential range** *of q, i.e., in the set*

$$\{z \in \mathbb{K} : \mu(\{x : |f(x) - z| < \epsilon\}) > 0 \text{ for all } \varepsilon \neq 0\}.$$

In this case, determine the inverse.

CHAPTER 3

Hilbert spaces

Definition 3.1. *Let H be a vector space over a field $\mathbb{K}$. An* ***inner product*** $(\cdot|\cdot)$ *on H is a mapping $H \times H \to \mathbb{K}$ such that for all $x, y \in H$ and all $\lambda \in \mathbb{K}$*
(1) $(x|x) \geq 0$ *and* $(x|x) = 0 \Leftrightarrow x = 0$*,*
(2) $(\lambda x|y) = \lambda(x|y)$*,*
(3) $(x|y+z) = (x|y) + (x|z)$*.*
Furthermore, it is required that
(4) $(x|y) = \overline{(y|x)}$ *for all $x, y \in H$*
if $\mathbb{K} = \mathbb{C}$. Then, $(H, (\cdot|\cdot)_H)$ is called a ***pre-Hilbert space****.*

Exercise 3.2. *Let $(H_1, (\cdot|\cdot)_1), (H_2, (\cdot|\cdot)_2)$ be pre-Hilbert spaces. Prove that*

$$((x_1, y_1)|(x_2, y_2)) := (x_1|x_2)_{H_1} + (y_1|y_2)_{H_2}, \qquad x_1, x_2 \in H_1,\ y_1, y_2 \in H_2,$$

defines an inner product on $H_1 \times H_2$.

Remark 3.3. *Observe that accordingly*

$$(x+y|x+y)_H = (x|x)_H + 2(x|y) + (y|y)_H \qquad \text{for all } x, y \in H$$

if H is a real pre-Hilbert space (or $(x+y|x+y)_H = (x|x)_H + 2\mathrm{Re}(x|y) + (y|y)_H$ if H is a complex pre-Hilbert space).

The following important estimate from above for inner products has been proved in 1821 by Augustin Louis Cauchy in the case of sums of scalars. It has been obtained in 1859 by Viktor Yakovlevich Bunyakovsky (and rediscovered in 1888 by Hermann Amandus Schwarz) in the case of integrals of products of functions.

Lemma 3.4 (Cauchy–Schwarz inequality)**.** *Let H be a pre-Hilbert space. Then for all $x, y \in H$ one has*

$$|(x|y)|^2 \leq (x|x)(y|y).$$

Proof. Let $x, y \in H$. If $y = 0$, the assertion is clear. Otherwise, for all $\lambda \in \mathbb{K} = \mathbb{C}$ we have

$$0 \leq (x - \lambda y|x - \lambda y)_H = (x|x)_H + |\lambda|^2 (y|y)_H - 2\mathrm{Re}(\overline{\lambda}(x|y)_H).$$

Setting $\lambda = \frac{\overline{(x|y)_H}}{(y|y)_H}$ we obtain

$$0 \leq (x|x)_H + \frac{|(x|y)_H|^2}{(y|y)_H} - 2\frac{|(x|y)_H|^2}{(y|y)_H} = (x|x)_H - \frac{|(x|y)_H|^2}{(y|y)_H}.$$

Accordingly,

$$(x|x)_H \geq \frac{|(x|y)_H|^2}{(y|y)_H},$$

whence the claim follows. (As usual, the case of $\mathbb{K} = \mathbb{R}$ can be discussed similarly.) □

Exercise 3.5. *Let H be a pre-Hilbert space and define a mapping $\|\cdot\|$ by*

$$\|x\| := \sqrt{(x|x)}, \qquad x \in H. \tag{3.1}$$

(1) Prove that $\|\cdot\|$ satisfies $\|x+y\|^2 \leq (\|x\| + \|y\|)^2$.
(2) Conclude that H is also a normed vector space with respect to the norm $\|\cdot\|$.

Definition 3.6. *If a pre-Hilbert space H is complete with respect to the canonical norm introduced in* (3.1) *(i.e., if each Cauchy sequence with respect to this norm is also convergent), then H is called a* ***Hilbert space****.*

Example 3.7. *The vector space $C([0,1])$ of continuous real-valued functions on $[0,1]$ is a pre-Hilbert space with respect to the inner product $(f|g) := \int_0^1 f(x)g(x)dx$. However, this pre-Hilbert space is* not *complete, i.e., it is not a Hilbert space. Still, this inner product induces a Hilbert space structure on the larger space $L^2(0,1)$ of square summable functions.*

Example 3.8. *The vector space ℓ^2 of all square-summable (complex-valued) sequences becomes a Hilbert space after introducing $(x|y)_2 := \sum_{n\in\mathbb{N}} x_n \overline{y_n}$, $x, y \in \ell^2$.*

Exercise 3.9. *Let H be a Hilbert space.*
(1) Show that each closed subspace of a Hilbert space H is again a Hilbert space with respect to the inner product induced by H.
(2) Deduce that the sets of square summable functions on $\mathbb{R}$ that are positive or even (see Exercise 3.18) are Hilbert spaces.

Exercise 3.10. *Let H be a pre-Hilbert space and $T \in \mathcal{L}(H)$. Define the* **quadratic form associated with** *T by $a(x) := (Tx|x)_H$.*
(1) Show that if $\mathbb{K} = \mathbb{C}$ and $a(x) = 0$ for all $x \in H$, then $T = 0$, i.e., $Tx = 0$ for all $x \in H$.
(2) Consider the linear mapping $T := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on $\mathbb{R}^2$ in order to show that $a(x)$ may vanish for all $x \in H$ even if $T \neq 0$ in the case $\mathbb{K} = \mathbb{R}$.

Exercise 3.11. *Let $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ be convergent sequences in a pre-Hilbert space H. Show that also the sequence of scalars $((x_n|y_n)_H)_{n\in\mathbb{N}}$ is convergent and*

$$\lim_{n\to\infty} (x_n|y_n) = \left(\lim_{n\to\infty} x_n \Big| \lim_{n\to\infty} y_n\right).$$

Definition 3.12. *Let H be a pre-Hilbert space. Two vectors $x, y \in H$ are said to be* **orthogonal to each other** *if $(x|y)_H = 0$, and we denote $x \perp y$.*

If two subsets A, B of H satisfy $(x|y)_H = 0$ for all $x \in A$ and all $y \in B$, also A, B are said to be **orthogonal to each other***. Moreover, the set of all vectors of H that are orthogonal to each vector in A is called* **orthogonal complement** *of A and is denoted by $A^\perp$.*

Exercise 3.13. *Let H be a pre-Hilbert space and $x \in H$, and $\lambda \in \mathbb{K}$. Then for all $y \in H$ such that $\|y\|_H = 1$ one has $(x - \lambda y) \perp y$ if and only if $\lambda = (x|y)_H$.*

Exercise 3.14. *Let H be a pre-Hilbert space. Prove the following assertions.*

(1) If x, y are orthogonal to each other, then $\|x\|_H^2 + \|y\|_H^2 = \|x + y\|_H^2$.
(2) More generally, $2\|x\|_H^2 + 2\|y\|_H^2 = \|x + y\|_H^2 + \|x - y\|_H^2$ for all $x, y \in H$.
(3) Also, for all $x, y \in H$ one has

$$4(x|y)_H = \|x + y\|_H^2 - \|x - y\|_H^2$$

if $\mathbb{K} = \mathbb{R}$, and

$$4(x|y)_H = \|x + y\|_H^2 + i\|x + iy\|_H^2 - \|x - y\|_H^2 - i\|x - iy\|_H^2$$

if $\mathbb{K} = \mathbb{C}$.
(4) If A is a subset of H, then $A^\perp$ is a closed subspace of H.
(5) If A is a subset of H, then $A \subset (A^\perp)^\perp$.
(6) The orthogonal complement $H^\perp$ agrees with $\{0\}$.

Remark 3.15. *The result in Exercise 3.14.(1) can be compared with Pythagoras' theorem). Beware: the converse is not true. Consider the vectors $(1, i), (0, 1) \in \mathbb{C}^2$: then the real part of their inner product vanishes, but they are not orthogonal to each other.*

The formulae presented in Exercise 3.14.(2)–(3) are usually known as **parallelogram law** and **polarisation identity**, respectively. Observe that in particular a normed vector space $(X, \|\cdot\|_X)$ is a pre-Hilbert space – i.e., $\|\cdot\|_X$ comes from an inner product – if and only if $\|\cdot\|_X$ satisfies the parallelogram law. Also observe that the polarisation identity implies joint continuity of the mapping $(\cdot|\cdot)_H : H \times H \to \mathbb{C}$.

One of the fundamental differences between Hilbert spaces and general Banach spaces is given in the following.

Theorem 3.16. *Let H be a Hilbert space. Let A be closed and convex subset of H and let $x_0 \in H$.*

1) Then there exists exactly one vector x of ***best approximation*** *to x_0 in A, i.e., there exists a unique vector $x \in A$ such that*

$$\|x - x_0\|_H = \inf_{y \in A} \|y - x_0\|_H.$$

2) Such a best approximation x of x_0 is characterized by the inequality

$$(x_0 - x|y - x) \leq 0 \qquad \text{for all } y \in K \text{ if } \mathbb{K} = \mathbb{R}, \text{ or by} \tag{3.2}$$

$$\operatorname{Re}(x_0 - x|y - x) \leq 0 \qquad \text{for all } y \in K \text{ if } \mathbb{K} = \mathbb{C}. \tag{3.3}$$

Such an x is usually denoted by $P_A(x_0)$ and called **orthogonal projection of** x_0 **onto** A. Somehow confusingly, also the operator P_A is commonly called **orthogonal projection of** H **onto** A.

PROOF. 1) One can assume without loss of generality that $x_0 = 0 \notin A$.

i) In order to prove existence of the vector of best approximation, let $z := \inf_{y\in A} \|y\|_H$ and consider a sequence $(y_n)_{n\in\mathbb{N}} \subset A$ such that $\lim_{n\to\infty} \|y_n\|_H = z$. Then by the parallelogram law (see Exercise 3.14.(2))

$$\lim_{m,n\to\infty} \left\|\frac{y_n + y_m}{2}\right\|_H^2 + \lim_{m,n\to\infty} \left\|\frac{y_n - y_m}{2}\right\|_H^2 = \lim_{m,n\to\infty} \frac{1}{2}\left(\|y_n\|_H^2 + \|y_m\|_H^2\right) = z^2.$$

Because A is convex $\frac{y_n+y_m}{2} \in A$, so that by definition $\|\frac{y_n+y_m}{2}\|_H^2 \geq z^2$. One concludes that

$$\lim_{m,n\to\infty} \left\|\frac{y_n - y_m}{2}\right\|_H^2 = 0,$$

i.e., $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. By completeness of H there exists a $x := \lim_{n\to\infty} y_n$, which belongs to A since A is closed. Clearly, $\|x\| = \lim_{n\to\infty} \|y_n\| = z$.

ii) In order to prove that a vector of best approximation is unique, assume that both x, x^* satisfy $\|x\|_H = \|x^*\|_H = z$. If $x \neq x^*$, then $\|x + x^*\|_H^2 < \|x + x^*\|_H^2 + \|x - x^*\|_H^2$, and by the parallelogram law

$$\left\|\frac{x + x^*}{2}\right\|_H^2 < \left\|\frac{x + x^*}{2}\right\|_H^2 + \left\|\frac{x - x^*}{2}\right\|_H^2 = \frac{1}{2}(\|x\|_H^2 + \|x^*\|_H^2) = z^2.$$

In other words, $\frac{x+x^*}{2}$ would be a better approximation of x. Since $\frac{x+x^*}{2} \in A$, this would contradict the construction of x as vector of best approximation of x_0 in A. Hence, $x = x^*$.

2) Let $w \in A$ and set $y_t := (1-t)x + tw$, where $t \in (0,1]$ is a scalar to be optimized in the following. Since A is convex, $y_t \in A$ and accordingly

$$\|x_0 - x\|_H < \|x_0 - y_t\|_H,$$

since y_t is not the (unique!) best approximation of x_0 in A. Accordingly,

$$\|x_0 - x\|_H < \|x_0 - (1-t)x - tw\|_H = \|(x_0 - x) + t(x - w)\|_H,$$

and squaring both sides we obtain by Remark 3.3

$$\|x_0 - x\|_H^2 < \|(x_0 - x)\|_H^2 + t^2\|(x - w)\|_H^2 - 2t\mathrm{Re}(x_0 - x|w - x).$$

It follows that $t^2\|(x - w)\|_H^2 > 2t\mathrm{Re}(x_0 - x|w - x)$ for all $t \in (0,1]$. Therefore, $0 \geq 2\mathrm{Re}(x_0 - x|w - x)$ in the limit $t \to 0$ and the claimed inequality holds.

Conversely, let x satisfy (3.3). Then for all $y \in A$

$$\|x - x_0\|_H^2 - \|y - x_0\|_H^2 = 2\mathrm{Re}(x_0 - x|y - x) - \|y - x\|_H^2 \leq 0,$$

i.e., $\|x - x_0\|_H^2 \leq \|y - x_0\|_H^2$. It follows that x is the best approximation of x_0 in A. (In the case $\mathbb{K} = \mathbb{R}$ the assertions can be proved in just the same way). □

Exercise 3.17. *Let H be a Hilbert space. Let A_1, A_2 be closed and convex subsets of H and denote by P_1, P_2 the orthogonal projections onto A_1, A_2, repsectively. Prove that $P_1A_2 \subset A_2$ if and only if $P_2A_1 \subset A_1$ if and only if P_1, P_2 commute, i.e., $P_1P_2x = P_2P_1x$ for all $x \in H$.*

Exercise 3.18. *Define $A_1, A_2 \subset L^2(\mathbb{R})$ as the sets of all square summable functions that are a.e. even and positive, respectively.*

(1) Show that A_1, A_2 are closed convex subsets of $L^2(\mathbb{R})$.
(2) Prove that the orthogonal projections P_{A_1}, P_{A_2} onto A_1, A_2 are given by

$$P_{A_1} f(x) = \frac{f(x) + f(-x)}{2} \qquad \text{and} \qquad P_{A_2} f(x) = \frac{|f(x)| + f(x)}{2} \qquad \text{for a.e. } x \in \mathbb{R}.$$

Exercise 3.19. *Let H be a Hilbert space and A be a closed convex subset of H.*

(1) Show that the orthogonal projection P_A is linear if and only if A is a closed subspace of H.
(2) Prove that P_A is Lipschitz continuous with Lipschitz constant 1.

Exercise 3.20. *Let H be a Hilbert space and Y be a closed subspace of H.*

(1) Show that if $Y \neq \{0\}$, then the orthogonal projection P_Y of H onto Y satisfies $\|P_Y\| = 1$ and $\mathrm{Ker} P_Y = Y^\perp$.
(2) Prove that each $x \in H$ admits a unique decomposition as $x = y + z$, where $y = P_Y x \in Y$ and $z = P_{Y^\perp} x \in Y^\perp$.

Recall that the linear span of two vector spaces Y, Z whose interection is $\{0\}$ is called their **direct sum**, and we write $Y \oplus Z$. Accordingly, under the assumptions of Exercise 3.20 H is the direct sum of Y and $Y^\perp$. We also denote by $x = y \oplus z$ the decomposition introduced in Exercise 3.20.(2).

The following is one of the fundamental results in functional analysis. It has been first proved by Eduard Helly in 1912, although it is commonly dubbed the Hahn–Banach Theorem, after Hans Hahn and Stefan Banach who rediscovered it independently at the end of the 1920s.

Theorem 3.21 (Hahn–Banach Theorem)**.** *Let H be a real Hilbert space and C a nonempty convex set. Let $x_0 \in H \setminus \overline{C}$. Then there exists $x^* \in H$, $x^* \neq 0$, such that*

$$(x^*|x)_H > (x^*|x_0) \qquad \text{for all } x \in C.$$

Proof. The desired vector x^* can be written down explicitly: it is given by $x^* := P_{\overline{C}} x_0 - x_0$, where the existence of the orthogonal projection $P_{\overline{C}}$ of H onto the closed convex set $\overline{C}$ is ensured by Theorem 3.16.(1). Observe that since $x_0 \notin C$, $P_{\overline{C}} x_0 \neq x_0$ and therefore $x^* \neq 0$.

In order to check that such an x^* has the claimed properties, observe that by Theorem 3.16.(2) for all $x \in \overline{C}$ we have

$$\begin{aligned} 0 &\leq (x^*|x - P_{\overline{C}} x_0)_H \\ &= (x^*|x - x_0 + x_0 - P_{\overline{C}} x_0)_H \\ &= (x^*|x - x_0)_H - (x^*|P_{\overline{C}} x_0 - x_0)_H \\ &= (x^*|x)_H - (x^*|x_0) - \|x^*\|_H^2. \end{aligned}$$

Accordingly, $(x^*|x)_H \geq \|x^*\|_H^2 + (x^*|x_0) > (x^*|x_0)$ and the claim follows. ☐

Exercise 3.22. *Mimic the proof of Theorem 3.21 in order to proof the following weaker result:*

Let H be a real Hilbert space and C a nonempty, open and convex set. Let $x_0 \in H \setminus C$. Then there exists $x^* \in H$ such that

$$\inf_{x \in C} (x^*|x)_H \geq (x^*|x_0).$$

Remark 3.23. *It is a fundamental fact of functional analysis that the Hahn–Banach theorem also holds in general Banach (rather than Hilbert) spaces. The proof is however much lengthier and we omit it. Observe that, unlike in the Hilbert space setting, the Banach space version of the Hahn–Banach theory relies upon the Axiom of choice.*

The above Hahn–Banach theorem has a manifold of more or less direct consequences in functional analysis. E.g., Chapter 1 of [**2**] is devoted to the Hahn–Banach *theorems*. Let us mention one of the most prominent.

Corollary 3.24. *Let H be a Hilbert space, A, B to convex, nonempty and disjoint subsets of H. If A is open, then there exists a* **closed hyperplane** *that* **separates** *A, B* **in a broad sense**, *i.e., there exist $x^* \in H$, $x^* \neq 0$, and $\alpha \in \mathbb{K}$ such that $(x^*|x)_H \geq \alpha$ for all $x \in A$ and $(x^*|x)_H \leq \alpha$ for all $x \in B$.*

Proof. Consider the convex set $C = A - B$. Then C is open, since it is the union of open sets (it can be written as $C = \bigcup_{y \in B} A - \{y\}$) and $0 \notin C$, since there is no $x \in A \cap B$. Then by Exercise 3.22 there exists $x \in H$ such that $\inf_{x \in C}(x^*|x)_H \geq 0$, i.e., such that $(x^*|x)_H \geq (x^*|y)$ for all $x \in A$ and all $y \in B$. The claim now follows setting $\alpha := \frac{\inf_{x \in A}(x^*|x)_H - \sup_{y \in B}(x^*|y)}{2}$. □

Exercise 3.25. *Consider an open subset $D \subset \mathbb{C}$, a Hilbert space H and a mapping $f : D \to H$. Deduce from the theorem of Hahn–Banach that f is holomorphic[1] if and only if $\langle \phi, f(\cdot) \rangle : D \to \mathbb{C}$ is holomorphic for all $\phi \in H$.*

---

[1] I.e., for all $z \in D$ there exists $x \in H$ such that

$$\lim_{h \to 0} \frac{f(z+h) - f(z) - hx}{h} = 0.$$

CHAPTER 4

Fourier analysis

Definition 4.1. *Let $(H, (\cdot|\cdot)_H)$ be a pre-Hilbert space. Then a family $\{e_n \in H \setminus \{0\} : n \in J\}$, $J \subset \mathbb{N}$, is called* **orthogonal** *if $(e_n|e_m)_H = 0$ for all $n \neq m$, and* **orthonormal** *if $(e_n|e_m)_H = \delta_{mn}$ for all $m, n \in J$, where δ_{mn} denotes the Kronecker delta.*

Moreover, $\{e_n \in H : n \in J\}$ is called **total** *if its linear span (i.e., the set of all* finite *linear combinations of elements of the family) is dense in H.*

An orthonormal and total family is called a Hilbert space basis of H, or simply a **basis**. *The smallest cardinality of a basis is called* **Hilbert space dimension of** *H.*

Definition 4.2. *A Hilbert space H is called* **separable** *if it contains a countable total family.*

Example 4.3. *The vector space $\mathbb{C}^n$ is a Hilbert space with respect to the inner product defined by $(x|y) := \sum_{k=1}^n x_k \overline{y_k}$, $x, y \in \mathbb{C}^n$. It is clearly separable.*

Example 4.4. *Let $\ell^2([0,1])$ denote the space of all functions $f : [0,1] \to \mathbb{R}$ that vanish on $[0,1] \setminus E_f$ (where the exceptional set E_f is countable), and such that $(f(x))_{x \in E_f}$ is a square summable sequence. Define a mapping $H \times H \to \mathbb{R}$ by*

$$\langle f, g \rangle = \sum_{x \in [0,1]} f(x) g(x)$$

(the sum is over a countable set!). Then $\langle \cdot, \cdot \rangle$ defines an inner product on H, and in fact one can check that $(H, \langle \cdot, \cdot \rangle)$ is a Hilbert space. However, H is not separable. In fact for the sequence $(f_n)_{n \in \mathbb{N}} \subset H$ there exists a function $f \in H$, $f \neq 0$, such that

$$\langle f, f_n \rangle = 0 \qquad \forall n \in \mathbb{N},$$

since the set $\{f_n(x) \neq 0 : n \in \mathbb{N}, x \in [0,1]\}$ is countable. Accordingly, the linear span of any sequence $(f_n)_{n \in \mathbb{N}}$ is not dense in H.

Example 4.5. *Each closed subspace of a separable Hilbert space is separable.*

Definition 4.6. *Let X be a normed space and $\{x_n \in X : n \in J\}$ a family of vectors. The associated series $\sum_{n \in J} x_n$ is called* **convergent** *if for some $x \in X$ and all $\epsilon > 0$ there exists a finite set $J_\epsilon \subset \mathbb{N}$ such that $\|x - \sum_{n \in \tilde{J}} x_n\|_X < \epsilon$ and all finite sets $\tilde{J}$ such that $J_\epsilon \subset \tilde{J} \subset J$. It is called* **absolutely convergent** *if $(\|x_n\|_X)_{n \in \mathbb{N}}$ belongs to ℓ^1.*

Remark 4.7. *Observe that if $J \subset \mathbb{N}$, then convergence of the series $\sum_{n \in J} x_n$ associated with $\{x_n \in X : n \in J\}$ is equivalent to the usual notion – i.e., convergence of the sequence of partial sums.*

Exercise 4.8. *Let X be a normed space. Then X is complete if and only if each absolutely convergent series is also convergent.*

Corollary 4.9. *Let H be a Hilbert space. Let $\{x_n \in H : n \in \mathbb{N}\}$ be an orthogonal family. Then the series $\sum_{n\in\mathbb{N}} x_n$ is convergent if and only if $(\|x_n\|_X)_{n\in\mathbb{N}} \subset \mathbb{R}$ is square summable.*

Proof. Applying Exercise 3.14.(1) repeatedly we deduce that

$$\Big\| \sum_{k\in J} x_k \Big\|_H^2 = \sum_{k\in J} \|x_k\|_H^2$$

for any finite subset J of $\mathbb{N}$. Taking the limit we can extend this equality to the case of series associated with infinite sequences. Hence, $\| \sum_{k\in\mathbb{N}} x_k\|_H < \infty$ if and only if

$$\Big\| \sum_{k\in\mathbb{N}} x_k \Big\|_H^2 = \sum_{k\in\mathbb{N}} \|x_k\|_H^2 < \infty.$$

This completes the proof. □

Theorem 4.10. *Let H be a pre-Hilbert space. Let $\{e_n \in H : n \in J\}$ be an orthonormal family. Then the following assertions hold.*

(1) $\sum_{n\in J} |(x|e_n)_H|^2 \leq \|x\|_H^2$ for all $x \in H$. If in particular H is complete, then the series $\sum_{n\in J}(x|e_n)_H e_n$ converges.

(2) If $J \subset \mathbb{N}$ and $x = \sum_{n\in J} a_n e_n$, then $\|x\|_H^2 = \sum_{n\in J} |a_n|^2$ and $a_n = (x|e_n)_H$ for all $n \in J$.

The assertions in (1) and (2) are usually called **Bessel's inequality** and **Parseval's identity identity**, after Friedrich Bessel and Marc-Antoine Parseval des Chênes. The scalars a_n in (2) are called **Fourier coefficients of** x.

Proof. (1) Upon going to the limit, it suffices to prove the claimed inequality for any *finite* orthonormal family $\{e_1, \dots, e_N\}$. Then one has

$$\begin{aligned} 0 &\leq \Big\| x - \sum_{n=1}^N (x|e_n)_H e_n \Big\|^2 \\ &= \|x\|_H^2 - \sum_{n=1}^N (x|e_n)_H (e_n|x)_H - \sum_{n=1}^N \overline{(x|e_n)_H} (e_n|x)_H + \Big(\sum_{n=1}^N (x|e_n) e_n \Big| \sum_{m=1}^N (x|e_m) e_m \Big) \\ &= \|x\|_H^2 - 2\sum_{n=1}^N |(x|e_n)_H|^2 + \sum_{n=1}^N |(x|e_n)_H|^2. \end{aligned}$$

If moreover H is complete, then convergence of $\sum_{n\in J}(x|e_n)_H e_n$ can be deduced showing its absolute convergence, i.e., applying Corollary 4.9 to the sequence $((x|e_n)_H e_n)_{n\in\mathbb{N}}$.

(2) If J is finite, say $J = \{1, \dots, N\}$, then the first assertion is clear: in fact, due to orthogonality one has

$$\Big\| \sum_{k=1}^N a_k e_k \Big\|_H^2 = \sum_{k=1}^N \|a_k e_k\|_H^2 = \sum_{k=1}^N |a_k|^2.$$

If J is infinite – and hence without loss of generality $J = \mathbb{N}$ – then

$$\|x\|_H^2 = \lim_{n\to\infty} \Big\| \sum_{k=1}^n a_k e_k \Big\|_H^2 = \lim_{n\to\infty} \sum_{k=1}^n |a_k|^2 = \sum_{n\in J} |a_n|^2.$$

Moreover, since $x = \sum_{n\in J} a_n e_n$, i.e., $\lim_{n\to\infty} \| \sum_{k=1}^n a_n e_n - x\|_H = 0$, one sees that for fixed $m \in \mathbb{N}$ and all $n \geq m$

$$\Big(\sum_{k=1}^n a_n e_n | e_m \Big)_H - (x|e_m)_H = \Big(\sum_{k=1}^n a_n e_n - x | e_m \Big)_H,$$

and by the Cauchy–Schwarz inequality

$$\Big| \Big(\sum_{k=1}^n a_n e_n - x | e_m \Big)_H \Big| \leq \Big\| \sum_{k=1}^n a_n e_n - x \Big\|_H \|e_m\|_H = \Big\| \sum_{k=1}^n a_n e_n - x \Big\|_H.$$

Accordingly,

$$\lim_{n\to\infty} \Big(\sum_{k=1}^n a_n e_n | e_m \Big)_H = (x|e_m).$$

Furthermore,

$$\Big(\sum_{k=1}^n a_k e_k | e_m \Big)_H = \sum_{k=1}^n (a_k e_k | e_m)_H = \sum_{k=1}^n a_k (e_k | e_m)_H = \sum_{k=1}^n a_k \delta_{km} = a_m.$$

This concludes the proof. □

Proposition 4.11. *Let H be a Hilbert space. An orthonormal family $\{e_n : n \in \mathbb{N}\}$ is in fact total (i.e., a basis) if and only if the only vector orthogonal to each e_n is 0.*

PROOF. Let $f \in H$. Let $\{e_n : n \in \mathbb{N}\}$ be a basis such that $(f|e_k)_H = 0$ for all $k \in \mathbb{N}$. Fix an $\epsilon > 0$. By totality of $\{e_n : n \in \mathbb{N}\}$ there exists a finite set $\{a_1, \ldots, a_n\}$ such that $\|f - \sum_{k=1}^n a_k e_k\|_H < \epsilon$. Accordingly,

$$\begin{aligned}
\|f\|_H^2 &= \Big| \|f\|_H^2 - \Big(f \Big| \sum_{k=1}^n a_k e_k \Big)_H \Big| \\
&= \Big| \Big(f \Big| f - \sum_{k=1}^n a_k e_k \Big)_H \Big| \\
&\leq \|f\|_H \Big\| f - \sum_{k=1}^n a_k e_k \Big\|_H \\
&< \epsilon \|f\|_H.
\end{aligned}$$

Therefore, $\|f\|_H < \epsilon$ for all $\epsilon > 0$, i.e., $\|f\|_H = 0$, hence f vanishes a.e.

Conversely, let the only vector orthogonal to each e_n be 0. Define $(s_n)_{n\in\mathbb{N}}$ by

$$s_n := \sum_{k=1}^{n} (f|e_k)_H,$$

which is a Cauchy sequence by Theorem 4.10.(1), hence convergent towards $g := \sum_{k=1}^{\infty} (f|e_k)_H e_k \in H$. By Theorem 4.10.(2), $(g|e_k)_H = (f|e_k)_H$ or rather $(g - f|e_k) = 0$ for all $k \in \mathbb{N}$. By Exercise 3.14.(6) this means that $f = g$. I.e., any $f \in H$ can be expressed as a Fourier series with respect to $(e_n)_{n\in\mathbb{N}}$. It follows that $(e_n)_{n\in\mathbb{N}}$ is total. □

Remark 4.12. *Completeness of H was only used to prove the latter implication. In fact, the former one holds for general pre-Hilbert spaces. On the other hand, the latter implication does fail in general pre-Hilbert spaces. This has been shown by J. Dixmier,*

Example 4.13. *The vector space ℓ^2 of all square-summable sequences is separable, since the family $(e_n)_{n\in\mathbb{N}}$ defined by setting $e_n := (\delta_{nm})_{m\in\mathbb{N}}$ (where $\delta_{n\cdot}$ denotes the Kronecker delta associated with n) is orthonormal and total. In particular, totality can be proved by Proposition 4.11.*

Exercise 4.14. *Let H be a separable Hilbert space with an orthornormal basis $(e_n)_{n\in J}$. Show that a sequence $(a_n)_{n\in J}$ of scalars belongs to ℓ^2 if and only if $\sum_{n\in J} a_n e_n$ converges.*

Lemma 4.15. *Let $(H, (\cdot|\cdot)_H)$ be a Hilbert space. Let $\{e_n \in H : n \in J\}$ be an orthonormal family and denote by Y its linear span. Then the orthogonal projection P_Y of H onto Y is given by*

$$P_Y x = \sum_{n\in J} (x|e_n)_H e_n, \qquad x \in H.$$

Observe that the above series converges by Bessel's inequality.

Proof. Let $x \in H$. Then clearly $P_Y x \in Y$ and moreover

$$(x - P_Y x|e_m)_H = (x|e_m)_H - \sum_{n\in J} (x|e_n)_H (e_n|e_m)_H = (x|e_m)_H - (x|e_m)_H$$

due to orthonormality, i.e., $(x - P_Y x|e_m)_H = 0$ for all $m \in J$. We deduce that $x - P_Y x \in Y^\perp$, and accordingly each $x \in H$ can be written as $x = P_Y x + (x - P_Y x)$, the unique sum of two terms in Y and $Y^\perp$, respectively. We conclude that $H = Y \oplus Y^\perp$, and the claim follows. □

Remark 4.16. *It follows from the above proof that the orthogonal projection on $Y^\perp$ is given by $P_Y^\perp = (\mathrm{Id} - P_Y)$.*

Theorem 4.17. *Let $(H, (\cdot|\cdot)_H)$ be a separable Hilbert space. Let $\{x_n \in H : n \in J\}$, $J \subset \mathbb{N}$, be a total family of linearly independent vectors. Set $e_n := \frac{f_n}{\|f_n\|_H}$, where the sequence $(f_n)_{n\geq 1}$ is defined recursively by*

$$f_1 := x_1, \qquad f_n := x_n - \sum_{k=1}^{n-1} (x_n|e_k)_H e_k, \qquad n = 1, 2, \ldots. \tag{4.1}$$

Then $\{e_n \in H : n \geq 1\}$ is a basis.

The recursive process described in (4.1) is called **Gram–Schmidt orthonormalisation**, after Jørgen Pedersen Gram and Erhard Schmidt who discovered it independently in 1883 and 1907, respectively (but it was already known to Pierre-Simon Laplace und Augustin Louis Cauchy).

PROOF. We only consider the case of $J = \mathbb{N}$, since if J is finite the assertion agrees with validity of the usual linear algebraic Gram–Schmidt orthonormalisation process in finite dimensional spaces.

Observe that by Lemma 4.14 $f_n = x_n - P_{n-1}x_n = P_{n-1}^{\perp}x_n$, where P_{n-1} denotes the orthogonal projection onto the space $\operatorname{span}\{x_k : k = 1, \ldots, n-1\}$ of all linear combinations of vectors $x_1, \ldots, x_{n-1}$. In other words, $f_n \in \operatorname{span}\{x_k : k = 1, \ldots, n-1\}^{\perp}$. This shows that the family is orthonormal.

It remains to show totality of $\{e_n \in H : n \in J\}$. Observe that each x_n lies in the linear span of $\{e_n \in H : n \in J\}$, and conversely one can prove by induction that each e_n lies in the linear span of $\{x_n \in H : n \in J\}$. Accordingly, the closure of the linear span of both $\{e_n \in H : n \in J\}$ and $\{x_n \in H : n \in J\}$ agree. □

Exercise 4.18. *Consider the subset $M = \{f_n : n \in \mathbb{N}\}$ of $L^2(-1,1)$, where $f_n(t) := t^n$ for a.e. $t \in (-1,1)$. Perform the Gram–Schmidt orthonormalisation process on M. The vectors of the basis obtained in this way are called* **Legendre polynomials***, as they have been introduced in 1784 by Adrien-Marie Legendre.*

Exercise 4.19. *Show that the* **Rademacher functions** *$\{r_n : n \in \mathbb{N}\}$ define an orthonormal family of $L^2(0,1)$, where $r_n(t) = \operatorname{sign}\sin(2^n\pi t)$ for a.e. $t \in (0,1)$. Rademacher functions have been proposed by Hans Adolph Rademacher in 1922.*

Exercise 4.20. *Let $\phi \in L^2(\mathbb{R})$. The countable family $\{\phi_{jk} : j,k \in \mathbb{Z}\}$ is called a* **wavelet** *if it defines a basis of $L^2(\mathbb{R})$, where ϕ_{jk} is defined by*

$$\phi_{jk}(t) := 2^{\frac{k}{2}}\phi(2^k t - j) \qquad \text{for a.e. } t \in (0,1).$$

Show that if $\phi(t) = \mathbb{1}_{[0,\frac{1}{2})} - \mathbb{1}_{[\frac{1}{2},1)}$ for a.e. $t \in (0,1)$, then $\{\phi_{jk} : j,k \in \mathbb{Z}\}$ is a wavelet, the so-called **Haar wavelet** *discovered in 1909 by Alfred Haar.*

Exercise 4.21. *Show that the family $\{1, \cos 2\pi n\cdot, \sin 2\pi m\cdot : n,m = 1,2,3,\ldots\}$ is orthonormal in $L^2(0,1;\mathbb{R})$. Finite linear combinations of elements of this family are called* **trigonometric polynomials***.*

Exercise 4.22. *Show that the family $\{e^{2\pi ikt} : k \in \mathbb{Z}\}$ is orthonormal in $L^2(0,1;\mathbb{C})$. Observe that trigonometric polynomials are also finite linear combinations of elements of this family.*

Remark 4.23. *It should be observed that $(r_n)_{n\in\mathbb{N}}$ does not define a basis of $L^2(0,1)$, since e.g. the function $t \mapsto \cos 2\pi t$ is orthogonal to all Rademacher functions. Also, remark that the Haar wavelet does not consist of continuous functions, unlike Legendre polynomials and the other most common bases of L^2.*

Remark 4.24. *One can easily generalize the above examples and show that each Lebesgue space $L^2(I)$ is separable, where $I \subset \mathbb{R}$ is an interval (either bounded or unbounded).*

Remark 4.25. *The notion of basis, i.e., of a subset $\{e_n : n \in \mathbb{N}\}$ such that for each vector x there is a unique sequence $(a_n)_{n\in\mathbb{N}}$ such that*

$$\lim_{n\to\infty} \left\| x - \sum_{k=1}^{n} a_k e_k \right\| = 0$$

can be defined also in the Banach space case – in this context it is then called a **Schauder basis**, *after Juliusz Schauder who first introduced them in an axiomatic way in 1927. Although most common spaces do have a Schauder basis, Per Enflo has exhibited in 1973 a Banach space without a Schauder basis. Observe that constructing general operators is almost impossible if no Schauder basis is available – this makes it extremely difficult to construct (nontrivial) counterexamples in Banach spaces without a Schauder basis.*

Proposition 4.26. *Each separable Hilbert space has a countable basis.*

Proof. Let H be a separable Hilbert space. Consider a dense countable set $M \subset H$. Without loss of generality we can assume the elements of M to be linearly independent. Apply the Gram–Schmidt orthonormalisation process to the elements of M to find a countable, total, orthonormal family. □

Corollary 4.27. *Each separable Hilbert space is isometric isomorphic to ℓ^2.*

Proof. Let H be a separable Hilbert space and $(f_n)_{n\in\mathbb{N}}$ be a countable basis of H, which exists by Proposition 4.25. Define an operator T by $Tf_n := e_n$ for all $n \in \mathbb{N}$, where e_n is the canonical basis of ℓ^2 introduced in Example 3.8. Then T is clearly an isomorphism, and moreover by construction $\|Tf\|_{\ell^2} = \|f\|_H$. □

Definition 4.28. *Let $f : [0,1] \to \mathbb{C}$. Introduce a sequence $(S_n f)_{n\in\mathbb{N}}$ of functions defined by*

$$(S_n f)(t) = \sum_{|k|\le n} (f|e_k)_{L^2(0,1)} e_k(t) = \sum_{|k|\le n} \int_0^1 f(x) e^{2\pi k i(t-x)} dx, \qquad t \in [0,1],$$

where the orthonormal family $(e_n)_{n\in\mathbb{N}}$ is defined as in Exercise 4.20. If $(S_n f)_{n\in\mathbb{N}}$ converges to f, then it is called the **Fourier series** *associated with f.*

The above definition of Fourier series is admittedly vague. What kind of convergence is the sequence $(S_n f)_{n\in\mathbb{N}}$ requested to satisfy? A natural guess would be the uniform or pointwise convergence. In fact a century ago it has been realize that a notion of convergence coarser than these was necessary in order to fully understand approximation property of Fourier series has been one of the earlier successes of Hilbert space theory. The following result, proven independently in 1907 by Frigyes Riesz and Ernst Sigismund Fischer, is the starting point of the modern theory of Fourier analysis.

In the proof we will exploit completeness of the space L^2, which we have mentioned in Remark 1.11.

Theorem 4.29. *The orthonormal family introduced in Exercise 4.21 is a basis of $L^2(0,1)$. Accordingly, for all $f \in L^2(0,1;\mathbb{C})$ the Fourier series associated with it converges to f with respect to the L^2-norm, i.e.,*

$$\lim_{n\to\infty} \int_0^1 \left| f(t) - \sum_{|k|\le n} \int_0^1 f(x) e^{2\pi i k(t-x)} dx \right|^2 dt = 0.$$

Proof. It has been already proved in Exercise 4.20 that the family $\{e_n := e^{2\pi i n \cdot} : n \in \mathbb{Z}\}$ is orthonormal in $L^2(0,1;\mathbb{R})$. By Proposition 4.11, it is a basis if and only if

$$(4.2) \qquad (f|e_n)_{L^2(0,1)} = 0 \text{ for all } n \in \mathbb{Z} \quad \text{implies that } f(x) = 0 \text{ for a.e. } x \in (0,1).$$

Observe that if $(f|e_n)_{L^2(0,1)} = 0$ for all $n \in \mathbb{Z}$, then we also have $(\overline{f}|e_n)_{L^2(0,1)} = 0$ for all $n \in \mathbb{Z}$ and summing these both relations we deduce that both the real and imaginary parts of f are orthgonal to all basis vectors. Accordingly, it suffices to check condition (4.2) for any real-valued function f.

We first consider a continuous function $f : [0,1] \to \mathbb{R}$ and observe that if $f \neq 0$, then there is $x_0 \in [0,1]$ such that $f(x_0) \neq 0$. Without loss of generality we can assume $f(x_0)$ to be a positive maximum of f. Due to continuity there exists a neighbourhood $(x_0 - \delta, x_0 + \delta)$ such that $2f(x) > f(x_0)$ for all $x \in [x_0 - \delta, x_0 + \delta]$. Consider a linear combination p of basis vectors such that

$$m \le p(y) \quad \text{for some } m > 1 \text{ and all } y \in [x_0 - \frac{\delta}{2}, x_0 + \frac{\delta}{2}]$$

and

$$|p(y)| \le 1 \quad \text{for all } y \notin [x_0 - \delta, x_0 + \delta].$$

(Such a function p surely exists, consider e.g. the trigonometric polynomial

$$p(x) := 1 - \cos 2\pi\delta + \cos 2\pi(x_0 - x)).$$

It follows from the assumption in (4.2) that f is orthogonal to p^n for each $n \in \mathbb{N}$, and hence

$$\begin{aligned} 0 &= \int_0^1 f(x)p^n(x)dx \\ &= \int_0^{x_0-\delta} f(x)p^n(x)dx + \int_{x_0-\delta}^{x_0+\delta} f(x)p^n(x)dx + \int_{x_0+\delta}^1 f(x)p^n(x)dx. \end{aligned}$$

Moreover, for all $n \in \mathbb{N}$ one has

$$\begin{aligned} \left| \int_0^{x_0-\delta} f(x)p^n(x)dx \right| + \left| \int_{x_0+\delta}^1 f(x)p^n(x)dx \right| &\le \int_0^{x_0-\delta} |f(x)p^n(x)|dx + \int_{x_0+\delta}^1 |f(x)p^n(x)|dx \\ &\le \int_0^{x_0-\delta} |f(x)|dx + \int_{x_0+\delta}^1 |f(x)|dx \\ &\le \|f\|_{L^1(0,1)} < \infty, \end{aligned}$$

since in particular $f \in L^1(0,1)$. Still, one sees that

$$\int_{x_0-\delta}^{x_0+\delta} f(x)p^n(x)dx \geq \int_{x_0-\frac{\delta}{2}}^{x_0+\frac{\delta}{2}} f(x)m^n dx \geq \frac{f(x_0)}{2} m^n \frac{\delta}{2},$$

and therefore $\lim_{n\to\infty} \int_{x_0-\delta}^{x_0+\delta} f(x)p^n(x)dx = \infty$, a contradiction to the assumption that $\int_0^1 f(x)p^n(x)dx = 0$ for all $n \in \mathbb{N}$.

Let us now consider a possibly discontinuous function $g \in L^2(0,1;\mathbb{R}) \subset L^1(0,1;\mathbb{R})$ and consider the continuous function $G := \int_0^\cdot g(x)dx$. Since by assumptions g is orthogonal to each function $e^{2\pi ik\cdot}$, $k \in \mathbb{Z}$, integrating by parts one clearly obtains that also $G - \int_0^1 G(x)dx$ is orthogonal to each function $e^{2\pi ik\cdot}$, $k \in \mathbb{Z}$. (In fact, the corrective term $\int_0^1 G(x)dx$ is needed since G is in general not orthogonal to $1 = e^0$). Due to continuity of $G - \int_0^1 G(x)dx$, we can apply the result obtained above and deduce that $G - \int_0^1 G(x)dx \equiv 0$, i.e., $g(x) = G'(x) \equiv 0$ for a.e. $x \in (0,1)$. Here we are using the well-known fact that if a function $h : [0,1] \to \mathbb{C}$ satisfies $\int_0^t h(t)dt = 0$ for all $t \in (0,1)$, then $h \equiv 0$.

Moreover, L^2-convergence of Fourier series is a direct consequence of Bessel's inequality, as already deduced in the proof of Proposition 4.11. □

Theorem 4.30. *Let $f : [0,1] \to \mathbb{C}$ be a continuously differentiable function such that $f(0) = f(1)$. Then $(S_n f)_{n\in\mathbb{N}}$ converges uniformly to f.*

Although the result deals with spaces of continuous functions and the sup-norm, the proof is based on properties of the Hilbert space $L^2(0,1)$.

Proof. Let $k \neq 0$ and observe that

$$(f|e_k)_{L^2(0,1)} = \int_0^1 f(t)e^{-2\pi kit}dt = \frac{1}{2k\pi i}\int_0^1 f'(t)e^{-2\pi kit}dt = \frac{1}{2\pi ki}(f'|e_k)_{L^2(0,1)}.$$

Accordingly, $(S_n f)_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_\infty$, because for all $n, m \in \mathbb{N}$ with $n < m$

$$\|S_n f - S_m f\|_\infty \leq \frac{1}{2\pi} \sum_{|k|=n+1}^{m} |(f'|e_k)_{L^2(0,1)}| \frac{1}{k} \leq \frac{1}{2\pi} \sum_{|k|=n+1}^{m} |(f'|e_k)_{L^2(0,1)}|^2 \sum_{|k|=n+1}^{m} \frac{1}{k^2}.$$

due to the Cauchy–Schwarz inequality: in fact, both sums in the right hand side define converge to 0 due to Bessel's inequality and the convergence of the geometric series, respectively. Denote by $g \in C([0,1])$ the uniform limit of $(S_n f)_{n\in\mathbb{N}}$. Then by Theorem 4.28 there holds

$$0 \leq \|g - f\|_{L^2(0,1)} = \lim_{n\to\infty} \|g - S_n f\|_2 \leq \lim_{n\to\infty} \|g - S_n f\|_\infty.$$

It follows that $f = g$ a.e., and the claim follows because $g \in C[0,1]$. □

Fourier series have been first developed in a fundamental book published by Joseph Fourier in 1822. Fourier claimed that such series always converge – more precisely, that $\lim_{n\to\infty} S_n f = f$ *uniformly* "for all functions f".

In fact, it is clear that if a Fourier series converges uniformly, then it necessarily converges to a continuous function (excluding several interesting but discontinuous functions). More generally, Johann Peter Gustav Lejeune Dirichlet showed in 1828 that the Fourier series associated with a piecewise continuous function converges pointwise on each closed interval not containing any point of discontinuity. This arises the question about convergence of Fourier series of functions having infinitely many discontinuities. Studying this problem led Georg Cantor to develop his theory of infinite sets, as nicely and thoroughly discussed in [**5**].

It became clear soon that Fourier's approach was ingenious but rather heuristical, and that a more detailed proof is in order. In fact, Paul Du Bois-Reymond in 1873 was the first mathematician who could find a continuous function whose associated Fourier series did not converge uniformly. It was then realized that the right approach was to weaken the notion of convergence as a tradeoff to be able to keep on representing functions by nice trigonometric series. This marked the birth of *harmonic analysis*. Several further results on Fourier series can be found in [**22**, § IV.2].

Lemma 4.31. *Let $f : [0,1] \to \mathbb{R}$ be a continuous function such that $f(0) = f(1)$. Extend f to $\mathbb{R}$ by periodicity and denote this extension again by f. Then*

$$(S_n f)(t) = \int_0^1 f(s+t) D_n(s)\,ds, \qquad t \in [0,1],\ n \in \mathbb{N},$$

where the **Dirichlet kernel** *is defined by*

$$D_n(t) := \frac{\sin\big((2n+1)\pi t\big)}{\sin \pi t}, \qquad t \in [0,1],\ n \in \mathbb{N}.$$

Proof. There holds

$$\begin{aligned} \frac{\sin\big((2n+1)\pi t\big)}{\sin \pi t} &= \frac{e^{(2n+1)\pi i t} - e^{-(2n+1)\pi i t}}{e^{\pi i t} - e^{-\pi i t}} \\ &= e^{2n\pi i t} \frac{e^{-(2n+1)2\pi i t} - 1}{e^{-2\pi i t} - 1} \\ &= e^{2n\pi i t} \sum_{k=0}^{2n} (e^{-2\pi i t})^k \\ &= \sum_{|k| \le n} e^{-2k\pi i t}. \end{aligned}$$

Therefore one has

$$\begin{aligned}(S_n f)(t) &= \sum_{|k|\leq n} \left(\int_0^1 f(s)e^{-2k\pi is}ds \right) e^{2k\pi it} \\ &= \sum_{|k|\leq n} \int_0^1 f(s+t)e^{-2k\pi is}ds \\ &= \int_0^1 f(s+t)D_n(s)ds.\end{aligned}$$

This completes the proof. □

In fact, continuity is not even sufficient to ensure *pointwise* convergence of the Fourier series.

Proposition 4.32. *There exists a continuous function* $f : [0,1] \to \mathbb{R}$ *such that* $f(0) = f(1)$ *whose associated Fourier series is divergent in (at least) one point* $x \in [0,1]$.

Proof. Define a family of operators $(T_n)_{n\in\mathbb{N}}$ from $\{f \in C([0,1]) : f(0) = f(1)\}$ to $\mathbb{R}$ by

$$T_n f := (S_n f)(0), \qquad n \in \mathbb{N},\ f \in C([0,1]),$$

where the operators S_n have been introduced in Definition 4.27. We are going to show that the linear operators T_n are bounded with $\lim_{n\to\infty} \|T_n\| = \infty$.

First of all, observe that

$$|T_n f| = \sup_{\|f\|_\infty \leq 1} \left| \int_0^1 f(t)D_n(t)dt \right| \leq \int_0^1 |D_n(t)|dt.$$

Moreover, consider some sequence $(g_k)_{k\in\mathbb{N}}$ of continuous functions from $[0,1]$ to $\mathbb{R}$ such that $g_k(0) = g_k(1)$ for all $k \in \mathbb{N}$ and such that

(1) $\|g_k\|_\infty \leq 1$ for all $k \in \mathbb{N}$ and
(2) $\lim_{k\to\infty} g_k(t)$ agrees with $\text{sign}D_n(t) \in \{-1, 1\}$ (the sign of $D_n(t)$) for all $t \in [0,1]$.

Then one even has $\|T_n\| = \int_0^1 |D_n(t)|dt$, since $\|T_n\| \geq \int_0^1 D_n(t)g_k(t)dt$ for all $k \in \mathbb{N}$ and therefore

$$\|T_n\| \geq \lim_{k\to\infty} \int_0^1 D_n(t)g_k(t)dt = \int_0^1 D_n(t)\text{sign}D_n(t) = \int_0^1 |D_n(t)|dt,$$

where the first identity follows from Lebesgue's Dominated Convergence Theorem. To complete the proof, observe that

$$\begin{aligned}\int_0^1 |D_n(t)|dt &\geq \int_0^1 \frac{|\sin\big((2n+1)\pi t|\big)}{\pi t}dt \\ &= \int_0^{2n+1} \frac{|\sin \pi s|}{\pi s}ds \\ &= \sum_{k=1}^{2n+1} \int_{k-1}^{k} \frac{|\sin \pi s|}{\pi s}ds \\ &\geq \sum_{k=1}^{2n+1} \frac{1}{k\pi} \int_{k-1}^{k} |\sin \pi s|ds \\ &= \frac{2}{\pi^2} \sum_{k=1}^{2n+1} \frac{1}{k}.\end{aligned}$$

Since this last series is divergent, we conclude by the uniform boundedness principle that also $\lim_{n\to\infty} |T_n f| = \lim_{n\to\infty} |S_n f(0)| = \infty$ for some $f \in C([0,1])$. ☐

It should be observed that the assumptions of Theorem 4.29 can be weakened and yield that in fact not only continuously differentiable functions, but also Lipschitz and more generally even Hölder continuous functions can be approximated *uniformly* by the associated Fourier series.

Karl Theodor Wilhelm Weierstraß proved in 1885 that every continuous function defined on a compact interval can be approximated in $\|\cdot\|_\infty$ by a polynomial. Marshall Harvey Stone proved in 1937 the following interesting generalisation. In the following, a subspace A of $C(K,\mathbb{K})$ is called a **subalgebra** if it is closed under all its operations (sum, scalar product, vector product), and carrying the induced operations.

Theorem 4.33 (Stone–Weierstrass Theorem)**.** *Let K be a compact subset of $\mathbb{R}^n$, A a subalgebra of $C(K,\mathbb{K})$, such that*

(1) A contains the constant functions of $C(K,\mathbb{K})$,
(2) if $f \in A$, then $\overline{f} \in A$, and
*(3) A **separates the points of** K, i.e., for all $x, y \in K$ such that $x \neq y$ there exists $g \in A$ such that $g(x) \neq g(y)$.*

Then A is dense in $C(K,\mathbb{K})$.

The second condition is clearly void if $\mathbb{K} = \mathbb{R}$.

In the proof of the Stone–Weierstrass Theorem we will need the following result, which we state without proof – see e.g. [**22**, Satz VI.4.6].

Lemma 4.34. *Let K be a compact subset of $\mathbb{R}^n$, $(f_n)_{n\in\mathbb{N}}$ a sequence in $C(K,\mathbb{R})$. Let this sequence be monotonically increasing, in the sense that $f_n(x) \leq f_{n+1}(x)$ for all $n \in \mathbb{N}$ and all $x \in K$. If the sequence converges pointwise to some $f \in C(K,\mathbb{R})$, then the convergence is also uniform.*

PROOF. Let us first assume that $\mathbb{K}=\mathbb{R}$.

To begin with, we proof an auxiliary result: The mapping $[0,1]\ni t\mapsto\sqrt{t}\in[0,1]$ is the limit (with respect to uniform convergence) of a sequence of polynomials.

This can be seen by defining a sequence $(p_n)_{n\in\mathbb{N}}$ recursively by

$$p_0(t):=0,\qquad p_{n+1}(t):=p_n(t)+\frac{1}{2}(t-p_n(t))^2,\qquad t\in[0,1],\ n\in\mathbb{N}.$$

Then one can prove by induction that $p_n(t)\le\sqrt{t}$ for all $t\in[0,1]$ and all $n\in\mathbb{N}$. This shows that $(p_n(t))_{n\in\mathbb{N}}$ is a bounded and monotonically increasing sequence for all $t\in[0,1]$, hence it converges – and in fact one sees that its limit is $\sqrt{t}$, cf. [**20**, Beispiel I.1.1]. Then the claim follows from Lemma 4.33.

Let us now prove that if $f\in A$, then $|f|:K\ni x\mapsto|f(x)|\in\mathbb{R}_+$ belongs to the closure $\overline{A}$ of A. This is clear if $f=0$. If $f\neq0$, let $(p_n)_{n\in\mathbb{N}}$ be the above introduced approximating sequence and observe that

$$|f|=\|f\|_\infty\sqrt{\frac{f^2}{\|f\|_\infty^2}}=\|f\|_\infty\lim_{n\to\infty}p_n\Big(\frac{f^2}{\|f\|_\infty^2}\Big).$$

This allows to show that if f,g belong to the closure $\overline{A}$ of A, then also $\inf\{f,g\}:=\frac{1}{2}(f+g-|f-g|)$ and $\sup\{f,g\}:=\frac{1}{2}(f+g+|f-g|)$ belong to $\overline{A}$.

Let now $x,y\in K$ with $x\neq y$ and $\alpha,\beta\in\mathbb{R}$. Then there exists $f\in A$ such that $f(x)=\alpha$ and $f(y)=\beta$ – simply take f defined by

$$f(z):=\alpha+(\beta-\alpha)\frac{g(z)-g(x)}{g(y)-g(x)},\qquad z\in K,$$

where g is an element of A that separates x,y, which exists by assumption.

Finally, we are in the position to prove density of A in $C(K,\mathbb{R})$. To this aim, take $\epsilon>0$ and $f\in C(K,\mathbb{R})$. Take moreover $x\in K$. Then for all $y\in K$ we can pick a function $g_y\in A$ such that $g_y(x)=f(x)$ and $g_y(y)=f(y)$. Consider a neighbourhood U_y of each such y such that $g_y\le f+\epsilon$ pointwise in U_y. Clearly $K=\bigcup_{y\in K}U_y$, and due to compactness of K by the Heine–Borel Theorem there exists a finite set $\{y_1,\dots,y_N\}$ such that $K=\bigcup_{k=1}^N U_{y_k}$. Let now $h_x:=\inf_{1\le k\le N}g_{y_k}$: as proved above, $h_x\in\overline{A}$. Then one sees that $h_x(x)=f(x)$ and moreover $h_x\le f+\epsilon$, and moreover $h_x\ge f-\epsilon$ pointwise in a neighborhood V_x of x. Again by the Heine–Borel Theorem there exists a finite set $\{x_1,\dots,x_M\}$ such that $K=\bigcup_{k=1}^M V_{x_k}$. Set now

$$h:=\sup_{1\le k\le M}h_{x_k}$$

in order to conclude the proof, since $h\in\overline{A}$.

Let us now consider the case $\mathbb{K}=\mathbb{C}$. Then the proof can be led back to that of the real-valued case. More precisely, observe that for all $f\in A$ also the real and imaginary parts of f – i.e., $\frac{f+\overline{f}}{2}$ and $\frac{f-\overline{f}}{2i}$ respectively – belong to A. Consider then $A_0:=\{f\in A:f=\overline{f}\}$: since A_0 clearly contains the (real) constant functions of $C(K,\mathbb{K})$ and separates the points of K, it follows that A_0 is dense in $C(K,\mathbb{R})$. The assertion now follows since $A=A_0+iA_0$ – i.e., each function of A can be represented by its real and imaginary parts, two real-valued continuous functions on K. $\square$

Exercise 4.35. The trigonometric polynomials are dense in $L^2(0,1)$, i.e., each function in $L^2(0,1)$ can be approximated (in $\|\cdot\|_2$-norm) by a trigonometric polynomial.

- *Prove this assertion as a consequence of Theorem 4.28.*
- *Prove this assertion as a consequence of Theorem 4.32.*

Observe that even if H is a Hilbert space, $\mathcal{L}(H)$ is in general *not* a Hilbert space. In order to overcome this problem, one sometimes consider a special class of operators, which indeed allows to apply Hilbert space methods again.

Definition 4.36. *Let H be a separable Hilbert space and denote by $(e_n)_{n\in\mathbb{N}}$ a basis of H. Let $T \in \mathcal{L}(H)$. If $(Te_n)_{n\in\mathbb{N}} \in \ell^2$, i.e., if $\sum_{n\in\mathbb{N}} \|Te_n\|^2 < \infty$, then T is called a* **Hilbert–Schmidt operator**. *The class of Hilbert–Schmidt operators on H is commonly denoted by $\mathcal{L}_2(H)$.*

The above definition makes sense because one can in fact show that an operator is Hilbert–Schmidt with respect to a given basis if and only if it is Hilbert–Schmidt with respect to any other basis.

Exercise 4.37. *Let H be a Hilbert space.*

(1) Show that the class $L_2(H)$ of Hilbert–Schmidt operators on a Hilbert space H is a vector space.

(2) Show that $L_2(H)$ becomes a Hilbert space with respect to the inner product

$$(T|S)_{\mathcal{L}_2(H)} := \sum_{n\in\mathbb{N}} (Te_n|Se_n)_H, \qquad T, S \in \mathcal{L}_2(H),$$

for any basis $(e_n)_{n\in\mathbb{N}}$ of H.

(3) Show that if $T \in L_2(H)$, then $\|T\| \le \|T\|_{L_2(H)}$, where as usual $\|T\|^2_{L_2(H)} = (T|T)_{L_2(H)}$.

Exercise 4.38. *Consider the Fredholm operator K introduced in Exercise 2.8 and assume that the kernel k is merely in $L^2((0,1)\times(0,1))$.*

(1) Show that $\|K\|^2 \le \|k\|^2_{L^2((0,1)\times(0,1))}$.

(2) Show that K is a Hilbert–Schmidt operator.

Further properties of Hilbert–Schmidt operators can be found in [**15**, § 30.8].

CHAPTER 5

Functionals and dual spaces

Definition 5.1. *Let X be a Banach space. A linear operator from X to $\mathbb{K}$ is called a* **linear functional on** *X. The space X' of all bounded functionals on X is called* **dual space** *of X.*

Traditionally, the evaluation of a bounded linear functional ϕ at $x \in X$ is denoted by $\langle \phi, x \rangle$, rather than $\phi(x)$ or ϕx. Observe that the operator norm $\|\phi\|$ of $\phi \in X'$ is given by

$$\|\phi\| := \sup_{\|x\|_H \leq 1} |\langle \phi, x \rangle|,$$

so that by definition

$$|\langle \phi, x \rangle| \leq \|\phi\| \|x\|_X \qquad \text{for all } x \in X.$$

Determining the dual space of a Banach space is in general not an easy task. We provide the following relevant example.

Proposition 5.2. *The sequence space ℓ^∞ is isometric isomorphic to the dual space of ℓ^1.*

Proof. Every $x \in \ell^1$ has a unique representation

$$x = \sum_{k \in \mathbb{N}} x_k e_k,$$

where e_k is the ℓ^1-sequence consisting of all zeroes, with the only exception of a 1 in the k-th position. Let now $\phi \in \ell^{1'}$, the dual space of ℓ^1. Since f is linear and bounded, necessarily

$$\langle \phi, x \rangle = \sum_{k \in \mathbb{N}} x_k \langle \phi, e_k \rangle.$$

We do not have many possibilities in order to define an ℓ^∞-sequence out of ϕ: in fact, we are going to take the sequence $(\langle \phi, e_k \rangle)_{k \in \mathbb{N}}$. Observe that

$$|\langle \phi, e_k \rangle| \leq \|\phi\| \|e_k\|_1 = \|\phi\|,$$

whence $(\langle \phi, e_k \rangle)_{k \in \mathbb{N}} \in \ell^\infty$. Thus, we can define an operator $T : \ell^{1'} \to \ell^\infty$ by

$$T\phi := (\langle \phi, e_k \rangle)_{k \in \mathbb{N}}.$$

It is clear that T is linear, and moreover it has been just shown that $\|T\| \leq 1$.

Moreover, we show that T is surjective: if $y \in \ell^\infty$, then we define $\phi \in \ell^{1'}$ by

$$\langle \phi, x \rangle := \sum_{k \in \mathbb{N}} x_k y_k, \qquad x = (x_k)_{n \in \mathbb{N}} \in \ell^1.$$

It is clear that ϕ is linear, and moreover

$$|\langle \phi, x\rangle| \leq \sup_{k\in\mathbb{N}} |y_k| \sum_{k\in\mathbb{N}} |x_k| = \|x\|_1 \|y\|_\infty.$$

Taking the supremum over all x in the unit ball of ℓ^1 shows that $\|\phi\| \leq \|y\|_\infty$, thus actually $\phi \in \ell^{1'}$. This also shows that $\|phi\| \leq \|y\|_\infty = \|T\phi\|_\infty$ for all $f \in \ell^{1'}$, hence T is an isometry, too. Summing up, T is a surjective, isometric bounded linear operator, hence an isometric isomorphism. ☐

Remark 5.3. *Likewise, for $1 < p < \infty$ one introduces the space ℓ^p of all sequences $x = (x_n)_{n\in\mathbb{N}}$ such that*

$$\|x\|_p := \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{\frac{1}{p}} < \infty$$

and shows that the dual space of ℓ^p is isometric isomorphic to ℓ^q provided that $\frac{1}{p} + \frac{1}{q} = 1$.

Similarly, one can show a similar assertion for the Lebesgue spaces: Let Ω be an open subset of $\mathbb{R}^n$. Then the dual space of $L^p(\Omega)$ is isometric isomorphic to $L^q(\Omega)$ provided that $\frac{1}{p} + \frac{1}{q} = 1$.

Example 5.4. *1) Let c denote the space of all sequences $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ that converge. Than $\lim : c \ni x \mapsto \lim_{n\to\infty} x_n \in \mathbb{R}$ is a functional on c.*

2) The mapping $\int : C([0,1]) \ni f \mapsto \int_0^1 \langle f(x) dx$ is a functional on $C([0,1])$.

3) Let H be a Hilbert space. Then for all $y \in H$ the mapping $H \ni x \mapsto (y|x)_H \in \mathbb{K}$ is a bounded (due to the Cauchy–Schwarz inequality) linear functional on H.

The following so-called **representation theorem** finds wide application in mathematics and physics. It has been proved in 1907 by Frigyes Riesz and independently also by Maurice René Fréchet.

Theorem 5.5. *Let H be a Hilbert space. For each bounded linear functional ϕ on H, i.e., for all $\phi \in H'$ there exists a unique $y_\phi \in H$ such that*

$$\langle \phi, x\rangle = (x|y_\phi)_H \qquad \text{for all } x \in H. \tag{5.1}$$

Moreover, the mapping $H \ni \phi \mapsto y_\phi \in H$ is an isometric isomorphism.

Proof. It suffices to prove that $\Phi : H \ni y \mapsto \phi_y := (\cdot|y) \in H'$ is an isometric isomorphism.

To begin with, we prove that Φ is isometric (and therefore injective, too). Clearly, by definition and the Cauchy–Schwarz inequality $|\langle \phi_y, x\rangle| = |(x|y)_H| \leq \|y\|_H \|x\|_H$, so that the norm of ϕ_y satisfies $\|\phi_y\| \leq \|y\|_H$. In order to check the equality in the non-trivial case of $y \neq 0$, take $x := \frac{y}{\|y\|_H}$ and observe that $\langle \phi_y, x\rangle = \|y\|_H$ by definition.

In order to prove surjectivity of Φ, take $\phi \in H'$. If $\text{Ker}\phi = H$, then $\phi = 0$ and the assertion is clear – so we can assume that $\text{Ker}\phi \neq H$ and (up to rescaling) that $\|\phi\| = 1$. By Remark 2.14, one has $H = \text{Ker}\phi \oplus \text{Ker}\phi^\perp$. Moreover, the closed subspace $\text{Ker}\phi^\perp$ has dimension 1, since the restriction of ϕ to $\text{Ker}\phi^\perp$ is an isomorphism from $\text{Ker}\phi^\perp$ to $\mathbb{K}$. Accordingly, there exists $\xi \in \text{Ker}\phi^\perp$ – which up to rescaling can be assumed to satisfy $\langle \phi, \xi\rangle = 1$ – such that each

$z \in \mathrm{Ker}\phi^\perp$ has the form $z = \lambda\xi$ for some $\lambda \in \mathbb{K}$, and in particular each $x \in H$ admits the decomposition $x = P_{\mathrm{Ker}\phi}x \oplus \lambda\xi$. Then

$$\langle \phi, x\rangle = \langle \phi, P_{\mathrm{Ker}\phi}x + \lambda\xi\rangle = \lambda\langle \phi, \xi\rangle = \lambda$$

as well as

$$(x|\xi)_H = (P_{\mathrm{Ker}\phi}x + \lambda\xi|\xi)_H = \lambda\|\xi\|_H^2,$$

where we have used the fact that $P_{\mathrm{Ker}\phi}x$ and ξ belong to subspaces that are orthogonal to each other. We deduce that

$$\langle \phi, x\rangle = \lambda = \frac{(x|\xi)_H}{\|\xi\|_H^2} =: (x|y_\phi)_H,$$

for all $x \in H$, and we conclude that Φ is sujective. This concludes the proof. □

Remark 5.6. *By the Representation theorem of Riesz–Fréchet, the dual H' of a Hilbert space H may always be identified with H. However, one does not necessarily have to do so – in fact, it is not always a smart idea to consider $H = H'$. An interesting discussion about pros and cons can be found in* [**2**, § V.2].

Example 5.7. *Let H be a separable Hilbert space. Then each functional in $L_2(H)'$ can be represented by a suitable Hilbert–Schmidt operator A, i.e., by*

$$L_2(H) \ni B \mapsto (A|B)_{L_2(H)} \in \mathbb{C}.$$

An interesting application of the Representation theorem of Riesz–Fréchet concerns the solvability of **elliptic equations** like

$$u(x) - u''(x) = f(x), \qquad x \in [0, 1], \tag{5.2}$$

with – say – Dirichlet boundary conditions

$$u(0) = u(1) = 0. \tag{5.3}$$

One of the most fruitful mathematical ideas of the last century is the weak formulation of differential equations. One weakens the notion of solution of a boundary/initial value differential problem, looks for a solution in a suitably larger class (which in turn allows to use standard Hilbert space methods, like the Representation theorem) and eventually proves that the obtained solution is in fact also a solution in a classical sense. This method is thoroughly explained in [**2**] and [**21**]. The essential idea behind this approach is that of *weak derivative*, one which is based on replacing the usual property of differentiability by one prominent quality of differentiable functions – the possibility to integrate by parts.

Definition 5.8. *Let $I \subset \mathbb{R}$ be an open interval. A function $f \in L^2(I)$ is said to be* **weakly differentiable** *if there exists $g \in L^2(I)$ such that*

$$\int_I f(x)\overline{h'(x)}dx = -\int_I g(x)\overline{h(x)} \qquad \text{for all } h \in C_c^1(I). \tag{5.4}$$

The set of weakly differentiable functions $f \in L^2(I)$ such that their weak derivative is also in $L^2(I)$ is denoted by $H^1(I)$ and called **first Sobolev space**. *They were introduced in 1936 by Sergei Lvovich Sobolev.*

Remark 5.9. *Observe that since any two continuosly differentiable functions f, h satisfy (5.4) (which is nothing but the usual formula of integration by parts), by definition $C^1(\overline{I}) \subset H^1(I)$ – i.e., each function in $C^1(\overline{I})$ is representative of a weakly differentiable L^2-function whose weak derivative is again in L^2.*

For an open interval $I \subset \mathbb{R}$ we have here denoted by $C^1_c(I)$ the vector space of continuously differentable functions with compact support, i.e., continuously differentiable functions $f : I \to \mathbb{K}$ such that $f(x) = f'(x) = 0$ for all x outside some compact subset of I.

Exercise 5.10. *Let $I \subset \mathbb{R}$ be an open interval. Let $f \in L^2(I)$. Show that if a function g satisfying (5.4) exists, then it is unique.*

This motivates to introduce the following.

Definition 5.11. *Let $I \subset \mathbb{R}$ be an open interval. Let $f \in L^2(I)$ be weakly differentiable. The unique function g introduced in Definition 5.8 is called the* **weak derivative** *of f and with an abuse of notation we write $f' = g$.*

Remark 5.12. *Consider the momentum operator $S : f \mapsto f'$. We have already remarked that S is not bounded on $C^1([0,1])$ – and in fact, also on $L^2(0,1)$. However, it is clear that S is a bounded linear operator from $H^1(0,1)$ to $L^2(0,1)$.*

Example 5.13. *Let $I = (-1,1)$. The prototypical case of a weakly differentiable function that does not admit a classical derivative in some point is given by*

$$f(x) := \frac{|x| + x}{2} = \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } x > 0. \end{cases}$$

Take some function $h \in C^1_c((-1,1))$ and observe that

$$\begin{aligned} \int_{-1}^{1} f(x)\overline{h'(x)}dx &= \int_{-1}^{0} f(x)\overline{h'(x)}dx + \int_{0}^{1} f(x)\overline{h'(x)}dx \\ &= \int_{0}^{1} x\overline{h'(x)}dx \\ &= \Big[xh(x)\Big]_0^1 - \int_0^1 \overline{h(x)}dx \\ &= -\int_0^1 \overline{h(x)}dx, \end{aligned}$$

where the last equality follows from compactness of support of h (whence $h(1) = 0$). In other words,

$$\int_{-1}^{1} f(x)\overline{h'(x)}dx = \int_{-1}^{1} H(x)\overline{h(x)}dx$$

for all $h \in C^1_c(I)$, where H is defined by

$$H(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0. \end{cases}$$

This shows that f is weakly differentiable with $f' = H$, where H is called the **Heaviside function** *after Oliver Heaviside.*

Integrating by parts one clearly sees that each continuously differentiable function is also weakly differentiable, i.e., $C^1(\overline{I}) \subset H^1(I)$. In general a function that is merely in $C^1(I)$ need not be in $L^2(I)$, but in fact each $u \in C^1(I) \cap L^2(I)$ such that $u' \in L^2(I)$ also belongs to $H^1(I)$.

Lemma 5.14. *Let I be an interval. Then the set $H^1(I)$ is a separable Hilbert space with respect to the inner product*

$$(f|g)_{H^1(I)} := (f|g)_{L^2(I)} + (f'|g')_{L^2(I)} = \int_0^1 f(x)\overline{g(x)}dx + \int_0^1 f'(x)\overline{g'(x)}dx.$$

Proof. It is easy to see that $(\cdot|\cdot)_{H^1(I)}$ is an inner product, since in particular $(\cdot|\cdot)_{L^2(I)}$ is. In order to show completeness, take a Cauchy sequence in $H^1(I)$, i.e., a sequence $(f_n)_{n\in\mathbb{N}}$ of weakly differentiable functions such that both $(f_n)_{n\in\mathbb{N}}$ and $(f'_n)_{n\in\mathbb{N}}$ are Cauchy in $L^2(I)$. By completeness of $L^2(I)$ both sequences converge, say to ϕ, ψ respectively. Furthermore,

$$\int_I f_n(x)\overline{h'(x)}dx = -\int_I f'_n(x)\overline{h(x)} \qquad \text{for all } h \in C^1_c(I),$$

so that

$$\int_I \phi(x)\overline{h'(x)}dx = \lim_{n\to\infty}\int_I f_n(x)\overline{h'(x)}dx = -\lim_{n\to\infty}\int_I f'_n(x)\overline{h(x)}dx = -\int_I \psi(x)\overline{h(x)}dx \qquad \text{for all } h \in C^1_c(I).$$

This shows that $\phi' = \psi$, so that $H^1(I)$ is in fact a Hilbert space.

Consider the bounded operator $H^1(I) \ni f \mapsto (f, f') \in L^2(I) \times L^2(I)$. It is apparent that this operator is bounded and in fact an isometry whose range is closed in the product Hilbert space $L^2(I) \times L^2(I)$ (cf. Exercise 3.2), hence by Exercise 4.5 a separable Hilbert space. ☐

Exercise 5.15. *Let I be an interval.*

(1) Let $f \in L^2(I)$ such that $\int_I f(x)\overline{h'(x)}dx = 0$ for all $h \in C^1_c(\overline{I})$. Show that there exists a constant $c \in \mathbb{K}$ such that $f(x) = c$ for a.e. $x \in I$.

(2) Let $g \in L^2(I)$ and $x_0 \in I$. Define $G : I \ni x \mapsto \int_{x_0}^x g(t)dt \in \mathbb{K}$. Show that $G \in C(I)$ and moreover $\int_I G(x)\overline{h'(x)}dx = -\int_I g(x)\overline{h(x)}dx$ for all $h \in C^1_c(I)$.

(3) Conclude that each $f \in H^1(I)$ has a continuous representative $\tilde{f}$ in $C(\overline{I})$ such that $f(x) = \tilde{f}(x)$ for a.e. $x \in I$. Moreover, $\|f\|_{C(\overline{I})} \le \|f\|_{H^1(I)}$ for all $f \in C([0,1])$.

Remark 5.16. *Accordingly, it is common (although a slight abuse of language) to say that* weakly differentiable functions are continuous. *It is worthwile to emphasize that this important property is exclusive of the 1-dimensional case, even though Sobolev spaces can also be introduced for functions acting on subsets of $\mathbb{R}^n$ for $n > 1$. In particular, this allows to talk about point evaluation of functions in $H^1(I)$.*

The converse is not true, i.e., there exist continuous functions that are not weakly differentiable. In fact, it can be proved that a weakly differentiable function is differentiable (in the classical sense) almost everywhere. Hence, a continuous but nowhere differentiable function *will do the job: this is the famous* **Weierstraß function***, introduced by Karl Theodor Wilhelm Weierstraß in 1872.*

Exercise 5.17. *Prove the so-called* **Poincaré inequality***: For all* $f \in H^1(0,1)$ *such that* $f(0) = f(1) = 0$ *there holds*

$$\|f\|_2^2 \leq \|f'\|_2^2.$$

Let $I \subset \mathbb{R}$ be an open interval, say $I = (a,b)$. Motivated by the above exercise we introduce the vector space

$$H_0^1(I) := \{f \in H^1(I) : f(a) = f(b) = 0\}.$$

Definition 5.18. *A function* $u \in H_0^1(I)$ *is called a* **weak solution** *of the elliptic problem* (5.2)–(5.3) *if*

$$\int_0^1 u(x)\overline{h(x)}dx + \int_0^1 u'(x)\overline{h'(x)}dx = \int_0^1 f(x)\overline{h(x)}dx \qquad \text{for all } h \in C_c^1(I). \tag{5.5}$$

This definition does not come out of the blue. In fact, each "classical" solution u of (5.2)–(5.3), once integrated "against a C^1-function" (i.e., upon multiplying it by any $h \in C_c^1(I)$ and the integrating over I) satisfies (5.5) after integration by parts.

Theorem 5.19. *For all* $f \in L^2(I)$ *the elliptic problem* (5.2)–(5.3) *has a unique weak solution.*

Proof. Consider the mapping $\phi : H_0^1(I) \ni h \mapsto (h|f)_{L^2(I)} \in \mathbb{K}$, which is clearly linear. This linear functional is bounded, since

$$|\langle \phi, h\rangle| = |(h|f)_{L^2(I)}| \leq \|h\|_{L^2(I)}\|f\|_{L^2(I)} \leq \|h\|_{H_0^1(I)}\|f\|_{L^2(I)}$$

for all $h \in H^1(I)$. Then by the Representation theorem of Riesz–Fréchet there exists a unique $u \in H_0^1(I)$ (continuously depending on f) such that $\langle \phi, h\rangle = (u|h)_{H^1(I)} = (u|h)_{L^2(I)} + (u'|h')_{L^2(I)}$. By definition of weak solution, this completes the proof. ☐

Sometimes it is necessary to deal with elliptic problems that do not have the nice symmetric structure assumed in the Representation theorem. However, it is often still possible to apply the following result.

Exercise 5.20. *Let* A *be a bounded linear operator on a Hilbert space* H *such that*

$$|(Ax|x)_H| \geq \alpha\|x\|^2 \qquad \text{for all } x \in H$$

for some $\alpha > 0$.

(1) Show that $\|Ax\| \geq \alpha\|x\|$ *for all* $x \in H$*. (Hint: use the Cauchy–Schwarz inequality.)*
(2) Show that A *is injective and* $A(H)$ *is closed.*
(3) Show that $\overline{A(H)} = H$*. (Hint: Show that* $(A(H))^\perp = \{0\}$*.)*
(4) Conclude that A *is invertible.*

The statement of the above exercise can be reformulated as follows. This result has been obtained in 1954 by Peter David Lax and Arthur Norton Milgram and is commonly known as the **Lax–Milgram Lemma**. In the following, a a (bi)linear mapping from $V \times V$ to $\mathbb{K}$, V a Hilbert space, is called **coercive** if there exists $c > 0$ such that $\mathrm{Re} a\,(u,u) \geq c\|u\|^2$ for all $u \in V$.

Exercise 5.21. *Let V be a Hilbert space and a a (bi)linear mapping from $V \times V$ to $\mathbb{K}$. Let a be bounded and coercive. Then, for any $\phi \in V'$ there is a unique solution $u =: T\phi \in V$ to $a(u,v) = \langle \phi, v\rangle$ – which also satisfies $\|u\| \leq \frac{1}{c}\|\phi\|_{V'}$. Moreover, T is an isomorphism from V' to V.*

Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. It is sometimes useful for applications to consider an operator T^* such that

$$(Tx|y)_{H_2} = (x|T^*y)_{H_1} \qquad \text{for all } x \in H_1 \text{ and } y \in H_2. \tag{5.6}$$

Proposition 5.22. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. Then there exists exactly one bounded linear operator T^* from H_2 to H_1 such that* (5.6) *holds. Moreover, there holds $\|T\| = \|T^*\|$.*

Proof. Let $y \in H_2$. Then $\phi_y : H_1 \ni x \mapsto (Tx|y)_{H_2} \in \mathbb{K}$ defines a bounded linear functional, since by the Cauchy–Schwarz inequality $|\phi_y(x)| \leq \|T\|\|x\|_{H_1}\|y\|_{H_2}$. Therefore, by Theorem 5.5 there exists a vector $T^*y \in (H_1)' \cong H_1$ such that $(Tx|y)_{H_2} = \langle \phi_y, x\rangle = (x|T^*y)_{H_1}$. This defines an operator $T^* : H_2 \ni y \mapsto T^*y \in H_1$.

To check linearity of T^*, take $y_1, y_2 \in H_2$ and observe that for all $x \in H_1$

$$\begin{aligned}(x|T^*(y_1+y_2) - T^*y_1 - T^*y_2)_{H_1} &= (x|T^*(y_1+y_2))_{H_2} - (x|T^*y_1)_{H_1} - (x|T^*y_2)_{H_1} \\ &= (Tx|y_1+y_2)_{H_1} - (Tx|y_1)_{H_2} - (Tx|y_2)_{H_2} = 0.\end{aligned}$$

Accordingly, $T^*(y_1+y_2) - T^*y_1 - T^*y_2$ belongs to $H^\perp = \{0\}$ for all y_1, y_2. Similarly, take $y \in H_2$ and $\lambda \in \mathbb{K}$ and observe that for all $x \in H_1$

$$\begin{aligned}(x|T^*(\lambda y) - \lambda T^*y)_{H_1} &= (x|T^*(\lambda y))_{H_2} - (x|\lambda T^*y)_{H_1} \\ &= (Tx|\lambda y)_{H_1} - \overline{\lambda}(Tx|y) = 0,\end{aligned}$$

i.e., $T^*(\lambda y) - \lambda T^*y$ belongs to $H^\perp = \{0\}$ for all $y \in H_2$ and all $\lambda \in \mathbb{K}$.

Boundedness of T^* follows by boundedness of T, since for all $y \in H_2$

$$\|T^*y\|_{H_1} \leq \|\phi_y\| \leq \|T\|\|y\|_{H_2}.$$

This shows that $\|T^*\| \leq \|T\|$. Conversely, take $x \in H_1$ with $\|x\|_{H_1} \leq 1$ and observe that

$$\|Tx\|^2_{H_2} = (Tx|Tx)_{H_2} = (T^*Tx|x)_{H_1} \leq \|T^*Tx\|_{H_1}\|x\|_{H_1} \leq \|T^*T\| \leq \|T^*\|\|T\|.$$

Accordingly, $\|T\|^2 \leq \|T^*\|\|T\|$ and in particular $\|T\| \leq \|T^*\|$. This completes the proof. □

Exercise 5.23. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. Show that $T^{**} = T$.*

Definition 5.24. *Let H_1, H_2 be Hilbert spaces and $T \in \mathcal{L}(H_1, H_2)$. The unique operator $T^* \in \mathcal{L}(H_2, H_1)$ that satisfies* (5.6) *is called the* **adjoint** *of T.*

If $H_1 = H_2$ and $T^* = T$, then the operator T is called **self-adjoint**.

Corollary 5.25. *Let H be a Hilbert space and T be a bounded linear self-adjoint operator on H. Then $(Tx|x)_H \in \mathbb{R}$ for all $x \in H$ and $\|T\| = \sup_{\|x\|_H=1} |(Tx|x)_H|$.*

Proof. One sees that $(Tx|x)_H \in \mathbb{R}$ for all $x \in H$ because

$$(Tx|x)_H = \overline{(x|Tx)_H} = \overline{(x|T^*x)_H} = \overline{(Tx|x)_H}.$$

To begin with, it clearly follows from the Cauchy–Schwarz inequality and the definition of norm that $\sup_{\|x\|_H=1} |(Tx|x)_H| \leq \|T\|$.

Conversely, one easily checks that $(T(x+y)|x+y)_H-(T(x-y)|x-y)_H = 2(Tx|y)_H+2(Ty|x)_H$ and $2\mathrm{Re}(Tx|y)_H = (Tx|y)_H + (Ty|x)_H$ for all $x, y \in H$. Accordingly, for all $x, y \in H$ such that $\|x\|_H = \|y\|_H = 1$ we have

$$4\mathrm{Re}(Tx|y)_H \leq \sup_{\|x\|_H=1} |(Tx|x)_H| \left(\|x+y\|_H^2 + \|x-y\|_H^2\right) \leq 2 \sup_{\|x\|_H=1} |(Tx|x)_H| \left(\|x\|_H^2 + \|y\|_H^2\right),$$

by the parallelogram law and because

$$|(Tx|x)_H| = |(T\frac{x}{\|x\|_H}|\frac{x}{\|x\|_H})_H|\|x\|_H^2 \leq \sup_{\|z\|_H=1} |(Tz|z)_H|\|x\|_H^2 \text{ for all } 0 \neq x \in H.$$

We conclude that for all $x, y \in H$ such that $\|x\|_H = \|y\|_H = 1$ we have

$$\mathrm{Re}(Tx|y)_H = \mathrm{Re}|(Tx|y)_H|e^{i\arg(Tx|y)_H} = |(Tx|y)_H|\mathrm{Re}(Tx|e^{-i\arg(Tx|y)_H}y)_H \leq \sup_{\|z\|_H=1} |(Tz|z)_H|.$$

Since this holds for all y, by linearity one concludes that

$$|(x|Ty)|_H = |(Tx|y)_H| \leq \sup_{\|z\|_H=1} (Tz|z) \text{ for all } y \in H \text{ such that } \|y\|_H \leq 1.$$

It follows from the Cauchy–Schwarz inequality that

$$\|T\| = \sup_{\|y\|_H\leq 1} \|Ty\| = \sup_{\|y\|_H\leq 1} \sup_{\|x\|_H=1} \|Ty\| \leq \sup_{\|x\|_H=1} (Tx|x).$$

This concludes the proof. □

Remark 5.26. *The set $N := \{(Tx|x)_H : x \in H,\ \|x\|_H = 1\}$ is called* **numerical range** *of $T \in \mathcal{L}(H)$. There is a rich theory on numerical ranges with several applications in numerical analysis, see e.g.* [**7**] *and* [**10**, Chapter 22].

Example 5.27. *If $H = \mathbb{C}^n$ and $T \in \mathcal{L}(H) = M_n(\mathbb{C})$, then $T^* = \overline{T^T}$, where T^T denotes the transposed matrix of T.*

Example 5.28. *The operators* 0 *and* Id *are both self-adjoint.*

An important class of self-adjoint operators is given by orthogonal projections.

Exercise 5.29. *Let H be a Hilbert space and $P \in \mathcal{L}(H)$ be a projection, i.e., $P^2 = P$. Show that P is the orthogonal projection of H onto some closed subspace $A \subset H$ if and only if P is self-adjoint.* (Hint: P is an orthogonal projection if and only if its null space and its range are orthogonal to each other.

Example 5.30. *For any essentially bounded measurable* $q : \Omega \to \mathbb{R}$ *the multiplication operator* M_q *is self-adjoint. More generally, for any essentially bounded measurable* $q : \Omega \to \mathbb{C}$ *the adjoint of* M_q *is given by* $M_q^* = M_{\overline{q}}$, *where* $\overline{q}$ *is defined by* $\overline{q}(x) := \overline{q(x)}$ *a.e.*

Example 5.31. *For any continuous* $k : [0,1] \times [0,1] \to \mathbb{R}$ *that is symmetric (i.e., such that* $k(x,y) = k(y,x)$*) the Fredholm operator* F_k *is self-adjoint. More generally, for any continuous* $k : [0,1] \times [0,1] \to \mathbb{C}$ *the adjoint of* F_k *is given by* $F_k^* = F_{k^*}$, *where* k^* *is defined by* $k^*(x,y) := \overline{k(y,x)}$ *a.e.*

Exercise 5.32. *Let* H_1, H_2, H_3 *be Hilbert spaces and* $T, S \in \mathcal{L}(H_1, H_2)$, $R \in \mathcal{L}(H_2, H_3)$. *Prove the following assertions.*
(1) $(T + S)^* = T^* + S^*$;
(2) $(\lambda T)^* = \overline{\lambda} T$ *for all* $\lambda \in \mathbb{K}$;
(3) $(RT)^* = T^* R^*$;
(4) $\operatorname{Ker}(\lambda \operatorname{Id} - T) = \operatorname{Ran}(\overline{\lambda} \operatorname{Id} - T^*)^\perp$;
(5) T^* *is invertible if* T *is, and in this case* $(T^*)^{-1} = (T^{-1})^*$.

Remark 5.33. *Beside the already discussed vector space structure, the set* $\mathcal{L}(H)$ *of all bounded linear operators on a Hilbert space* H *has also other significant properties. In particular, since it is closed under composition of its elements (cf. Remark 2.3) it qualifies as an algebra. We have seen that* $\mathcal{L}(H)$ *is even closed under a further operation, the so-called* involution $A \mapsto A^*$. *These both properties (together with a handful of more technical ones) turn* $\mathcal{L}(H)$ *into a* **C^*-algebra**, *a notion introduced by Israel Gelfand and Mark Neumark in 1943 and simplified by Irving Kaplansky in 1952. Nowadays the theory of* C^**-algebras is one of the richest and most vital fields of functional analysis, but a thorough description of its main theorems goes far beyond the scope of this course.*

Exercise 5.34. *Let* H_1, H_2 *be Hilbert spaces and* T *be a bounded linear operator from* H_1 *to* H_2.
(1) Show that the **graph** *of* T, *i.e.,* $\{(f,g) \in H_1 \times H_2 : Tf = g\}$ *is a closed subspace of* $H_1 \times H_2$.
(2) Prove that the subspace of $H_1 \times H_2$ *orthogonal to the graph of* T *is given by* $\{(f,g) \in H_1 \times H_2 : f = -T^* g\}$.
(3) Deduce from Exercise 5.20 that both $I + TT^*$ *and* $I + T^*T$ *are invertible.*
(4) Conclude that the orthogonal projection of $H_1 \times H_2$ *onto the graph of* T *is given by the operator matrix*

$$\begin{pmatrix} (I + T^*T)^{-1} & T^*(I + TT^*)^{-1} \\ T(I + T^*T)^{-1} & I - (I + TT^*)^{-1} \end{pmatrix},$$

in the sense of Example 2.7.

The above formula for the projection onto the graph of an operator has been first obtained in 1950 by mathematician, quantum physicist and early computer scientist John von Neumann in [**17**].

CHAPTER 6

Compactness and spectral theory

We recall the following.

Definition 6.1. *Let (X, d) be a metric space. A subset $M \subset X$ is called sequentially compact, or simply* **compact***, if each sequence in M has a subsequence that converges to an element of M. A subset of X is called* **precompact** *if its closure is compact.*

Remark 6.2. *It can be proved that a subset M of a metric space X is precompact if and only if for any $\epsilon > 0$ there exists a finite subset $M' \subset M$ such that $X = \bigcup_{x \in M'} B_\epsilon(x)$. This is sometimes called* **Heine–Borel condition***, after Émile Borel and Eduard Heine.*

Exercise 6.3. *Show that a subset of a metric space is compact if and only if it is closed and precompact.*

The following assertion is a special case of **Tychonoff's theorem**, proved by Andrey Nikolayevich Tychonoff in 1930 in the case of general topological spaces.

Exercise 6.4. *Let (X, d) be a metric space. Let K_1, K_2 be compact subsets of X. Show that $K_1 \times K_2$ is compact in the product space $X \times X$ with respect to the metric defined by*

$$d_\times((x_1, x_2), (y_1, y_2)) := d(x_1, y_1) + d(x_2, y_2), \qquad x_1, x_2, y_1, y_2 \in X.$$

Exercise 6.5. *Let H be a Hilbert space and $A, B \subset H$ be nonempty, convex and disjoint. Let additionally A, B be closed and compact, respectively. Deduce from the Hahn–Banach Theorem that there exists $x^* \in H$ and some $\alpha, \beta \in \mathbb{R}$ such that $(x^*|x)_H \leq \alpha < \beta \leq (x^*|y)_H$ for all $x \in A$ and all $y \in B$.*

The following fixed point theorem is a consequence of the above corollary of the Hahn–Banach Theorem. It is a special case of the fixed point theorem introduced by Juliusz Schauder in 1927. In Schauder's general version, the relevant function T is not required to be affine.

Exercise 6.6. *Consider a Hilbert space H and a convex compact nonempty subset $K \subset H$. Let a mapping $T : K \to K$ be continuous and affine, i.e., assume that $T(\lambda x + (1 - \lambda)y) = \lambda Tx + (1 - \lambda)Ty$ for all $x, y \in K$ and all $\lambda \in [0, 1]$. Show that there exists at least one* **fixed point** *of T, i.e., one $x \in K$ such that $Tx = x$.*

(Hint: Observe that $\{(x, x) \in K \times K : x \in K\}$ and $\{(x, Tx) \in K \times K : x \in K\}$ are convex and compact by Tychonoff's theorem. Show that if T had no fixed point, then these sets could be separated by the theorem of Hahn–Banach. Show that this leads to the construction of $x^* \in H$ such that the sequence $((x^*|T^n x)_H)_{n \in \mathbb{N}}$ is unbounded for some $x \in K$).

In order to characterize spaces in which all bounded closed subsets are compact we need the following result, proved by Frigyes Riesz in 1918. It is commonly known as Riesz' Lemma.

Lemma 6.7. *Let X be a normed vector space and Y be a closed subspace different from either $\{0\}$ or X. Then for all $r < 1$ there exists $x_{(r)} \in X$ with $\|x_{(r)}\|_X = 1$ such that $\inf_{y\in Y} \|x_{(r)} - y\|_X > r$.*

Proof. Since there exists $x \in X \setminus Y$, and hence with $\inf_{y\in Y} \|x - y\|_X > 0$ due to closedness of Y, there exists $y_0 \in Y$ such that $\|x - y_0\| < \frac{\inf_{y\in Y} \|x-y\|}{r}$ (observe that $\inf_{y\in Y} \|x - y\| < \frac{\inf_{y\in Y} \|x-y\|}{r}$, hence $\frac{\inf_{y\in Y} \|x-y\|}{r}$ is not a lower bound of $\{\|x - y\| : y \in Y\}$). Now, it suffices to set $x_{(r)} := \frac{x-y_0}{\|x-y_0\|_X}$, since in particular

$$\inf_{y\in Y} \|x_{(r)} - y\|_X = \inf_{y\in Y} \left\| \frac{x - y_0}{\|x - y_0\|_X} - y \right\|_X = \frac{1}{\|x - y_0\|_X} \inf_{y\in Y} \|x - y_0 - y\|_X$$

(in the last step we have used the fact that Y is a linear subspace). Since $y_0 \in Y$, we conclude that

$$\inf_{y\in Y} \|x_{(r)} - y\|_X = \frac{1}{\|x - y_0\|_X} \inf_{y\in Y} \|x - y\|_X > r$$

by construction. □

The next result follows directly.

Proposition 6.8. *Let X be a normed vector space. Then the following assertions are equivalent.*

(i) The closed unit ball $\overline{B_1(0)}$ is compact.
(ii) Each bounded closed subset of X is compact.
(iii) X is isomorphic to $\mathbb{K}^n$ for some $n \in \mathbb{N}$.

Proof. *(i)* ⇒ *(iii)* We are going to show that if X is not isomorphic to $\mathbb{K}^n$ for any $n \in \mathbb{N}$, i.e., if X is not finite dimensional, then $\overline{B_1(0)}$ is not compact, i.e., there exists a sequence $(x_n)_{n\in\mathbb{N}}$ with $\|x_n\|_X \le 1$ such that $\|x_n - x_m\|_X \ge \frac{1}{2}$ for all $n, m \in \mathbb{N}$. Take $x_1 \in X$ with $\|x_1\|_X = 1$ and consider the linear span $Y_1 = \{\alpha_1 x_1 \in X : \alpha_1 \in \mathbb{K}\}$ of $\{x_1\}$. By Lemma 6.7 there exists $x_2 \in X$ such that $\|x_2\|_X = 1$ and $\|\alpha_1 x_1 - x_2\| \ge \frac{1}{2}$ for all $\alpha_1 \in \mathbb{K}$. Let us now consider the linear span $Y_2 = \{\alpha_1 x_1 + \alpha_2 x_2 \in X : \alpha_1, \alpha_2 \in \mathbb{K}\}$ of $\{x_1, x_2\}$. Again by Lemma 6.7 there exists $x_3 \in X$ such that $\|x_3\|_X = 1$ and $\|\alpha_1 x_1 + \alpha_2 x_2 - x_3\| \ge \frac{1}{2}$ for all $\alpha_1, \alpha_2 \in \mathbb{K}$. Proceding in this way, we obtain a sequence $(x_n)_{n\in\mathbb{N}}$ with $\|x_n\|_X \le 1$ such that $\|x_n - x_m\|_X \ge \frac{1}{2}$ for all $n, m \in \mathbb{N}$, i.e., such that no subsequence satisfies the Cauchy condition.

This completes the proof, since *(ii)* ⇒ *(i)* is clear and *(iii)* ⇒ *(ii)* is well-known Heine–Borel Theorem. □

The following fundamental result has been established by Giulio Ascoli in 1883 and by Cesare Arzelà in 1895, respectively. It is commonly known as **theorem of Ascoli–Arzelà**.

Theorem 6.9. *Let $K \subset \mathbb{R}^n$ be compact. Consider $M \subset C(K)$ such that*

- *M is* **pointwise bounded**, *i.e., $\sup_{f\in M} |f(x)| < \infty$ for all $x \in K$ and*
- *M is* **uniformly equicontinuous**, *i.e., for all $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $f \in M$ and all $x, y \in K$ with $|x - y| < \delta$.*

Then M is a precompact subset of $C(K)$.

Proof. Observe that $\mathbb{Q}^n \cap K$ is a countable dense subset of K, hence we define $\{x_m : m \in \mathbb{N}\} := \mathbb{Q}^n \cap K$. We take a sequence $(f_n)_{n\in\mathbb{N}}$ and want to show that it contains a convergent subsequence. Now observe that $(f_n(x_1))_{n\in\mathbb{N}}$ is a bounded sequence, due to pointwise boundedness of M. Hence, by the Theorem of Bolzano–Weierstraß it contains a convergent subsequence, say $(f_{n_k^1}(x_1))_{k\in\mathbb{N}}$. Likewise, also $(f_n(x_2))_{n\in\mathbb{N}}$ contains a convergent subsequence, say $(f_{n_k^2}(x_2))_{k\in\mathbb{N}}$, and so on. This suggests to apply Cantor's diagonal method: Take the sequence $(f_{n_k^k})_{k\in\mathbb{N}}$ of functions in M and recall that by construction

$$(6.1) \qquad (f_{n_k^k}(x_m))_{k\in\mathbb{N}} \quad \text{converges for all } m \in \mathbb{N}.$$

It now suffices to show that $(f_{n_k^k})_{k\in\mathbb{N}}$ is a Cauchy sequence.

Let $\epsilon > 0$. Since M is by assumption equicontinuous, there exists $\delta > 0$ such that for all $f \in M$ one has $|f(x) - f(y)| < \epsilon$ for all $x, y \in K$ such that $|x - y| < \delta$. By the Theorem of Heine–Borel there exists a finite covering of K consisting of open balls $V_1, \ldots, V_p$ with radius $\frac{\delta}{2}$. Since $\{x_m : m \in \mathbb{N}\}$ is a countable dense subset of K, each x_m is contained in at least one of these open balls – say, m_1 is the least index such that $x_{m_1} \in V_1$, m_2 is the least index larger than m_1 such that $x_{m_2} \in V_2$ and so on, until $x_{m_h} \in V_h$. Now the assertion follows from a 3ϵ-argument: let $p, q \in \mathbb{N}$ large enough, one has for an arbitrary $x \in K$ – say, $x \in V_h$ – that

$$|f_{n_p^p}(x) - f_{n_q^q}(x)| \le |f_{n_p^p}(x) - f_{n_p^p}(x_{m_h})| + |f_{n_p^p}(x_{m_h}) - f_{n_q^q}(x_{m_h})| + |f_{n_q^q}(x_{m_h}) - f_{n_q^q}(x)| < 3\epsilon,$$

for p, q large enough: the first and third terms in the right hand side can be estimated using uniform equicontinuity (due to the fact that $|x - x_{m_h}| < \delta$) and the second term by (6.1). This concludes the proof. ☐

It is worth to remark that the class of pointwise bounded, equicontinuous function sets exhausts in fact the precompact subsets of $C(K)$.

Exercise 6.10. *Let $K \subset \mathbb{R}^n$ be compact and let $M \subset C(K)$.*

(1) Assume that M is **equicontinuous***, i.e., for all $\epsilon > 0$ and all $x \in K$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $f \in M$ and all $y \in K$ with $|x - y| < \delta$. Show that M is uniformly equicontinuous.*

(2) Show that if M is precompact, then it is equicontinuous.

(3) Conclude that the converse of Theorem 6.9 holds, i.e., the Ascoli–Arzelà condition is in fact a characterization.

Exercise 6.11. *Let $I \subset \mathbb{R}$ be a bounded open interval. Show that the sets*

- $C^1(\overline{I})$ *and*
- $\{f \in C^1(\overline{I}) : \|f\|_\infty + \|f'\|_\infty < \infty\}$

are precompact in $C(\overline{I})$. Is either of them also closed, hence compact?

Definition 6.12. *Let X, Y be Banach spaces. An operator T from X to Y is called* **compact** *if $TB_1(0)$ is precompact, i.e., if the closure of the image of the unit ball of X under T is compact in Y.*

In other words, an operator is compact if and only if it maps each bounded sequences into sequences that contain a convergent subsequence – i.e., if it maps bounded sets into precompact sets. In particular, *linear* compact operators will play an important role in spectral theory.

Exercise 6.13. *Let $X = \mathbb{K}^n$ and $Y = \mathbb{K}^m$. Show that a compact operator from X to Y is simply a continuous mapping.*

Proposition 6.14. *Let X, Y, W, Z be Banach spaces and T be a compact linear operator from Y to W. Let additionally S, R be bounded linear operators from X to Y and from W to Z, respectively. Then both TS and RT are compact (from X to W and from Y to Z, respectively).*

Proof. First, observe that since S is bounded, also $SB_1(0)$ is bounded in Y (where $B_1(0)$ denotes the unit ball in X), and hence by definition of compact operator $TSB_1(0)$ is precompact in W, yielding the claim.

Let now $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in Y. Then there exists a subsequence $(Tx_{n_k})_{k\in\mathbb{N}}$ that converges in W. By continuity of R, also $(RTx_{n_k})_{k\in\mathbb{N}}$ converges. This completes the proof. □

Exercise 6.15. *Let X, Y be Banach spaces and T be a linear operator from X to Y.*

(1) Show that if T is compact, then it is also bounded.
(2) Show that if T is bounded and X or Y are isomorphic to $\mathbb{K}^n$, then T is also compact.

Exercise 6.16. *Let X, Y be Banach spaces. Show that the set of all compact linear operators from X to Y is a vector space.*

The vector space of compact linear operators between Banach spaces X, Y is commonly denoted by $\mathcal{K}(X, Y)$.

Example 6.17. *Let X be a normed space. The zero operator on X is clearly compact. By Proposition 6.8, the identity on X is compact if and only if X is isomorphic to $\mathbb{R}^n$.*

Exercise 6.18. *Let $I \subset \mathbb{R}$ be a bounded interval. Show that the* **embedding** *– i.e., the identity considered as an operator from $H^1(I)$ into $C(\overline{I})$ (in the sense of Exercise 5.15) – is compact.*

Remark 6.19. *Observe that compactness of the embedding of $H^1(I)$ into $C(\overline{I})$ only means that $H^1(I)$ is a* precompact *subset of $C(\overline{I})$,* not *a compact subset. In fact, $H^1(I)$ contains the polynomials on I and hence by the theorem of Stone–Weierstraß it is dense (hence not closed) in $C(\overline{I})$.*

Exercise 6.20. *Consider the Fredholm operator K with kernel k introduced in Exercise 2.8.*

(1) Show that if $k \in L^2((0,1) \times (0,1))$, then K is a compact operator on $L^2(0,1)$.
(2) Show that if $k \in C([0,1] \times [0,1])$, then K is a compact operator from $C([0,1])$.

Definition 6.21. *Let X be normed spaces. An* **eigenvalue** *of a bounded linear operator T is a number $\lambda \in \mathbb{K}$ such that $\lambda \operatorname{Id} - T$ is not injective, i.e., such that there exists $u \in X$, $u \neq 0$, such that $Tu = \lambda u$. In this case u is called* **eigenvector associated with** λ *and the vector space* $\operatorname{Ker}(\lambda \operatorname{Id} - T)$ **eigenspace associated with** λ*. The set of all eigenvalues of T is called* **point**

spectrum of *T and is denoted by* $\mathrm{P}\sigma(T)$*. The set of all* $\lambda \in \mathbb{K}$ *such that* $\lambda \operatorname{Id} - T$ *is not bijective is called* **spectrum of** *T and denoted by* $\sigma(T)$*. If* $\lambda \notin \sigma(T)$*, then* $R(\lambda, T) := (\lambda \operatorname{Id} - T)^{-1}$ *is called* **resolvent operator of** T **at** λ*.*

The notions in Definition 6.21 were first defined (if not introduced) by David Hilbert in 1904 in his study of integral equations. They had such an impact on the newborn functional analysis that they were adopted in the english literature without even changing the prefix *eigen-*. Also the notion of spectrum has been introduced by Hilbert, in 1912.

In our introduction of spectral theory we follow [**16**, Kap. 2]. Throughout the following we assume that $\mathbb{K} = \mathbb{C}$.

Example 6.22. *The zero operator on a normed space* X *has* 0 *as its only eigenvalue, since* $\lambda \operatorname{Id}$ *is clearly injective (and even invertible) for all* $\lambda \neq 0$*, since so is the identity operator. Accordingly,* X *is the only eigenspace.*

Similarly, 1 *is the only eigenvalue of the identity operator with associated eigenspace* X*. More generally, the spectrum of each multiple of the identity* $\mu \operatorname{Id}$ *consists of* $\mu \in \mathbb{C}$ *only.*

Exercise 6.23. *Deduce from Proposition 6.14 that the resolvent operator of a bounded linear operator on a Banach space can never be compact unless the Banach space is isomorphic to* $\mathbb{K}^n$*.*

Exercise 6.24. *Let* X *be a Banach space and* T *be a bounded linear operator on* X*.*

- *Show that* $\lambda \notin \sigma(T)$ *if and only if* $1 \notin \sigma(\frac{T}{\lambda})$ *and in this case* $R(1, \frac{T}{\lambda}) = \lambda R(\lambda, T)$*.*
- *Let* S *be a further bounded linear operator on* X *and* $\lambda \notin \sigma(T)$*. Show that* $\lambda \notin \sigma(T+S)$ *if and only if* $1 \notin R(\lambda, T)S$*. (Hint: Write* $\lambda \operatorname{Id} - (T+S) = (\lambda - T)(\operatorname{Id} - R(\lambda, T)S)$*.)*
- *Let* X *be a Hilbert space. Show that the spectrum of the adjoint* T^* *is given by* $\{\overline{\lambda} \in \mathbb{C} : \lambda \in \sigma(T)\}$

Lemma 6.25. *Let* X *be a Banach space and* T *be a bounded linear operator on* X*. If* $\|T\| < 1$*, then* $1 \notin \sigma(T)$ *and*

$$R(1, T) = \sum_{n=0}^{\infty} T^n \qquad \text{with} \qquad \|R(1, T)\| \leq \frac{1}{1 - \|T\|}. \tag{6.2}$$

The series in (6.2) is called the **Neumann series**, as it has been introduced in 1877 by Carl Gottfried Neumann. More generally, it can be proved that if the Neumann series converges, then $\operatorname{Id} - T$ is invertible and its inverse is actually given by the series.

Proof. Due to submultiplicativity of the norm and by convergence of the geometric series (in $\mathbb{R}$), one sees that

$$\sum_{n=0}^{\infty} \|T^n\| \leq \sum_{n=0}^{\infty} \|T\|^n = \frac{1}{1 - \|T\|}. \tag{6.3}$$

Since X is complete, by Exercise 4.8 this implies convergence of $\sum_{n=0}^{m} T^n =: S_m$ towards $\sum_{n=0}^{\infty} T^n =: S \in \mathcal{L}(H)$ as $m \to \infty$. Hence,

$$S_m(\mathrm{Id} - T) = S_m - S_m T = S_m - T\sum_{n=0}^{m} T^n = S_m - \sum_{n=1}^{m+1} T^n = S_m - (S_{m+1} - \mathrm{Id}) = \mathrm{Id} - T^{m+1}$$

and accordingly

$$S(\mathrm{Id} - T) = \lim_{m\to\infty} S_m(\mathrm{Id} - T) = \lim_{m\to\infty} S_m - S_m T = \lim_{m\to\infty} S_m - (S_m - \mathrm{Id}) = \mathrm{Id} - \lim_{m\to\infty} T^{m+1} = \mathrm{Id},$$

since $\lim_{m\to\infty} \|T^{m+1}\| \leq \lim_{m\to\infty} \|T\|^{m+1} = 0$ because $\|T\| < 1$. One proves likewise that $(\mathrm{Id} - T)S = \mathrm{Id}$. We conclude that $R(1,T) = S = \sum_{n=0}^{\infty} T^n$, and the estimate in (6.2) follows from (6.3) and the triangle inequality. □

Lemma 6.26. *Let X be a Banach space. Then the following assertions hold.*

(1) The set $GL_X := \{T \in \mathcal{L}(X) : 0 \notin \sigma(T)\}$ of all invertible operators on X is open in the normed space $(\mathcal{L}(X), \|\cdot\|)$.
(2) The (nonlinear) inversion operator $T \mapsto T^{-1}$ is continuous on the set GL_X.
(3) The spectrum of T is contained in $B_{\|T\|}(0)$.
(4) The spectrum of a bounded operator is closed.

In the proof of this result we will need the notion of *Fréchet differentiability*, which lies at the basis of the theory of differentiation in Banach spaces. We refer to [**11**, § 175] for some basic definitions and results.

Proof. (1) Let $T \in GL_X$ and $S \in \mathcal{L}(X)$ such that $\|S - T\| < \|T^{-1}\|^{-1}$. Then it follows that

$$\|\mathrm{Id} - T^{-1}S\| = \|T^{-1}(T - S)\| \leq \|T^{-1}\|\|T - S\| < 1.$$

Accordingly, by Lemma 6.25 $T^{-1}S$, and therefore also S are invertible, i.e., $S \in GL_X$.

(2) We are going to prove even more, namely that the inversion operator is Fréchet differentiable at each $T \in GL_X$, i.e., that for all $T \in GL_X$ there exists a bounded linear operator D_T on the Banach space $\mathcal{L}(X)$ such that

$$\lim_{R\to 0} \frac{\|(T + R)^{-1} - T^{-1} - D_T R\|}{\|R\|} = 0.$$

Let again $T \in GL_X$. By (1), there is a neighbourhood of T in $\mathcal{L}(X)$ consisting of invertible operators. In fact, take $R \in \mathcal{L}(X)$ such that

$$\|R\| < \frac{1}{2}\|T^{-1}\|^{-1}, \tag{6.4}$$

i.e., $\|R\|\|T^{-1}\|\frac{1}{2}$. Then as in (1) $T - R = T(\mathrm{Id} - T^{-1}R)$ and therefore $T - R \in GL_X$ with

$$(T - R)^{-1} = R(1, T^{-1}R)T^{-1} = \sum_{n=0}^{\infty} (T^{-1}R)^n T^{-1} = T^{-1} + \sum_{n=1}^{\infty} (T^{-1}R)^n T^{-1}.$$

In order to check Fréchet differentiability it clearly suffices to show that the linear operator $\frac{1}{\|R\|}(T+R)^{-1} - T^{-1}$ is bounded by a constant not depending on R. This follows after observing that

$$\begin{aligned}
\|(T-R)^{-1} - T^{-1}\| &\leq \|\sum_{n=1}^{\infty}(T^{-1}R)^n T^{-1}\| \\
&\leq \sum_{n=1}^{\infty} \|T^{-1}\|^{n+1}\|R\|^n\| \\
&\leq \|T^{-1}\|^2\|R\| \sum_{n=1}^{\infty} \|T^{-1}\|^{n-1}\|R\|^{n-1} \\
&\leq \|T^{-1}\|^2\|R\| \sum_{n=0}^{\infty} \|T^{-1}\|^{n}\|R\|^{n} \\
&\leq \frac{\|T^{-1}\|^2\|R\|}{1-\|T^{-1}\|\|R\|} \\
&\leq \frac{1}{2}\|T^{-1}\|^2\|R\|
\end{aligned}$$

by (6.4) and due to convergence of the (scalar) geometric series.

(3) Take $\lambda \in \mathbb{C}$ with $|\lambda| > \|T\|$, hence in particular $\lambda \neq 0$. Then $\|\operatorname{Id} - (\operatorname{Id} - \lambda^{-1}T)\| < 1$, hence by Lemma 6.25 $\operatorname{Id} - \lambda^{-1}T \in GL_X$. Multiplying this by λ yields $\lambda \operatorname{Id} - T \in GL_X$ which by definition means that $\lambda \notin \sigma(T)$.

(4) Let $\lambda \notin \sigma(T)$, so that $\lambda \operatorname{Id} - T \in GL_X$. Since GL_X is open in $\mathcal{L}(X)$, there exists a neighourhood of $\lambda \operatorname{Id} - T$ contained in GL_X, i.e., there exists $\mu \in \mathbb{C}$ with $|\lambda - \mu|$ small enough that $\mu \operatorname{Id} - T \in GL_X$, and hence that $\mu \notin \sigma(T)$. Accordingly, $\sigma(T)$ is closed. □

Lemma 6.27. *Let X be a Banach space. Let S, T be bounded linear operators on X, $\lambda, \mu \in \mathbb{C}$. Then the following assertions hold.*

(1) $R(\lambda, T) - R(\mu, T) = (\mu - \lambda)R(\lambda, T)R(\mu, T)$ for all $\lambda, \mu \notin \sigma(T)$, and in particular all the resolvent operators of T commute.
(2) $R(\lambda, T) - R(\lambda, S) = R(\lambda, T)(T - S)R(\lambda, S)$ for all $\lambda \notin \sigma(T) \cup \sigma(S)$.
(3) $R(\cdot, T) : \mathbb{C} \setminus \sigma(T) \to \mathbb{C}$ is holomorphic and its derivative is given by

$$\frac{d}{d\lambda} R(\lambda, T) = -R(\lambda, T)^2, \qquad \lambda \notin \sigma(T).$$

Proof. (1) Since $(\mu \operatorname{Id} - T) - (\lambda \operatorname{Id} - T) = \mu \operatorname{Id} - \lambda \operatorname{Id}$, composing both sides with $R(\lambda, T)$ from the left and $R(\mu, T)$ from the right.

(2) Similarly, since $(\lambda \operatorname{Id} - S) - (\lambda \operatorname{Id} - T) = T - S$, the assertion follows composing both sides with $R(\lambda, S)$ from the left and $R(\lambda, T)$ from the right.

(3) Take $\lambda, \mu \notin \sigma(T)$ and deduce from (1) that

$$\frac{R(\lambda,T) - R(\mu,T)}{\lambda - \mu} = -R(\lambda,T)R(\mu,T).$$

Due to continuity of the inversion operator (cf. Lemma 6.26.(3)) It follows that

$$\lim_{\mu\to\lambda} \frac{R(\lambda,T) - R(\mu,T)}{\lambda - \mu} = -R(\lambda,T)^2,$$

hence $R(\cdot,T)$ is complex differentiable in λ and its derivative is $-R(\lambda,T)^2$. □

Theorem 6.28. *Let H be a (complex!) Hilbert space. Then each bounded linear operator T on H has nonempty spectrum and* $\lim_{|\lambda|\to\infty} \|R(\lambda,T)\| = 0$.

Proof. The proof goes in two steps.

(1) We prove an analogon of Liouville's theorem: Each H-valued function f that is bounded and holomorphic on all $\mathbb{C}$ is already constant.

(2) We deduce the claimed assertion.

(1) Consider functionals $(u|f(\cdot))_H : \mathbb{C} \to \mathbb{C}$, $u \in H$. By Exercise 3.25 also these functionals are holomorphic, and they are clearly bounded and holomorphic on all $\mathbb{C}$ by the Cauchy–Schwarz inequality. Hence by the well-known Liouville's theorem (scalar-valued case!) we deduce that $(u|f(\cdot))_H : \mathbb{C} \to \mathbb{C}$ is constant for all $u \in H$, say $(u|f(z))_H \equiv c_u$ for all $z \in \mathbb{C}$. It remains to prove that in fact $f(z_1) = f(z_2)$ for all $z_1, z_2 \in \mathbb{C}$. Take $z_1, z_2 \in \mathbb{C}$ and observe that $(u|f(z_1) - f(z_2))_H = 0$ for all $u \in H$, hence by Exercise 3.14.(6) $f(z_1) = f(z_2)$.

(2) Let finally $\lambda \in \mathbb{C}$ with $|\lambda| > \|T\|$, so that by Lemma 6.25 the bounded linear operator $\frac{1}{\lambda}T$ is invertible with inverse given by the Neumann series and $\|R(1, \frac{1}{\lambda}T)\| \leq \frac{1}{1-\|\frac{T}{\lambda}\|}$. Thus, by Exercise 6.24.(1) we deduce that

$$\|R(\lambda,T)\| = \left\|\frac{1}{\lambda}R\Big(1, \frac{T}{\lambda}\Big)\right\| \leq \frac{1}{|\lambda|(1 - \|\frac{T}{\lambda}\|)} = \frac{1}{|\lambda| - \|T\|},$$

whence

$$\lim_{|\lambda|\to\infty} \|R(\lambda,T)\| = 0. \tag{6.5}$$

Because of its holomorphy, the resolvent mapping $\lambda \mapsto R(\lambda,T)$ is bounded on each bounded subset of $\mathbb{C} \setminus \sigma(T)$. Assume now $\sigma(T)$ to be empty. Then by Lemma 6.27 $R(\cdot,T)$ is bounded and holomorphic on all $\mathbb{C}$, and by (1) it is constant. Because of (6.5), such a constant value is necessarily 0, i.e., $R(\lambda,T) \equiv 0$ for all $\lambda \in \mathbb{C}$. Since however the zero operator is not invertible, this is a contradiction and the proof is completed. □

Remark 6.29. *We have needed H to be a Hilbert (rather than merely a Banach) space only to apply the holomorphy of the resolvent operator, which we have characterized by the Hahn–Banach theorem in Hilbert spaces, cf. Exercise 3.25. However, as already remarked the Hahn–Banach theorem actually holds in Banach spaces, too, and in fact Theorem 6.28 is also valid for bounded linear operators on general Banach spaces.*

As we have seen, the proof of the previous fundamental result is deeply rooted in the application of ideas and methods from complex analysis. This explains why we always assume that $\mathbb{K} = \mathbb{C}$ whenever discussing spectral theory.

While the spectrum of a bounded linear operator is always nonempty, its point spectrum may indeed be empty.

Exercise 6.30. *Define the* **Volterra operator** V *by*

$$Vf := \int_0^{\cdot} f(s)ds, \qquad f \in C([0,1]).$$

Show that V is a compact linear operator on $C([0,1])$ with no eigenvalues. Such an operator was introduced by Vito Volterra – at the beginning of 20^{th} century he was studying integral equations similar to those that led to the introduction of Fredholm operators.

However, this is never the case if the operator is compact and self-adjoint on a Hilbert space.

Proposition 6.31. *Let H be a Hilbert space and T be a compact self-adjoint linear operator on H. Then at least one element of $\{-\|T\|, \|T\|\}$ is an eigenvalue of T.*

Proof. Assume without loss of generality that $T \neq 0$, the claim being otherwise obvious. Since by Corollary 5.25 $\|T\| = \sup_{\|x\|=1} |(Tx|x)_H|$, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\|x_n\|_H = 1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} |(Tx_n|x_n)_H| = \|T\|$. Upon considering a subsequence, we may assume that all real numbers $(Tx_n|x_n)_H$ have the same sign, i.e., $\lim_{n\to\infty}(Tx_n|x_n)_H = \lambda \neq 0$ with either $\lambda = \|T\|$ or $\lambda = -\|T\|$. Accordingly,

$$\begin{aligned} 0 &\leq \|(T - \lambda \operatorname{Id})x_n\|_H^2 \\ &= \|Tx_n\|_H^2 - 2\lambda(Tx_n|x_n)_H + \lambda^2 \\ &\leq \|T\|^2 - 2\lambda(Tx_n|x_n)_H + \lambda^2 \\ &= 2\|T\|^2 - 2\lambda(Tx_n|x_n)_H. \end{aligned}$$

Since

$$\lim_{n\to\infty} \lambda(Tx_n|x_n)_H = \lambda \lim_{n\to\infty} (Tx_n|x_n)_H = \lambda^2 = \|T\|^2,$$

it follows that $\lim_{n\to\infty}(T - \lambda \operatorname{Id})x_n = 0$. By compactness of T we may assume (upon considering a subsequence) that $(Tx_n)_{n\in\mathbb{N}}$ converges towards some $x \in H$. Clearly,

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{(\lambda \operatorname{Id} - T)x_n + Tx_n}{\lambda} = \lim_{n\to\infty} \frac{(\lambda \operatorname{Id} - T)x_n}{\lambda} + \lim_{n\to\infty} \frac{Tx_n}{\lambda} = \frac{x}{\lambda}.$$

Observe now that

$$x = \lim_{n\to\infty} Tx_n = T \lim_{n\to\infty} x_n = \frac{Tx}{\lambda},$$

i.e., $Tx = \lambda x$. It remains to prove that $x \neq 0$, and this can be done passing to the norm and observing that $\|x\|_H = |\lambda| \|x_n\| = \|T\| \neq 0$, $n \in \mathbb{N}$. □

Proposition 6.32. *Let H be a separable Hilbert space and T be a compact linear operator on H. If $\lambda \neq 0$ is an eigenvalue of T, then the associated eigenspace is finite dimensional.*

Proof. Consider a basis $\{e_n : n \geq 1\}$ of the null space of $\lambda \operatorname{Id} - T$, $\lambda \neq 0$. This is surely possible by the Gram–Schmidt orthonormalisation process, because such a null space is a closed subspace of H and hence itself a separable Hilbert space. Assume the eigenspace associated with λ to be infinite dimensional, i.e., assume the above basis to have the cardinality of $\mathbb{N}$, i.e., it is $\{e_n : n \in \mathbb{N}\}$. Since this sequence is bounded, $\{Te_n : n \in \mathbb{N}\}$ is a precompact set and it admits a convergent subsequence $\{Te_{n_k} : k \in \mathbb{N}\}$.

However, this contradicts the fact that for indices $k \neq h$ Parseval's identity implies

$$\|Te_{n_k} - Te_{n_h}\|^2 = \|\lambda e_{n_k} - \lambda e_{n_h}\|^2 = 2|\lambda|^2 > 0,$$

i.e., $\{Te_{n_k} : k \in \mathbb{N}\}$ is not a Cauchy sequence. □

Exercise 6.33. *Let $\Omega \subset \mathbb{R}^n$ and $q \in L^\infty(\Omega)$. Consider the multiplication operator introduced in Exercise 2.25. Show that the spectrum of M_q agrees with the essential range of q.*

The following **Spectral theorem** is one of the fundamental results obtained by David Hilbert. It was published in 1906, in a generalization of his previous investigations dating back to 1904.

Theorem 6.34. *Let H be a Hilbert space and T be a linear compact self-adjoint operator on H. Then the following assertions hold.*

(1) The point spectrum of T is (at most) a countable set $\{\lambda_n : n \geq 1\}$ in $\mathbb{C}$ whose only possible accumulation point is 0.

(2) Denoting by P_n the orthogonal projection onto the (closed) eigenspace associated with λ_n, there holds $T = \sum_{n\geq 1} \lambda_n P_n$.

(3) Finally, H has an basis $\{u_n : n \geq 1\}$ consisting of eigenvectors of T and such that

$$Tx = \sum_{n\geq 1} \lambda_n (x|u_n)_H u_n \qquad \text{for all } x \in H.,$$

provided that H is separable.

The Spectral theorem can be extended to the case of a bounded linear operator, although both the formulation and the proof become much more involved, see e.g. [**20**, Satz 3.3.3].

Proof. (1) If $T = 0$ the assertion is trivial. Let therefore $T_1 := T \neq 0$ and observe that by Proposition 6.31 we can pick an eigenvalue λ_1 of T whose absolute value agrees with $\|T_1\|$. Clearly, T_1 leaves invariant both the eigenspace associated with λ_1 and its orthogonal complement, i.e., $T\operatorname{Ker}(\lambda_1 \operatorname{Id} - T) \subset \operatorname{Ker}(\lambda_1 \operatorname{Id} - T)$ and $T\operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp \subset \operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp$. If $\operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp = \{0\}$, i.e., if $\operatorname{Ker}(\lambda_1 \operatorname{Id} - T) = H$, then the eigenvalue λ_1 is the only element of $\sigma(T)$ and the assertion holds.

If however $\operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp \neq \{0\}$, we consider the restriction T_2 of T to $\operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp$ as a compact linear operator on the closed subspace (with respect to the induced inner product) and hence Hilbert space in its own right $\operatorname{Ker}(\lambda_1 \operatorname{Id} - T)^\perp$. Moreover, T_2 is self-adjoint (why?), and again we consider the two cases $T_2 = 0$ (and then the assertion follows) or $T_2 \neq 0$, in which case we pick an eigenvalue λ_2 of T_2 whose absolute value agrees with $\|T_2\|$. We observe that

$$\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2) = \operatorname{Ker}(\lambda_2 \operatorname{Id} - T_1) = \operatorname{Ker}(\lambda_2 \operatorname{Id} - T),$$

since $\{0\} \neq \operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2) \subset \operatorname{Ker}(\lambda_2 \operatorname{Id} - T_1)$ with $\lambda_1 \neq \lambda_2$. It follows that $\lambda_1 \neq \lambda_2$, because otherwise we would have additionally that $\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2) \subset \operatorname{Ker}(\lambda_1 \operatorname{Id} - T_1)$, which contradicts the fact that $\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2)$ and $\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_1)$ are orthogonal. Observe that $|\lambda_2| \leq |\lambda_1|$, since by construction $\|T_2\| \leq \|T_1\|$.

Again, consider the direct sum $\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2) \oplus \operatorname{Ker}(\lambda_1 \operatorname{Id} - T_1)$, its orthogonal subspace

$$(\operatorname{Ker}(\lambda_2 \operatorname{Id} - T_2) \oplus \operatorname{Ker}(\lambda_1 \operatorname{Id} - T_1))^\perp$$

and the restriction T_3 of T_2 to it, which as above is compact and such that $\|T_3\| \leq \|T_2\|$, linear and self-adjoint operator, so that we can consider an eigenvalue λ_3 of T_3 (and in fact of T) with $|\lambda_3| = \|T_3\|$ and the associated eigenspace $\operatorname{Ker}(\lambda_3 \operatorname{Id} - T)$, and so on...

In this way we can inductively define a sequence $\{\lambda_n : n \geq 1\}$ of eigenvalues with decreasing absolute value – this defining process possibly stops whenever

$$\bigoplus_{k=1}^{n} \operatorname{Ker}(\lambda_k \operatorname{Id} - T) = H \qquad \text{or} \qquad T_n = 0$$

for some n.[1] If however this process does not stop at any finite step n, then we can consider a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in \operatorname{Ker}(\lambda_n \operatorname{Id} - T)$ and $\|x_n\| = 1$. By compactness of T, the sequence $(Tx_n)_{n\in\mathbb{N}}$ admits a convergent subsequence $(Tx_{n_k})_{k\in\mathbb{N}}$. Observe however that since each x_n is an eigenvector associated with the eigenvalue λ_n (i.e., $\lim_{k\to\infty} Tx_{n_k} = \lim_{k\to\infty} \lambda_{n_k} x_{n_k}$), we have

$$\|Tx_{n_k} - Tx_{n_h}\|^2 = \|\lambda_{n_k} x_{n_k} - \lambda_{n_h} x_{n_h}\|^2 = \|\lambda_{n_k} x_{n_k}\|^2 + \|\lambda_{n_h} x_{n_h}\|^2 = |\lambda_{n_k}|^2 + |\lambda_{n_h}|^2$$

by Exercise 3.14.(1) since the eigenspaces $\operatorname{Ker}(\lambda_{n_k} \operatorname{Id} - T)$ are by construction pairwise orthogonal, and therefore $\lim_{h,k\to\infty} |\lambda_{n_k}|^2 + |\lambda_{n_h}|^2 = 0$. In other words, the unique possible accumulation point of $(\lambda_n)_{n\in\mathbb{N}}$ is 0.

(2) In order to prove the claimed representation of T, take a finite n, pick $x \in \operatorname{Ker}(\lambda_k \operatorname{Id} - T)$ for some $k \leq n$ and observe that

$$\Big(T - \sum_{k=1}^{n} \lambda_n P_n\Big)x = Tx - \lambda_k x = 0,$$

where the equalities follow from the facts that the eigenspaces are pairwise orthogonal and that x is an eigenvector of T, respectively. This means that each of the (pairwise orthogonal) subspaces $\operatorname{Ker}(\lambda_k \operatorname{Id} - T)$ is contained in the null space of the operator $\Big(T - \sum_{n\geq 1} \lambda_n P_n\Big)$, $k \leq n$, and hence so is their direct sum. On the other hand, if we pick

$$x \in \left(\bigoplus_{k=1}^{n} \operatorname{Ker}(\lambda_k \operatorname{Id} - T)\right)^\perp$$

[1] Observe in particular that by Exercise 3.17 we deduce

$$P_n P_m = P_m P_n = \begin{cases} P_n & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases}$$

we obtain by construction

$$\Big(T - \sum_{k=1}^{n} \lambda_n P_n\Big) x = Tx,$$

and accordingly

$$\lim_{n\to\infty} \Big\| T - \sum_{k=1}^{n} \lambda_n P_n \Big\| = \lim_{n\to\infty} |\lambda_{n+1}|.$$

The claim follows because by (1) the unique possible accumulation point of the decreasing sequence

$$\left(\|T - \sum_{k=1}^{n} \lambda_n P_n\| \right)_{n\in\mathbb{N}}$$

is 0.

(3) Finally, consider bases $\{u_j : j \in J_0\}$ of $\operatorname{Ker}(T)$ and $\{u_j : j \in J_n\}$ of $\operatorname{Ker}(\lambda_n \operatorname{Id} - T)$, $n \geq 1$. This can be done by separability of H. Consider their union $\bigcup_{n\geq 0}\{u_j : j \in J_n\}$ and observe that this defines a basis of H, since $\operatorname{Ker}(T)$ is orthogonal to the range of T, i.e., to

$$\bigoplus_{n\geq 1} \operatorname{Ker}(\lambda_n \operatorname{Id} - T),$$

and H is direct sum of $\operatorname{Ker}(T)$ and $\overline{\operatorname{Ran}(T)}$, because $\|Tx\|^2 = \sum_{n\geq 1} |\lambda_n|^2 \|P_n x\|^2$ for all $x \in H$ by Exercise 3.14.(1). □

Theorem 6.35. *Let H be a Hilbert space and T be a linear compact operator on H. Then the following assertions hold.*

1. *Let $\lambda \in \mathbb{C}$. If $\lambda \operatorname{Id} - T$ is injective, then $\lambda \operatorname{Id} - T$ is surjective.*
2. *With the possible exception of 0, each element of $\sigma(T)$ is also an element of $P\sigma(T)$ – i.e., an eigenvalue.*

Proof. (1) Assume $\lambda \operatorname{Id} - T$ to be injective but its range $H_1 := (\lambda \operatorname{Id} - T)H$ to be different from H. Observe that T maps H_1 into itself. Since H_1 is closed in H and hence a Hilbert space (why?) and since the restriction of T to H_1 is a compact operator on H_1 (why?), we deduce similarly that $H_2 := (\lambda \operatorname{Id} - T)H_1 \subset H_1$ is a Hilbert space. Moreover, $H_1 \neq H_2$ because T is injective. In this way, we can define recursively a sequence of Hilbert spaces $(H_n)_{n\in\mathbb{N}}$, with H_n strictly included in H_m whenever $n > m$. By Lemma 6.7 there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in H_n$, $\|x_n\| = 1$, and $\inf_{y\in H_{n+1}} \|y - x_n\|_H \geq \frac{1}{2}$ for all $n \in \mathbb{N}$. Observe that for all $n, m \in \mathbb{N}$ one has

$$Tx_n - Tx_m = -(\lambda x_n - Tx_n) + (\lambda x_m - Tx_m) + (\lambda x_n - \lambda x_m),$$

and accordingly for all $n > m$

$$-(\lambda x_n - Tx_n) + (\lambda x_m - Tx_m) + \lambda x_n \in H_{m+1}.$$

Accordingly, $\|Tx_n - Tx_m\|_H \geq \frac{1}{2}$ for all $n > m$ and hence $(Tx_n)_{n\in\mathbb{N}}$ has no convergent subsequence, in spite of the fact that $(x_n)_{n\in\mathbb{N}}$ is bounded – a contradiction to compactness of T.

(2) We can assume that H is not isomorphic to $\mathbb{K}^n$ for any $n \in \mathbb{N}$, since otherwise the assertion is trivial. Let therefore $\lambda \in \sigma(T)$ and assume λ not to be an eigenvalue, i.e., $\lambda \operatorname{Id} - T$ to be injective. Then by (1) $\lambda \operatorname{Id} - T$ is surjective and therefore $\lambda \notin \sigma(T)$, a contradiction. □

Exercise 6.36. *Let H be a Hilbert space and T be a bounded linear operator on H. Show that T is compact if and only if T^* is compact. Combine this with Exercise 5.32.(4) and deduce that the converse of Theorem 6.35.(1) also holds.*

Exercise 6.37. *Show that the Fredholm operator from Exercise 4.37 is compact. Deduce the validity of the so-called* **Fredholm alternative***: Given $0 \neq \lambda \in \mathbb{C}$, either there exists $0 \neq f \in L^2(0,1)$ such that*

$$\lambda f(x) = \int_0^1 k(x,y) f(y) dy \qquad \text{for a.e. } x \in (0,1),$$

or for all $g \in L^2(0,1)$ there exists a unique $f \in L^2(0,1)$ such that

$$\lambda f(x) - \int_0^1 k(x,y) f(y) dy = g(x) \qquad \text{for a.e. } x \in (0,1).$$

Exercise 6.38. *Let H be a separable Hilbert space with a basis $\{e_n : n \in \mathbb{N}\}$. A linear compact self-adjoint operator T on H is called* **positive definite** *if $(Te_n|e_n)_H \geq 0$ for all $n \in \mathbb{N}$. Show that T is positive definite if and only if all eigenvalues of T are positive real numbers.*

Remark 6.39. *Hilbert's original formulation of the Spectral theorem was very technical. It was Paul Richard Halmos who showed in 1974 in* [**8**] *how the theorem can be reformulated in a much simpler way, bringing to light its close relation to the result on diagonalisability of symmetric matrices. In his formulation, the Spectral theorem says that for each linear compact self-adjoint operator T on a Hilbert space H there is a measure space (X, Σ, μ) and an essentially bounded measurable function $q : X \to \mathbb{R}$ such that A is unitarily equivalent to the multiplication operator M_q, i.e., $U^* M_q U = T$ for some unitary*[2] *operator U from H to $L^2(X, \mu)$.*

[2] A linear operator on a Hilbert space is called **unitary** if it is invertible and its inverse coincides with its adjoint.

Index

Bibliography

[1] Jürgen Appell, Martin Väth, *Elemente der Funktionalanalysis*, Vieweg, Wiesbaden, 2009.

[2] Haïm Brezis, *Analyse Fonctionelle*, Masson, Paris, 1984.

[3] John B. Conway, *A Course in Functional Analysis*, Springer-Verlag, New York, 1985.

[4] Lokenath Debnath, Piotr Mikusiński, *Introduction to Hilbert spaces with applications*, Academic Press, San Diego, 2005.

[5] David Forster Wallace, *A Compact History of ∞ – Everything and More*, Norton, New York, 2003.

[6] Alfred Göpfert, Thomas Riedrich, Christiane Tammer, *Angewandte Funktionalanalysis*, Vieweg+Teubner, Wiesbaden, 2009.

[7] Karl E. Gustafson, Duggirala K. M. Rao, *Numerical Range*, Springer-Verlag, New York, 1996.

[8] Paul R. Halmos, *What does the spectral theorem say?*, Am. Math. Mon. **70** (1974), 241–247.

[9] Paul R. Halmos, *Finite-Dimensional Vector Spaces*, Springer-Verlag, New York, 1974.

[10] Paul R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York, 1982.

[11] Harro Heuser, *Lehrbuch der Analysis – Teil 2*, Teubner, Stuttgart, 2004.

[12] Harro Heuser, *Funktionalanalysis: Theorie und Anwendung*, Teubner, Stuttgart, 2006.

[13] Erwin Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, New York, 1978.

[14] Serge Lang, *Real and Functional Analysis*, Springer-Verlag, New York, 1993.

[15] Peter D. Lax, *Functional Analysis*, Wiley, Hoboken, NJ, 2002.

[16] Martin Mathieu, *Funktionalanalysis*, Spektrum, Heidelberg, 1999.

[17] John von Neumann, *Functional Operators. Vol. II. The Geometry of Orthogonal Spaces*, Princeton University Press, Princeton, 1950.

[18] Walter Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

[19] Karen Saxe, *Beginning Functional Analysis*, Springer-Verlag, New York, 2002.

[20] Herbert Schröder, *Funktionalanalysis*, Akademie Verlag, Berlin, 1997.

[21] Ralph E. Showalter, Hilbert Space Methods for Partial Differential Equations, Pitman, London, 1977. Freely available at `http://www.emis.ams.org/journals/EJDE/Monographs/01/toc.html`

[22] Dirk Werner, *Funktionalanalysis*, Springer-Verlag, Berlin, 2006.