Chapter 1

Inner Product Spaces

The main objects of study in functional analysis are *function spaces*, i.e., vector spaces of real or complex-valued functions on certain sets. Although much of the theory can be done in the context of real vector spaces, at certain points it is very convenient to have vector spaces over the complex number field $\mathbb{C}$. So we introduce the generic notation $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$. Background on linear algebra is collected in Appendix A.7.

We begin by introducing two central examples of function spaces.

The Space $\mathbb{K}^d$. This is the set of all tuples $x = (x_1, \dots, x_d)$ with components $x_1, \dots, x_d \in \mathbb{K}$:

$$\mathbb{K}^d := \{x = (x_1, \dots, x_d) \mid x_1, \dots, x_d \in \mathbb{K}\}.$$

It is a vector space over $\mathbb{K}$ with the obvious (i.e., componentwise) operations:

$$(x_1, \dots, x_d) + (y_1, \dots, y_d) := (x_1 + y_1, \dots, x_d + y_d),$$
$$\lambda(x_1, \dots, x_d) := (\lambda x_1, \dots, \lambda x_d).$$

The Space $\mathrm{C}[a, b]$. We let $[a, b]$ be any closed interval of $\mathbb{R}$ of positive length. Let us define

$$\mathcal{F}[a, b] := \{f \mid f : [a, b] \longrightarrow \mathbb{K}\},$$
$$\mathrm{C}[a, b] := \{f \mid f : [a, b] \longrightarrow \mathbb{K},\ \text{continuous}\}.$$

If $\mathbb{K} = \mathbb{C}$, then $f : [a, b] \to \mathbb{C}$ can be written as $f = \operatorname{Re} f + \mathrm{i} \operatorname{Im} f$ with $\operatorname{Re} f$, $\operatorname{Im} f$ being real-valued functions; and f is continuous if and only if both $\operatorname{Re} f$, $\operatorname{Im} f$ are continuous.

Let us define the sum and the scalar multiple of functions pointwise, i.e.,

$$(f + g)(t) := f(t) + g(t), \quad (\lambda f)(t) := \lambda f(t)$$

where $f, g : [a, b] \to \mathbb{K}$ are functions, $\lambda \in \mathbb{K}$ and $t \in [a, b]$. This turns the set $\mathcal{F}[a, b]$ into a vector space over $\mathbb{K}$ (see also Appendix A.7).

If $f, g \in \mathrm{C}[a, b]$ and $\lambda \in \mathbb{K}$, then we know from elementary analysis that $f + g, \lambda f \in \mathrm{C}[a, b]$ again. Since $\mathrm{C}[a, b]$ is certainly not the empty set, it is therefore a subspace of $\mathcal{F}[a, b]$, and hence a vector space in its own right.

We use the notation $\mathrm{C}[a, b]$ for the generic case, leaving open whether $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$. If we want to stress a particular choice of $\mathbb{K}$, we write $\mathrm{C}([a, b]; \mathbb{C})$ or $\mathrm{C}([a, b]; \mathbb{R})$. If we use the notation $\mathrm{C}[a, b]$ in concrete situations, it is always tacitly assumed that we have the more general case $\mathbb{K} = \mathbb{C}$. Similar remarks apply to $\mathcal{F}[a, b]$ and all other function spaces.

There is an analogy between these two examples. Namely, note that each vector $(x_1, \dots, x_d) \in \mathbb{K}^d$ defines a map

$$x : \{1, \dots, d\} \longrightarrow \mathbb{K} \quad \text{by} \quad x(j) := x_j \quad (j = 1, \dots, d).$$

Conversely, each such map x determines exactly one vector $(x(1), \dots, x(d))$. Apart from a set-theoretical point of view, there is no difference between the vector and the corresponding function, and we will henceforth identify them. So we may write

$$\mathbb{K}^d = \mathcal{F}(\{1, \dots, d\}; \mathbb{K}).$$

A short look will convince you that the addition and scalar multiplication in vector notation coincides precisely with the pointwise sum and scalar multiplication of functions.

How far can we push the analogy between $\mathbb{K}^d$ and $\mathrm{C}[a, b]$? Well, the first result is negative:

Theorem 1.1. *The space $\mathbb{K}^d$ has a basis consisting of precisely d vectors, hence is finite-dimensional. The space $\mathrm{C}[a, b]$ is not finite-dimensional. For example, the set of monomials $\{1, t, t^2, \dots\}$ is an infinite linearly independent subset of $\mathrm{C}[a, b]$.*

Proof. The first assertion is known from linear algebra. Let us turn to the second. Let

$$p(t) := a_n t^n + \cdots + a_1 t + a_0$$

be a finite linear combination of monomials, i.e., $a_0, \dots, a_n \in \mathbb{K}$. We suppose that not all coefficients a_j are zero, and we have to show that then p cannot be the zero function.

Now, if $p(c) = 0$, then by long division we can find a polynomial q such that $p(t) = (t - c)q(t)$ and $\deg q < \deg p$. If one applies this repeatedly, one may write

$$p(t) = (t - c_1)\,(t - c_2)\ldots(t - c_k)\,q(t)$$

for some $k \le n$ and some polynomial q that has no zeroes in $[a, b]$. But that means that p can have only finitely many zeroes in $[a, b]$. Since the interval $[a, b]$ has infinitely many points, we are done. (See Exercise 1.1 for an alternative proof.) □ Ex.1.1

1.1. Inner Products

We now come to a positive result. The **standard inner product** of two vectors $x, y \in \mathbb{K}^d$ is defined by

$$\langle x, y\rangle := x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_d\overline{y_d} = \sum\nolimits_{j=1}^{d} x_j\overline{y_j}.$$

If $\mathbb{K} = \mathbb{R}$, this is the usual scalar product known from undergraduate courses; for $\mathbb{K} = \mathbb{C}$ this is a natural extension of it.

Analogously, we define the **standard inner product** on $\mathrm{C}[a, b]$ by

$$\langle f, g\rangle := \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$$

for $f, g \in \mathrm{C}[a, b]$. There is a general notion behind these examples.

Definition 1.2. Let E be a vector space. A mapping

$$E \times E \longrightarrow \mathbb{K}, \quad (f, g) \longmapsto \langle f, g\rangle$$

is called an **inner product** or a **scalar product** if it is *sesquilinear*:

$$\begin{aligned}&\langle \lambda f + \mu g, h\rangle = \lambda\,\langle f, h\rangle + \mu\,\langle g, h\rangle\,,\\ &\langle h, \lambda f + \mu g\rangle = \overline{\lambda}\,\langle h, f\rangle + \overline{\mu}\,\langle h, g\rangle \qquad (f, g, h \in E,\ \lambda, \mu \in \mathbb{K}),\end{aligned}$$

symmetric:

$$\langle f, g\rangle = \overline{\langle g, f\rangle} \qquad (f, g \in E),$$

positive:

$$\langle f, f\rangle \ge 0 \qquad (f \in E),$$

and *definite*:

$$\langle f, f\rangle = 0 \quad \Longrightarrow \quad f = 0 \qquad (f \in E).$$

A vector space E together with an inner product on it is called an **inner product space** or a **pre-Hilbert space**.[1]

[1] David Hilbert (1862–1943), German mathematician.

There are different symbols used to denote inner products; for example,

$$\langle f, g \rangle, \quad (f \mid g), \quad \langle f | g \rangle \quad \text{or simply} \quad (f, g).$$

The latter has the disadvantage that it is the same as for the *ordered pair* (f, g). To avoid confusion, we stick to the notation $\langle f, g \rangle$ in this book.

The proof that the standard inner product on $\mathrm{C}[a, b]$ is sesquilinear,
Ex.1.2 symmetric and positive is an exercise. The definiteness is more interesting and derives from the following fact.

Lemma 1.3. *Let $f \in \mathrm{C}[a, b]$, $f \geq 0$. If $\int_a^b f(t)\,\mathrm{d}t = 0$, then $f = 0$.*

Proof. To prove the statement, suppose towards a contradiction that $f \neq 0$. Then there is $t_0 \in (a, b)$ where $f(t_0) \neq 0$, i.e., $f(t_0) > 0$. By continuity, there are $\epsilon, \delta > 0$ such that

$$|t - t_0| \leq \delta \quad \Rightarrow \quad f(t) \geq \epsilon.$$

But then

$$\int_a^b f(t)\,\mathrm{d}t \geq \int_{t_0-\delta}^{t_0+\delta} f(t)\,\mathrm{d}t \geq 2\delta\epsilon > 0,$$

which contradicts the hypothesis. □

Using this lemma we prove definiteness as follows: Suppose that $f \in \mathrm{C}[a, b]$ is such that $\langle f, f \rangle = 0$. Then

$$0 = \langle f, f \rangle = \int_a^b f(t)\overline{f(t)}\,\mathrm{d}t = \int_a^b |f(t)|^2\,\mathrm{d}t.$$

Since $|f|^2$ is also a continuous function, the previous lemma applies and yields $|f|^2 = 0$, but this is equivalent to $f = 0$.

Note that by "$f = 0$" we actually mean "$f(t) = 0$ for all $t \in [a, b]$", and we use $|f|$ as an abbreviation of the *function* $t \mapsto |f(t)|$.

Inner products endow a vector space with a geometric structure that allows one to measure lengths and angles. If $(E, \langle \cdot, \cdot \rangle)$ is an inner product space, then the **length** of $f \in E$ is given by

$$\|f\| := \sqrt{\langle f, f \rangle}.$$

The following properties are straightforward from the definition:

$$\|f\| \geq 0, \quad \|\lambda f\| = |\lambda|\,\|f\|, \quad \|f\| = 0 \iff f = 0.$$

The mapping $\|\cdot\| : E \to \mathbb{R}$ is called the **norm induced by** or **associated with** the inner product $\langle \cdot, \cdot \rangle$. We will learn more about norms in the next chapter. Ex.1.3

Example 1.4. For the standard inner product on $\mathbb{K}^d$, the associated norm is

$$\|x\|_2 := \sqrt{\langle x, x\rangle} = \Big(\sum\nolimits_{j=1}^{d} x_j \overline{x_j} \Big)^{1/2} = \Big(\sum\nolimits_{j=1}^{d} |x_j|^2 \Big)^{1/2}$$

and is called the **2-norm** or **Euclidean**[2] **norm** on $\mathbb{K}^d$.

For the standard inner product on $\mathrm{C}[a, b]$ the associated norm is given by

$$\|f\|_2 := \sqrt{\langle f, f\rangle} = \Big(\int_a^b f(t)\overline{f(t)}\,\mathrm{d}t \Big)^{1/2} = \Big(\int_a^b |f(t)|^2\,\mathrm{d}t \Big)^{1/2}$$

and is also called the **2-norm**. Ex.1.4

Let us turn to some "geometric properties" of the norm.

Lemma 1.5. *Let $(E, \langle \cdot, \cdot \rangle)$ be an inner product space. Then the following identities hold for all $f, g \in E$:*

a) $\|f + g\|^2 = \|f\|^2 + 2\,\mathrm{Re}\,\langle f, g\rangle + \|g\|^2$,

b) $\|f + g\|^2 - \|f - g\|^2 = 4\,\mathrm{Re}\,\langle f, g\rangle$ **(polarization identity)**, Ex.1.5

c) $\|f + g\|^2 + \|f - g\|^2 = 2\,\|f\|^2 + 2\,\|g\|^2$ **(parallelogram law).** Ex.1.6

Proof. The sesquilinearity and symmetry of the inner product yields

$$\begin{aligned}\|f + g\|^2 &= \langle f + g, f + g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle \\ &= \|f\|^2 + \langle f, g\rangle + \overline{\langle f, g\rangle} + \|g\|^2 = \|f\|^2 + 2\,\mathrm{Re}\,\langle f, g\rangle + \|g\|^2,\end{aligned}$$

since $z + \overline{z} = 2\,\mathrm{Re}\,z$ for every complex number $z \in \mathbb{C}$; this is a). Replacing g by $-g$ yields

$$\|f - g\|^2 = \|f\|^2 - 2\,\mathrm{Re}\,\langle f, g\rangle + \|g\|^2$$

and adding this to a) yields c). Subtracting it leads to b). □

If $\mathbb{K} = \mathbb{R}$, the polarization identity reads

$$\langle f, g\rangle = \frac{1}{4}\left(\|f + g\|^2 - \|f - g\|^2 \right) \qquad (f, g \in E).$$

[2] Euclid (around 280 BC), Greek mathematician.

Consequently, the inner product is completely determined by its associated norm. Geometrically, this means that lengths determine angles. More precisely: if $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ are two inner products on a real vector space E with associated norms $\|\cdot\|_1$, $\|\cdot\|_2$ satisfying

$$\|f\|_1 = \|f\|_2 \qquad \text{for all } f \in E,$$

then $\langle f, g\rangle_1 = \langle f, g\rangle_2$ for all $f, g \in E$.

The same statement is true in the case $\mathbb{K} = \mathbb{C}$; to prove this, one can use the extended polarization identity of Exercise 1.5.

1.2. Orthogonality

As in the case of geometry in three-dimensional Euclidean space, one can use the inner product to define angles between vectors. However, in this book we shall need only right angles, so we confine ourselves to these.

Definition 1.6. Let $(E, \langle\cdot,\cdot\rangle)$ be an inner product space. Two elements $f, g \in E$ are called **orthogonal**, written $f \perp g$, if $\langle f, g\rangle = 0$. For a subset $S \subseteq E$ we let

$$S^\perp := \{f \in E \mid f \perp g \text{ for every } g \in S\}$$

and we also write $f \perp S$ in place of $f \in S^\perp$.

By the symmetry of the inner product we have $f \perp g \iff g \perp f$. The definiteness of the inner product translates into the useful fact

$$f \perp E \quad\Longleftrightarrow\quad f = 0,$$

or, in short, $E^\perp = \{0\}$.

Example 1.7. In the (standard) inner product space $\mathrm{C}[a,b]$ we denote by $\mathbf{1}$ the function which is constantly equal to 1, i.e., $\mathbf{1}(t) := 1$, $t \in [a,b]$. Then for $f \in \mathrm{C}[a,b]$ one has

$$\langle f, \mathbf{1}\rangle = \int_a^b f(t)\overline{\mathbf{1}(t)}\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

Hence $\{\mathbf{1}\}^\perp = \{f \in \mathrm{C}[a,b] \mid \int_a^b f(t)\,\mathrm{d}t = 0\}$.

Let us note a useful lemma.

Lemma 1.8. *Let $(E, \langle\cdot,\cdot\rangle)$ be an inner product space, and let $S \subseteq E$ be any subset. Then $S^\perp$ is a linear subspace of E.*

Proof. Clearly $0 \in S^\perp$. If $f, g \in S^\perp$, $\lambda \in \mathbb{K}$, then

$$\langle \lambda f + g, s\rangle = \lambda \langle f, s\rangle + \langle g, s\rangle = \lambda \cdot 0 + 0 = 0$$

for arbitrary $s \in S$. Consequently, $\lambda f + g \in S^\perp$, and this had to be shown. □

The following should seem familiar from elementary geometry.

Lemma 1.9 (Pythagoras[3])**.** *Let $(E, \langle\cdot,\cdot\rangle)$ be an inner product space with associated norm $\|\cdot\|$. Let $f_1, \dots, f_d \in E$ be pairwise orthogonal, i.e., $f_i \perp f_j$ whenever $i \neq j$. Then*

$$\|f_1 + \cdots + f_d\|^2 = \|f_1\|^2 + \cdots + \|f_d\|^2 .$$

Proof. For $d = 2$ this follows from Lemma 1.5.a); the rest is induction. □ Ex.1.7

Let I be any nonempty index set. A collection of vectors $(e_j)_{j\in I}$ in an inner product space E is called an **orthonormal system** if

$$\langle e_i, e_j\rangle = \delta_{ij} := \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases}$$

Given an orthonormal system $(e_j)_{j\in I}$ in an inner product space E and a vector $f \in E$, we call the scalar

$$\langle f, e_j\rangle$$

the j-th **abstract Fourier coefficient**[4] and the formal(!) series

$$\sum\nolimits_{j\in I} \langle f, e_j\rangle\, e_j$$

the **abstract Fourier series** of f with respect to $(e_j)_{j\in I}$. We shall keep things simple and confine our study to *finite* orthonormal systems for the moment. It is one of the major achievements of functional analysis to make sense of such expressions when I is not finite; cf. Section 8.4 and Appendix F.

Lemma 1.10. *Let $(E, \langle\cdot,\cdot\rangle)$ be an inner product space with associated norm $\|\cdot\|$, and let $e_1, \dots, e_n \in E$ be a finite orthonormal system.*

a) *Let $g = \sum_{j=1}^n \lambda_j e_j$ (with $\lambda_1, \dots, \lambda_n \in \mathbb{K}$) be any linear combination of the e_j. Then*

$$\langle g, e_k\rangle = \sum\nolimits_{j=1}^n \lambda_j \langle e_j, e_k\rangle = \lambda_k \qquad (k = 1, \dots, n)$$

and

$$\|g\|^2 = \sum\nolimits_{j=1}^n |\lambda_j|^2 = \sum\nolimits_{j=1}^n |\langle g, e_j\rangle|^2 .$$

[3] Pythagoras (570–510(?) BC), Greek mathematician and religious leader.
[4] Joseph Fourier (1768–1830), French mathematician.

b) *For* $f \in E$ *let* $Pf := \sum_{j=1}^{n} \langle f, e_j \rangle e_j$. *Then*

$$f - Pf \perp \operatorname{span}\{e_1, \dots, e_n\} \quad \textit{and} \quad \|Pf\| \le \|f\|.$$

Proof. a) is just sesquilinearity and Pythagoras' lemma. For the proof of b) note that by a) we have $\langle Pf, e_j \rangle = \langle f, e_j \rangle$, i.e.,

$$\langle f - Pf, e_j \rangle = \langle f, e_j \rangle - \langle Pf, e_j \rangle = 0 \quad \text{for all } j = 1, \dots, n.$$

By Lemma 1.8 it follows that $f - Pf \perp \operatorname{span}\{e_j \mid j = 1, \dots, n\} =: F$. In particular, since $Pf \in F$ we have $f - Pf \perp Pf$ and

$$\|f\|^2 = \|(f - Pf) + Pf\|^2 = \|f - Pf\|^2 + \|Pf\|^2 \ge \|Pf\|^2$$

by Pythagoras' lemma. □

Let us abbreviate $F := \operatorname{span}\{e_1, \dots, e_n\}$. The mapping

$$P : E \longrightarrow E, \qquad Pf = \sum_{j=1}^{n} \langle f, e_j \rangle e_j$$

is called the **orthogonal projection** onto the subspace F. The mapping P is *linear*, i.e., it satisfies

$$P(f + g) = Pf + Pg, \quad P(\lambda f) = \lambda Pf \qquad (f, g \in E, \lambda \in \mathbb{K}).$$

By Exercise 1.8.b), P does only depend on the subspace F and not on the
Ex.1.8 chosen orthonormal basis of F used in the construction of P.

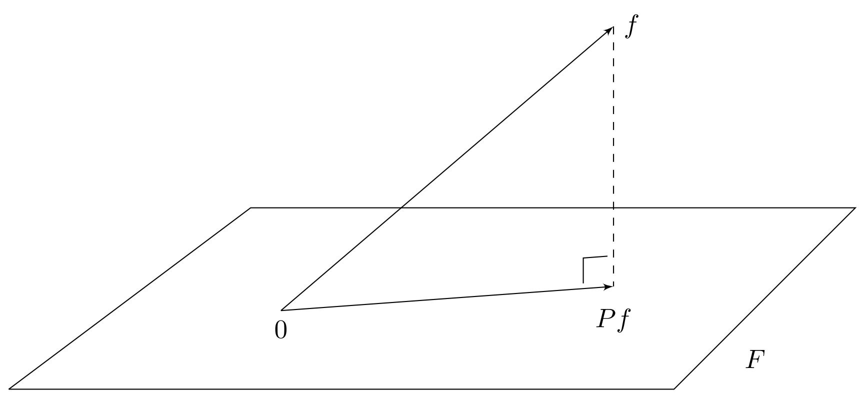

Figure 1. The orthogonal projection onto $F = \operatorname{span}\{e_1, \dots, e_n\}$.

Combining a) and b) of Lemma 1.10 one obtains **Bessel's inequality**[5]

$$\sum_{j=1}^{n} |\langle f, e_j \rangle|^2 = \|Pf\|^2 \le \|f\|^2 \qquad (f \in E). \tag{1.1}$$

[5] Friedrich Wilhelm Bessel (1784–1846), German mathematician and astronomer.

Orthogonal projections are an indispensable tool in Hilbert space theory and its applications. We shall see in Chapter 8 how to construct them in the case that the range space is no longer finite-dimensional.

By Lemma 1.10.a) each orthonormal system is a linearly independent set, i.e., a *basis* for its linear span. Assume for the moment that this span is already the whole space, that is,

$$E = \operatorname{span}\{e_1, \dots, e_n\}.$$

Now consider the (linear!) mapping

$$T : E \longrightarrow \mathbb{K}^n, \qquad Tf := (\langle f, e_1\rangle, \dots, \langle f, e_n\rangle).$$

By Lemma 1.10.a) T is exactly the **coordinatization mapping** associated with the algebraic basis $\{e_1, \dots, e_n\}$. Hence it is an *algebraic isomorphism.* However, more is true:

$$\langle Tf, Tg\rangle_{\mathbb{K}^n} = \langle f, g\rangle_E \qquad (f, g \in E) \tag{1.2}$$

where $\langle\cdot,\cdot\rangle_{\mathbb{K}^n}$ denotes the standard inner product on $\mathbb{K}^n$. As a consequence, one obtains that Ex.1.9

$$\|Tf\|_{2,\mathbb{K}^n} = \|f\|_E \qquad \text{for all } f \in E.$$

This means that T maps members of E onto members of $\mathbb{K}^n$ of equal length, and is therefore called a (linear) **isometry**.

The next, probably already well-known result shows that one can always find an orthonormal basis in an inner product space with finite or countable algebraic basis.

Lemma 1.11 (Gram[6]–Schmidt[7]). *Let $N \in \mathbb{N}\cup\{\infty\}$ and let $(f_n)_{1\le n<N}$ be a linearly independent set of vectors in an inner product space E. Then there is an orthonormal system $(e_n)_{1\le n<N}$ in E such that*

$$\operatorname{span}\{e_j \mid 0 \le j < n\} = \operatorname{span}\{f_j \mid 0 \le j < n\} \qquad \textit{for all } n \le N.$$

Proof. The construction is recursive. By the linear independence, f_1 cannot be the zero vector, so $e_1 := (1/\|f_1\|)f_1$ has norm one. Let $g_2 := f_2 - \langle f_2, e_1\rangle\, e_1$. Then $g_2 \perp e_1$. Since f_1, f_2 are linear independent, $g_2 \neq 0$ and so $e_2 := (1/\|g_2\|)g_2$ is the next unit vector.

[6] Jørgen Pedersen Gram (1850–1916), Danish mathematician.
[7] Erhard Schmidt (1876–1959), German mathematician.

Suppose that we have already constructed pairwise orthogonal unit vectors $e_1, \dots, e_{n-1}$ such that $\operatorname{span}\{e_1, \dots, e_{n-1}\} = \operatorname{span}\{f_1, \dots, f_{n-1}\}$. If $n = N$, we are done. Otherwise let

$$g_n := f_n - \sum\nolimits_{j=1}^{n-1} \langle f_n, e_j \rangle \, e_j.$$

Then $g_n \perp e_j$ for all $1 \le j < n$ (Lemma 1.10). Moreover, by the linear independence of the f_j and the construction of the e_j so far, $g_n \neq 0$. Hence
Ex.1.10 $e_n := (1/\|g_n\|) g_n$ is the next unit vector in the orthonormal system. □

As a corollary we obtain that for each finite-dimensional subspace F of an inner product space E, there exists the orthogonal projection from E onto F. In Chapter 8 we shall be occupied with the extension of this statement to the infinite-dimensional case.

1.3. The Trigonometric System

We now come to an important example of an orthonormal system in the inner product space $E = \mathrm{C}([0,1];\mathbb{C})$. Consider the functions

$$e_n(t) := \mathrm{e}^{2n\pi \mathrm{i} t} \qquad (t \in [0,1],\, n \in \mathbb{Z}),$$

where e is **Euler's constant**[8]. If $n \neq m$, then

$$\begin{aligned} \langle e_n, e_m \rangle &= \int_0^1 e_n(t)\overline{e_m(t)} \,\mathrm{d}t = \int_0^1 \mathrm{e}^{2\pi \mathrm{i}(n-m)t} \,\mathrm{d}t \\ &= \frac{\mathrm{e}^{2\pi \mathrm{i}(n-m)t}}{2\pi \mathrm{i}(n-m)}\bigg|_0^1 = \frac{1-1}{2\pi \mathrm{i}(n-m)} = 0 \end{aligned}$$

by the fundamental theorem of calculus. On the other hand,

$$\|e_n\|_2^2 = \int_0^1 \left|\mathrm{e}^{2\pi \mathrm{i} n t}\right|^2 \,\mathrm{d}t = \int_0^1 \mathbf{1} \,\mathrm{d}t = 1.$$

This shows that $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal system in the complex space $\mathrm{C}([0,1];\mathbb{C})$, the so-called **trigonometric system**. One can construct from this an orthonormal system in the real space $\mathrm{C}([0,1];\mathbb{R})$; see Exercise 1.14.

The number

$$\widehat{f}(n) := \langle f, e_n \rangle = \int_0^1 f(t)\overline{e_n(t)} \,\mathrm{d}t = \int_0^1 f(t)\mathrm{e}^{-2\pi \mathrm{i} n t} \,\mathrm{d}t$$

is called the n-th **Fourier coefficient** of f. Note that n ranges over the whole set of integers $\mathbb{Z}$. Bessel's inequality in this context reads

$$\sum\nolimits_{n=-N}^{N} |\widehat{f}(n)|^2 \le \|f\|_2^2 = \int_0^1 |f(t)|^2 \,\mathrm{d}t. \tag{1.3}$$

[8]Leonhard Euler (1707–1783), Swiss mathematician and physicist.

For $f \in \mathrm{C}[0,1]$ the abstract series

$$f \sim \sum\nolimits_{n=-\infty}^{\infty} \langle f, e_n \rangle \, \mathrm{e}^{2\pi i n t}$$

with respect to the trigonometric system is called its **Fourier series**.

Exercises 1A

Exercise 1.1. Here is a different way of proving Theorem 1.1. Suppose first that 0 is in the interior of $[a, b]$. Then prove the theorem by considering the derivatives $p^{(j)}(0)$ for $j = 0, \dots, n$. In the general case, find $a < c < b$ and use the change of variables $t = s - c$.

Exercise 1.2. Show that $\langle \cdot, \cdot \rangle : \mathrm{C}[a, b] \times \mathrm{C}[a, b] \to \mathbb{K}$ defined above is indeed sesquilinear, positive and symmetric on $\mathrm{C}[a, b]$.

Exercise 1.3. Show that in an inner product space $\|\lambda f\| = |\lambda| \, \|f\|$ for every $f \in E$ and $\lambda \in \mathbb{K}$. Treat complex scalars explicitly!

Exercise 1.4. a) Compute $\|\cdot\|_2$ of the monomials t^n, $n \in \mathbb{N}$, in the inner product space $\mathrm{C}[a, b]$ with standard inner product.

b) Let $E := \mathrm{P}[0, \infty)$ be the space of all polynomials, considered as functions on the half-line $[0, \infty)$. Define $\|p\|$ by

$$\|p\|^2 = \int_0^\infty |p(t)|^2 \, \mathrm{e}^{-t} \, \mathrm{d}t$$

for $p \in E$. Show that $\|p\|$ is a norm associated with an inner product on E. Prove all your claims.

Exercise 1.5. Let $(E, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{K} = \mathbb{C}$. Prove that for all $f, g \in E$ one has

$$\|f + ig\|^2 - \|f - ig\|^2 = 4 \operatorname{Im} \langle f, g \rangle \, .$$

Then conclude that the **polarization identity**

$$\langle f, g \rangle = \frac{1}{4} \left(\|f + g\|^2 - \|f - g\|^2 + i \, \|f + ig\|^2 - i \, \|f - ig\|^2 \right)$$

holds for all $f, g \in E$.

Exercise 1.6. Make a drawing that helps you to understand why the parallelogram law carries its name.

Exercise 1.7. Work out the induction proof of Pythagoras' lemma.

Exercise 1.8. Let $\{e_1, \dots, e_n\}$ be a finite orthonormal system in an inner product space $(E, \langle \cdot, \cdot \rangle)$, let $F := \operatorname{span}\{e_1, \dots, e_n\}$ and let $P : E \to F$ be the orthogonal projection onto F. Show that the following assertions hold:

a) $PPf = Pf$ for all $f \in E$.

b) If $f, g \in E$ are such that $g \in F$ and $f - g \perp F$, then $g = Pf$.

c) Each $f \in E$ has a *unique* representation as a sum $f = u + v$, where $u \in F$ and $v \in F^\perp$. (In fact, $u = Pf$.)

d) If $f \in E$ is such that $f \perp F^\perp$, then $f \in F$. (Put differently: $(F^\perp)^\perp = F$.)

e) Let $Qf := f - Pf$, $f \in E$. Show that $QQf = Qf$ and $\|Qf\| \le \|f\|$ for all $f \in E$.

Exercise 1.9. Prove the identity (1.2).

Exercise 1.10. Apply the Gram-Schmidt procedure to the polynomials $1, t, t^2$ in the inner product space $\mathrm{C}[-1,1]$ to construct an orthonormal basis of $F = \{p \in \mathrm{P}[-1,1] \mid \deg p \le 2\}$. (Continuing this for $t^3, t^4 \dots$ yields the sequence of so-called **Legendre polynomials**.[9])

Exercises 1B

Exercise 1.11. Apply the Gram–Schmidt procedure to the monomials $1, t, t^2$ in the inner product space $\mathrm{P}[0,\infty)$ with inner product

$$\langle f, g\rangle := \int_0^\infty f(t)\overline{g(t)}e^{-t}\,\mathrm{d}t.$$

Exercise 1.12. Let us call a function $f : [1,\infty) \to \mathbb{K}$ *mildly decreasing* if there is a constant $c = c(f)$ such that $|f(t)| \le ct^{-1}$ for all $t \ge 1$. Let

$$E := \{f : [1,\infty) \longrightarrow \mathbb{K} \mid f \text{ is continuous and mildly decreasing}\}.$$

a) Show that E is a linear subspace of $\mathrm{C}[1,\infty)$.

b) Show that

$$\langle f, g\rangle := \int_1^\infty f(t)\overline{g(t)}\,\mathrm{d}t$$

defines an inner product on E.

c) Apply the Gram–Schmidt procedure to the functions t^{-1}, t^{-2}.

Exercise 1.13. Let E be the space of polynomials of degree at most 2. On E define

$$\langle f, g\rangle := f(-1)\overline{g(-1)} + f(0)\overline{g(0)} + f(1)\overline{g(1)} \qquad (f, g \in E).$$

a) Show that this defines an inner product on E.

b) Describe $\{t^2 - 1\}^\perp$.

c) Show that the polynomials $t^2 - 1, t^2 - t$ are orthogonal, and find a nonzero polynomial $p \in E$ that is orthogonal to both of them.

Exercise 1.14. Let $(e_n)_{n\in\mathbb{Z}}$ be any orthonormal system in $\mathrm{C}([a,b];\mathbb{C})$, with standard inner product. Suppose further that $e_{-n} = \overline{e_n}$ for all $n \in \mathbb{Z}$. Show that

$$\{e_0\} \cup \{\sqrt{2}\,\mathrm{Re}\, e_n \mid n \in \mathbb{N}\} \cup \{\sqrt{2}\,\mathrm{Im}\, e_n \mid n \in \mathbb{N}\}$$

is an orthnormal system in the *real* inner product space $\mathrm{C}([a,b];\mathbb{R})$.

[9] Adrien-Marie Legendre (1752–1833), French mathematician.

Exercise 1.15. Each vector space E over $\mathbb{C}$ is also a vector space over $\mathbb{R}$. Show that if $(E, \langle\cdot,\cdot\rangle)$ is a complex inner product space, then

$$\langle f, g\rangle_r := \operatorname{Re}\langle f, g\rangle \qquad (f, g \in E)$$

is a real inner product on E satisfying $\langle \mathrm{i}f, \mathrm{i}g\rangle_r = \langle f, g\rangle_r$ for all $f, g \in E$.

Conversely, show that if $\langle\cdot,\cdot\rangle$ is a real inner product on the $\mathbb{C}$-vector space E such that $\langle \mathrm{i}f, \mathrm{i}g\rangle = \langle f, g\rangle$ for all $f, g \in E$, then

$$\langle f, g\rangle_c := \langle f, g\rangle + \mathrm{i}\,\langle f, \mathrm{i}g\rangle$$

is the unique complex inner product on E with $\langle\cdot,\cdot\rangle_{cr} = \langle\cdot,\cdot\rangle$.

Exercise 1.16. With the terminology from the previous exercise, let $(E, \langle\cdot,\cdot\rangle)$ be a complex inner product space. Show that $(e_j)_{j\in I}$ is a $\langle\cdot,\cdot\rangle$-orthonormal system in E if and only if $e_j, \mathrm{i}e_j$ $(j \in I)$ is an $\langle\cdot,\cdot\rangle_r$-orthonormal system in E.

Chapter 15

Baire's Theorem and Its Consequences

In this chapter we encounter two of the most important and useful results from functional analysis: the uniform boundedness principle and the open mapping theorem. The proofs need nothing more than a little metric space theory and some basics about Banach spaces and bounded operators, and hence the systematic part of this chapter could have been placed quite in the beginning (after Chapter 5, say).

15.1. Baire's Theorem

Both, the uniform boundedness principle and the open mapping theorem rest on a relatively simple but very powerful result about metric spaces called *Baire's theorem*[1]. Recall that an *open ball* in a metric space (Ω, d) is any set of the form

$$\mathrm{B}(x, r) = \{y \in \Omega \mid d(x, y) < r\}$$

for some "center" $x \in \Omega$ and "radius" $r > 0$. The *closed ball* with the same center and radius is

$$\mathrm{B}[x, r] = \{y \in \Omega \mid d(x, y) \leq r\}.$$

We have seen in Lemma 4.2 that an open ball is indeed an open, and a closed ball is indeed a closed subset of Ω. Note that

$$0 < s < r \quad \Rightarrow \quad \mathrm{B}[x, s] \subseteq \mathrm{B}(x, r) \subseteq \mathrm{B}[x, r]$$

[1] René-Louis Baire (1874–1932), French mathematician.

for any $x \in \Omega$. So every open ball contains a closed ball with the same
Ex.15.1 center and positive radius.

In its simplest form, Baire's theorem states that if a complete metric space is exhausted by a union of countably many closed subsets, then at least one of these subsets must contain an open ball. More precisely, we have the following.

Theorem 15.1 (Baire). *Let (Ω, d) be a nonempty complete metric space and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of closed subsets of Ω such that*

$$\Omega = \bigcup\nolimits_{n\in\mathbb{N}} A_n.$$

Then there is $n \in \mathbb{N}$ and $x \in \Omega, r > 0$ with $\mathrm{B}(x, r) \subseteq A_n$.

Theorem 15.1 is not exactly what is known as "Baire's theorem" in the literature, but a little weaker. For the full version see [**Rud87**, 5.6] and Exercise 15.11 below.

Ex.15.2 The proof of Baire's theorem will become quite perspicuous if one knows the following lemma.

Lemma 15.2 (Principle of Nested Balls). *Let (Ω, d) be a complete metric space, and let*

$$\mathrm{B}[x_1, r_1] \supseteq \mathrm{B}[x_2, r_2] \supseteq \mathrm{B}[x_3, r_3] \supseteq \dots$$

be a nested *sequence of closed balls in it. If $r_n \to 0$, then $x := \lim_{n\to\infty} x_n$ exists and*

$$\bigcap\nolimits_{n\in\mathbb{N}} \mathrm{B}[x_n, r_n] = \{x\}. \tag{15.1}$$

Proof. Note first that if $y \in \mathrm{B}[x_n, r_n]$ for all $n \in \mathbb{N}$, then $d(y, x_n) \le r_n \to 0$. Since limits are unique, the intersection $\bigcap_{n\in\mathbb{N}} \mathrm{B}[x_n, r_n]$ contains at most one point.

Now fix $\epsilon > 0$ and find $N \in \mathbb{N}$ such that $r_N \le \epsilon/2$. If $n, m \ge N$ we have $x_n, x_m \in \mathrm{B}[x_N, r_N]$, and hence

$$d(x_n, x_m) \le d(x_n, x_N) + d(x_N, x_m) \le r_N + r_N = 2r_N \le \epsilon.$$

Hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. By completeness, $x := \lim_{n\to\infty} x_n$ exists. For any $N \in \mathbb{N}$, the set $\mathrm{B}[x_N, r_N]$ is closed and contains all x_n with $n \ge N$. Consequently, $x \in \mathrm{B}[x_N, r_N]$, and this proves (15.1). □

Proof of Theorem 15.1. We suppose that no A_n contains an open ball, and claim that then there exists $x \in \Omega$ which is not contained in any A_n. To find that x we are going to construct a sequence of nested balls.

In the first step, pick any $x_1 \in \Omega \setminus A_1$. This must exist, otherwise $A_1 = \Omega$, which trivially contains an open ball. Since A_1 is closed, if $r_1 > 0$ is small enough one has

$$A_1 \cap \mathrm{B}[x_1, r_1] = \emptyset.$$

By hypothesis, A_2 does not contain $\mathrm{B}(x_1, r_1)$, so there is $x_2 \in \mathrm{B}(x_1, r_1)$ but $x_2 \notin A_2$. Since A_2 is closed and $\mathrm{B}(x_1, r_1)$ is open, for $r_2 > 0$ small enough we have

$$A_2 \cap \mathrm{B}[x_2, r_2] = \emptyset \quad \text{and} \quad \mathrm{B}[x_2, r_2] \subseteq \mathrm{B}(x_1, r_1).$$

Again by hypothesis, the set A_3 does not contain the ball $\mathrm{B}(x_2, r_2)$, and hence we find $x_3 \in \mathrm{B}(x_2, r_2)$ but $x_3 \notin A_3$. Since A_3 is closed and $\mathrm{B}(x_2, r_2)$ is open, for small enough $r_3 > 0$ we have

$$A_3 \cap \mathrm{B}[x_3, r_3] = \emptyset \quad \text{and} \quad \mathrm{B}[x_3, r_3] \subseteq \mathrm{B}(x_2, r_2).$$

In this manner we find a nested sequence of balls $\mathrm{B}[x_n, r_n]$ such that

$$\mathrm{B}[x_n, r_n] \cap A_n = \emptyset \qquad \text{for all } n \in \mathbb{N}. \tag{15.2}$$

Since in each step we can make the radius r_n as small as we like, we can arrange it such that $r_n \to 0$. By the principle of nested balls, the centers $(x_n)_{n\in\mathbb{N}}$ converge to $x \in \bigcap_{n\in\mathbb{N}} \mathrm{B}[x_n, r_n]$. By (15.2), $x \notin A_n$ for each $n \in \mathbb{N}$, and the proof is complete. □

Baire's theorem has plenty of interesting applications, in particular to the theory of real functions. For instance, one can employ it to show that the set of nowhere differentiable functions is dense in $\mathrm{C}[a, b]$; see [**Boa96**, Section 10]. However, going deeper here would lead us too far astray, so we restrict ourselves to (some) applications of Baire's theorem in functional analysis.

15.2. The Uniform Boundedness Principle

Recall that a linear mapping $T : E \to F$, where E, F are normed spaces, is *bounded* if there is a number $c \geq 0$ such that

$$\|Tf\| \leq c \|f\| \qquad \text{for all } f \in E.$$

In this case, the *operator norm*

$$\|T\| := \sup\{\|Tf\| \mid \|f\| \leq 1\}$$

is a finite number, and one has $\|Tf\| \leq \|T\| \, \|f\|$ for all $f \in E$ (see Chapter 2). In this section we consider whole collections $\mathcal{T}$ of bounded linear mappings.

Definition 15.3. Let E, F be normed linear spaces. A collection $\mathcal{T}$ of linear mappings from E to F is called **uniformly bounded** if there is a $c \geq 0$ such that

$$\|Tf\| \leq c\,\|f\| \qquad \text{for all } f \in E \text{ and all } T \in \mathcal{T}.$$

In other words, $\mathcal{T}$ is uniformly bounded if each $T \in \mathcal{T}$ is bounded and

$$\sup\{\|T\| \mid T \in \mathcal{T}\} < \infty,$$

i.e., $\mathcal{T}$ is a bounded subset of the normed space $\mathcal{L}(E;F)$.

We remark that we encountered uniformly bounded *sequences* of operators already in Section 9.6 in connection with the "strong convergence lemma" (Corollary 9.23) and Fejér's theorem.

Suppose that $\mathcal{T} \subseteq \mathcal{L}(E;F)$ is a uniformly bounded collection of linear operators. Then for each $f \in E$ one has

$$\|Tf\| \leq \|T\|\,\|f\| \leq \big(\sup_{S\in\mathcal{T}} \|S\|\big)\ \|f\|$$

for all $T \in \mathcal{T}$, and hence $\sup_{T\in\mathcal{T}} \|Tf\| < \infty$. We say that the operator family $\mathcal{T}$ is *pointwise bounded*. The uniform boundedness principle asserts that in case that E is complete, i.e., a Banach space, one has the converse implication.

Theorem 15.4 (Uniform Boundedness Principle)**.** *Let E be a Banach space, let F be a normed space, and let $\mathcal{T}$ be a collection of bounded linear operators from E to F. Then $\mathcal{T}$ is uniformly bounded if and only if it is pointwise bounded.*

One implication was already mentioned, and is trivial. For the proof of the nontrivial implication we need a little lemma.

Lemma 15.5. *Let $(E, \|\cdot\|)$ be a normed space and let $K \subseteq E$ be a subset with the following properties:*

1) *K is "midpoint-convex", i.e., if $f, g \in K$, then also $\frac{1}{2}(f+g) \in K$;*
2) *K is "symmetric", i.e., if $f \in K$, then also $-f \in K$.*

Ex.15.3 Ex.15.4 *Then, if K contains some open ball of radius $r > 0$, it also contains* $\mathrm{B}(0,r)$.

Proof. (See also Figure 22.) Suppose that $r > 0$ and $g \in E$ are such that $\mathrm{B}(g,r) \subseteq K$. By 2), K contains also $-\mathrm{B}(g,r)$ and by 1) it then must also contain $\frac{1}{2}(\mathrm{B}(g,r) - \mathrm{B}(g,r))$. But the latter contains $\mathrm{B}(0,r)$, as each $h \in E$ with $\|h\| < r$ can be written as

$$h = \frac{2h+g-g}{2} = \frac{(g+h)-(g-h)}{2} \in \tfrac{1}{2}\big(\mathrm{B}(g,r) - \mathrm{B}(g,r)\big). \qquad \square$$

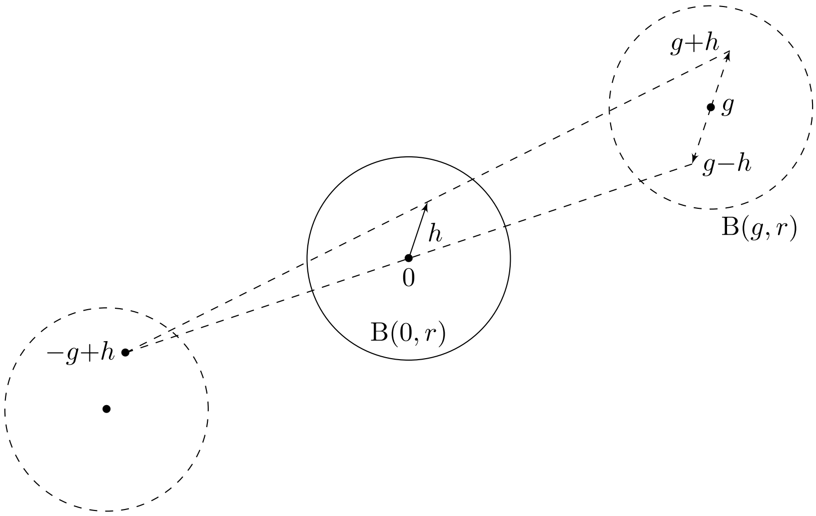

Figure 22. $\mathrm{B}(0,r) \subseteq {}^1\!/\!_2\big(\mathrm{B}(g,r) - \mathrm{B}(g,r)\big)$.

Proof of Theorem 15.4. Let $\mathcal{T} \subseteq \mathcal{L}(E;F)$ be pointwise bounded. For $n \in \mathbb{N}$, define

$$K_n := \{f \in E \mid \|Tf\| \le n \text{ for all } T \in \mathcal{T}\}.$$

Since each $T \in \mathcal{T}$ is bounded and the norm mapping is continuous, K_n is a closed subset of E. By hypothesis, every $f \in E$ is contained in at least one K_n, so

$$E = \bigcup\nolimits_{n\in\mathbb{N}} K_n.$$

Since E is complete, Baire's theorem applies and yields $n \in \mathbb{N}$, $r > 0$, and $g \in E$ with $\mathrm{B}(g,r) \subseteq K_n$. By straightforward arguments, K_n is midpoint-convex and symmetric. Hence Lemma 15.5 implies that $\mathrm{B}(0,r) \subseteq K_n$, and since K_n is closed, we have even $\mathrm{B}[0,r] \subseteq K_n$, by Exercise 15.1.

Now take $f \in E$ with $\|f\| \le 1$. Then $rf \in \mathrm{B}[0,r] \subseteq K_n$, which means that $r\,\|Tf\| = \|T(rf)\| \le n$ for each $T \in \mathcal{T}$. Dividing by r yields

$$\|Tf\| \le {}^n\!/\!_r \qquad \text{for all } T \in \mathcal{T},$$

and hence $\|T\| \le {}^n\!/\!_r$ for all $T \in \mathcal{T}$. □

We continue with an important consequence, complementing the strong convergence lemma (Corollary 9.23).

Theorem 15.6 (Banach–Steinhaus[2])**.** *Let E, F be Banach spaces, and let $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(E;F)$ be a sequence such that*

$$Tf := \lim_{n\to\infty} T_n f$$

exists for every $f \in E$. Then T is a bounded operator, $(T_n)_{n\in\mathbb{N}}$ is uniformly bounded, and

$$\|T\| \le \liminf_{n\to\infty} \|T_n\|\,.$$

Proof. For each $f \in E$, since $(T_n f)_{n\in\mathbb{N}}$ converges, also $(\|T_n f\|)_{n\in\mathbb{N}}$ converges, and therefore $\sup_{n\in\mathbb{N}} \|T_n f\| < \infty$. By the uniform boundedness principle, $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$. If $f \in E$ with $\|f\| \le 1$, then by the continuity of the norm,

$$\|Tf\| = \lim_{n\to\infty} \|T_n f\| = \liminf_{n\to\infty} \|T_n f\| \le \liminf_{n\to\infty} \|T_n\|\,.$$

Taking the supremum over all such f concludes the proof. □

15.3. Nonconvergence of Fourier Series

Let us turn to a nice application of the uniform boundedness principle to the theory of classical Fourier series.

Theorem 15.7 (Du Bois-Reymond)**.** *There exists a function $f \in \mathrm{C}_{\mathrm{per}}[0,1]$ such that its partial Fourier series at $t = 0$,*

$$S_n f(0) = \sum_{k=-n}^{n} \widehat{f}(k) e^{2\pi i k t}\Big|_{t=0} = \sum_{k=-n}^{n} \widehat{f}(k) \qquad (n \in \mathbb{N})$$

does not converge to $f(0)$.

Proof. We consider the linear functionals

$$T_n : \mathrm{C}_{\mathrm{per}}[0,1] \longrightarrow \mathbb{C}, \qquad T_n f := (S_n f)(0)$$

for $n \in \mathbb{N}$. Then

$$T_n f = \sum_{k=-n}^{n} \int_0^1 \mathrm{e}^{2\pi i k s} f(s)\,\mathrm{d}s = \int_0^1 \frac{\sin(2n+1)\pi s}{\sin \pi s} f(s)\,\mathrm{d}s$$

for all $f \in \mathrm{C}_{\mathrm{per}}[0,1]$. We now consider $E := \mathrm{C}_{\mathrm{per}}[0,1]$ with the supremum norm. Then each T_n is bounded and E is a Banach space. By the uniform boundedness principle it remains to prove that $\sup_{n\in\mathbb{N}} \|T_n\| = \infty$ to conclude the existence of a function $f \in \mathrm{C}_{\mathrm{per}}[0,1]$ such that $\sup_{n\in\mathbb{N}} |S_n f(0)| = \infty$. In particular, $\lim_{n\to\infty} S_n f(0)$ does not exist.

[2] Hugo Steinhaus (1887–1972), Polish mathematician.

Define the **Dirichlet kernel**

$$D_n(s) := \frac{\sin(2n+1)\pi s}{\sin \pi s},$$

so that $T_n f = \int_0^1 D_n(s) f(s)\, \mathrm{d}s$ for $f \in E$. We claim that

$$\|T_n\| = \int_0^1 |D_n(s)|\, \mathrm{d}s.$$

For C[0, 1] in place of $\mathrm{C}_{\mathrm{per}}[0, 1]$, this has been done in Example 2.27. But the argument there can be adapted, because D_n is periodic, and hence also the approximants $\overline{D_n}/(|D_n| + \epsilon)$ are periodic, i.e., contained in E.

Finally, we use the inequality $0 \le \sin s \le s$ for $s \in [0, \pi]$ to estimate

$$\begin{aligned}\int_0^1 |D_n(s)|\, \mathrm{d}s &= \int_0^1 \frac{|\sin(2n+1)\pi s|}{\sin \pi s}\, \mathrm{d}s = \frac{1}{\pi}\int_0^{\pi} \frac{|\sin(2n+1)s|}{\sin s}\, \mathrm{d}s \\ &\ge \frac{1}{\pi}\int_0^{\pi} \frac{|\sin(2n+1)s|}{s}\, \mathrm{d}s = \frac{1}{\pi}\int_0^{(2n+1)\pi} \frac{|\sin s|}{s}\, \mathrm{d}s \\ &= \sum_{k=1}^{2n+1} \frac{1}{\pi}\int_{(k-1)\pi}^{k\pi} \frac{|\sin s|}{s}\, \mathrm{d}s \ge \sum_{k=1}^{2n+1} \frac{1}{\pi^2 k}\int_{(k-1)\pi}^{k\pi} |\sin s|\, \mathrm{d}s \\ &= \frac{1}{\pi^2}\left(\int_0^{\pi} |\sin s|\, \mathrm{d}s\right)\sum_{k=1}^{2n+1}\frac{1}{k} = \frac{2}{\pi^2}\sum_{k=1}^{2n+1}\frac{1}{k}\end{aligned}$$

for all $n \in \mathbb{N}$. Hence

$$\sup_{n\in\mathbb{N}} \int_0^1 |D_n(s)|\, \mathrm{d}s = \infty$$

since the harmonic series diverges, and this concludes the proof. □

15.4. The Open Mapping Theorem

The second major result based on Baire's theorem is the so-called open mapping theorem. Actually, it is not just one theorem, but a collection of results all of the same flavor. Here is the most central formulation:

Theorem 15.8 (Open Mapping Theorem)**.** *Let E, F be Banach spaces and let $T : E \to F$ be a bounded linear mapping which is surjective. Then there is $a > 0$ such that for each $g \in F$ there is $f \in E$ with $\|f\| \le a\,\|g\|$ and $Tf = g$.*

The name "open mapping theorem" stems from the fact that the conclusion of the theorem is equivalent to saying that T maps open subsets of E onto open subsets of F. But at least for our purposes we can entirely avoid this equivalent formulation.

If T is not only surjective, but also injective, then it is algebraically invertible and Theorem 15.8 yields $\|T^{-1}g\| \le a\,\|g\|$ for all $g \in F$. But this just means that the inverse T^{-1} is bounded, with operator norm $\|T^{-1}\| \le a$. Hence we have established the following important corollary.

Corollary 15.9 (Bounded Inverse Theorem)**.** *If E, F are Banach spaces, and if $T \in \mathcal{L}(E; F)$ is bijective, then the inverse operator T^{-1} is also* Ex.15.5 *bounded.*

This theorem is by no means obvious, and it is false if one drops the completeness of one of the spaces. E.g., if $E = (\mathrm{C}[0,1], \|\cdot\|_\infty)$ and $F = (\mathrm{C}[0,1], \|\cdot\|_2)$, then the identity operator

$$\mathrm{I} : (\mathrm{C}[0,1], \|\cdot\|_\infty) \longrightarrow (\mathrm{C}[0,1], \|\cdot\|_2)$$

is bounded and bijective, but its inverse is not bounded.

Since boundedness of an operator is the same as continuity, Corollary 15.9 helps us to understand Theorem 15.8. Mere surjectivity of T just tells us that to any $g \in F$ there is always a pre-image of g, but it tells us nothing about the error we make in picking such a pre-image if g is slightly perturbed. Theorem 15.8 now says that this error can be *controlled.* In fact, it is amplified by a factor of at most a. However, the theorem does not tell how large a is, and thus it is of limited use in those situations when such information matters.

The following corollary sheds new light on our discussions of the supremum norm, L^2-norm and L^1-norm on $\mathrm{C}[a,b]$ in previous chapters.

Corollary 15.10. *Let E be a linear space that is a Banach space with respect to either one of two given norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on E. If there is $M \ge 0$ with*

$$\|f\|_2 \le M\,\|f\|_1 \qquad \textit{for all } f \in E,$$

then the two norms are equivalent.

Proof. The hypothesis just says that $\mathrm{I} : (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$ is bounded. If both are Banach spaces, Corollary 15.9 applies, and hence $\mathrm{I} = \mathrm{I}^{-1}$:

$(E, \|\cdot\|_2) \to (E, \|\cdot\|_2)$ is bounded, too. This yields a constant $M' \geq 0$ such that $\|f\|_1 \leq M' \|f\|_1$ for every $f \in E$, and hence both norms are equivalent. $\square$

The proof of the open mapping theorem is based on Baire's theorem — but not exclusively. Another ingredient is the following rather technical result, interesting in its own right. Roughly speaking, it says that if an operator is "approximately surjective" and one can pick "approximate pre-images" in a controlled way, then the mapping must be surjective. Here is the precise formulation.

Theorem 15.11. *Let E, F be Banach spaces, and let $T \in \mathcal{L}(E; F)$. Suppose that there exist $0 \leq q < 1$ and $a \geq 0$ such that for every $g \in F$ with $\|g\| \leq 1$ there is $f \in E$ such that*

$$\|f\| \leq a \quad \text{and} \quad \|Tf - g\| \leq q.$$

Then for each $g \in F$ there is $f \in E$ such that $Tf = g$ and $\|f\| \leq \frac{a}{1-q} \|g\|$.

Proof. It suffices to prove the statement in the case $\|g\| \leq 1$, because for $g = 0$ the assertion is trivial, and if $g \neq 0$ we may replace g by $(1/\|g\|)g$.

Let $A := \{Tf \mid f \in E,\ \|f\| \leq a\}$. Then the hypothesis asserts that

$$h \in F,\ \|h\| \leq 1 \quad \Longrightarrow \quad \exists\ h' \in A\ :\ \|h - h'\| \leq q.$$

We shall recursively construct a sequence $(g_n)_{n\in\mathbb{N}}$ in A such that

$$\left\| g - \sum\nolimits_{k=1}^{n} q^{k-1} g_k \right\| \leq q^n \tag{15.3}$$

for $n \geq 1$. The case $n = 1$ is just the hypothesis, since $\|g\| \leq 1$. Suppose that $g_1, \dots, g_n$ are already constructed such that (15.3) holds. Then by hypothesis there is $g_{n+1} \in A$ such that

$$\left\| q^{-n}\Big(g - \sum\nolimits_{k=1}^{n} q^{k-1} g_k\Big) - g_{n+1} \right\| \leq q,$$

and multiplying by q^n yields (15.3) with n replaced by $n + 1$.

By definition of A, for each $n \in \mathbb{N}$ we find $f_n \in E$ with $\|f_n\| \leq a$ and $Tf_n = g_n$. Since $0 \leq q < 1$ and E is a Banach space, the series $f := \sum_{n=1}^{\infty} q^{n-1} f_n$ converges absolutely in E, with

$$\|f\| \leq \sum\nolimits_{k=1}^{\infty} q^{k-1} \|f_k\| \leq \sum\nolimits_{k=1}^{\infty} a q^{k-1} = \frac{a}{1-q}.$$

Since T is bounded, one has

$$Tf = T\Big(\sum\nolimits_{k=1}^{\infty} q^{k-1} f_k\Big) = \sum\nolimits_{k=1}^{\infty} q^{k-1} T f_k = \sum\nolimits_{k=1}^{\infty} q^{k-1} g_k = g,$$

by (15.3). This concludes the proof. $\square$

Proof of Theorem 15.8. Let E, F be Banach spaces and let $T \in \mathcal{L}(E; F)$ be surjective. For $n \in \mathbb{N}$ define

$$B_n := \{g \in F \mid \exists\, f \in E \text{ s.t. } \|f\| \le n,\ \ Tf = g\} = T(\mathrm{B}_E[0, n]).$$

Note that B_n is midpoint-convex and symmetric, hence — by Exercise 15.3 — $A_n := \overline{B_n}$ has the same properties. Moreover, $F = \bigcup_{n\in\mathbb{N}} B_n$, by the surjectivity of T, and hence

$$F = \bigcup\nolimits_{n\in\mathbb{N}} A_n.$$

Since all the A_n are closed, and F is complete, Baire's theorem applies and we find an $n \in \mathbb{N}$ such that A_n contains an open ball. But A_n is midpoint-convex and symmetric, whence by Lemma 15.5 there is $r > 0$ such that $\mathrm{B}_E(0, r) \subseteq A_n$. Since A_n is closed, we have even

$$\mathrm{B}_F[0, r] \subseteq A_n = \overline{B_n} = \overline{T(\mathrm{B}_E[0, n])}. \tag{15.4}$$

Dividing by r yields

$$\mathrm{B}_F[0, 1] \subseteq \overline{T(\mathrm{B}_E[0, {}^{n}\!/_{r}])},$$

and this means that the hypotheses of Theorem 15.11 are satisfied with $a = {}^{n}\!/_{r}$ and any $q \in (0, 1)$. The conclusion of Theorem 15.11 yields exactly what we want. □

The Closed Graph Theorem. Let E, F be two normed spaces. A linear mapping $T : E \to F$ is said to have a **closed graph** if

$$f_n \to f, \quad Tf_n \to g \quad \Longrightarrow \quad Tf = g$$

holds for all sequences $(f_n)_{n\in\mathbb{N}}$ in E and all $f \in E$, $g \in F$. That means, T has closed graph if its graph

$$\mathrm{graph}(T) = \{(f, Tf) \mid f \in E\}$$

is closed in the normed space $E \times F$; see Exercise 4.20.

Theorem 15.12 (Closed Graph Theorem). *If E, F are Banach spaces and $T : E \to F$ is a linear mapping, then T is bounded if and only if it has a closed graph.*

Proof. Define the new norm $|||f||| := \|f\|_E + \|Tf\|_F$ for $f \in E$. Then $\|f\|_E \le |||f|||$ for each $f \in E$. The closedness of the graph of T and since both E, F are complete implies that E is complete with respect to this new
Ex.15.6 norm. Hence by Corollary 15.10 there must be a constant $c > 0$ such that $\|Tf\| \le |||f||| \le c\,\|f\|$ for all $f \in E$. □

15.5. Applications with a Look Back

In this section we sketch some situations where the results of this chapter apply.

Countable Algebraic Bases. A vector space E with a countable algebraic basis $(e_j)_{j\in\mathbb{N}}$ is linearly isomorphic to the space c_{00}, the space of finite sequences. We have seen many norms on it, but none was complete. This is actually a general fact.

Theorem 15.13. *A normed space with a countable algebraic basis is never a Banach space.*

Proof. Apply Baire's theorem to $A_n := \operatorname{span}\{e_1, \dots, e_n\}$. □ Ex.15.7

Boundedness and Convergence. Let $\alpha = (\alpha_j)_{j\in\mathbb{N}}$ be a scalar sequence. If $\alpha \in \ell^\infty$, then clearly

$$\sum\nolimits_{j=1}^{\infty} \alpha_j x_j$$

converges for every $x \in \ell^1$. We claim that the converse holds. Indeed, for each $n \in \mathbb{N}$ the linear functional

$$\varphi_n : \ell^1 \longrightarrow \mathbb{C}, \qquad \varphi_n(x) := \sum\nolimits_{j=1}^{n} \alpha_j x_j$$

is clearly bounded, and it is easy to see that

$$\|\varphi_n\| = \max\{|\alpha_1|, \dots, |\alpha_n|\}.$$

By assumption, $\lim_{n\to\infty} \varphi_n(x)$ exists for each $x \in \ell^1$, whence the sequence $(\varphi_n)_n$ is pointwise bounded. Since ℓ^1 is a Banach space, the uniform boundedness principle applies and yields that

$$\|\alpha\|_\infty = \sup\nolimits_{n\in\mathbb{N}} \max\{|\alpha_1|, \dots, |\alpha_n|\} = \sup\nolimits_{n\in\mathbb{N}} \|\varphi_n\| < \infty$$

as claimed. Can you find a direct proof without using the uniform boundedness principle? Ex.15.8 Ex.15.9

Nonequivalence and Noncompleteness. Recall the trivial inequality

$$\|f\|_2 \le \sqrt{b-a}\, \|f\|_\infty \qquad (f \in \mathrm{C}[a,b])$$

for the L^2-norm and the supremum norm on $\mathrm{C}[a,b]$. Consider the statements

(i) $\|\cdot\|_\infty$ and $\|\cdot\|_2$ are not equivalent.

(ii) $\mathrm{C}[a,b]$ is not complete with respect to $\|\cdot\|_2$.

(iii) $\mathrm{C}[a,b]$ is complete with respect to $\|\cdot\|_\infty$.

Each statement has been proved individually (Example 3.14, Theorem 5.8, Example 5.13). But the open mapping theorem (or better: Corollary 15.10) tells us that (i) and (ii) are equivalent in the presence of (iii). This means we could have saved some work on the expense of having only an abstract argument in place of a concrete counterexample.

Well-Posedness. Recall our discussion in Section 11.2 of the well-posedness of an equation on the basis of the Poisson problem. The solution operator there is $-\Delta_D^{-1} : \mathrm{L}^2(a,b) \to \mathrm{H}^2(a,b)$, and for well-posedness of the original equation one needs the boundedness of this operator, where one has to take the H^2-norm on the target space. This was established in Exercise 11.5, but the closed graph theorem renders this superfluous. Namely, one has the following corollary.

Corollary 15.14. *Let E, F, G be Banach spaces with $F \subseteq G$ and this inclusion is continuous. Let $T : E \to F$ be a bounded linear operator such that* $\operatorname{ran}(T) \subseteq F$. *Then $T : E \to F$ is bounded.*

Proof. That the inclusion is continuous means that $\|\cdot\|_G$ is weaker on F than $\|\cdot\|_F$. By the closed graph theorem it suffices to show that $T : E \to F$ has a closed graph. So suppose that $f_n, f \in E$ and $g \in F$ such that $\|f_n - f\|_E \to 0$ and $\|Tf_n - g\|_F \to 0$. Since T is continuous into G we obtain $\|Tf_n - Tf\|_G \to 0$, and since the inclusion $F \subseteq G$ is continuous we
Ex.15.10 have $\|Tf_n - g\|_G \to 0$. But limits are unique, whence $g = Tf$. □

Now, since we knew that $-\Delta_D^{-1}$ is a Hilbert–Schmidt operator — and hence bounded — on $\mathrm{L}^2(a,b)$, and since $\mathrm{H}^2(a,b)$ is a Banach space, Corollary 15.14 yields that $-\Delta_D^{-1}$ is bounded from L^2 to H^2. Similarly one can weaken the assumptions in Theorem 12.21, the criterion for being an abstract Hilbert–Schmidt operator; see Remark 12.22.

Invertible Operators and Spectral Theory. In our definition of an *invertible operator* $T : E \to F$ apart from the bijectivity of T we required also the boundedness of T^{-1}. If E, F are Banach spaces, this is automatic, by Corollary 15.9. In particular, if $T \in \mathcal{L}(E)$, and E is a Banach space, then $\lambda \in \mathbb{C}$ is in the spectrum of T if $\lambda\mathrm{I} - T$ is not bijective. This is a simpler definition of the spectrum $\sigma(T)$ than the one given on page 232.

The bounded inverse theorem also accounts for another proof of the well-posedness of the Poisson problem. Namely, $\operatorname{dom}(\Delta_D)$ is a closed subspace of $\mathrm{H}^2(a,b)$ (why?), hence a Hilbert space. Since $\mathrm{L}^2(a,b)$ is a Hilbert space and Δ_D is obviously bounded, its inverse must be bounded, which is the well-posedness.

Tietze's Theorem. Let (Ω, d) be a metric space. Any subset $A \subseteq \Omega$ is a metric space with respect to the induced metric, and if $f \in \mathrm{C_b}(\Omega)$ is a bounded continuous function, one can consider its restriction

$$Tf := f|_A \in \mathrm{C_b}(A)$$

to the set A. The operator $T : \mathrm{C_b}(\Omega) \to \mathrm{C_b}(A)$ is linear with $\|T\| \le 1$. Tietze's theorem states that if A is closed, then T is surjective.

Theorem 15.15 (Tietze[3]). *Let (Ω, d) a metric space, $A \subseteq \Omega$ a closed subset and $g \in \mathrm{C_b}(A; \mathbb{R})$. Then there is $h \in \mathrm{C_b}(\Omega; \mathbb{R})$ such that $h|_A = g$ and $\|h\|_\infty = \|g\|_\infty$.*

Note: For the case $\Omega = \mathbb{R}$ and $A = [a, b]$ the result is straighforward. But think of A being a "weird" closed set, e.g., the Cantor set!

Before we enter the actual proof, we need to note a general fact about metric spaces, interesting in its own right. Namely, if A and B are *disjoint* closed subsets of Ω, then the function

$$f(x) := \frac{d(x, A) - d(x, B)}{d(x, A) + d(x, B)} \qquad (x \in \Omega)$$

(see Definition 3.15 and Exercise 8.9) has the property that it is continuous with

$$f(\Omega) \subseteq [-1, 1], \quad f = -1 \text{ on } A \quad \text{and} \quad f = 1 \text{ on } B.$$

By adding constants and scaling one can modify f in such a way that for given $a < b$ one has $f(\Omega) \subseteq [a, b]$, $f = a$ on A and $f = b$ on B.

Proof of Tietze's theorem. We want to apply Theorem 15.11 with $E = \mathrm{C_b}(\Omega)$, $F = \mathrm{C_b}(A)$, T the restriction mapping as above. Let $g : A \to [-1, 1]$ be continuous. Then the sets $g^{-1}[-1, -1/3]$ and $g^{-1}[1/3, 1]$ are closed in Ω (why?). By the remarks above we can find a continuous function $f : \Omega \to [-1/3, 1/3]$ such that $f = -1/3$ on $g^{-1}[-1, -1/3]$ and $f = 1/3$ on $g^{-1}[1/3, 1]$. This implies that

$$\|f\|_\infty \le \frac{1}{3} =: a \quad \text{and} \quad \|f|_A - g\|_\infty \le \frac{2}{3} =: q$$

Theorem 15.11 applies and yields $h \in \mathrm{C_b}(\Omega)$ such that $h|_A = g$ and

$$\|h\|_\infty \le \frac{a}{1-q} \|g\|_\infty = \|g\|_\infty \le \|h\|_\infty$$

as claimed. □

[3] Heinrich Tietze (1880–1964), Austrian mathematician.

Exercises 15A

Exercise 15.1. Let E be a normed space. Show that

$$\overline{\mathrm{B}(f,r)} = \mathrm{B}[f,r]$$

for every $f \in E$ and $r > 0$.

Exercise 15.2 (Reformulation of Baire's Theorem)**.** Let (Ω, d) be a nonempty complete metric space, and let $O_n \subseteq \Omega$ be open in Ω for $n \in \mathbb{N}$. Show, using Baire's theorem 15.1, the implication

$$\overline{O_n} = \Omega \quad \text{for all } n \in \mathbb{N} \quad \Longrightarrow \quad \bigcap_{n\in\mathbb{N}} O_n \neq \emptyset.$$

Exercise 15.3. Let E be a normed space, and let $A \subseteq E$ be any subset. Show that if A is midpoint-convex/symmetric, then $\overline{A}$ is also midpoint-convex/symmetric.

Exercise 15.4. Let E be a normed space, and let $A \subseteq E$ be a *closed* midpoint-convex subset of E. Show that then E is *convex*, i.e., whenever $f, g \in A$ and $0 < \lambda < 1$, then also $\lambda f + (1-\lambda)g \in A$. [Hint: Consider first dyadic rationals $\lambda = j/2^n$ via successively taking midpoints, then employ an approximation argument.]

Exercise 15.5. Let E, F be Banach spaces, and let $T : E \to F$ be a bounded linear mapping such that $\ker(T) = \{0\}$. Show that if $\operatorname{ran}(T)$ is closed, then there is $c > 0$ such that $\|f\| \le c\,\|Tf\|$ for all $f \in E$.

Exercise 15.6. Let E, F be Banach spaces, let $T \in \mathcal{L}(E; F)$, and define $|||f||| := \|f\|_E + \|Tf\|_F$ for $f \in E$. Show that E is a Banach space with respect to this norm.

Exercise 15.7. Fill in the details in the proof of Theorem 15.13.

Exercise 15.8. Let $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ be a scalar sequence. Suppose that

$$\sum_{j=1}^{\infty} \alpha_j x_j$$

converges for every $x = (x_j)_{j\in\mathbb{N}}$ in c_0. Show that $\alpha \in \ell^1$.

Exercise 15.9. Let $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ be a scalar sequence. Suppose that

$$\sum_{j=1}^{\infty} \alpha_j x_j$$

converges for every $x = (x_j)_{j\in\mathbb{N}}$ in ℓ^2. Show that $\alpha \in \ell^2$.

Exercise 15.10. Prove the following generalization of Corollary 15.14: Let E, F, G be Banach space, let $T : E \to F$ and $S : F \to G$ be linear operators such that ST is bounded, S is bounded and injective. Then T is bounded, too.

Exercises 15B

Exercise 15.11 (Baire's Theorem, Full Version)**.** One can strengthen the result of Exercise 15.2 by making use of it. Let, as in that exercise, (Ω, d) be a nonempty complete metric space, and let $O_n \subseteq \Omega$ be open in Ω for $n \in \mathbb{N}$. Show the implication

$$\overline{O_n} = \Omega \quad \text{for all } n \in \mathbb{N} \quad \Longrightarrow \quad \overline{\bigcap_{n\in\mathbb{N}} O_n} = \Omega.$$

[Hint: Fix $x \in \Omega, r > 0$ and apply Exercise 15.2 to the metric space $\Omega' := \mathrm{B}[x, r]$ and the subsets $O'_n := O_n \cap \Omega'$, open in Ω'.]

Exercise 15.12. For a linear operator $T : E \to F$ between normed spaces E, F, show that

$$r\,\|T\| = \sup_{f\in\mathrm{B}[0,r]} \|Tf\| \le \sup_{f\in\mathrm{B}[g,r]} \|Tf\|$$

for every $g \in E$ and $r > 0$. [Hint: Consider $K = \{f \in E \mid \|Tf\| \le \sup_{h\in\mathrm{B}[g,r]} \|Th\|\}$ and apply Lemma 15.5.]

Exercise 15.13. Let E be a Banach space and let $P : E \to E$ be a *projection*, i.e., a linear operator satisfying $P^2 = P$. Show that P is bounded if and only if $F := \mathrm{ran}(P)$ is closed.

Exercise 15.14 (Hellinger–Toeplitz Theorem[4,5])**.** Let H be a Hilbert space, and let $T : H \to H$ be a linear operator such that

$$\langle Tf, g\rangle = \langle f, Tg\rangle \quad \text{for all } f, g \in H.$$

Show that T is bounded.

Exercise 15.15. Let E, F, G be Banach spaces, and let $B : E \times F \to G$ be a bilinear mapping. Suppose that B is separately continuous, i.e., for each $f \in E$ and $g \in F$ the mappings

$$B(f, \cdot) : F \to G \quad \text{and} \quad B(\cdot, g) : E \to G$$

are continuous. Show that there is $c \ge 0$ such that

$$\|B(f, g)\|_G \le c\,\|f\|_E\,\|g\|_G \quad \text{for all } f \in E, g \in F.$$

Exercise 15.16. Let E, F be Banach spaces and let $T \in \mathcal{L}(E; F)$ be surjective. Show that for $S \in \mathcal{L}(E; F)$ such that $\|S - T\|$ is small enough, then S is also surjective. [Hint: Let a be as in Theorem 15.8 and $\|S - T\| < 1/a$. Then Theorem 15.11 can be applied to S with $q = \|S - T\|\,a$].

Exercise 15.17. Can one modify the proof of Tietze's theorem in such a way that one can replace the space $\mathrm{C_b}$ of bounded continuous functions everywhere by $\mathrm{UC_b}$, the space of bounded *uniformly* continuous functions? (See Exercise 5.14.)

[4]Ernst Hellinger (1883–1950), German mathematician.
[5]Otto Toeplitz (1881–1940), German mathematician.

Exercises 15C

Exercise 15.18. a) Consider, for $f \in \mathrm{C}_{\text{per}}[0,1]$, the convolution operator

$$T : \mathrm{L}^1(0,1) \to \mathrm{L}^1(0,1), \qquad Tg = f * g.$$

Show that $\|T\| = \|f\|_1$. [Hint: Fejér's theorem.]

b) Use a) to show that there is at least one function $f \in \mathrm{L}^1(0,1)$ such that the sequence of partial Fourier sums

$$S_n f = \sum_{k=-n}^{n} \widehat{f}(k) \mathrm{e}^{2\pi \mathrm{i} k\, t}$$

does not converge with respect to $\|\cdot\|_1$.

Remark: It can be shown that $\|S_n f - f\|_p \to 0$ as $n \to \infty$ for each $f \in \mathrm{L}^p(0,1)$, $1 < p < \infty$. The case $p = 2$ is "simple" (Corollary 9.15), but for $p \neq 2$ this is rather involved; see [**Kat04**, II.1.5].