Linear Analysis (Fall 2001)

Volker Runde

August 22, 2003

Contents

Introduction

These are the TeXed and polished notes of the course Math 516 (Linear Analysis[1]) as I taught it in the fall terms 2000 and 2001. The most distinctive feature of these notes is their complete lack of originality: Everything can be found in one textbook or another. The book that is probably closest in spirit is

1. J. B. Conway, *A Course in Functional Analysis*. Springer Verlag, 1985.

Other recommended books are:

2. B. Bollobás, *Linear Analysis. An Introductory Course*, Second Edition. Cambridge University Press, 1999.

3. N. Dunford and J. T. Schwartz, *Linear Operators*, I. Wiley-Interscience, 1988.

4. G. K. Pedersen, *Analysis Now*. Springer Verlag, 1989.

5. W. Rudin, *Functional Analysis*, Second Edition. McGraw-Hill, 1991.

The notes, which are generally kept in a rather brutal theorem-proof style, are not intended to replace any of these books, but rather to supplement them by relieving the students from the necessity of taking notes and thus allowing them to devote their full attention to the lecture.

Volker Runde, Edmonton — August 22, 2003

[1]called "baby functional analysis" by some

Chapter 1

Basic concepts

In this chapter, we introduce the main objects of study in this course:

- normed linear spaces, in particular Banach spaces, and
- the bounded linear maps between them.

1.1 Normed spaces and Banach spaces

All linear spaces considered in these notes are supposed to be over a field $\mathbb{F}$, which can be $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1.1 Let E be a linear space. A *norm* on E is a map $\|\cdot\| \colon E \to [0,\infty)$ such that

(a) $\|x\| = 0 \quad \iff \quad x = 0 \qquad (x \in E)$;

(b) $\|\lambda x\| = |\lambda|\|x\| \qquad (\lambda \in \mathbb{F},\, x \in E)$;

(c) $\|x + y\| \leq \|x\| + \|y\| \qquad (x, y \in E)$.

A linear space equipped with a norm is called a *normed space*.

Examples 1. There are several canonical norms on the linear space $E := \mathbb{F}^N$. For $x = (\lambda_1, \ldots, \lambda_N)$, let:

$$\begin{aligned} \|x\|_1 &:= \sum_{j=1}^{N} |\lambda_j|; \\ \|x\|_2 &:= \left(\sum_{j=1}^{N} |\lambda_j|^2\right)^{\frac{1}{2}}; \\ \|x\|_\infty &:= \max\{|\lambda_1|, \ldots, |\lambda_N|\}. \end{aligned}$$

Then $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are norms on E which satisfy the inequality

$$\|x\|_\infty \leq \|x\|_1 \leq \sqrt{N}\|x\|_2 \leq N\|x\|_\infty \qquad (x \in E). \tag{1.1}$$

2. Let $S \neq \varnothing$ be a set and define

$$\ell^\infty(S, \mathbb{F}) := \left\{ f : S \to \mathbb{F} : \sup_{s \in S} |f(s)| < \infty \right\}.$$

For $f \in \ell^\infty(S, \mathbb{F})$, define

$$\|f\|_\infty := \sup_{s \in S} |f(s)|.$$

This turns $\ell^\infty(S, \mathbb{F})$ into a normed space.

3. In fact, there is a norm on *any* linear space E. Let S be a Hamel basis (see Exercise 1.2 below) for E. Let $x \in E$. Then there are (necessarily unique) $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ along with $s_1, \ldots, s_n \in S$ such that $x = \sum_{j=1}^n \lambda_j s_j$. Define

$$\|x\| := \sum_{j=1}^n |\lambda_j|.$$

This defines a norm on E.

The last example emphasizes that a normed space is not just a linear space that can be equipped with a norm, but a linear space equipped with a particular norm.

Exercise 1.1 Justify (1.1).

Exercise 1.2 Let $E \neq \{0\}$ be a (possibly infinite-dimensional) linear space. A *Hamel basis* for E is a set S of elements of E with the following properties:

- S is linearly independent.
- Each element of E is a linear combination of elements of S.

Show that E has a Hamel basis by proceeding as follows:

(i) Let $\mathcal{S} := \{T \subset E : T \text{ is linearly independent}\}$. Show that $\mathcal{S} \neq \varnothing$.

(ii) Let $\mathcal{T}$ be a non-empty subset of $\mathcal{S}$ which is totally ordered by set inclusion. Show that $\bigcup\{T : T \in \mathcal{T}\}$ belongs again to $\mathcal{S}$.

(iii) Use Zorn's lemma to conclude that $\mathcal{S}$ has maximal elements.

(iv) Show that any such maximal element is a Hamel basis for E.

Exercise 1.3 Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

(i) Show that

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{x}{p} + \frac{y}{q} \qquad (x, y > 0). \tag{1.2}$$

(*Hint*: Apply the logarithm to (1.2) and prove that inequality first.)

For $x = (\lambda_1, \ldots, \lambda_N) \in \mathbb{F}^N$, let

$$\|x\|_p := \left(\sum_{j=1}^{N} |\lambda_j|^p \right)^{\frac{1}{p}}.$$

(ii) *Hölder's inequality*. Show that, for $x = (\lambda_1, \ldots, \lambda_N), y = (\mu_1, \ldots, \mu_N) \in \mathbb{F}^N$, we have

$$\sum_{j=1}^{N} |\lambda_j \mu_j| \leq \|x\|_p \|y\|_q.$$

Which known inequality do you obtain for $p = q = 2$?

(iii) *Minkowski's inequality*. Show that

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \qquad (x, y \in \mathbb{F}^N).$$

(iv) Conclude that $\| \cdot \|_p$ is a norm on $\mathbb{F}^N$.

Exercise 1.4 Let F be a linear subspace of a normed space E. Show that the closure of F in E is also a linear subspace of E.

Exercise 1.5 A *seminorm* on a linear space E is a map $p : E \to [0, \infty)$ with the following properties:

- $p(\lambda x) = |\lambda| p(x) \qquad (\lambda \in \mathbb{F},\, x \in E)$;
- $p(x + y) \leq p(x) + p(y) \qquad (x, y \in E)$.

What is missing from the definition of a norm?

(i) Show that $F := \{x \in E : p(x) = 0\}$ is a linear subspace of E.

(ii) For $x \in E$, define $\|x + F\| := p(x)$. Show that $\||\cdot\||$ is a norm on E/F.

If $(E, \|\cdot\|)$ is a normed space, then

$$d \colon E \times E \to [0, \infty), \quad (x, y) \mapsto \|x - y\|$$

is a metric. We may thus speak of convergence, etc., in normed spaces.

Definition 1.1.2 A normed space E is called a *Banach space* if the corresponding metric space is complete, i.e. every Cauchy sequence in E converges in E.

Exercise 1.6 Let E be a normed space, and let F be a linear subspace of E. Show that, if F is a Banach space, then it is closed in F. Conversely, show that, if E is a Banach space and if F is closed in F, then F is a Banach space.

Examples 1. $(\mathbb{F}^N, \|\cdot\|_j)$ is a Banach space for $j = 1, 2, \infty$.

2. Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $(\ell^\infty(S,\mathbb{F}), \|\cdot\|_\infty)$. For each $s \in S$, we have

$$|f_n(s) - f_m(s)| \le \|f_n - f_m\|_\infty \qquad (n, m \in \mathbb{N}).$$

Hence, $(f_n(s))_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{F}$. Define $f\colon S \to \mathbb{F}$ by letting

$$f(s) := \lim_{n\to\infty} f_n(s).$$

We claim that $f \in \ell^\infty(S,\mathbb{F})$ and that $\|f_n - f\|_\infty \to 0$. We prove both claims simultaneously. Let $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that

$$\|f_n - f_m\| < \epsilon \qquad (n, m \ge N).$$

It follows that

$$|f_n(s) - f_m(s)| \le \|f_n - f_m\| < \epsilon \qquad (n, m \ge N,\ s \in S).$$

We thus obtain for $n \ge N$ and $s \in S$:

$$\begin{aligned} |f_n(s) - f(s)| &= \lim_{m\to\infty} |f_n(s) - f_m(s)| \\ &\le \limsup_{m\to\infty} \|f_n - f_m\| \\ &\le \epsilon. \end{aligned} \tag{1.3}$$

In particular, we obtain

$$|f(s)| \le |f_N(s)| + \epsilon \le \|f_N\|_\infty + \epsilon,$$

so that $f \in \ell^\infty(S,\mathbb{F})$. Taking the supremum over $s \in S$ in (1.3) yields

$$\|f_n - f\|_\infty \le \epsilon$$

3. Let X be a topological space, and define

$$\mathcal{C}_b(X,\mathbb{F}) := \{f \in \ell^\infty(X,\mathbb{F}) : f \text{ is continuous}\}.$$

By Theorem A.2.4, $\mathcal{C}_b(X,\mathbb{F})$ is a closed subspace of $\ell^\infty(X,\mathbb{F})$ and thus a Banach space by Exercise 1.4.

4. Let X be a locally compact Hausdorff space, and let $\mathcal{C}_0(X,\mathbb{F})$ be defined as in Definition A.3.13. Let $(f_n)_{n=1}^\infty$ be a sequence in $\mathcal{C}_0(X,\mathbb{F})$ converging to $f \in \mathcal{C}_b(X,\mathbb{F})$ (with respect to $\|\cdot\|_\infty$). Let $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that

$$\|f_n - f\|_\infty < \frac{\epsilon}{2} \qquad (n \ge N).$$

It follows that

$$|f_N(x) - f(x)| \leq \|f_N - f\|_\infty < \frac{\epsilon}{2}$$

and consequently

$$|f(x)| \leq |f_N(x)| + \frac{\epsilon}{2}.$$

Since $f_N \in \mathcal{C}_0(X, \mathbb{F})$, there is a compact set $K \subset X$ such that $\sup_{x \in X \setminus K} |f(x)| < \frac{\epsilon}{2}$. This implies

$$\sup_{x \in X \setminus K} |f(x)| < \epsilon.$$

Hence, $\mathcal{C}_0(X, \mathbb{F})$ is a closed subspace of the Banach space $\mathcal{C}_b(X, \mathbb{F})$ and thus a Banach space itself.

5. For $N \in \mathbb{N}$, define

$$\mathcal{C}^N([0,1], \mathbb{F}) := \{f \colon [0,1] \to \mathbb{F} : f \text{ is } N\text{-times continuously differentiable}\}.$$

Define

$$f \colon [0,1] \to \mathbb{F}, \quad x \mapsto \begin{cases} \frac{1}{2} - x, & x \in \left[0, \frac{1}{2}\right], \\ x - \frac{1}{2}, & x \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Then f belongs to $\mathcal{C}([0,1])$[1], but not to $\mathcal{C}^N([0,1])$. Since f is piecewise continuously differentiable, it can be uniformly approximated by the sequence $(s_n)_{n=1}^\infty$ of the partial sums of its Fourier series. Since $s_n \in \mathcal{C}^N([0,1])$ for each $n \in \mathbb{N}$, we obtain that $\mathcal{C}^N([0,1])$ is not closed in $(\mathcal{C}([0,1]), \|\cdot\|_\infty)$. Hence, $(\mathcal{C}^N([0,1]), \|\cdot\|_\infty)$ is not a Banach space.

Define another norm on $\mathcal{C}^N([0,1])$:

$$\|f\|_N := \sum_{j=1}^N \|f^{(j)}\|_\infty \qquad (f \in \mathcal{C}^N([0,1]).$$

Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $(\mathcal{C}^N([0,1]), \|\cdot\|_N)$. Then $(f^{(j)})_{n=1}^\infty$ is a Cauchy sequence in $(\mathcal{C}([0,1]), \|\cdot\|_\infty)$ for $j = 0, 1, \ldots, N$. Since $\mathcal{C}([0,1])$ is a Banach space, there are $g_0, g_1, \ldots, g_N \in \mathcal{C}([0,1])$ such that

$$\|f_n^{(j)} - g_j\|_\infty \to 0 \qquad (j = 0, 1, \ldots, N).$$

We claim that

$$g_{j+1} = g_j' \qquad (j = 0, \ldots, N-1).$$

[1]If the field $\mathbb{F}$ is obvious or irrelevant, we often write $\mathcal{C}_0(x)$, $\mathcal{C}^N([0,1])$, etc., without the symbol $\mathbb{F}$.

This, however, follows immediately from

$$\begin{aligned} g_j(x) &= \lim_{n\to\infty} f_n^{(j)}(x) \\ &= \lim_{n\to\infty} \left[\int_0^x f^{(j+1)}(t)\,dt + f_n^{(j)}(0)\right] \\ &= \int_0^x g_{j+1}(t)\,dt + g_j(0) \qquad (x \in [0,1],\ j = 0,\ldots,N-1). \end{aligned}$$

Let $f := g_0$. Then we obtain inductively that $f^{(j)} = g_j$ for $j = 0,1,\ldots,N$. This yields

$$\|f_n - f\|_N = \sum_{j=0}^{N} \|f_n^{(j)} - f^{(j)}\|_\infty = \sum_{j=0}^{N} \|f_n^{(j)} - g_j\|_\infty \to 0.$$

Hence, $(\mathcal{C}^N([0,1]), \|\cdot\|_N)$ is a Banach space.

The last example shows that a linear space equipped with one norm may fail to be a Banach space, but can be a Banach space with respect to another norm.

Exercise 1.7 We write shorthand $\ell^\infty(\mathbb{F})$ for $\ell^\infty(\mathbb{N},\mathbb{F})$. Let $c(\mathbb{F})$ denote the subspace of $\ell^\infty(\mathbb{F})$ consisting of all convergent sequences in $\mathbb{F}$, let $c_0(\mathbb{F})$ be the subspace of all sequences in $\mathbb{F}$ that converge to zero, and let $c_{00}(\mathbb{F})$ consist of all sequences $(\lambda_n)_{n=1}^\infty$ in $\mathbb{F}$ such that $\lambda_n = 0$ for all but finitely many $n \in \mathbb{N}$.

(i) Show that $(c(\mathbb{F}), \|\cdot\|_\infty)$ and $(c_0(\mathbb{F}), \|\cdot\|_\infty)$ are Banach spaces.

(ii) Show that $(c_{00}(\mathbb{F}), \|\cdot\|_\infty)$ is not a Banach space.

Our first theorem is an often useful characterization of completeness in normed spaces in terms of convergent series:

Definition 1.1.3 Let $(x_n)_{n=1}^\infty$ be a sequence in the normed space E. We say that the series $\sum_{n=1}^\infty x_n$ *converges* in E if the sequence $\left(\sum_{n=1}^N x_n\right)_{N=1}^\infty$ of its partial sums converges. We say that $\sum_{n=1}^\infty x_n$ *converges absolutely* if $\sum_{n=1}^\infty \|x_n\| < \infty$.

Exercise 1.8 For $n \in \mathbb{N}$, let

$$x_n \colon \mathbb{N} \to \mathbb{F}, \quad m \mapsto \begin{cases} \dfrac{1}{n}, & m = n, \\ 0, & \text{else}. \end{cases}$$

(i) Show that, for every permutation $\pi \colon \mathbb{N} \to \mathbb{N}$, the series $\sum_{n=1}^\infty x_{\pi(n)}$ converges in c_0 to the same limit. Which is it?

(ii) Show that the series $\sum_{n=1}^\infty x_n$ is not absolutely convergent.

Theorem 1.1.4 (Riesz–Fischer theorem) *Let E be a normed space. Then the following are equivalent:*

(i) *E is a Banach space.*

(ii) *Every absolutely converging series in E converges.*

Proof (i) $\Longrightarrow$ (ii): This is proven in the same fashion as for series in $\mathbb{R}$.

(ii) $\Longrightarrow$ (i): Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence. Choose a subsequence $(x_{n_k})_{k=1}^\infty$ of $(x_n)_{n=1}^\infty$ such that

$$\|x_{n_k} - x_{n_{k+1}}\| < \frac{1}{2^k} \qquad (k \in \mathbb{N}).$$

Let

$$y_k := x_{n_k} - x_{n_{k+1}} \qquad (k \in \mathbb{N}).$$

Then the series $\sum_{k=1}^\infty y_k$ converges absolutely and thus converges in E. Since

$$\sum_{k=1}^K y_k = (x_{n_1} - x_{n_2}) + (x_{n_2} - x_{n_3}) + \cdots + (x_{n_K} - x_{n_{K+1}}) = x_{n_1} - x_{n_{K+1}} \qquad (K \in \mathbb{N}),$$

it follows that $(x_{n_k})_{k=1}^\infty$ is also convergent, with limit x, say. Let $\epsilon > 0$, and choose $K, N \in \mathbb{N}$ with $n_K \geq N$

$$\|x_n - x_m\| < \frac{\epsilon}{2} \quad (n, m \geq N) \qquad \text{and} \qquad \|x_{n_k} - x\| < \frac{\epsilon}{2} \quad (k \geq K).$$

For $n \geq \max\{N, K\}$, this means

$$\|x_n - x\| \leq \|x_n - x_{n_k}\| + \|x_{n_k} - x\| < \epsilon,$$

so that $x = \lim_{n\to\infty} x_n$. $\square$

Example Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, let $p \in [1, \infty)$, and let

$$\|f\|_p := \left(\int_\Omega |f(\omega)|^p \, d\mu(\omega) \right)^{\frac{1}{p}}$$

for each measurable function $f \colon \Omega \to \mathbb{F}$. Define

$$\mathcal{L}^p(\Omega, \mathfrak{S}, \mu) := \{f \colon \Omega \to \mathbb{F} : f \text{ is measurable with } \|f\|_p < \infty\}.$$

Using Hölder's and Minkowski's inequalities (compare Exercise 1.3), it can be shown that $\mathcal{L}^p(\Omega, \mathfrak{S}, \mu)$ is a linear space, and $\|\cdot\|_p$ is a seminorm on it. Let

$$\mathcal{N}_p := \{f \in \mathcal{L}^p(\Omega, \mathfrak{S}, \mu) : \|f\|_p = 0\},$$

and define

$$L^p(\Omega, \mathfrak{S}, \mu) := \mathcal{L}^p(\Omega, \mathfrak{S}, \mu)/\mathcal{N}_p.$$

By Exercise 1.5, $\|\cdot\|_p$ induces a norm on $L^p(\Omega, \mathfrak{S}, \mu)$, which we denote by $\|\cdot\|_p$ as well. We claim that $(L^p(\Omega, \mathfrak{S}, \mu), \|\cdot\|_p)$ is a Banach space.

We will use Theorem 1.1.4. Let $(f_n)_{n=1}^\infty$ be a sequence in $\mathcal{L}^p(\Omega, \mathfrak{S}, \mu)$ such that $\sum_{n=1}^\infty \|f_n\|_p < \infty$. Define

$$g \colon \Omega \to [0, \infty], \quad \omega \mapsto \left(\sum_{n=1}^\infty |f_n(\omega)| \right)^p.$$

We claim that g is integrable. To see this, note that

$$\left(\int_\Omega \left(\sum_{n=1}^N |f_n(\omega)| \right)^p d\mu(\omega) \right)^{\frac{1}{p}} = \left\| \sum_{n=1}^N |f_n| \right\|_p \leq \sum_{n=1}^N \|f_n\|_p \qquad (N \in \mathbb{N}).$$

By the monotone convergence theorem (Theorem B.3.1), this means

$$\int_\Omega g(\omega)\, d\mu(\omega) = \lim_{N\to\infty} \int_\Omega \left(\sum_{n=1}^N |f_n(\omega)| \right)^p d\mu(\omega) \leq \left(\sum_{n=1}^\infty \|f_n\|_p \right)^p < \infty,$$

so that g is indeed integrable. Hence, for almost all $\omega \in \Omega$, the series $\sum_{n=1}^\infty f(\omega)$ converges absolutely. Define

$$f \colon \Omega \to \mathbb{F}, \quad \omega \mapsto \begin{cases} \sum_{n=1}^\infty f_n(\omega), & \text{if } g(\omega) < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

Then f is measurable with $|f|^p \leq g$. For almost all $\omega \in \Omega$, we have

$$\lim_{N\to\infty} \left| \sum_{n=1}^N f_n(\omega) - f(\omega) \right| = 0 \qquad \text{and} \qquad \left| \sum_{n=1}^N f_n(\omega) - f(\omega) \right|^p \leq g(\omega).$$

For the dominated convergence theorem (Theorem B.3.2) it is then easily inferred that $\lim_{N\to\infty} \left\| \sum_{n=1}^N f_n - f \right\|_p = 0$, i.e. $f = \sum_{n=1}^\infty f_n$.

The following exercise is the measure theory free version of the preceding example:

Exercise 1.9 Let $p \in [1, \infty)$. Let ℓ^p be the set of all sequences $(x_n)_{n=1}^\infty$ in $\mathbb{F}$ with $\sum_{n=1}^\infty |x_n|^p < \infty$.

(i) For $x = (x_n)_{n=1}^\infty \in \ell^p$ define

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

Show that $\|\cdot\|_p$ is a norm on ℓ^p.

(ii) Show that $(\ell^p, \|\cdot\|_p)$ is a Banach space.

Example Let $(\Omega, \mathfrak{S}, \mu)$ be a σ-finite measure space. A measurable function $f : \Omega \to \mathbb{F}$ is called *essentially bounded* if there is $C \geq 0$ such

$$\{\omega \in \Omega : |f(\omega)| \geq C\} \tag{1.4}$$

is a μ-zero set. Let $\mathcal{L}^\infty(\Omega, \mathfrak{S}, \mu)$ denote the set of all essentially bounded functions on Ω. For $f \in \mathcal{L}^\infty(\Omega, \mathfrak{S}, \mu)$ define

$$\|f\|_\infty := \inf\{C \geq 0 : (1.4) \text{ is a } \mu\text{-zero set}\}.$$

It is easy to see that $\|\cdot\|_\infty$ is a seminorm on $\mathcal{L}^\infty(\Omega, \mathfrak{S}, \mu)$. Let $\mathcal{N}_\infty := \{f \in \mathcal{L}^\infty(\Omega, \mathfrak{S}, \mu) : \|f\|_\infty = 0\}$, and define

$$L^\infty(\Omega, \mathfrak{S}, \mu) := \mathcal{L}^\infty(\Omega, \mathfrak{S}, \mu)/\mathcal{N}_\infty.$$

Then $L^\infty(\Omega, \mathfrak{S}, \mu)$ equipped with the norm induced by $\|\cdot\|_\infty$ — likewise denoted by $\|\cdot\|_\infty$ — is a Banach space.

1.2 Finite-dimensional spaces

We have seen in the previous section, that there may be different norms on one linear space: $(\mathcal{C}^N([0,1]), \|\cdot\|_\infty)$ is not a Banach space whereas $(\mathcal{C}^N([0,1]), \|\cdot\|_1)$ isn't. On the other hand, the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ on $\mathbb{F}^N$ are related by (1.1), so that the resulting topologies are identical. The following definition captures this phenomenon:

Definition 1.2.1 Let E be a linear space. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are called *equivalent* (in symbols: $\|\cdot\|_1 \sim \|\cdot\|_2$) if there is $C \geq 0$ such that

$$\|x\|_1 \leq C\|x\|_2 \quad \text{and} \quad \|x\|_2 \leq C\|x\|_1 \qquad (x \in E).$$

Exercise 1.10 Verify that the equivalence of norms is indeed an equivalence relation, i.e. it is reflexive, symmetric, and transitive.

Exercise 1.11 Let E be a linear space, and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two equivalent norms on E. Verify in detail that $(E, \|\cdot\|_1)$ is a Banach space if and only if $(E, \|\cdot\|_2)$ is a Banach space.

Examples

1. $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ on $\mathbb{F}^N$ are equivalent by (1.1).
2. for $x = (\lambda_n)_{n=1}^\infty \in c_{00}(\mathbb{F})$ we have

   $$\|x\|_\infty \leq \sum_{n=1}^\infty |\lambda_n| =: \|x\|_1.$$

   On the other hand, let

   $$x_n := (\underbrace{1, \ldots, 1}_{n\text{-times}}, 0, \ldots).$$

Then

$$\|x_n\|_\infty = 1, \quad \text{but} \quad \|x_n\|_1 = n \qquad (n \in \mathbb{N}),$$

so that $\|\cdot\|_\infty \not\sim \|\cdot\|_1$.

3. $\|\cdot\|_\infty$ and $\|\cdot\|_N$ are not equivalent on $\mathcal{C}^N([0,1])$. (Why?)

The next theorem shows that there is only one equivalence class of norms on a finite-dimensional vector space.

Theorem 1.2.2 *Let E be a finite-dimensional linear space. Then all norms on E are equivalent.*

Proof Let $x_1, \dots, x_N \in E$ be a basis for E. For $x = \lambda_1 x_1 + \cdots + \lambda_N x_N$, let

$$|||x||| := \max\{|\lambda_1|, \dots, |\lambda_N|\}.$$

It is sufficient to show that $|||\cdot||| \sim \|\cdot\|$ for every other norm on E.

Let $x \in E$. Then we have:

$$\begin{aligned} \|x\| &\leq |\lambda_1|\|x_1\| + \cdots + |\lambda_N|\|x_N\| \\ &\leq |||x|||\underbrace{(\|x_1\| + \cdots + \|x_N\|)}_{=:C_1}. \end{aligned}$$

It remains to be shown that there is $C_2 \geq 0$ with $|||x||| \leq C_2\|x\|$ for all $x \in E$.

Assume otherwise. Then there is a sequence $(x^{(n)})_{n=1}^\infty$ in E with $|||x^{(n)}||| > n\|x^{(n)}\|$. Let

$$y^{(n)} := \frac{x^{(n)}}{|||x^{(n)}|||} \qquad (n \in \mathbb{N}).$$

For each $n \in \mathbb{N}$, there are unique $\lambda_1^{(n)}, \dots, \lambda_N^{(n)} \in \mathbb{F}$ with $y^{(n)} = \sum_{j=1}^N \lambda_j^{(n)} x_j$. It follows that

$$\left\| \left(\lambda_1^{(n)}, \dots, \lambda_N^{(n)} \right) \right\|_\infty = |||y^{(n)}||| = 1 \qquad (n \in \mathbb{N}).$$

By the Heine–Borel theorem, the sequence $\left(\left(\lambda_1^{(n)}, \dots, \lambda_N^{(n)} \right) \right)_{n=1}^\infty$ has a convergent subsequence $\left(\left(\lambda_1^{(n_k)}, \dots, \lambda_N^{(n_k)} \right) \right)_{k=1}^\infty$. Let $\mu_j := \lim_{k\to\infty} \lambda_j^{(n_k)}$ for $j = 1, \dots, n$, and define $y := \sum_{j=1}^N \mu_j x_j$. It follows that

$$|||y||| = \|(\mu_1, \dots, \mu_N)\|_\infty = 1,$$

so that, in particular, $y \neq 0$, and

$$\|y^{(n_k)} - y\| \leq C_1|||y^{(n_k)} - y||| = C_1 \left\| \left(\lambda_1^{(n^k)} - \mu_1, \dots, \lambda_N^{(n_k)} - \mu_N \right) \right\|_\infty \to 0.$$

On the other hand, the choice of $(x^{(n)})_{n=1}^\infty$ implies $n\|y^{(n)}\| < 1$ so that $\lim_{n\to\infty} \|y^{(n)}\| = 0$. But this means that $y = 0$, which is impossible. $\square$

Theorem 1.2.2 does not mean that finite-dimensional normed spaces are uninteresting: It says nothing about the constant C showing up in Definition 1.2.1. To find optimal values for C for concrete norms can be quite challenging. We won't pursue this, however.

Corollary 1.2.3 *Every finite-dimensional normed space is a Banach space.*

Exercise 1.12 Prove Corollary 1.2.3.

Corollary 1.2.4 *Every finite-dimensional subspace of a normed space is closed.*

As an immediate consequence, each finite-dimensional normed space can be identified with $\mathbb{F}^N$, so that theorems for $\mathbb{F}^N$ carry over to arbitrary finite-dimensional normed spaces. In particular, a finite-dimensional space has the Bolzano–Weierstraß property: Each bounded sequence has a convergent subsequence. As we shall now see, this property even characterizes the finite-dimensional normed spaces.

Lemma 1.2.5 (Riesz' lemma) *Let E be a normed space, and let F be a closed, proper, i.e. $F \neq E$, subspace of E. Then, for each $\theta \in (0,1)$, there is $x_\theta \in E$ with $\|x_\theta\| = 1$, and $\|x - x_\theta\| \geq \theta$ for all $x \in F$.*

Proof Let $x \in E \setminus F$, and let $\delta := \inf\{\|x - y\| : y \in F\}$. Then there is a sequence $(x_n)_{n=1}^\infty$ in F with $\lim_{n\to\infty} \|x - x_n\| = \delta$. If $\delta = 0$, the closedness of F implies $x \in F$, which is a contradiction. Hence, $\delta > 0$ must hold. Since $\theta \in (0,1)$, we have $\delta < \frac{\delta}{\theta}$. Choose $y \in F$ with $0 < \|x - y\| < \frac{\delta}{\theta}$. Let

$$x_\theta := \frac{y - x}{\|y - x\|},$$

so that trivially $\|x_\theta\| = 1$. For any $z \in F$, we then have:

$$\begin{aligned}
\|z - x_\theta\| &= \left\| z - \frac{y - x}{\|y - x\|} \right\| \\
&= \left\| z - \frac{y}{\|y - x\|} + \frac{x}{\|y - x\|} \right\| \\
&= \frac{1}{\|x - y\|} \underbrace{\|(\underbrace{\|x - y\|z + y}_{\in F}) - x\|}_{\geq \delta} \\
&\geq \frac{\theta}{\delta}\delta \\
&= \theta.
\end{aligned}$$

This completes the proof. $\square$

Theorem 1.2.6 *For a normed space E, the following are equivalent:*

(i) *Every bounded sequence in E has a convergent subsequence.*

(ii) $\dim E < \infty$.

Proof (ii) $\Longrightarrow$ (i) is elementary.

(i) $\Longrightarrow$ (ii): Suppose that $\dim E = \infty$. Choose $x_1 \in E$ with $\|x_1\| = 1$. Suppose that $x_1, \ldots, x_n$ have already been chosen such that

$$\|x_j\| = 1 \qquad (j = 1, \ldots, n)$$

and

$$\|x_j - x_k\| \geq \frac{1}{2} \qquad (j, k = 1, \ldots, n \; j \neq k).$$

Let $F := \operatorname{lin}\{x_1, \ldots, x_n\}$. Since $\dim F < \infty$, F is a proper and automatically closed subspace of E. By Riesz' lemma, there is $x_{n+1} \in E$ with

$$\|x_{n+1}\| = 1 \qquad \text{and} \qquad \|x - x_{n+1}\| \geq \frac{1}{2} \quad (x \in F).$$

Inductively, we thus obtain a sequence $(x_n)_{n=1}^{\infty}$ of unit vectors such that

$$\|x_n - x_m\| \geq \frac{1}{2} \qquad (n \neq m). \tag{1.5}$$

By (1.5), $(x_n)_{n=1}^{\infty}$ has no Cauchy subsequence. $\square$

Theorem 1.2.6 is the first example for the many subtle and often surprising links between algebra and analysis that surface in this course: A purely algebraic property — a linear space has finite dimension — turns out to be equivalent to the purely analytic Bolzano–Weierstraß property.

1.3 Linear operators

One of the major topics in linear algebra is the study of linear maps between finite-dimensional linear spaces. A considerable part of this course will be devoted to the study of linear maps between (possibly, but not necessarily) infinite-dimensional spaces.

Definition 1.3.1 Let E and F be linear spaces. A map $T \colon E \to F$ is called *linear* if

$$T(\lambda x + \mu y) = \lambda T x + \mu T y \qquad (x, y \in E, \; \lambda, \mu \in \mathbb{F}).$$

Linear maps are also called *linear operators*. A linear operator from E to $\mathbb{F}$ is called a *linear functional*.

Examples 1. Let $E := \mathbb{F}^N$, let $F := \mathbb{F}^M$, and let A be an $M \times N$-matrix. Then

$$T_A \colon E \to F, \quad x \mapsto Ax$$

is a linear operator.

2. Let $\varnothing \neq \Omega \subset \mathbb{R}^N$ be open, and let $\mathcal{C}^M(\Omega)$ denote the linear space of all functions $f\colon \Omega \to \mathbb{F}$ for which all partial derivatives of order at most M exist and are continuous. For each multiindex $\alpha \in \mathbb{N}_0^N$ with $|\alpha| := \alpha_1 + \cdots + \alpha_N \leq M$, let $f_\alpha \in \mathcal{C}(\Omega)$. Then

$$D \colon \mathcal{C}^M(\Omega) \to \mathcal{C}(\Omega), \quad f \mapsto \sum_{|\alpha| \leq M} f_\alpha \frac{\partial^\alpha f}{\partial x^\alpha}$$

is linear. Operators of this type are called *linear (partial) differential operators*.

3. Let $k : [0,1] \times [0,1] \to \mathbb{F}$ be continuous. For $f \in \mathcal{C}([0,1])$, define $Tf : [0,1] \to \mathbb{F}$ through

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

We claim that $Tf \in \mathcal{C}([0,1])$. Fix $x_0 \in [0,1]$, and let $\epsilon > 0$. Since $[0,1] \times [0,1]$ is compact, k is uniformly continuous. Hence, there is $\delta > 0$ such that

$$|k(x,y) - k(x',y')| < \frac{\epsilon}{\|f\|_\infty + 1}$$

for all $(x,y),(x',y') \in [0,1] \times [0,1]$ with $\|(x,y) - (x',y')\|_2 < \delta$. Let $x \in [0,1]$ such that $|x - x_0| < \delta$. It follows that

$$\|(x,y) - (x_0,y)\|_2 = |x - x_0| < \delta \qquad (y \in [0,1]).$$

We thus obtain:

$$\begin{aligned} |(Tf)(x) - (Tf)(x_0)| &= \left| \int_0^1 f(y)[k(x,y) - k(x_0,y)]\,dy \right| \\ &\leq \int_0^1 \underbrace{|f(y)|}_{\leq \|f\|_\infty} \underbrace{[k(x,y) - k(x_0,y)]}_{< \frac{\epsilon}{\|f\|_\infty + 1}}\,dy \\ &\leq \epsilon. \end{aligned}$$

Hence, Tf is continuous.

It is immediately checked that

$$T \colon \mathcal{C}([0,1]) \to \mathcal{C}([0,1]), \quad f \mapsto Tf$$

is a linear operator, the *Fredholm operator with kernel k*. Fredholm operators are part of the larger class of *linear integral operators*.

The first examples suggests that we may view linear operators as generalizations of matrices.

Exercise 1.13 Let E be a linear space with Hamel basis s, let F be another linear space, and let $(y_\alpha)_{s\in S}$ be an arbitrary family of elements of F. Show that there is a unique linear operator $T\colon E \to F$ such that $Ts = y_s$ for all $s \in S$.

There is virtually nothing of substance that can be said on linear operators between arbitrary linear spaces. We have to confine ourselves to the setting of normed spaces — preferably Banach spaces — and continuous linear operators.

Theorem 1.3.2 *Let E and F be normed spaces. Then the following are equivalent for a linear operator $T\colon E \to F$:*

(i) *T is continuous at 0.*

(ii) *T is continuous.*

(iii) *There is $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in E$.*

(iv) $\sup\{\|Tx\| : x \in E,\ \|x\| \leq 1\} < \infty$.

Operators satisfying these equivalent conditions are called bounded.

Proof (i) $\Longrightarrow$ (ii): Let $x \in E$, and let $(x_n)_{n=1}^\infty$ be a sequence in E such that $x_n \to x$. Then

$$\|Tx_n - Tx\| = \|T(\underbrace{x_n - x}_{\to 0})\| \to 0$$

holds, which proves (ii).

(ii) $\Longrightarrow$ (iii): Assume that (ii) holds, but that (iii) is false. Then there is a sequence $(x_n)_{n=1}^\infty$ in E such that $\|Tx_n\| > n\|x_n\|$ for all $n \in \mathbb{N}$. Let

$$y_n := \frac{x_n}{\|Tx_n\|} \qquad (n \in \mathbb{N}).$$

Since $1 > n\|y_n\|$ for all $n \in \mathbb{N}$, it follows that $y_n \to 0$. On the other hand, we have $\|Tx_n\| = 1$, which is impossible if T is continuous at 0.

(iii) $\Longrightarrow$ (iv): Clearly, (iii) implies

$$\sup\{\|Tx\| : x \in E,\ \|x\| \leq 1\} \leq C.$$

(iv) $\Longrightarrow$ (i): Assume that T is not continuous at 0. Then there is a sequence $(x_n)_{n=1}^\infty$ in E such that $x_n \to 0$, but $\delta := \inf \|Tx_n\| > 0$. Let

$$y_n := \frac{x_n}{\|x_n\|} \qquad (n \in \mathbb{N}),$$

so that $\|y_n\| = 1$ for all $n \in \mathbb{N}$. On the other hand, we have

$$\|Ty_n\| = \frac{1}{\|x_n\|}\|Tx_n\| > \frac{\delta}{\|x_n\|} \to \infty,$$

which contradicts (iv). $\square$

Exercise 1.14 Show that the following are equivalent for a normed space E:

(a) $\dim E = \infty$.

(b) For each normed space $F \neq \{0\}$, there is an unbounded linear operator $T \colon E \to F$.

(c) There is an unbounded linear functional on E.

Exercise 1.15 Let E be a normed space. Show that the following are equivalent for a linear functional $\phi \colon E \to \mathbb{F}$:

(a) $\phi \notin E^*$.

(b) $\phi(\{x \in E : \|x\| \leq 1\}) = \mathbb{F}$.

(c) $\ker \phi = \phi^{-1}(\{0\})$ is dense in E.

Exercise 1.16 Let E and F be Banach spaces, and let $T \in \mathcal{B}(E, F)$ be such that there is $C \geq 0$ with

$$\|x\| \leq C\|Tx\| \qquad (x \in E).$$

Show that T is injective and has closed range.

Examples 1. Let E and F be normed spaces, and let $T : E \to F$ be linear. Suppose that $\dim E < \infty$. Define

$$|||x||| := \max\{\|x\|, \|Tx\|\} \qquad (x \in E).$$

Then $|||\cdot|||$ is a norm on E. By Theorem 1.2.2, $|||\cdot|||$ and $\|\cdot\|$ are equivalent, so that there is $C \geq 0$ with

$$\|Tx\| \leq |||x||| \leq C\|x\| \qquad (x \in E).$$

Hence, T is bounded.

2. Let

$$T \colon \mathcal{C}^1([0,1]) \to \mathcal{C}([0,1]), \quad f \mapsto f',$$

and let both $\mathcal{C}^1([0,1])$ and $\mathcal{C}([0,1])$ be equipped with $\|\cdot\|_\infty$. For $n \in \mathbb{N}$, define

$$f_n(x) := x^n \qquad (x \in [0,1]),$$

so that $\|f_n\|_\infty = 1$ for $n \in \mathbb{N}$. However, since

$$f_n'(x) = \frac{nx^{n-1}}{n} \qquad (n \in \mathbb{N},\, x \in [0,1]),$$

we have $\|Tf_n\|_\infty = n$, so that T is not bounded. If, however, $\mathcal{C}^1([0,1])$ is equipped with the $\mathcal{C}^1$-norm $\|\cdot\|_1$, T becomes bounded:

$$\|Tf\|_\infty = \|f'\|_\infty \leq \|f\|_\infty + \|f'\|_\infty = \|f\|_1 \qquad (f \in \mathcal{C}^1([0,1])).$$

3. Let $k\colon [0,1] \times [0,1] \to \mathbb{F}$ be continuous, and let $T\colon \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ be given by

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

For each $f \in \mathcal{C}([0,1])$, we have:

$$\|Tf\|_\infty \leq \sup_{x\in[0,1]} \int_0^1 |f(y)||k(x,y)|\,dy \|f\|_\infty \|k\|_\infty.$$

Hence, T is bounded.

4. Let $(\Omega, \mathfrak{S}, \mu)$ be a σ-finite measure space, let $p \in [1,\infty]$, and let $\phi \in L^\infty(\Omega, \mathfrak{S}, \mu)$. Define $M_\phi\colon L^p(\Omega, \mathfrak{S}, \mu) \to L^p(\Omega, \mathfrak{S}, \mu)$ through

$$M_\phi f := \phi f \qquad (f \in L^p(\Omega, \mathfrak{S}, \mu)).$$

It is easy to see that

$$\|M_\phi f\|_\infty \leq \|\phi\|_\infty \|f\|_p \qquad (f \in L^p(\Omega, \mathfrak{S}, \mu)).$$

The first of these examples shows that the requirement of boundedness is vacuous for any operator between finite-dimensional spaces.

Given two normed spaces, we shall now see that the collection of all bounded linear operators between them is again a normed space in a natural manner:

Definition 1.3.3 Let E and F be normed spaces.

(a) The set of all bounded linear operators from E to F is denoted by $\mathcal{B}(E,F)$. If $E = F$, let $\mathcal{B}(E,F) =: \mathcal{B}(E)$; if $F = \mathbb{F}$, let $\mathcal{B}(E,F) =: E^*$.

(b) For $T \in \mathcal{B}(E,F)$, the *operator norm* of T is defined as

$$\|T\| := \sup\{\|Tx\| : x \in E,\, \|x\| \leq 1\}.$$

Proposition 1.3.4 *Let E and F be normed space. Then:*

(i) *$\mathcal{B}(E, F)$ equipped with the operator norm is a normed space.*

(ii) *For $T \in \mathcal{B}(E, F)$, $\|T\|$ is the smallest number $C \geq 0$ such that*

$$\|Tx\| \leq C\|x\| \qquad (x \in E). \tag{1.6}$$

(iii) *If G is another normed space, then for $T \in \mathcal{B}(E, F)$ and $S \in \mathcal{B}(F, G)$, we have $ST \in \mathcal{B}(E, G)$ such that*

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

Proof (i): It is straightforward to see that $\mathcal{B}(E, F)$ is a linear space. We have to check the norm axioms:

(a) Let $T \in \mathcal{B}(E, F)$ be such that $\|T\| = 0$. Let $x \in E \setminus \{0\}$. Then we have

$$\frac{1}{\|x\|}\|Tx\| = \left\| T\left(\frac{x}{\|x\|}\right)\right\| \leq \|T\| = 0,$$

so that $Tx = 0$. Since x was arbitrary, this means $T = 0$.

(b) It is routine to see that

$$\|\lambda T\| = |\lambda|\|T\| \qquad (\lambda \in \mathbb{F},\, T \in \mathcal{B}(E, F)).$$

(c) Let $x \in E$ be with $\|x\| \leq 1$. Then we have

$$\|Sx + Tx\| \leq \|Sx\| + \|Tx\| \leq \|S\| + \|T\| \qquad (S, T \in \mathcal{B}(E, F))$$

and thus

$$\|S + T\| \leq \|S\| + \|T\| \qquad (S, T \in \mathcal{B}(E, F)).$$

(ii): Let $x \in E \setminus \{0\}$. Since

$$\frac{1}{\|x\|}\|Tx\| = \left\| T\left(\frac{x}{\|x\|}\right)\right\| \leq \|T\|,$$

it follows that $\|Tx\| \leq \|T\|\|x\|$. Let $C \geq 0$ be any other number such that (1.6) holds. Then

$$\sup\{\|Tx\| : x \in E,\, \|x\| \leq 1\} \leq C.$$

(iii): Let $x \in E$. Then we have

$$\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|.$$

From (ii), it follows that $\|ST\| \leq \|S\|\|T\|$. $\square$

Exercise 1.17 Let $p \in [1,\infty)$, and let $L, R \colon \ell^p \to \ell^p$ be defined through

$$L(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$$
$$\text{and} \quad R(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots) \qquad ((x_1, x_2, x_3, \dots) \in \ell^p).$$

Show that $L, R \in \mathcal{B}(\ell^p)$ and calculate $\|L\|$ and $\|R\|$.

Exercise 1.18 Let $\mathbb{F}^N$ and $\mathbb{F}^M$ be equipped with $\|\cdot\|_\infty$, and let $A = [a_{j,k}]_{\substack{j=1,\dots,M \\ k=1,\dots,N}}$ be an $M \times N$-matrix over $\mathbb{F}$. Show that

$$\|T_A\| = \max_{j=1,\dots,M} \sum_{k=1}^{N} |a_{j,k}|$$

Theorem 1.3.5 *Let E be a normed space, and let F be a Banach space. Then $\mathcal{B}(E,F)$ is a Banach space.*

Proof Let $(T_n)_{n=1}^\infty$ be a Cauchy sequence in $\mathcal{B}(E,F)$. Let $x \in E$. Then

$$\|T_n x - T_m x\| \le \|T_n - T_m\| \|x\| \qquad (n, m \in \mathbb{N}),$$

so that $(T_n x)_{n=1}^\infty$ is a Cauchy sequence — and thus convergent — in F. Define

$$T \colon E \to F, \quad x \mapsto \lim_{n\to\infty} T_n x.$$

Clearly, T is linear.

We claim that $T \in \mathcal{B}(E,F)$ and that $\|T_n - T\| \to 0$. Let $x \in E$ with $\|x\| \le 1$, and let $\epsilon > 0$. There is $N \in \mathbb{N}$ — independent of x — such that

$$\|T_n x - T_m x\| \le \|T_n - T_m\| < \epsilon \qquad (n, m \ge N).$$

For $n \ge N$, this entails that

$$\|T_n x - Tx\| = \lim_{m\to\infty} \|T_n x - T_m x\| \le \limsup_{m\to\infty} \|T_n - T_m\| \le \epsilon. \tag{1.7}$$

In particular, we have $\|Tx\| \le \|T_N\| + \epsilon$, so that $T \in \mathcal{B}(E,F)$. Taking the supremum over $x \in E$ with $\|x\| \le 1$ in (1.7), we see that $\|T_n - T\| \le \epsilon$ for $n \ge N$. $\square$

Corollary 1.3.6 *For every normed space E, its* dual space *E^* is a Banach space.*

As turns out, arbitrary bounded linear operators between Banach spaces are still very general objects. To obtain stronger results, we have to look at a smaller class of operators:

Definition 1.3.7 Let E and F be normed space. A linear operator $T \colon E \to F$ is called *compact* if $T(\{x \in E : \|x\| \le 1\})$ is relatively compact in F. The set of all compact operators from E to F is denoted by $\mathcal{K}(E,F)$.

Proposition 1.3.8 *Let E and F be normed spaces. Then:*

(i) *$\mathcal{K}(E,F)$ is a subspace of $\mathcal{B}(E,F)$.*

(ii) *$T \in \mathcal{B}(E,F)$ is compact if and only if, for each bounded sequence $(x_n)_{n=1}^{\infty}$ in E, the sequence $(Tx_n)_{n=1}^{\infty}$ has a convergent subsequence.*

Proof (i): The set $T(\{x \in E : \|x\| \leq 1\})$ is relatively compact and thus bounded in F. This proves $\mathcal{K}(E,F) \subset \mathcal{B}(E,F)$. From (ii), it follows easily, that $\mathcal{K}(E,F)$ is a sub*space* of $\mathcal{B}(E,F)$.

(ii): We have:

$$\begin{aligned} T \text{ is compact} \quad &\iff \quad T(\{x \in E : \|x\| \leq 1\}) \text{ is relatively compact} \\ &\iff \quad rT(\{x \in E : \|x\| \leq 1\}) \text{ is relatively compact for each } r > 0 \\ &\iff \quad T(\{x \in E : \|x\| \leq r\}) \text{ is relatively compact for each } r > 0. \end{aligned}$$

This implies (ii). $\square$

Exercise 1.19 Let E, F, and G, be normed linear spaces, and let $T \in \mathcal{B}(E,F)$ and $S \in \mathcal{B}(F,G)$. Show that $ST \in \mathcal{K}(E,G)$ if T or S is compact.

Exercise 1.20 Let E be a linear space. A linear operator $P \colon E \to E$ is called a *projection* if $P^2 = P$.

(i) Show that a linear map $P \colon E \to E$ is a projection if and only if its restriction to PE is the identity.

(ii) Let E be normed, and let $P \in \mathcal{B}(E)$ be a projection. Show that P has closed range.

(iii) Let E be normed. Show that a projection $P \in \mathcal{B}(E)$ is compact if and only if it has finite rank.

Exercise 1.21 Is one of the operators L and R from Exercise 1.17 compact?

Examples 1. Let $\dim E = \infty$. Then $\mathrm{id}_E \colon E \to E$ is not compact.

2. Let $T \in \mathcal{B}(E,F)$ have finite rank, i.e. $\dim TE < \infty$. Then T is compact. To see this, let $(x_n)_{n=1}^{\infty}$ be a bounded sequence in E. Then $(Tx_n)_{n=1}^{\infty}$ is a bounded sequence in TE and thus has a convergent subsequence by Theorem 1.2.2.

3. Let $k \colon [0,1] \times [0,1] \to \mathbb{F}$ be continuous, and let $T \in \mathcal{B}(\mathcal{C}([0,1]))$ be the corresponding Fredholm operator. Let $(f_n)_{n=1}^{\infty}$ be a bounded sequence, and let $C := \sup_{n \in \mathbb{N}} \|f_n\|_\infty$. Let $\epsilon > 0$, and choose $\delta > 0$ such that

$$|k(x,y) - k(x',y')| < \frac{\epsilon}{C+1}$$

whenever $\|(x,y)-(x',y')\|_2 < \delta$. Let $x, x' \in [0,1]$ with $|x-x'| < \delta$. It follows that

$$|(Tf_n)(x) - (Tf_n)(x')| \leq \int_0^1 \underbrace{|f_n(y)|}_{\leq C} \underbrace{|k(x,y) - k(x',y')|}_{< \frac{\epsilon}{C+1}} \, dy < \epsilon.$$

Hence, $(Tf_n)_{n=1}^\infty$ is bounded an equicontinuous. By the Arzelà–Ascoli theorem, $(Tf_n)_{n=1}^\infty$ thus has a uniformly convergent subsequence.

Theorem 1.3.9 *Let E be a normed space, and let F be a Banach space. Then $\mathcal{K}(E,F)$ is a closed, linear subspace of F.*

Proof Let $(T_n)_{n=1}^\infty$ be a sequence in $\mathcal{K}(E,F)$, and let $T \in \mathcal{B}(E,F)$ be such that $\|T_n - T\| \to 0$. Assume that $T \notin \mathcal{K}(E,F)$. Then there is a bounded sequence $(x_n)_{n=1}^\infty$ in E such that $(Tx_n)_{n=1}^\infty$ has no convergent, i.e. Cauchy, subsequence. Passing to a subsequence, we may thus suppose that

$$\delta := \inf_{n \neq m} \|Tx_n - Tx_m\| > 0. \tag{1.8}$$

Let $C := \sup_{n \in \mathbb{N}} \|x_n\|$, and choose $N \in \mathbb{N}$ so large that $\|T - T_N\| < \frac{\delta}{3C+1}$. For $n, m \in \mathbb{N}$, we then have:

$$\begin{aligned} \|Tx_n - Tx_m\| &\leq \underbrace{\|Tx_n - T_N x_n\|}_{\leq \|T - T_N\| \|x_n\|} + \|T_N x_n - T_N x_m\| + \underbrace{\|T_N x_m - Tx_m\|}_{\leq \|T_N - T\| \|x_m\|} \\ &< \frac{2}{3}\delta + \|T_N x_n - T_N x_m\|. \end{aligned}$$

Since $T_N \in \mathcal{K}(E,F)$, the sequence $(T_N x_n)_{n=1}^\infty$ has a Cauchy subsequence. In particular, there are $n, m \in \mathbb{N}$, $n \neq m$, such that $\|T_N x_n - T_N x_m\| < \frac{\delta}{3}$. It follows that $\|Tx_n - Tx_m\| < \delta$ contradicting (1.8). $\square$

1.4 The dual space of a normed space

We now focus on a particular space of bounded linear operators:

Definition 1.4.1 For a normed space E, the Banach space $E^* (= \mathcal{B}(E, \mathbb{F}))$ is called the *dual space* or *dual* of E.

We want to give concrete descriptions of some dual spaces.

Definition 1.4.2 Let E and F be normed spaces.

(a) An *isomorphism* of E and F is a linear map $T \in \mathcal{B}(E,F)$ such that $S \in \mathcal{B}(F,E)$ exists with $ST = \mathrm{id}_E$ and $TS = \mathrm{id}_F$. If there is an isomorphism between E and F, we call E and F *isomorphic* (in symbols: $E \cong F$).

(b) A isometry from E to F is a linear map $T\colon E \to F$ such that

$$\|Tx\| = \|x\| \qquad (x \in E).$$

If there is an isomorphism of E and F which is also an isometry, then E and F are called *isometrically isomorphic* (in symbols: $E = F$).

Exercise 1.22 Let E and F be normed spaces, and let $T\colon E \to F$ be an isometry.

(i) Show that T is injective.

(ii) Suppose that T is surjective. Show that there is an isometry $S \in \mathcal{B}(F, E)$ exists with $ST = \mathrm{id}_E$ and $TS = \mathrm{id}_F$.

Exercise 1.23 Let $c_{00} := c_{00}(\mathbb{F})$ be equipped with the following norms:

$$\|x\|_\infty := \sup_{n \in \mathbb{N}} |x(n)| \quad \text{and} \quad \|x\|_1 := \sum_{n=1}^\infty |x(n)| \qquad (x \in c_{00}).$$

Show that the identity map $\mathrm{id}\colon (c_{00}, \|\cdot\|_1) \to (c_{00}, \|\cdot\|_\infty)$ is a continuous bijection, but not an isomorphism.

Examples 1. Let E and F be normed spaces with $\dim E = \dim F < \infty$. Then $E \cong F$.

2. Let $p \in (1, \infty)$, and let $q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Define $T\colon \ell^q \to (\ell^p)^*$ by letting for $x = (x_n)_{n=1}^\infty \in \ell^q$ and $Y = (y_n)_{n=1}^\infty \in \ell^p$:

$$(Tx)(y) := \sum_{n=1}^\infty x_n y_n.$$

Since

$$|(Tx)(y)| \leq \lim_{N\to\infty} \sum_{n=1}^N |x_n y_n| \leq \lim_{N\to\infty} \left(\sum_{n=1}^N |x_n|^q\right)^{\frac{1}{q}} \left(\sum_{n=1}^N |y_n|^p\right)^{\frac{1}{p}} = \|x\|_q \|y\|_p,$$

the map T is well defined and satisfies $\|T\| \leq 1$. Let $x \in \ell^q$ with $\|x\|_q = 1$, and define $(y_n)_{n=1}^\infty$ as follows:

$$y_n := \begin{cases} \frac{|x_n|^q}{x_n}, & \text{if } x_n \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

We have

$$\sum_{n=1}^\infty |y_n|^p = \sum_{\substack{n=1 \\ x_n \neq 0}}^\infty \frac{|x_n|^{pq}}{|x_n|^p} = \sum_{n=1}^\infty |x_n|^{pq-p} = \|x\|_q^q = 1,$$

so that $\|y\|_p = 1$. It follows that

$$\|Tx\| \geq |(Tx)(y)| = \sum_{n=1}^\infty |x_n|^q = \|x\|_q^q = 1.$$

For arbitrary $x \in \ell^q \setminus \{0\}$, we thus have:

$$\|Tx\| = \|x\| \left\| T\left(\frac{x}{\|x\|}\right) \right\| \geq \|x\|.$$

Hence, T is an isometry.

We claim that T is surjective. Let $\phi \in (\ell^p)^*$. For each $n \in \mathbb{N}$, define $e^{(n)} \in \ell^p$ through

$$e_m^{(n)} := \begin{cases} 1, & n = m, \\ 0, & n \neq m \end{cases}$$

Define $x = (x_n)_{n=1}^\infty$ by letting $x_n := \phi(e^{(n)})$ for $n \in \mathbb{N}$. *If* $x \in \ell^q$, then $Tx = \phi$ (Why?). For $N \in \mathbb{N}$, define $x^{(N)} \in \ell^q$ through

$$x_n^{(N)} := \begin{cases} x_n, & n \leq N, \\ 0, & n > N. \end{cases}$$

For any $y = (y_n)_{n=1}^\infty \in \ell^p$ define $z = (z_n)_{n=1}^\infty$ through

$$z_n := \begin{cases} \frac{|y_n x_n|}{x_n}, & \text{if } n \leq N \text{ and } x_n \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

If is clear that $z \in \ell^p$ with $\|z\|_p \leq \|y\|_p$. We now have:

$$\begin{aligned} |(Tx^{(N)})(y)| &\leq \sum_{n=1}^N |x_n y_n| \\ &= \sum_{n=1}^\infty z_n x_n \\ &= \sum_{n=1}^\infty z_n \phi(e^{(n)}) \\ &= \phi(z) \\ &= |\phi(z)| \\ &\leq \|\phi\| \|z\|_p \\ &\leq \|\phi\| \|y\|_p. \end{aligned}$$

It follows that

$$\|x\|_q^q = \lim_{N \to \infty} \sum_{n=1}^N |x_n|^q = \lim_{N \to \infty} \|x^{(N)}\|_q^q = \lim_{N \to \infty} \|Tx^{(N)}\|_q^q \leq \|\phi\|^q,$$

so that $x \in \ell^q$ with $\|x\|_q \leq \|\phi\|$.

All in all, we have $(\ell^p)^* = \ell^q$.

3. Similarly (Exercise 1.24 below), we have $(\ell^1)^* = \ell^\infty$ and $(c_0)^* = \ell^1$.
4. If $(\Omega, \mathfrak{S}, \mu)$ is any measure space, and if $p, q \in (1, \infty)$ are such that $\frac{1}{p} + \frac{1}{q} = 1$, we have an isometric isomorphism $T \colon L^q(\Omega, \mathfrak{S}, \mu) \to L^p(\Omega, \mathfrak{S}, \mu)^*$ given by
$$(Tf)(g) := \int_\Omega f(\omega) g(\omega) \, d\mu(\omega),$$
so that $L^p(\Omega, \mathfrak{S}, \mu)^* = L^q(\Omega, \mathfrak{S}, \mu)$.
5. For σ-finite $(\Omega, \mathfrak{S}, \mu)$, we also have $L^1(\Omega, \mathfrak{S}, \mu)^* = L^\infty(\Omega, \mathfrak{S}, \mu)$.

Exercise 1.24 Show that $(\ell^1)^* = \ell^\infty$ and $(c_0)^* = \ell^1$.

We now have concrete descriptions of E^* for a few normed spaces E. But what can we say about E^* for a general normed space E? So far, the only linear function on E of which we positively know that it's in E^* is the zero-functional. Are there any others? The answer to this question is "yes", as we shall see in the next chapter.

Chapter 2

The fundamental principles of functional analysis

In this chapter, we prove four fundamental theorems of functional analysis:

- the Hahn–Banach theorem;
- Baire's theorem;
- the open mapping theorem;
- the closed graph theorem.

We illustrate the power of each theorem with application, e.g. to complex variables and initial value problems.

2.1 The Hahn–Banach theorem

Given an arbitrary normed space E with dual E^*, we cannot tell right now if E^* contains any non-zero elements. This will change in this section: We will prove the Hahn–Banach theorem, which implies that there are enough functionals in E^* to separate the points of E.

Roughly speaking, the Hahn–Banach theorem asserts that, if we have a linear functional on a subspace of a linear space whose growth can somehow be controlled, then this functional can be extended to the whole space such that the growth remains under control.

Definition 2.1.1 Let E be a linear space. A map $p : E \to \mathbb{R}$ is called a *sublinear functional* if

$$p(x+y) \leq p(x) + p(y) \qquad (x, y \in E)$$

and

$$p(\lambda x) = \lambda p(x) \qquad (x \in E,\, \lambda \in \mathbb{R},\, \lambda \geq 0).$$

Lemma 2.1.2 *Let E be a linear space over $\mathbb{R}$, let F be a subspace of E, let $x_0 \in E \setminus F$, and let $p\colon E \to \mathbb{R}$ be a sublinear functional. Suppose that $\phi\colon F \to \mathbb{R}$ is a linear functional such that*

$$\phi(x) \leq p(x) \qquad (x \in F).$$

Then there is a linear functional $\tilde{\phi}\colon F + \mathbb{R}x_0 \to \mathbb{R}$ which extends ϕ and satisfies

$$\tilde{\phi}(x) \leq p(x) \qquad (x \in F + \mathbb{R}x_0).$$

Proof We need to show that there is $\alpha \in \mathbb{R}$ such that

$$\phi(x) + t\alpha \leq p(x + tx_0) \qquad (x \in F,\, t \in \mathbb{R}).$$

If this is done, we can define $\tilde{\phi}$ by letting $\tilde{\phi}(x + tx_0) := \phi(x) + t\alpha$ for all $x \in F$ and $t \in \mathbb{R}$.

For any $x, y \in F$, we have

$$\phi(x) + \phi(y) = \phi(x + y) \leq p(x - x_0 + x_0 + y) \leq p(x - x_0) + p(x_0 + y)$$

and thus

$$\phi(x) - p(x - x_0) \leq p(y + x_0) - \phi(y) \qquad (x, y \in F). \tag{2.1}$$

Let $\alpha := \sup\{\phi(x) - p(x - x_0) : x \in F\}$. It follows from (2.1) that

$$\phi(x) - p(x - x_0) \leq \alpha \leq p(y + x_0) - \phi(y) \qquad (x, y \in F)$$

and thus

$$\phi(x) - \alpha \leq p(x - x_0) \qquad (x \in F) \tag{2.2}$$

and

$$\phi(x) + \alpha \leq p(x + x_0) \qquad (x \in F). \tag{2.3}$$

Let $t \in \mathbb{R}$, and let $x \in F$. If $t > 0$, we obtain from (2.3):

$$\phi(x) + t\alpha = t\left(\phi\left(\frac{1}{t}x\right) + \alpha\right) \leq t\,p\left(\frac{1}{t}x + x_0\right) = p(x + tx_0).$$

For $t < 0$, inequality (2.2) yields:

$$\phi(x) + t\alpha = -t\left(\phi\left(\frac{1}{-t}x\right) - \alpha\right) \leq -t\,p\left(\frac{1}{-t}x - x_0\right) = p(x + tx_0).$$

This completes the proof. $\square$

Theorem 2.1.3 (Hahn–Banach theorem) *Let E be a linear space over $\mathbb{R}$, let F be a subspace of E, and let $p\colon E \to \mathbb{R}$ be a sublinear functional. Suppose that $\phi\colon F \to \mathbb{R}$ is a linear functional such that*

$$\phi(x) \leq p(x) \qquad (x \in F).$$

Then there is a linear functional $\tilde{\phi}\colon E \to \mathbb{R}$ which extends ϕ and satisfies

$$\tilde{\phi}(x) \leq p(x) \qquad (x \in E).$$

Proof Let $\mathcal{S}$ be the collection of all pairs (X, ψ) with the following properties:

- X is a subspace of E with $F \subset X$;
- $\psi\colon X \to \mathbb{R}$ is linear with $\psi|_F = \phi$;
- $\psi(x) \leq p(x) \qquad (x \in X)$.

Clearly, $(F, \phi) \in \mathcal{S}$.

Define an order $\prec$ on $\mathcal{S}$:

$$(X_1, \psi_1) \prec (X_2, \phi_2) \quad :\Longleftrightarrow \quad X_1 \subset X_2 \text{ and } \psi_2|_{X_1} = \psi_1.$$

Let $\mathcal{T} \subset \mathcal{S}$ be totally ordered. Define

$$\tilde{X} := \bigcup \{X : (X, \psi) \in \mathcal{T}\}.$$

Then $\tilde{X}$ is a subspace of E with $F \subset \tilde{X}$. Define $\tilde{\psi}: \tilde{X} \to \mathbb{R}$ by letting $\tilde{\psi}(x) := \psi(x)$ if $x \in X$ for $(X, \psi) \in \mathcal{T}$. Then $\tilde{\psi}\colon \tilde{X} \to \mathbb{R}$ is well-defined, linear, extends ϕ, and satisfies

$$\tilde{\psi}(x) \leq p(x) \qquad (x \in \tilde{X}).$$

It follows that $(\tilde{X}, \tilde{\psi}) \in \mathcal{S}$ is an upper bound for $\mathcal{T}$. Hence, by Zorn's lemma, $\mathcal{S}$ has a maximal element $(X_{\max}, \psi_{\max})$. We claim that $X_{\max} = E$. Otherwise, there is $x_0 \in E \setminus X_{\max}$. By Lemma 2.1.2, there is a linear extension $\tilde{\psi}_{\max}\colon X_{\max} + \mathbb{R}x_0 \to \mathbb{R}$ of $\psi_{\max}$ such that

$$\tilde{\psi}_{\max}(x) \leq p(x) \qquad (x \in X_{\max} + \mathbb{R}x_0).$$

This contradicts the maximality of $(X_{\max}, \psi_{\max})$. $\square$

Exercise 2.1 A *Banach limit* on $\ell^\infty(\mathbb{R})$ is a linear functional $\mathrm{Lim}\colon \ell^\infty(\mathbb{R}) \to \mathbb{R}$ such that for any sequence $(x_n)_{n=1}^\infty$(we write $\mathrm{Lim}_{n\to\infty} x_n$ instead of $\mathrm{Lim}((x_n)_{n=1}^\infty)$):

(a) $\liminf_{n\to\infty} x_n \leq \mathrm{Lim}_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$;

(b) $\mathrm{Lim}_{n\to\infty} x_{n+k} = \mathrm{Lim}_{n\to\infty} x_n$ for all $k \in \mathbb{N}$.

What is $\operatorname{Lim}_{n\to\infty} x_n$ if $(x_n)_{n=1}^\infty$ converges?

(i) Show that $\|\operatorname{Lim}\| = 1$.

(ii) Show that Banach limits do exist. (*Hint*: Let F be the subspace of $\ell^\infty(\mathbb{R})$ consisting of those sequences $(x_n)_{n=1}^\infty$ for which $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n x_n$ exists; define Lim on F to be that limit, and apply the Hahn–Banach theorem.)

Exercise 2.2 Let E be a $\mathbb{C}$-linear space.

(i) Let $\phi\colon E \to \mathbb{C}$ be $\mathbb{C}$-linear. Show that

$$\phi(x) = \operatorname{Re}\phi(x) - i\operatorname{Re}\phi(ix) \qquad (x \in E).$$

(ii) Let $\psi\colon E \to \mathbb{R}$ be $\mathbb{R}$-linear. Show that $\phi\colon E \to \mathbb{C}$ defined through

$$\phi(x) := \psi(x) - i\psi(ix) \qquad (x \in E)$$

is $\mathbb{C}$-linear.

We will rarely apply the Hahn–Banach theorem directly, but rather one of the following corollaries:

Corollary 2.1.4 *Let E be a linear space, let F be subspace of E, and let $p\colon E \to [0,\infty)$ be a seminorm. Suppose that $\phi\colon F \to \mathbb{F}$ is linear such that*

$$|\phi(x)| \le p(x) \qquad (x \in F).$$

Then ϕ has a linear extension $\tilde{\phi}\colon E \to \mathbb{F}$ such that

$$|\tilde{\phi}(x)| \le p(x) \qquad (x \in E).$$

Proof Suppose first that $\mathbb{F} = \mathbb{R}$. By Theorem 2.1.3, we have a linear extension $\tilde{\phi}\colon E \to \mathbb{R}$ such that

$$\tilde{\phi}(x) \le p(x) \qquad (x \in E).$$

If $\tilde{\phi}(x) \le 0$, then

$$-\tilde{\phi}(x) = \tilde{\phi}(-x) \le p(-x) = p(x),$$

so that

$$-p(x) \le \tilde{\phi}(x) \le p(x) \qquad (x \in E).$$

Now consider the case where $\mathbb{F} = \mathbb{C}$. Define $\psi\colon F \to \mathbb{R}$ through

$$\psi(x) := \operatorname{Re}\phi(x) \qquad (x \in F).$$

By Exercise 2.2(i), we then have

$$\phi(x) = \psi(x) - i\phi(ix) \qquad (x \in F).$$

By the first case, ψ has an $\mathbb{R}$-linear extension $\tilde{\psi}\colon E \to \mathbb{R}$ with

$$|\tilde{\psi}(x)| \leq p(x) \qquad (x \in E).$$

Define $\tilde{\phi}\colon E \to \mathbb{R}$ by letting

$$\tilde{\phi}(x) := \tilde{\psi}(x) - i\tilde{\phi}(ix) \qquad (x \in E).$$

By Exercise 2.2(ii), $\tilde{\phi}$ is $\mathbb{C}$-linear and clearly extends ϕ. Let $x \in E$, and choose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ such that $\tilde{\phi}(x) = \lambda|\tilde{\phi}(x)|$. We obtain:

$$|\tilde{\phi}(x)| = \bar{\lambda}\tilde{\phi}(x) = \tilde{\phi}(\bar{\lambda}x) = \tilde{\psi}(\bar{\lambda}x) \leq p(\bar{\lambda}x) = p(x).$$

This proves the claim in the complex case. $\square$

Corollary 2.1.5 *Let E be a normed space, let F be subspace, and let $\phi \in F^*$. Then ϕ has an extension $\tilde{\phi} \in E^*$ with $\|\tilde{\phi}\| = \|\phi\|$.*

Proof Apply Corollary 2.1.4 with $p(x) := \|\phi\|\|x\|$ for $x \in E$. $\square$

Exercise 2.3 Let $S \neq \varnothing$ be a set, let E be a normed space, and let F be a subspace of E. Show that any operator $T \in \mathcal{B}(F, \ell^\infty(S))$ has an extension $\tilde{T} \in \mathcal{B}(E, \ell^\infty(S))$ with $\|\tilde{T}\| = \|T\|$.

Corollary 2.1.6 *Let E be a normed space, let F be a closed subspace of E, and let $x_0 \in E \setminus F$. Then there is $\phi \in E^*$ with $\|\phi\| = 1$, $\phi|_F = 0$, and $\phi(x_0) = \operatorname{dist}(x_0, F)$.*

Proof Define

$$p\colon E \to [0, \infty), \quad x \mapsto \operatorname{dist}(x, F)(:= \inf\{x - y\| : y \in F\})$$

and

$$\phi\colon F + \mathbb{F}x_0, \quad x + \lambda x_0 \mapsto \lambda \operatorname{dist}(x_0, F).$$

It follows that

$$|\phi(x)| \leq \operatorname{dist}(x, F) \qquad (x \in F + \mathbb{F}x_0).$$

By Corollary 2.1.4, ϕ has a linear extension $\tilde{\phi}$ to all of E with

$$|\tilde{\phi}(x)| \leq \operatorname{dist}(x, F) \leq \|x\| \qquad (x \in E),$$

so that, in particular, $\|\tilde{\phi}\| \leq 1$. Let $\epsilon > 0$. Then there is $y \in F$ with $\|x_0 - y\| \leq \operatorname{dist}(x_0, F) + \epsilon$. Let $z := \frac{x_0 - y}{\|x_0 - y\|}$, so that $\|z\| = 1$. It follows that

$$|\tilde{\phi}(z)| = \frac{\phi(x_0 - y)}{\|x_0 - y\|} = \frac{\phi(x_0)}{\|x_0 - y\|} \geq \frac{\operatorname{dist}(x_0, F)}{\operatorname{dist}(x_0, F) + \epsilon}.$$

Since $\epsilon > 0$ was arbitrary, this means $\|\tilde{\phi}\| \geq 1$. $\square$

Corollary 2.1.6 can be used to prove approximation theorems: Let x_0 be an element of a normed space E and assume that it is not in the closure of a subspace F. Then Corollary 2.1.6 yields $\phi \in E^*$ which vanishes on F but not in x_0. Since for many spaces E we have a concrete description of E^*, this may then be used to arrive at a contradiction, so that x_0 must lie in the closure of F.

Exercise 2.4 A metric space is called *separable* if it has a countable dense subset (any subset of a separable metric space is again separable).

(i) Show that c_0 as well as ℓ^p for $p \in [1, \infty)$ are separable.

(ii) Show that ℓ^∞ is not separable (*Hint*: Show that the subset of ℓ^∞ consisting of those $f \in \ell^\infty$ such that $f(\mathbb{N}) \subset \{0, 1\}$ is uncountable and conclude that, for this reason, it cannot be separable.)

Exercise 2.5 Let E be a normed space such that E^* is separable. Show that E must be separable as well. Proceed as follows:

- Let $\{\phi_n : n \in \mathbb{N}\}$ be a dense subset of $\{\phi \in E^* : \|\phi\| = 1\}$. For each $n \in \mathbb{N}$ pick $x_n \in E$ with $\|x_n\| \leq 1$ and $|\phi_n(x_n)| \geq \frac{1}{2}$.
- Use the Hahn–Banach theorem to show that the linear span of $\{x_n : n \in \mathbb{N}\}$ is dense in E.

Does, conversely, the separability of E imply that E^* is separable? Can $(\ell^\infty)^* = \ell^1$ hold?

Corollary 2.1.7 *Let E be a normed space, and let $x \in E$. Then there is $\phi \in E^*$ with $\|\phi\| = 1$ and $\phi(x) = \|x\|$.*

Proof Apply Corollary 2.1.6 with $F = \{0\}$. $\square$

2.2 Applications of the Hahn–Banach theorem

We now present several application of the Hahn–Banach theorem.

2.2.1 The bidual of a normed space

Definition 2.2.1 The *bidual* E^{**} of a normed space E is defined as $(E^*)^*$.

There is a canonical map $J \colon E \to E^{**}$ defined by

$$(Jx)(\phi) := \phi(x) \qquad (x \in E,\ \phi \in E^*).$$

By Corollary 2.1.7 we have:

$$\|Jx\| = \sup\{|\phi(x)| : \phi \in E^*,\ \|\phi\| \leq 1\} = \|x\| \qquad (x \in E).$$

Hence, J is an isometry and we may identify JE with E. In particular, every normed space “is” the subspace of a Banach space.

2.2.2 Transpose operators

Definition 2.2.2 Let E and F be normed spaces, and let $T \in \mathcal{B}(E, F)$. The *transpose* $T^* \colon F^* \to E^*$ of T is defined through

$$(T^*\phi)(x) := \phi(Tx) \qquad (x \in E,\ \phi \in F^*)$$

Exercise 2.6 Let E, F and G be normed spaces, let $S, T \in \mathcal{B}(E, F)$, $R \in \mathcal{B}(F, G)$, and $\lambda, \mu \in \mathbb{F}$. Show that:

(i) $(\lambda T + \mu S)^* = \lambda S^* + \mu T^*$;

(ii) $(RT)^* = T^* R^*$.

Theorem 2.2.3 *Let E and F be normed spaces, and let $T \in \mathcal{B}(E, F)$. Then $T^* \in \mathcal{B}(F^*, E^*)$ with $\|T^*\| = \|T\|$.*

Proof For $x \in E$ and $\phi \in F^*$, we have

$$|(T^*\phi)(x)| = |\phi(Tx)| \leq \|\phi\|\|T\|\|x\|,$$

therefore

$$\|T^*\phi\| \leq \|T\|\|\phi\|,$$

and eventually $\|T^*\| \leq \|T\|$

Consider $T^{**} \colon E^{**} \to F^{**}$. We have $\|T^{**}\| \leq \|T^*\|$. On the other hand, we have for $x \in E$ and $\phi \in F^*$:

$$(T^{**}Jx)(\phi) = (Jx)(T^*\phi) = (T^*\phi)(x) = \phi(Tx) = (JTx)(\phi).$$

Hence, T^{**} extends T, so that, in particular, $\|T^{**}\| \geq \|T\|$. $\square$

The following theorem has nothing to do with the Hahn–Banach theorem, but we will need it later and it fits into the discussion of transpose operators.

Theorem 2.2.4 (Schauder's theorem) *Let E and F be normed spaces, and let $T \in \mathcal{K}(E, F)$. Then $T^* \in \mathcal{K}(F^*, E^*)$.*

Proof Let $(\phi_n)_{n=1}^\infty$ be a sequence in F^* bounded by $C \geq 0$. Let

$$K := \overline{T(\{x \in E : \|x\| \leq 1\})}.$$

Then K is a compact metric space. For $y, z \in K$ and $n \in \mathbb{N}$, we have

$$|\phi_n(y) - \phi_n(z)| \leq \|y - z\|.$$

Consequently, the sequence $(\phi_n|_K)_{n=1}^\infty$ in $\mathcal{C}(K)$ is bounded and equicontinuous. By the Arzelà–Ascoli theorem, there is a subsequence $(\phi_{n_k}|_K)_{k=1}^\infty$ converging uniformly to some function in $\mathcal{C}(K)$. In particular, $(\phi_{n_k}|_K)_{k=1}^\infty$ is a Cauchy sequence with respect to the uniform norm. For $k, l \in \mathbb{N}$, however, we have:

$$\begin{aligned} \|\phi_{n_k}|_K - \phi_{n_l}|_K\|_\infty &= \sup_{y \in K} |\phi_{n_k}(y) - \phi_{n_l}(y)| \\ &\geq \sup\{|\phi_{n_k}(Tx) - \phi_{n_l}(Tx)| : x \in E, \, \|x\| \leq 1\} \\ &= \|T^*\phi_{n_k} - T^*\phi_{n_l}\|. \end{aligned}$$

Hence, $(T^*\phi_{n_k})_{k=1}^\infty$ is a Cauchy sequence in E^* and thus convergent. $\square$

2.2.3 Quotient spaces and duals

Definition 2.2.5 Let E be a normed space, and let F be a closed subspace of F. The *quotient norm* on E/F is defined as

$$\|x + F\| := \inf\{\|x - y\| : y \in F\} \qquad (x \in E).$$

Theorem 2.2.6 *Let E be a normed space, and let F be a closed subspace. Then E/F equipped with the quotient norm is normed space. If E is a Banach space, then so is E/F.*

Proof It is routine to verify that the quotient norm is indeed a norm.

Let $(x_n)_{n=1}^\infty$ be a sequence in E such that $\sum_{n=1}^\infty \|x_n + F\| < \infty$. For each $n \in \mathbb{N}$, choose $y_n \in F$ such that $\|x_n - y_n\| < \frac{1}{2^n}$. It follows that $\sum_{n=1}^\infty \|x_n - y_n\| < \infty$. Since E is a Banach space, $\sum_{n=1}^\infty (x_n - y_n)$ converges in E to x E. It is clear that $x + F = \sum_{n=1}^\infty (x_n + F)$. $\square$

Exercise 2.7 Let E be a normed space, and let F be a closed subspace. Let G be another normed space, and let $T \in \mathcal{B}(E, G)$ vanish on F. Show that

$$\tilde{T}(x + F) := Tx \qquad (x \in E)$$

defines $\tilde{T} \in \mathcal{B}(E/F, G)$ with $\|\tilde{T}\| = \|T\|$.

Definition 2.2.7 Let E be a normed space. For any subset S of E, we define

$$S^\perp := \{\phi \in E^* : \phi|_S = 0\}.$$

Theorem 2.2.8 *Let E be a normed space, and let F be a closed subspace of E. Then*

$$T : E^* \to F^*, \quad \phi \mapsto \phi|_F$$

induces an isometric isomorphism of $E^*/F^\perp$ and F^*.

Proof Clearly, $\|T\| \leq 1$ and $\ker T = F^\perp$, so that T, by Exercise 2.7, T induces an injective map $\tilde{T} \colon E^*/F^\perp \to F^*$ with $\|\tilde{T}\| \leq 1$. We claim that this map is surjective and an isometry.

Let $\psi \in F^*$. By Corollary 2.1.5, there is $\phi \in E^*$ with $\|\phi\| = \|\psi\|$ extending ψ. We thus have $T\phi = \psi$ and

$$\|\phi + F^\perp\| \geq \|\tilde{T}(\phi + F^\perp)\| = \|\psi\| = \|\phi\| \geq \|\phi + F^\perp\|,$$

which completes the proof. $\square$

Theorem 2.2.9 *Let E be a normed space, and let F be a normed subspace. Then $T : F^\perp \mapsto (E/F)^*$ defined by*

$$(T\phi)(x + F) := \phi(x) \qquad (\phi \in F^\perp,\, x \in E)$$

is an isometric isomorphism of $F^\perp$ and $(E/F)^*$.

Proof It is routinely checked that T is well-defined and linear.

Let $\pi : E \to E/F$ be the quotient map. For any $\psi \in (E/F)^*$, the functional $\psi \circ \pi$ belongs to $F^\perp$ such that $T(\pi \circ \psi) = \psi$. Hence, T is surjective. From Exercise 2.7, it follows that T is an isometry. $\square$

2.2.4 The dual space of $\mathcal{C}([0,1])$

Definition 2.2.10 A function $\alpha \colon [0,1] \to \mathbb{F}$ is said to be of *bounded variation* if

$$\|\alpha\|_{BV} := \sup \left\{ \sum_{j=1}^{n} |\alpha(x_j) - \alpha(x_{j-1})| : n \in \mathbb{N},\, 0 = x_0 < x_1 < \cdots < x_n = 1 \right\} < \infty.$$

We define:

$$BV([0,1]) := \{\alpha \colon [0,1] \to \mathbb{F} : \alpha \text{ is of bounded variation}\}$$

The following are easily checked:

- $BV([0,1])$ is a linear space;
- $\|\cdot\|_{BV}$ is a seminorm on $BV([0,1])$;
- $\|\alpha\|_{BV} = 0 \iff \alpha$ is constant.

We let

$$BV_0([0,1]) := \{\alpha \in BV([0,1]) : \alpha(0) = 0\}.$$

Then $\|\cdot\|_{BV}$ is a normed space.

Theorem 2.2.11 *The linear map $T\colon BV_0([0,1]) \to \mathcal{C}([0,1])^*$ defined by*

$$(T\alpha)(f) := \int_0^1 f(x)\, d\alpha(x)$$

is an isometric isomorphism.

Proof It is obvious that $\|T\alpha\| \leq \|\alpha\|_{BV}$ for all $\alpha \in BV_0([0,1])$.

Conversely, let $\phi \in \mathcal{C}([0,1])^*$. Let $u_0 :\equiv 0$, and for any $x \in (0,1]$, define $u_x\colon [0,1] \to \mathbb{R}$ by letting

$$u_x(t) := \begin{cases} 1, & 0 \leq t \leq x, \\ 0, & x < t \leq 1. \end{cases}$$

By Corollary 2.1.5, there is an extension $\tilde{\phi} \in \ell^\infty([0,1])^*$ of ϕ with $\|\tilde{\phi}\| = \|\phi\|$. Define

$$\alpha\colon [0,1] \to \mathbb{F}, \quad x \mapsto \tilde{\phi}(u_x).$$

We claim that $\alpha \in BV_0([0,1])$. Let $0 = x_0 < x_1 < \cdots < x_n = 1$, and define, for $j = 1, \ldots, n$:

$$\sigma_j := \begin{cases} \frac{|\alpha(x_j)-\alpha(x_{j-1})|}{\alpha(x_j)-\alpha(x_{j-1})}, & \text{if } \alpha(x_j) \neq \alpha(x_{j-1}), \\ 0, & \text{otherwise.} \end{cases}$$

We then have:

$$\begin{aligned} \sum_{j=1}^n |\alpha(x_j) - \alpha(x_{j-1})| &= \sum_{j=1}^n \sigma_j(\alpha(x_j) - \alpha(x_{j-1})) \\ &= \sum_{j=1}^n \sigma_j(\tilde{\phi}(u_{x_j}) - \tilde{\phi}(u_{x_{j-1}})) \\ &= \underbrace{\tilde{\phi}\left(\sum_{j=1}^n \sigma_j(u_{x_j} - u_{x_{j-1}})\right)}_{\|\cdot\|_\infty \leq 1} \\ &\leq \|\tilde{\phi}\| \\ &= \|\phi\|. \end{aligned}$$

Hence, α is a bounded variation such that $\|\alpha\|_{BV} \leq \|\phi\|$.

Next, we claim that $T\alpha = \phi$ (this establishes at once that T is a surjective isometry and thus an isometric isomorphism). For any $f \in \mathcal{C}([0,1])$ and any partition $\mathcal{P} = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$ of $[0,1]$, let

$$\delta(\mathcal{P}) = \sup_{j=1,\ldots,n} |x_j - x_{j-1}|$$

and

$$S(f,\mathcal{P}) := \sum_{j=1}^{n} f(x_j)(\alpha(x_j) - \alpha(x_{j-1})).$$

From the properties of the Riemann–Stieltjes integral, we know that

$$\lim_{\delta(\mathcal{P})\to 0} S(f,\mathcal{P}) = \int_0^1 f(x)\,d\alpha(x).$$

Define

$$f_{\mathcal{P}} := \sum_{j=1}^{n} f(x_j)(u_{x_j} - u_{x_{j-1}})$$

From the uniform continuity of f, we infer that $\lim_{\delta(\mathcal{P})\to 0} \|f_{\mathcal{P}} - f\|_\infty = 0$. We thus have:

$$\begin{aligned}
\phi(f) &= \lim_{\delta(\mathcal{P})\to 0} \tilde{\phi}(f_{\mathcal{P}}) \\
&= \lim_{\delta(\mathcal{P})\to 0} \sum_{j=1}^{n} f(x_j)(\tilde{\phi}(u_{x_j}) - \tilde{\phi}(u_{x_{j-1}})) \\
&= \lim_{\delta(\mathcal{P})\to 0} \sum_{j=1}^{n} f(x_j)(\alpha(x_j) - \alpha(x_{j-1})) \\
&= \lim_{\delta(\mathcal{P})\to 0} S(f,\mathcal{P}) \\
&= \int_0^1 f(x)\,d\alpha(x).
\end{aligned}$$

This establishes the claim and thus completes the proof. □

This result is only a rather special case of Riesz' representation theorem (Theorem B.3.8).

Exercise 2.8 Let $C \geq 0$, $c_1, c_2, \ldots \in \mathbb{F}$, and $f_1, f_2, \ldots \in \mathcal{C}([0,1])$ be given. Show that the following are equivalent:

(a) There is $\alpha \in BV[0,1]$ with $\|\alpha\|_{BV} \leq C$ such that

$$c_n = \int_0^1 f_n(t)\,d\alpha(t) \qquad (n \in \mathbb{N}).$$

(b) For all $n \in \mathbb{N}$, and for all $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, we have

$$\left|\sum_{k=1}^{n} \lambda_k c_k\right| \leq C \left\|\sum_{k=1}^{n} \lambda_k f_k\right\|_\infty .$$

2.2.5 Runge's approximation theorem

We now use Corollary 2.1.6 to prove an approximation theorem from complex analysis:

Theorem 2.2.12 (Runge's approximation theorem) *Let $K \subset \mathbb{C}$ be compact, and let $E \subset \mathbb{C}_\infty \setminus K$ be such that E has at least one point in common with each component of $\mathbb{C}_\infty \setminus K$. Let $U \subset \mathbb{C}$ be an open set containing K, and let $f : U \to \mathbb{C}$ be holomorphic. Then, for each $\epsilon > 0$, there is a rational function with poles in E such that*

$$\sup_{z \in K} |f(z) - r(z)| < \epsilon.$$

Proof Note that $f|_K \in \mathcal{C}(K)$. Let

$$\mathcal{R}_E(K) := \overline{\{r|_K : r \text{ is a rational function with poles in } E\}}.$$

We need to show that $f|_K \in \mathcal{R}_E(K)$.

Assume that this is not true. By Corollary 2.1.6, there is $\phi \in \mathcal{C}(K)^*$ with

$$\phi|_{\mathcal{R}_E(K)} = 0 \qquad \text{and} \qquad \phi(f|_K) \neq 0.$$

By the Riesz' representation theorem (Theorem B.3.8), there is $\mu \in M(K)$ such that

$$\phi(g) = \int_K g(z)\, d\mu(z) \qquad (g \in \mathcal{C}(K)).$$

Define

$$\hat{\mu} \colon \mathbb{C} \setminus K \to \mathbb{C}, \quad w \mapsto \int_K \frac{d\mu(z)}{z - w}.$$

We claim that $\hat{\mu} \equiv 0$. Let V be a component of $\mathbb{C}_\infty \setminus K$, and let $p \in E \cap V$.

Case 1: $p \neq \infty$. Then choose $r > 0$ such that $B_r(p) \subset V$. For fixed $w \in B_r(p)$, we then have uniformly in $z \in K$:

$$\begin{aligned}
\frac{1}{z-w} &= \frac{1}{(z-p)-(w-p)} \\
&= \frac{1}{z-p}\frac{1}{1-\frac{w-p}{z-p}} \\
&= \frac{1}{z-p}\sum_{n=0}^{\infty}\left(\frac{w-p}{z-p}\right)^n \\
&= \sum_{n=0}^{\infty}\frac{(w-p)^n}{(z-p)^{n+1}}.
\end{aligned}$$

Hence, the function $\frac{1}{z-w}$ of z belongs to $\mathcal{R}_E(K)$. It follows that $\hat{\mu}(w) = 0$. Since $w \in B_r(p)$ was arbitrary, the identity theorem yields $\hat{\mu}|_V \equiv 0$.

Case 2: $p = \infty$. Choose $r > 0$ so large that $|z| < r$ for all $z \in K$, and let $w \in \mathbb{C}$ with $|w| > r$. Then we have uniformly in $z \in K$:

$$\frac{1}{z-w} = \frac{1}{w}\frac{1}{\frac{z}{w}-1} = -\sum_{n=0}^{\infty}\frac{z^n}{w^{n+1}}.$$

It follows again that $\hat{\mu}(w) = 0$ for $|w| > r$ and thus $\hat{\mu}|_V \equiv 0$.

Let Γ by a cycle in U whose winding number around each point in K is 1 and around each point in $\mathbb{C} \setminus U$ is zero. By the Cauchy integral formula, we have

$$f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f(w)}{z-w}\,dw \qquad (z \in K).$$

But this yields

$$\begin{aligned}
\int_K f(z)\,d\mu(z) &= \frac{1}{2\pi i}\int_K \left[\int_\Gamma \frac{f(w)}{z-w}\,dw\right] d\mu(z) \\
&= \frac{1}{2\pi i}\int_\Gamma \left[\int_K \frac{f(w)}{z-w}\,d\mu(z)\right] dw \\
&= \frac{1}{2\pi i}\int_\Gamma f(w)\left[\int_K \frac{1}{z-w}\,d\mu(z)\right] dw \\
&= \frac{1}{2\pi i}\int_\Gamma f(w)\hat{\mu}(w)dw \\
&= 0,
\end{aligned}$$

which contradicts the choice of ϕ. $\square$

The advantage of this functional analytic proof is its brevity and its elegance. The drawback is that it is not constructive: It depends on the Hahn–Banach theorem and therefore on Zorn's lemma.

2.3 Baire's theorem

Theorem 2.3.1 (Baire's theorem) *Let X be a complete metric space, and let $(U_n)_{n=1}^\infty$ be a sequence of dense open subsets of X. Then $\bigcap_{n=1}^\infty U_n$ is dense in X.*

Proof Assume that the theorem is wrong. Then there are $x_0 \in X$ and $\epsilon > 0$ such that

$$B_\epsilon(x_0) \subset X \setminus \bigcap_{n=1}^\infty U_n.$$

Let $V_0 := B_\epsilon(x_0)$. Since U_1 is dense in X, there is $x_1 \in U_1 \cap V_0$. Choose $r_1 \in (0,1)$ so small that

$$\overline{B_{r_1}(x_1)} \subset U_1 \cap V_0.$$

Let $V_1 := U_{r_1}(x_1)$. Suppose that open subsets $V_0, V_1, \ldots, V_n$ of X have already been constructed such that

- $\overline{V_{j+1}} \subset U_{j+1} \cap V_j$ for $j = 0, \ldots, n-1$, and
- $\operatorname{diam} \overline{V_j} \leq \frac{2}{j}$ for $j = 1, \ldots, n$.

Since U_{n+1} is dense in X, there is $x_{n+1} \in U_{n+1} \cap V_n$. Choose $r_{n+1} \in \left(0, \frac{1}{n+1}\right)$ so small that

$$\overline{B_{r_{n+1}}(x_{n+1})} \subset U_{n+1} \cap V_n,$$

and let $V_{n+1} := B_{r_{n+1}}(x_{n+1})$. Continue inductively.

Since $\operatorname{diam} \overline{V_n} \leq \frac{2}{n}$ for all $n \in \mathbb{N}$, and since X is complete, there is $x \in \bigcap_{n=1}^{n} \overline{V_n}$. By construction, however,

$$x \in \bigcap_{n=1}^{n} \overline{V_n} \subset \bigcap_{n=1}^{\infty} U_n \qquad \text{and} \qquad x \in \bigcup_{n=1}^{n} \overline{V_n} \subset V_0 \subset X \setminus \bigcap_{n=1}^{\infty} U_n,$$

which is impossible. $\square$

Corollary 2.3.2 *Let X be a complete metric space, and let $(F_n)_{n=1}^{\infty}$ be a sequence of closed subsets of X such that $\bigcup_{n=1}^{\infty} F_n$ has an interior point. Then at least one F_n has an interior point.*

Proof Let $U_n := X \setminus F_n$. $\square$

Example Let E be a Banach space with a countable Hamel basis. We claim that $\dim E < \infty$.

Assume that E has a Hamel basis $\{x_n : n \in \mathbb{N}\}$. For $n \in \mathbb{N}$, let

$$E_n := \operatorname{lin}\{x_1, \ldots, x_n\}.$$

Then E_n is a closed subspace of E. Since $E = \bigcup_{n=1}^{\infty} E_n$, Corollary 2.3.2 yields that there is $N \in \mathbb{N}$ such that E_N has interior points, i.e. there is $x_0 \in E_N$ and $\epsilon > 0$ such that $B_\epsilon(x_0) \subset E_N$. Let $x := x_0 + \frac{\epsilon}{2} \frac{x_{N+1}}{\|x_{N+1}\|}$. Then $x \in B_\epsilon(x_0)$, but $x \notin E_N$.

In particular, there is *no* norm on c_{00} turning it into a Banach space.

For our next application of Baire's theorem, we need the following approximation theorem:

Theorem 2.3.3 (Weierstraß' approximation theorem) *Let $a, b \in \mathbb{R}$, $a < b$, let $f \in \mathcal{C}([a,b])$, and let $\epsilon > 0$. Then there is a polynomial p such that $\|f - p\|_\infty < \epsilon$.*

Proof Without loss of generality, let $a = 0$, $b = 1$, and $\mathbb{F} = \mathbb{R}$.

For each $g \in \mathcal{C}([0,1])$, let

$$B_n(g; t) := \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} g\left(\frac{k}{n}\right) \qquad (t \in [0,1]).$$

be its *n-th Bernstein polynomial*. It is routinely checked that

$$\begin{aligned} B_n(1,t) &= \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} \\ &= (t + (1-t))^n \\ &= 1, \\ B_n(x,t) &= \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} \frac{k}{n} \\ &= \sum_{k=1}^{n} \binom{n-1}{k-1} t^k (1-t)^{n-k} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} t^{k+1} (1-t)^{n-(k+1)} \\ &= t \sum_{k=0}^{n-1} \binom{n-1}{k} t^k (1-t)^{(n-1)-k} \\ &= t(t + (1-t))^{n-1} \\ &= t, \end{aligned}$$

and

$$\begin{aligned} B_n(x^2,t) &= \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} \left(\frac{k}{n}\right)^2 \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} t^{k+1} (1-t)^{n-(k+1)} \frac{k+1}{n} \\ &= \frac{t}{n} + \sum_{k=0}^{n-1} \binom{n-1}{k} t^k (1-t)^{(n-1)-k} \frac{k}{n} t \\ &= \frac{t}{n} + \frac{n-1}{n} t^2 \\ &= \frac{t(1-t)}{n} + t^2. \end{aligned}$$

Since f is uniformly continuous, there is $\delta > 0$ such that $|f(s) - f(t)| < \epsilon$ for all $s, t \in [0,1]$ with $|s-t| < \sqrt{\delta}$. Let $C := \frac{2\|f\|_\infty}{\delta}$. We claim that

$$|f(s) - f(t)| \leq \epsilon + C(t-s)^2 \qquad (s, t \in [0,1]). \tag{2.4}$$

This is clear if $|s-t| < \sqrt{\delta}$; if $|s-t| \geq \sqrt{\delta}$, it follows from

$$\epsilon + C(t-s)^2 > \epsilon + 2\|f\|_\infty > |f(s)| + |f(t)| \geq |f(s) - f(t)|.$$

Fix $t \in [0,1]$, and let $f_t(s) := (t-s)^2$. Then (2.4) implies

$$-\epsilon - Cf_t \leq f - f(t) \leq \epsilon + Cf_t$$

and thus

$$-B_n(\epsilon + Cf_t, \cdot) = B_n(-\epsilon - Cf_t, \cdot) \leq B_n(f - f(t), \cdot) \leq B(\epsilon + Cf_t, \cdot).$$

Since $B_n(f - f(t), \cdot) = B_n(f, \cdot) - f(t)$, we obtain for $p_n := B_n(f, \cdot)$:

$$|p_n(s) - f(t)| \leq B_n(\epsilon + Cf_t, s) = \epsilon + Ct^2 - 2Cts + C\left(\frac{s(1-s)}{n} + s^2\right) \qquad (s \in [0,1]).$$

Letting $s = t$, this yields

$$|p_n(t) - f(t)| \leq \epsilon + C\frac{t(1-t)}{n} \leq \epsilon + \frac{C}{n}.$$

Since $\epsilon > 0$ and $t \in [0,1]$ are arbitrary, this yields $\|p_n - f\|_\infty \to 0$. $\square$

There are proofs of Theorem 2.3.3 that use the Hahn–Banach theorem, but they are in no way easier than the one given here.

Example We will now use Baire's theorem to prove that there is a continuous function on $[0,1]$ which is nowhere differentiable.

For $n \in \mathbb{N}$, let

$$F_n := \left\{ f \in \mathcal{C}([0,2]) : \text{there is } t \in [0,1] \text{ such that } \sup_{h \in (0,1)} \frac{|f(t+h) - f(t)|}{h} \leq n \right\}.$$

Obviously, if $f \in \mathcal{C}([0,2])$ is differentiable at some $t \in [0,1]$, then $\sup_{h \in (0,1)} \frac{|f(t+h)-f(t)|}{h} < \infty$, so that $f \in \bigcup_{n=1}^\infty F_n$.

Let $(f_k)_{k=1}^\infty$ be a sequence in F_n such that $\|f_k - f\|_\infty \to 0$ for some $f \in \mathcal{C}([0,2])$. For each $k \in \mathbb{N}$, there is $t_k \in [0,1]$ such that

$$\sup_{h \in (0,1)} \frac{|f_k(t_k + h) - f_k(t_k)|}{h} \leq n.$$

Suppose without loss of generality that $(t_k)_{k=1}^\infty$ converges to some $t \in [0,1]$. Fix $h \in (0,1)$ and $\epsilon > 0$, and choose $K \in \mathbb{N}$ so large that

$$\left.\begin{array}{c} |f(t+h) - f(t_k + h)| \\ \|f - f_k\|_\infty \\ |f(t_k) - f(t)| \end{array}\right\} < \frac{\epsilon}{4}h \qquad (k \geq K).$$

For $k \geq K$, this implies

$$\begin{aligned} & |f(t+h) - f(t)| \\ \leq \; & \underbrace{|f(t+h) - f(t_k+h)|}_{<\frac{\epsilon}{4}h} + \underbrace{|f(t_k+h) - f_k(t_k+h)|}_{<\frac{\epsilon}{4}h} + \underbrace{|f_k(t_k+h) - f_k(t_k)|}_{\leq nh} \\ & + \underbrace{|f(t_k) - f_k(t_k)|}_{<\frac{\epsilon}{4}h} + \underbrace{|f(t) - f(t_k)|}_{<\frac{\epsilon}{4}h} \\ \leq \; & nh + \epsilon h, \end{aligned}$$

so that $\frac{|f(t+h)-f(t)|}{h} \leq n+\epsilon$. Since h and ϵ were arbitrary, this means that $f \in F_n$. Hence, F_n is closed.

Assume that every $f \in \mathcal{C}([0,2])$ is differentiable at some point in $[0,1]$. Then $\mathcal{C}([0,2]) = \bigcup_{n=1}^\infty F_n$, so that, by Corollary 2.3.2, there are $N \in \mathbb{N}$, $f \in \mathcal{C}([0,2])$, and $\epsilon > 0$ such that $B_\epsilon(f) \subset F_N$. By Theorem 2.3.3, $B_\epsilon(f)$ contains at least one polynomial. Without loss of generality, we may thus suppose that $f \in \mathcal{C}^1([0,2])$.

For $m \in \mathbb{N}$ and $j = 0, \ldots, m$, let $t_j := \frac{2j}{m}$. Define $g_m \colon [0,2] \to \mathbb{F}$ by letting

$$g_m(t) := \begin{cases} \frac{\epsilon}{2} m(t - t_{j-1}), & t \in \left[t_{j-1}, t_{j-1} + \frac{1}{m}\right], \\ \frac{\epsilon}{2} m(t_{j-1} - t), & t \in \left[t_j - \frac{1}{m}, t_j\right]. \end{cases}$$

Then $g_m \in \mathcal{C}([0,2])$ with $\|g_m\| = \frac{\epsilon}{2}$, but

$$\sup_{h \in (0,1)} \frac{|g_m(t+h) - g_m(t)|}{h} \geq \frac{\epsilon}{2} m \tag{2.5}$$

holds for any $t \in [0,1]$. Since $f + g_m \in B_\epsilon(f) \subset F_N$, there is $t \in [0,1]$ such that

$$\sup_{h \in (0,1)} \frac{|(f+g_m)(t+h) - (f+g_m)(t)|}{h} \leq N.$$

This, however, yields

$$\begin{aligned} & \sup_{h \in (0,1)} \frac{|g_m(t+h) - g_m(t)|}{h} \\ \leq \;& \sup_{h \in (0,1)} \frac{|(f+g_m)(t+h) - (f+g_m)(t)|}{h} + \sup_{h \in (0,1)} \frac{|f(t+h) - f(t)|}{h} \\ = \;& N + \|f'\|_\infty, \end{aligned}$$

which contradicts (2.5) if we choose $m \in \mathbb{N}$ so large that $\frac{\epsilon}{2} m > N + \|f'\|_\infty$.

Exercise 2.9 Let $(f_k)_{k=1}^\infty$ be a sequence in $\mathcal{C}([0,1])$ which converges *pointwise* to a function $f \colon [0,1] \to \mathbb{F}$.

(i) For $\theta > 0$ and $n \in \mathbb{N}$ let

$$F_n := \{t \in [0,1] : |f_n(t) - f_k(t)| \leq \theta \text{ for all } k \geq n\}.$$

Show that F_n is closed, and that $[0,1] = \bigcup_{n=1}^\infty F_n$.

(ii) Let $\epsilon > 0$, and let I be a non-degenerate, closed subinterval of $[0,1]$. Show that there is a non-degenerate, closed subinterval J of I such that

$$|f(t) - f(s)| \leq \epsilon \qquad (t, s \in J).$$

(*Hint*: Apply (a) with $\theta := \frac{\epsilon}{3}$ and Baire's theorem.)

(iii) Let I be a non-degenerate, closed subinterval of $[0,1]$. Show that there is a decreasing sequence of non-degenerate, closed subintervals of I such that

- the length of I_n is at most $\frac{1}{n}$, and
- $|f(t) - f(s)| \leq \frac{1}{n}$ for all $s, t \in I_n$.

What can be said about f at all points in $\bigcap_{n=1}^{\infty} I_n$?

(iv) Conclude that the set of points in $[0,1]$ at which f is continuous is dense in $[0,1]$.

2.4 The uniform boundedness principle

Theorem 2.4.1 (uniform boundedness principle) *Let E be a Banach space, let $(F_\alpha)_\alpha$ be a family of normed spaces, and let $T_\alpha \in \mathcal{B}(E, F_\alpha)$ be such that*

$$\sup_\alpha \|T_\alpha x\| < \infty \qquad (x \in E). \tag{2.6}$$

Then $\sup_\alpha \|T_\alpha\| < \infty$ holds.

Proof For $n \in \mathbb{N}$, let

$$E_n := \left\{ x \in E : \sup_\alpha \|T_\alpha x\| \leq n \right\}.$$

Then (2.6) implies that $E = \bigcup_{n=1}^{\infty}$. Let $(x_k)_{k=1}^{\infty}$ be a sequence in E_n such that $x_k \to x \in E$. For any index α, we have

$$\|T_\alpha x\| = \lim_{k \to \infty} \|T_\alpha x_k\| \leq n.$$

It follows that $x \in E$, so that E_n is closed.

By Corollary 2.3.2 there are thus $N \in \mathbb{N}$, $x_0 \in E$, and $\epsilon > 0$ such that $\overline{U_\epsilon(x_0)} \subset E_N$. Let $x \in E$ be such that $\|x\| \leq 1$. It follows that $\epsilon x + x_0 \in E_N$. Hence, we have for all α:

$$\epsilon \|T_\alpha x\| = \|T_\alpha(\epsilon x)\| \leq \|T_\alpha(\epsilon x + x_0)\| + \|T_\alpha x_0\| \leq 2N,$$

and consequently

$$\|T_\alpha x\| \leq \frac{2N}{\epsilon}.$$

It follows that $\sup_\alpha \|T_\alpha\| \leq \frac{2N}{\epsilon}$. $\square$

Examples 1. Let E be a Banach space, let F be a normed space, and let $(T_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{B}(E, F)$ such that $\lim_{n \to \infty} T_n x$ exists in F for each $x \in E$. We claim that $T \colon E \to F$ defined through

$$Tx := \lim_{n \to \infty} T_n x \qquad (x \in E).$$

Since $(T_n x)_{n=1}^\infty$ is convergent and thus bounded for each $x \in E$, Theorem 2.4.1 implies that $C := \sup_{n\in\mathbb{N}} \|T_n\| < \infty$. Let $x \in E$. Then there is $N \in \mathbb{N}$ such that $\|T_N x - Tx\| \leq \|x\|$. This yields

$$\|Tx\| \leq \|T_N x\| + \|T_N x - Tx\| \leq C\|x\| + \|x\| = (C+1)\|x\|.$$

2. For each continuous function $f\colon [-\pi, \pi] \to \mathbb{R}$, its Fourier series is

$$\frac{a_0}{2} + \sum_{k=1}^\infty (a_k \cos(kx) + b_k \sin(kx)), \tag{2.7}$$

where

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx)\, dx \qquad (k \in \mathbb{N}_0)$$

and

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx)\, dx \quad (k \in \mathbb{N}).$$

We will show that there are $f \in \mathcal{C}([-\pi, \pi])$ for which (2.7) does not converge in every point $x \in [-\pi, \pi]$.

For $n \in \mathbb{N}$, let

$$D_n(x) := \frac{\sin\left(\left(n + \frac{1}{2}\right) x\right)}{2 \sin \frac{x}{2}} \qquad (x \in [-\pi, \pi]),$$

and define

$$s_n \colon \mathcal{C}([-\pi, \pi]) \to \mathbb{R}, \quad f \mapsto \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) D_n(x)\, dx.$$

It can be shown that, for any $f \in \mathcal{C}([-\pi, \pi])$ with Fourier series (2.7), we have

$$s_n(f) = \frac{a_0}{2} + \sum_{k=1}^n a_k,$$

i.e. $s_n(f)$ is the n-th partial sum of (2.7) at $x = 0$. It is easy to see that $(s_n)_{n=1}^\infty$ is a sequence in $\mathcal{C}([-\pi, \pi], \mathbb{R})^*$ such that

$$\|s_n\| \leq \frac{1}{\pi} \int_{-\pi}^{\pi} |D_n(x)|\, dx \qquad (n \in \mathbb{N}).$$

Let $\epsilon > 0$, and let $x_1, \ldots, x_m \in [-\pi, \pi]$ be the zeros of D_n in $[-\pi, \pi]$. Choose $\delta > 0$ such that

$$\frac{2\delta m \|D_n\|_\infty}{\pi} \leq \frac{\epsilon}{2}.$$

Define $f \in \mathcal{C}([-\pi,\pi])$ through

$$f(x) := \begin{cases} \frac{|D_n(x)|}{D_n(x)}, & x \notin \bigcup_{j=1}^m (x_j - \delta, x_j + \delta) =: I_\delta, \\ \text{linear in between.} \end{cases}$$

It is clear that $\|f\|_\infty \leq 1$. It follows that

$$\begin{aligned} \|s_n\| &\geq |s_n(f)| \\ &= \frac{1}{\pi}\left|\int_{[-\pi,\pi]\setminus I_\delta} |D_n(x)|\,dx + \int_{I_\delta} f(x)D_n(x)\,dx\right| \\ &\geq \frac{1}{\pi}\int_{[-\pi,\pi]\setminus I_\delta} |D_n(x)|\,dx - \underbrace{\frac{1}{\pi}\int_{I_\delta} |D_n(x)|\,dx}_{\leq\frac{\epsilon}{2}} \\ &\geq \frac{1}{\pi}\int_{[-\pi,\pi]\setminus I_\delta} |D_n(x)|\,dx - \frac{\epsilon}{2}, \end{aligned}$$

consequently

$$\begin{aligned} \frac{\epsilon}{2} + \|s_n\| &\geq \frac{1}{\pi}\int_{[-\pi,\pi]\setminus I_\delta} |D_n(x)|\,dx \\ &= \frac{1}{\pi}\int_{[-\pi,\pi]} |D_n(x)|\,dx - \underbrace{\frac{1}{\pi}\int_{I_\delta} |D_n(x)|\,dx}_{\leq\frac{\epsilon}{2}} \\ &\geq \frac{1}{\pi}\int_{[-\pi,\pi]} |D_n(x)|\,dx - \frac{\epsilon}{2}, \end{aligned}$$

and finally

$$\epsilon + \|s_n\| \geq \frac{1}{\pi}\int_{[-\pi,\pi]} |D_n(x)|\,dx.$$

Since $\epsilon > 0$, this yields

$$\|s_n\| = \frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(x)|\,dx = \frac{2}{\pi}\int_0^{\pi} |D_n(x)|\,dx.$$

Since

$$\begin{aligned} \int_0^\pi \frac{\left|\sin\left(\left(n+\frac{1}{2}\right)x\right)\right|}{\sin\frac{x}{2}}\,dx &\geq \int_0^\pi \frac{\left|\sin\left(\left(n+\frac{1}{2}\right)x\right)\right|}{\frac{x}{2}}\,dx \\ &= 2\int_0^{\left(n+\frac{1}{2}\right)\pi} \frac{|\sin y|}{y}\,dy \\ &\to \infty, \end{aligned}$$

it follows that $\|s_n\| \to \infty$. Hence, Theorem 2.4.1 implies that there is $f \in \mathcal{C}([-\pi,\pi])$ such that $\sup_{n\in\mathbb{N}} |s_n(f)| = \infty$. In particular, the Fourier series of f diverges at $x = 0$.

Exercise 2.10 Prove the *Banach–Steinhaus theorem*: Let E and F be Banach spaces, let $T \in \mathcal{B}(E,F)$, and let $(T_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{B}(E,F)$. Then the following are equivalent:

(a) $Tx = \lim_{n\to\infty} T_n x$ for each $x \in E$;

(b) $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$ and $Tx = \lim_{n\to\infty} T_n x$ for all x in some dense subspace of E.

Exercise 2.11 A family $(x_\alpha)_\alpha$ of elements in a normed space E is called *weakly bounded* if

$$\sup_\alpha |\phi(x_\alpha)| < \infty$$

for all $\phi \in E^*$. Show that a family $(x_\alpha)_\alpha$ in a normed space is bounded if and only if it is weakly bounded.

Exercise 2.12 Let E, F, and G be Banach spaces, and let $T\colon E \times F \to G$ be a bilinear map which is continuous in each variable. Show that there is $C \geq 0$ such that

$$\|T(x,y)\| \leq C\|x\|\|y\| \qquad (x \in E,\ y \in F).$$

2.5 The open mapping theorem

Definition 2.5.1 Let E and F be normed spaces. A linear map $T\colon E \to F$ is called *open* if TU is open in F for every open subset U of E.

Theorem 2.5.2 (open mapping theorem) *Let E and F be Banach spaces, and let $T \in \mathcal{B}(E,F)$ be surjective. Then T is open.*

Proof We claim that, for each $r > 0$, the zero vector 0 is an interior point of $\overline{TB_r(0)}$.

Since T is surjective, we have

$$F = \bigcup_{n=1}^{\infty} \overline{TB_{\frac{nr}{2}}(0)}.$$

By Corollary 2.3.2, there is $N \in \mathbb{N}$ such that

$$\overline{TB_{\frac{Nr}{2}}(0)} = N\,\overline{TB_{\frac{r}{2}}(0)}$$

has an interior point, say x_0. Hence, there is $\epsilon > 0$ such that $B_\epsilon(x_0) \subset \overline{TB_{\frac{r}{2}}(0)}$. For any $x \in B_\epsilon(0)$, we then have

$$x = x + x_0 - x_0 \in \overline{TB_{\frac{r}{2}}(0)} + \overline{TB_{\frac{r}{2}}(0)} \subset \overline{TB_r(0)}.$$

This proves the first claim.

Secondly, we claim that $\overline{TB_{\frac{r}{2}}(0)} \subset TB_r(0)$ for all $r > 0$.

Fix $y_1 \in \overline{TB_{\frac{r}{2}}(0)}$. Since 0 is an interior point of $\overline{TB_{\frac{r}{4}}(0)}$, it follows that

$$(y_1 - \overline{TB_{\frac{r}{4}}(0)}) \cap TB_{\frac{r}{2}}(0) \neq \varnothing.$$

Choose $x_1 \in B_{\frac{r}{2}}(0)$ such that $Tx_1 \in y_1 - \overline{TB_{\frac{r}{4}}(0)}$. Then choose $y_2 \in \overline{TB_{\frac{r}{4}}(0)}$ such that $Tx_1 = y_1 - y_2$. Continuing in this fashion, we obtain sequences $(x_n)_{n=1}^\infty$ in E and $(y_n)_{n=1}^\infty$ in F such that

$$\left\{ \begin{array}{c} x_n \in B_{\frac{r}{2^n}}(0), \\ y_n \in \overline{TB_{\frac{r}{2^n}}(0)}, \\ y_{n+1} = y_n - Tx_n. \end{array} \right\} \qquad (n \in \mathbb{N}).$$

Since $\sum_{n=1}^\infty \|x_n\| < \infty$, the series $\sum_{n=1}^\infty x_n$ converges to some $x \in E$ with $\|x\| < r$. Moreover, we have

$$\begin{aligned} Tx &= \sum_{n=1}^\infty Tx_n \\ &= \sum_{n=1}^\infty (y_n - y_{n+1}) \\ &= \lim_{N\to\infty} \sum_{n=1}^N (y_n - y_{n+1}) \\ &= \lim_{N\to\infty} (y_1 - y_{N+1}) \\ &= y_1, \end{aligned}$$

so that $y_1 \in TB_r(0)$.

Finally, we deduce that T is indeed open.

Let $U \subset E$ be open, and let $x \in U$. Choose $r > 0$ such that $B_r(x) \subset U$. Since 0 is an interior point of $TB_r(0)$, it follows that Tx is an interior point of $TB_r(x)$ and thus of TU. Since $x \in U$ was arbitrary, this means that TU is open. $\square$

Exercise 2.13 Let E and F be Banach spaces, and let $T \in \mathcal{B}(E, F)$ be such that $\dim F/TE < \infty$. Show that T has closed range. (*Hint*: Choose a finite-dimensional subspace G of F with $F = TE + G$ and $TE \cap G = \{0\}$, and consider

$$S \colon E \oplus G \to F, \quad (x, y) \mapsto Tx + y.$$

Apply the open mapping theorem.)

Exercise 2.14 Let E and F be normed spaces, and let $T \colon E \to F$ be an open linear map. Show that T is surjective. (*Warning* Trick question.)

Corollary 2.5.3 *Let E and F be Banach spaces, and let $T \in \mathcal{B}(E, F)$ be bijective. Then T is an isomorphism, i.e. $T^{-1} \in \mathcal{B}(F, E)$.*

Exercise 2.15 Let E and F be Banach spaces, and let $T \in \mathcal{B}(E, T)$ be surjective. Show that there is $C \geq 0$ such that, for each $y \in F$, there is $x \in E$ with $\|x\| \leq C\|y\|$ such that $Tx = y$.

Exercise 2.16 Let E be a Banach space. An operator $T \in \mathcal{B}(E)$ is called *quasi-nilpotent* if

$$\lim_{n\to\infty} \sqrt[n]{\|T^n\|} = 0.$$

Show that a quasi-nilpotent operator can never be surjective unless $E = \{0\}$. (*Hint*: Previous problem.)

Exercise 2.17 Let E and F be Banach spaces. Show that the following are equivalent for $T \in \mathcal{B}(E, F)$:

(a) T is injective and has closed range.

(b) There is $C \geq 0$ such that

$$\|x\| \leq C\|Tx\| \qquad (x \in E).$$

Examples 1. Let E be a Banach space. A closed subspace F is called *complemented* in E if there is another closed subspace G of E with $E = F + G$ and $F \cap G = \{0\}$. We claim that, if F is complemented, then the canonical projection $P\colon E \to F$ with $\ker P = G$ is continuous.

Clearly,

$$\tilde{E} = F \oplus G$$

with

$$\|(x, y)\| = \max\{\|x\|, \|y\|\| \qquad (x \in F,\, y \in G)$$

is a Banach space. Let $\pi : \tilde{E} \to F$ be the projection onto the first coordinate, and let

$$T\colon \tilde{E} \to E, \quad (x, y) \mapsto x + y.$$

Then T is continuous and bijective and thus has a continuous inverse. Since $P = \pi \circ T^{-1}$ this shows that P is continuous.

2. Given $f_0, f_1 \in \mathcal{C}([0,1])$, the initial value problem

$$y'' + f_1 y' + f_0 y = g, \qquad y(0) = y_1, \quad y'(0) = y_2 \tag{2.8}$$

has a unique solution in $\mathcal{C}^2([0,1])$ for all $g \in \mathcal{C}([0,1])$ and $y_1, y_2 \in \mathbb{R}$. Let $E = \mathcal{C}^2([0,1])$ (equipped with $\|\cdot\|_2$, and let

$$F := \mathcal{C}([0,1]) \oplus \mathbb{R} \oplus \mathbb{R}$$

be equipped with

$$\|(f, x_1, x_2)\| := \max\{\|f\|_\infty, |x_1|, |x_2|\} \qquad (f \in \mathcal{C}([0,1]),\, x_1, x_2 \in \mathbb{R}).$$

Define $T\colon E \to F$ through

$$T\phi := (\phi'' + f_1\phi' + f_0\phi, \phi(0), \phi'(0)) \qquad (\phi \in \mathcal{C}^2([0,1])).$$

Then T is linear such that, for any $\phi \in \mathcal{C}^2([0,1])$:

$$\begin{aligned} \|T\phi\| &= \max\{\|\phi'' + f_1\phi' + f_0\phi\|_\infty, |\phi(0)|, |\phi'(0)|\} \\ &\leq \max\{1, \|f_1\|_\infty, \|f_0\|_\infty\} \underbrace{\sum_{j=0}^{2} \|\phi^{(j)}\|_\infty}_{=\|\phi\|_2}. \end{aligned}$$

Hence, T is bounded. The existence and uniqueness of the solutions of (2.8) imply that T is a bijection. By Corollary 2.5.3, T^{-1} is also continuous. Hence, the solutions of (2.8) depend continuously on the data g, y_1. and y_2.

Exercise 2.18 Show that c_0 is not complemented in ℓ^∞:

(i) Show that there is an uncountable family $(S_\alpha)_\alpha$ of infinite subsets of $\mathbb{N}$ such that $S_\alpha \cap S_\beta$ is finite for all $\alpha \neq \beta$. (*Hint*: Replace $\mathbb{N}$ by $\mathbb{Q}$ (you can do that because they have the same cardinality), use $\mathbb{R}$ as your index set, and utilize the fact that every real number is the limit of a sequence in $\mathbb{Q}$.)

(ii) There is no countable subset Φ of $(\ell^\infty/c_0)^*$ such that for each non-zero $f \in \ell^\infty/c_0$ there is $\phi \in \Phi$ with $\phi(f) \neq 0$. (*Hint*: Choose $(S_\alpha)_\alpha$ as in (i) and consider the family $(f_\alpha)_\alpha$ of the cosets in ℓ^∞/c_0 of the indicator functions of the sets S_α; show that, for fixed $\phi \in (\ell^\infty/c_0)^*$, the set $\{f_\alpha : \phi(f_\alpha) \neq 0\}$ is at most countable.)

(iii) Conclude from (i) and (ii) that c_0 is not complemented in ℓ^∞.

2.6 The closed graph theorem

Definition 2.6.1 Let E and F be normed spaces.

(a) A *partially defined operator* from E to F is a linear map $T\colon \mathcal{D}_T \to F$, where $\mathcal{D}_T$ is a subspace of F.

(b) A partially defined operator is called *closed* if its graph

$$\mathrm{Gr}\ T := \{(x, Tx) : x \in \mathcal{D}_T\}$$

is closed in $E \oplus F$.

Example Let $E = F := \mathcal{C}([0,1])$, let $\mathcal{D}_T := \mathcal{C}^1([0,1])$, and let

$$T\colon \mathcal{D}_T \to F, \quad f \mapsto f'.$$

Let $((f_n, Tf_n))_{n=1}^{\infty}$ be a sequence in Gr T with $(f_n, Tf_n) \to (g, h) \in E \oplus F$, i.e.

$$\|f_n - g\|_\infty \to 0 \qquad \text{and} \qquad \|f_n' - h\|_\infty \to 0.$$

It follows that $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{C}^1([0,1])$ with respect to $\|\cdot\|_1$. Let $f \in \mathcal{C}^1([0,1])$ be the limit of $(f_n)_{n=1}^{\infty}$, i.e.

$$\|f_n - f\|_\infty \to 0 \qquad \text{and} \qquad \|f_n' - f'\|_\infty \to 0.$$

It follows that $g = f$ and $h = f'$, so that $(g, h) \in \text{Gr } T$. Consequently, T is closed (although not continuous).

Theorem 2.6.2 (closed graph theorem) *Let E and F be Banach spaces, and let $T \colon E \to F$ be closed. Then T is continuous.*

Proof Define

$$\pi_1 \colon \text{Gr } T \to E, \quad (x, Tx) \mapsto x.$$

Then T is a continuous bijection and thus its inverse

$$\iota \colon E \to \text{Gr } T, \quad x \mapsto (x, Tx)$$

is continuous as well. Let

$$\pi_1 \colon \text{Gr } T \to F, \quad (x, Tx) \mapsto Tx.$$

Then π_2 is continuous, and so is $T = \pi_2 \circ \iota$. $\square$

Corollary 2.6.3 *Let E and F be Banach spaces, and let $T \colon E \to F$ be linear with the following property:*

> *If $(x_n)_{n=1}^{\infty}$ is a sequence in E and y is a vector in F such that $x_n \to 0$ and $Tx_n \to y \in F$, then $y = 0$.*

Then T is continuous.

Proof Let $(x_n)_{n=1}^{\infty}$ be a sequence in E, and let $x \in E$ and $y \in F$ be such that

$$\|x_n - x\| \to 0 \qquad \text{and} \qquad \|Tx_n - y\|_\infty \to 0.$$

It follows that

$$x_n - x \to 0 \qquad \text{and} \qquad T(x_n - x) = Tx_n - Tx \to y - Tx.$$

The hypothesis on T implies that $y - Tx = 0$, so that $(x, y) \in \text{Gr}, T$. Hence, Gr T is closed, and T is continuous by Theorem 2.6.2. $\square$

Exercise 2.19 Let E and F be Banach spaces, and let $T\colon E \to F$ be linear. The *separating space* of T is defined as

$$\mathfrak{S}(T) := \{y \in F : \text{there is a sequence } (x_n)_{n=1}^{\infty} \text{ in } E \text{ with } x_n \to 0 \text{ and } Tx_n \to y\}$$

(i) Show that $\mathfrak{S}(T)$ is a closed, linear subspace of F.

(ii) Show that $\mathfrak{S}(T) = \{0\}$ if and only if $T \in \mathcal{B}(E, F)$.

Example Let X be a locally compact Hausdorff space, and let $\phi\colon X \to \mathbb{F}$ be such that $\phi f \in \mathcal{C}_0(X)$ for all $f \in \mathcal{C}_0(X)$. Define

$$M_\phi\colon \mathcal{C}_0(X) \to \mathcal{C}_0(X), \quad f \mapsto \phi f.$$

Let $(f_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{C}_0(X)$, and let $g \in \mathcal{C}_0(X)$ be such that

$$f_n \to 0 \qquad \text{and} \qquad Tf_n \to g.$$

For all $x \in X$, we have that

$$g(x) = \lim_{n\to\infty} \phi(x) f_n(x) = 0,$$

so that $g = 0$. With Corollary 2.6.3, it follows that M_ϕ is bounded. (It can then be shown that $\phi \in \mathcal{C}_b(X)$.)

Exercise 2.20 Let X be a locally compact Hausdorff space. A *multiplier* of $\mathcal{C}_0(X)$ is a linear map $T\colon \mathcal{C}_0(X) \to \mathcal{C}_0(X)$ such that

$$T(fg) = fTg \qquad (f, g \in \mathcal{C}_0(X))$$

Show that T is continuous.

Chapter 3

Spectral theory of bounded linear operators

In this chapter, we develop the basics of the spectral theory of bounded, linear operators on Banach spaces.

3.1 The spectrum of a bounded linear operator

The spectrum can be thought of as the appropriate infinite-dimensional analogue of the set of eigenvalues of a matrix.

Definition 3.1.1 Let E be a Banach space. We let

$$\operatorname{Inv}\mathcal{B}(E) := \{T \in \mathcal{B}(E) : T \text{ is invertible}\}.$$

Definition 3.1.2 Let E be a Banach space, and let $T \in \mathcal{B}(E)$. Then

$$\sigma(T) := \{\lambda \in \mathbb{F} : \lambda - T \notin \operatorname{Inv}\mathcal{B}(E)\}$$

is called the *spectrum* of T. The complement $\rho(T) := \mathbb{C} \setminus \sigma(T)$ is the *resolvent set* of T.

Examples 1. Let $\dim E < \infty$. Then:

$$\begin{aligned} \lambda \in \sigma(T) \quad &\iff \quad \lambda - T \text{ is not bijective} \\ &\iff \quad \lambda - T \text{ is not injective} \\ &\iff \quad \text{there is } x \in E \setminus \{0\} \text{ such that } Tx = \lambda x \\ &\iff \quad \lambda \text{ is an eigenvalue of } T. \end{aligned}$$

2. Let $E = \mathbb{R}^2$, and let $T = T_A$ for $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then $\sigma(T) = \varnothing$.

3. Let $\phi \in \mathcal{C}([0,1])$ and let

$$M_\phi \colon \mathcal{C}([0,1]) \to \mathcal{C}([0,1]), \quad f \mapsto \phi f.$$

Let $\lambda \in \mathbb{F} \setminus \phi([0,1])$. Then

$$\psi(x) := \frac{1}{\lambda - \phi(x)} \qquad (x \in [0,1])$$

defines an element of $\mathcal{C}([0,1])$. We have

$$M_\psi(\lambda - M_\phi)f = (\lambda M_\phi) - M_\psi f = \frac{\lambda - \phi}{\lambda - \phi} f = f \qquad (f \in \mathcal{C}([0,1])),$$

so that $M_\psi = (\lambda - M_\phi)^{-1}$ and $\lambda \notin \sigma(M_\phi)$. It follows that $\sigma(T) \subset \phi([0,1])$.

Conversely, let $\lambda \in \phi([0,1])$ and assume that $\lambda \notin \sigma(T)$. Let $\psi := (\lambda - M_\phi)^{-1}1$. Then we obtain

$$(\lambda - \phi)\psi = (\lambda - M_\phi)\psi = 1,$$

which is impossible.

Exercise 3.1 Let E be a Banach space, and let $P \in \mathcal{B}(E)$ be a projection. Show that $\sigma(P) \subset \{0,1\}$.

Exercise 3.2 Let $T \in \mathcal{B}(\mathcal{C}([0,1]))$ be defined through

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

(i) Show that T has no eigenvalues.

(ii) What is $\sigma(T)$?

Exercise 3.3 Let E be a Banach space, and let $T \in \mathcal{B}(T)$. Show that $\sigma(T) = \sigma(T^*)$.

Exercise 3.4 Let $p \in [1,\infty]$, and let $L, R \colon \ell^p \to \ell^p$ be defined through

$$L(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$$
$$\text{and} \quad R(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots) \qquad ((x_1, x_2, x_3, \ldots) \in \ell^p).$$

(i) Show that every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$ is an eigenvalue of R.

(ii) Conclude that $\sigma(R) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$.

(iii) Show that L has no eigenvalues, but $\sigma(L) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$. (*Hint*: Take adjoints.)

As Exercise 3.2 shows, a bounded linear operator on an infinite-dimensional Banach space may have an empty set of eigenvalues. As we shall see in the remainder of this section, the spectrum of a bounded, linear operator is a compact set, which is non-empty if $\mathbb{F} = \mathbb{C}$.

Lemma 3.1.3 *Let E be a Banach space, and let $T \in \mathcal{B}(E)$ be such that $\|\mathrm{id}_E - T\| < 1$. Then $T \in \operatorname{Inv}\mathcal{B}(E)$.*

Proof Let

$$S := \sum_{n=0}^{\infty} (\mathrm{id}_E - T)^n.$$

It follows that

$$\begin{aligned} S - TS &= (\mathrm{id}_E - T)S \\ &= \sum_{n=0}^{\infty} (\mathrm{id}_E - T)^{n+1} \\ &= \sum_{n=1}^{\infty} (\mathrm{id}_E - T)^n \\ &= S - \mathrm{id}_E, \end{aligned}$$

so that $TS = \mathrm{id}_E$. In a similar way, $ST = \mathrm{id}_E$ is proven. $\square$

Corollary 3.1.4 *Let E be a Banach space, and let $T \in \mathcal{B}(E)$. Then $\sigma(T)$ is bounded by $\|T\|$.*

Proof Let $\lambda \in \mathbb{F}$ be such that $|\lambda| > \|T\|$. Then

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

so that

$$\lambda - T = \lambda \left(1 - \frac{T}{\lambda}\right) \in \operatorname{Inv}\mathcal{B}(E)$$

by Lemma 3.1.3. This means that $\lambda \in \rho(T)$. $\square$

Corollary 3.1.5 *Let E be a Banach space. Then $\operatorname{Inv}\mathcal{B}(E)$ is open in $\mathcal{B}(E)$.*

Proof Let $T \in \operatorname{Inv}\mathcal{B}(E)$, and let $S \in \mathcal{B}(E)$ be such that $\|S - T\| < \frac{1}{\|T^{-1}\|}$. It follows that

$$\|1 - T^{-1}S\| = \|T^{-1}(T - S)\| < 1,$$

so that $T^{-1}S \in \operatorname{Inv}\mathcal{B}(E)$ by Lemma 3.1.3. $\square$

Corollary 3.1.6 *Let E be a Banach space, and let $T \in \mathcal{B}(E)$. Then $\sigma(T)$ is closed in $\mathbb{F}$.*

Proof Let $\lambda \in \rho(T)$, i.e. $\lambda - T \in \operatorname{Inv} \mathcal{B}(E)$. By Corollary 3.1.5, there is $\epsilon > 0$ such that $S \in \operatorname{Inv} \mathcal{B}(E)$ whenever $\|\lambda - T - S\| < \epsilon$. For $\mu \in \mathbb{F}$ with $|\lambda - \mu| < \epsilon$, we then have

$$\|(\lambda - T) - (\mu - T)\| = |\lambda - \mu| < \epsilon,$$

so that $\mu - T \in \operatorname{Inv} \mathcal{B}(E)$. $\square$

Lemma 3.1.7 *Let E be a Banach space, and let $(T_n)_{n=1}^\infty$ be a sequence in* $\operatorname{Inv} \mathcal{B}(E)$ *which converges to* $T \in \operatorname{Inv} \mathcal{B}(E)$. *Then* $T_n^{-1} \to T^{-1}$.

Proof We first show that $\sup_{n\in\mathbb{N}} \|T_n^{-1}\| < \infty$. Since $T_n T^{-1} \to \mathrm{id}_E$, there is $N \in \mathbb{N}$ such that

$$\|\mathrm{id}_E - T_n T^{-1}\| < \frac{1}{2} \qquad (n \geq N).$$

In the proof of Lemma 3.1.7, we saw that

$$(T_n T^{-1})^{-1} = \sum_{k=0}^{\infty} (\mathrm{id}_E - T_n T^{-1})^k \qquad (n \geq N),$$

so that

$$\|T T_n^{-1}\| = \|(T_n T^{-1})^{-1}\| \leq \sum_{k=0}^{\infty} \|\mathrm{id}_E - T_n T^{-1}\|^k \leq 2 \qquad (n \geq N).$$

Consequently,

$$\|T_n^{-1}\| \leq \|T^{-1}\| \|T T_n^{-1}\| \leq 2\|T^{-1}\| \qquad (n \geq N).$$

Since

$$\|T_n^{-1} - T^{-1}\| = \|T_n^{-1}(T - T_n)T^{-1}\| \leq \|T_n^{-1}\| \|T - T_n\| \|T^{-1}\| \qquad (n \in \mathbb{N}),$$

it follows that $\lim_{n\to\infty} T_n^{-1} = T^{-1}$. $\square$

Theorem 3.1.8 *Let $E \neq \{0\}$ be a Banach space over $\mathbb{C}$, and let $T \in \mathcal{B}(E)$. Then $\sigma(T)$ is a non-empty, compact subset of $\mathbb{C}$.*

Proof In view of Corollaries 3.1.4 and 3.1.6, it is clear that $\sigma(T)$ is compact (this does not require that the Banach space be over $\mathbb{C}$).

All we have to show is therefore that $\sigma(T) \neq \varnothing$. Assume towards a contradiction that $\sigma(T) = \varnothing$, i.e. $\lambda - T \in \operatorname{Inv} \mathcal{B}(E)$ for all $\lambda \in \mathbb{C}$. Let $\phi \in \mathcal{B}(E)^*$, and define

$$f \colon \mathbb{C} \to \mathbb{C}, \quad \lambda \mapsto \phi((\lambda - T)^{-1}).$$

Let $h \in \mathbb{C} \setminus \{0\}$. Then we have:

$$\begin{aligned}
\frac{f(\lambda+h)-f(\lambda)}{h} &= \frac{1}{h}\phi((\lambda+h-T)^{-1}-(\lambda-T)^{-1}) \\
&= \frac{1}{h}\phi((\lambda+h-T)^{-1}[(\lambda-T)-(\lambda+h-T)](\lambda-T)^{-1}) \\
&= -\frac{1}{h}\phi((\lambda+h-T)^{-1}[(\lambda-T)-(\lambda+h-T)](\lambda-T)^{-1}) \\
&\overset{h\to 0}{\to} -\phi((\lambda-T)^{-2}).
\end{aligned}$$

Hence, f is holomorphic. Moreover, since

$$f(\lambda) = \frac{1}{\lambda}\phi((1-\lambda^{-1}T)^{-1}) \overset{|\lambda|\to\infty}{\to} 0, \tag{3.1}$$

the function f is also bounded. By Liouville's theorem, this means that f is constant. In conjunction with (3.1), this means $f \equiv 0$. In particular, $0 = f(0) = \phi(T^{-1})$. Since $\phi \in \mathcal{B}(E)^*$ was arbitrary, Corollary 2.1.7 yields $T^{-1} = 0$, which is impossible. □

Exercise 3.5 Let $\varnothing \neq K \subset \mathbb{C}$ be compact. Show that there are a Banach space E over $\mathbb{C}$ and $T \in \mathcal{B}(E)$ such that $\sigma(T) = K$.

Exercise 3.6 Let $p \in [1,\infty]$, and let $L, R \colon \ell^p \to \ell^p$ be defined through

$$\begin{aligned}
&L(x_1,x_2,x_3,\dots) = (0,x_1,x_2,\dots) \\
&\text{and}\quad R(x_1,x_2,x_3,\dots) = (x_2,x_3,x_4,\dots) \qquad ((x_1,x_2,x_3,\dots) \in \ell^p).
\end{aligned}$$

(i) Show that every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$ is an eigenvalue of R.

(ii) Conclude that $\sigma(R) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$.

(iii) Show that L has no eigenvalues, but $\sigma(L) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$. (*Hint*: Take adjoints.)

Exercise 3.7 Let E be a Banach space over $\mathbb{C}$, and let $T \in \operatorname{Inv}\mathcal{B}(E)$.

(i) Show that $\lambda \in \sigma(T)$ if and only if $\lambda^{-1} \in \sigma(T^{-1})$.

(ii) Suppose further that T is an isometry. Show that $\sigma(T) \subset \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

Exercise 3.8 Let E be a Banach space over $\mathbb{C}$. Let $T \in \mathcal{B}(E) \setminus \operatorname{Inv}\mathcal{B}(E)$ be such that there is a sequence $(T_n)_{n=1}^\infty$ in $\operatorname{Inv}\mathcal{B}(E)$ such that $T = \lim_{n\to\infty} T_n$. Show that $\lim_{n=1} \|T_n^{-1}\| = \infty$.

Exercise 3.9 Let E be a Banach space over $\mathbb{C}$. An element $\lambda \in \mathbb{C}$ is called an *approximate eigenvalue* for T if

$$\inf\{\|(\lambda-T)x\| : x \in E,\ \|x\| = 1\} = 0.$$

Show that

$$\partial\sigma(T) \subset \{\text{approximate eigenvalues of } T\} \subset \sigma(T).$$

As Exercise 3.5, there is nothing more that can be said about the spectra of bounded, linear operators on Banach spaces over $\mathbb{C}$ except that they are non-empty subsets of $\mathbb{C}$. To get more detailed information, we need to look at a smaller class of operators.

3.2 Spectral theory for compact operators

In this section, all spaces are over $\mathbb{C}$.

Lemma 3.2.1 *Let E be a Banach space, let F be a closed subspace of E, let $T \in \mathcal{K}(E)$, and let $\lambda \in \mathbb{C} \setminus \{0\}$ be such that*

$$\inf\{\|(\lambda - T)x\| : x \in F,\ \|x\| = 1\} = 0.$$

Then $F \cap \ker(\lambda - E) \neq \{0\}$.

Proof Let $(x_n)_{n=1}^\infty$ be a sequence in F with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $(\lambda - T)x_n \to 0$. Since T is compact, $(Tx_n)_{n=1}^\infty$ has a convergent subsequence $(Tx_{n_k})_{k=1}^\infty$. Hence $(\lambda x_{n_k})_{k=1}^\infty$ converges and since $\lambda \neq 0$, so does $(x_{n_k})_{k=1}^\infty$. Let $x := \lim_{k\to\infty} x_{n_k}$. Then $x \in F$ with $\|x\| = 1$ belongs to $\ker(\lambda - T)$. □

Proposition 3.2.2 *Let E be a Banach space, let $T \in \mathcal{K}(E)$, and let $\lambda \in \sigma(T) \setminus \{0\}$. Then the following hold:*

(i) $\dim \ker(\lambda - T) < \infty$.

(ii) *$(\lambda - T)E$ is closed and has finite codimension.*

(iii) *There is $n \in \mathbb{N}$ such that $\ker(\lambda - T)^n = \ker(\lambda - T)^{n+1}$ and $(\lambda - T)^n E = (\lambda - T)^{n+1}$.*

Proof For (i), observe that

$$\mathrm{id}|_{\ker(\lambda - T)} = \frac{1}{\lambda} T|_{\ker(\lambda - T)}$$

is compact. This implies $\dim \ker(\lambda - T) < \infty$.

For (ii) choose that a closed subspace F of E such that $E = \ker(\lambda - T) \oplus F$. It follows that $(\lambda - T)F = (\lambda - T)E$. Since $F \cap \ker(\lambda - T) = \{0\}$, Lemma 3.2.1 implies that

$$C := \inf\{\|(\lambda - T)x\| : x \in F,\ \|x\| = 1\} > 0.$$

Hence,

$$\|(\lambda - T)x\| \geq C\|x\| \qquad (x \in F),$$

so that $(\lambda - T)F = (\lambda - T)E$ is closed.

To see that $(\lambda - T)E$ has finite codimension, note that

$$\begin{aligned} (\lambda - T)E^\perp &= \{\phi \in E^* : \phi((\lambda - T)x) = 0 \text{ for all } x \in E\} \\ &= \{\phi \in E^* : ((\lambda - T)\phi)(x) = 0 \text{ for all } x \in E\} \\ &= \ker(\lambda - T^*). \end{aligned}$$

Since T^* is also compact by Theorem 2.2.4, (i) yields $\dim(\lambda - T)E^\perp < \infty$, so that $(\lambda - T)E$ has finite codimension (since it is closed).

We only prove the first statement of (iii) in detail (the second one is established analogously). Assume towards a contradiction that

$$E_{n+1} := \ker(\lambda - T)^{n+1} \supsetneq \ker(\lambda - T)^n =: E_n \qquad (n \in \mathbb{N}).$$

For each $n \in \mathbb{N}$ choose $x_n \in E_{n+1}$ such that $\|x_n\| = 1$ and $\operatorname{dist}(x_n, E_n) \geq \frac{1}{2}$. Let $n > m \geq 2$. Since

$$Tx_n - Tx_m = \lambda x_n - \underbrace{(\lambda - T)x_n}_{\in E_n} + \underbrace{(\lambda - T)x_m - \lambda x_m}_{\in E_m \subset E_n},$$

we have

$$\|Tx_n - Tx_m\| \geq \frac{|\lambda|}{2}.$$

Hence, $(Tx_n)_{n=1}^\infty$ has no convergent subsequence. To prove the second statement, proceed similarly, first noting that

$$(\lambda - T)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^k (-T)^{n-k} = \lambda^n - \underbrace{\sum_{k=0}^{n-1} (-1)\lambda^k (-T)^{n-k}}_{\in \mathcal{K}(E)} \qquad (n \in \mathbb{N}),$$

so that $(\lambda - T)^n E$ is closed for each $n \in \mathbb{N}$. $\square$

Lemma 3.2.3 *Let E be a Banach space, let $T \in \mathcal{K}(E)$, and let $\lambda \in \sigma(T) \setminus \{0\}$. Then λ is an eigenvalue of T or of T^*.*

Proof Suppose that λ is not an eigenvalue of T. By Lemma 3.2.1, this means that

$$\inf\{\|(\lambda - T)x\| = 0 : x \in E, \ \|x\| = 1\} > 0.$$

As in the proof of Proposition 3.2.2, we conclude that $\lambda - T$ is injective and has closed range. Since $\lambda \in \sigma(T)$, we have $(\lambda - T)E \subsetneq E$. Choose $\phi \in E^* \setminus \{0\}$ such that $\phi \in (\lambda - T)E^\perp$. Again as in the proof of Proposition 3.2.2, we see that $\phi \in \ker(\lambda - T)^*$. $\square$

Lemma 3.2.4 *Let E be a Banach space, let $T \in \mathcal{K}(E)$, and let $(\lambda_n)_{n=1}^\infty$ be a sequence of pairwise distinct eigenvalues of T. Then $\lim_{n\to\infty} \lambda_n = 0$.*

Proof Without loss of generality suppose that $\lambda_n \neq 0$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, choose $x_n \in \ker(\lambda_n - T) \setminus \{0\}$. Let

$$E_n := \operatorname{lin}\{x_1, \ldots, x_n\},$$

so that $E_1 \subsetneq \cdots \subsetneq E_n \subsetneq E_{n+1} \subsetneq \cdots$. For each $n \geq 2$, choose $y_n \in E_n$ such that

$$\|y_n\| = 1 \qquad \text{and} \qquad \operatorname{dist}(y_n, E_{n-1}) \geq \frac{1}{2}.$$

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be such that $y_n = \alpha_1 x_1 + \cdots + \alpha_n x_n$. It follows that

$$(\lambda_n - T)y_n = \alpha_1(\lambda_n - \lambda_1)x_1 + \cdots + \alpha_{n-1}(\lambda_n - \lambda_{n-1})x_{n-1}.$$

For $n > m \geq 2$, this yields:

$$T(\lambda_n^{-1}y_n) - T(\lambda_m^{-1}y_m) = \underbrace{\lambda_n^{-1}(\lambda_n - T)y_n - \lambda_m^{-1}(\lambda_m - T)y_m + y_m}_{\in E_{n-1}} - y_n.$$

Consequently,

$$\|T(\lambda_n^{-1}y_n) - T(\lambda_m^{-1}y_m)\| \geq \frac{1}{2} \qquad (n \neq m,\ n, m \geq 2),$$

so that $(T(\lambda_n^{-1}y_n))_{n=1}^{\infty}$ has no convergent subsequence. Since T is compact, this means that $(\lambda_n^{-1}y_n)_{n=1}^{\infty}$ has no bounded subsequence, i.e. $\|\lambda_n^{-1}y_n\| = |\lambda_n^{-1}| \to \infty$ and thus $\lambda_n \to 0$. $\square$

Lemma 3.2.5 *Let E be a Banach space, and let $\lambda \in \sigma(T) \setminus \{0\}$. Then λ is an isolated point of $\sigma(T)$.*

Proof Assume that there is a sequence $(\lambda_n)_{n=1}^{\infty}$ in $\sigma(T) \setminus \{\lambda\}$ such that $\lambda_n \to \lambda$. Without loss of generality, we may suppose that the λ_ns are pairwise distinct. By Lemma 3.2.3, there is a subsequence $(\lambda_{n_k})_{k=1}^{\infty}$ of $(\lambda_n)_{n=1}^{\infty}$ such that

(a) each λ_{n_k} is an eigenvalue of T, or

(b) each λ_{n_k} is an eigenvalue of T^*.

However, (a) contradicts Lemma 3.2.4, and (b) leads equally to a contradiction with Lemma 3.2.4 if we apply that lemma with T^* instead of T. $\square$

Theorem 3.2.6 *Let E be a Banach space with $\dim E = \infty$, and let $T \in \mathcal{K}(E)$. Then one of the following holds:*

(i) *$\sigma(T) = \{0\}$;*

(ii) *$\sigma(T) = \{0, \lambda_1, \ldots, \lambda_n\}$, where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of T such that $\dim\ker(\lambda_j - T) < \infty$ for $j = 1, \ldots, n$.*

(iii) *$\sigma(T) = \{0, \lambda_1, \lambda_2, \ldots\}$, where $\lambda_1, \lambda_2, \ldots$ are eigenvalues of T such that $\dim\ker(\lambda_n - T) < \infty$ for $n \in \mathbb{N}$ and $\lim_{n\to\infty} \lambda_n = 0$.*

Proof Since $\dim E = \infty$ and $T \in \mathcal{K}(E)$, certainly $0 \in \sigma(T)$. By Lemma 3.2.5, $\sigma(T)$ is at most countably finite, by Lemma 3.2.4, $\lambda_n \to 0$ whenever $(\lambda_n)_{n=1}^{\infty}$ is a sequence of pairwise distinct eigenvalues of T, and Proposition 3.2.2(i) ensures that $\dim \ker(\lambda - T) < \infty$ for each non-zero eigenvalue λ. We are thus done once we have shown that every $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue.

Let $\lambda \in \sigma(T) \setminus \{0\}$. By Proposition 3.2.2(iii), there is $n \in \mathbb{N}$ such that $(\lambda - E)^n E = (\lambda - E)^{n+1} E$. Since

Proposition 3.2.2(ii) yields that $(\lambda - T)^n E$ is closed and has finite codimension. Assume that λ is not an eigenvalue of T, so that $\lambda - T$ is injective. Consequently, $(\lambda - T)|_{(\lambda - T)^n E}$ is bijective. Let $S \in \mathcal{B}((\lambda - T)^n E)$ be the inverse of $(\lambda - T)|_{(\lambda - T)^n E}$, and define

$$P \colon E \to E, \quad x \mapsto S^n (\lambda - T)^n x.$$

It follows that

$$P^2 = S^n (\lambda - T)^n S^n (\lambda - T)^n = S^n (\lambda - T)^n = P.$$

We also have:

$$\begin{aligned} (\lambda - T)P &= (\lambda - T) S^n (\lambda - T)^n \\ &= S^{n-1} (\lambda - T)^n \\ &= S^{n-1} S (\lambda - T)(\lambda - T)^n \\ &= S^n (\lambda - T)^{n+1} \\ &= P(\lambda - T). \end{aligned}$$

Since $\lambda - T$ is not bijective, but $\lambda - T$ is assumed to be injective, there is $x \in E \setminus (\lambda - T)E$, so that $y := x - Px \neq 0$. On the other hand, we have

$$(\lambda - T)^n y = P(\lambda - T)^n y = (\lambda - T)^n P y = (\lambda - T)^n (Px - P^2 x) = 0,$$

so that $\ker(\lambda - T)^n \neq \{0\}$ and thus $\ker(\lambda - T) \neq \{0\}$. $\square$

Exercise 3.10 Let $\phi \in \ell^\infty$. Show that $M_\phi \in \mathcal{B}(\ell^\infty)$ is compact if and only if $\phi \in c_0$.

Corollary 3.2.7 (Fredholm alternative) *Let E be a Banach space, let $T \in \mathcal{K}(E)$, and let $\lambda \in \mathbb{C} \setminus \{0\}$. Then the following are equivalent:*

(i) $\lambda - T$ *is bijective.*

(ii) $\lambda - T$ *is injective.*

(iii) $\lambda - T$ *is surjective.*

Proof (i) $\Longrightarrow$ (ii), (iii) is trivial.

(ii) $\Longrightarrow$ (i): clear by Theorem 3.2.6.

(iii) $\Longrightarrow$ (i): Assume that $\lambda - T$ is not bijective, i.e. $\lambda \in \sigma(T) = \sigma(T^*)$. Hence, λ is an eigenvalue of T^*. Hence, there are non-zero elements in $\ker(\lambda - T^*) = (\lambda - T)E^{\perp}$. Consequently, $(\lambda - T)E \neq E$ must hold, contradicting the surjectivity of $\lambda - T$. $\square$

Example Let $\lambda \in \mathbb{C} \setminus \{0\}$, and let $k \in \mathcal{C}([0,1] \times [0,1])$, and consider, for $f, g \in \mathcal{C}([01,])$ the integral equations

$$\lambda f(x) - \int_0^1 f(y)k(x,y)\,dy = g(x) \qquad (x \in [0,1]) \tag{3.2}$$

and

$$\lambda f(x) - \int_0^1 f(y)k(x,y)\,dy = 0 \qquad (x \in [0,1]). \tag{3.3}$$

Then there are the following alternatives:

(i) (3.2) has a unique solution $f \in \mathcal{C}([0,1])$ for each $g \in \mathcal{C}([0,1])$; in particular, (3.3) only has the trivial solution $f \equiv 0$.

(ii) There is $g \in \mathcal{C}([0,1])$ such that (3.2) has no solution $f \in \mathcal{C}([0,1])$. In this case, (3.3) has non-trivial solutions $f \in \mathcal{C}([0,1])$, which form a finite-dimensional subspace of $\mathcal{C}([0,1])$.

Since the Fredholm operator on $\mathcal{C}([0,1])$ with kernel k is compact, this is an immediate consequence of Corollary 3.2.7.

To make stronger assertions on the spectral theory of compact operators, we need to leave the general Banach space framework.

3.3 Hilbert spaces

Hilbert spaces are, in a certain sense, the infinite-dimensional spaces which behave most like finite-dimensional Euclidean space.

3.3.1 Inner products

Definition 3.3.1 A *semi-inner product* on a vector space E is map $[\cdot,\cdot]\colon E \times E \to \mathbb{F}$ such that

(a) $[\lambda x + \mu y, z] = \lambda[x,z] + \mu[y,z] \qquad (\lambda, \mu \in \mathbb{F},\ x, y, z \in E)$;

(b) $[z, \lambda x + \mu y] = \overline{\lambda}[z,x] + \overline{\mu}[z,y] \qquad (\lambda, \mu \in \mathbb{F},\ x, y, z \in E)$;

(c) $[x,x] \geq 0 \qquad (x \in E)$;

(d) $\overline{[x,y]} = [y,x] \qquad (x,y \in E)$.

A semi-inner product is called an *inner product* if

$$[x,x] = 0 \quad \Longleftrightarrow \quad x = 0 \qquad (x \in E).$$

Example Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space. For $f, g \in \mathcal{L}^2(\Omega, \mathfrak{S}, \mu)$, define

$$[f,g] := \int_\Omega f(\omega)\overline{g(\omega)}\, d\mu(\omega). \tag{3.4}$$

Then $[\cdot,\cdot]$ is a semi-inner product.

Proposition 3.3.2 (Cauchy–Schwarz inequality) *Let E be a vector space with a semi-inner product $[\cdot,\cdot]$ on it. Then*

$$|[x,y]|^2 \leq [x,x][y,y] \qquad (x,y \in E)$$

holds.

Proof Let $x, y \in E$. For all $\lambda \in \mathbb{F}$, we have:

$$0 \leq [x - \lambda y, x - \lambda y] = [x,x] - \lambda[y,x] - \overline{\lambda}[x,y] + |\lambda|^2[y,y]. \tag{3.5}$$

Choose $\mu \in \mathbb{F}$ with $|\mu| = 1$ such that $[y,x] = \mu|[y,x]|$. For $t \in \mathbb{R}$ and $\lambda = t\overline{\mu}$ we obtain from (3.5):

$$\begin{aligned} 0 &\leq [x,x] - t\overline{\mu}[y,x] - t\mu[x,y] + t^2[y,y] \\ &= [x,x] - t\overline{\mu}\mu|[y,x]| - t\mu\overline{\mu}|[x,y]| + t^2[y,y] \\ &= [x,x] - 2t|[y,x]| + t^2[y,y] \\ &=: q(t). \end{aligned}$$

Then q is a quadratic polynomial in t with at most one zero in $\mathbb{R}$. Hence, the discriminant of q must be less than or equal to zero:

$$0 \geq 4|[y,x]|^2 - 4[x,x][y,y].$$

This yields the claim. $\square$

Corollary 3.3.3 *Let E be a vector space with a semi-inner product $[\cdot,\cdot]$. Then*

$$\|x\| := [x,x]^{\frac{1}{2}} \qquad (x \in E)$$

defines a seminorm on E, which is a norm if and only if $[\cdot,\cdot]$ is a inner product.

Proof Only the triangle inequality needs proof. For $x, y \in E$, we have:

$$\begin{aligned}
\|x+y\|^2 &= [x+y, x+y] \\
&= [x,x] + [x,y] + \underbrace{[y,x]}_{=\overline{[x,y]}} + [y,y] \\
&= [x,x] + 2\mathrm{Re}\,[x,y] + [y,y] \\
&\leq [x,x] + 2|[x,y]| + [y,y] \\
&\leq [x,x] + 2[x,x]^{\frac{1}{2}}[y,y]^{\frac{1}{2}} + [y,y], \qquad \text{by Proposition 3.3.2,} \\
&= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Taking roots yields the triangle inequality. $\square$

Exercise 3.11 Let E be a linear space, and let $[\cdot,\cdot]$ be a semi-inner product on E.

(i) Show that $F := \{x \in E : [x,x] = 0\}$ is a linear subspace of E.

(ii) Show that

$$\langle x+F, y+F \rangle := [x,y] \qquad (x, y \in E)$$

defines an inner product on E/F.

Definition 3.3.4 A vector space $\mathfrak{H}$ equipped with an inner product $\langle\cdot,\cdot\rangle$ is called a *Hilbert space* if $\mathfrak{H}$ equipped with the norm

$$\|\xi\| := \langle \xi, \xi \rangle^{\frac{1}{2}} \qquad (\xi \in \mathfrak{H})$$

is a Banach space.

Example Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space. Then (3.4) induces a inner product on $L^2(\Omega, \mathfrak{S}, \mu)$ turning it into a Hilbert space. In particular, for each index set $I \neq \varnothing$, the space

$$\ell^2(I) := \left\{ f : I \to \mathbb{F} : \sum_{i \in I} |f(i)|^2 < \infty \right\}$$

equipped with

$$\langle f, g \rangle := \sum_{i \in I} f(i)\overline{g(i)} \qquad (f, g \in \ell^2(I))$$

is a Hilbert space.

3.3.2 Orthogonality and self-duality

Definition 3.3.5 Let $\mathfrak{H}$ be a Hilbert space. We say that $\xi, \eta \in \mathfrak{H}$ are orthogonal — in symbols: $\xi \perp \eta$ — if $\langle \xi, \eta \rangle = 0$.

Exercise 3.12 Let $\mathfrak{H}$ be a Hilbert space, and let $\xi_1, \ldots, \xi_n \in \mathfrak{H}$ be pairwise orthogonal. Show that

$$\|\xi_1 + \cdots + \xi_n\|^2 = \|\xi_1\|^2 + \cdots + \|\xi_n\|^2.$$

How do you interpret this geometrically?

Lemma 3.3.6 (parallelogram law) *Let $\mathfrak{H}$ be a Hilbert space. Then we have:*

$$\|\xi + \eta\|^2 + \|\xi - \eta\|^2 = 2(\|\xi\|^2 + \|\eta\|^2) \qquad (\xi, \eta \in \mathfrak{H}).$$

Proof We have

$$\|\xi + \eta\|^2 = \langle \xi + \eta, \xi + \eta \rangle = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\,\langle \xi, \eta \rangle$$

and

$$\|\xi - \eta\|^2 = \langle \xi + \eta, \xi + \eta \rangle = \|\xi\|^2 + \|\eta\|^2 - 2\mathrm{Re}\,\langle \xi, \eta \rangle.$$

Adding both equations yields the claim. $\square$

Theorem 3.3.7 *Let K be a closed, convex, non-empty subset of a Hilbert space $\mathfrak{H}$. Then, for each $\xi \in \mathfrak{H}$, there is a unique $\eta \in K$ such that $\|\xi - \eta\| = \mathrm{dist}(\xi, K)$.*

Proof Let $\xi \in \mathfrak{H}$, and let $\delta := \mathrm{dist}(\xi, K)$, so that there is a sequence $(\eta_n)_{n=1}^\infty$ in K such that $\|\xi - \eta_n\| \to \delta$. Note that, for $n, m \in \mathbb{N}$, we have

$$\begin{aligned}
&\left\| \frac{1}{2}(\eta_n - \eta_m) \right\|^2 \\
&= \left\| \frac{1}{2}[(\eta_n - \xi) - (\eta_m - \xi)] \right\|^2 \\
&= \frac{1}{2}\left(\|\eta_n - \xi\|^2 + \|\eta_m - \xi\|^2\right) - \underbrace{\left\| \underbrace{\frac{1}{2}(\eta_n + \eta_m)}_{\in K} - \eta \right\|^2}_{\geq \delta^2}, \qquad \text{by Lemma 3.3.6.} \quad (3.6)
\end{aligned}$$

Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$\|\eta_n - \xi\|^2 < \delta^2 + \frac{1}{4}\epsilon^2 \qquad (n \geq N).$$

Then (3.6) yields

$$\left\|\frac{1}{2}(\eta_n - \eta_m)\right\|^2 < \frac{1}{2}\left(2\delta^2 + \frac{1}{2}\epsilon^2\right) - \delta^2 = \frac{1}{4}\epsilon^2 \qquad (n, m \geq N)$$

and thus

$$\|\eta_n - \eta_m\| < \epsilon \qquad (n, m \geq N).$$

Thus, $(\eta_n)_{n=1}^\infty$ is a Cauchy sequence which therefore converges to some $\eta \in K$. It is obvious that $\|\xi - \eta\| = \delta$.

To prove the uniqueness of η, let $\eta_1, \eta_2 \in K$ be such that $\|\xi - \eta_j\| = \delta$ for $j = 1, 2$. Since $\frac{1}{2}(\eta_1 + \eta_2) \in K$, we have

$$\delta \leq \left\|\xi - \frac{1}{2}(\eta_1 + \eta_2)\right\| = \left\|\frac{1}{2}(\xi - \eta_1) + \frac{1}{2}(\xi - \eta_2)\right\| \leq \frac{1}{2}(\|\xi - \eta_1\| + \|\xi - \eta_2\|) = \delta,$$

so that $\left\|\xi - \frac{1}{2}(\eta_1 + \eta_2)\right\| = \delta$ as well. This, in turn, implies that

$$\begin{aligned} \delta^2 &= \left\|\frac{1}{2}(\xi - \eta_1) + \frac{1}{2}(\xi - \eta_2)\right\|^2 \\ &= 2\underbrace{\left\|\frac{1}{2}(\xi - \eta_1)\right\|^2}_{=\frac{1}{4}\delta^2} + 2\underbrace{\left\|\frac{1}{2}(\xi - \eta_2)\right\|^2}_{=\frac{1}{4}\delta^2} - \left\|\frac{1}{2}(\eta_1 - \eta_2)\right\|^2 \end{aligned}$$

and thus

$$\delta^2 = \delta^2 - \left\|\frac{1}{2}(\eta_1 - \eta_2)\right\|^2.$$

This means that $\eta_1 = \eta_2$. $\square$

Lemma 3.3.8 *Let $\mathfrak{H}$ be a Hilbert space, let $\mathfrak{K}$ be a closed subspace, and let $\xi \in \mathfrak{H}$. Then the following are equivalent for $\eta \in \mathfrak{K}$:*

(i) $\|\xi - \eta\| = \operatorname{dist}(\xi, \mathfrak{K})$;

(ii) $\xi - \eta \perp \mathfrak{K}$, *i.e.* $\xi - \eta \perp \tilde{\eta}$ *for each* $\tilde{\eta} \in \mathfrak{K}$.

Proof (i) $\Longrightarrow$ (ii): Let $\tilde{\eta} \in \mathfrak{K}$. Then

$$\|\xi - \eta\|^2 \leq \|\xi - (\eta + \tilde{\eta})\|^2 = \|\xi - \eta\|^2 - 2\mathrm{Re}\,\langle \xi - \eta, \tilde{\eta}\rangle + \|\tilde{\eta}\|^2 \qquad (3.7)$$

holds, so that $2\mathrm{Re}\,\langle \xi - \eta, \tilde{\eta}\rangle \leq \|\tilde{\eta}\|^2$. Choose $\lambda \in \mathbb{F}$ such that $|\lambda| = 1$ and $\langle \xi - \eta, \tilde{\eta}\rangle = \lambda|\langle \xi - \eta, \tilde{\eta}\rangle|$. Replacing $\tilde{\eta}$ in (3.7) with $t\lambda\tilde{\eta}$ for $t \in \mathbb{R}$ thus yields

$$\underbrace{2\mathrm{Re}\,\langle \xi - \eta, t\lambda\tilde{\eta}\rangle}_{=2t|\langle \xi - \eta, \tilde{\eta}\rangle|} \leq t^2\|\tilde{\eta}\| \qquad (t \in \mathbb{R}),$$

i.e.

$$2|\langle \xi - \eta, \tilde{\eta} \rangle| \leq t\|\tilde{\eta}\|^2 \qquad (t > 0).$$

Letting $t \to 0$ yields $\langle \xi - \eta, \tilde{\eta} \rangle = 0$, i.e. $\xi - \eta \perp \tilde{\eta}$.

(ii) $\Longrightarrow$ (i): Let $\tilde{\eta} \in \mathfrak{K}$, so that $\xi - \eta \perp \eta - \tilde{\eta}$. It follows that

$$\|\xi - \tilde{\eta}\|^2 = \|(\xi - \eta) + (\eta - \tilde{\eta})\|^2 = \|\xi - \eta\|^2 + \|\eta - \tilde{\eta}\|^2 \geq \|\xi - \eta\|^2,$$

which proves the claim. $\square$

As you saw in Exercise 2.18, a closed subspace of a Banach space need not be complemented. This situation is different for Hilbert spaces:

Theorem 3.3.9 *Let $\mathfrak{H}$ be a Hilbert space, and let $\mathfrak{K}$ be a closed subspace of $\mathfrak{H}$. Then there is a unique $P \in \mathcal{B}(\mathfrak{H})$ with the following properties:*

(i) $P\mathfrak{H} = \mathfrak{K}$;

(ii) $P^2 = P$;

(iii) $\ker P = \mathfrak{K}^\perp := \{\xi \in \mathfrak{H} : \xi \perp \mathfrak{K}\}$;

(iv) $\|P\| \leq 1$.

This map P is called the orthogonal projection *onto $\mathfrak{K}$.*

Proof For $\xi \in \mathfrak{H}$, define

$$P\xi := \text{the unique } \eta \in \mathfrak{K} \text{ such that } \|\xi - \eta\| = \operatorname{dist}(\xi, \mathfrak{K}).$$

It is then clear that $P \colon \mathfrak{H} \to \mathfrak{H}$ satisfies (i). By Lemma 3.3.8, this means that

$$P\xi := \text{the unique } \eta \in \mathfrak{K} \text{ such that } \xi - \eta \perp \mathfrak{K}.$$

This yields immediately that P is linear, is the identity on $\mathfrak{K}$, i.e. satisfies (ii), and also satisfies (iii).

Since $\xi - P\xi \perp \mathfrak{K}$ for all $\xi \in \mathfrak{H}$, we have

$$\|\xi\|^2 = \|(\xi - P\xi) + P\xi\|^2 = \|\xi - P\xi\|^2 + \|P\xi\|^2 \geq \|P\xi\|^2 \qquad (\xi \in \mathfrak{H}),$$

which yields (iv). $\square$

Exercise 3.13 Let $\mathfrak{H}$ be a Hilbert space, and let $\mathfrak{K}$ be a closed subspace of $\mathfrak{H}$. Show that $\mathfrak{H} \cong \mathfrak{K} \oplus \mathfrak{K}^\perp$.

Corollary 3.3.10 *Let $\mathfrak{H}$ be a Hilbert space, and let $\mathfrak{K}$ be a closed subspace of $\mathfrak{H}$. Then $\mathfrak{K}$ is complemented in $\mathfrak{H}$.*

Theorem 3.3.11 *Let $\mathfrak{H}$ be a Hilbert space, and let $\phi \in \mathfrak{H}^*$. Then there is a unique $\eta \in \mathfrak{H}$ such that*

$$\phi(\xi) = \langle \xi, \eta \rangle \qquad (\xi \in \mathfrak{H}). \tag{3.8}$$

Moreover, η satisfies $\|\phi\| = \|\eta\|$.

Proof The uniqueness of η is clear.

For the proof of existence, we may suppose without loss of generality that $\|\phi\| = 1$. Let $\mathfrak{K} := \ker \phi$, and let $P \in \mathcal{B}(\mathfrak{H})$ denote the orthogonal projection onto $\mathfrak{K}$. Choose $\eta_0 \in \mathfrak{H} \setminus \mathfrak{K}$. Then $\eta_0 - P\eta_0 \perp \mathfrak{K}$ and $\eta_0 - P\eta_0 \neq 0$, so that

$$\tilde{\eta} := \frac{\eta_0 - P\eta_0}{\|\eta_0 - P\eta_0\|}$$

is well-defined. Define $\psi \in \mathfrak{H}^*$ as

$$\psi \colon \mathfrak{H} \to \mathbb{F}, \quad \xi \mapsto \langle \xi, \tilde{\eta} \rangle,$$

so that $\ker \phi = \mathfrak{K} \subset \ker \psi$. This means that there is $\lambda \in \mathbb{F}$ such that $\phi = \lambda \psi$. Since

$$|\psi(\xi)| = |\langle \xi, \tilde{\eta} \rangle| \leq \|\tilde{\eta}\| \|\xi\| = \|\xi\| \qquad (\xi \in \mathfrak{H})$$

by the Cauchy–Schwarz inequality, we have

$$1 \geq \|\psi\| \geq |\psi(\tilde{\eta})| = \|\tilde{\eta}\|^2 = 1,$$

which implies $|\lambda| = 1$. Letting $\eta := \overline{\lambda}\tilde{\eta}$, we obtain (3.8). $\square$

Corollary 3.3.12 *Let $\mathfrak{H}$ be a Hilbert space, and define, for each $\eta \in \mathfrak{H}$, a functional $\phi_\eta \in \mathfrak{H}^*$ by letting*

$$\phi_\eta(\xi) := \langle \xi, \eta \rangle \qquad (\xi \in \mathfrak{H}).$$

Then the map

$$\mathfrak{H} \to \mathfrak{H}^*, \quad \eta \mapsto \phi_\eta$$

is a conjugate linear isometry.

As an application of this so-called self-duality of Hilbert spaces, we will now give a Hilbert space theoretic proof of the Radon–Nikodým theorem from measure theory.

Definition 3.3.13 Let $(\Omega, \mathfrak{S})$ be a measurable space, and let μ and ν be measures on $(\Omega, \mathfrak{S})$. We say that ν is *absolutely continuous* with respect to μ — in symbols: $\nu \ll \mu$ — if $\nu(N) = 0$ for every $N \in \mathfrak{S}$ such that $\mu(N) = 0$.

Example Let $(\Omega, \mathfrak{S}, \mu)$ be any measure space, and let $f \colon \Omega \to [0, \infty]$ be measurable. Then $\nu \colon \mathfrak{S} \to [0, \infty]$ defined by

$$\nu(S) := \int_S f(\omega)\, d\mu(\omega) \qquad (S \in \mathfrak{S})$$

is absolutely continuous with respect to μ.

Theorem 3.3.14 (Radon–Nikodým theorem) *Let $(\Omega, \mathfrak{S})$ be a measurable space, and let μ and ν be finite measures on $(\Omega, \mathfrak{S})$ such that $\nu \ll \mu$. Then there is a non-negative $h \in L^1(\Omega, \mathfrak{S}, \mu)$ such that*

$$\nu(S) = \int_S h(\omega)\, d\mu(\omega) \qquad (S \in \mathfrak{S})$$

Proof Let $\lambda := \mu + \nu$, and define $\phi \in L^2(\Omega, \mathfrak{S}, \lambda; \mathbb{R})^*$ by letting

$$\phi(f) := \int_\Omega f(\omega)\, d\nu(\omega) \qquad (f \in L^2(\Omega, \mathfrak{S}, \lambda)).$$

By Theorem 3.3.11, there is a unique $g \in L^2(\Omega, \mathfrak{S}, \lambda)$ such that

$$\phi(f) := \int_\Omega f(\omega) g(\omega)\, d\lambda(\omega) \qquad (f \in L^2(\Omega, \mathfrak{S}, \lambda)),$$

so that

$$\int_\Omega (1 - g(\omega)) f(\omega)\, d\nu(\omega) = \int_\Omega g(\omega) f(\omega)\, d\mu(\omega) \qquad (f \in L^2(\Omega, \mathfrak{S}, \lambda)).$$

Let $A := \{\omega \in \Omega : g(\omega) < 0\}$. It follows that

$$0 \geq \int_A g(\omega)\, d\mu(\omega) = \int_\Omega g(\omega)\chi_A(\omega)\, d\mu(\omega) = \int_\Omega (1 - g(\omega))\chi_A(\omega)\, d\nu(\omega) \geq 0,$$

so that $\mu(A) = 0$ and hence $\nu(A) = 0$. Let $B := \{\omega \in \Omega : g(\omega) \geq 1\}$. Similarly, we have

$$0 \geq \int_B (1 - g(\omega))\, d\nu(\omega) = \int_B g(\omega)\, d\nu(\omega) \geq 0,$$

so that $\mu(B) = \nu(B) = 0$ as well. We may therefore suppose without loss of generality that $g(\omega) \in [0, 1)$ for all $\omega \in \Omega$.

For $S \in \mathfrak{S}$ and $n \in \mathbb{N}$, define $f_n := (1 + g + \cdots + g^n)\chi_S$. It follows that

$$\begin{aligned}
\int_S (1 - g(\omega)^{n+1})\, d\nu(\omega) &= \int_\Omega (1 - g(\omega)) f_n(\omega)\, d\nu(\omega) \\
&= \int_\Omega g(\omega) f_n(\omega)\, d\mu(\omega) \\
&= \int_S (g(\omega) + g(\omega)^2 + \cdots + g(\omega)^{n+1})\, d\mu(\omega).
\end{aligned}$$

Let $h := \sum_{n=1}^\infty g^n$. Since $g(\Omega) \subset [0,1)$, this series converges; for the same reason, we have $1 - g^{n+1} \to 1$ pointwise. All in all, we have:

$$\begin{aligned}
\nu(S) &= \int_S d\nu(\omega) \\
&= \lim_{n\to\infty} \int_S (1 - g(\omega)^{n+1})\, d\nu(\omega), \qquad \text{by dominated convergence,} \\
&= \lim_{n\to\infty} \int_S (g(\omega) + g(\omega)^2 + \cdots + g(\omega)^{n+1})\, d\mu(\omega) \\
&= \int_S h(\omega)\, d\mu(\omega), \qquad \text{by monotone convergence.}
\end{aligned}$$

This completes the proof. $\square$

3.3.3 Orthonormal bases

Definition 3.3.15 Let $\mathfrak{H}$ be a Hilbert space. A family $(e_\alpha)_\alpha$ of vectors in $\mathfrak{H}$ is called *orthonormal* if $\|e_\alpha\| = 1$ and $e_\alpha \perp e_\beta$ for $\alpha \neq \beta$.

Examples 1. Let $\mathfrak{H} = \mathbb{F}^N$, and let $e_j := (0, \ldots, 0, 1, 0, \ldots, 0)$ for $j = 1, \ldots, N$, where the non-zero entry is in the j-th position. Then $(e_j)_{j=1}^N$ is orthonormal.

2. More generally, let $\mathfrak{H} = \ell^2(I)$ for $I \neq \varnothing$, and define, for $i \in I$, a vector $e_i \colon I \to \mathbb{F}$ by letting $e_i(j) := \delta_{i,j}$ for $j \in I$. Then $(e_j)_{j\in I}$ is orthonormal.

3. Let $\mathfrak{H} = L^2([0, 2\pi]; \mathbb{C})$. For $n \in \mathbb{N}$, define $e_n \in L^2([0, 2\pi])$ by letting

$$e_n(x) := \frac{1}{\sqrt{2\pi}} e^{inx} \qquad (x \in [0, 2\pi]).$$

Then, clearly, $\|e_n\| = 1$. For $n \neq m$, we have

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)x}\, dx = \frac{1}{2i\pi(n-m)} e^{i(n-m)x} \Big|_0^{2\pi} = 0.$$

Hence, $(e_n)_{n=1}^\infty$ is orthonormal.

Lemma 3.3.16 (Bessel's inequality) *Let $\mathfrak{H}$ be a Hilbert space, and let $(e_\alpha)_\alpha$ be an orthonormal family in $\mathfrak{H}$. Then*

$$\sum_\alpha |\langle \xi, e_\alpha \rangle|^2 \leq \|\xi\|^2 \qquad (\xi \in \mathfrak{H})$$

holds.

Proof For any finite number of indices $\alpha_1, \ldots, \alpha_n$, let $\eta := \xi - \sum_{j=1}^n \langle \xi, e_{\alpha_j} \rangle e_{\alpha_j}$. It follows that $\eta \perp e_{\alpha_j}$ for $j = 1, \ldots, n$. Consequently, we have

$$\begin{aligned} \|\xi\|^2 &= \|\eta\|^2 + \left\| \sum_{j=1}^n \langle \xi, e_{\alpha_j} \rangle e_{\alpha_j} \right\|^2 \\ &= \|\eta\|^2 + \sum_{j=1}^n |\langle \xi, e_{\alpha_j} \rangle|^2 \\ &\geq \sum_{j=1}^n |\langle \xi, e_{\alpha_j} \rangle|^2. \end{aligned}$$

Since $\alpha_1, \ldots, \alpha_n$ were arbitrary, this yields the claim. $\square$

Lemma 3.3.17 *Let $\mathfrak{H}$ be a Hilbert space, and let $(e_\alpha)_\alpha$ be an orthonormal family in $\mathfrak{H}$. Then $\sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha$ converges for every $\xi \in \mathfrak{H}$.*

Proof By Lemma 3.3.16, the set $\left\{ \alpha : |\langle \xi, e_\alpha \rangle| \geq \frac{1}{n} \right\}$ is finite for each $n \in \mathbb{N}$. Hence, there are only countably many α such that $\langle \xi, e_\alpha \rangle \neq 0$. Without loss of generality, we can therefore suppose that we are dealing with a sequence $(e_n)_{n=1}^\infty$.

For $n > m$, we have:

$$\left\| \sum_{k=1}^n \langle \xi, e_k \rangle e_k - \sum_{k=1}^m \langle \xi, e_k \rangle e_k \right\|^2 = \left\| \sum_{k=m+1}^n \langle \xi, e_k \rangle e_k \right\|^2 = \sum_{k=m+1}^n |\langle \xi, e_k \rangle|^2.$$

Let $\epsilon > 0$. Since $\sum_{k=1}^\infty |\langle \xi, e_k \rangle|^2 \leq \|\xi\|^2 < \infty$ by Lemma 3.3.16, there is $N \in \mathbb{N}$ such that

$$\sum_{k=m+1}^n |\langle \xi, e_k \rangle|^2 < \epsilon^2 \qquad (n, m \geq N),$$

so that

$$\left\| \sum_{k=1}^n \langle \xi, e_k \rangle e_k - \sum_{k=1}^m \langle \xi, e_k \rangle e_k \right\| < \epsilon \qquad (n, m \geq N).$$

Hence, $\left(\sum_{k=1}^n \langle \xi, e_k \rangle e_k \right)_{n=1}^\infty$ is Cauchy and thus converges. $\square$

Theorem 3.3.18 *Let $\mathfrak{H}$ be a Hilbert space, and let $(e_\alpha)_\alpha$ be an orthonormal family in $\mathfrak{H}$. Then the following are equivalent:*

(i) *$(e_\alpha)_\alpha$ is maximal.*

(ii) *If $\xi \perp e_\alpha$ for all α, then $\xi = 0$.*

(iii) *$\mathfrak{H} = \overline{\operatorname{lin} \{e_\alpha : \alpha\}}$.*

(iv) *If* $\xi \in \mathfrak{H}$, *then*

$$\xi = \sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha$$

holds.

(v) *If* $\xi, \eta \in \mathfrak{H}$, *then*

$$\langle \xi, \eta \rangle = \sum_\alpha \langle \xi, e_\alpha \rangle \langle e_\alpha, \eta \rangle$$

holds.

(vi) Parseval's identity *holds for each* $\xi \in \mathfrak{H}$*:*

$$\|\xi\|^2 = \sum_\alpha |\langle \xi, e_\alpha \rangle|^2.$$

If $(e_\alpha)_\alpha$ *satisfies these conditions, it is called an* orthonormal basis *of* $\mathfrak{H}$*.*

Proof (i) $\Longrightarrow$ (ii): Assume that there is $\xi \in \mathfrak{H} \setminus \{0\}$ such that $\xi \perp e_\alpha$ for all α. Then we may add the vector $\frac{\xi}{\|\xi\|}$ to the family $(e_\alpha)_\alpha$ and thus obtain an orthonormal family strictly larger than $(e_\alpha)_\alpha$.

(ii) $\Longrightarrow$ (iii): Assume that $\mathfrak{K} := \overline{\operatorname{lin}\{e_\alpha : \alpha\}} \subsetneq \mathfrak{H}$. By Corollary 2.1.6, there is $\phi \in \mathfrak{H}^* \setminus \{0\}$ such that $\phi|_{\mathfrak{K}} = 0$. By Theorem 3.3.11, there is $\eta \in \mathfrak{H}$ such that

$$\phi(\xi) = \langle \xi, \eta \rangle \qquad (\xi \in \mathfrak{H}).$$

It follows that $0 = \phi(e_\alpha) = \langle e_\alpha, \eta \rangle$, so that $\eta \perp e_\alpha$ for all α and hence $\eta = 0$. This, however, contradicts $\phi \neq 0$.

(iii) $\Longrightarrow$ (iv): Let $\xi \in \mathfrak{H}$, and define $\eta := \xi - \sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha$, which is well defined by Lemma 3.3.17. It follows that

$$\langle \eta, e_\beta \rangle = \langle \xi, e_\beta \rangle - \sum_\alpha \langle \xi, e_\alpha \rangle \langle e_\alpha, e_\beta \rangle = \langle \xi, e_\beta \rangle - \langle \xi, e_\beta \rangle = 0$$

for any index β. Since $\mathfrak{H} = \overline{\operatorname{lin}\{e_\alpha : \alpha\}}$, this implies $\langle \eta, \eta \rangle = 0$ and thus $\eta = 0$.

(iv) $\Longrightarrow$ (v): Let $\xi, \eta \in \mathfrak{H}$. By (iv), we have

$$\xi = \sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha \qquad \text{and} \qquad \eta = \sum_\alpha \langle \eta, e_\alpha \rangle e_\alpha.$$

This implies

$$\begin{aligned} \langle \xi, \eta \rangle &= \left\langle \sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha, \sum_\beta \langle \eta, e_\beta \rangle e_\beta \right\rangle \\ &= \sum_\alpha \sum_\beta \langle \xi, e_\alpha \rangle \overline{\langle \eta, e_\beta \rangle} \langle e_\alpha, e_\beta \rangle \\ &= \sum_\alpha \langle \xi, e_\alpha \rangle \langle e_\alpha, \eta \rangle. \end{aligned}$$

(v) $\Longrightarrow$ (vi): Let $\eta = \xi$.

(vi) $\Longrightarrow$ (i): Let $(f_\beta)_\beta$ be an orthonormal family such that $(e_\alpha)_\alpha$ is a proper subfamily. Then there is one f_{β_0} such that $f_{\beta_0} \perp e_\alpha$ for all α. It follows that

$$\sum_\alpha |\langle f_{\beta_0}, e_\alpha \rangle|^2 = 0 \neq 1 = \|f_{\beta_0}\|^2,$$

which contradicts (vi). $\square$

Corollary 3.3.19 *Let $\mathfrak{H} \neq \{0\}$ be a Hilbert space. Then $\mathfrak{H}$ has an orthonormal basis.*

Proof Use Zorn's lemma to obtain a maximal orthonormal family in $\mathfrak{H}$. $\square$

Exercise 3.14 Let $\mathfrak{H}$ be a Hilbert space, let $\mathfrak{K}$ be a closed subspace, and let $(e_\alpha)_\alpha$ be an orthonormal basis for $\mathfrak{K}$. Show that the orthogonal projection P onto $\mathfrak{K}$ is given by

$$P\xi = \sum_\alpha \langle \xi, e_\alpha \rangle e_\alpha \qquad (\xi \in \mathfrak{H}).$$

Exercise 3.15 Show that every orthonormal basis for a separable, infinite-dimensional Hilbert space is countably infinite.

Lemma 3.3.20 *Let $(e_\alpha)_\alpha$ and $(f_\beta)_\beta$ be orthonormal bases for a Hilbert space $\mathfrak{H}$. Then $(e_\alpha)_\alpha$ and $(f_\beta)_\beta$ have the same cardinality.*

Proof Let κ be the cardinality of $(e_\alpha)_\alpha$, and let λ be the cardinality of $(f_\beta)_\beta$.

Case 1: κ is finite.

In this case, $\dim \mathfrak{H} < \infty$, and $(e_\alpha)_\alpha$ is a Hamel basis. It is easy to see that orthonormal families are always linearly independent. Hence, λ must be finite, too. Since $(f_\beta)_\beta$ spans $\mathfrak{H}$, we have that $(f_\beta)_\beta$ is also a Hamel basis for $\mathfrak{H}$. It follows that $\lambda = \dim \mathfrak{H} = \kappa$.

Case 2: κ is infinite.

By the first case, this means that λ is also infinite. For any index α, define $\mathbb{B}_\alpha := \{\beta : \langle e_\alpha, f_\beta \rangle \neq 0\}$; by Bessel's inequality, $\mathbb{B}_\alpha$ is countable. It follows that

$$\lambda = \left| \bigcup_\alpha \mathbb{B}_\alpha \right| \leq \aleph_0 \cdot \kappa = \kappa.$$

Similarly, one sees that $\kappa \leq \lambda$. $\square$

Theorem 3.3.21 *The following are equivalent for two Hilbert spaces $\mathfrak{H}$ and $\mathfrak{K}$:*

(i) *Every orthonormal basis of $\mathfrak{H}$ has the same cardinality as every orthonormal basis of $\mathfrak{K}$.*

(ii) *There are an orthonormal basis of $\mathfrak{H}$ and an orthonormal basis of $\mathfrak{K}$ having the same cardinality.*

(iii) *There is a surjective operator* $U\colon \mathfrak{H} \to \mathfrak{K}$ *such that*

$$\langle U\xi, U\eta\rangle = \langle \xi, \eta\rangle \qquad (\xi, \eta \in \mathfrak{H}).$$

If these conditions are satisfied, $\mathfrak{H}$ *and* $\mathfrak{K}$ *are called* unitarily equivalent.

Proof (i) $\Longrightarrow$ (ii) is obvious, and (ii) $\Longrightarrow$ (i) follows with Lemma 3.3.20.

(ii) $\Longrightarrow$ (iii): Let $(e_\alpha)_\alpha$ and $(f_\alpha)_\alpha$ be orthonormal bases of $\mathfrak{H}$ and $\mathfrak{K}$, respectively, with the same cardinality, i.e. we may use the same index set. Define

$$U\colon \mathfrak{H} \to \mathfrak{K}, \quad \xi \mapsto \sum_\alpha \langle \xi, e_\alpha\rangle f_\alpha.$$

The same argument as in the proof of Lemma 3.3.17 shows that this is well-defined. We then have for $\xi, \eta \in \mathfrak{H}$:

$$\begin{aligned} \langle U\xi, U\eta\rangle &= \left\langle \sum_\alpha \langle \xi, e_\alpha\rangle f_\alpha, \sum_\beta \langle \eta, e_\beta\rangle f_\beta \right\rangle \\ &= \sum_\alpha \sum_\beta \langle \xi, e_\alpha\rangle \langle e_\beta, \eta\rangle \langle f_\alpha, f_\beta\rangle \\ &= \sum_\alpha \langle \xi, e_\alpha\rangle \langle e_\alpha, \eta\rangle \\ &= \langle \xi, \eta\rangle. \end{aligned}$$

In particular, $\|U\xi\| = \|\xi\|$ holds for each $\xi \in \mathfrak{H}$, so that U is an isometry and thus has closed range.

Assume that U is not surjective. By Corollary 2.1.6 and Theorem 3.3.11, we can then find $\eta \in \mathfrak{K} \setminus \{0\}$ such that $\eta \perp U\mathfrak{H}$. This means in particular that $\eta \perp Ue_\alpha = f_\alpha$ for all α, which is impossible.

(iii) $\Longrightarrow$ (ii): Let $(e_\alpha)_\alpha$ be an orthonormal basis for $\mathfrak{H}$. It follows that $(Ue_\alpha)_\alpha$ is orthonormal in $\mathfrak{K}$ such that $\mathfrak{K} = \overline{\{Ue_\alpha : \alpha\}}$. Hence, $(Ue_\alpha)_\alpha$ is an orthonormal basis for $\mathfrak{K}$. Clearly, $(e_\alpha)_\alpha$ and $(Ue_\alpha)_\alpha$ have the same cardinality. $\square$

Corollary 3.3.22 *Let* $\mathfrak{H} \neq \{0\}$ *be a Hilbert space. Then* $\mathfrak{H}$ *is unitarily equivalent to* $\ell^2(I)$ *for an appropriate index set* $I \neq \varnothing$.

Corollary 3.3.23 *Up to unitary equivalence, there is only one separable, infinite-dimensional Hilbert space.*

Exercise 3.16 Show that the Hilbert spaces ℓ^2, $L^2(\mathbb{R})$, are $L^2([0,1])$ all separable and thus unitarily equivalent.

Exercise 3.17 Let $\mathfrak{H}$ be a Hilbert space, let $T \in \mathcal{K}(\mathfrak{H})$, and let $(e_n)_{n=1}^\infty$ be an orthonormal sequence in $\mathfrak{H}$. Show that $\|Te_n\| \to 0$.

3.3.4 Operators on Hilbert spaces

As we say in the previous subsection, Hilbert spaces are essentially "dull" objects. This dullness, however, forces their bounded linear operators to be much more tractable than in a general Banach space setting.

Theorem 3.3.24 *Let $\mathfrak{H}$ and $\mathfrak{K}$ be Hilbert spaces, and let $T \in \mathcal{B}(\mathfrak{H}, \mathfrak{K})$. Then there is a unique operator $T^* \in \mathcal{B}(\mathfrak{K}, \mathfrak{H})$ such that*

$$\langle T\xi, \eta \rangle = \langle \xi, T^*\eta \rangle \qquad (\xi \in \mathfrak{H},\ \eta \in \mathfrak{K}).$$

The operator T^* is called the adjoint *of T.*

Exercise 3.18 Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{B}(\mathfrak{H})$. Show that $\ker T = (T^*\mathfrak{H})^\perp$.

Proof For fixed $\eta \in \mathfrak{K}$, define $\phi \in \mathfrak{H}^*$ by letting

$$\phi(\xi) := \langle T\xi, \eta \rangle \qquad (\xi \in \mathfrak{H}).$$

By Theorem 3.3.11, there is a unique $T^* \in \mathfrak{H}$ such that

$$\langle T\xi, \eta \rangle = \phi(\xi) = \langle \xi, T^*\eta \rangle \qquad (\xi \in \mathfrak{H}).$$

It is easy to see that $\mathfrak{K} \ni \eta \mapsto T^*\eta$ is a bounded, linear operator from $\mathfrak{K}$ to $\mathfrak{H}$. $\square$

Remark Despite the use of the same symbol, the adjoint of T is not to be confused with the transpose defined earlier for operators between Banach spaces. Nevertheless, in many ways, taking the adjoint operator behaves very much like taking the transpose.

Examples 1. Let $\mathfrak{H} = \mathbb{F}^N$, let $\mathfrak{K} = \mathbb{F}^M$, and let $T = T_A$ for

$$A = \begin{bmatrix} a_{1,1}, & \cdots, & a_{1,N} \\ \vdots & & \vdots \\ a_{M,1}, & \cdots, & a_{M,N} \end{bmatrix}.$$

Then $T^* = T_{A^*}$, where

$$A^* = \begin{bmatrix} \overline{a_{1,1}}, & \cdots, & \overline{a_{M,1}} \\ \vdots & & \vdots \\ \overline{a_{1,N}}, & \cdots, & \overline{a_{M,N}} \end{bmatrix}.$$

2. Let $(\Omega, \mathfrak{S}, \mu)$ be a σ-finite measure space, and let $\phi \in L^\infty(\Omega, \mathfrak{S}, \mu)$. Then $M_\phi^* = M_{\overline{\phi}}$.

Proposition 3.3.25 *Let $\mathfrak{H}$ be a Hilbert space, let $T, S \in \mathcal{B}(\mathfrak{H})$, and let $\lambda, \mu \in \mathbb{F}$. Then we have:*

(i) $(\lambda T + \mu S)^* = \overline{\lambda} T^* + \overline{\mu} S^*$;

(ii) $(ST)^* = T^* S^*$;

(iii) $T^{**} = T$;

(iv) $\|T\|^2 = \|T^*\|^2 = \|T^* T\|$.

Proof (i), (ii), and (iii), are straightforward.

For (iv), let $\xi \in \mathfrak{H}$ be such that $\|\xi\| \leq 1$. It follows that

$$\begin{aligned} \|T\xi\|^2 &= \langle T\xi, T\xi \rangle \\ &= \langle T^* T\xi, \xi \rangle \\ &= \|T^* T\xi\| \|\xi\| \\ &\leq \|T^* T\| \\ &\leq \|T^*\| \|T\| \end{aligned}$$

and hence

$$\|T\|^2 \leq \|T^* T\| \leq \|T^*\| \|T\|.$$

It follows, in particular, that $\|T\| \leq \|T^*\|$. On the other hand, (iii) yields that $\|T^*\| \leq \|T^{**}\| = \|T\|$. Hence, $\|T\| = \|T^*\|$ holds and also $\|T\|^2 = \|T^* T\|$. $\square$

Definition 3.3.26 Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{B}(\mathfrak{H})$. Then:

(a) T is *self-adjoint* if $T = T^*$.

(b) T is *normal* if $T^* T = T T^*$.

Exercise 3.19 Let $\mathfrak{H}$ be a Hilbert space. Show that a projection $P \in \mathcal{B}(\mathfrak{H})$ is self-adjoint if and only if it is an orthogonal projection.

Exercise 3.20 Let $\mathfrak{H}$ be a Hilbert space, and let $T \colon \mathfrak{H} \to \mathfrak{H}$ be linear such that

$$\langle T\xi, \eta \rangle = \langle \xi, T\eta \rangle \qquad (\xi, \eta \in \mathfrak{H}).$$

Show that T is bounded and self-adjoint.

Exercise 3.21 Let $\mathfrak{H}$ be a Hilbert space over $\mathbb{C}$. Show that $N \in \mathcal{B}(\mathfrak{H})$ is normal if and only if $\|N\xi\| = \|N^*\xi\|$ for all $\xi \in \mathfrak{H}$.

Exercise 3.22 Let $\mathfrak{H}$ be a Hilbert space over $\mathbb{C}$, let $N \in \mathcal{B}(\mathfrak{H})$ be normal, and let $\xi \in \mathfrak{H}$ and $\lambda \in \mathbb{C}$ be such that $N\xi = \lambda\xi$. Show that $N^*\xi = \bar{\lambda}\xi$.

Exercise 3.23 Let $\mathfrak{H}$ be a Hilbert space over $\mathbb{C}$. For $T \in \mathcal{B}(\mathfrak{H})$ define

$$\operatorname{Re} T := \frac{1}{2}(T + T^*) \qquad \text{and} \qquad \operatorname{Im} T = \frac{1}{2i}(T - T^*).$$

Show that $N \in \mathcal{B}(\mathfrak{H})$ is normal if and only if $\operatorname{Re} N$ and $\operatorname{Im} N$ commute.

Proposition 3.3.27 *Let $\mathfrak{H}$ be a $\mathbb{C}$-Hilbert space. Then the following are equivalent for $T \in \mathcal{B}(\mathfrak{H})$:*

(i) *T is self-adjoint.*

(ii) *$\langle T\xi, \xi\rangle \in \mathbb{R}$ for $\xi \in \mathfrak{H}$.*

Proof (i) $\Longrightarrow$ (ii) is clear because

$$\langle T\xi, \xi\rangle = \langle \xi, T\xi\rangle = \overline{\langle T\xi, \xi\rangle} \qquad (\xi \in \mathfrak{H}).$$

(ii) $\Longrightarrow$ (i): Let $\xi, \eta \in \mathfrak{H}$, and let $\lambda \in \mathbb{C}$. Then

$$\langle T(\xi + \lambda\eta), \xi + \lambda\eta\rangle = \langle T\xi, \xi\rangle + \overline{\lambda}\langle T\xi, \eta\rangle + \lambda\langle T\eta, \xi\rangle + |\lambda|^2\langle T\eta, \eta\rangle$$

is real and thus equal to its complex conjugate. This, in turn, implies that

$$\begin{aligned} \lambda\langle T\eta, \xi\rangle + \overline{\lambda}\langle T\xi, \eta\rangle &= \overline{\lambda}\langle \xi, T\eta\rangle + \lambda\langle \eta, T\xi\rangle \\ &= \overline{\lambda}\langle T^*\xi, \eta\rangle + \lambda\langle T^*\eta, \xi\rangle. \end{aligned}$$

Letting $\lambda = 1$ and $\lambda = i$, respectively, we obtain:

$$\begin{cases} \langle T\eta, \xi\rangle + \langle T\xi, \eta\rangle = \langle T^*\xi, \eta\rangle + \langle T^*\eta, \xi\rangle, \\ i\langle T\eta, \xi\rangle - i\langle T\xi, \eta\rangle = -i\langle T^*\xi, \eta\rangle + i\langle T^*\eta, \xi\rangle. \end{cases}$$

Dividing the second of those equations by i, we get

$$\begin{cases} \langle T\eta, \xi\rangle + \langle T\xi, \eta\rangle = \langle T^*\xi, \eta\rangle + \langle T^*\eta, \xi\rangle, \\ \langle T\eta, \xi\rangle - \langle T\xi, \eta\rangle = -\langle T^*\xi, \eta\rangle + \langle T^*\eta, \xi\rangle, \end{cases}$$

and adding them yields $2\langle T\eta, \xi\rangle = 2\langle T^*\eta, \xi\rangle$, i.e. $T = T^*$. $\square$

Proposition 3.3.28 *Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{B}(\mathfrak{H})$ be self-adjoint. Then*

$$\|T\| = \sup\{|\langle T\xi, \xi\rangle| : \xi \in \mathfrak{H},\ \|\xi\| \le 1\}. \tag{3.9}$$

Proof Let M denote the supremum in (3.9). It is clear that $M \le \|T\|$.

Let $\xi, \eta \in \mathfrak{H}$ with $\|\xi\|, \|\eta\| \le 1$. It follows that

$$\begin{aligned} \langle T(\xi \pm \eta), \xi \pm \eta\rangle &= \langle T\xi, \xi\rangle \pm \langle T\xi, \eta\rangle \pm \langle T\eta, \xi\rangle + \langle T\eta, \eta\rangle \\ &= \langle T\xi, \xi\rangle \pm \langle T\xi, \eta\rangle \pm \langle \eta, T^*\xi\rangle + \langle T\eta, \eta\rangle \\ &= \langle T\xi, \xi\rangle \pm 2\mathrm{Re}\,\langle T\xi, \eta\rangle + \langle T\eta, \eta\rangle. \end{aligned}$$

Subtraction yields

$$4\mathrm{Re}\,\langle T\xi,\eta\rangle = \langle T(\xi+\eta),\xi+\eta\rangle - \langle T(\xi-\eta),\xi-\eta\rangle.$$

It follows that

$$\begin{aligned} 4\mathrm{Re}\,\langle T\xi,\eta\rangle &\leq M(\|\xi+\eta\|^2 + \|\xi-\eta\|^2) \\ &= 2M(\|\xi\|^2 + \|\eta\|^2) \\ &\leq 4M. \end{aligned}$$

Choose $\lambda \in \mathbb{F}$ with $|\lambda| = 1$ such that $\langle T\xi,\eta\rangle = \lambda|\langle T\xi,\eta\rangle|$. Replacing ξ in the previous argument by $\overline{\lambda}\xi$ yields

$$|\langle T\xi,\eta\rangle| = \overline{\lambda}\langle T\xi,\eta\rangle = \langle T(\overline{\lambda}\xi),\eta\rangle \leq M.$$

For any $\xi \in \mathfrak{H}$ with $\|\xi\| \leq 1$, we thus have

$$\|T\xi\| = \sup\{|\langle T\xi,\eta\rangle| : \eta \in \mathfrak{H},\ \|\eta\| \leq 1\} \leq M$$

and therefore $\|T\| \leq M$. $\square$

Corollary 3.3.29 *Let $\mathfrak{H}$ be a $\mathbb{C}$-Hilbert space, and let $T \in \mathcal{B}(\mathfrak{H})$ be such that $\langle T\xi,\xi\rangle = 0$ for all $\xi \in \mathfrak{H}$. Then T is zero.*

Proof First, note that

$$\langle T^*\xi,\xi\rangle = \langle \xi, T\xi\rangle = \overline{\langle T\xi,\xi\rangle} \qquad (\xi \in \mathfrak{H}).$$

Let

$$\mathrm{Re}\,T := \frac{1}{2}(T+T^*) \qquad \text{and} \qquad \mathrm{Im}\,T := \frac{1}{2i}(T-T^*).$$

Then $\mathrm{Re}\,T$ and $\mathrm{Im}\,T$ are self-adjoint such that $T = \mathrm{Re}\,T + i\mathrm{Im}\,T$ and $\langle S\xi,\xi\rangle = 0$ for all $\xi \in \mathfrak{H}$, where $S = \mathrm{Re}\,T$ or $S = \mathrm{Im}\,T$. By Proposition 3.3.28, this means $\mathrm{Re}\,T = \mathrm{Im}\,T = 0$ and thus $T = 0$. $\square$

Remark Proposition 3.3.28 is false for Hilbert spaces over $\mathbb{R}$ (take $\mathfrak{H} = \mathbb{R}$ and $T = T_A$, where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$).

Exercise 3.24 Let $\mathfrak{H}$ be a Hilbert space, and let $\mathcal{F}(\mathfrak{H})$ be the family of all finite-dimensional subspaces of $\mathfrak{H}$. For each $\mathfrak{K} \in \mathcal{F}(\mathfrak{H})$, let $P_{\mathfrak{K}}$ be the orthogonal projection onto $\mathfrak{K}$. Show that $(P_{\mathfrak{K}})_{\mathfrak{K}\in\mathcal{F}(\mathfrak{H})}$ is a net such that

$$\|P_{\mathfrak{K}}T - T\| \to 0 \qquad \text{and} \qquad \|TP_{\mathfrak{K}} - T\| \to 0$$

for all $T \in \mathcal{K}(\mathfrak{H})$.

3.4 The spectral theorem for compact, self-adjoint operators

Throughout this section, all spaces are again over $\mathbb{C}$.

In linear algebra, it is shown that a self-adjoint matrix can be diagonalized. This means that such a matrix can be completely described once its eigenvalues and its eigenspaces are known. In this section, we extend this theorem on matrices to compact, self-adjoint operators on Hilbert space.

Lemma 3.4.1 *Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{K}(\mathfrak{H})$ be self-adjoint. Then $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof By Proposition 3.3.28, there is a sequence $(\xi_n)_{n=1}^\infty$ in $\mathfrak{H}$ such that $\|\xi_n\| = 1$ for all $n \in \mathbb{N}$ and $|\langle T\xi_n, \xi_n\rangle| \to \|T\|$. Passing to a subsequence, we may suppose that $(\langle T\xi_n, \xi_n\rangle)_{n=1}^\infty$ converges to $\lambda \in \mathbb{R}$. It follows necessarily that $|\lambda| = \|T\|$. Note that

$$\begin{aligned} \|(\lambda - T)\xi_n\|^2 &= \lambda^2 - 2\lambda\langle T\xi_n, \xi_n\rangle + \|T\xi_n\|^2 \\ &\leq 2\lambda^2 - 2\lambda\langle T\xi_n, \xi_n\rangle \\ &\to 0. \end{aligned}$$

With Lemma 3.2.1, we conclude that λ is an eigenvalue of T. $\square$

Corollary 3.4.2 *Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{K}(\mathfrak{H})$ be self-adjoint such that $\sigma(T) = \{0\}$. Then $T = 0$.*

Lemma 3.4.3 *Let $\mathfrak{H}$ be a Hilbert space, let $N \in \mathcal{K}(\mathfrak{H})$ be normal, and let $\lambda \neq \mu$ be eigenvalues of N. Then $\ker(\lambda - N) \perp \ker(\mu - N)$.*

Proof Let $\xi \in \ker(\lambda - N)$ and $\eta \in \ker(\mu - N)$, and note that

$$\lambda\langle \xi, \eta\rangle = \langle \lambda\xi, \eta\rangle = \langle N\xi, \eta\rangle = \langle \xi, N^*\eta\rangle = \langle \xi, \overline{\mu}\eta\rangle = \mu\langle \xi, \eta\rangle.$$

It follows that $\langle \xi, \eta\rangle = 0$. $\square$

Lemma 3.4.4 *Let $\mathfrak{H}$ be a Hilbert space, and let $T \in \mathcal{K}(\mathfrak{H})$ be self-adjoint. Then $\sigma(T) \subset \mathbb{R}$.*

Proof Let $\lambda \in \sigma(T) \setminus \{0\}$, so that λ is an eigenvalue of T. Hence, there is $\xi \in \mathfrak{H} \setminus \{0\}$ such that $T\xi = \lambda\xi$. It follows that $T\xi = T^*\xi = \overline{\lambda}\xi$ and thus $\xi \in \ker(\lambda - T) \cap \ker(\overline{\lambda} - T)$. By Lemma 3.4.3, this is possible only if $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$. $\square$

Theorem 3.4.5 (spectral theorem for compact, self-adjoint operators) *Let $\mathfrak{H}$ be a Hilbert space, let $T \in \mathcal{K}(\mathfrak{H})$ be self-adjoint, let $\{\lambda_1, \lambda_2, \ldots\}$ be the distinct, non-zero eigenvalues of T, and let P_n denote the orthogonal projection onto $\ker(\lambda_n - T)$. Then the following hold true:*

(i) $\{\lambda_1, \lambda_2, \dots\} \subset \mathbb{R}$;

(ii) $P_n P_m = P_m P_n = 0 \qquad (n \neq m)$;

(iii) $T = \sum_{n=1}^{\leq\infty} \lambda_n P_n$.

Proof (i) follows from Lemma 3.4.4.

For (ii), let $\xi \in \mathfrak{H}$, and note that, by Lemma 3.4.3, for $n \neq m$

$$P_n \xi \in \ker(\lambda_n - T) \subset \ker(\lambda_m - T)^\perp = \ker P_m$$

and thus $P_m P_n \xi = 0$. It follows that $P_m P_n = 0$.

(iii): Choose an eigenvalue λ_1 of T such that $|\lambda_1| = \|T\|$ (this is possible by Lemma 3.4.1). Let $\mathfrak{H}_1 := \ker(\lambda_1 - T)$ and let P_1 denote the orthogonal projection onto $\mathfrak{H}_1$. Let $\mathfrak{K}_2 := \mathfrak{H}_1^\perp (= \ker P_1)$. Let $\xi \in \mathfrak{K}_2$ and $\eta \in \mathfrak{H}_1$ and note that

$$\langle T\xi, \eta \rangle = \langle \xi, T\eta \rangle = \langle \xi, \lambda_1 \eta \rangle = \lambda_1 \langle \xi, \eta \rangle = 0.$$

It follows that $T\mathfrak{K}_2 \subset \mathfrak{K}_2$. If $T|_{\mathfrak{K}_2} = 0$, finish. Otherwise — since $T_2 := T|_{\mathfrak{K}_2}$ is also compact — and self-adjoint, there is an eigenvalue λ_2 of T_2 such that $|\lambda| = \|T_2\|$. Let $\mathfrak{H}_2 := \ker(\lambda_2 - T_2)$. It is easy to see that $\mathfrak{H}_2 = \ker(\lambda_2 - T)$. Since $\mathfrak{H}_1 \perp \mathfrak{H}_2$ by definition, we have $\lambda_1 \neq \lambda_2$. Let P_2 be the orthogonal projection onto $\mathfrak{H}_2$, and let $\mathfrak{K}_3 := (\mathfrak{H}_1 \oplus \mathfrak{H}_2)^\perp$.

Continue inductively and obtain:

(a) A (possibly finite) sequence $\{\lambda_1, \lambda_2, \dots\}$ of distinct eigenvalues of T which satisfies $|\lambda_1| \geq |\lambda_2| \geq \cdots$.

(b) A sequence of pairwise orthogonal closed subspaces $\mathfrak{H}_n$ of $\mathfrak{H}$ such that

$$\mathfrak{H}_n = \ker(\lambda_n - T) \qquad \text{and} \qquad |\lambda_{n+1}| = \|T|_{(\mathfrak{H}_1 \oplus \cdots \oplus \mathfrak{H}_n)^\perp}\|.$$

Fix $n \in \mathbb{N}$, and let $\xi \in \mathfrak{H}_k$ with $k \in \{1, \dots, n\}$. It follows that

$$T\xi - \sum_{j=1}^{n} \lambda_j P_j \xi = \lambda_k \xi - \lambda_k \xi = 0.$$

Let $\xi \in (\mathfrak{H}_1 \oplus \cdots \oplus \mathfrak{H}_n)^\perp$. Then $P_k \xi = 0$ for $k = 1, \dots, n$ and thus

$$T\xi - \sum_{j=1}^{n} \lambda_j P_j \xi = T\xi \in (\mathfrak{H}_1 \oplus \cdots \oplus \mathfrak{H}_n)^\perp.$$

It follows that

$$\begin{aligned}\left\|T-\sum_{j=1}^{n}\lambda_j P_j\right\| &= \sup\left\{\left\|T\xi-\sum_{j=1}^{n}\lambda_j P_j\xi\right\| : \xi\in\mathfrak{H},\ \|\xi\|\le 1\right\}\\ &= \sup\left\{\left\|T\xi-\sum_{j=1}^{n}\lambda_j P_j\xi\right\| : \xi\in(\mathfrak{H}_1\oplus\cdots\oplus\mathfrak{H}_n)^\perp,\ \|\xi\|\le 1\right\}\\ &= \sup\left\{\|T\xi\| : \xi\in(\mathfrak{H}_1\oplus\cdots\oplus\mathfrak{H}_n)^\perp,\ \|\xi\|\le 1\right\}\\ &= |\lambda_{n+1}|\\ &\to 0\end{aligned}$$

and thus

$$T=\sum_{n=1}^{\le\infty}\lambda_n P_n.$$

Assume that T has an eigenvalue λ which does not occur in the sequence $\{\lambda_1,\lambda_2,\ldots\}$. Since $\ker(\lambda-T)\perp\ker(\lambda_n-T)$ for $n\in\mathbb{N}$, this means that

$$T\xi=\sum_{n=1}^{\le\infty}\lambda_n P_n\xi=0,$$

i.e. $\lambda=0$. $\square$

With the exception of (i), the assertions of Theorem 3.4.5 still hold for compact, normal operators.

Exercise 3.25 Read paragraphs II.6 and II.7 in Conway's book (where the spectral theorem for compact, normal operators is proven).

Remark The spectral theorem generalizes to arbitrary, not necessarily compact, bounded, normal operators on Hilbert spaces. Since such operators need no longer have eigenvalues, the projections onto the eigenspaces have to be replaces by a more general object: Given a Hilbert space $\mathfrak{H}$ and a normal operator $N\in\mathcal{B}(\mathfrak{H})$, there is a unique so-called spectral measure E — a measure whose values are orthogonal projections on $\mathfrak{H}$ — on $\sigma(N)$ such that

$$N=\int_{\sigma(N)} z\,dE(z).$$

Chapter 4

Fixed point theorems and locally convex spaces

This last chapter of the lecture notes deals with fixed point theorems, i.e. theorems that guarantee that a certain map of a certain family of maps has a fixed point. Fixed point theorems are important because many problems concerning the solvability of equations can be formulated as fixed point problems. In order to present these fixed point theorems in sufficient generality, we develop the theory of locally convex vector spaces to some extent.

4.1 Banach's fixed point theorem

Banach's fixed point theorem is one of the most elegant and most widely applicable theorems in all of analysis:

Theorem 4.1.1 (Banach's fixed point theorem) *Let X be a complete metric space, and let $T\colon X \to X$ be a map such that, for some $\theta \in (0,1)$,*

$$d(T(x), T(y)) \leq \theta d(x,y) \qquad (x, y \in X).$$

Then T has a unique fixed point in X.

Proof Let $x, y \in X$ be fixed points of X. Since

$$d(x,y) = d(T(x), T(y)) \leq \theta d(x,y)$$

and $\theta \in (0,1)$, it follows that $d(x,y) = 0$, i.e. $x = y$. This proves the uniqueness of the fixed point.

For the proof of existence, choose $x_0 \in X$ arbitrary, and define inductively $x_n := T(x_{n-1})$ for $n \in \mathbb{N}$. We claim that the sequence $(x_n)_{n=0}^{\infty}$ converges. An easy induction on

n shows that

$$d(x_{n-1}, x_n) \leq \theta^{n-1} d(x_0, x_1) \qquad (n \in \mathbb{N}).$$

For $n > m$, we have

$$d(x_m, x_n) \leq \sum_{k=m+1}^{n} d(x_{k-1}, x_k) \leq d(x_0, x_1) \sum_{k=m+1}^{n} \theta^{k-1}. \tag{4.1}$$

Since $\theta \in (0,1)$, the geometric series $\sum_{n=1}^{\infty} \theta^{n-1}$ converges. Hence, given $\epsilon > 0$, there is $N \in \mathbb{N}$ such that

$$\sum_{k=m+1}^{n} \theta^{k-1} < \frac{\epsilon}{d(x_0, x_1) + 1} \qquad (n, m \geq N).$$

Together with (4.1), this shows that $d(x_m, x_n) < \epsilon$ for all $n, m \geq N$. Hence, $(x_n)_{n=0}^{\infty}$ is a Cauchy sequence, and $x := \lim_{n \to \infty} x_n$ exists. Since T is clearly (uniformly) continuous, we have

$$Tx = \lim_{n \to \infty} Tx_n = \lim_{n \to \infty} Tx_{n-1} = \lim_{n \to \infty} x_n = x,$$

i.e. x is a fixed point of T. $\square$

Remark Banach's fixed point theorem not only guarantees that a fixed point exsists — it also provides an effective way of computing such a fixed point along with error estimates.

Exercise 4.1 Is the following fixed point theorem true or not?

> *Let X be a complete metric space, and let $T\colon X \to X$ be a map such that*
>
> $$d(T(x), T(y)) < d(x, y) \qquad (x, y \in X,\, x \neq y).$$
>
> *Then T has a unique fixed point in X.*

Give a proof or a counterexample.

Exercise 4.2 Let (X, d) be a complete metric space, and let $T\colon X \to X$ be a map such that there are $\theta \in (0,1)$ and $n \in \mathbb{N}$ with

$$d(T^n(x), T^n(y)) \leq \theta\, d(x, y) \qquad (x, y \in X).$$

Show that T has a unique fixed point.

We apply Banach's fixed point theorem to initial value problems:

Theorem 4.1.2 (Picard–Lindelöf theorem) *Let $I = [a, b]$, and let $f : I \times \mathbb{R} \to \mathbb{R}$ be continuous such that there is $C \geq 0$ with*

$$|f(x, y_1) - f(x, y_2)| \leq C|y_1 - y_2| \qquad (x \in I,\, y_1, y_2 \in \mathbb{R}).$$

Then the initial value problem

$$y' = f(x, y), \quad y(a) = y_0 \tag{4.2}$$

has a unique solution $\phi \in \mathcal{C}^1(I)$ for each $y_0 \in \mathbb{R}$.

Proof If (4.2) has a solution ϕ, it satisfies

$$\phi(x) = \int_a^x f(t, \phi(t))\, dt + y_0. \tag{4.3}$$

Conversely, every function $\phi \in \mathcal{C}(I)$ satisfying (4.3) is automatically in $\mathcal{C}^1(I)$ and a solution of (4.2).

Define $T : \mathcal{C}(I) \to \mathcal{C}(I)$ by letting

$$(T\phi)(x) \int_a^x f(t, \phi(t))\, dt + y_0.$$

Then $\phi \in \mathcal{C}(I)$ solves (4.3) and hence (4.2) if and only if it is a fixed point of T.

Let $\phi_1, \phi_2 \in \mathcal{C}(I)$. Then we have

$$\begin{aligned} |(T\phi_1)(x) - (T\phi_2)(x)| &= \left| \int_a^x (f(t, \phi_1(t)) - f(t, \phi_2(t)))\, dt \right| \\ &\leq C \int_a^x |\phi_1(t) - \phi_2(t)|\, dt \\ &= C(b-a)\|\phi_1 - \phi_2\|_\infty \qquad (x \in I), \end{aligned}$$

and therefore

$$\|T\phi_1 - T\phi_2\|_\infty \leq C(b-a)\|\phi_1 - \phi_2\|_\infty.$$

The problem that arises at this point is that $C(b-a)$ may not belong to $(0, 1)$, so that Theorem 4.1.1 is not directly applicable. We circumvent this problem by replacing $\|\cdot\|_\infty$ by an equivalent norm $\|\cdot\|_\alpha$ for some parameter $\alpha > 0$. Define for $\alpha > 0$

$$\|\phi\|_\alpha := \sup\{|\phi(x)|e^{-\alpha x} : x \in I\} \qquad (\phi \in \mathcal{C}(I)).$$

Then, for $\phi \in \mathcal{C}(I)$,

$$\|\phi\|_\alpha \leq \|\phi\|_\infty \sup\{e^{-\alpha x} : x \in I\}$$

and

$$\|\phi\|_\infty = \sup\{|\phi(x)|e^{-\alpha x}e^{\alpha x} : x \in I\} \leq \|\phi\|_\alpha \sup\{e^{\alpha x} : x \in I\},$$

so that $\|\cdot\|_\infty$ and $\|\cdot\|_\alpha$ are equivalent. Moreover, we have for $\phi_1, \phi_2 \in \mathcal{C}(I)$ and $x \in I$:

$$\begin{aligned}
|(T\phi_1)(x) - (T\phi_2)(x)|e^{-\alpha x} &= e^{-\alpha x}\left|\int_a^x (f(t,\phi_1(t)) - f(t,\phi_2(t)))\,dt\right| \\
&\leq Ce^{-\alpha x}\int_a^x |\phi_1(t) - \phi_2(t)|\,dt \\
&= Ce^{-\alpha x}\int_a^x |\phi_1(t) - \phi_2(t)|e^{-\alpha t}e^{\alpha t}\,dt \\
&\leq Ce^{-\alpha x}\|\phi_1 - \phi_2\|_\alpha \int_a^x e^{\alpha t}\,dt \\
&\leq Ce^{-\alpha x}\|\phi_1 - \phi_2\|_\alpha \frac{e^{\alpha x}}{\alpha} \\
&= \frac{C}{\alpha}\|\phi_1 - \phi_2\|_\alpha.
\end{aligned}$$

It follows that

$$\|T\phi_1 - T\phi_2\|_\alpha \leq \frac{C}{\alpha}\|\phi_1 - \phi_2\|_\alpha \qquad (\phi_1, \phi_2 \in \mathcal{C}(I)).$$

Choosing $\alpha > 0$ so large that $\frac{C}{\alpha} < 1$ and applying Banach's fixed point theorem, we obtain a unique fixed point of T and thus a unique solution of (4.2). $\square$

4.2 Locally convex vector spaces

The next fixed point theorem we are going to cover — Schauder's fixed point theorem — requires a compactness hypothesis for its domain. Since compactness in normed, infinite-dimensional spaces is rather the exception than the rule, we have to leave the framework of normed spaces and work in a more general context in order to obtain fixed point theorems of sufficient generality.

Definition 4.2.1 A linear space E is called *locally convex* if it is equipped with a family $\mathcal{P}$ of seminorms on E such that

$$\bigcap_{p \in \mathcal{P}} \{x \in E : p(x) = 0\} = \{0\}.$$

Example Let X be a topological space, and let $\mathcal{C}(X)$ denote the vector space of *all* continuous functions on X. Let $\mathcal{K}$ be the collection of all compact subsets of X. For $K \in \mathcal{K}$, define

$$p_K(f) := \sup\{|f(x)| : x \in K\} \qquad (f \in \mathcal{C}(X)).$$

Then $\mathcal{C}(X)$ equipped with $(p_K)_{K \in \mathcal{K}}$ is a locally convex vector space.

Definition 4.2.2 Let E be a locally convex vector space. A subset U of E is defined as *open* if, for each $x_0 \in U$, there are $\epsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{P}$ such that

$$\left\{ x \in E : \max_{j=1,\ldots,n} p_j(x - x_0) < \epsilon \right\} \subset U.$$

Proposition 4.2.3 *Let E be a locally convex vector space. Then the collection of open subsets of E in the sense of Definition* 4.2.2 *is a topology, i.e.*

(i) *$\varnothing$ and E are open;*

(ii) *if $(U_\alpha)_\alpha$ is a family of open sets, then $\bigcup_\alpha U_\alpha$ is open;*

(iii) *if $U_1, \ldots, U_n$ are open, then so is $U_1 \cap \cdots \cap U_n$.*

Proof We only prove (iii).

Let $x_0 \in U_1 \cap \cdots \cap U_n$. For each $k = 1, \ldots, n$, there are $\epsilon_k > 0$ and $p_1^{(k)}, \ldots, p_{n_k}^{(k)} \in \mathcal{P}$ such that

$$V_k := \left\{ x \in E : \max_{j=1,\ldots,n_k} p_j^{(k)}(x - x_0) < \epsilon_k \right\} \subset U_k.$$

Let $\epsilon := \min\{\epsilon_1, \ldots, \epsilon_k\}$, and note that

$$\left\{ x \in E : \max_{k=1,\ldots,n} \max_{j=1,\ldots,n_k} p_j^{(k)}(x - x_0) < \epsilon \right\} \subset V_1 \cap \cdots \cap V_n \subset U_1 \cap \cdots \cap U_n.$$

By Definition 4.2.2, this means that $U_1 \cap \cdots \cap U_n$ is open. $\square$

Proposition 4.2.4 *Let E be a locally convex vector space. Then a net $(x_\alpha)_\alpha$ in E converges to $x_0 \in E$ in the topology if and only if $p(x_\alpha - x_0) \to 0$ for each $p \in \mathcal{P}$.*

Proof Suppose that $x_\alpha \to x_0$ in the topology. Fix $\epsilon > 0$ and $p \in \mathcal{P}$. Then $U := \{x \in E : p(x - x_0) < \epsilon\}$ is an open neighborhood of x_0. Hence, there is an index α_0 such that $x_\alpha \in U$, i.e. $p(x_\alpha - x_0) < \epsilon$ for all $\alpha \succ \alpha_0$. Hence, $p(x_\alpha - x_0) \to 0$.

Conversely, suppose that $p(x_\alpha - x_0) \to 0$ for all $p \in \mathcal{P}$. Let U be a neighborhood of x_0, i.e. there is an open set $V \subset U$ with $x_0 \in V$. By Definition 4.2.2, there are $\epsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{P}$ such that

$$\left\{ x \in E : \max_{j=1,\ldots,n} p_j(x - x_0) < \epsilon \right\} \subset V.$$

Since $p_j(x_\alpha - x_0) \to 0$ for $j = 1, \ldots, n$, there is an index α_0 such that

$$p_j(x_\alpha - x_0) < \epsilon \qquad (j = 1, \ldots, n, \, \alpha \succ \alpha_0).$$

This means, however, that $x_\alpha \in V \subset U$ for all $\alpha \succ \alpha_0$. $\square$

Exercise 4.3 Let E be a locally convex vector space, and let F be a finite-dimensional subspace. Show that the relative topology on F is induced by a norm.

4.2.1 Weak and weak* topologies

There is a canonical locally convex topology on each normed space:

Definition 4.2.5 Let E be a normed space. Then $\{p_\phi : \phi \in E^*\}$ with

$$p_\phi(x) = |\phi(x)| \qquad (x \in E,\ \phi \in E^*)$$

is a family of seminorms on E such that $\bigcap_{\phi \in E^*}\{x \in E : p_\phi(x) = 0\}$. The corresponding topology on E is called the *weak topology* on E.

Lemma 4.2.6 *Let E be a linear space, and let $\phi, \phi_1, \ldots, \phi_n : E \to \mathbb{F}$ be linear. Then the following are equivalent:*

(i) *there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that $\phi = \lambda_1\phi_1 + \cdots + \lambda_n\phi_n$;*

(ii) $\bigcap_{j=1}^n \ker\phi_j \subset \ker\phi$.

Exercise 4.4 Prove Lemma 4.2.6.

Theorem 4.2.7 *Let E be a normed space. Then the following are equivalent:*

(i) $\dim E < \infty$*;*

(ii) *the weak topology and the norm topology coincide;*

(iii) *the weak topology is metrizable.*

Proof (i) $\Longrightarrow$ (ii): If $\dim E < \infty$, then $\dim E^* < \infty$. Let $\phi_1, \ldots, \phi_n \in E^*$ be a Hamel basis for E^*. Define

$$|x| := \max_{j=1,\ldots,n} |\phi_j(x)| \qquad (x \in E).$$

Then E is a norm on E, so that $|\cdot| \sim \|\cdot\|$. Let $U \subset E$ be norm open. This means that, for every $x_0 \in U$, there is $\epsilon > 0$ such that

$$\{x \in E : |x - x_0| < \epsilon\} = \left\{x \in E : \max_{j=1,\ldots,n} |\phi_j(x - x_0)| < \epsilon\right\} \subset U.$$

From the definition of the weak topology, it follows that U is also weakly open. Since the weak topology is coarser than the norm topology, every weakly open subset of E is norm open.

(ii) $\Longrightarrow$ (iii) is trivial.

(iii) $\Longrightarrow$ (i): Suppose that there is a metric d on E which induces the weak topology. Hence, for all $n \in \mathbb{N}$,

$$U_n := \left\{x \in E : d(x, 0) < \frac{1}{n}\right\}$$

is weakly open. By the definition of the weak topology, there are, for each $n \in \mathbb{N}$, a number $\epsilon_n > 0$ and functionals $\phi_1^{(n)}, \ldots, \phi_{k_n}^{(n)} \in E^*$ such that

$$\left\{ x \in E : \max_{j=1,\ldots,k_n} \left| \phi_j^{(n)}(x - x_0) \right| < \epsilon_n \right\} \subset U_n.$$

Let $\phi \in E^*$. Then ϕ is continuous with respect to the weak topology, i.e. if $(x_\alpha)_\alpha$ is a net in E with $x_\alpha \stackrel{\text{weakly}}{\to} 0$, then $\phi(x_\alpha) \to 0$. Assume that ϕ is unbounded on each U_n. This means that, for each $n \in \mathbb{N}$, an element $x_n \in U_n$ with $|\phi(x_n)| \geq n$. It follows that $d(x_n, 0) \leq \frac{1}{n} \to 0$ and thus $x_n \stackrel{\text{weakly}}{\to} 0$, whereas $\phi(x_n) \not\to 0$. It follows that there is $N \in \mathbb{N}$ such that $\sup\{|\phi(x)| : x \in U_N\} < \infty$. Since

$$\bigcap_{j=1}^{k_N} \ker \phi_j^{(N)} \subset \left\{ x \in E : \max_{j=1,\ldots,k_N} \left| \phi_j^{(N)}(x) \right| < \epsilon_n \right\} \subset U_N,$$

the functional ϕ must be bounded on $\bigcap_{j=1}^{k_N} \ker \phi_j^{(N)}$, which, in turn, is possible only if $\bigcap_{j=1}^{k_N} \ker \phi_j^{(N)} \subset \ker \phi$. By Lemma 4.2.6, there are thus $\lambda_1, \ldots, \lambda_{k_N} \in \mathbb{F}$ such that $\phi = \lambda_1 \phi_1^{(N)} + \cdots + \lambda_{k_N} \phi_{k_N}^{(N)}$.

It follows that E^* is the linear span of the countable set $\left\{ \phi_j^{(n)} : n \in \mathbb{N},\, j = 1, \ldots, k_n \right\}$. Hence, E^* has a countable Hamel basis. Since E^* is always a Banach space, this means that $\dim E^* < \infty$ and, consequently, $\dim E < \infty$. $\square$

In analogy with the weak topology, one defines an even weaker topology on the dual of a normed space:

Definition 4.2.8 Let E be a normed space. Then $\{p_x : x \in E\}$ with

$$p_x(\phi) = |\phi(x)| \qquad (\phi \in E^*,\, x \in E)$$

is a family of seminorms on E^* such that $\bigcap_{x \in E} \{\phi \in E^* : p_x(\phi) = 0\}$. The corresponding topology on E^* is called the *weak* topology* on E^*.

Exercise 4.5 Let E be a separable Banach space. Show that the relative topology of the weak*-topology on the closed unit ball of E^* is metrizable. (*Hint*: Let $\{x_n : n \in \mathbb{N}\}$ be a countable dense subset of E, and define

$$d(\phi, \psi) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|\phi(x_n) - \psi(x_n)|}{|\phi(x_n) - \psi(x_n)| + 1} \qquad (\phi, \psi \in E^*).$$

Show that d is a metric on E^* which, on the closed unit ball of E^*, defines the weak*-topology.)

Exercise 4.6 Let E be a Banach space. Show that the following are equivalent:

(a) $\dim E < \infty$;

(b) the weak*-topology and the norm topology coincide;

(c) the weak*-topology is metrizable.

How does this go together with Exercise 4.5?

Theorem 4.2.9 (Alaoglu–Bourbaki theorem) *Let E be a normed space. Then the closed unit ball of E^* is compact in the weak* topology on E^*.*

Proof For each $x \in E$, let

$$K_x := \{\lambda \in \mathbb{F} : |\lambda| \leq \|x\|\}.$$

Since each K_x is closed and bounded, it is compact. By Tychonoff's theorem (Theorem A.7.4), $\prod_{x \in E} K_x$ is compact in the product topology. Embed the closed unit ball of E^* into $\prod_{x \in E} K_x$ via

$$B_1[0] \to \prod_{x \in E} K_x, \quad \phi \mapsto (\phi(x))_{x \in E}.$$

Let $(\phi_\alpha)_\alpha$ be a net in the closed unit ball of E^*; we will show that it has a convergent subnet. By Theorem A.6.6, the net $((\phi_\alpha(x))_{x \in E})_\alpha$ has a subnet $((\phi_\alpha(x))_{x \in E})_\alpha$ that converges in the product topology, i.e. for each $x \in E$, there is $\lambda_x \in K_x$ such that

$$\lambda_x = \lim_\beta \phi_\beta(x).$$

Define $\phi \colon E \to \mathbb{F}$ by letting $\phi(x) := \lambda_x$ for $x \in E$. For $x, y \in E$ and $\mu \in \mathbb{F}$, we have

$$\phi(x + y) = \lambda_{x+y} = \lim_\beta \phi_\beta(x + y) = \lim_\beta \phi_\beta(x) + \lim_\beta \phi_\beta(y) = \lambda_x + \lambda_y = \phi(x) + \phi(y)$$

and

$$\phi(\mu x) = \lambda_{\mu x} = \lim_\beta \phi_\beta(\mu x) = \mu \lim_\beta \phi_\beta = \mu \lambda_x = \mu \phi(x).$$

Hence, ϕ is linear. Moreover, note that, for $x \in E$ with $\|x\| \leq 1$,

$$|\phi(x)| = |\lambda_x| \leq \|x\| \leq 1$$

because $\lambda_x \in K_x$. It follows that $\phi \in E^*$ lies in the closed unit ball. From the definition of the weak* topology, it is clear that $\phi_\alpha \stackrel{\text{weak}^*}{\to} \phi$. Theorem A.6.6 eventually yields the weak* compactness of the closed unit ball of E^*. □

Exercise 4.7 Let E be a normed space. Show that there is a compact Hausdorff space X and an isometry $\iota \colon E \to \mathcal{C}(X)$.

Exercise 4.8 A Banach space E is called *reflexive* if the canonical map $J : E \to E^{**}$ is an isomorphism.

(i) Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, and let $p \in (1, \infty)$. Show that $L^p(\Omega, \mathfrak{S}, \mu)$ is reflexive.

(ii) Conclude that every Hilbert space is reflexive.

(iii) Argue that ℓ^∞ is not reflexive.

Exercise 4.9 Let E be a reflexive Banach space. Show that the closed unit ball of E is compact in the weak topology.

4.3 Schauder's fixed point theorem

We need the following fact from algebraic topology:

Theorem 4.3.1 *There is no continuous map from the closed unit ball of* $\mathbb{R}^N$ *to* $S^{N-1} := \{x \in \mathbb{R}^N : \|x\|_2 = 1\}$ *whose restriction to* S^{N-1} *is the identity.*

Theorem 4.3.2 (Brouwer's fixed point theorem) *Let* $K := \{x \in \mathbb{R}^N : \|x\|_2 \leq 1\}$, *and let* $f \colon K \to K$ *be continuous. Then* f *has a fixed point in* K.

Proof Assume that the theorem is false, i.e. $f(x) \neq x$ for all $x \in K$. Define $\phi \colon K \to \mathbb{R}^N$ by letting

$$\phi(x) := \text{the unique intersection point of the line from } f(x) \text{ through } x \text{ with } S^{N-1}.$$

Then ϕ is continuous with $\phi(K) \subset S^{N-1}$ and $\phi|_{S^{N-1}} = \mathrm{id}$, which is impossible by Theorem 4.3.1. $\square$

Corollary 4.3.3 *Let* E *be a finite-dimensional normed space, let* $\varnothing \neq K \subset E$ *be compact and convex, and let* $f \colon K \to K$ *be continuous. Then* f *has a fixed point in* K.

Proof Without loss of generality, let $E = \mathbb{R}^N$. Choose $r > 0$ such that

$$K \subset \{x \in \mathbb{R}^N : \|x\|_2 \leq r\} =: B.$$

Define $\phi \colon B \to K$ by letting

$$\phi(x) := \text{the unique point } y \in K \text{ such that } \|x - y\|_2 = \operatorname{dist}(x, K) \qquad (x \in B).$$

Then ϕ is continuous, and $\phi(x) = x$ for $x \in K$. Hence, $f \circ \phi$ is continuous and maps B into B. By Theorem 4.3.2, there is $x \in B$ such that $f(\phi(x)) = x$. Since $\phi(B) \subset K$, we have $x \in K$, so that $f(x) = f(\phi(x)) = x$. $\square$

Exercise 4.10 Use the intermediate value theorem to prove Corollary 4.3.3 in the one-dimensional case: If $f \colon [a, b] \to [a, b]$ is continuous, then f has a fixed point.

Schauder's fixed point theorem is the infinite-dimensional generalization of Corollary 4.3.3.

Definition 4.3.4 A subset S of a linear space E is called *balanced* if $\{\lambda x : x \in S, \lambda \in \mathbb{F}, |\lambda| \leq 1\} \subset S$.

Definition 4.3.5 Let E be a linear space, and let $\varnothing \neq K \subset E$ be convex and balanced. Then the *Minkowski functional* μ_K of K is defined by

$$\mu_K(x) := \inf\{t > 0 : x \in tK\} \qquad (x \in E).$$

Proposition 4.3.6 *Let E be a locally convex vector space, and let $\varnothing \neq U \subset E$ be open, convex, and balanced. Then μ_U is a continuous seminorm on E such that*

$$U = \{x \in E : \mu_U(x) < 1\}. \tag{4.4}$$

Proof Since U is balanced, we have $0 \in U$. Let $x \in E$. Since $\frac{1}{n}x \to 0$, and since U is a neighborhood of 0, there is $n \in \mathbb{N}$ such that $\frac{1}{n}x \in U$. Hence, $\mu_U(x) < \infty$. Clearly, $\mu_U(0) = 0$. Let $x \in E$, and $\lambda \in \mathbb{F} \setminus \{0\}$. We have that

$$\begin{aligned}
\lambda x \in tU \quad &\Longleftrightarrow \quad x \in t\lambda^{-1}U \\
&\Longleftrightarrow \quad x \in t\lambda^{-1}\frac{\lambda}{|\lambda|}U, \qquad \text{since } \frac{\lambda}{|\lambda|}U = U, \\
&\Longleftrightarrow \quad x \in t|\lambda^{-1}|U \qquad (t > 0),
\end{aligned}$$

and hence

$$\mu_U(\lambda x) = \inf\{t > 0 : x \in t|\lambda^{-1}|U\} = |\lambda| \inf\{t > 0 : x \in tU\} = |\lambda|\mu_U(x).$$

Let $x, y \in E$, and let $\epsilon > 0$. Choose $t, s > 0$ such that $x \in tU$, $y \in sU$, $t < \mu_U(x) + \frac{\epsilon}{2}$, and $s < \mu_U(y) + \frac{\epsilon}{2}$. It follows that

$$\frac{x}{t+s} \in \frac{t}{t+s}U \qquad \text{and} \qquad \frac{y}{t+s} \in \frac{s}{t+s}U$$

and therefore

$$\frac{x+y}{t+s} \in \frac{t}{t+s}U + \frac{s}{t+s}U \subset U$$

because U is convex. It follows that

$$\mu_U(x+y) \leq t + s \leq \mu_U(x) + \frac{\epsilon}{2} + \mu_U(y) + \frac{\epsilon}{2} = \mu_U(x) + \mu_U(y) + \epsilon$$

and therefore

$$\mu_U(x+y) \leq \mu_U(x) + \mu_U(y).$$

All in all, μ_U is a seminorm.

Let $x \in U$. Since

$$(0, \infty) \mapsto E, \quad t \mapsto t^{-1}x$$

is continuous, and since U is open, there is $t \in (0,1)$ such that $t^{-1}x \in U$. It follows that $\mu_U(x) < 1$. Since U is balanced, $\mu_U(x) \geq 1$ holds trivially for $x \notin U$. This proves (4.4).

Since U is open and contains 0, there are $\epsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{P}$ such that

$$\left\{ x \in E : \max_{j=1,\ldots,n} p_j(x) \right\} \subset U.$$

Let $(x_\alpha)_\alpha$ be a net in E such that $x_\alpha \to 0$, i.e. $p_j(x_\alpha) \to 0$ for $j = 1, \ldots, n$. Let $\delta > 0$. Then there is α_0 such that $p_j(x_\alpha) < \epsilon\delta$ for $j = 1, \ldots, n$ and $\alpha \succ \alpha_0$. This means that $x_\alpha \in \delta U$ for all $\alpha \succ \alpha_0$ and thus $\mu_U(x_\alpha) < \delta$ for all $\alpha \succ \alpha_0$. It it follows that $\mu_U(x_\alpha) \to 0$. Finally, let $(x_\alpha)_\alpha$ be a net in E such that $x_\alpha \to x \in E$. Since $x_\alpha - x \to 0$, we have

$$|\mu_U(x_\alpha) - \mu_U(x)| \leq \mu_U(x_\alpha - x) \to 0.$$

Hence, μ_U is continuous. $\square$

Lemma 4.3.7 *Let E be a locally convex vector space, let $\varnothing \neq K \subset E$ be compact, let $f : K \to K$ be continuous and suppose that $f(x) \neq x$ for all $x \in K$. Then there is an open, convex, balanced set W such that*

$$(\{(x, f(x)) : x \in K\} + W \times W) \cap \{(x, x) : x \in E\} = \varnothing.$$

Proof Let

$$\operatorname{Gr} f := \{(x, f(x)) : x \in K\} \qquad \text{and} \qquad \Delta := \{(x, x) : x \in E\}.$$

Let $(x, f(x)) \in \operatorname{Gr} f$. Then there are open subset U and V of E with $x \in U$, $f(x) \in V$, and $(U \times V) \cap \Delta = \varnothing$. Without loss of generality, we may suppose that there are $\epsilon > 0$ and $p_1, \ldots, p_n, q_1, \ldots, q_m \in \mathcal{P}$ such that

$$U = \left\{ y \in E : \max_{j=1,\ldots,n} p_j(x - y) < \epsilon \right\}$$

and

$$V = \left\{ y \in E : \max_{j=1,\ldots,m} q_j(f(x) - y) < \epsilon \right\}.$$

Define

$$W_x := \left\{ y \in E : \max_{\substack{j=1,\ldots,n \\ k=1,\ldots,m}} \{p_j(y), q_k(y) < \epsilon \right\}.$$

Then W_x is an open, convex, balanced set with $((x, f(x)) + W_x \times W_x) \cap \Delta = \varnothing$. Since K is compact, there are $x_1, \ldots, x_k \in K$ such that

$$K \subset \bigcup_{j=1}^{k} ((x_j, f(x_j)) + W_{x_j} \times W_{x_j}).$$

Define $W := \bigcap_{j=1}^{k} W_{x_j}$. $\square$

Exercise 4.11 Let E be a linear space, and let S be a non-empty subset of E. The *convex hull* $\operatorname{conv} S$ of S is defined as the intersection of all convex subsets of E containing S. Show that $\operatorname{conv} S$ consists of all elements of E of the form $\sum_{j=1}^{n} t_j s_j$, where $n \in \mathbb{N}$, $s_1, \ldots, s_n \in S$, and $t_1, \ldots, t_n > 0$ with $\sum_{j=1}^{n} t_j = 1$.

Exercise 4.12 Let E be a locally convex vector space, and let $x_1, \ldots, x_n \in E$. Show that the convex hull of $\{x_1, \ldots, x_n\}$ is compact.

Theorem 4.3.8 (Schauder's fixed point theorem) *Let E be a locally convex space, let $\varnothing \neq K \subset E$ be compact, and let $f\colon K \to K$ be continuous. Then f has a fixed point in K.*

Proof Assume that the theorem is false, i.e. the hypotheses of Lemma 4.3.7 are satisfied. Choose W as specified in Lemma 4.3.7, i.e., in particular,

$$f(x) \notin x + W \qquad (x \in K). \tag{4.5}$$

By Proposition 4.3.6, μ_W is a continuous seminorm on E such that $W = \{x \in E : \mu_W(x) < 1\}$. Define

$$\alpha\colon E \to \mathbb{R}, \quad x \mapsto \max\{0, 1 - \mu_W(x)\}.$$

Choose $x_1, \ldots, x_n \in K$ such that $K \subset \bigcup_{j=1}^{n} x_j + W$. For $j = 1, \ldots, n$, define $\alpha_j\colon E \to \mathbb{R}$ and $\beta_j\colon K \to \mathbb{R}$ by letting

$$\alpha_j(x) := \alpha(x - x_j) \qquad (x \in E,\ j = 1, \ldots n)$$

and

$$\beta_j(x) := \frac{\alpha_j(x)}{\alpha_1(x) + \cdots + \alpha_n(x)} \qquad (x \in K,\ j = 1, \ldots n).$$

Let F be the linear span of $\{x_1, \ldots, x_n\}$, and let $C := \operatorname{conv}\{x_1, \ldots, x_n\}$. Then $C \subset K$ is a compact, convex subset of the finite-dimensional (normed) space F. Define

$$g\colon K \to H, \quad x \mapsto \sum_{j=1}^{n} \beta_j(x) x_j.$$

Since $g \circ f$ maps H into itself, Corollary 4.3.3 yields $x_0 \in H$ such that $g(f(x_0)) = x_0$. Since $\beta_j(x) = 0$ for $x \notin x_j + W$, we have

$$x - g(x) = \sum_{j=1}^{n} \beta_j(x)(x - x_j) \in W \qquad (x \in K).$$

In particular,

$$f(x_0) - x_0 = f(x_0) - g(f(x_0)) \in W$$

holds. This, however, contradicts (4.5). $\square$

Exercise 4.13 Let B be the closed unit ball in ℓ^2, and define $f : B \to \ell^2$ by letting, for $x = (x_n)_{n=1}^{\infty}$,

$$f(x) = ((1 - \|x\|^2), x_1, x_2, \ldots).$$

Show that f is continuous with $f(B) \subset B$, but has no fixed point.

4.3.1 Peano's theorem

Like Banach's fixed point theorem, Schauder's fixed point theorem can be applied to initial value problems:

Lemma 4.3.9 (Mazur's theorem) *Let E be a Banach space, and let $K \subset E$ be compact. Then $\overline{\operatorname{conv} K}$ is also compact.*

Proof Let $\epsilon > 0$. Choose $x_1, \ldots, x_n \in K$ such that

$$K \subset \bigcup_{j=1}^{n} B_{\frac{\epsilon}{3}}(x_j).$$

Let $C := \operatorname{conv}\{x_1, \ldots, x_n\}$. Since C is compact, there are $y_1, \ldots, y_m \in C$ such that

$$C \subset \bigcup_{j=1}^{m} B_{\frac{\epsilon}{3}}(y_j).$$

Let $z \in \overline{\operatorname{conv} K}$. Then there is $w \in \operatorname{conv} K$ such that $\|w - z\| < \frac{\epsilon}{3}$. Let $t_1, \ldots, t_k > 0$ with $t_1 + \cdots + t_k = 1$ and $v_1, \ldots, v_k \in K$ such that $w = \sum_{j=1}^{k} t_j v_j$. For each $j = 1, \ldots, k$, there is $\nu(j) \in \{1, \ldots, n\}$ such that $\|v_j - x_{\nu(j)}\| < \frac{\epsilon}{3}$. It follows that

$$\begin{aligned} \left\| w - \sum_{j=1}^{k} t_j x_{\nu(j)} \right\| &= \left\| \sum_{j=1}^{k} t_j (v_j - x_{\nu(j)}) \right\| \\ &\leq \sum_{j=1}^{k} t_j \|v_j - x_{\nu(j)}\| \\ &< \frac{\epsilon}{3}. \end{aligned}$$

Since $\sum_{j=1}^{k} t_j x_{\nu(j)} \in C$, there is $j_0 \in \{1, \ldots, m\}$ such that

$$\left\| \sum_{j=1}^{k} t_j x_{\nu(j)} - y_{j_0} \right\| < \frac{\epsilon}{3}.$$

All in all, we have that

$$\begin{aligned}
\|z - y_{j_0}\| &\leq \|z - w\| + \left\| w - \sum_{j=1}^{k} t_j x_{\nu(j)} \right\| + \left\| \sum_{j=1}^{k} t_j x_{\nu(j)} - y_{j_0} \right\| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon.
\end{aligned}$$

Since $z \in \overline{\operatorname{conv} K}$ was arbitrary, this means that

$$\overline{\operatorname{conv} K} \subset \bigcup_{j=1}^{m} B_\epsilon(y_j),$$

so that $\overline{\operatorname{conv} K}$ is totally bounded and thus compact. $\square$

Theorem 4.3.10 (Peano's theorem) *Let $I = [a, b]$, and let $f \in \mathcal{C}_b(I \times \mathbb{R})$. Then the initial value problem*

$$y' = f(x, y), \quad y(a) = y_0$$

has a solution $\phi \in \mathcal{C}^1(I)$.

Proof As in the proof of Theorem 4.1.2, we need to find $\phi \in \mathcal{C}(I)$ such that

$$\phi(x) = \int_a^x f(t, \phi(t))\, d + y_0. \tag{4.6}$$

For the sake of simplicity, suppose that $I = [0, 1]$, $y_0 = 0$, and $\|f\|_\infty \leq 1$. Again as in the proof of Theorem 4.1.2, define $T \colon \mathcal{C}(I) \to \mathcal{C}(I)$ by letting

$$(T\phi)(x) := \int_0^x f(t, \phi(t))\, dt \qquad (\phi \in \mathcal{C}(I),\ x \in I).$$

Let $B := \{\phi \in \mathcal{C}(I) : \|\phi\|_\infty \leq 1\}$. Then T maps B into itself. Note that

$$|(T\phi)(y) - (T\phi)(x)| \leq \int_x^y |f(t, \phi(t))|\, dt \leq y - x \qquad (y \geq x).$$

The family $\{T\phi : \phi \in B\}$ is therefore equicontinuous and thus relatively compact. Let $K := \overline{\operatorname{conv} TB}$. By Lemma 4.3.9, K is compact and clearly $TK \subset K$. By Theorem 4.3.8, T has a fixed point in K, i.e. a solution of (4.6). $\square$

Exercise 4.14 Give two different solutions of the initial value problem

$$y' = \sqrt{y}, \qquad y(0) = 0.$$

4.3.2 Lomonosov's theorem

We now give a second application of Schauder's fixed point theorem to a famous open problem in operator theory.

Definition 4.3.11 Let E be a Banach space, and let $T \in \mathcal{B}(E)$. A closed subspace F of E is called *invariant* for T if

(a) $\{0\} \subsetneq F \subsetneq E$, and

(b) $TF \subset F$.

If F is invariant for every $S \in \mathcal{B}(E)$ commuting with T, it is called *hyperinvariant*

The *invariant subspace problem* — posed by J. von Neumann — is the following question:

> *Let $\mathfrak{H}$ be a (separable, infinite-dimensional) Hilbert space over $\mathbb{C}$, and let $T \in \mathcal{B}(\mathfrak{H})$. Does T have an invariant subspace?*

For operators on Banach spaces, the answer is negative: Counterexamples have been constructed by P. Enflo and C. J. Read. Read's construction even yields an operator $T \in \mathcal{B}(\ell^1)$ without invariant subspace.

Exercise 4.15 Let $\mathfrak{H}$ be a Hilbert space such that either $2 \leq \dim \mathfrak{H} < \infty$ or that $\mathfrak{H}$ is not separable. Show that every bounded linear operator on $\mathfrak{H}$ has an invariant subspace.

Theorem 4.3.12 (Lomonosov's theorem) *Let E be an infinite-dimensional Banach space over $\mathbb{C}$, and let $T \in \mathcal{K}(E) \setminus \{0\}$. Then T has a hyperinvariant subspace.*

Proof Assume towards a contradiction that the claim is wrong. Without loss of generality suppose that $\|T\| = 1$. Fix $x_0 \in E$ with $\|Tx_0\| > 1$, and let $B := \{x \in E : \|x - x_0\| \leq 1\}$, so that

$$0 \notin B \qquad \text{and} \qquad 0 \notin \overline{TB}.$$

For any $x \in E \setminus \{0\}$ define

$$F_x := \overline{\{Sx : S \in \mathcal{B}(E),\, ST = TS\}}.$$

Then $F_x \neq \{0\}$ is a closed subspace of E with $SF_x \subset F_x$ for all $S \in \mathcal{B}(E)$ with $ST = TS$. By assumption, this means that $F_x = E$ for all $x \in E \setminus \{0\}$. Hence, for each $y \in \overline{TB}$, there is $S_y \in \mathcal{B}(E)$ commuting with T such that $\|S_y y - x_0\| < 1$. Since $K := \overline{TB}$ is compact, there are $S_1, \ldots, S_n \in \mathcal{B}(E)$ commuting with T such that

$$K \subset \bigcup_{j=1}^{n} \{y \in E : \|S_j y - x_0\| < 1\}.$$

For $y \in K$ and $j = 1, \ldots, n$, let

$$\alpha_j(y) := \max\{0, 1 - \|S_j y - x_0\|\}.$$

For $j = 1, \ldots, n$, define

$$\beta_j \colon K \to \mathbb{R}, \quad y \mapsto \frac{\alpha_j(y)}{\alpha_1(y) + \cdots + \alpha_n(y)},$$

and let

$$f \colon B \to E, \quad x \mapsto \sum_{j=1}^{n} \beta_j(Tx) S_j Tx,$$

so that f is continuous. Let $x \in B$, so that $Tx \in K$. If $\beta_j(Tx) > 0$, then $\alpha_j(Tx) > 0$ and therefore $\|S_j Tx - x_0\| < 1$, i.e. $S_j Tx \in B$. The convexity of B yields $f(B) \subset B$. The compactness of T yields that $\overline{f(B)}$ is compact. Let $C := \overline{\operatorname{conv} f(B)}$. By Lemma 4.3.9, C is compact, and clearly $f(C) \subset C$. By Theorem 4.3.8, there is a fixed point y_0 of f in $C \subset B$. Let

$$R := \sum_{j=1}^{n} \beta_j(Ty_0) S_j.$$

Then T commutes with T and satisfies $RTy_0 = f(y_0) = y_0$. Since $y_0 \in B \subset E \setminus \{0\}$, this means that $F_0 := \ker(RT - \mathrm{id}_E) \neq \{0\}$. Since $RT \in \mathcal{K}(E)$, we also have that $\dim F_0 < \infty$ and thus $F_0 \subsetneq E$. Clearly, F_0 is invariant form T. Let $\lambda \in \mathbb{C}$ be an eigenvalue of $T|_{F_0}$, so that $F := \ker(\lambda - T) \neq \{0\}$. If $\lambda = 0$, then $F = \ker T \neq E$, since $T \neq 0$. If $\lambda \neq 0$, then $\dim F < \infty$, so that $F \neq E$ as well in this case. Clearly, F is hyperinvariant for T. $\square$

4.4 The Markoff–Kakutani fixed point theorem

We present yet another fixed point theorem, this time not for a single function, but for a whole family of maps.

Definition 4.4.1 Let K be a convex subset of a linear space. A map $T \colon K \to K$ is called *affine* if

$$T(tx + (1-t)y) = tTx + (1-t)Ty \qquad (x, y \in K,\ t \in [0,1]).$$

Theorem 4.4.2 (Markoff–Kakutani fixed point theorem) *Let E be a locally convex vector space, let $\varnothing \neq K \subset E$ be compact and convex, and let $\mathcal{S}$ be an abelian semigroup of continuous affine maps on K, i.e.*

(a) *each $S \in \mathcal{S}$ is a continuous affine map on K,*

(b) *$ST \in \mathcal{S}$ for all $S, T \in \mathcal{S}$, and*

(c) *$ST = TS$ for all $S, T \in \mathcal{S}$.*

Then there is $x_0 \in K$ such that $Sx_0 = x_0$ for all $S \in \mathcal{S}$.

Proof For $n \in \mathbb{N}$ and $S \in \mathcal{S}$, let

$$S_n := \frac{1}{n} \sum_{k=0}^{n-1} S^k.$$

It is easy to see that

$$S_n T_m = T_m S_n \qquad (n, m \in \mathbb{N},\ S, T \in \mathcal{S}).$$

Let

$$\mathcal{K} := \{S_n K : n \in \mathbb{N},\ S \in \mathcal{S}\}.$$

Then $\mathcal{K}$ consists of non-empty, compact, convex subsets of K. Let $n_1, \ldots, n_k \in \mathbb{N}$ and $S^{(1)}, \ldots, S^{(k)} \in \mathcal{S}$ and note that

$$\underbrace{\left(S^{(1)}_{n_1} \cdots S^{(k)}_{n_k}\right)(K)}_{\neq \varnothing} \subset \bigcap_{j=1}^{k} S^{(j)}_{n_j} K.$$

Hence, every finite family of sets in $\mathcal{K}$ has non-empty intersection. Since K is compact, this means that

$$\bigcap \{S_n K : n \in \mathbb{N},\ S \in \mathcal{S}\} \neq \varnothing$$

Let x_0 be any point in this intersection. Fix $S \in \mathcal{S}$ and $n \in \mathbb{N}$. Then there is $x \in K$ such that

$$x_0 = S_n x = \frac{1}{n}(x + Sx + \cdots + S^{n-1}x).$$

It follows that

$$\begin{aligned} Sx_0 - x_0 &= \frac{1}{n}(Sx + S^2x + \cdots + S^n x) - \frac{1}{n}(x + Sx + \cdots + S^{n-1}x) \\ &= \frac{1}{n}(S^n x - x) \\ &\in \frac{1}{n}(K - K). \end{aligned}$$

Let $\epsilon > 0$, and let $p_1, \ldots, p_n \in \mathcal{P}$. Let

$$U := \left\{ x \in E : \max_{j=1,\ldots,n} p_j(x) < \epsilon \right\}.$$

Since K is compact, so is $K - K$. Hence, there is $N \in \mathbb{N}$ such that $\frac{1}{n}(K - K) \subset U$ for all $n \geq N$. It follows that $\lim_{n\to\infty}(Sx_0 - x_0) = 0$, i.e. $Sx_0 = x_0$. $\square$

Definition 4.4.3 Let G be a group.

(a) A *mean* on $\ell^\infty(G)$ is a functional $m \in \ell^\infty(G)^*$ such that $m(\phi) \geq 0$ for $\phi \geq 0$ and $m(1) = 1$.

(b) A mean on $\ell^\infty(G)$ is called *left translation invariant* if

$$m(L_g\phi) = m(\phi) \qquad (g \in G,\, \phi \in \ell^\infty(G)),$$

where

$$(L_g\phi)(h) := \phi(gh) \qquad (g, h \in G,\, \phi \in \ell^\infty(G)).$$

(c) G is called *amenable* if there is a left translation invariant mean on $\ell^\infty(G)$.

Examples 1. If G is finite, then it is amenable: Define

$$m(\phi) := \frac{1}{|G|}\sum_{g \in G} \phi(g) \qquad (\phi \in \ell^\infty(G)).$$

2. The free group $\mathbb{F}_2$ in two generators, say a and b, is not amenable. To see this, let, for $x \in \{a, b, a^{-1}, b^{-1}\}$,

$$W(x) := \{w \in \mathbb{F}_2 : w \text{ starts with } x\},$$

so that

$$\mathbb{F}_2 = W(a) \cup W(b) \cup W(a^{-1}) \cup W(b^{-1}) \cup \{e\}, \tag{4.7}$$

the union being disjoint. In terms of indicator functions, (4.7) becomes

$$1 = \chi_{W(a)} + \chi_{W(b)} + \chi_{W(a^{-1})} + \chi_{W(b^{-1})} + \chi_{\{e\}}.$$

Let $w \in \mathbb{F}_2 \setminus W(a)$. Then we have $a^{-1}w \in W(a^{-1})$ and thus $w \in aW(a^{-1})$. Hence, we have the union

$$\mathbb{F}_2 = W(a) \cup aW(a^{-1})$$

which means, in terms of indicator functions, that

$$1 \leq \chi_{W(a)} + \chi_{aW(a^{-1})} = \chi_{W(a)} + L_{a^{-1}}\chi_{W(a^{-1})};$$

similarly, we obtain

$$1 = \chi_{W(b)} + L_{b^{-1}}\chi_{W(b^{-1})}.$$

Assume now that we have a left translation invariant mean $m \in \ell^\infty(G)^*$. We have:

$$\begin{aligned}
1 &= m(1) \\
&\geq m(\chi_{W(a)} + \chi_{W(b)} + \chi_{W(a^{-1})} + \chi_{W(b^{-1})}) \\
&= m(\chi_{W(a)}) + m(\chi_{W(b)}) + m(\chi_{W(a^{-1})}) + m(\chi_{W(b^{-1})})) \\
&= m(\chi_{W(a)}) + m(\chi_{W(b)}) + m(L_{a^{-1}}\chi_{W(a^{-1})}) + m(L_{b^{-1}}\chi_{W(b^{-1})})) \\
&= m(\chi_{W(a)} + L_{a^{-1}}\chi_{W(a^{-1})}) + m(\chi_{W(b)} + L_{b^{-1}}\chi_{W(b^{-1})})) \\
&\geq m(1) + m(1) \\
&= 2.
\end{aligned}$$

This is impossible.

3. Let G be an abelian group. Let

$$K := \{m \in \ell^\infty(G)^* : \|m\| \leq 1 \text{ and } m \text{ is a mean}\}.$$

Clearly, K is convex. Since K is weak*-closed in the closed unit ball of $\ell^\infty(G)^*$, it is weak* compact by Theorem 4.2.9. Clearly, if $m \in K$, then so is $L_g^* m$ for all $g \in G$. The family $\{L_g^* : g \in G\}$ is an abelian semigroup of continuous affine mappings on K. By Theorem 4.4.2, there is $m_0 \in K$ such that $L_g^* m_0 = m_0$ for all $g \in G$, i.e.

$$m_0(\phi) = (L_g^* m_0)(\phi) = m_0(L_g \phi) \qquad (g \in G,\ \phi \in \ell^\infty(G)).$$

Hence, G is amenable.

4.5 A geometric consequence of the Hahn–Banach theorem

Lemma 4.5.1 *Let E be a locally convex vector space, and let U be a neighborhood of 0. Then any linear functional $\phi\colon E \to \mathbb{F}$ with $\sup\{|\phi(x)| : x \in U\} < \infty$ is continuous.*

Proof Let $\epsilon > 0$, and let $C := \sup\{|\phi(x)| : x \in U\}$. Then

$$V := \left\{ \frac{\epsilon}{C+1} x : x \in U \right\}$$

is a neighborhood of 0 such that $|\phi(x)| < \epsilon$ for all $x \in V$. Hence, ϕ is continuous at 0 and thus everywhere. $\square$

Lemma 4.5.2 *Let E be a locally convex vector space, and let U and K be non-empty convex subsets of E with $U \cap K = \varnothing$ and U open. Then there are $\phi \in E^*$ and $c \in \mathbb{R}$ such that*

$$\operatorname{Re} \phi(x) < c \leq \operatorname{Re} \phi(y) \qquad (x \in U,\ y \in K).$$

Proof Consider first the case where $\mathbb{F} = \mathbb{R}$.

Fix $x_0 \in U$ and $y_0 \in K$. Let $z_0 := y_0 - x_0$, and define

$$V := U - K + z_0.$$

Then V is an open, convex neighborhood of 0. Let μ_V be the corresponding Minkowski functional. As in the proof of Proposition 4.3.6, one sees that μ_V is a sublinear functional on E with

$$V = \{x \in E : \mu_V(x) < 1\}.$$

In particular, $\mu_V(z_0) \geq 1$ holds. Let $F = \mathbb{R}z_0$, and define $\psi \colon F \to \mathbb{R}$ by letting $\psi(tz_0) = t$ for $t \in \mathbb{R}$. It follows that

$$\psi(tz_0) = \begin{cases} t \leq t\mu_V(z_0) = \mu_V(tz_0) & (t \geq 0) \\ t < 0 \leq \mu_V(tz_0) & (t < 0). \end{cases}$$

The Hahn–Banach theorem (Theorem 2.1.3) then yields $\phi : E \to \mathbb{R}$ with $\phi|_F = \psi$ and $\phi(x) \leq \mu_V(x)$ for all $x \in E$. Let $x \in U$ and $y \in K$, and note that

$$\phi(x) - \phi(y) + 1 = \phi(x - y + z_0) \leq \mu_V(\underbrace{x - y + z_0}_{\in V}) < 1;$$

it follows that

$$\phi(x) < \phi(y) \qquad (x \in U,\, y \in K). \tag{4.8}$$

Since $\phi(x) \leq 1$ for $x \in V$, we have $\phi(x) \geq -1$ for $x \in -V$ and thus $|\phi(x)| \leq 1$ for $x \in V \cap (-V)$. By Lemma 4.5.1, this means that $\phi \in E^*$. By (4.8), $\phi(U)$ and $\phi(K)$ are disjoint convex subsets of $\mathbb{R}$. Let $c := \sup_{x \in U} |\phi(x)|$. It follows that

$$\phi(x) \leq c \leq \phi(y) \qquad (x \in U,\, y \in K). \tag{4.9}$$

It is easy to see that $\phi(U) \subset \mathbb{R}$ is open. Hence, the first inequality in (4.9) must be strict.

Next, consider the case where $\mathbb{F} = \mathbb{C}$. Find $c \in \mathbb{R}$ and a continuous, $\mathbb{R}$-linear functional $\tilde{\phi} \colon E \to \mathbb{R}$ such that

$$\tilde{\phi}(x) < c \leq \tilde{\phi}(y) \qquad (x \in U,\, y \in K).$$

Define $\phi \in E^*$ by letting

$$\phi(x) = \tilde{\phi}(x) - i\tilde{\phi}(ix) \qquad (x \in E).$$

This completes the proof. $\square$

Theorem 4.5.3 *Let E be a locally convex vector space, let F and K be non-empty, disjoint, convex subsets of E such that F is closed and K is compact. Then there are $\phi \in E^*$ and $c_1, c_2 \in \mathbb{R}$ such that*

$$\operatorname{Re}\phi(x) \leq c_1 < c_2 \leq \operatorname{Re}\phi(y) \qquad (x \in K,\, y \in F).$$

Proof As in the proof of Lemma 4.3.7, we find an open, convex, balanced neighborhood V of 0 such that $(K + V) \cap F = \varnothing$. By Lemma 4.5.2, there are $c \in \mathbb{R}$ and $\phi \in E^*$ such that

$$\operatorname{Re}\phi(x) < c \leq \operatorname{Re}\phi(y) \qquad (x \in K + V,\, y \in F).$$

Let

$$c_1 := \sup_{x \in K} \operatorname{Re}\phi(x) \qquad \text{and} \qquad c_2 := \sup_{x \in K+V} \operatorname{Re}\phi(x).$$

Since $\phi(K + V)$ and $\phi(F)$ are disjoint convex subsets of $\mathbb{R}$ with $\phi(K + V)$ open and to the left of $\phi(F)$, we have

$$c_2 \leq \operatorname{Re}\phi(y) \qquad (y \in F).$$

Since $\phi(K) \subset \phi(K + V)$ is compact,

$$\phi(x) \leq c_1 < c_2 \qquad (x \in K)$$

follows. $\square$

4.6 The Kreĭn–Milman theorem

Definition 4.6.1 Let E be a linear space, and let $K \subset E$ be convex. A convex set $\varnothing \neq S \subset K$ is called an *extremal subset* of K if $tx + (1-t)y \in S$ with $x, y \in K$ and $t \in (0, 1)$ only if $x, y \in S$. If $x \in K$ is such that $\{x\}$ is an extremal subset of K, the point x is called an *extremal point* of K. The set of all extremal points of K is denoted by $\operatorname{ext} K$.

Lemma 4.6.2 *Let E be a locally convex vector space, let $\varnothing \neq K \subset E$ be compact and convex, let $\phi \in E^*$, and let $C := \sup_{x \in K} \operatorname{Re}\phi(x)$. Then*

$$K_\phi := \{x \in K : \operatorname{Re}\phi(x) = C\}$$

is an extremal, compact, convex subset of K.

Proof Clearly, K_ϕ is compact and convex.

Let $x, y \in K$ and $t \in (0,1)$ be such that

$$\operatorname{Re}\phi(tx + (1-t)y) = C.$$

On the other hand, we have

$$C = tC + (1-t)C \geq t\operatorname{Re}\phi(x) + (1-t)\operatorname{Re}\phi(y) = \operatorname{Re}\phi(tx + (1-t)y) = C,$$

which is possible only if $\operatorname{Re}\phi(x) = \operatorname{Re}\phi(y) = C$. $\square$

The Kreĭn–Milman theorem asserts that under certain conditions extremal points exist in abundance:

Theorem 4.6.3 (Kreĭn–Milman theorem) *Let E be a locally convex vector space, and let $\varnothing \neq K \subset E$ be compact and convex. Then $K := \overline{\operatorname{conv}(\operatorname{ext} K)}$.*

Proof Let $\mathcal{K}$ be the collection of all extremal, compact, convex subsets of K. By Lemma 4.6.2, $\mathcal{K}$ is not empty.

For $K_0 \in \mathcal{K}$, let $\mathcal{K}_0$ consist of those sets in $\mathcal{K}$ contained in K_0. Let $\mathcal{K}_0$ be ordered by set inclusion. By Zorn's lemma, $\mathcal{K}_0$ has a minimal element, say S. Let $x \in S$, and assume that there is $y \in S \setminus \{x\}$. Choose $\phi \in E^*$ with $\operatorname{Re}\phi(y) < \operatorname{Re}\phi(x)$. By Lemma 4.6.2, $S_\phi \subsetneq S$ is an extremal compact, convex subset of S. Since S is an extremal subset of K, this means that S_ϕ is also an extremal subset of K. This contradicts the minimality of S, so that $S = \{x\}$. In particular, we see that

$$K_0 \cap \operatorname{ext} K \neq \varnothing. \tag{4.10}$$

Clearly,

$$\tilde{K} := \operatorname{conv}\left(\overline{\operatorname{ext} K}\right) \subset K$$

holds, so that $\tilde{K}$ is compact. Assume that there is $x_0 \in K \setminus \tilde{K}$. By Theorem 4.5.3, there is $\phi \in E^*$ such that

$$\sup_{x \in \tilde{K}} \operatorname{Re}\phi(x) < \operatorname{Re}\phi(x_0) =: C. \tag{4.11}$$

But then, by Lemma 4.5.2 again,

$$K_\phi = \{x \in E : \operatorname{Re}\phi(x) = C\}$$

belongs to $\mathcal{K}$. By (4.11), we have $K_\phi \cap \tilde{K} = \varnothing$. This, however, contradicts (4.10) (with $K_0 = K_\phi$). $\square$

4.6.1 The Stone–Weierstraß theorem

We conclude these notes with a generalization of Theorem 2.3.3 whose proof makes use of a number of powerful theorems we have encountered:

Theorem 4.6.4 (Stone–Weierstraß theorem) *Let X be a compact Hausdorff space, and let $\mathfrak{A}$ be a closed subalgebra of $\mathcal{C}(X)$ such that:*

(a) $1 \in \mathfrak{A}$;

(b) *if $f \in \mathfrak{A}$, then $\bar{f} \in \mathfrak{A}$;*

(c) *if $x, y \in X$ with $x \neq y$, there there is $f \in \mathfrak{A}$ such that $f(x) \neq f(y)$.*

Then $\mathfrak{A} = \mathcal{C}(X)$.

Proof Assume that $\mathfrak{A} \subsetneq \mathcal{C}(X)$. By Corollary 2.1.6, there is $\psi \in \mathcal{C}(X)^*$ such that $\|\psi\| = 1$, but $\psi|_{\mathfrak{A}} = 0$. It follows that

$$K := \{\phi \in \mathcal{C}(X)^* : \|\phi\| \leq 1 \text{ and } \phi|_{\mathfrak{A}} = 0\} \neq \{0\}.$$

Clearly, K is convex and closed in the weak* topology. By Theorem 4.2.9, this means that K is weak* compact and Theorem 4.6.3 implies that $\operatorname{ext} K \neq \varnothing$. Let $\phi \in \operatorname{ext} K$; it is easy to see that $\|\phi\| = 1$. By Theorem B.3.8, there is a unique $\mu \in M(X)$ such that

$$\phi(f) = \int_X f(x)\, d\mu(x) \qquad (f \in \mathcal{C}(X)).$$

Let $X_0 := \operatorname{supp} \mu$, and let $x_0 \in X_0$.

We claim that $X_0 = \{x_0\}$. Let $x \in X \setminus \{x_0\}$. By (c), there is $f_1 \in \mathfrak{A}$ such that $\beta := f_1(x) \neq f_1(x_0)$. By (a), we have $\beta \in \mathfrak{A}$ and therefore $f_2 := f_1 - \beta \in \mathfrak{A}$. It follows that $f_2(x_0) \neq 0 = f_2(x)$. Let $f_3 = |f_2|^2 = f_2 \bar{f}_2 \in \mathfrak{A}$. Then (c) implies that $f_3 \in \mathfrak{A}$, and it is clear that $f_3(x) = 0 < f_3(x_0)$. Finally, let

$$f := \frac{1}{\|f_3\|_\infty + 1} f_3,$$

so that

$$f(x) = 0, \quad f(x_0) > 0, \quad \text{and} \quad 0 \leq f < 1.$$

Since $\mathfrak{A}$ is an algebra, we have $fg, (1-f)g \in \mathfrak{A}$ for all $g \in \mathfrak{A}$ and therefore

$$0 = \int fg\, d\mu = \int (1-f)g\, d\mu \qquad (g \in \mathfrak{A}).$$

It follows that $f\mu, (1-f)\mu \in K$. Let

$$\alpha := \|f\mu\| = \int f\, d|\mu|.$$

Since $f(x_0) > 0$, there are $\epsilon > 0$ and a neighborhood U of x_0 such that $f \geq \epsilon$ on U. It follows that

$$\alpha = \int f \, d|\mu| \geq \int_U f \, d|\mu| \geq \epsilon |\mu|(U) > 0$$

because $U \cap X_0 \neq \varnothing$. In a similar fashion, one shows that $\alpha < 1$. We also have

$$1 - \alpha = 1 - \int f \, d|\mu| = \int (1 - f) \, d|\mu| = \|(1 - f)\mu\|.$$

Hence,

$$\mu = \alpha \frac{f\mu}{\|f\mu\|} + (1 - \alpha) \frac{(1 - f)\mu}{\|(1 - f)\mu\|}$$

holds. Since $\mu \in \text{ext } K$, this means that

$$\mu = \frac{f\mu}{\|f\mu\|},$$

i.e. $\frac{f}{\|\mu f\|} = 1$ $|\mu|$-alsmost everywhere. Since f is continuuos, f equals α on X_0. Since $x_0 \in X_0$, we have

$$\alpha = f(x_0) > f(x) = 0,$$

so that $x \notin X_0$. Consequently, $X_0 = \{x_0\}$. We thus have $\lambda \in \mathbb{F}$ with $|\lambda| = 1$ such that $\mu = \lambda \delta_{x_0}$. Since

$$\int 1 \, d(\lambda \delta_{x_0}) = \lambda \neq 0,$$

this contradicts the choice of μ. $\square$

Corollary 4.6.5 *Let $\varnothing \neq K \subset \mathbb{R}^N$ be compact, and let $f \in \mathcal{C}(K)$. Then, for each $\epsilon > 0$, there is a polynomial p in N variables such that $\|f - p\| < \epsilon$.*

Proof Apply Theorem 4.6.4 with

$$\mathfrak{A} := \overline{\{p|_K : p \text{ is a polynomial in } N \text{ variables}\}}.$$

It follows that $\mathfrak{A} = \mathcal{C}(K)$. $\square$

Definition 4.6.6 Let X and Y be compact Hausdorff spaces, and let $f \in \mathcal{C}(X)$ and $g \in \mathcal{C}(Y)$. Then $f \otimes g \in \mathcal{C}(X \times Y)$ is defined through

$$(f \otimes g)(x, y) := f(x)g(y) \qquad (x \in X, \, y \in Y).$$

The *tenor product* $\mathcal{C}(X) \otimes \mathcal{C}(Y)$ of $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ is defined as the linear span in $\mathcal{C}(X \times Y)$ of the set $\{f \otimes g : f \in \mathcal{C}(X), \, g \in \mathcal{C}(Y)\}$.

Corollary 4.6.7 *Let X and Y be compact Hausdorff spaces, and let $f \in \mathcal{C}(X \times Y)$. Then $\mathcal{C}(X) \otimes \mathcal{C}(Y)$ is dense in $\mathcal{C}(X \times Y)$.*

Proof Let $\mathfrak{A} := \overline{\mathcal{C}(X) \otimes \mathcal{C}(Y)}$, and apply Theorem 4.6.4. $\square$

Appendix A

Point set topology

We need point set topology in this course for two reasons:

- spaces of continuous functions are important examples of Banach spaces;
- later in this course, we need to consider topological vector spaces which are not normed.

This appendix contains the necessary background in point set topology. For most statements, I have included proofs.

A.1 Open and closed sets

A topological space is a set that has just enough structure, so that we can speak sensibly of continuous functions on it. The notion of an open set is crucial:

Definition A.1.1 A *topological space* is a non-empty set X together with a family τ of subsets of X such that the following properties are satisfied:

(i) $\varnothing, X \in \tau$;

(ii) if $\mathcal{U}$ is a family of sets in τ, then $\bigcup\{U : U \in \mathcal{U}\} \in \tau$;

(iii) if $U_1, \ldots, U_n \in \tau$, then $U_1 \cap \cdots \cap U_n \in \tau$.

The family τ is called the *topology* of X.

Examples 1. Let (X, d) be a metric space. Define $U \subset X$ as open if, for each $x \in U$, there is $\epsilon > 0$ such that

$$B_\epsilon(x) := \{y \in X : d(x, y) < \epsilon\} \subset U.$$

It is well known that the collection of open subsets of X forms a topology. Different metrics can induce the same topology.

2. The collection of all subsets of any non-empty set is a topology. This topology is called the *discrete topology* of X.

3. For any non-empty set X, the collection $\{\varnothing, X\}$ is a topology. This topology is called the *chaotic topology* of X.

Exercise A.1 Verify the statement made in the first example. Show that, if (X, d) is any metric space, then

$$X \times X \to [0, \infty), \quad (x, y) \mapsto \frac{d(x, y)}{1 + d(x, y)}$$

is also a metric and induces the same topology as d.

Definition A.1.2 Let (X, τ) be a topological space.

(i) A subset U of X is *open* if $U \in \tau$.

(ii) A subset F of X is *closed* if $X \setminus F$ is open.

Passing to complements, the following is an immediate consequence of Definitions A.1.1 and A.1.2.

Theorem A.1.3 *Let X be a topological space. Then the following are true:*

(i) *$\varnothing$ and X are closed;*

(ii) *if $\mathcal{F}$ is a family of closed subsets of X, the $\bigcap\{F : F \in \mathcal{F}\}$ is closed;*

(iii) *if $F_1, \ldots, F_n$ are closed, then $F_1 \cup \cdots \cup F_n$ is closed.*

A.2 Continuity

In order to define continuity for functions between arbitrary topological spaces, we first introduce the notion of a neighborhood:

Definition A.2.1 Let X be a topological space, and let $x \in X$. A set $U \subset X$ is called a *neighborhood* of x if there is an open set $V \subset U$ such that $x \in V$.

Exercise A.2 Show that a subset of a topological space is open if and only if its a neighborhood of each of its points.

Definition A.2.2 Let X and Y be topological spaces. A function $f \colon X \to Y$ is *continuous at $x_0 \in X$* if $f^{-1}(U)$ is a neighborhood of x_0 for each neighborhood U of $f(x_0)$. If f is continuous at each $x \in X$, we simply say that f is *continuous*.

Exercise A.3 Let X and Y be topological spaces. Show that $f : X \to Y$ is continuous if and only if $f^{-1}(U)$ is open for each open $U \subset Y$ if and only if $f^{-1}(F)$ is closed for each closed $F \subset Y$.

For metric spaces, Definition A.2.2 is (equivalent to) the usual definition:

Proposition A.2.3 *Let (X, d) and (Y, Δ) be metric spaces, and let $x_0 \in X$. Then $f : X \to Y$ is continuous at x_0 in the usual sense if and only if f is continuous at x_0 in the sense of Definition* A.2.2.

Proof Suppose that f is continuous at x_0 in the usual sense. Let U be a neighborhood of $f(x_0)$. By the definition of a neighborhood, there is an open set $V \subset U$ such that $f(x_0) \in V$. From the definition of an open set in a metric space, there is $\epsilon > 0$ such that $B_\epsilon(f(x_0)) \in V$. From the definition of continuity in the context of metric spaces, there is $\delta > 0$ such that, for all $x \in X$,

$$d(x_0, x) < \delta \quad \Longrightarrow \quad \Delta(f(x_0), f(x)) < \epsilon, \tag{A.1}$$

i.e.

$$B_\delta(x_0) \subset f^{-1}(B_\epsilon(f(x_0))) \subset f^{-1}(V) \subset f^{-1}(U).$$

Since $B_\delta(x_0)$ is an open set containing x_0, this means that $f^{-1}(U)$ is a neighborhood of x_0.

Suppose conversely that f is continuous at x_0 in the sense of Definition A.2.2. Let $\epsilon > 0$. Then $B_\epsilon(f(x_0))$ is a neighborhood of $f(x_0)$. By Definition A.2.2, $f^{-1}(B_\epsilon(f(x_0)))$ is a neighborhood of x_0, so that there is an open set $V \subset f^{-1}(B_\epsilon(f(x_0)))$ with $x_0 \in V$. From the definition of open sets in metric spaces, there is $\delta > 0$ such that

$$B_\delta(x_0) \subset V \subset f^{-1}(B_\epsilon(f(x_0))).$$

But this just states that (A.1) holds for all $x \in X$. $\square$

Exercise A.4 Show that, if X is a non-empty set equipped with the discrete topology, then each function from X into any topological space is continuous.

Exercise A.5 Let X be a non-empty set equipped with its chaotic topology. Describe the continuous functions on X into a metric space.

Theorem A.2.4 *Let X be a topological space, and let $(f_n)_{n=1}^\infty$ be a sequence of $\mathbb{F}$-valued continuous functions on X that converges uniformly to a function f on X. Then f is continuous.*

Proof Let $x_0 \in X$ be arbitrary, and let $U \subset \mathbb{F}$ be a neighborhood of $f(x_0)$. Without loss of generality (Why?), we may suppose that $U = B_\epsilon(f(x_0))$ for some $\epsilon > 0$. Since $f_n \to f$ uniformly on X, there is $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \frac{\epsilon}{3} \qquad (x \in X,\, n \geq N). \tag{A.2}$$

This means, in particular, that $f_N(x_0) \in U$. Let $V := B_{\frac{\epsilon}{3}}(f_N(x_0))$. Then V is a neighborhood of $f_N(x_0)$. By Definition A.2.2, this means that $f_N^{-1}(V)$ is a neighborhood of x_0. Hence, there is an open set $W \subset f_N^{-1}(V)$ with $x_0 \in W$.

For $x \in W$, we have

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \frac{\epsilon}{3} + |f_N(x) - f_N(x_0)| + \frac{\epsilon}{3}, \qquad \text{by (A.2)}, \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3}, \qquad \text{by the choice of } W, \\ &= \epsilon, \end{aligned}$$

i.e. $f(x) \in B_\epsilon(f(x_0))$. Hence, $W \subset f^{-1}(B_\epsilon(f(x_0)))$, so that $f^{-1}(B_\epsilon(f(x_0)))$ is a neighborhood of x_0. $\square$

A.3 (Local) compactness

You probably have already encountered the notion of compactness in the context of metric spaces. Since we have a notion of openness, in arbitrary topological spaces, we can define compact topological spaces through a finite covering property:

Definition A.3.1 A topological space K is called *compact* if for any family $\mathcal{U}$ of open subsets of K such that $K = \bigcup\{U : U \in \mathcal{U}\}$ there are $U_1, \ldots, U_n \in \mathcal{U}$ such that $K = U_1 \cup \cdots \cup U_n$.

Exercise A.6 Show that a topological space X is compact if and only if X has the *finite intersection property*, i.e. for each family $\mathcal{F}$ of closed subsets of X such that $\bigcap\{F : F \in \mathcal{F}\} = \varnothing$, there are $F_1, \ldots, F_n \in \mathcal{F}$ such that $F_1 \cap \cdots \cap F_n = \varnothing$.

The following theorem, which we state without proof, characterizes the compact metric spaces:

Theorem A.3.2 *Let X be a metric space. Then the following are equivalent:*

(i) *X is compact.*

(ii) *Every sequence in X has a convergent subsequence.*

Example A subset of $\mathbb{F}^N$ is compact precisely when it is closed and bounded (this is immediate from the Heine–Borel theorem).

We also need to speak of compactness of subsets of arbitrary topological spaces:

Definition A.3.3 Let (X, τ) be a topological space, and let $Y \subset X$ be non-empty. Then

$$\tau|_Y := \{Y \cap U : U \in \tau\}$$

is the *relative topology* on X induced by τ.

Exercise A.7 Let (X, τ) be a topological space, and let $Y \subset X$ be non-empty. Show that $(Y, \tau|_Y)$ is a topological space.

When saying that a certain subset of a topological space is compact (or has some other topological property), we just mean that it is compact (or has that other property) with respect to its relative topology.

Theorem A.3.4 *Let K be a compact, topological space, let Y be a topological space, and let $f\colon K \to Y$ be continuous. Then $f(K)$ is compact.*

Proof Let $\mathcal{U}$ be a family of open sets of X such that $f(K) \subset \bigcup\{U : U \in \mathcal{U}\}$. Then $\{f^{-1}(U) : U \in \mathcal{U}\}$ is a family of open subsets of K such that $K = \bigcup\{f^{-1}(U) : U \in \mathcal{U}\}$. Since K is compact, there are $U_1, \ldots, U_n \in \mathcal{U}$ such that

$$K = f^{-1}(U_1) \cup \cdots \cup f^{-1}(U_n)$$

and hence

$$f(K) \subset U_1 \cup \cdots \cup U_n.$$

From the definition of the relative topology on $f(K)$, this means that $f(K)$ is compact (in its relative topology). $\square$

Corollary A.3.5 *Let K be a compact topological space, and let $f\colon K \to \mathbb{R}$ be continuous. Then f is bounded and attains both its minimum and its maximum on K.*

Exercise A.8 Derive Corollary A.3.5 for Theorem A.3.4.

In a metric space, we can always separate two distinct points through disjoint open balls. The following definition is a generalization of this property of metric spaces:

Definition A.3.6 A topological space X is called a *Hausdorff space* if, for any $x, y \in X$ with $x \neq y$, there are open sets $U, V \subset X$ with $U \cap V = \varnothing$, $x \in U$, and $y \in V$.

Example Every metric space is a Hausdorff space.

Exercise A.9 Give an example of a compact topological space which is not a Hausdorff space.

Proposition A.3.7 *Let K be a compact topological space and let $F \subset K$ be closed. Then F is compact.*

Proof Let $\mathcal{U}$ be a family of open sets of K such that $F \subset \bigcup\{U : U \in \mathcal{U}\}$. It follows that

$$K = \bigcup\{U : U \in \mathcal{U}\} \cup (K \setminus F).$$

Since K is compact and $K \setminus F$ is closed, there are $U_1, \ldots, U_n \in \mathcal{U}$ such that

$$K = U_1 \cup \cdots \cup U_n \cup (K \setminus F)$$

and thus

$$F \subset U_1 \cup \cdots \cup U_n.$$

This complete the proof. $\square$

Proposition A.3.8 *Let X be a Hausdorff space, and let $K \subset X$ be compact. Then K is closed in X.*

Proof Let $x \in X \setminus K$. For each $y \in K$, there are thus open subsets U_y and V_y of X with $U_y \cap V_y = \varnothing$ and $x \in U_y$ and $y \in V_y$. Since $K \subset \bigcup\{V_y : y \in K\}$ and K is compact, there are $y_1, \ldots, y_n \in K$ such that $K = V_{y_1} \cup \cdots \cup V_{y_n}$. Let

$$W_x := U_{y_1} \cap \cdots \cap U_{y_n}.$$

Then W_x is open such that $x \in W_x$ and $U_x \subset X \setminus K$.

Since $x \in X \setminus K$ was arbitrary, we can define W_x for each such x. Consequently,

$$X \setminus K = \bigcup\{W_x : x \in X \setminus K\}$$

is open, so that K is closed. $\square$

Exercise A.10 Does Proposition A.3.8 remain true from compact subsets of non-Hausdorff s-paces.

The following theorem is often useful when it comes to establishing the continuity of a function:

Theorem A.3.9 *Let K be a compact topological space, let X be a Hausdorff space, and let $f\colon K \to X$ be continuous and bijective. Then $f^{-1}\colon X \to K$ is also continuous.*

Proof Let $F \subset K$ be closed. We have to show that $(f^{-1})^{-1}(F) = f(F)$ is closed. By Proposition A.3.7, F is compact and so is $f(F)$ by Theorem A.3.4. Since X is Hausdorff, this means that $f(F)$ is closed. $\square$

Exercise A.11 Give an example that shows that Theorem A.3.9 becomes false if we drop the demand that X be Hausdorff.

Definition A.3.10 Let X be a topological space, and let $S \subset X$ be arbitrary. Then the *closure* of S in X is defined as

$$\overline{S} = \bigcap \{F : F \subset X \text{ is closed with } S \subset F\}.$$

Definition A.3.11 Let X be a topological spaces. Then $S \subset X$ is called *relatively compact* if $S = \varnothing$ or if $\overline{S}$ is compact.

Definition A.3.12 A topological space X is called *locally compact* if for each point in X has a relatively compact neighborhood.

Example $\mathbb{F}^N$ is locally compact, but not compact.

Definition A.3.13 Let X be a locally compact Hausdorff space. A continuous function $f : X \to \mathbb{F}$ is said to *vanish at infinity* if, for each $\epsilon > 0$, there is a compact subset K of X such that $|f(x)| < \epsilon$ for all $x \in X \setminus K$. We write $\mathcal{C}_0(X, \mathbb{F})$ for the linear space of all continuous function on X into $\mathbb{F}$ that vanish at infinity.

Exercise A.12 Show that for $X = \mathbb{R}$ this yields the usual definition of a function vanishing at infinity, i.e.

$$f \in \mathcal{C}_0(\mathbb{R}, \mathbb{F}) \quad \iff \quad \lim_{t \to \infty} f(t) = 0.$$

We state the following theorem without proof; (i) is a consequence of Urysohn's lemma, and (ii) is proved on page 4.

Theorem A.3.14 *Let X be a locally compact Hausdorff space. Then:*

(i) *For each compact subset K of X and each closed subset F of X such that $K \cap F = \varnothing$, there is $f \in \mathcal{C}_0(X, \mathbb{R})$ such that $f|_F \equiv 0$ and $f|_K \equiv 1$.*

(ii) *$(\mathcal{C}_0(X, \mathbb{F}), \|\cdot\|_\infty)$ is a Banach space.*

Exercise A.13 Prove Theorem A.3.14(ii).

A.4 Convergence of nets

In metric spaces, topological concepts such as closedness and continuity can be characterized through convergent sequences. This is not possible anymore for topological spaces (see below), but there is an appropriate substitute:

Definition A.4.1 A non-empty set $\mathbb{A}$ is called *directed* if there is an oder relation $\prec$ on $\mathbb{A}$ such that:

(a) If $\alpha \prec \beta$ and $\beta \prec \gamma$, then $\alpha \prec \gamma$.

(b) If $\alpha \prec \beta$ and $\beta \prec \alpha$, then $\alpha = \beta$.

(c) For any $\alpha, \beta \in \mathbb{A}$, there is $\gamma \in \mathbb{A}$ such that $\alpha \prec \gamma$ and $\beta \prec \gamma$.

Definition A.4.2 A *net* in a non-empty set S is a function from a directed set into S.

Example All sequences are nets.

If $\mathbb{A}$ is a directed set and $s \colon \mathbb{A} \to S$ is a net, we denote s by $(s_\alpha)_{\alpha\in\mathbb{A}}$ and write s_α for $s(\alpha)$; if the index set $\mathbb{A}$ is obvious or irrelevant, we often write $(s_\alpha)_\alpha$.

Definition A.4.3 Let X be a topological space, let $x \in X$, and let $(x_\alpha)_\alpha$ be a net in X. Then $(x_\alpha)_\alpha$ *converges* to x — in symbols $x = \lim_\alpha x_\alpha$ or $x_\alpha \stackrel{\alpha}{\to} x$ — if, for each neighborhood U of x, there is an index α such that $x_\beta \in U$ for each index β with $\alpha \prec \beta$.

Proposition A.4.4 *Let X a Hausdorff space, let $x, y \in X$, and let $(x_\alpha)_\alpha$ be a net in X that converges to both x and y. Then $x = y$.*

Proof Assume that $x \neq y$. Then there are disjoint open sets $U, V \subset X$ such that $x \in U$ and $y \in V$. Since $x = \lim_\alpha x_\alpha$, there is α_x such that $x_\alpha \in U$ for all α such that $\alpha_x \prec \alpha$. Since $y = \lim_\alpha x_\alpha$, there is α_y such that $x_\alpha \in V$ for all α such that $\alpha_y \prec \alpha$. Choose β such that $\alpha_x \prec \beta$ and $\alpha_y \prec \beta$. Then $x_\alpha \in U \cap V$ for all $\alpha \succ \beta$, which is impossible since $U \cap V = \varnothing$. $\square$

Exercise A.14 Let X be a nonempty set equipped with the chaotic topology. Show that every net in X converges to every point of X.

A.5 Closedness and continuity via nets

Theorem A.5.1 *Let X be a topological space, and let S be a non-empty subset of X. Then the following are equivalent for $x \in X$:*

(i) $x \in \overline{S}$;

(ii) *there is a net* $(x_\alpha)_\alpha$ *in* S *such that* $x = \lim_\alpha x_\alpha$.

Proof (i) $\Longrightarrow$ (ii): Let $\mathcal{N}_x$ denote the collection of all neighborhoods of x. For $U, V \in \mathcal{N}_x$ define:

$$U \prec V \quad :\Longleftrightarrow \quad U \supset V.$$

Then $\mathcal{N}_x$ is directed. By the definition of $\overline{S}$, there is, for each $U \in \mathcal{N}_x$, an element $x_U \in U \cap S$. Then $(x_U)_{U \in \mathcal{N}_x}$ is a net in S such that $x = \lim_U x_U$.

(ii) $\Longrightarrow$ (i): Let $(x_\alpha)_\alpha$ be a net in S such that $x = \lim_\alpha x_\alpha$, and assume that $x \in U := X \setminus \overline{S}$. Then there is α such that $x_\beta \in U \subset X \setminus S$ for $\beta \succ \alpha$, which is impossible. $\square$

This theorem is wrong for sequences:

Example Let X be an uncountable set, and define a subset of X as open if it is empty or has countable complement. It follows that a closed subset of X is the whole space or countable. Pick $x \in X$. Then the only closed set containing $S := X \setminus \{x\}$ is X, so that $\overline{S} = X$. Assume that there is a sequence $(x_n)_{n=1}^\infty$ in S such that $x_n \stackrel{n}{\to} x$. Then $U := X \setminus \{x_1, x_2, \ldots\}$ is an open neighborhood of x, but $x_n \notin U$ for all $n \in \mathbb{N}$.

Corollary A.5.2 *Let* X *be a topological space. Then the following are equivalent for a non-empty subset* F *of* X*:*

(i) F *is closed;*

(ii) *for each net* $(x_\alpha)_\alpha$ *in* F *that converges to* $x \in X$*, we have* $x \in F$.

Proof (i) $\Longrightarrow$ (ii): Let $(x_\alpha)_\alpha$ be a net in F with limit $x \in X$. Assume that $x \notin F$, i.e. $x \in U := X \setminus F$. Since U is a neighborhood of x, there is α such that $x_\beta \in U$ for $\beta \succ \alpha$. But this is impossible, since $(x_\alpha)_\alpha$ is a net in F.

(ii) $\Longrightarrow$ (i): It follows immediately from Theorem A.5.1 that $\overline{F} = F$, so that F is closed. $\square$

Theorem A.5.3 *Let* X *and* Y *be topological spaces, and let* $x \in X$. *Then the following are equivalent for a map* $f \colon X \to Y$*:*

(i) f *is continuous at* x*;*

(ii) *for each net* $(x_\alpha)_\alpha$ *in* X *such that* $x_\alpha \stackrel{\alpha}{\to} x$*, we have* $f(x_\alpha) \stackrel{\alpha}{\to} f(x)$.

Proof (i) $\Longrightarrow$ (ii): Let U be a neighborhood of $f(x)$. Then $f^{-1}(V) \subset f^{-1}(U)$ is open, so that $f^{-1}(U)$ is a neighborhood of x. Since $x = \lim_\alpha x_\alpha$, there is an index α such that $x_\beta \in f^{-1}(U)$ for $\beta \succ \alpha$. But this means that $f(x_\beta) \in U$ for $\beta \succ \alpha$.

(ii) $\Longrightarrow$ (i): Let U be a neighborhood of $f(x)$, and assume towards a contradiction that $f^{-1}(U)$ is not a neighborhood of x. Hence, $V \not\subset f^{-1}(U)$ for each open subset V of X with x in V. Let $\mathcal{V}_x$ denote the collection of all open subsets of X containing X. Then $\mathcal{V}_x$ is directed in a natural way. By assumptions, we can choose $x_V \in V \setminus f^{-1}(U)$ for each $V \in \mathcal{V}_x$. It is clear that $\lim_V x_V = x$ (Why?). But since $f(x_V) \notin U$ for all $V \in \mathcal{V}_x$, it follows that $f(x_V) \not\to f(x)$. $\square$

A.6 Compactness via nets

Definition A.6.1 Let $\mathbb{A}$ and $\mathbb{B}$ be directed sets. A map $\phi \colon \mathbb{B} \to \mathbb{A}$ is called *cofinal* if, for each $\alpha \in \mathbb{A}$, there is $\beta \in \mathbb{B}$ such that $\phi(\beta) \succ \alpha$.

Definition A.6.2 Let X be a non-empty set, and let $(x_\alpha)_{\alpha\in\mathbb{A}}$ and $(y_\beta)_{\beta\in\mathbb{B}}$ be nets in X. Then $(y_\beta)_{\beta\in\mathbb{B}}$ is a *subnet* of $(x_\alpha)_{\alpha\in\mathbb{A}}$ if $y_\beta = x_{\phi(\beta)}$ for a cofinal map $\phi \colon \mathbb{A} \to \mathbb{B}$.

Exercise A.15 Does a subnet of a sequence have to be again a sequence?

Proposition A.6.3 *Let X be a topological space, let $(x_\alpha)_\alpha$ be a net in X, and let $x \in X$ be a limit of $(x_\alpha)_\alpha$. Then each subnet of $(x_\alpha)_\alpha$ converges to x.*

Exercise A.16 Prove Proposition A.6.3.

Definition A.6.4 Let X be a topological space, and let $(x_\alpha)_\alpha$ be a net in X. A point $x \in X$ is an *cluster point* of $(x_\alpha)_\alpha$ if, for each α and for each neighborhood U of x, there is $\beta \succ \alpha$ such that $x_\beta \in U$.

Proposition A.6.5 *Let X be a topological space, and let $(x_\alpha)_\alpha$ be a net in X. Then the following are equivalent for $x \in X$:*

(i) *x is an cluster point of $(x_\alpha)_\alpha$;*

(ii) *there is a subnet of $(x_\alpha)_\alpha$ converging to x.*

Proof (i) $\Longrightarrow$ (ii): Let $\mathcal{N}_x$ denote the collection of all neighborhoods of x. Let $\mathbb{B} := \mathbb{A}\times\mathcal{N}_x$. For $(\alpha_1, U_1), (\alpha_2, U_2) \in \mathbb{B}$ define:

$$(\alpha_1, U_1)\tilde{\prec}(\alpha_2, U_2) \quad :\Longleftrightarrow \quad \alpha_1 \prec \alpha_2 \text{ and } U_1 \supset U_2.$$

This turns $\mathbb{B}$ into a directed set. Let $(\alpha, U) \in \mathbb{B}$. By the definition of an cluster point, there is $\phi(\alpha, U) \in \mathbb{A}$ with $\phi(\alpha, U) \succ \alpha$ such that $x_{\phi(\alpha,U)} \in U$. The map $\phi : \mathbb{B} \to \mathbb{A}$ is cofinal, and the net $(x_{\phi(\alpha,U)})_{(\alpha,U)\in\mathbb{B}}$ converges to x.

(ii) $\Longrightarrow$ (i): Clear by definition. $\square$

We can now prove the analogue of Theorem A.3.2 for general topological spaces:

Theorem A.6.6 *For a topological space X the following are equivalent:*

(i) *X is compact;*

(ii) *each net in X has a convergent subnet.*

Proof (i) $\Longrightarrow$ (ii): Let $(x_\alpha)_\alpha$ be a net in X. By Proposition A.6.5, it is sufficient to show that $(x_\alpha)_\alpha$ has an cluster point. Assume that $(x_\alpha)_\alpha$ has no cluster point. Then, for each $x \in X$, there is a neighborhood U_x of x (which we can choose to be open) and an index α_x such that $x_\beta \notin U_x$ for $\beta \succ \alpha_x$. The family $(U_x)_{x\in X}$ is an open cover of X and thus has a finite subcover $\{U_{x_1}, \ldots, U_{x_n}\}$. Chose an index α such that $\alpha \succ \alpha_j$ for $j = 1, \ldots, n$. Hence, for $\beta \succ \alpha$

$$x_\beta \notin U_{x_1} \cup \cdots \cup U_{x_n} = X,$$

which is absurd.

(ii) $\Longrightarrow$ (i): Assume that K is not compact. Then there is an open cover $\mathfrak{U}$ which has no finite subcover. Let $\mathcal{F}(\mathfrak{U})$ be the collection of all finite subsets of $\mathfrak{U}$ ordered by set inclusion. For each $\mathcal{U} \in \mathcal{F}(\mathfrak{U})$, there is

$$x_{\mathcal{U}} \in X \setminus \bigcup\{U : U \in \mathcal{U}\} = \bigcap\{X \setminus U : U \in \mathcal{U}\}$$

(otherwise, $\mathfrak{U}$ would have a finite subcover). By hypothesis, the net $(x_{\mathcal{U}})_{\mathcal{U}\in\mathcal{F}(\mathfrak{U})}$ has an cluster point $x \in X$. For any $U \in \mathfrak{U}$, and for any open neighborhood V of x, there is $\mathcal{V} \succ \{U\}$ with $x_{\mathcal{V}} \in V$ and thus $V \cap X \setminus U \neq \varnothing$. Assume that $x \notin X \setminus U$. Then $x \in U$, so that U would be an open neighborhood of x; by the foregoing, we would have $U \cap X \setminus U \neq \varnothing$, which is absurd. It follows that $x \in X \setminus U$. Since $U \in \mathfrak{U}$ is arbitrary, we have

$$x \in \bigcap\{X \setminus U : U \in \mathfrak{U}\} = X \setminus \bigcup\{U : U \in \mathfrak{U}\} = \varnothing,$$

which is again absurd. $\square$

A.7 Tychonoff's theorem

Tychonoff's theorem is possibly the deepest theorem in point set topology. It states that compactness is preserved under arbitrary Cartesian products.

If τ and σ are two topologies, the τ is called *coarser* than σ if τ has fewer open sets then σ (σ is then called *finer* than τ).

Definition A.7.1 Let $(X_i)_{i\in\mathbb{I}}$ be a family of topological spaces, let $X := \prod_{i\in\mathbb{I}} X_i$, and let $\pi_i \colon X \to X_i$ be the projection onto the i-th coordinate. The *product topology* on X is the coarsest topology such that the projections π_i are all continuous.

Lemma A.7.2 *Let $(X_i)_{i\in\mathbb{I}}$ be a family of topological spaces, and let $X := \prod_{i\in\mathbb{I}} X_i$. Then the open subsets of X in the product topology are exactly the unions of sets of the form*

$$\pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_n}^{-1}(U_{i_n}), \tag{A.3}$$

where $n \in \mathbb{N}$, $i_1, \ldots, i_n \in \mathbb{I}$, and $U_{i_1} \subset X_{i_1}, \ldots, U_{i_n} \subset X_{i_n}$ are open.

Proof The collection of all union of sets of the form (A.3) is a topology that makes the projections continuous; hence, each open subset of X in the product topology is of that form.

Conversely, any topology making the projections continuous, must contain the sets of the form (A.3) and thus their arbitrary unions. □

Proposition A.7.3 *Let $(X_i)_{i\in\mathbb{I}}$ be a family of topological spaces, and let $X := \prod_{i\in\mathbb{I}} X_i$. The the following are equivalent for a net $(x_\alpha)_\alpha$ in X and a point $x \in X$:*

(i) $x_\alpha \xrightarrow{\alpha} x$ *in the product topology;*

(ii) $\pi_i(x_\alpha) \xrightarrow{\alpha} \pi_i(x)$ *for each $i \in \mathbb{I}$.*

Proof (i) $\Longrightarrow$ (ii): This is clear by Theorem A.5.3, since the projections are continuous.

(ii) $\Longrightarrow$ (i): Let U be a neighborhood of $x \in X$. By Lemma A.7.2, there are $n \in \mathbb{N}$, $i_1, \ldots, i_n \in \mathbb{I}$, and open sets $U_{i_1} \subset X_{i_1}, \ldots, U_{i_n} \subset X_{i_n}$ with

$$x \in \pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_n}^{-1}(U_{i_n}) \subset U.$$

By hypothesis, there is α such that $\pi_{i_j}(x_\beta) \in U_{i_j}$ for $\beta \succ \alpha$ and $j = 1, \ldots, n$. This, however, means that

$$x_\beta \in \pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_n}^{-1}(U_{i_n}) \subset U$$

for $\beta \succ \alpha$. □

Exercise A.17 Let $(X_i)_{i\in\mathbb{I}}$ be a family of Hausdorff spaces. Show that $\prod_{i\in\mathbb{I}} X_i$ equipped with the product topology is also Hausdorff.

Theorem A.7.4 (Tychonoff's theorem) *Let $(X_i)_{i\in\mathbb{I}}$ be a family of compact topological spaces. Then $X := \prod_{i\in\mathbb{I}} X_i$ equipped with the product topology is also compact.*

Proof Let $(x_\alpha)_\alpha$ be a net in X. Let $\mathbb{J} \subset \mathbb{I}$; we call an element $x \in X$ a $\mathbb{J}$-*partial cluster point* of $(x_\alpha)_\alpha$ if $x|_\mathbb{J}$ is an cluster point of $(x_\alpha|_\mathbb{J})_\alpha$ in $\prod_{i \in \mathbb{J}} X_i$. We call $x \in X$ a *partial cluster point* of $(x_\alpha)_\alpha$ if it is a $\mathbb{J}$-partial cluster point of for some $\mathbb{J} \subset \mathbb{I}$; $\mathbb{J}$ is called the *domain* of x.

Let $\mathcal{P}$ be the set of all partial cluster points of $(x_\alpha)_\alpha$. Let $x_1, x_2 \in \mathcal{P}$. We define

$$x_1 \leq x_2 \quad :\Longleftrightarrow \quad \text{domain of } x_1 \subset \text{domain of } x_2 \text{ and } x_2|_{\text{domain of } x_1} = x_1.$$

Since each X_i is compact, $(x_\alpha)_\alpha$ has $\{i\}$-partial cluster points for each $i \in \mathbb{I}$.

Let $\mathcal{K}$ be a totally ordered subset of $\mathcal{P}$. Let $\mathbb{J} := \bigcup\{\text{domain of } x : x \in \mathcal{K}\}$. Define $y \in \prod_{j \in \mathbb{J}} X_j$ by letting $y(j) := x(j)$ if $j \in$ domain of x. Since $\mathcal{K}$ is totally ordered, y is well defined. We claim that y is a $\mathbb{J}$-partial cluster point of $(x_\alpha)_\alpha$. Let $U \subset \prod_{j \in \mathbb{J}} X_j$ be a neighborhood of y. By Lemma A.7.2, we may suppose that

$$U = \pi_{j_1}^{-1}(U_{j_1}) \cap \cdots \cap \pi_{j_n}^{-1}(U_{j_n}),$$

where $n \in \mathbb{N}$, $j_1, \ldots, j_n \in \mathbb{J}$, and $U_{j_1} \subset X_{j_1}, \ldots, U_{j_n} \subset X_{j_n}$ are open. Clearly, for each α there is $\beta \succ \alpha$ such that

$$x_\beta(j_k) = \pi_j(x_\beta) \in U_{j_k} \qquad (k = 1, \ldots, n),$$

so that $x_\beta \in U$.

By Zorn's lemma, $\mathcal{P}$ has a maximal element x with domain $\mathbb{J}$. Assume there is $i \in \mathbb{I} \setminus \mathbb{J}$. There is a subnet $(x_{\alpha_\beta})_\beta$ of $(x_\alpha)_\alpha$ such that $\pi_j(x_{\alpha_\beta}) \xrightarrow{\beta} \pi_j(x)$ for each $j \in \mathbb{J}$. Since X_i is compact, we may find a subnet $(x_{\alpha_{\beta_\gamma}})_\gamma$ of $(x_{\alpha_\beta})_\beta$ such that $\pi_i(x_{\alpha_{\beta_\gamma}})_\gamma$ converges to some x_i in X_i. Define $\tilde{x} \in \prod_{j \in \mathbb{J} \cup \{i\}} X_j$ by letting $\tilde{x}|_\mathbb{J} = x$ and $\tilde{x}(i) = x_i$. It follows that $\tilde{x}$ is a $\mathbb{J} \cup \{i\}$-partial cluster point of $(x_\alpha)_\alpha$, which contradicts the maximality of x. $\square$

Appendix B

Measure and integration

Like point set topology, measure theory is an important source of examples in functional analysis.

In this appendix, I have collected the definitions and results we need. Proofs are not given.

B.1 Measure spaces

Definition B.1.1 Let Ω be a set. A *σ-algebra* over Ω is collection $\mathfrak{S}$ of subsets of Ω such that the following are satisfied:

(a) $\Omega \in \mathfrak{S}$;

(b) if $A \in \mathfrak{S}$, then $A^c \in \mathfrak{S}$;

(c) if $(A_n)_{n=1}^\infty$ is a sequence in $\mathfrak{S}$, then $\bigcup_{n=1}^\infty A_n \in \mathfrak{S}$.

The pair $(\Omega, \mathfrak{S})$ is called a *measurable space*.

Examples 1. $\mathfrak{P}(\Omega)$ is a σ-algebra.

2. $\{A \subset \Omega : A \text{ or } A^c \text{ is countable}\}$ is a σ-algebra.

3. $\{A \subset \Omega : A \text{ or } A^c \text{ is finite}\}$ is not a σ-algebra if Ω is infinite.

4. If $\mathcal{S} \subset \mathfrak{P}(\Omega)$ is arbitrary, there is a smallest σ-algebra over Ω containing $\mathcal{S}$; this σ-algebra is called the σ-algebra generated by $\mathcal{S}$. If Ω is a topological space, the σ-algebra generated by its open subsets is called the *Borel σ-algebra* over Ω; we denote it by $\mathfrak{B}(\Omega)$.

Definition B.1.2 Let $(\Omega, \mathfrak{S})$ be a measurable space. A (positive) *measure* on $(\Omega, \mathfrak{S})$ is a function $\mu\colon \mathfrak{S} \to [o, \infty)$ such that:

(a) $\mu(\varnothing) = 0$;

(b) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ for each sequence $(A_n)_{n=1}^{\infty}$ of pairwise disjoint sets in $\mathfrak{S}$.

The triple $(\Omega, \mathfrak{S}, \mu)$ is called a *measure space*.

Examples 1. *Counting measure*: Ω any set; $\mathfrak{S} = \mathcal{P}(\Omega)$; $\mu(A) := |A|$.

2. *Dirac measure*: Ω any set with $\omega \in \Omega$ fixed; $\mathfrak{S} = \mathcal{P}(\Omega)$; $\mu = \delta_\omega$, i.e. $\mu(A) = 1$ if $\omega \in A$, otherwise $\mu(A) = 0$.

3. *N-dimensional Lebesgue measure*: $\Omega = \mathbb{R}^N$; $\mathfrak{S} = \mathfrak{B}(\mathbb{R}^N)$; $\mu = \lambda^N$, i.e. N-dimensional Lebesgue measure.

Definition B.1.3 A measure space $(\Omega, \mathfrak{S}, \mu)$ (or rather the measure μ) is called:

(a) *σ-finite* if there is a sequence $(A_n)_{n=1}^{\infty}$ in $\mathfrak{S}$ such that $\Omega = \bigcup_{n=1}^{\infty} A_n$ and $\mu(A_n) < \infty$ for each n;

(b) *finite* if $\mu(\Omega) < \infty$;

(c) a *probability space* (or rather *probability measure*) if $\mu(\Omega) = 1$.

Examples 1. N-dimensional Lebesgue measure is σ-finite, but not finite.

2. Any Dirac measure is a probability measure.

3. Counting measure is finite if and only if Ω is finite, and σ-finite if and only if Ω is countable.

Definition B.1.4 Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space.

(a) A set $N \in \mathfrak{S}$ is called a *μ-zero set* if $\mu(N) = 0$.

(b) The *completion of $\mathfrak{S}$ with respect to μ* is defined as

$$\{S \subset \Omega : \text{there are } A, N \in \mathfrak{S} \text{ with } A \subset S \subset A \cup N \text{ and } N \text{ is a } \mu\text{-zero set}\}.$$

(c) A property is said to hold *μ-almost everywhere* (short: μ-a.e.) on Ω if there is a μ-zero set $N \in \mathfrak{S}$ such that the property in question holds on $\Omega \setminus N$.

Exercise B.1 Show that the completion of a σ-algebra with respect to a measure is again a σ-algebra.

Example The completion of $\mathfrak{B}(\mathbb{R}^N)$ with respect to λ^N is the σ-algebra of Lebesgue measurable sets.

B.2 Definition of the integral

Definition B.2.1 Let $(\Omega, \mathfrak{S})$ be a measurable space. A function $f : \Omega \to \mathbb{R}$ is called *elementary* if there are $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathfrak{S}$ such that

$$f = \sum_{k=1}^{n} \alpha_k \chi_{A_k}.$$

Definition B.2.2 Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, and let $f : \Omega \to \mathbb{R}$ be an elementary function. The integral of f with respect to μ is defined as

$$\int f \, d\mu := \sum_{k=1}^{n} \alpha_k \mu(A_k),$$

where $f = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$ with $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathfrak{S}$.

Remark It can be shown that the value $\int f \, d\mu$ is independent of the representation $f = \sum_{k=1}^{n} \alpha_k \chi_{A_k}$.

Definition B.2.3 Let $(\Omega, \mathfrak{S})$ be a measurable space. A function $f : \Omega \to \mathbb{R} \cup \{\infty\}$ is called $\mathfrak{S}$-*measurable* if $\{\omega \in \Omega : f(\omega) \leq \alpha\rangle\} \in \mathfrak{S}$ for each $\alpha \in \mathbb{R}$.

Examples

1. Every elementary function is measurable.
2. If Ω is any topological space, then every continuous function $f : \Omega \to \mathbb{R}$ is Borel-measurable.
3. Every increasing function $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is Borel-measurable.
4. Measurability is preserved under taking *pointwise* suprema, infima, and limits.

Proposition B.2.4 *Let $(\Omega, \mathfrak{S})$ be a measurable space, and let $f : \Omega \to [0, \infty]$ be $\mathfrak{S}$-measurable. Then there is an increasing sequence $(f_n)_{n=1}^{\infty}$ of $\mathfrak{S}$-measurable functions on Ω that converges to f pointwise; in case f is bounded, we can even chose that sequence in such a way that we have uniform convergence.*

Definition B.2.5 Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, let $f : \Omega \to [0, \infty]$ be $\mathfrak{S}$-measurable, and let $(f_n)_{n=1}^{\infty}$ be as in Proposition B.2.4. Then the integral of f with respect to μ is defined as

$$\int f \, d\mu := \lim_{n \to \infty} \int f_n \, d\mu.$$

Remarks

1. It can be shown that the value $\int f \, d\mu$ is independent of the choice of the sequence $(f_n)_{n=1}^{\infty}$.

2. As the limit of an increasing sequence $\int f\,d\mu$ always exists, but may be ∞.

3. If $\int f\,d\mu < \infty$, then $\{\omega \in \Omega : f(\omega) = \infty\}$ is a μ-zero set.

4. We have always $\int f\,d\mu \geq 0$, and $\int f\,d\mu = 0$ if and only if $f = 0$ μ-almost everywhere.

Definition B.2.6 Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space. A measurable function $f : \Omega \to \mathbb{R} \cup \{\infty\}$ is called *μ-integrable* if $\int f_+\,d\mu < \infty$ and $\int f_-\,d\mu < \infty$. The integral of f with respect to μ is defined as

$$\int f\,d\mu := \int f_+\,d\mu - \int f_-\,d\mu.$$

Remark Treating real and imaginary part separately, one can also define the integral of $\mathbb{C}$-valued functions.

Proposition B.2.7 *Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, and let*

$$\mathcal{L}^1(\Omega, \mathfrak{S}, \mu) := \{f\colon \Omega \to \mathbb{R} : f \text{ is integrable}\}.$$

Then:

(i) *$\mathcal{L}^1(\Omega, \mathfrak{S}, \mu)$ is a linear space;*

(ii) *the integral is linear on $\mathcal{L}^1(\Omega, \mathfrak{S}, \mu)$, i.e.*

$$\int (\alpha f + \beta g)\,d\mu = \alpha \int f\,d\mu + \beta \int g\,d\mu \qquad (\alpha, \beta \in \mathbb{R},\ f, g \in \mathcal{L}^1(\Omega, \mathfrak{S}, \mu));$$

(iii) *the integral is positive, i.e. if $f \geq 0$ μ-a.e. for some $f \in \mathcal{L}^1(\Omega, \mathfrak{S}, \mu)$, then $\int f\,d\mu \geq 0$.*

Examples 1. For $(\mathbb{R}^N, \mathfrak{B}(\mathbb{R}^N), \lambda^N)$ we get the familiar N-dimensional Lebesgue integral.

2. For $(\Omega, \mathfrak{P}(\Omega), \delta_\omega)$, every function $f\colon \Omega \to \mathbb{R}$ is integrable, and we have

$$\int f\,\delta_\omega = f(\omega) \qquad (f \in \mathcal{L}^1(\Omega, \mathfrak{S}, \mu)).$$

3. For $(\mathbb{N}, \mathfrak{P}(\mathbb{N}), \mu)$ with μ counting measure, a function $f\colon \mathbb{N} \to \mathbb{R}$ is integrable if and only if the series $\sum_{n=1}^{\infty} f(n)$ converges absolutely; in this case, we have

$$\int f\,d\mu = \sum_{n=1}^{\infty} f(n).$$

B.3 Theorems about the integral

The main advantage the Lebesgue integral has over the Riemann integral is the ease with which it can be interchanged with pointwise limits. These limit theorems hold in a more abstract measure theoretic context:

Theorem B.3.1 (monotone convergence theorem) *Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, let $(f_n)_{n=1}^\infty$ be an increasing sequence of $[0,\infty]$-valued, $\mathfrak{S}$-measurable functions on Ω, and let $f\colon \Omega \to [0,\infty]$ be their pointwise limit. Then*

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Theorem B.3.2 (dominated convergence theorem) *Let $(\Omega, \mathfrak{S}, \mu)$ be a measure space, let $(f_n)_{n=1}^\infty$ be a sequence of $\mathbb{R} \cup \{\infty\}$-valued, μ-integrable functions functions on Ω, and let $f, g\colon \Omega \to \mathbb{R} \cup \{\infty\}$ be such that:*

(a) *$f = \lim_{n\to\infty} f_n$ μ-a.e.;*

(b) *g is μ-integrable;*

(c) *$|f_n| \leq g$ μ-a.e. for all $n \in \mathbb{N}$.*

Then f is μ-integrable with

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Definition B.3.3 Let $(\Omega, \mathfrak{S})$ be a measurable space, and let μ and ν be measures on $(\Omega, \mathfrak{S})$. Then ν is said to be *absolutely continuous* with respect to μ (in symbols: $\nu \ll \mu$) if every μ-zero set is already a ν-zero set.

Examples

1. Let $(\Omega, \mathfrak{S})$ be a measure space, let μ be a measure on $(\Omega, \mathfrak{S})$, and let $f\colon \Omega \to [0,\infty]$ be $\mathfrak{S}$-measurable. Define $\nu\colon \mathfrak{S} \to [0,\infty]$ through

$$\nu(A) := \int f\chi_A \, d\mu \qquad (A \in \mathfrak{S}).$$

Then ν is a measure on $(\Omega, \mathfrak{S})$ (which is finite if and only if f is μ-integrable) such that $\nu \ll \mu$.

2. Let $\alpha \in BV[a,b]$. Then there is a unique measure μ on $([a,b], \mathfrak{B}([a,b]))$ such that

$$\mu([c,d)) = \alpha(d) - \alpha(c) \qquad\qquad (c, d \in [a,b])$$

(the integral with respect to μ is just the Riemann–Stieltjes integral with respect to α). The measure μ is absolutely continuous with respect to Lebesgue measure if and only if α is absolutely continuous.

Theorem B.3.4 (Radon–Nikodým theorem) *Let $(\Omega, \mathfrak{S})$ be a measurable space, and let μ and ν be measures on $(\Omega, \mathfrak{S})$ such that $\nu \ll \mu$ and μ is σ-finite. Then there is a $\mathfrak{S}$-measurable function $f \colon \Omega \to [0, \infty]$ such that*

$$\nu(A) := \int f \chi_A \, d\mu \qquad (A \in \mathfrak{S}).$$

Any two such functions must be equal μ-a.e..

Remark It is necessary that μ be σ-finite in the Radon–Nikodým theorem: A straightforward counterexample for non-σ-finite μ is $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ with counting measure as μ and Lebesgue measure as ν.

Definition B.3.5 Let $(\Omega, \mathfrak{S})$ be a measurable space. A *complex measure* on $(\Omega, \mathfrak{S})$ is a function $\mu \colon \mathfrak{S} \to \mathbb{C}$ such that:

(a) $\mu(\varnothing) = 0$;

(b) $\mu\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu(A_n)$ for each sequence $(A_n)_{n=1}^\infty$ of pairwise disjoint sets in $\mathfrak{S}$.

Remark Every complex measure μ has a so called *Jordan decomposition*, i.e. there are (with some strings attached) uniquely determined finite measures μ_1, μ_2, μ_3, and μ_4 such that $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$. A function is said to be μ-integrable if it is μ_j-integrable for $j = 1, \ldots, 4$. The integral of a μ-integrable function f is then defined as

$$\int f \, d\mu := \sum_{j=1}^4 \int g \, d\mu_j.$$

Definition B.3.6 Let $(\Omega, \mathfrak{S})$ be a measurable space, and let μ be a complex measure on $(\Omega, \mathfrak{S})$. The *total variation* of μ is defined as

$$\|\mu\| := \sup \left\{ \sum_{j=1}^n |\mu(A_j)| : n \in \mathbb{N},\, A_j \cap A_k = \varnothing \text{ for } j \neq k,\, \Omega = \bigcup_{j=1}^n A_j \right\}.$$

Definition B.3.7 Let Ω be a locally compact space. A (positive) measure μ on $(\Omega, \mathfrak{B}(\Omega))$ is called *regular* if

(a) $\mu(K) < \infty$ for all compact $K \subset \Omega$,

(b) $\mu(A) = \inf\{\mu(U) : A \subset U \subset \Omega \text{ with } U \text{ open}\}$ for all $A \in \mathfrak{B}(\Omega)$, and

(c) $\mu(U) = \sup\{\mu(K) : K \subset U \text{ is compact}\}$ for all open $U \subset \Omega$.

A complex measure is called regular if all the measures occurring in its Jordan decomposition are regular. The collection of all regular, complex measures on $(\Omega, \mathfrak{B}(\Omega))$ is denoted by $M(\Omega)$.

Example N-dimensional Lebesgue measure is regular.

Theorem B.3.8 (Riesz representation theorem) *Let Ω be a locally compact space. Then $T\colon M(\Omega) \to \mathcal{C}_0(\Omega)^*$ with*

$$(T\mu)(f) := \int f\, d\mu \qquad (\mu \in M(\Omega),\ f \in \mathcal{C}_0(\Omega))$$

is a linear bijection such that

$$\|T\mu\| = \|\mu\| \qquad (\mu \in M(\Omega)).$$

Remark The Riesz representation theorem is also valid over $\mathbb{R}$.