Introduction to Functional Analysis

Michael Müger

24.02.2024

Abstract

These are notes for my Bachelor course Inleiding in de Functionaalanalyse (14×90 min.). They are also recommended as background for my Master courses on Operator Algebras.

Some familiarity with metric and topological spaces is assumed, and the last lecture (Section 18) will use some measure theory. Complex analysis is not used.

Contents

1 Introduction

We will begin with a quick delineation of what we will discuss – and what not!

- "Classical analysis" is concerned with 'analysis in finitely many dimensions'. 'Functional analysis' is the generalization or extension of classical analysis to infinitely many dimensions. Before one can try to make sense of this, one should make the first sentence more precise. Since the creation of general topology, one can talk about convergence and continuity in very general terms. As far as I see it, this is not analysis, even if infinite sums (=series) are studied. Analysis proper starts as soon as one talks about differentiation and/or integration. Differentiation has to do with approximating functions locally by linear ones, and for this one needs the spaces considered to be vector spaces (at least locally). This is the reason why most of classical analysis considers functions between the vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ (or subsets of them). (In a second step, one can then generalize to spaces that look like $\mathbb{R}^n$ only locally by introducing topological and smooth manifolds and their generalizations, but the underlying model of $\mathbb{R}^n$ remains important.) On the other hand, integration, at least in the sense of the modern theory, can be studied much more generally, i.e. on arbitrary sets equipped with a measure (defined on some σ-algebra). Such a set can be very far from being a vector space or manifold, for example by being totally disconnected.
- In view of the above, it is not surprising that functional analysis is concerned with (possibly) infinite-dimensional vector spaces and continuous maps between them. (Again, one can then generalize to spaces that look like a vector space only locally, but this would be considered infinite-dimensional geometry, not functional analysis.) In addition to the vector space structure one needs a topology, which naturally leads to topological vector spaces, which we will define soon.

- The importance of topologies is not specific to infinite dimensions. The point rather is that $\mathbb{R}^n, \mathbb{C}^n$ have unique topologies making them Hausdorff topological vector spaces. This is no more true in infinite dimensions!
- Actually, 'functional analysis' most often studies only linear maps between topological vector spaces so that this domain of study should be called 'linear functional analysis', but this is done only rarely, e.g. [145]. Allowing non-linear maps leads to non-linear functional analysis. This course will discuss only linear functional analysis. Thorough mastery of the latter is needed anyway before one can think about non-linear FA or infinite-dimensional geometry. For the simplest result of non-linear functional analysis, see Section B.15. For more, you could have a look at, e.g., [164, 37, 115]. There even is a five volume treatise [177]!
- The restriction to linear maps means that the notion of differentiation becomes point-less, the derivative of $f(x) = Ax$ being just A everywhere. But there are many non-trivial connections between linear FA and integration (and measure) theory. For example, every measure space $(X, \mathcal{A}, \mu)$ gives rise to a family of topological vector spaces $\mathcal{L}^p(X, \mathcal{A}, \mu)$, $p \in (0, \infty]$, and integration provides linear functionals. Proper appreciation of these matters requires some knowledge of measure and integration theory, cf. e.g. [29, 146]. I will not suppose that you have followed a course on this subject (but if you haven't, I strongly that you do so on the next occasion or, at least, read the appendix in MacCluer's book [101].). Yet, one can get a reasonably good idea by focusing on sequence spaces, for which no measure theory is required, see Section 4.
- One should probably consider linear functional analysis as an infinite-dimensional and topological version of linear algebra rather than as a branch of analysis! This might lead one to suspect linear FA to be slightly boring, but this would be wrong for many reasons:
  - Functional analysis (linear or not) leads to very interesting (and arbitrarily challenging) technical questions (most of which reduce to very easy ones in finite dimensions).
  - Linear FA is essential for non-linear FA, like variational calculus, and the theory of differential equations – not only linear ones!
  - Quantum theory [92] is a linear theory and cannot be done properly without functional analysis, despite the fact that many physicists think so! Conversely, many developments in FA were directly motivated by quantum theory.
- The above could give the impression that functional analysis arose from the wish of generalizing analysis to infinitely many dimensions. This may have played a role for some of its creators, but its beginnings (and much of what is being done now) were mostly motivated by finite-dimensional "classical"[1] analysis: If $U \subseteq \mathbb{R}^n$, the set of functions (possibly continuous, differentiable, etc.) from U to $\mathbb{R}^m$ is a vector space as soon as we put $(cf + dg)(x) = cf(x) + dg(x)$. Unless U is a finite set, this vector space will be infinite-dimensional. Now one can consider certain operations on such vector spaces, like differentiation $C^\infty(U) \to C^\infty(U), f \mapsto f'$ or integration $f \mapsto \int_U f$. This sort of considerations provided the initial motivation for the development functional analysis, and indeed FA now is a very important tool for the study of ordinary and partial differential equations on finite-dimensional spaces. See e.g. [23, 52]. The relevance of FA is even more obvious if one studies differential equations in infinitely many dimensions. In fact, it is often useful to study a partial differential equation (like heat or wave equation) by singling out one of the

[1]Ultimately, it is quite futile to try and draw a neat line between "classical" and "modern" or functional analysis, in particular since many problems in the former require methods from the latter for their proper treatment.

variables (typically 'time') and studying the equation as an ordinary differential equation in an infinite-dimensional space of functions. FA is also essential for variational calculus (which in a sense is just a branch of differential calculus in infinitely many dimensions).

- In view of the above, FA studies abstract topological vector spaces as well as 'concrete' spaces, whose elements are functions. In order to obtain a proper understanding of FA, one needs some familiarity with both aspects.

Before we delve into technicalities, some further general remarks:

- The history of functional analysis is quite interesting, cf. e.g. the article [15], [120, Chapter 4] and the books [40, 105, 122]. But clearly it makes little sense to study it before one has some technical knowledge of FA. It is surprisingly intertwined with the development of linear algebra. One would think that (finite-dimensional) vector spaces, linear maps etc. were defined much earlier than, e.g., Banach spaces, but this is not what has happened. In fact, Banach's[2] book [9], based on his 1920 PhD thesis, is one of the first references containing the modern definition of a vector space. Some mathematicians, mostly Italian ones, like Peano, Pincherle and Volterra, essentially had the modern definition already in the last decades of the 19th century, but their work had little impact since the usefulness of an abstract/axiomatic approach was not yet widely appreciated. Cf. [85, Chapter 5] or [41, 106].

  Here I limit myself to mentioning that the basics of functional analysis (Hilbert and Banach spaces and bounded linear maps between them) were developed in the period 1900-1930. Nevertheless, many important developments (locally convex spaces, distributions, operator algebras) took place in 1930-1960. After that, functional analysis has split up into many quite specialized subfields that interact quite little with each other. The very interesting article [166] ends with the conclusion that 'functional analysis' has ceased to exist as a coherent field of study!

- The study of functional analysis requires a solid background in general topology. It may well be that you'll have to refresh and likely also extend yours. In Appendix A I have collected brief accounts of the topics that – sadly – you are most likely not to have encountered before. All of them are treated in [56] or [142] (written by a functional analyst!), but my favorite (I'm admittedly biased) reference is [108]. You should have seen Weierstrass' theorem, but those of Tietze and Arzelà-Ascoli tend to vanish in the (pedagogical, not factual) gap between general topology and functional analysis.

- If you find these notes too advanced, you might want to have a look at less ambitious books like [116, 145, 148]. On the other hand, if you want more, [118] is a good place to start, followed by [128, 94, 30, 141]. (The Dutch MasterMath course currently uses [30].)

One word about notation (without guarantee of always sticking to it): General vector spaces, but also normed spaces, are denoted $V, W, \ldots$, normed spaces also as $E, F, \ldots$. Vectors in such spaces are $e, f, \ldots, x, y, \ldots$. Linear maps are always denoted $A, B, \ldots$, except linear functionals $V \to \mathbb{F}$, which are φ, ψ. Algebras are usually denoted $\mathcal{A}, \mathcal{B}, \ldots$ and their elements $a, b, \ldots$. (For $\mathcal{A} = B(E)$ this leads to inconsistency, but I cannot bring myself to using capital letters for abstract algebra elements.)

Note for experts. Our treatment deviates from the beaten path at a number of points. E.g. we

[2]Stefan Banach (1892-1945). Polish mathematician and pioneer of functional analysis. Also known for B. algebras, B.'s contraction principle, the B.-Tarski paradox and the Hahn-B. and B.-Steinhaus theorems, etc.

- emphasize that absolute and unconditional convergence of series are not the same thing in infinite-dimensional spaces, including a proof the Dvoretzky-Rogers theorem. Very strangely, most introductory books on functional analysis fail to point this out.
- following [61] we simplify the lengthy *ad hoc* argument in the standard proof of the open mapping theorem by using a lemma that also gives Tietze's extension theorem.
- include a fairly recent (2017) proof [53] of the uniform boundedness theorem (weak version) that only uses the axiom of countable choice, thus neither Baire's theorem nor the equivalent axiom of countable dependent choice (used by all previous 'elementary' proofs).
- and also show how Baire's theorem gives a stronger version of uniform boundedness, which is old but ignored by most authors.
- give a slick proof, inspired by [99], of the Hahn-Banach theorem using only Tychonov's theorem for Hausdorff spaces (equivalent to the Ultrafilter Lemma) instead of Zorn's lemma.
- follow [72] in proving the Arzelà-Ascoli theorem for complete metric spaces as target spaces and the Kolmogorov-Riesz compactness theorem (but only for ℓ^p).
- prove Pitt's compactness theorem without using bases, following [38].
- introduce characters of a Banach algebra, albeit without the weak-$*$ topology, at a relatively early stage. Among other things, this allows for a relatively painless extension of the continuous functional calculus for self-adjoint elements of a C^*-algebra to normal ones. This approach seems to be new. (Doing this via the full blown Gelfand isomorphism as e.g. in [101] seems inappropriate in a first course.)
- follow Rickart [130] in proving the Beurling-Gelfand formula for the spectral radius in a Banach algebra without using complex analysis (one reason being that the author cannot assume that all his students have been exposed to complex analysis). Since we also do the same on a few other occasions, like the Fuglede-Putnam theorem, the text does not assume any knowledge of complex analysis (apart from the elementary fact that power series converge on disks, which is not genuine complex analysis).
- we give a purely elementary construction of the Riesz projectors associated with isolated points of the spectrum (which we apply to compact non-normal operators and in the discussion of discrete and essential spectra).
- we similarly limit the use of measure theory to the absolute minimum until it becomes unavoidable in discussing the spectral theorem for normal operators. We prove all standard results on the Lebesgue spaces L^p for discrete measure spaces, i.e. sets S equipped with the counting measure, not limiting ourselves to $S = \mathbb{N}$. We then indicate how most proofs generalize to arbitrary measure spaces, while for the duality $(L^p)^* \cong L^q$ in the general case we give a proof using uniform convexity.
- touch upon, in Appendix B, a number of somewhat more advanced topics, for which there is no time in the author's lecture course, but which are very closely related to the core material. In particular, we go slightly further into Banach space theory and operator theory than many introductory texts.

Acknowledgment. I thank Bram Balkema, Victor Hissink Muller and Niels Vooijs for corrections, but in particular Tim Peters for a huge number of them.

Part I: Fundamentals

2 Setting the stage

2.1 Topological algebra: Topological groups, fields, vector spaces

As said in the Introduction, functional analysis (even most of the non-linear theory) is concerned with vector spaces, allowing infinite-dimensional ones. Large parts of linear algebra of course work equally well for finite and infinite-dimensional spaces. One aspect where problems arise in infinite dimensions is the description of linear maps by matrices, for example since multiplication of infinite matrices involves infinite summations, which require the introduction of topologies. (Actually, in some restricted contexts infinite matrices still are quite useful.)

We begin with the following

2.1 Definition *A topological group is a group $(G, \cdot, 1)$ equipped with a topology τ such that the group operations $G \times G \to G, (g,h) \mapsto gh$ and $G \to G, g \mapsto g^{-1}$ are continuous (where $G \times G$ is given the product topology). (For abelian groups, one often denotes the binary operation by $+$ instead of $\cdot$.)*

2.2 Example 1. If $(G, \cdot, 1)$ is any group then it becomes a topological group by putting $\tau = \tau_{\text{disc}}$, the discrete topology on G. (This is true since every function from a discrete space to a topological space is continuous.)

2. The group $(\mathbb{R}, +, 0)$, where $\mathbb{R}$ is equipped with its standard topology, is easily seen to be a topological group.

3. If $n \in \mathbb{N}$[3] and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ then the set $GL(n, \mathbb{F}) = \{A \in M_{n\times n}(\mathbb{F}) \mid \det(A) \neq 0\}$ of invertible $n \times n$ matrices is a group w.r.t. matrix product and inversion and in fact a topological group when equipped with the subspace topology induced from $M_{n\times n}(\mathbb{F}) \cong \mathbb{F}^{n^2}$.

2.3 Definition *A topological field is a field $(\mathbb{F}, +, 0, \cdot, 1)$ equipped with a topology on $\mathbb{F}$ such that $(\mathbb{F}, +, 0)$ and $(\mathbb{F}\backslash\{0\}, \cdot, 1)$ are topological groups. (Equivalently, all field operations are continuous.) Usually we just write $\mathbb{F}$ and denote the topology by $\tau_{\mathbb{F}}$.*

Again, if $\mathbb{F}$ is any field then $(\mathbb{F}, \tau_{\text{disc}})$ is a topological field.

2.4 Exercise Prove that $\mathbb{R}$ and $\mathbb{C}$ are topological fields when equipped with their standard topologies induced by the metric $d(c, c') = |c - c'|$.

2.5 Definition *Let $\mathbb{F}$ be a topological field. Then a topological vector space (TVS) over $\mathbb{F}$ is an $\mathbb{F}$-vector space V equipped with a topology τ_V (to be distinguished, obviously, from the topology $\tau_{\mathbb{F}}$ on $\mathbb{F}$) such that the maps $V \times V \to V, (x, y) \mapsto x + y$ and $\mathbb{F} \times V \to V, (c, x) \mapsto cx$ are continuous.*

(These conditions imply that $V \to V, x \mapsto -x$ is continuous, so that $(V, +, 0)$ is a topological group, but not conversely.)

2.6 Exercise Let $\mathbb{F}$ be a topological field.

(i) Prove that every $\mathbb{F}$-vector space is a a TVS over $\mathbb{F}$ if equipped with the indiscrete topology.

(ii) If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, prove that an $\mathbb{F}$-vector space $V \neq \{0\}$ equipped with the discrete topology is a TVS over $\mathbb{F}$ if and only if $\mathbb{F}$ is discrete.

[3] Throughout these notes, $\mathbb{N} = \{1, 2, 3, \ldots\}$, thus $0 \notin \mathbb{N}$.

(iii) If S is any set, let $V = \{f : S \to \mathbb{F}\} = \mathbb{F}^S$. With pointwise addition and scalar multiplication, V is an $\mathbb{F}$-vector space. Let $\tau_{\mathbb{F}}^S$ be the product topology on $\mathbb{F}^S$. Prove that $(V, \tau_{\mathbb{F}}^S)$ is a TVS over $\mathbb{F}$.

In this course, the only topological fields considered are $\mathbb{R}$ and $\mathbb{C}$. When a result holds for either of the two, we will write $\mathbb{F}$. (In part I of these notes it will hardly ever matter whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, but much of part II will require $\mathbb{F} = \mathbb{C}$.) Note however that one can consider topological vector spaces over other topological fields, like the p-adic ones $\mathbb{Q}_p$ [60]. (This said, the resulting p-adic functional analysis is quite different in some respects from the 'usual' one, cf. the comments in Section B.1 and the literature, e.g. [138, 125].) Now:

2.7 Definition *Functional analysis (ordinary, as opposed to p-adic) is concerned with topological vector spaces over $\mathbb{R}$ or $\mathbb{C}$ and continuous maps between them. Linear functional analysis considers only linear maps.*

As it turns out, the notion of topological vector spaces is a bit too general to base a satisfactory and useful theory upon it. We'll prove only one result (Proposition 2.29) in this setting. Just as in topology it is often (but by no means always!) sufficient to work with metric spaces, for most purposes it is usually sufficient to consider certain subclasses of topological vector spaces. The following diagram illustrates some of these classes and their relationships:

$$\begin{array}{ccccccc} \text{topological vector sp.} & \supset & \text{metrized/F-sp.} & & & & \\ \cup & & \cup & & & & \\ \text{locally convex sp.} & \supset & \text{Fréchet sp.} & \supset & \text{normed/Banach sp.} & \supset & \text{(pre-)Hilbert sp.} \end{array}$$

(Note that F-spaces, Fréchet[4], Banach and Hilbert[5] spaces are assumed complete but one also has the non-complete versions. There is no special name for Fréchet spaces with completeness dropped other than metrizable locally convex spaces. In the other cases, one speaks of metrized, normed and pre-Hilbert spaces.)

The most useful of these classes are those in the bottom row. In fact, locally convex (vector) spaces are general enough for almost all applications. They are thoroughly discussed in the MasterMath course on functional analysis, while we will only briefly touch upon them. Most of the time, we will be discussing Banach and Hilbert spaces. There is much to be said for studying them in some depth before turning to locally convex spaces (or more general) spaces. (Some books on functional analysis, like [141], begin with general topological vector spaces and then turn to some special classes, but for a first encounter this does not seem appropriate. This said, the author doesn't see the point of beginning by proving many results on Hilbert spaces that literally generalize to Banach spaces.)

2.2 Translation-invariant metrics. Normed and Banach spaces

I assume that you remember the notion of a metric on a set X: A map $d : X \times X \to [0, \infty)$ satisfying $d(x, y) = d(y, x)$ and $d(x, z) \le d(x, y) + d(y, z)$ for all $x, y, z \in X$ and $d(x, y) = 0 \Leftrightarrow x = y$. Every metric d on X defines a topology τ_d on X, the smallest topology τ containing all open balls $B(x, r) = \{y \in X \mid d(x, y) < r\}$. (The open balls then form a base[6], not just

[4] Maurice Fréchet (1878-1973). French mathematician. Introduced metric spaces in his 1906 PhD thesis.

[5] David Hilbert (1862-1943). Eminent German mathematician who worked on many different subjects. Considered the strongest and most influential mathematician in the decades around 1900, only Poincaré coming close.

[6] We write 'base' for the notion in topology and 'basis' for the one in linear algebra. The plural for both is 'bases'.

a subbase, for τ.) A topology τ on X is called metrizable if there exists a metric d on X (not necessarily unique) such that $\tau = \tau_d$. Metrizable topologies automatically have many nice properties, like e.g. normality and, a fortiori, the Hausdorff property. I also assume as familiar the notion of completeness of a metric space and the fact that that every metric space can be completed, i.e. embedded isometrically into a complete metric space (unique up to isometry) as a dense subspace.

On a vector space, it is natural and common to consider only metrics of a special type:

2.8 Definition *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.*

(i) A metric d on an $\mathbb{F}$-vector space V is called translation-invariant if it satisfies the equivalent statements

$$d(x,y) = d(x-z, y-z) \quad \forall x,y,z \in V \quad \Leftrightarrow \quad d(x,y) = d(x-y,0) \quad \forall x,y \in V. \tag{2.1}$$

(ii) A topological $\mathbb{F}$-vector space (V,τ) is called (completely) metrizable if there exists a translation-invariant (and complete) metric d on V such that $\tau = \tau_d$. (Completely metrizable TVS are also called F-spaces.)

2.9 Lemma *Let V be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and d a translation-invariant metric on V. Then addition $V \times V \to V, (x,y) \mapsto x+y$ and inversion $V \to V, x \mapsto -x$ are continuous, thus (V, τ_d) is a topological abelian group.*

Proof. If $x, x', y, y' \in V$ we have

$$d(x+y, x'+y') = d(x, x'+y'-y) \leq d(x,x') + d(x', x'+y'-y) = d(x,x') + d(y,y'),$$

where we used translation invariance twice and the triangle inequality once. This implies that the operation of addition $+ : V \times V \to V$ is jointly continuous. Continuity of the inverse operation $x \mapsto -x$ follows from $d(-x,-y) = d(0, x-y) = d(x-y, 0) = d(x,y)$. ■

2.10 Remark 1. If d is a translation-invariant metric on a vector space V, (V, τ_d) need not be a TVS since the scalar multiplication $\mathbb{F} \times V \to V$ can fail to be continuous! (This follows from Exercise 2.6(ii) and the fact that the discrete topology is metrizable by $d(x,y) = 1$ whenever $x \neq y$.) With $d(cx, c'x') \leq d(cx, cx') + d(cx', c'x') = d(c(x-x'), 0) + d((c-c')x', 0)$ we find that continuity of the scalar action $\mathbb{F} \times V \to V$ is equivalent to the condition that $d(cx, 0) \to 0$ whenever $c \to 0$ (with x fixed) or $d(x,0) \to 0$ (with c fixed). This problem does not arise for topologies coming from a norm, to which we turn soon.

2. Since metric spaces are Hausdorff and first countable, we have many examples of non-metrizable TVS: All non-Hausdorff ones, like the indiscrete ones, cf. Exercise 2.6(i). On the other hand, $\mathbb{R}^S$ is Hausdorff and a TVS by Exercise 2.6(iii), but for uncountable S it doesn't have countable neighborhood bases, thus is not metrizable.

3. The necessary condition for metrizability given in 2. can be proven to be sufficient, see [141, Theorem 1.24], where it suffices to have a countable[7] open neighborhood base at zero. □

2.11 Exercise Let (V, τ) be a topological vector space.

(i) Let d be a translation invariant metric on V such that $\tau = \tau_d$. Show that a sequence $\{x_n\}$ in V is Cauchy w.r.t. d if and only if for every open neighborhood U of 0 there is an $N \in \mathbb{N}$ such that $n, m \geq N$ implies $x_n - x_m \in U$.

[7]'Countable' always means 'at most countable', otherwise we'll say 'countably infinite'.

(ii) If d_1, d_2 are translation-invariant metrics on V such that $\tau_{d_1} = \tau_{d_2}$, show that d_1 is complete if and only if d_2 is complete.

2.12 REMARK 1. The analogue of (ii) for topological spaces is false: There can be equivalent metrics of which only one is complete!

2. ⋆ The equivalent characterization of Cauchy sequences in (i) makes sense in arbitrary TVS V: A net $\{x_\iota\}_{\iota \in I}$ in V is called a Cauchy net if for every open neighborhood U of 0 there is a $\iota_0 \in I$ such that $\iota, \iota' \geq \iota_0$ implies $x_\iota - x_{\iota'} \in U$. Now V is called complete if every Cauchy net converges. □

2.13 DEFINITION *Let V be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. A seminorm on V is a map $V \to [0, \infty)$, $x \mapsto \|x\|$ (thus $\|x\| = \infty$ is not allowed!) such that*

$$\begin{aligned} \|x + y\| &\leq \|x\| + \|y\| \quad \forall x, y \in V. \qquad \text{(Subadditivity)} \\ \|cx\| &= |c| \, \|x\| \quad \forall c \in \mathbb{F}, x \in V. \qquad \text{(absolute homogeneity)} \end{aligned}$$

A norm is a seminorm satisfying also $\|x\| = 0 \;\Rightarrow\; x = 0$.

A normed space over $\mathbb{F}$ is a pair $(V, \|\cdot\|)$, where V is an $\mathbb{F}$-vector space and $\|\cdot\|$ a norm on V.

It is immediate that every seminorm satisfies $\|0\| = 0$ and $\|-x\| = \|x\|$.

2.14 EXERCISE If $\|\cdot\|$ is a (semi)norm on V, prove $\big|\|x\| - \|y\|\big| \leq \|x - y\| \; \forall x, y \in V$, the 'reverse triangle inequality.

2.15 LEMMA *Let $(V, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space, and define $d(x, y) = \|x - y\|$. Then*

(i) d is a translation-invariant metric on V.

(ii) (V, τ_d) is a Hausdorff topological vector space.

(iii) The map $V \to [0, \infty)$, $x \mapsto \|x\|$ is τ_d-continuous.

Proof. (i) That norms give rise to metrics is probably known: $d(x, y) \geq 0$ follows from $\|x\| \geq 0$, and $d(x, y) = 0 \Leftrightarrow x = y$ follows from the norm axiom $\|x\| = 0 \Leftrightarrow x = 0$. Furthermore,

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

where we used $\|-x\| = \|x\|$, a special case of the second seminorm axiom. Finally,

$$d(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z).$$

Now translation invariance of d is obvious: $d(x, y) := \|x - y\| = d(x - y, 0)$.

(ii) All metric spaces are Hausdorff. We have seen earlier that the topology coming from a translation invariant metric makes addition jointly continuous. And if $x, x' \in V$ and $c, c' \in \mathbb{F}$ then

$$\begin{aligned} d(cx, c'x') &= \|cx - c'x'\| = \|cx - cx' + cx' - c'x'\| \leq \|c(x - x')\| + \|(c - c')x'\| \\ &= |c| \, \|x - x'\| + |c - c'| \, \|x'\| = |c| d(x, x') + |c - c'| \|x'\|. \end{aligned}$$

This implies joint continuity of the scalar action $\mathbb{F} \times V \to V$.

(iii) This is immediate by the inequality proven in Exercise 2.14. ■

2.16 Definition • *A norm $\|\cdot\|$ on a vector space V is called complete if the metric $d(x,y) = \|x-y\|$ is complete.*

- *A topological vector space (V,τ) is called (completely) normable if there exists a (complete) norm $\|\cdot\|$ on V such that $\tau = \tau_d$ with $d(x,y) = \|x-y\|$.*
- *A complete normed space is called Banach space.*
- *A normed space $(V, \|\cdot\|)$ is called separable if the associated norm topology τ is separable. (I.e. V has a countable τ-dense subset.)*

2.17 Remark 1. Obviously every normable TVS is metrizable. If d is a translation-invariant metric on V and we put $\|x\| = d(x,0)$ then clearly $\|x\| = 0 \Rightarrow x = 0$ and the computation $\|x+y\| = d(x+y,0) = d(x,-y) \le d(x,0) + d(0,-y) = d(x,0) + d(y,0) = \|x\| + \|y\|$ proves subadditivity. But there is no reason why $\|cx\| = |c|\|x\|$ should hold. We will later see examples of metrizable TVS that are not normable, see Section 4.2. (Such spaces, while somewhat better behaved than general TVS, can still be rather pathological. The subclass of Fréchet spaces, see below, is much better.)

2. We will soon prove that every finite-dimensional Hausdorff TVS is normable.

3. In every normed space $(V, \|\cdot\|)$ the balls $B(0,r) = \{x \in V \mid \|x\| < r\}$ are bounded, open and convex (see Definition 5.22). The first two properties are obvious, and convexity follows from $\|tx + (1-t)y\| \le t\|x\| + (1-t)\|y\| < t + (1-t) = 1$ for all $x, y \in B(0,r)$ and $t \in [0,1]$. Conversely, one can show that every TVS in which zero has a bounded convex open neighborhood is normable! (Note that in a general TVS one needs a new definition of boundedness of sets since neither norm nor metric are available a priori. See Appendix B.6.2 for definition and proof.)

4. Just as complete metric spaces are 'better behaved' (in the sense of allowing stronger theorems) than general metric spaces, Banach spaces are 'better' than normed spaces. We'll meet some applications of completeness in Section 3.2, and more will follow.

5. Separability is a somewhat annoying restriction that we will avoid as much as possible. (An opposite philosophy, cf. e.g. [116], restricts to separable spaces from the beginning in order to make do with weak versions of the axiom of choice.) □

2.18 Example 0. Clearly $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is a vector space over itself and $\|c\| := |c|$ defines a norm, making $\mathbb{F}$ a complete normed $\mathbb{F}$-vector space.

1. Let X be a compact topological space and $V = C(X, \mathbb{F})$. Clearly, V is an $\mathbb{F}$-vector space. Now $\|f\| = \sup_{x \in X} |f(x)|$ is a norm on V. You probably know that the normed space $(V, \|\cdot\|)$ is complete. (See Lemma A.30 for a proof.) One can prove that it is separable if and only if X is second countable, see Proposition A.48.

If X is non-compact then $\|f\|$ can be infinite, but replacing $C(X, \mathbb{F})$ by

$$C_b(X, \mathbb{F}) = \{f \in C(X, \mathbb{F}) \mid \|f\| < \infty\},$$

$\|\cdot\|$ again is a norm with which $C_b(X, \mathbb{F})$ is complete.

2. Let $n \in \mathbb{N}$ and $V = \mathbb{C}^n$. For $x \in V$ and $1 \le p < \infty$ (NB: p does not stand for prime!), define

$$\begin{aligned} \|x\|_\infty &= \max_{i=1,\dots,n} |x_i|, \\ \|x\|_p &= \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}. \end{aligned}$$

(Note that all these $\|\cdot\|_p$ including $p=\infty$ coincide if $n=1$.) It is quite obvious that for each $p\in[1,\infty]$ we have $\|x\|_p=0 \Leftrightarrow x=0$ and $\|cx\|_p=|c|\,\|x\|_p$. For $p=1$ and $p=\infty$ also the subadditivity is trivial to check using only $|c+d|\le|c|+|d|$. Subadditivity also holds for $1<p<\infty$, but is harder to prove. You have probably seen the proof for $p=2$, which relies on the Cauchy-Schwarz inequality. The proof for $1<p<2$ and $2<p<\infty$ is similar, using the inequality of Hölder instead.

2.19 Exercise Prove that the norms $\|\cdot\|_p$ on $\mathbb{F}^n$ are complete for all $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$, $n\in\mathbb{N}$, $p\in[1,\infty]$.

3. The above examples are easily generalized to infinite dimensions. Let S be any set. For a function $f:S\to\mathbb{F}$ and $1\le p<\infty$ define

$$\|f\|_\infty=\sup_{s\in S}|f(s)|, \qquad \|f\|_S=\Big(\sum_{s\in S}|f(s)|^p\Big)^{1/p}$$

with the understanding that $(+\infty)^{1/p}=+\infty$. For the definition of infinite sums like $\sum_{s\in S}f(s)$ see Appendix A.1. Now let

$$\ell^p(S,\mathbb{F})=\{f:S\to\mathbb{F}\mid \|f\|_p<\infty\}.$$

Now one can prove that $\|\cdot\|_p$ is a complete norm on $(\ell^p(S,F),\|\cdot\|_p)$ for each $p\in[1,\infty]$. We will do this in Section 4.

4. Let $(X,\mathcal{A},\mu)$ be a measure space, $f:X\to\mathbb{F}$ measurable and $1\le p<\infty$. We define

$$\begin{aligned}\|f\|_p &= \left(\int_X |f|^p d\mu\right)^{1/p},\\ \|f\|_\infty &= \inf\{M>0\mid \mu(\{x\in X\mid |f(x)|>M\})=0\}\end{aligned}$$

and

$$\mathcal{L}^p(X,\mathcal{A},\mu;\mathbb{F})=\{f:X\to\mathbb{F}\text{ measurable}\mid \|f\|_p<\infty\}.$$

Then $\|f\|_p=(\int_X|f|^pd\mu)^{1/p}$ is a seminorm on $\mathcal{L}^p(X,\mathcal{A},\mu;\mathbb{F})$ for all $1\le p<\infty$.

However, in general $\|\cdot\|_p$ it is not a norm since $\|f\|_p$ vanishes whenever f is zero almost everywhere, i.e. $\mu(f^{-1}(\mathbb{C}\backslash\{0\}))=0$, which may well happen even if f is not identically zero. In order to obtain a normed space one puts $V_0=\{f\in\mathcal{L}^p\mid \|f\|_p=0\}$, which is a linear subspace, and considers the quotient space $L^p(X,\mathcal{A},\mu)=\mathcal{L}^p(X,\mathcal{A},\mu)/V_0$, to which the norms $\|\cdot\|_p$ descend. Going into the details would require too much measure theory. See [101, Appendices A.1-A.3] for a crash course or [29, 146] for the full story.

There is an instructive special case: If S is a set, $\mathcal{A}=P(S)$ and $\mu(A)=\#A$ (the counting measure) then for every $f:S\to\mathbb{F}$ we have $\int_S f(s)\,d\mu(s)=\sum_{s\in S}f(s)$, where the integral, like the (unordered) sum, exists if and only if $\sum_{s\in S}|f(s)|<\infty$. Thus $\mathcal{L}^p(S,\mathcal{A},\mu;\mathbb{F})=\ell^p(S,\mathbb{F})$[8].

Note that the norm of a normable space (V,τ) never is unique (unless $V=\{0\}$): If $\|\cdot\|$ is a norm compatible with τ then the same holds for $c\|\cdot\|$ for every $c>0$. Thus the choice of a norm on a vector space is an extra piece of structure. If $\|\cdot\|_1,\|\cdot\|_2$ are different norms on V then $(V,\|\cdot\|_1),(V,\|\cdot\|_2)$ are different as normed spaces even if the norms give rise to the same topology!

[8]In view of these facts, which we cannot prove without going deeper into measure and integration theory, it certainly isn't unreasonable to ask that you understand the much simpler unordered summation.

2.20 Definition *Let V be an $\mathbb{F}$-vector space. Two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are called equivalent if $\tau_{d_1} = \tau_{d_2}$, where $d_i(x, y) = \|x - y\|_i$.*

This definition is a special case of the notion of equivalence of two metrics d_1, d_2 on a set, also defined by $\tau_{d_1} = \tau_{d_2}$. In that general situation one can prove criteria for equivalence, cf. e.g. [108], but for normed spaces one has a much simpler one:

2.21 Proposition *Two norms $\|\cdot\|_1, \|\cdot\|_2$ on an $\mathbb{F}$-vector space V are equivalent if and only if there are $0 < c' \leq c$ such that $c'\|x\|_1 \leq \|x\|_2 \leq c\|x\|_1$ for all $x \in V$.*

Proof. Since the norm topologies τ_i are defined in terms of the translation invariant metrics d_i, for them to coincide it suffices that every d_1-open ball around zero contains a d_2-open ball around zero and vice versa. By the absolute homogeneity of the norms, this is equivalent to the existence of $s, s' > 0$ such that $B^{\|\cdot\|_1}(0, s) \subseteq B^{\|\cdot\|_2}(0, 1)$ and $B^{\|\cdot\|_2}(0, s') \subseteq B^{\|\cdot\|_2}(0, 1)$, which means that

$$\|x\|_1 < s \Rightarrow \|x\|_2 < 1 \quad \text{and} \quad \|x\|_2 < s' \Rightarrow \|x\|_1 < 1. \tag{2.2}$$

This clearly is implied by the statement $c'\|x\|_1 \leq \|x\|_2 \leq c\|x\|_1$ with $c' > 0$. On the other hand, by continuity of the norms (2.2) implies $\|x\|_1 \leq s \Rightarrow \|x\|_2 \leq 1$ and $\|x\|_2 \leq s' \Rightarrow \|x\|_1 \leq 1$ which, using homogeneity again, gives $\|x\|_2 \leq s^{-1}\|x\|_1$ and $\|x\|_1 \leq s'^{-1}\|x\|_2$, i.e. our condition (with $c' = s', c = s^{-1}$). ■

2.22 Example Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}, p \in [1, \infty), n \in \mathbb{N}$. Then for $x \in \mathbb{F}^n$ we have

$$\|x\|_\infty^p = \max_i |x_i|^p \leq \|x\|_p^p = \sum_{i=1}^n |x_i|^p \leq n\|x\|_\infty^p,$$

Thus $\|x\|_\infty \leq \|x\|_p \leq n^{1/p}\|x\|_\infty$, so that $\|\cdot\|_\infty$ is equivalent to $\|\cdot\|_p$ for all $p < \infty$ *if n is finite*. This clearly implies that all $\|\cdot\|_p$, $p \in [1, \infty]$ are mutually equivalent. (You probably know that all norms on $\mathbb{F}^n$ are equivalent, not only those of the form $\|\cdot\|_p$. We will prove the even stronger result that there is only one Hausdorff topology on $\mathbb{F}^n$ making it a TVS.)

Later (Section 7.1) we will also prove the following deeper and quite surprising result:

2.23 Theorem (Two norm theorem) *If V is a vector space that is complete w.r.t. each of the norms $\|\cdot\|_1, \|\cdot\|_2$ and $\|\cdot\|_2 \leq c\|\cdot\|_1$ for some $c > 0$ then also $\|\cdot\|_1 \leq c'\|\cdot\|_2$ for some $c' > 0$, thus the two norms are equivalent.*

2.24 Example Let S be an infinite set, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $V = \ell^1(S, \mathbb{F})$. Then $\|f\|_\infty \leq \|f\|_1$ for all $f \in V$, but the norms are not equivalent, for example since $(V, \|\cdot\|_1)$ is complete while $(V, \|\cdot\|_\infty)$ is not. This also is the reason why there is no contradiction with the above theorem.

2.25 Exercise Prove: If V is a vector space and $\|\cdot\|_1, \|\cdot\|_2$ are equivalent norms on V then completeness of $(V, \|\cdot\|_1)$ is equivalent to completeness of $(V, \|\cdot\|_2)$.

2.26 Exercise Let $(V, \|\cdot\|)$ be a normed space. Put $d(x, y) = \|x - y\|$ and let $(\widehat{V}, \widehat{d})$ be the completion of the metric space (V, d). Prove that $\widehat{V}$ is a Banach space (in particular a vector space!) and give its norm.

2.27 Exercise Let $(V_1, \|\cdot\|_1), (V_2, \|\cdot\|_2)$ be normed spaces.

(i) Prove that $\|(x_1,x_2)\|_s = \|x_1\|_1 + \|x_2\|_2$ and $\|(x_1,x_2)\|_m = \max(\|x_1\|_1, \|x_2\|_2)$ are equivalent norms on $V_1 \oplus V_2$.

(ii) Prove that $(V_1 \oplus V_2, \|\cdot\|_{s/m})$ is complete if and only if $(V_1, \|\cdot\|_1), (V_2, \|\cdot\|_2)$ both are complete.

2.28 Exercise (i) Let $\{(V_i, \|\cdot\|_i)\}_{i\in I}$ be a family of normed spaces, where I is any set. Put

$$\bigoplus_{i\in I} V_i = \{\{x_i\}_{i\in I} \mid \sum_{i\in I} \|f(i)\|_i < \infty\}.$$

(Technically, this is a subset of $\prod_i V_i = \{f : I \mapsto \bigcup_j V_j \mid f(i) \in V_i \ \forall i \in I\}$.) Prove that this is a linear space and $\|f\| = \sum_i \|f(i)\|_i$ a norm on it.

(ii) Prove that $(\bigoplus_{i\in I} V_i, \|\cdot\|)$ is complete if all the V_i are complete. Hint: The proof is an adaptation of the one for $\ell^1(S)$ given in Section 4.3.

If $V_i = \mathbb{F}$ for all $i \in I$ with the usual norm, we have $\bigoplus_{i\in I} V_i \cong \ell^1(I, \mathbb{F})$.

2.3 A glimpse beyond normed spaces

Normed spaces are an extremely versatile notion that has many applications. Below we will prove that all finite-dimensional Hausdorff TVS are normable. But non-normable spaces exist, and we have already met one generalization of normable spaces, the metrizable TVS. We will quickly look at a generalization into a different direction, the locally convex TVS. The spaces that are both locally convex and metrizable, called Fréchet spaces, are almost as 'good' as Banach spaces.

2.3.1 Finite-dimensional TVS

2.29 Proposition *Let V be a finite-dimensional vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then there is a unique Hausdorff topology τ on V making (V, τ) a topological $\mathbb{F}$-vector space.*

Proof. We give the proof for $\mathbb{F} = \mathbb{R}$ and leave it to the reader to adapt it to $\mathbb{F} = \mathbb{C}$.

Let $\tau_\mathbb{F}$ be the standard topology on $\mathbb{F}$ coming from $d(c, c') = |c - c'|$. Let V be a finite-dimensional $\mathbb{F}$-vector space, put $n = \dim_\mathbb{F} V$ and choose a basis $E = \{e_1, \ldots, e_n\}$ for V. Now the product topology $\tau_\mathbb{F}^n = \tau_\mathbb{F} \times \cdots \times \tau_\mathbb{F}$ on $\mathbb{F}^n$ is Hausdorff, and $(\mathbb{F}^n, \tau_\mathbb{F}^n)$ is a TVS by Exercise 2.6(iii) or by Lemma 2.15 and the fact that any of the norms $\|\cdot\|_p$, $p \in [1, \infty]$ on $\mathbb{F}^n$ induce the product topology $\tau_\mathbb{F}^n$. Since the map $\alpha_E : \mathbb{F}^n \to V,\ (x_1, \ldots, x_n) \mapsto \sum_{i=1}^n x_i e_i$ is a bijection, $\tau = \{\alpha_E(U) \mid U \in \tau_\mathbb{F}^n\}$ is a Hausdorff topology on V and one easily checks that (V, τ) a TVS. This proves existence of τ. For uniqueness, let τ be an arbitrary Hausdorff topology on V making (V, τ) a TVS. Then the maps $\cdot : \mathbb{F} \times V \to V$ and $+ : V \times V \to V$ are continuous w.r.t. τ. This implies that $\alpha_E : (\mathbb{F}^n, \tau_\mathbb{F}^n) \to (V, \tau)$ is continuous, thus $\tau' := \{\alpha_E^{-1}(U) \mid U \in \tau\} \subseteq \tau_\mathbb{F}^n$. Since α_E is a bijection, τ' is a Hausdorff topology on $\mathbb{F}^n$ such that $\cdot : \mathbb{F} \times (\mathbb{F}^n, \tau') \to (\mathbb{F}^n, \tau')$ is continuous. Thus if we prove that every such $\tau' \subseteq \tau_\mathbb{F}^n$ coincides with $\tau_\mathbb{F}^n$, the uniqueness of τ follows.

Let $S = \{x \in \mathbb{F}^n \mid \|x\|_2 = 1\}$ be the euclidean unit sphere, which is $\tau_\mathbb{F}^n$-compact. Now $\tau' \subseteq \tau_\mathbb{F}^n$ implies that S is also τ'-compact[9], and therefore τ'-closed (since τ' is Hausdorff). Thus $\mathbb{F}^n \backslash S \in \tau'$. Since the scalar action $\mathbb{F} \times \mathbb{F}^n \to \mathbb{F}^n$ is continuous (w.r.t. the metric topology $\tau_\mathbb{F}$ on $\mathbb{F}$ and τ' on $\mathbb{F}^n$), the pre-image $V = \cdot^{-1}(\mathbb{F}^n \backslash S) \subseteq \mathbb{F} \times \mathbb{F}^n$, which contains 0×0, is $\tau_\mathbb{F} \times \tau'$-open. By

[9] An open cover by elements of τ' is an open cover by elements of $\tau_\mathbb{F}^n$, thus has a finite subcover.

definition of the product topology, there are $\varepsilon > 0$ and $0_{\mathbb{F}^n} \in U' \in \tau'$ such that $B(\varepsilon, 0_{\mathbb{F}}) \times U' \subseteq V$. In other words, $c \in \mathbb{F}, |c| < \varepsilon$ and $x \in U'$ imply $cx \notin S$, which is equivalent to $\|cx\|_2 \neq 1$ and to $\|x\|_2 \neq 1/|c|$. Since $1/|c|$ may assume any value larger than $1/\varepsilon$, we find that $x \in U'$ implies $\|x\|_2 \leq 1/\varepsilon$. Replacing x by $d^{-1}x$, where $d > 0$, we find $x \in dU' \Rightarrow \|x\|_2 \leq d/\varepsilon$, so that the map $\mathrm{id} : (\mathbb{F}^n, \tau') \to (\mathbb{F}^n, \tau^n_{\mathbb{F}})$ is continuous at zero, thus everywhere by linearity. This means $\tau^n_{\mathbb{F}} \subseteq \tau'$, completing the proof of $\tau' = \tau^n_{\mathbb{F}}$. ■

2.30 Remark 1. The choice of the basis E in the above proof does not matter. Why?

2. The example of the indiscrete topology, which turns every vector space into a TVS, shows that there is no uniqueness if one omits the Hausdorff assumption.

3. The above proof is quite typical for proofs in topological algebra. Luckily, as soon as we have (semi)norms at our disposal, proofs tend to be less point set topological. For example explicit invocations of the product topology are rare. □

2.31 Corollary *Every finite-dimensional Hausdorff TVS (V, τ) over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is normable.*

Proof. Pick a basis $E = \{e_1, \ldots, e_n\}$ for V and define a norm on V by $\|x\| = \|\alpha_E^{-1}(x)\|_1$. (Thus $\|\sum_i c_i e_i\| = \sum_i |c_i|$.) By Lemma 2.15 the topology on V induced by this norm is Hausdorff, thus coincides with τ by Proposition 2.29. Thus τ is normable. ■

2.32 Corollary *On a finite-dimensional vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, all norms are equivalent.*

Proof. Let $\|\cdot\|_1, \|\cdot\|_2$ be norms on V. They give rise to topologies τ_1, τ_2 such that (V, τ_i) is a Hausdorff TVS for $i = 1, 2$. Proposition 2.29 implies $\tau_1 = \tau_2$, so that in view of Definition 2.20 the norms are equivalent. ■

2.33 Exercise Prove that every finite-dimensional normed space $(V, \|\cdot\|)$ over $\mathbb{R}$ or $\mathbb{C}$ is complete.

2.3.2 Locally convex and Fréchet spaces

We have seen that every norm on a vector space gives rise to a translation-invariant metric and a TVS structure. Analogously, if $\|\cdot\|$ is seminorm on V, but not a norm, then $d(x, y) = \|x - y\|$ defines only a pseudo-metric, and τ_d is not Hausdorff (if $x \neq 0$ is such that $\|x\| = 0$ then there are no disjoint open sets U, V containing $x, 0$, respectively).

2.34 Definition *If V is an $\mathbb{F}$-vector space and $\mathcal{F}$ is a family of seminorms on V then the topology $\tau_{\mathcal{F}}$ is the smallest topology on V containing the balls*

$$B_{\|\cdot\|}(x, r) = \{y \in V \mid \|x - y\| < r\}$$

for all $x \in V, r > 0, \|\cdot\| \in \mathcal{F}$.

More explicitly, $\tau_{\mathcal{F}}$ consists of all unions of finite intersections of such balls, i.e. the latter form a subbase for $\tau_{\mathcal{F}}$. Now a sequence or net $\{x_\iota\}$ in V converges to $z \in V$ if and only if $\|x_\iota - z\|$ converges to zero for each $\|\cdot\| \in \mathcal{F}$.

2.35 Definition *We say that $\mathcal{F}$ is separating if for any non-zero $x \in V$ there is a $\|\cdot\| \in \mathcal{F}$ such that $\|x\| \neq 0$.*

The property of being separating is important since one usually is only interested in Hausdorff topological vector spaces and the following holds:

2.36 Lemma *The topology $\tau_{\mathcal{F}}$ induced by a family $\mathcal{F}$ of seminorms on V is Hausdorff if and only if $\mathcal{F}$ is separating.*

Proof. $\Rightarrow$ Assume $\mathcal{F}$ is not separating. Then there is $0 \neq x \in V$ such that $\|x\| = 0$ for all $\|\cdot\| \in \mathcal{F}$. Then by definition of $\tau_{\mathcal{F}}$, every open set containing 0 also contains x and vice versa, so that $\tau_{\mathcal{F}}$ is not Hausdorff.

$\Leftarrow$ Assume $x \neq y$. By assumption there is a $\|\cdot\| \in \mathcal{F}$ such that $c = \|x-y\| > 0$. Let $U = B_{\|\cdot\|}(x, c/2), V = B_{\|\cdot\|}(y, c/2)$. Then U, V are open sets containing x, y, respectively, and existence of $z \in U \cap V$ would imply the contradiction $d(x,y) \leq d(x,z) + d(z,y) < c/2 + c/2 = r = d(x,y)$. Thus $U \cap V = \emptyset$, so that $\tau_{\mathcal{F}}$ is Hausdorff. ■

If V is an $\mathbb{F}$-vector space and $\mathcal{F}$ is a family of seminorms on V, one can prove that V is a topological vector space when equipped with the topology $\tau_{\mathcal{F}}$. The proof is not much more complicated than for the case of one (semi)norm considered above.

2.37 Definition *A topological vector space (V, τ) over $\mathbb{F}$ is called locally convex if there exists a separating family $\mathcal{F}$ of seminorms on V such that $\tau = \tau_{\mathcal{F}}$.*

2.38 Remark 1. Note that with this definition, locally convex spaces are Hausdorff.

2. For an equivalent, more geometric way of defining local convexity of a TVS see the supplementary Section B.6.2, and for more on locally convex spaces see, e.g., [94, 30, 141].

3. Locally convex spaces, introduced in 1935 by von Neumann[10], have many applications. In these notes, we will encounter the weak and weak-$*$ topologies on (duals of) Banach spaces and the strong and weak operator topologies. There are many others, as in distribution theory, relevant for the theory of (partial) differential equations. □

If the separating family $\mathcal{F}$ has just one element, we are back at the notion of a normed, possibly Banach, space. If $\mathcal{F}$ is finite, i.e. $\mathcal{F} = \{\|\cdot\|_1, \ldots, \|\cdot\|_n\}$, then $\|\cdot\| = \sum_{i=1}^n \|\cdot\|_i$ is a seminorm, and it is a norm if and only if $\mathcal{F}$ is separating. Thus the case of finite $\mathcal{F}$ again gives a normed space. Thus $\mathcal{F}$ must be infinite in order for interesting things to happen.

If $\mathcal{F}$ is infinite, we cannot obtain a norm by putting $\|x\|' = \sum_{\|\cdot\| \in \mathcal{F}} \|x\|$, since the r.h.s. has no reason to converge for all $x \in V$. But if the family $\mathcal{F}$ of seminorms on V is countable, we can label the elements of $\mathcal{F}$ as $\|\cdot\|_n$, $n \in \mathbb{N}$ and define

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} \min(1, \|x-y\|_n).$$

Now each term $\min(1, \|x-y\|_n)$ is a translation-invariant pseudometric [defined like a metric, but without the requirement $d(x,y) = 0 \Rightarrow x = y$] on V that is bounded by 1, and the sum converges to a translation-invariant metric on V. With just a bit more work one shows that $\tau_{\mathcal{F}} = \tau_d$, thus $(V, \tau_{\mathcal{F}})$ is metrizable. (Note that we could not have defined $\|x\| = \sum_{i=1}^{\infty} 2^{-i} \|x\|_i$ since this again may fail to converge, thus need not be a norm.) If such a space is complete, it is called a Fréchet space. Clearly, every Fréchet space is an F-space.

[10]John von Neumann (1903-1957). Hungarian, later American, mathematician. Countless contributions mostly to foundational matters and analysis, e.g. the theory of unbounded operators and the spectral theorem, von Neumann algebras, locally convex spaecs, but also to applied mathematics and computer science.

Here is an example of a Fréchet space: For $f \in C^\infty(\mathbb{R}, \mathbb{C})$ and $n, m \in \mathbb{N}_0$, define

$$\|f\|_{n,m} = \sup_{x \in \mathbb{R}} |x|^n |f^{(m)}(x)|,$$

where $f^{(m)}$ is the m-th derivative of f. These $\|\cdot\|_{n,m}$ are seminorms. Since the family $\mathcal{F} = \{\|\cdot\|_{n,m} \mid n, m \in \mathbb{N}_0\}$ is countable, the space

$$\mathcal{S} = \{f \in C^\infty(\mathbb{R}, \mathbb{C}) \mid \|f\|_{n,m} < \infty \ \forall n, m \in \mathbb{N}_0\}$$

equipped with the topology $\tau_\mathcal{F}$ is a Fréchet space. Its elements are called Schwartz[11] functions. They are infinitely differentiable functions that, together with all their derivatives, vanish as $|x| \to \infty$ faster than $|x|^{-n}$ for any n. (This definition is easily generalized to functions of several variables.) Note that the seminorm $\|\cdot\|_{0,0}$ alone already separates the elements of $\mathcal{S}$, thus is a norm, but having the other seminorms around gives rise to a finer topology, one that is not normable.

3 Normed and Banach space basics

3.1 Linear maps: bounded ⇔ continuous

If E, F are vector spaces over $\mathbb{F}$, a map $A : E \to F$ is called linear if $A(x+y) = Ax + Ay$ for all $x, y \in E$ and and $A(cx) = cAx$ for all $x \in E, c \in \mathbb{F}$. (NB: A map of the form $x \mapsto Ax + b$, where $A : E \to F$ is linear and $b \in F \backslash \{0\}$, is *not* called a linear map, but an affine one!) As in linear algebra, we mostly write Ax instead of $A(x)$.

3.1 Definition *A linear map $A : V \to W$ of normed spaces is called an isometry if $\|Ax\| = \|x\|$ for all $x \in V$.*

Recall that a linear map $A : V \to W$ is injective if and only if its kernel $\ker A = A^{-1}(0)$ is $\{0\}$. It follows that an isometry is automatically injective. Furthermore, if $A : V \to W$ is a surjective isometry then it is invertible, and its inverse also is an isometry. Then A is called an isometric isomorphism of normed spaces, and we write $V \cong W$. Normed spaces that are isometrically isomorphic are essentially indistinguishable.

3.2 Definition *Let E, F be normed spaces and $A : E \to F$ a linear map. Then the norm $\|A\| \in [0, \infty]$ is defined by*

$$\|A\| = \sup_{0 \neq e \in E} \frac{\|Ae\|}{\|e\|} = \sup_{\substack{e \in E \\ \|e\| = 1}} \|Ae\|.$$ [12]

(The equality of the second and third expression is due to linearity of A and homogeneity of the norms.) If $\|A\| < \infty$ then A is called bounded.

[11] Laurent Schwartz (1915-2002). French mathematician who invented 'distributions', an important notion in functional analysis.

[12] It should be clear that writing $\sup_{\substack{e \in E \\ \|e\| \leq 1}} \|Ae\|$ instead would not change the result.

3.3 Remark 1. Every isometry has norm one, but not every norm one map is an isometry.

2. The definition implies $\|Ax\| \leq \|A\|\|x\| \ \forall x \in E$, and $\|Ax\| \leq C\|x\| \ \forall x$ implies $\|A\| \leq C$.

3. Linear maps are also called linear operators, but linear maps $A : E \to \mathbb{F}$ are called linear functionals (whence the term 'functional analysis').

4. If $V = C^\infty(\mathbb{R}, \mathbb{R})$ with norm $\|f\| = \sup_x |f(x)|$ and $f_c(x) = \sin(cx)$ then for all $c \in \mathbb{R}$ we have $f_c \in V$, $\|f_c\| = 1$ and $\|f_c'\| = c$. Thus $A : V \to V$, $f \mapsto f'$ is unbounded. (The same holds for essentially all differential operators.) While this operator is defined on all of V, for unbounded operators it often is too restrictive to require them to be defined on the whole space. Cf. also Remark 7.33.

5. In fact, every infinite-dimensional space admits unbounded linear maps, cf. Exercise 3.8 below. □

3.4 Exercise If E, G, H are normed spaces and $S : E \to G$, $T : G \to H$ are linear maps, prove that $\|TS\| \leq \|S\|\|T\|$. (We write TS for the composite map $T \circ S : E \to H$.)

3.5 Lemma *Let E, F be normed spaces and $A : E \to F$ a linear map. Then the following are equivalent:*

(i) A is bounded.

(ii) A is continuous (w.r.t. the norm topologies).

(iii) A is continuous at $0 \in E$.

Proof. (i)⇒(ii) For $x, y \in E$ we have $\|Ax - Ay\| = \|A(x-y)\| \leq \|A\| \, \|x-y\|$. Thus $d(Ax, Ay) \leq \|A\| \, d(x,y)$, and with $\|A\| < \infty$ we have (uniform) continuity of E.

(ii)⇒(iii) This is obvious.

(iii)⇒(i) $B^F(0,1) \subseteq F$ is an open neighborhood of $0 \in F$. Since A is continuous at 0, the preimage $A^{-1}(B^F(0,1)) \subseteq E$, which clearly contains 0, is open. Thus there exists $\varepsilon > 0$ such that $B^E(0,\varepsilon) \subseteq A^{-1}(B^F(0,1))$. In other words, $\|x\| < \varepsilon$ implies $\|Ax\| < 1$. A fortiori, $\|x\| \leq \varepsilon/2 \Rightarrow \|Ax\| \leq 1$. By linearity of A and absoloute homegeneity of the norms, this is equivalent to $\|x\| \leq 1 \Rightarrow \|Ax\| \leq 2/\varepsilon$, thus $\|A\| \leq 2/\varepsilon < \infty$. (More precisely, $\|A\| = (\sup\{\varepsilon > 0 \mid B^E(0,\varepsilon) \subseteq A^{-1}(B^F(0,1))\})^{-1}$.) ■

The above motivates the following important notion:

3.6 Definition *Let V, W be normed spaces. A linear map $A : V \to W$ is called isomorphism if it satisfies the equivalent statements*

(i) A is a bijection and a homeomorphism, i.e. A and A^{-1} are continuous.

(ii) A is surjective and there are $C_1, C_2 > 0$ such that $C_1\|x\| \leq \|Ax\| \leq C_2\|x\| \ \forall x \in V$.

Clearly A is an isometric isomorphism if and only if $C_1 = C_2 = 1$. If an isomorphism $V \to W$ exists, we write $V \simeq W$ (not to be confused with $\cong$ for isometric isomorphism).

In particular, if $\|\cdot\|_1, \|\cdot\|_2$ are norms on V then $\mathrm{id}_V : (V, \|\cdot\|_1) \to (V, \|\cdot\|_2)$ is an isomorphism (isometric isomorphism) if and only if the two norms are equivalent (equal).

3.7 Exercise Let V, W be normed spaces, where V is finite-dimensional. Prove that every linear map $V \to W$ is bounded.

3.8 Exercise Let V be an infinite-dimensional Banach space.

(i) Show that there exists an unbounded linear map $\varphi : V \to \mathbb{F}$. (Hint: Use a Hamel basis[13])

(ii) If W is a non-zero normed space, show that there is an unbounded linear map $A : V \to W$.

For linear functionals, i.e. linear maps from an $\mathbb{F}$-vector space to $\mathbb{F}$, there is another characterization of boundedness/continuity:

3.9 Exercise Let $(V, \|\cdot\|)$ be a normed $\mathbb{F}$-vector space and $\varphi : V \to \mathbb{F}$ a linear functional. Prove that φ is continuous if and only if $\ker\varphi = \varphi^{-1}(0) \subseteq V$ is closed.

Hint: For $\Leftarrow$, pick a ball $B(x,r) \subseteq V\backslash\ker\varphi$ and prove that $\varphi(B(0,r))$ is bounded.

If V is a finite-dimensional normed space of dimension d, by linear algebra there exists a basis $E = \{e_1, \ldots, e_d\}$. We can normalize its elements so that $\|e_i\| = 1$ for all i. Once E is fixed, there is a unique family $F = \{\varphi_1, \ldots, \varphi_d\}$ of linear functionals $V \to \mathbb{F}$ such that $\varphi_i(e_j) = \delta_{i,j}\ \forall i, j$. (One easily checks that F is a basis for V^*.) By Exercise 3.7, the φ_i are bounded, and $\|\varphi_i\| \geq |\varphi(e_i)|/\|e_i\| = 1/1 = 1$. Less trivially one has:

3.10 Proposition (Auerbach 1929) [14] *Every finite-dimensional normed space admits a normalized basis E, called Auerbach basis, such that also the dual basis F is normalized, i.e. $\|\varphi_i\| = 1\ \forall i$.*

Proof. Since every finite-dimensional vector space is isomorphic to $\mathbb{R}^n$ for some n, it suffices to prove this for $W = \mathbb{R}^n$ (but with arbitrary norm). Let $X = \{w \in W \mid \|w\| = 1\}$ be the unit sphere, which is compact (since it is bounded and closed and the Heine-Borel theorem[15] applies since the topology on $W = \mathbb{R}^n$ is the Euclidean one by Corollary 2.32). Consider the map $f : X^n \to \mathbb{R}, E = (e_1, \ldots, e_n) \mapsto \det(e_1|e_2|\cdots|e_n)$, the matrix on the right having the e_i as columns. This is a continuous function, so $\mu = \sup_{E\in X^n} f(E)$ is finite, positive (since f changes sign under exchange of two columns) and is assumed by some $E = (e_1, \ldots, e_n)$. Defining $\varphi_i : x \mapsto \det(e_1, \ldots, e_{i-1}, x, e_{i+1}, \ldots, e_n)/\mu$, we have $\varphi_i(e_j) = \delta_{ij}$ (since the determinant vanishes if two columns are equal). The definition of μ implies $\det(e_1, \ldots, e_{i-1}, x, e_{i+1}, \ldots, e_n) \leq \mu$ for all $x \in X$, thus $\|\varphi_i\| = \sup_{x\in X} |\varphi_i(x)| \leq 1$, and we are done. ■

3.11 Exercise (Banach-Mazur distance) [16] Let V, W be Banach spaces. Define $D(V, W) \in [0, \infty]$ by $D(V, W) = +\infty$ if V and W are not isomorphic, i.e. $V \not\simeq W$, and otherwise by

$$D(V, W) = \inf\{\|A\|\,\|A^{-1}\| \mid A : V \to W \text{ bounded linear isomorphism}\}.$$

Prove:

(i) $D(V, W) = D(W, V) \geq 1$.

(ii) If $V \cong W$ (isometric isomorphism) then $D(V, W) = 1$. In particular $D(V, V) = 1$.

(iii) If $V \simeq W \simeq Z$ then $D(V, Z) \leq D(V, W)D(W, Z)$.

[13] Recall from linear algebra that a Hamel basis for V is a subset $E \subset V$ such that every $x \in V$ is a linear combination of finitely many elements of E, in a unique way.

[14] Herman Auerbach (1901-1942). Polish mathematician. Born in Tarnopol (then Austria-Hungary, now Ukraine). Murdered in the Belzec extermination camp.

[15] A subset of $\mathbb{R}^n$, equipped with its standard topology, is compact if and only if it is closed and bounded. While the name attributes the theorem to Eduard Heine (1821-1881) and Emile Borel (1871-1956), the real history is very complicated. See [3].

[16] Stanisław Mazur (1905-1981). Polish mathematician. Also known for the Gelfand-M. theorem and others.

(iv) Restricted to a set of mutually isomorphic Banach spaces, $d(V,W) = \log D(V,W)$ is a pseudometric.

If V and W are finite-dimensional normed spaces, one can prove $D(V,W) = 1$, thus they are isometrically isomorphic, but in infinite dimensions this is not true!

3.2 Why we care about completeness

As you (should) know from topology, completeness of a metric space is convenient since it leads to results that are not necessarily true without it, like Cantor's intersection theorem and the contraction principle (or Banach's fixed point theorem). The same holds for normed spaces. In this section we present three important applications of completeness, each of which will be used repeatedly. Later we will encounter others.

3.2.1 Extension of bounded linear maps

The following application of completeness will be used several times:

3.12 LEMMA *Let V be a normed space, $W \subseteq V$ a dense linear subspace, Z a Banach space and $A : W \to Z$ a bounded linear map. Then there is a unique bounded linear map $\widehat{A} : V \to Z$ with $A = \widehat{A}_{|W}$*[17]*. It satisfies $\|\widehat{A}\| = \|A\|$. If A is an isometry, so is $\widehat{A}$.*

Proof. Let $x \in V$. Then there is a sequence $\{w_n\}$ in W such that $\|w_n - x\| \to 0$. Then $\{w_n\} \subseteq W$ is a Cauchy sequence, and so is $\{Aw_n\} \subseteq Z$ by boundedness of A. The latter converges since Z is complete. If $\{w'_n\}$ is another sequence converging to x then $\|A(w_n - w'_n)\| \to 0$, so that $\lim Aw'_n = \lim Aw_n$. It thus is well-defined to put $\widehat{A}x = \lim_{n\to\infty} Aw_n$. We omit the easy proof of linearity of $\widehat{A}$. If $x \in W$ then we can put $w_n = x \ \forall n$, obtaining $\widehat{A}x = Ax$, thus $\widehat{A}_{|W} = A$. By density of $W \subseteq V$, any other continuous extension of A coincides with $\widehat{A}$. We have $\|\widehat{A}x\| = \lim \|Aw_n\| \le \|A\|\|x\|$. Thus $\|\widehat{A}\| \le \|A\|$, and the converse inequality is obvious. If A is an isometry then $\|\widehat{A}x\| = \lim_n \|Aw_n\| = \lim_n \|w_n\| = \|x\|$, so that $\widehat{A}$ is an isometry. ■

3.13 EXERCISE Let X, Y be Banach spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let $\{x_i\}_{i\in I} \subseteq X$, $\{y_i\}_{i\in I} \subseteq Y$ be families of vectors such that $\mathrm{span}_{\mathbb{F}}\{x_i \mid i \in I\} \subseteq X$ is dense. Show that there is an $A \in B(X,Y)$ satisfying $Ax_i = y_i \ \forall i \in I$ if and only if there exists a $C \in [0,\infty)$ such that

$$\Big\| \sum_{i\in J} c_i y_i \Big\| \le C \Big\| \sum_{i\in J} c_i x_i \Big\|$$

for all finite subsets $J \subseteq I$ and numbers $\{c_i\}_{i\in J}$ in $\mathbb{F}$. Show that this A is uniquely determined.

3.2.2 Convergence of series

3.14 DEFINITION *Let $(V, \|\cdot\|)$ be a normed space and $\{x_n\}_{n\in\mathbb{N}} \subset V$ a sequence. The series $\sum_{n=1}^\infty x_n$ is called*

- *convergent if the sequence $\{S_n\}$ of partial sums $S_n = \sum_{k=1}^n x_k$ converges to some $s \in V$.*
- *unconditionally convergent if $\sum_{k=1}^\infty x_{\sigma(k)}$ converges for each permutation σ of $\mathbb{N}$.*

[17] If $f : X \to Y$ is a function and $Z \subseteq X$, then according to typographical convenience we write either $f_{|Z}$ or $f\restriction Z$ for the restriction of f to Z, which is a map $Z \to Y$. But it if $f : X \to X$ maps $Y \subseteq Y$ into itself, $f\restriction Y = f_{|Y}$ usually is meant as a map $Y \to Y$, not $Y \to X$.

- *conditionally convergent if it converges, but not unconditionally.*
- *absolutely convergent if* $\sum_{n=1}^\infty \|x_n\| < \infty$.

3.15 Proposition *Let* $(V, \|\cdot\|)$ *be a normed space over* $\mathbb{F}$. *Then*

(i) *If* $\sum_{n=1}^\infty x_n$ *converges then* $\|\sum_{n=1}^\infty x_n\| \le \sum_{n=1}^\infty \|x_n\|$.

(ii) *Unconditional convergence implies convergence, but the converse is false (unless* $V = \{0\}$*).*

(iii) *If* $(V, \|\cdot\|)$ *is complete then every absolutely convergent series* $\sum_{n=1}^\infty x_n$ *in* V *converges unconditionally, all sums* $\sum_{k=1}^\infty x_{\sigma(k)}$ *being equal.*

(iv) *If every absolutely convergent series in* V *converges then* $(V, \|\cdot\|)$ *is complete.*

Proof. (i) The subadditivity of the norm gives $\|\sum_{k=1}^n x_k\| \le \sum_{k=1}^n \|x_k\|$ for all $n \in \mathbb{N}$. As $n \to \infty$, the l.h.s. converges to $\|\sum_n x_n\|$ and the r.h.s. to $\sum_n \|x_n\|$ (which may be infinite, in which case the inequality trivially holds).

(ii) The first part is obvious. If $V \ne \{0\}$, pick $x \in V\backslash\{0\}$ and put $x_n = \frac{(-1)^n}{n}x$. Then $\sum_{n=1}^\infty x_n$ converges (to $x\sum_n \frac{(-1)^n}{n} = -x\ln 2$), but $\sum_n \|x_n\| = \sum_n 1/n = \infty$.

(iii) Assume V to be complete and $\sum_n x_n$ to be absolutely convergent. Let $S_n = \sum_{k=1}^n x_k$ and $T_n = \sum_{k=1}^n \|x_k\|$. For all $n > m$ we have (by subadditivity of the norm)

$$\|S_n - S_m\| = \left\|\sum_{k=m+1}^n x_k\right\| \le \sum_{k=m+1}^n \|x_k\| = T_n - T_m.$$

Since the sequence $\{T_n\}$ is convergent by assumption, thus Cauchy, the above implies that $\{S_n\}$ is Cauchy, thus convergent by completeness of V. That the convergence is unconditional follows from the fact that $\sum_{n=1}^\infty \|x_{\sigma(n)}\|$ is independent of σ, thus finite. The statement that then also $\sum_{n=1}^\infty x_{\sigma(n)}$ is independent of σ is proven as for series of real or complex numbers in a standard analysis course.[18]

(iv) Assume that every absolutely convergent series in V converges, and let $\{y_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in V. We can find (why?) a subsequence $\{z_k\}_{k\in\mathbb{N}} = \{y_{n_k}\}$ such that $\|z_k - z_{k-1}\| \le 2^{-k}\ \forall k \ge 2$. Now put $z_0 = 0$ and define $x_k = z_k - z_{k-1}$ for $k \ge 1$. Now

$$\sum_{k=1}^\infty \|x_k\| = \sum_{k=1}^\infty \|z_k - z_{k-1}\| \le \|z_1\| + \sum_{k=2}^\infty 2^{-k} < \infty.$$

Thus $\sum_{k=1}^\infty x_k$ is absolutely convergent, and therefore convergent by the hypothesis. To wit, $\lim_{n\to\infty} S_n$ exists, where $S_n = \sum_{k=1}^n x_k = \sum_{k=1}^n (z_k - z_{k-1}) = z_n$. Thus $z = \lim_{k\to\infty} z_k = \lim_{k\to\infty} y_{n_k}$ exists. Now the sequence $\{y_n\}$ is Cauchy and has a convergent subsequence $\{y_{n_k}\}$. This implies (why?) that the whole sequence $\{y_n\}$ converges to the limit of the subsequence. ■

The following is well-known (and we will re-prove it later):

3.16 Proposition *Every unconditionally convergent series in* $\mathbb{R}$ *or* $\mathbb{C}$ *is absolutely convergent.*[19]

[18] For completeness, here is a proof: The assumption $\sum_{n=1}^\infty \|x_n\| < \infty$ implies that for each $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $\sum_{n=N+1}^\infty \|x_n\| < \varepsilon$. Picking $M = \max(N, \sigma^{-1}(1), \ldots, \sigma^{-1}(N))$, the sums $\sum_{n=1}^M x_n$ and $\sum_{n=1}^M x_{\sigma(n)}$ both comprise $x_1, \ldots, x_N$ so that their difference is a finite linear combination of $x_{N+1}, x_{N+2}, \ldots$ with coefficients in $\{0, 1, -1\}$. Thus $\|\sum_{n=1}^M x_n - \sum_{n=1}^M x_{\sigma(n)}\| \le \sum_{n=N+1}^\infty \|x_n\| < \varepsilon$. Taking the limit $M \to \infty$ gives $\|\sum_{n=1}^\infty x_n - \sum_{n=1}^\infty x_{\sigma(n)}\| \le \varepsilon$, and since $\varepsilon > 0$ was arbitrary, we have $\sum_{n=1}^\infty x_n = \sum_{n=1}^\infty x_{\sigma(n)}$.

[19] This was proven (for real series) in 1854 by Bernhard Riemann (1826-1866), German mathematician.

3.17 Exercise Let $(V, \|\cdot\|)$ be a finite-dimensional normed space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Prove that every unconditionally convergent series $\sum_{n=1}^{\infty} x_n$ in V is absolutely convergent.

3.18 Remark 1. The result of the preceding exercise fails in infinite dimensions! In Exercise 4.16 we will encounter series in infinite-dimensional Banach spaces that are unconditionally but not absolutely convergent. In fact, by the remarkable Dvoretzky-Rogers theorem (1950) *every* infinite-dimensional Banach space contains unconditionally convergent series that are not absolutely convergent! (See Section B.2.1 for a proof.)

2. We will later prove in two different ways (Corollary 9.11 and Theorem A.4) that the independence of $\sum_{n=1}^{\infty} x_{\sigma(n)}$ of σ holds for all unconditionally convergent series, not only the absolutely convergent ones.

3. There is no characterization of unconditionally convergent series in terms of $\{\|x_n\|\}$, but see the results of Appendices A.2 and B.2. For example, we prove that $\sum_n x_n$ converges unconditionally $\Leftrightarrow$ $\sum_n c_n x_n$ converges for all bounded sequences $\{c_n\} \subseteq \mathbb{F}$ $\Leftrightarrow$ $\sum_i x_{n_i}$ converges for all $n_1 < n_2 < \cdots$. The moral is that unconditional convergence does not rely on 'cancellations'.

4. Proposition 3.15(iv) shows that it is somewhat careless to write "$\sum_n \|x_n\| < \infty$, thus $\sum_n x_n$ converges" since the completeness condition is indispensable. □

3.19 Exercise Let $(V, \|\cdot\|)$ be a Banach space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $E \subset V$ a Hamel basis. By linear algebra there are unique linear functionals $\{\varphi_e : V \to \mathbb{F}\}_{e \in E}$ such that for each $x \in X$ we have $x = \sum_{e \in E} \varphi_e(x) e$ for each $x \in X$, where $\{e \in E \mid \varphi_e(x) \neq 0\}$ is finite. Prove:

(i) If V is finite-dimensional then all φ_e are continuous.

(ii) If V is infinite-dimensional then $\{e \in E \mid \varphi_e \text{ is continuous}\}$ is finite. Hint: Argue by contradiction, using Proposition 3.15 to construct an $x \in V$ for which $\{e \in E \mid \varphi_e(x) \neq 0\}$ is infinite.

Part (ii) shows that Hamel bases are not very well suited for infinite-dimensional Banach spaces (other than for constructing counterexamples). This will be reinforced by Exercise 7.21, where we will see that every Hamel basis of a separable Banach space has cardinality $\mathfrak{c} = \#\mathbb{R}$ rather than $\aleph_0 = \#\mathbb{N}$.

3.2.3 Closedness vs. completeness of subspaces

Closedness and completeness of subsets of a metric space are related. We recall from topology (if you haven't seen this, prove it!):

3.20 Lemma *Let (X, d) be a metric space and $Y \subseteq X$. Then (instead of $d_{|Y}$ we just write d)*

(i) If (X, d) is complete and $Y \subseteq X$ is closed (w.r.t. τ_d, of course) then (Y, d) is complete.

(ii) If (Y, d) is complete then $Y \subseteq X$ is closed (whether or not (X, d) is complete).

The above should be compared with the fact that a closed subset of a compact space is compact and that a compact subset of a Hausdorff space is closed. In the above, completeness works as a weak substitute of compactness, an interpretation that is reinforced by the fact that every compact metric space is complete.

The above lemma readily specializes to normed spaces:

3.21 Lemma *Let $(V, \|\cdot\|)$ be a normed space and $W \subseteq V$ a <u>linear</u> subspace. Then*

(i) If V is complete (=Banach) and a $W \subseteq V$ is closed then W is Banach.

(ii) If W is complete then $W \subseteq V$ is closed (whether or not V is complete).

Note: We will often omit 'linear' from 'linear subspace'. When arbitrary, possibly non-linear, subsets are intended we will make this clear.

3.22 Exercise Prove that every finite-dimensional subspace of a normed space is closed.

The result of the preceding exercise is not at all true for infinite-dimensional subspaces! For example, let $V = \ell^1(\mathbb{N}) = \{f : \mathbb{N} \to \mathbb{R} \mid \sum_{n=1}^\infty |f(n)| < \infty\}$. Now the infinite-dimensional linear subspace $W = \{f : \mathbb{N} \to \mathbb{R} \mid \#\{n \in \mathbb{N} \mid f(n) \neq 0\} < \infty\} \subseteq V$ is non-closed as follows from the easy facts $W \neq V$ and $\overline{W} = V$.

The following (to be generalized in Lemma 7.39) gives closedness of the image of an isometry:

3.23 Corollary *Let V be a Banach space and W a normed space. If $A : V \to W$ is a linear isometry then the linear subspace $AV \subseteq W$ is closed.*

Proof. The map $A : V \to AV \subseteq W$ is an isometric bijection and therefore an isometric isomorphism of normed spaces. Thus $(AV, \|\cdot\|)$ is complete, thus closed in W by Lemma 3.21. ■

3.3 Spaces of bounded linear maps. First glimpse of Banach algebras

3.24 Definition *Let E, F be normed $\mathbb{F}$-vector spaces. The set of bounded linear maps from E to F is denoted $B(E, F)$. Instead of $B(E, E)$ and $B(E, \mathbb{F})$ one also writes $B(E)$ and E^*, respectively. E^* is called the dual space of E.*[20]

Clearly $B(E, F)$ should not be confused with notation $B(x, r)$ for open balls!

3.25 Proposition *Let E, F be normed spaces. Then*

(i) $B(E, F)$ is a vector space and $B(E, F) \to [0, \infty), A \mapsto \|A\|$ is a norm in the sense of Definition 2.13.

(ii) If F is complete (=Banach) then so is $B(E, F)$. In particular, E^* is always Banach.

Proof. (i) If $T : E \to F$ is a linear map, it is clear that $\|\alpha T\| = |\alpha|\|T\|$ and that $\|T\| = 0$ if and only if $T = 0$. If $S, T \in B(E, F)$ and $x \in E$ then $\|(S+T)x\| \leq \|Sx\| + \|Tx\| \leq (\|S\| + \|T\|)\|x\|$, so that $\|S + T\| \leq \|S\| + \|T\|$. This implies that $B(E, F)$ is a vector space.

(ii) Assume F is complete, and let $\{T_n\} \subseteq B(E, F)$ be a Cauchy sequence. Then there is n_0 such that $m, n \geq n_0 \Rightarrow \|T_m - T_n\| < 1$, in particular $T_m \in B(T_{n_0}, 1)$ for all $n \geq n_0$. Thus with $M = \max(\|T_1\|, \ldots, \|T_{n_0-1}\|, \|T_{n_0}\| + 1) < \infty$ we have $\|T_n\| \leq M$ for all n. If now $x \in E$ then $\|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\|$, so that $\{T_n x\}$ is a Cauchy sequence in F and therefore convergent by completeness of F. Now define $T : E \to F$ by $Tx = \lim_{n\to\infty} T_n x$. It is straightforward to check that T is linear. Finally, since $\|T_n x\| \leq M\|x\|$ for all n, we have $\|Tx\| = \|\lim_{n\to\infty} T_n x\| = \lim_{n\to\infty} \|T_n x\| \leq M\|x\|$, so that $T \in B(E, F)$. The second statement follows from the completeness of $\mathbb{R}$ and $\mathbb{C}$. ■

3.26 Exercise Let $(V_1, \|\cdot\|_1), (V_2, \|\cdot\|_2)$ be normed spaces. Prove $(V_1 \oplus V_2, \|\cdot\|_s)^* \cong (V_1^* \oplus V_2^*, \|\cdot\|_m)$ and $(V_1 \oplus V_2, \|\cdot\|_m)^* \cong (V_1^* \oplus V_2^*, \|\cdot\|_s)$. (See Exercise 2.27 for the definitions of $\|\cdot\|_s, \|\cdot\|_m$.)

[20] More generally, if V is any topological vector space we write V^* for the space of continuous linear maps $V \to \mathbb{F}$.

If E is a normed $\mathbb{F}$-vector space, the same holds for $B(E) = B(E, E)$, and by Exercise 3.4, we have $\|ST\| \le \|S\|\|T\|$ for all $S, T \in B(E)$. This motivates the following definition:

3.27 Definition *If $\mathbb{F}$ is a field, an $\mathbb{F}$-algebra is an $\mathbb{F}$-vector space $\mathcal{A}$ together with an associative bilinear operation $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$, $(a, b) \mapsto a \cdot b$, the 'multiplication'.*

Examples: (i) $\mathcal{A} = M_{n\times n}(\mathbb{F})$ with matrix product as multiplication,
(ii) $\mathcal{A} = C(X, \mathbb{F})$ with pointwise product of functions, i.e. $(f \cdot g)(x) := f(x)g(x)\ \forall x \in X$.

3.28 Definition *A normed $\mathbb{F}$-algebra is a $\mathbb{F}$-algebra $\mathcal{A}$ equipped with a norm $\|\cdot\|$ such that $\|ab\| \le \|a\|\|b\|$ for all $a, b \in \mathcal{A}$ (submultiplicativity). A Banach algebra is a normed algebra that is complete (as a normed space). An algebra $\mathcal{A}$ is called unital if it has a unit $\mathbf{1} \neq 0$. (In fact, if $\mathcal{A} \neq \{0\}$ then $\mathbf{1} = 0$ would imply the contradiction $\|a\| = \|\mathbf{1}a\| \le \|\mathbf{1}\|\|a\| = 0\ \forall a \in A$.)*

3.29 Remark 1. If $\mathcal{A}$ is a normed algebra then for all $a, a', b, b' \in \mathcal{A}$ we have

$$\|ab - a'b'\| = \|ab - ab' + ab' - a'b'\| \le \|a\|\|b - b'\| + \|a - a'\|\|b'\|.$$

This proves that the multiplication map $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is jointly continuous.

2. If $\mathcal{A}$ is a normed algebra with unit $\mathbf{1}$ then $\mathbf{1} = \mathbf{1}^2$, thus $\|\mathbf{1}\| = \|\mathbf{1}^2\| \le \|\mathbf{1}\|^2$. With $\|\mathbf{1}\| \neq 0$ this implies $1 \le \|\mathbf{1}\|$. Some authors require all unital normed algebras to satisfy $\|\mathbf{1}\| = 1$, but we don't. Of course this does hold for $B(E)$. □

By the above, $B(E)$ is a normed algebra for every normed space E, and by Proposition 3.25(ii), $B(E)$ is a Banach algebra whenever E is a Banach space. There is another standard class of examples:

3.30 Example Let X be a compact topological space and $\mathcal{A} = C(X, \mathbb{F})$.[21] We already know that $\mathcal{A}$, equipped with the norm $\|f\| = \sup_{x\in X} |f(x)|$ is a Banach space. The pointwise product $(fg)(x) = f(x)g(x)$ of functions is bilinear, associative and clearly satisfies $\|fg\| \le \|f\|\|g\|$. This makes $(\mathcal{A}, \|\cdot\|, \cdot)$ a Banach algebra. An analogous result holds for the algebra $C_b(X, \mathbb{F})$ of bounded continuous functions on a not necessarily compact space X.

We will have much more to say about Banach algebras later in the course.

Before we go on developing the general theory of Banach spaces, it is instructive to study in some depth an important class of spaces, the spaces $\ell^p(S, \mathbb{F})$, where everything can be done very explicitly, in particular the dual spaces can be determined.

4 The sequence spaces and their dual spaces

In this section we will consider in some detail the spaces $\ell^p(S, \mathbb{F})$, which we call sequence spaces even though strictly speaking this is correct only for $S = \mathbb{N}$. They are worth studying for several reasons:

- They provide a first encounter with the more general Lebesgue spaces $L^p(X, \mathcal{A}, \mu)$ without the measure and integration theoretic baggage needed for the latter.
- They can be studied quite completely and have their dual spaces identified.
- We will see that every Hilbert space is isometrically isomorphic to $\ell^2(S, \mathbb{F})$ for some S.

[21] Some authors, mostly operator algebraists, write $C(X)$ for $C(X, \mathbb{C})$, whereas topologists put $C(X) = C(X, \mathbb{R})$. Due to this ambiguity we avoid the notation $C(X)$.

4.1 Basics. $1 \le p \le \infty$: Hölder and Minkowski inequalities

4.1 Definition (ℓ^p-Spaces) *If $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, $0 < p < \infty$, S is a set and $f : S \to \mathbb{F}$, define*

$$\|f\|_\infty = \sup_{s \in S} |f(s)| \in [0, \infty], \qquad \|f\|_p = \left(\sum_{s \in S} |f(s)|^p \right)^{1/p} \in [0, \infty],$$

where $\infty^{1/p} = \infty$ and we use the notion of unordered sums, cf. Appendix A.1. Now for all $p \in (0, \infty]$ put

$$\ell^p(S, \mathbb{F}) := \{ f : S \to \mathbb{F} \mid \|f\|_p < \infty \}.$$

Here are some immediate observations:

- We have $|f(s)| \le \|f\|_p$ for all $s \in S$ and $p \in (0, \infty]$.
- $\|f\|_p = 0$ if and only if $f = 0$.
- For all $c \in \mathbb{F}$ we have $\|cf\|_p = |c|\|f\|_p$ (with the understanding that $0 \cdot \infty = 0$).
- If S is finite then $\ell^p(S, \mathbb{F}) = \{f : S \to \mathbb{F}\} = \mathbb{F}^S$. If $\#S = 1$ then all the $\|\cdot\|_p$ coincide.
- If $\#S = \infty$ then $\|\cdot\|_p$ is not a norm on $\mathbb{F}^S$ in the sense of Definition 2.13 since we can have $\|f\|_p = +\infty$. But we'll see that the restriction of $\|\cdot\|_p$ to $\ell^p(S, \mathbb{F})$ is a norm if $p \in [1, \infty]$.

4.2 Lemma *(i) $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are subadditive, thus norms on $\ell^1(S, \mathbb{F})$ and $\ell^\infty(S, \mathbb{F})$, resp.*

(ii) If $f \in \ell^1(S, \mathbb{F})$ and $g \in \ell^\infty(S, \mathbb{F})$ then

$$\left| \sum_{s \in S} f(s) g(s) \right| \le \|fg\|_1 \le \|f\|_1 \|g\|_\infty.$$

(iii) $(\ell^p(S, \mathbb{F}), \|\cdot\|_p)$ are vector spaces for all $p \in (0, \infty]$.

Proof. (i) In view of the preceding observations, it remains to prove subadditivity:

$$\begin{aligned} \|f+g\|_\infty &= \sup_s |f(s)+g(s)| \le \sup_s |f(s)| + \sup_s |g(s)| = \|f\|_\infty + \|g\|_\infty, \\ \|f+g\|_1 &= \sum_s |f(s)+g(s)| \le \sum_s (|f(s)| + |g(s)|) = \|f\|_1 + \|g\|_1. \end{aligned}$$

(ii) This just is $|\sum_{s \in S} f(s)g(s)| \le \sum_{s \in S} |f(s)||g(s)| \le \|g\|_\infty \sum_{s \in S} |f(s)| = \|g\|_\infty \|f\|_1$.

(iii) It remains to show that $f+g \in \ell^p(S, \mathbb{F})$ whenever $f, g \in \ell^p(S, \mathbb{F})$. For $p = \infty$ this follows from $\|\cdot\|_\infty$ being a norm. The map $\mathbb{R}_+ \to \mathbb{R}_+$, $t \mapsto t^p$ is monotonous for all $p \in (0, \infty)$. Thus

$$|a+b|^p \le (|a|+|b|)^p \le (2\max(|a|,|b|))^p = 2^p \max(|a|^p, |b|^p) \le 2^p(|a|^p + |b|^p), \tag{4.1}$$

so that with $f, g \in \ell^p(S, \mathbb{F})$ we have

$$\|f+g\|_p^p = \sum_s |f(s)+g(s)|^p \le 2^p \sum_s \left(|f(s)|^p + |g(s)|^p \right) = 2^p(\|f\|_p^p + \|g\|_p^p) < \infty.$$

Thus $\|f+g\|_p < \infty$ and $f + g \in \ell^p(S, \mathbb{C})$. ■

In order to obtain analogues of (i), (ii) for $1 < p < \infty$, we put

4.3 Definition *If $p, q \in [1, \infty]$ we say that p and q are dual (or conjugate) to each other if $\frac{1}{p} + \frac{1}{q} = 1$, with the understanding $\frac{1}{\infty} = 0$, $\frac{1}{0} = \infty$.*

One easily checks that every $p \in [1, \infty]$ has a unique conjugate $q \in [1, \infty]$. And for $1 < p, q < \infty$ conjugacy is equivalent to $pq = p + q$.

4.4 PROPOSITION *Let $1 < p < \infty$ and let q be conjugate to p, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Then*

(i) For all $f, g : S \to \mathbb{F}$ we have $\|fg\|_1 \leq \|f\|_p \|g\|_q$. (Inequality of Hölder[22] (1889))

(ii) For all $f, g : S \to \mathbb{F}$ we have $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. (Inequality of Minkowski[23] (1896))

Proof. (i) The inequality is trivially true if $\|f\|_p$ or $\|g\|_q$ is zero or infinite. Thus we assume $\|f\|_p$ and $\|g\|_q$ to be finite and non-zero. The exponential function $\mathbb{R} \to \mathbb{R}, x \mapsto e^x$ is convex[24], so that with of $\frac{1}{p} + \frac{1}{q} = 1$ we have

$$e^{a/p} e^{b/q} = \exp\left(\frac{a}{p} + \frac{b}{q}\right) \leq \frac{e^a}{p} + \frac{e^b}{q} \qquad \forall a, b \in \mathbb{R}.$$

With the substitutions $e^a = u^p, e^b = v^q$, where $u, v > 0$ this becomes

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q} \qquad \forall u, v \geq 0. \tag{4.2}$$

(The validity also for $u = 0$ or $v = 0$ is obvious.)

Putting $u = |f(s)|$, $v = |g(s)|$ in (4.2), we have $|f(s)g(s)| \leq p^{-1}|f(s)|^p + q^{-1}|g(s)|^p$, so that summing over s gives $\|fg\|_1 \leq p^{-1}\|f\|_p^p + q^{-1}\|g\|_q^q$. If $\|f\|_p = \|g\|_q = 1$ then this reduces to $\|fg\|_1 \leq \frac{1}{p} + \frac{1}{p} = 1$. With $f' = f/\|f\|_p$, $g' = g/\|g\|_q$ we have $\|f'\|_p = 1 = \|g'\|_q$, so that the above implies $\|f'g'\|_1 \leq 1$. Inserting the definitions of f', g' herein gives $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

(ii) We may assume that $f, g \in \ell^p$, thus $\|f\|_p, \|g\|_p < \infty$, since otherwise the inequality is trivially true. With $h = f + g$ this implies $h \in \ell^p$ by Lemma 4.2(iii). If q is conjugate to p, we have $pq = p + q$, and we find $\sum_s |h(s)|^{(p-1)q} = \sum_s |h(s)|^p < \infty$, so that the function $s \mapsto |h(s)|^{p-1}$ is in ℓ^q with $\||h|^{p-1}\|_q = \|h\|_p^{p/q}$. Now

$$\begin{aligned} \|f + g\|_p^p &= \sum_s |f(s) + g(s)|^p = \sum_s |f(s) + g(s)|\, |f(s) + g(s)|^{p-1} \\ &\leq \sum_s (|f(s)| + |g(s)|)\ |f(s) + g(s)|^{p-1}. \end{aligned} \tag{4.3}$$

Since $f \in \ell^p$ and $|h|^{p-1} \in \ell^q$, we can apply Hölder's inequality, obtaining $\sum_s |f(s)||h(s)|^{p-1} \leq \|f\|_p \||h|^{p-1}\|_q = \|f\|_p \|h\|_p^{p/q}$. Analogously, $\sum_s |g(s)||h(s)|^{p-1} \leq \|g\|_p \|h\|_p^{p/q}$. Plugging this into (4.3) we obtain

$$\|f + g\|_p^p \leq \Big(\|f\|_p + \|g\|_p\Big) \|f + g\|_p^{p/q}.$$

If $\|f + g\|_p \neq 0$ we divide by $\|f + g\|_p^{p/q}$, and using $p - p/q = p(1 - 1/q) = p\frac{1}{p} = 1$ we obtain

$$\|f + g\|_p = \|f + g\|_p^{p - p/q} \leq \|f\|_p + \|g\|_p.$$

Since this clearly also holds if $\|f + g\|_p = 0$, we are done. ■

---

[22] Otto Hölder (1859-1937). German mathematician. Important contributions to analysis and algebra.

[23] Hermann Minkowski (1864-1909). German mathematician. Contributions to number theory, relativity and other fields. We'll encounter M.-functionals.

[24] $f : [a, b] \to \mathbb{R}$ is convex if $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ for all $x, y \in [a, b]$ and $t \in [0, 1]$ and strictly convex if the inequality is strict whenever $x \neq y, 0 < t < 1$. See, e.g., [57, Vol. 1, Section 7.2].

For $p = q = 2$, the inequality of Hölder is known as the Cauchy-Schwarz inequality. We will also call the trivial inequalities of Lemma 4.2 for $\{p, q\} = \{1, \infty\}$ Hölder and Minkowski inequalities. Now the analogue of Lemma 4.2 for $1 < p < \infty$ is clear:

4.5 Corollary *Let $1 < p < \infty$. Then*

(i) $(\ell^p(S, \mathbb{F}), \|\cdot\|_p)$ *is a normed vector space.*[25]

(ii) If q is conjugate to p and $f \in \ell^p(S, \mathbb{F})$ and $g \in \ell^q(S, \mathbb{F})$ then

$$\left| \sum_{s \in S} f(s) g(s) \right| \le \|fg\|_1 \le \|f\|_p \|g\|_q.$$

4.2 ⋆ Aside: The translation-invariant metric d_p for $0 < p < 1$

For $s \in S$, let $\delta_s : S \to \mathbb{F}$ be the function defined by $\delta_s(t) = \delta_{s,t}$ (which is 1 for $s = t$ and zero otherwise).

4.6 Proposition *If $0 < p < 1$ then*

(i) $\|\cdot\|_p$ *violates subadditivity whenever $\#S \ge 2$, thus is not a norm.*

(ii) Nevertheless, $\ell^p(S, \mathbb{F})$ *is a vector space.*

(iii) Restricted to $\ell^p(S, \mathbb{F})$,

$$d_p(f, g) = \sum_{s \in S} |f(s) - g(s)|^p$$

defines a translation-invariant metric. (Note the absence of the p-th root present in $\|\cdot\|_p$!)

(iv) $\ell^p(S, \mathbb{F})$ *a topological vector space when given the metric topology* τ_{d_p}.

Proof. (i) Pick $s, t \in S$, $s \ne t$ and put $f = \delta_s, g = \delta_t$. Now $\|f\|_p = \|g\|_p = 1$ and

$$2 < 2^{1/p} = \|f + g\|_p \not\le \|f\|_p + \|g\|_p = 2$$

since $1/p > 1$. Thus $\|\cdot\|_p$ is not subadditive and therefore not a norm.

(ii) The proof of Lemma 4.2(iii) included the case $0 < p < 1$.

(iii) That $d_p(f, g) < \infty$ for all $f, g \in \ell^p(S, \mathbb{F})$ follows from ℓ^p being a vector space. Translation invariance of d_p and the axioms $d_p(f, g) = d_p(g, f)$ and $d_p(f, g) = 0 \Leftrightarrow f = g$ are all evident from the definition. We claim that

$$0 < p < 1,\ a, b \ge 0 \quad \Rightarrow \quad (a + b)^p \le a^p + b^p.$$

Believing this for a minute, we have

$$\begin{aligned} d_p(f, h) &= d_p(f - h, 0) = \sum_s |f(s) - h(s)|^p \le \sum_s (|f(s) - g(s)| + |g(s) - h(s)|)^p \\ &\le \sum_s (|f(s) - g(s)|^p + |g(s) - h(s)|^p) \\ &= d_p(f - g, 0) + d_p(g - h, 0) = d_p(f, g) + d_p(g, h), \end{aligned}$$

[25] If you wonder why the L^p-spaces with $p \notin \{1, 2, \infty\}$ are studied at all, the short answer is that they arise in many areas of analysis, in particular harmonic analysis and PDE theory. For longer answers, have a look at [180].

as wanted, where first used the triangle inequality and then the claim.

Turning to our claim $(a+b)^p \le a^p + b^p$, it is clear that this holds if $a = 0$. For $a = 1$ it reduces to $(1+b)^p \le 1 + b^p \ \forall b \ge 0$. For $b = 0$ this is true, and for all $b > 0$ it follows from the fact that

$$\frac{d}{db}(1 + b^p - (1+b)^p) = p(b^{p-1} - (b+1)^{p-1}) > 0$$

due to $p - 1 < 0$. If now $a > 0$ then

$$(a+b)^p = a^p(1 + (b/a))^p \le a^p(1 + (b/a)^p) = a^p + b^p,$$

and we are done. (By almost the same argument, for $p \ge 1$ we have $(a+b)^p \ge a^p + b^p$.)

(iv) By the above, d_p is a translation-invariant metric, so that the addition operation $+ : \ell^p \times \ell^p \to \ell^p$ is jointly continuous by Lemma 2.9. Furthermore,

$$d_p(cf, 0) = \sum_s |cf(s)|^p = |c|^p \sum_s |f(s)|^p = |c|^p d_p(f, 0),$$

which tends to zero if $c \to 0$ for fixed f or $f \to 0$ (in the sense of $d_p(f, 0) \to 0$) for fixed c. By Remark 2.10.3 this implies joint continuity of the scalar action $\mathbb{F} \times \ell^p \to \ell^p$. ■

4.7 Exercise Let S be an infinite set and $0 < p < 1$.

(i) Prove that $\ell^p(S, \mathbb{F})$ does not contain a bounded convex open neighborhood of 0.

(ii) Conclude that the metrizable TVS $(\ell^p(S, \mathbb{F}), d_p)$ are not normable if $0 < p < 1$.

In fact, the ℓ^p spaces with $0 < p < 1$ are not even locally convex. This leads to strange behavior. For example the dual space $\ell^p(S, \mathbb{F})^*$ is unexpected, cf. [108]. This strangeness is even more pronounced for the continuous versions $L^p(X, \mathcal{A}, \mu)$: For $X = [0, 1]$ equipped with Lebesgue measure, one has $L^p(X, \mathcal{A}, \mu)^* = \{0\}$, which cannot happen for a non-zero Banach (or locally convex) space in view of the Hahn-Banach theorem.

4.3 c_{00} and c_0. Completeness of $\ell^p(S, \mathbb{F})$ and $c_0(S, \mathbb{F})$

In what follows we put

$$d_p(f,g) = \begin{cases} \|f-g\|_\infty = \sup_s |f(s) - g(s)| & \text{if} \quad p = \infty \\ \|f-g\|_p = (\sum_s |f(s)-g(s)|^p)^{1/p} & \text{if} \quad 1 \le p < \infty \\ \|f-g\|_p^p = \sum_s |f(s)-g(s)|^p & \text{if} \quad 0 < p < 1 \end{cases}$$

which is a metric in all cases. For a function $f : S \to \mathbb{F}$ we define $\operatorname{supp} f = \{s \in S \mid f(s) \neq 0\}$.

4.8 Lemma *Let $p \in (0, \infty]$ and $d_p(x, y) = \|x - y\|_p$. Then $(\ell^p(S, \mathbb{F}), d_p)$ is complete for every set S and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.*

Proof. For all $p \in (0, \infty]$ we have $|f(s) - g(s)| \le \|f - g\|_p$. Thus for $p \ge 1$ we have $|f(s) - g(s)| \le d_p(f, g)$, while for $p \in (0, 1)$ we have $|f(s) - g(s)| \le \|f - g\|_p = d_p(f, g)^{1/p}$. In either case $d(f, g) \to 0$ implies $f(s) - g(s) \to 0$ for all $s \in S$. Thus if $\{f_n\} \subseteq \ell^p(S, \mathbb{F})$ is a Cauchy sequence w.r.t. d_p then $\{f_n(s)\}$ is a Cauchy sequence in $\mathbb{F}$, thus convergent for each $s \in S$. Defining $g(s) = \lim_{n\to\infty} f_n(s)$, it remains to prove $g \in \ell^p(S, \mathbb{F})$ and $d_p(f_n, g) \to 0$.

For $p = \infty$ and $\varepsilon > 0$ we can find n_0 such that $n, m \ge n_0$ implies $\|f_n - f_m\|_\infty < \varepsilon$, which readily gives $\|f_m\|_\infty \le \|f_{n_0}\|_\infty + \varepsilon$ for all $m \ge n_0$. Thus also $\|g\|_\infty \le \|f_{n_0}\|_\infty + \varepsilon < \infty$. Taking $m \to \infty$ in $\sup_s |f_n(s) - f_m(s)| < \varepsilon$ gives $\sup_s |f_n(s) - g(s)\| \le \varepsilon$, whence $\|f_n - g\|_\infty \to 0$.

For $0 < p < \infty$ we give a uniform argument. Since $\{f_n\}$ is Cauchy w.r.t. d_p, for $\varepsilon > 0$ we can find n_0 such that $n, m \geq n_0$ implies $d_p(f_n, f_m) < \varepsilon$. Applying the dominated convergence theorem (in the simple case of an infinite sum rather than a general integral, cf. Proposition A.3) gives $d_p(g, f_m) = \lim_{n\to\infty} d_p(f_n, f_m) \leq \varepsilon$. This implies both $g \in \ell^p(S, \mathbb{F})$ and $d_p(g, f_m) \to 0$ as $m \to \infty$. ■

4.9 Definition *For a set S and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ we define*

$$\begin{aligned} c_{00}(S, \mathbb{F}) &= \{f : S \to \mathbb{F} \mid \#(\operatorname{supp} f) < \infty\}, \\ c_0(S, \mathbb{F}) &= \{f : S \to \mathbb{F} \mid \varepsilon > 0 \ \Rightarrow \ \#\{s \in S \mid |f(s)| \geq \varepsilon\} < \infty\}. \end{aligned}$$

(The elements of c_0 are the functions that 'tend to zero at infinity'.)

4.10 Lemma *If $0 < p \leq q < \infty$, we have the set-theoretic inclusions*

$$c_{00}(S, \mathbb{F}) \subseteq \ell^p(S, \mathbb{F}) \subseteq \ell^q(S, \mathbb{F}) \subseteq c_0(S, \mathbb{F}) \subseteq \ell^\infty(S, \mathbb{F}),$$

where $\|f\|_\infty \leq \|f\|_q \leq \|f\|_p$. All the inclusion maps except the first have norm one.

Proof. If $f \in c_{00}(S, \mathbb{F})$ then clearly $\|f\|_p < \infty$ for all $p \in (0, \infty]$. And $f \in c_0(S, \mathbb{F})$ implies boundedness of f. This gives the first and last inclusion. The map $c_0(S, \mathbb{F}) \hookrightarrow \ell^\infty(S, \mathbb{F})$ is an isometry since both spaces have the norm $\|\cdot\|_\infty$.

If $f \in \ell^q(S, \mathbb{F})$ with $q \in (0, \infty)$ then finiteness of $\|f\|_q^q = \sum_{s\in S} |f(s)|^q$ implies that $\{s \in S \mid |f(s)| \geq \varepsilon\}$ is finite for each $\varepsilon > 0$, thus $f \in c_0(S, \mathbb{F})$. Since $|f(s)| \leq \|f\|_q$ for all s, we have $\|f\|_\infty = \sup_{s\in S} |f(s)| \leq \|f\|_q$.

Now let $0 < p < q < \infty$ and $f \in \ell^p(S, \mathbb{F})$ with $\|f\|_p = 1$. Then $|f(s)| \leq 1 \ \forall s$, thus

$$\|f\|_q^q = \sum_{s\in S} |f(s)|^q = \sum_{s\in S} |f(s)|^{\frac{q}{p}\cdot p} \leq \sum_{s\in S} |f(s)|^p = \|f\|_p^p = 1, \tag{4.4}$$

where we used $q/p > 1$ and $|f(s)| \leq 1 \ \forall s$. Thus $\|f\|_q \leq 1$. Applying this to $f = g/\|g\|_p$ gives $\|g\|_q \leq \|g\|_p$ for all $g \in \ell^p(S, \mathbb{F})$.

We now have that all the inclusion maps have norm ≤ 1. Taking $f = \delta_s$ and using $\|\delta_s\|_p = 1$ for all $p \in (0, \infty]$ gives that the inclusion maps all have norm one. ■

4.11 Remark While we have found continuous maps between them, the spaces ℓ^p $(1 \leq p < \infty)$ and c_0 are mutually non-isomorphic. See Corollary B.32 for an even stronger statement. □

4.12 Lemma *(i) We have*

$$\overline{c_{00}(S, \mathbb{F})}^{\|\cdot\|_p} = \begin{cases} \ell^p(S, \mathbb{F}) & \text{if } 0 < p < \infty \\ c_0(S, \mathbb{F}) & \text{if } p = \infty \end{cases}$$

(ii) $(c_0(S, \mathbb{F}), \|\cdot\|_\infty)$ is complete.

Proof. (i) Let $0 < p < \infty$ and $f \in \ell^p(S, \mathbb{F})$. Then $\sum_{s\in S} |f(s)|^p = \|f\|_p^p$ implies that for each $\varepsilon > 0$ there is a finite $F \subseteq S$ such that $\|f\|_p^p - \sum_{s\in F} |f(s)|^p < \varepsilon$. Putting $g(s) = f(s)\chi_F(s)$, we have $g \in c_{00}(S, \mathbb{F})$ and $\|f - g\|_p^p = \sum_{s\in S\setminus F} |f(s)|^p < \varepsilon$. Since $\varepsilon > 0$ is arbitrary, $c_{00} \subseteq \ell^p$ is dense.

If $f \in c_0(S, \mathbb{F})$ and $\varepsilon > 0$ then $F = \{s \in S \mid |f(s)| \geq \varepsilon\}$ is finite. Now $g = f\chi_F$ is in $c_{00}(S, \mathbb{F})$ and $\|f - g\|_\infty < \varepsilon$, proving $f \in \overline{c_{00}(S, \mathbb{F})}^{\|\cdot\|_\infty}$. And $f \in \overline{c_{00}(S, \mathbb{F})}^{\|\cdot\|_\infty}$ means that for each $\varepsilon > 0$

there is a $g \in c_{00}(S,\mathbb{F})$ with $\|f-g\|_\infty < \varepsilon$. But this means $|f(s)| < \varepsilon$ for all $s \in S\backslash F$, where $F = \text{supp}(g)$ is finite. Thus $f \in c_0(S,\mathbb{F})$.

(ii) Being the closure of $c_{00}(S,\mathbb{F})$ in $\ell^\infty(S,\mathbb{F})$, $c_0(S,\mathbb{F})$ is closed, thus complete by completeness of $\ell^\infty(S,\mathbb{F})$, cf. Lemmas 4.8 and 3.21. ■

While the finitely supported functions are not dense in $(\ell^\infty(S,\mathbb{F}), \|\cdot\|_\infty)$ (for infinite S), the finite-image functions are:

4.13 Lemma *The set $\{f : S \to \mathbb{F} \mid \#f(S) < \infty\}$ of functions assuming only finitely many values, equivalently, the set of finite linear combinations $\sum_{k=1}^K c_k \chi_{A_k}$ of characteristic functions, is dense in $(\ell^\infty(S,\mathbb{F}), \|\cdot\|_\infty)$.*

Proof. We prove this for $\mathbb{F} = \mathbb{R}$, from which the case $\mathbb{F} = \mathbb{C}$ is easily deduced. Let $f \in \ell^\infty(S,\mathbb{F})$ and $\varepsilon > 0$. For $k \in \mathbb{Z}$ define $A_k = f^{-1}([k\varepsilon, (k+1)\varepsilon))$. Define $K = \lceil \frac{\|f\|_\infty}{\varepsilon} \rceil + 1$ and $g = \varepsilon \sum_{|k| \le K} k \chi_{A_k}$. Then g has finite image and $\|f-g\|_\infty < \varepsilon$. ■

4.14 Exercise Let S be an infinite set and $0 < p < q < \infty$.

(i) Prove that all inclusions in Lemma 4.10(i) are strict.

(ii) For $p < q$, show that $\ell^p(S)$ is not a closed subspace of $(\ell^q(S), \|\cdot\|_q)$ and analogously for $\ell^q(S) \subseteq (c_0(S), \|\cdot\|_\infty)$. Why is this compatible with Corollary 3.23?

4.15 Exercise For $f : S \to \mathbb{C}$, prove

$$\lim_{p\to\infty} \|f\|_p = \begin{cases} \|f\|_\infty & \text{if } \|f\|_p < \infty \text{ for some } p \in [1,\infty) \\ \infty & \text{otherwise} \end{cases}$$

4.16 Exercise Let $1 < p \le \infty$ and $V = \ell^p(\mathbb{N},\mathbb{R})$. Define $\delta_n \in V$ by $\delta_n(m) = \delta_{n,m}$ and $x_n = n^{-\alpha}\delta_n$, where $\alpha > 0$.

(i) For which $\alpha > 0$ is $\sum_{n=1}^\infty x_n$ absolutely convergent?

(ii) For which $\alpha > 0$ is $\sum_{n=1}^\infty x_n$ unconditionally convergent? (Cf. Remark A.5.)

(iii) Use (i),(ii) to give examples of series that are unconditionally convergent, but not absolutely convergent, in each $\ell^p(\mathbb{N},\mathbb{F})$ with $1 < p \le \infty$.

(iv) BONUS: As (iii), but for $p = 1$. (NB: Just invoking Corollary B.3 is not enough!)

4.4 Separability of $\ell^p(S,\mathbb{F})$ and $c_0(S,\mathbb{F})$

4.17 Proposition *Let $p \in (0,\infty)$. The metric space $(\ell^p(S,\mathbb{F}), d_p)$, where $d_p(f,g) = \|f-g\|_p$, is separable ($\Leftrightarrow$ second countable) if and only if the set S is countable.*

Proof. We prove this for $\mathbb{F} = \mathbb{R}$, from which the claim for $\mathbb{F} = \mathbb{C}$ is easily deduced. For $f : S \to \mathbb{R}$, let $\text{supp}(f) := \{s \in S \mid f(s) \neq 0\} \subseteq S$ be the support of f. Now, if S is countable, then $Y = \{g : S \to \mathbb{Q} \mid \#(\text{supp}(g)) < \infty\} \subseteq \ell^p(S,\mathbb{R})$ is countable, and we claim that $\overline{Y} = \ell^p(S,\mathbb{R})$. To prove this, let $f \in \ell^p(S,\mathbb{R})$ and $\varepsilon > 0$. Since $\|f\|_p^p = \sum_{s\in S} |f(s)|^p < \infty$, there is a finite subset $T \subseteq S$ such that $\sum_{s\in S\backslash T} |f(s)|^p < \varepsilon/2$. On the other hand, since $\mathbb{Q}^{\#T} \subseteq \mathbb{R}^{\#T}$ is dense, we can choose $g : T \to \mathbb{Q}$ such that $\sum_{t\in T} |f(t)-g(t)|^p < \varepsilon/2$. Defining g to be zero on $S\backslash T$, we have $g \in Y$ and

$$\|f-g\|^p = \sum_{t\in T} |f(t)-g(t)|^p + \sum_{s\in S\backslash T} |f(s)|^p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and since $\varepsilon > 0$ was arbitrary, $Y \subseteq \ell^p(S,\mathbb{R})$ is dense.

For the converse, assume that S is uncountable. By Proposition A.2(iii), $\mathrm{supp}(f)$ is countable for every $f \in \ell^p(S,\mathbb{R})$. Thus if $Y \subseteq \ell^p(S,\mathbb{R})$ is countable then $T = \bigcup_{f\in Y} \mathrm{supp}(f) \subseteq S$ is a countable union of countable sets and therefore countable. Thus all functions $f \in Y$ vanish on $S\backslash T \neq \emptyset$, and the same holds for $f \in \overline{Y}$ since the coordinate maps $f \mapsto f(s)$ are continuous in view of $|f(s)| \leq \|f\|$. Thus Y cannot be dense. ■

4.18 Exercise With $d_\infty(f,g) = \|f-g\|_\infty$ prove

(i) The space $(\ell^\infty(S,\mathbb{F}), d_\infty)$ is separable if and only if S is finite.

(ii) The space $(c_0(S,\mathbb{F}), d_\infty)$ is separable if and only if S is countable.

Hint: For (i), consider $\{0,1\}^S \subseteq \ell^\infty(S)$.

4.5 Dual spaces of $\ell^p(S,\mathbb{F})$, $1 \leq p < \infty$, and $c_0(S,\mathbb{F})$

If $(V, \|\cdot\|)$ is a normed vector space over $\mathbb{F}$ and $\varphi : V \to \mathbb{F}$ is a linear functional, Definition 3.2 specializes to

$$\|\varphi\| = \sup_{0\neq x\in V} \frac{|\varphi(x)|}{\|x\|} = \sup_{\substack{x\in V\\ \|x\|\leq 1}} |\varphi(x)|.$$

Recall that the dual space $V^* = \{\varphi : V \to \mathbb{F} \text{ linear} \mid \|\varphi\| < \infty\}$ is a Banach space with norm $\|\varphi\|$. The aim of this section is to concretely identify $\ell^p(S,\mathbb{F})^*$ for $1 \leq p < \infty$ and $c_0(S,\mathbb{F})^*$. (We will have something to say about $\ell^\infty(S,\mathbb{F})^*$, but the complete story would lead us too far.)

For the purpose of the following proof, it will be useful to define $\mathrm{sgn} : \mathbb{C} \to \mathbb{C}$ by $\mathrm{sgn}(0) = 0$ and $\mathrm{sgn}(z) = z/|z|$ otherwise. Then $z = \mathrm{sgn}(z)|z|$ and $|z| = \overline{\mathrm{sgn} z}\, z$ for all $z \in \mathbb{C}$.

4.19 Theorem *(i) Let $p \in [1,\infty]$ with conjugate value q. Then for each $g \in \ell^q(S,\mathbb{F})$ the map $\varphi_g : \ell^p(S,\mathbb{F}) \to \mathbb{F},\ f \mapsto \sum_{s\in S} f(s)g(s)$ satisfies $\|\varphi_g\| \leq \|g\|_q$, thus $\varphi_g \in \ell^p(S,\mathbb{F})^*$. And the map $\iota : \ell^q(S,\mathbb{F}) \to \ell^p(S,\mathbb{F})^*,\ g \mapsto \varphi_g$, called the canonical map, is linear with $\|\iota\| \leq 1$.*

(ii) For all $1 \leq p \leq \infty$ the canonical map $\ell^q(S,\mathbb{F}) \to \ell^p(S,\mathbb{F})^*$ is isometric.

(iii) If $1 \leq p < \infty$, the canonical map $\ell^q(S,\mathbb{F}) \to \ell^p(S,\mathbb{F})^*$ is surjective, thus $\ell^p(S,\mathbb{F})^* \cong \ell^q(S,\mathbb{F})$.

(iv) The canonical map $\ell^1(S,\mathbb{F}) \to c_0(S,\mathbb{F})^*$ is an isometric bijection, thus $c_0(S,\mathbb{F})^* \cong \ell^1(S,\mathbb{F})$.

(v) If S is finite, the canonical map $\ell^1(S,\mathbb{F}) \to \ell^\infty(S,\mathbb{F})^*$ is surjective. But its image is a proper closed subspace of $\ell^\infty(S,\mathbb{F})^*$ whenever S is infinite.

Proof. (i) For all $p \in [1,\infty]$ and conjugate q we have

$$|\sum_s f(s)g(s)| \leq \sum_{s\in S} |f(s)g(s)| \leq \|f\|_p\|g\|_q < \infty \quad \forall f \in \ell^p, g \in \ell^q$$

by Hölder's inequality. In either case, the absolute convergence for all f,g implies that $(f,g) \mapsto \sum_s f(s)g(s)$ is bilinear.

(ii) If $\|g\|_\infty \neq 0$ and $\varepsilon > 0$ there is an $s \in S$ with $|g(s)| > \|g\|_\infty - \varepsilon$. If $f = \delta_s : t \mapsto \delta_{s,t}$, we have $|\varphi_g(f)| = |g(s)| > \|g\|_\infty - \varepsilon$. Since $\|f\|_1 = 1$, this proves $\|\varphi_g\| > \|g\|_\infty - \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have $\|\varphi_g\| \geq \|g\|_\infty$.

If $\|g\|_1 \neq 0$, define $f(s) = \overline{\mathrm{sgn}(g(s))}$. Then $\|f\|_\infty = 1$ and $\sum_s f(s)g(s) = \sum_s |g(s)| = \|g\|_1$. This proves $\|\varphi_g\| \geq \|g\|_1$.

If $1 < p, q < \infty$ and $\|g\|_q \neq 0$, define $f(s) = \overline{\operatorname{sgn}(g(s))}|g(s)|^{q-1}$. Then

$$\begin{aligned} \sum_s f(s)g(s) &= \sum_s |g(s)|^q = \|g\|_q^q, \\ \|f\|_p^p &= \sum_s |f(s)|^p = \sum_{s,\, g(s)\neq 0} |g(s)|^{(q-1)p} = \sum_s |g(s)|^q = \|g\|_q^q, \end{aligned}$$

where we used $p + q = pq$, whence $(q-1)p = q$. The above gives

$$\|\varphi_q\| \geq \frac{|\sum_s f(s)g(s)|}{\|f\|_p} = \frac{\|g\|_q^q}{\|f\|_p} = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q^{q(1-1/p)} = \|g\|_q.$$

We thus have proven $\|\varphi_g\| \geq \|g\|_q$ in all cases and since the opposite inequality is known from (i), $g \mapsto \varphi_g$ is isometric.

(iii) Let $0 \neq \varphi \in \ell^1(S, \mathbb{F})^*$. Define $g : S \to \mathbb{F}$ by $g(s) = \varphi(\delta_s)$. With $\|\delta_s\|_1 = 1$, we have $|g(s)| = |\varphi(\delta_s)| \leq \|\varphi\|$ for all $s \in S$, thus $\|g\|_\infty \leq \|\varphi\|$. If $f \in \ell^1(S, \mathbb{F})$ and $F \subseteq S$ is finite, we have $\varphi(f\chi_F) = \varphi(\sum_{s\in F} f(s)\delta_s) = \sum_{s\in F} f(s)g(s)$. In the limit $F \nearrow S$ this becomes $\varphi(f) = \sum_{s\in S} f(s)g(s) = \varphi_g(f)$ (since $fg \in \ell^1$, thus the r.h.s. is absolutely convergent, and $\|f(1-\chi_F)\|_1 \to 0$ and φ is $\|\cdot\|_1$-continuous). This proves $\varphi = \varphi_g$ with $g \in \ell^\infty(S, \mathbb{F})$.

Now let $1 < p, q < \infty$, and let $0 \neq \varphi \in \ell^p(S, \mathbb{F})^*$. Since $\ell^1(S, \mathbb{F}) \subseteq \ell^p(S, \mathbb{F})$ by Lemma 4.10, we can restrict φ to $\ell^1(S, \mathbb{F})^*$, and the preceding argument gives a $g \in \ell^\infty(S, \mathbb{F})$ such that $\varphi(f) = \sum_{s\in S} f(s)g(s)$ for all $f \in \ell^1(S, \mathbb{F})$. The arguments in the proof of (ii) also show that for $1 < p, q < \infty$ and *any* function $g : S \to \mathbb{F}$ we have

$$\|g\|_q = \sup \left\{ |\sum_{s\in S} f(s)g(s)| \mid f \in c_{00}(S, \mathbb{F}), \|f\|_p \leq 1 \right\}.$$

Using this and $\varphi(f) = \sum_s f(s)g(s)$ for all $f \in c_{00}(S, \mathbb{F})$ we have

$$\|g\|_q = \sup\{|\varphi(f)| \mid f \in c_{00}(S, \mathbb{F}), \|f\|_p \leq 1\} = \|\varphi\| < \infty.$$

Now $\varphi(f) = \sum_{s\in S} f(s)g(s) = \varphi_g(f)$ for all $f \in \ell^p(S, \mathbb{F})$ follows as before from $fg \in \ell^1$ and $\|f(1-\chi_F)\|_p \to 0$ as $F \nearrow S$ and the $\|\cdot\|_p$-continuity of φ.

(iv) Let $0 \neq g \in \ell^1(S, \mathbb{F})$. Then $\varphi_g \in \ell^\infty(S, \mathbb{F})^*$, which we can restrict to $c_0(S, \mathbb{F})$. For finite $F \subseteq S$ define $f_F = f\chi_F$ with $f(s) = \overline{\operatorname{sgn}(g(s))}$. Then $f_F \in c_{00}(S, \mathbb{F})$ with $\|f_F\|_\infty = 1$ (provided $F \cap \operatorname{supp} g \neq \emptyset$) and $\varphi(f_F) = \sum_{s\in F} |g(s)|$. Thus $\|\varphi\| \geq \sum_{s\in F} |g(s)|$ for all finite F intersecting $\operatorname{supp} g$, and this implies $\|\varphi_g\| \geq \|g\|_1$. The opposite being known, we have proven that $\ell^1(S, \mathbb{F}) \to c_0(S, \mathbb{F})^*$ is isometric.

To prove surjectivity, let $0 \neq \varphi \in c_0(S, \mathbb{F})^*$ and define $g : S \to \mathbb{F}, s \mapsto \varphi(\delta_s)$. If now $f \in c_0(S, \mathbb{F})$ and $F \subseteq S$ is finite, we have $f\chi_F = \sum_{s\in F} f(s)\delta_s$, thus $\varphi(f\chi_F) = \sum_{s\in F} f(s)g(s)$. In particular with $f(s) = \overline{\operatorname{sgn}(g(s))}$ we have $\varphi(f\chi_F) = \sum_{s\in F} f(s)g(s) = \sum_{s\in F} |g(s)|$. Again we have $\|f\chi_f\|_\infty \leq \|f\|_\infty = 1$, thus $|\varphi(f\chi_F)| \leq \|\varphi\|$, and combining these observations gives $\|g\|_1 \leq \|\varphi\| < \infty$, thus $g \in \ell^1(S, \mathbb{F})$. As $F \nearrow S$, we have $\|f(1-\chi_F)\|_\infty = \|f\chi_{S\setminus F}\|_\infty \to 0$ since $f \in c_0$, thus with $\|\cdot\|_\infty$-continuity of φ

$$\varphi(f) = \lim_{F\nearrow S} \varphi(f\chi_F) = \lim_{F\nearrow S} \sum_{s\in F} f(s)g(s) = \sum_{s\in S} f(s)g(s) = \varphi_g(f),$$

where we again used $fg \in \ell^1$. Thus $\varphi = \varphi_g$, so that $\ell^1(S, \mathbb{F}) \to c_0(S, \mathbb{F})^*$ is an isometric bijection.

(v) It is clear that $\iota : \ell^1(S,\mathbb{F}) \to \ell^\infty(S,\mathbb{F})^*$ is surjective if S is finite. Closedness of the image of ι always follows from the completeness of $\ell^1(S,\mathbb{F})$ and the fact that ι is an isometry, cf. Corollary 3.23. The failure of surjectivity is deeper than the results of this section so far, so that it is illuminating to give two proofs.

First proof: If S is infinite, the closed subspace $c_0(S,\mathbb{F}) \subseteq \ell^\infty(S,\mathbb{F})$ is proper since $1 \in \ell^\infty(S,\mathbb{F})\backslash c_0(S,\mathbb{F})$. Thus the quotient space $Z = \ell^\infty(S,\mathbb{F})/c_0(S,\mathbb{F})$ is non-trivial. In Section 6.1 we will show that Z is a Banach space, thus admits non-zero bounded linear maps $\psi : Z \to \mathbb{F}$ by the Hahn-Banach theorem (Section 9), and that the quotient map $P : \ell^\infty(S,\mathbb{F}) \to Z$ is bounded. Thus $\varphi = \psi \circ P$ is a non-zero bounded linear functional on $\ell^\infty(S,\mathbb{F})$ that vanishes on the closed subspace $c_0(S,\mathbb{F})$. By (iv), the canonical map $\ell^1(S,\mathbb{F}) \to c_0(S,\mathbb{F})^*$ is isometric, thus φ_g with $g \in \ell^1(S,\mathbb{F})$ vanishes identically on $c_0(S,\mathbb{F})$ if and only if $g = 0$. Thus $\varphi \neq \varphi_g$ for all $g \in \ell^1(S,\mathbb{F})$.

Second proof: (This proof uses no (as yet) unproven results from functional but the Stone-Čech compactification from general topology. Cf. Appendix A.6.6 and [108].) Since S is discrete, $\ell^\infty(S,\mathbb{F}) = C_b(S,\mathbb{F}) \cong C(\beta S,\mathbb{F})$, where βS is the Stone-Čech compactification of S. The isomorphism is given by the unique continuous extension $C_b(S,\mathbb{F}) \to C(\beta X,\mathbb{F}), f \mapsto \widehat{f}$ with the restriction map $C(\beta S,\mathbb{R}) \to C_b(S,\mathbb{R})$ as inverse. Since S is discrete and infinite, thus non-compact, $\beta S \neq S$. If $f \in c_0(S,\mathbb{F})$ then $\widehat{f}(x) = 0$ for every $x \in \beta S\backslash S$. (Proof: Let $x \in \beta S\backslash S$. Since $\overline{X} = \beta X$, we can find a net $\{x_\iota\}$ in X such that $x_\iota \to x$. Since $x \notin X$, x_ι leaves every finite subset of X. Now $f \in c_0(S)$ and continuity of $\widehat{f}$ imply $\widehat{f}(x) = \lim \widehat{f}(x_\iota) = \lim f(x_\iota) = 0$.) Thus for such an x, the evaluation map $\psi_x : C(\beta S,\mathbb{F}) \to \mathbb{F}$, $\widehat{f} \mapsto \widehat{f}(x)$ gives rise to a non-zero bounded linear functional (in fact character) $\varphi(f) = \widehat{f}(x)$ on $C_b(S,\mathbb{F}) = \ell^\infty(S,\mathbb{F})$ that vanishes on $c_0(S,\mathbb{F})$. Now we conclude as in the first proof that $\varphi \neq \varphi_g$ for all $g \in \ell^1(S,\mathbb{F})$. ■

4.20 Remark 1. The two proofs given above for the non-surjectivity of the canonical map $\ell^1(S,\mathbb{F}) \to \ell^\infty(S,\mathbb{F})^*$ for infinite S are both non-constructive: The first used the Hahn-Banach theorem, which we will prove using Zorn's lemma, equivalent to AC. The second used the Stone-Čech compactification βS whose usual construction relies on Tychonov's theorem, which also is equivalent to the axiom of choice. (But both the Hahn-Banach theorem and the existence of the Stone-Čech compactification can be proven using only the 'ultrafilter lemma', which is strictly weaker than AC. For Hahn-Banach see Appendix B.5, for Stone-Čech e.g. [149, 108].)

2. The dual space of $\ell^\infty(S,\mathbb{F})$ can be determined quite explicitly, but it is not a space of functions on S as are the spaces $c_0(S,\mathbb{F})^*$ and $\ell^p(S,\mathbb{F})^*$. It is the space $\mathrm{ba}(S,\mathbb{F})$ of 'finitely additive $\mathbb{F}$-valued measures on S'. A discussion of this can be found in the supplementary Section B.3.2.

3. The non-constructiveness mentioned above is unavoidable: There are set theoretic frameworks without the ultrafilter lemma (but with DC_ω) in which $\ell^\infty(\mathbb{N})^* \cong \ell^1(\mathbb{N})$, see [149, §23.10]. (In this situation, all finitely additive measures on $\mathbb{N}$ are countably additive!)

4. For all $p \in (0,1)$, the dual space $\ell^p(S,\mathbb{F})^*$ equals $\{\varphi_g \mid g \in \ell^\infty(S,\mathbb{F})\} = \ell^1(S,\mathbb{F})^*$. See [108, Appendix F.6]. Thus there is no p-dependence despite the fact that the $\ell^p(S,\mathbb{F})$ are mutually non-isomorphic! □

4.6 The Banach algebras $(\ell^\infty(S,\mathbb{F}),\cdot)$ and $(\ell^1(\mathbb{Z},\mathbb{F}),\star)$

4.21 Definition *If $f,g \in \ell^\infty(S,\mathbb{F})$ we define $f \cdot g$ by $(f\cdot g)(s) = f(s)g(s)$ (pointwise product). If $f,g \in \ell^1(\mathbb{Z},\mathbb{F})$ we define the 'convolution product' $f \star g$ by*

$$(f\star g)(n) = \sum_{m\in\mathbb{Z}} f(m)g(n-m) = \sum_{\substack{k,l\in\mathbb{Z}\\ k+l=n}} f(k)g(l). \tag{4.5}$$

4.22 LEMMA *(i) If $f, g \in \ell^\infty(S,\mathbb{F})$ then $\|f \cdot g\|_\infty \leq \|f\|_\infty \|g\|_\infty$, thus $f \cdot g \in \ell^\infty(S,\mathbb{F})$.*

(ii) If $f, g \in \ell^1(\mathbb{Z},\mathbb{F})$ then $\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$, thus $f \star g \in \ell^1(\mathbb{Z},\mathbb{F})$.

(iii) The maps $\cdot : \ell^\infty(S,\mathbb{F}) \times \ell^\infty(S,\mathbb{F}) \to \ell^\infty(S,\mathbb{F})$ and $\star : \ell^1(\mathbb{Z},\mathbb{F}) \times \ell^1(\mathbb{Z},\mathbb{F}) \to \ell^1(\mathbb{Z},\mathbb{F})$ are bilinear, commutative and associative. A unit for $\cdot$ is the constant function $\mathbf{1} \in \ell^\infty(S,\mathbb{F})$, while $\delta_0(n) = \delta_{n,0}$ is a unit for $\star$. Thus $(\ell^\infty(S,\mathbb{F}), \cdot, \mathbf{1})$ and $(\ell^1(\mathbb{Z},\mathbb{F}), \star, \delta_0)$ are commutative unital Banach algebras with $\|\mathbf{1}\| = 1$.

(iv) $c_0(S,\mathbb{F}) \subseteq \ell^\infty(S,\mathbb{F})$ is a closed two-sided ideal.

Proof. (i) If $f, g \in \ell^\infty(S,\mathbb{F})$ then $\sup_{s\in S} |f(s)g(s)| \leq \sup_{s\in S} |f(s)| \sup_{s\in S} |g(s)| < \infty$.

(ii) The second claim clearly follows from the first, which is seen by

$$\begin{aligned} \|f \star g\|_1 &= \sum_{n\in\mathbb{Z}} \Big| \sum_{m\in\mathbb{Z}} f(m)g(n-m) \Big| \leq \sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} |f(m)g(n-m)| \\ &= \sum_m \Big(|f(m)| \sum_n |g(n-m)| \Big) = \sum_n |f(n)| \sum_m |g(m)| = \|f\|_1 \|g\|_1. \end{aligned}$$

(iii) Bilinearity of both maps is obvious, as are commutativity and associativity of $\cdot$. Commutativity of $*$ is clear from the rightmost expression in (4.5). The latter is easily seen to imply

$$((f \star g) \star h)(n) = \sum_{\substack{k,l,m\in\mathbb{Z} \\ k+l+m=n}} f(k)g(l)h(m) = (f \star (g \star h))(n),$$

thus associativity of $\star$. The statements about units are easy, (i),(ii) give submultiplicativity of the norms, and completeness was proven earlier. In both cases it is obvious that $\|\mathbf{1}\| = 1$.

(iv) Since $c_0(S,\mathbb{F}) \subseteq \ell^\infty(S,\mathbb{F})$ is a closed linear subspace, it remains to show that $f \in c_0(S,\mathbb{F}), g \in \ell^\infty(S,\mathbb{F})$ implies $fg, gf \in c_0(S,\mathbb{F})$. This is obvious. ■

4.23 REMARK 1. For $1 < p < \infty$ there is no natural way of defining a bilinear product on $\ell^p(S,\mathbb{F})$ turning it into a Banach algebra.

2. If G is a discrete group, the definition of $\ell^1(\mathbb{Z},\mathbb{F})$ is easily adapted to $\ell^1(G,\mathbb{F})$ by putting $(f \star g)(k) = \sum_{l\in G} f(l)g(l^{-1}k)$. This again gives rise to a Banach algebra, but it is commutative if and only if G is abelian. □

4.7 Outlook on general L^p-spaces

For an arbitrary measure space $(X, \mathcal{A}, \mu)$ one can define normed spaces $L^p(X, \mathcal{A}, \mu; \mathbb{F})$ in a broadly analogous fashion. (We will usually omit the $\mathbb{F}$.) Since integration on measure spaces is not among the formal prerequisites of these notes, we only sketch the basic facts referring to, e.g., [29, 146] for details. If $f : X \to \mathbb{F}$ is a measurable function and $0 < p < \infty$, then $\|f\|_p = (\int |f(x)|^p d\mu(x))^{1/p} \in [0, \infty]$. If $p = \infty$, put[26]

$$\|f\|_\infty = \operatorname{ess\,sup}_\mu |f| = \inf\{\lambda > 0 \mid \mu(\{x \in X \mid |f(x)| > \lambda\}) = 0\}.$$

Now $\mathcal{L}^p(X, \mu) = \{f : X \to \mathbb{F} \text{ measurable} \mid \|f\|_p < \infty\}$ is an $\mathbb{F}$-vector space for all $p \in (0, \infty]$. For $1 \leq p \leq \infty$, the proofs of the inequalities of Hölder and Minkowski extend to the present setting without any difficulties, so that the $\|\cdot\|_p$ are seminorms on $\mathcal{L}^p(X, \mathcal{A}, \mu)$. But the latter fails to be a norm whenever there exists $\emptyset \neq Y \in \mathcal{A}$ with $\mu(Y) = 0$ since then $\|\chi_Y\|_p = 0$.

[26] Warning: [29] defines $\|\cdot\|_\infty$ using locally null sets instead of null sets, which is very non-standard.

For this reason we define $L^p(X,\mathcal{A},\mu) = \mathcal{L}^p(X,\mathcal{A},\mu)/\{f \mid \|f\|_p = 0\}$. Now it is straightforward to prove that $L^p(X,\mu) = \mathcal{L}^p(X,\mu)/\sim$ is a normed space, and in fact complete. The proof now uses Proposition 3.15. If S is a set and μ is the counting measure, we have $\ell^p(S,\mathbb{F}) = \mathcal{L}^p(S,P(S),\mu,\mathbb{F}) = L^p(S,P(S),\mu,\mathbb{F})$.

A measurable function is called simple if it assumes only finitely many values. Equivalently it is of the form $f(x) = \sum_{k=1}^K c_k \chi_{A_k}(x)$, where $A_1,\ldots,A_k$ are measurable sets. Now one proves that the simple functions are dense in L^p for all $p \in [1,\infty]$. If X is locally compact and μ is nice enough, the set $C_c(X,\mathbb{F})$ of compactly supported continuous functions is dense in $L^p(X,\mathcal{A},\mu;\mathbb{F})$ for $1 \le p < \infty$, while its closure in L^∞ is $C_0(X,\mathbb{F})$.

The inclusion $\ell^p \subseteq \ell^q$ for $p \le q$ (Lemma 4.10) is false for general measure spaces! In fact, if $\mu(X) < \infty$ then one has the reverse inclusion $p \le q \Rightarrow L^q(X,\mathcal{A},\mu) \subseteq L^p(X,\mathcal{A},\mu)$, while for general measure spaces there is no inclusion relation between the L^p with different p.

If $1 < p,q < \infty$ are conjugate, the canonical map $\varphi : L^q(X,\mathcal{A},\mu) \to L^p(X,\mathcal{A},\mu)^*$ is an isometric bijection for all measure spaces. That φ is an isometry is proven just as for the spaces ℓ^p: Hölder's inequality gives $\|\varphi_g\| \le \|g\|_q$, and equality is proven as in Theorem 4.19(ii) by showing $|\varphi_g(f)| \ge \|f\|_p\|g\|_1$, where the $f \in L^p$ are the same as before. However, isometry of $L^\infty(X,\mathcal{A},\mu) \to (L^1(X,\mathcal{A},\mu))^*$ is not automatic, as the measure space $X = \{x\}$, $\mathcal{A} = P(X) = \{\emptyset, X\}$ and $\mu : \emptyset \mapsto 0,\ X \mapsto +\infty$ shows, for which $L^1(X,\mathcal{A},\mu,\mathbb{F}) \cong \{0\}$ and $\mathbb{F} \cong L^\infty(X,\mathcal{A},\mu,\mathbb{F}) \not\cong L^1(X,\mathcal{A},\mu,\mathbb{F})^*$. It is not hard to show that $L^\infty \to (L^1)^*$ is isometric if and only if $(X,\mathcal{A},\mu)$ is semifinite, i.e.

$$\mu(Y) = \sup\{\mu(Z) \mid Z \in \mathcal{A},\ Z \subseteq Y,\ \mu(Z) < \infty\} \quad \forall Y \in \mathcal{A}.$$

If $1 < p < \infty$, one still has surjectivity of $L^p \to (L^q)^*$ for all measure spaces $(X,\mathcal{A},\mu)$, but the standard proof is outside our scope since it requires the Radon-Nikodym theorem. (For a more functional-analytic proof see Section B.6.8.) In order for $L^\infty \to (L^1)^*$ to be an isometric bijection, the measure space must be 'localizable', cf. [146]. This condition subsumes semifiniteness and is implied by σ-finiteness, to which case many books limit themselves.

Since we relegated the dual spaces $\ell^\infty(S,\mathbb{F})^*$ to an appendix, we only remark that also in general $L^\infty(X,\mathcal{A},\mu)^*$ is a space of finitely additive measures with fairly similar proofs, see [43]. For $0 < p < 1$, the dual spaces $(L^p)^*$ behave even stranger than $(\ell^p)^*$. For example, $L^p([0,1],\lambda;\mathbb{R})^* = \{0\}$.

5 Basics of Hilbert spaces

5.1 Inner products. Cauchy-Schwarz inequality

We have seen that every bounded linear functional φ on $\ell^p(S,\mathbb{F})$, where $1 \le p < \infty$ is of the form $\varphi_g : f \mapsto \sum_{s\in S} f(s)g(s)$ for a certain unique $g \in \ell^q(S,\mathbb{F})$. Here the conjugate exponent $q \in (1,\infty]$ is determined by $\frac{1}{p} + \frac{1}{q} = 1$. Clearly we have $p = q$ if and only if $p = 2$. In this space we have self-duality: $\ell^2(S,\mathbb{F})^* \cong \ell^2(S,\mathbb{F})$. The map

$$\ell^2(S,\mathbb{F}) \times \ell^2(S,\mathbb{F}) \to \mathbb{F},\ (f,g) \mapsto \sum_{s\in S} f(s)g(s)$$

is bilinear and symmetric. Furthermore, it satisfies $|\sum_{s\in S} f(s)g(s)| \le \|f\|_2\|g\|_2$. Defining $\overline{g}(s) = \overline{g(s)}$, we have $\|\overline{g}\| = \|g\|$, so that also $|\sum_{s\in S} f(s)\overline{g(s)}| \le \|f\|_2\|g\|_2$, which is the Cauchy-Schwarz inequality (in its incarnation for $\ell^2(S,\mathbb{C})$).

For the development of a general, abstract theory it is better to adopt a slightly different definition:

5.1 Definition *Let V be an $\mathbb{F}$-vector space. An inner product on V is a map $V \times V \to \mathbb{F}, (x,y) \mapsto \langle x,y\rangle$ such that*

(i) The map $x \mapsto \langle x,y\rangle$ is linear for each choice of $y \in V$.

(ii) $\langle y,x\rangle = \overline{\langle x,y\rangle}\ \forall x,y \in V$.

(iii) $\langle x,x\rangle \geq 0\ \forall x$, and $\langle x,x\rangle = 0 \ \Rightarrow\ x = 0$.

5.2 Remark 1. Many authors write (x,y) instead of $\langle x,y\rangle$, but this often leads to confusion with the notation for ordered pairs. We will use pointed brackets throughout.

2. Combining the first two axioms one finds that the map $y \mapsto \langle x,y\rangle$ is anti-linear for each choice of x. This means $\langle x, cy + c'y'\rangle = \overline{c}\langle x,y\rangle + \overline{c'}\langle x,y'\rangle$ for all $y,y' \in V$ and $c,c' \in \mathbb{F}$. Of course this reduces to linearity if $\mathbb{F} = \mathbb{R}$. A map $V \times V \to \mathbb{C}$ that is linear in the first variable and anti-linear in the second is called sesquilinear.

3. A large minority of authors, mostly (mathematical) physicists, defines inner products to be linear in the second and anti-linear in the first argument. We follow the majority practice.

4. If $\mathbb{F} = \mathbb{R}$ then $\langle y,x\rangle = \overline{\langle x,y\rangle} = \langle x,y\rangle\ \forall x,y$. Thus $\langle \cdot,\cdot\rangle$ is bilinear and symmetric.

5. The first two axioms together already imply $\langle x,x\rangle \in \mathbb{R}$ for all x, but not the positivity assumption.

6. If $\langle x,y\rangle = 0$ for all $y \in H$ then $x = 0$. To see this, it suffices to take $y = x$. □

5.3 Example 1. If $V = \mathbb{C}^n$ then $\langle x,y\rangle = \sum_{i=1}^n x_i\overline{y_i}$ is an inner product and the corresponding norm (see below) is $\|\cdot\|_2$, which is complete.

2. Let S be any set and $V = \ell^2(S,\mathbb{C})$. Then $\langle f,g\rangle = \sum_{s\in S} f(s)\overline{g(s)}$ converges for all $f,g \in V$ by Hölder's inequality and is easily seen to be an inner product. Of course, 1. is a special case of 2.

3. If $(X,\mathcal{A},\mu)$ is any measure space then $\langle f,g\rangle = \int_X f(x)\overline{g(x)}\,d\mu(x)$ is an inner product on $L^2(X,\mathcal{A},\mu;\mathcal{F})$ turning it into a Hilbert space. (Here we allow ourselves a standard sloppiness: The elements of L^p are not functions, but equivalence classes of functions. The inner product of two such classes is defined by picking arbitrary representers.)

4. Let $V = M_{n\times n}(\mathbb{C})$. For $a,b \in V$, define $\langle a,b\rangle = \mathrm{Tr}(b^*a) = \sum_{i,j=1}^n a_{ij}\overline{b_{ij}}$, where $(b^*)_{ij} = \overline{b_{ji}}$. That this is an inner product turning V into a Hilbert space follows from 1. upon the identification $M_{n\times n}(\mathbb{C}) \cong \mathbb{C}^{n^2}$.

In view of $\langle x,x\rangle \geq 0$ for all x, and we agree that $\langle x,x\rangle^{1/2}$ always is the positive root.

5.4 Lemma (Abstract Cauchy-Schwarz inequality) [27] *If $\langle \cdot,\cdot\rangle$ is an inner product on V then*

$$|\langle x,y\rangle| \leq \langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2} \quad \forall x,y \in V. \tag{5.1}$$

(This even holds if one drops the assumption that $\langle x,x\rangle = 0 \Rightarrow x = 0$.)

Equality holds in (5.1) if and only if one of the vectors is zero or $x = cy$ for some $c \in \mathbb{F}$.

5.5 Exercise Prove Lemma 5.4 along the following lines:

1. Prove it for $y = 0$, so that we may assume $y \neq 0$ from now on.
2. Define $x_1 = \|y\|^{-2}\langle x,y\rangle y$ and $x_2 = x - x_1$ and prove $\langle x_1,x_2\rangle = 0$.

[27] Augustin-Louis Cauchy (1789-1857). French mathematician with many important contributions to analysis. Karl Hermann Armandus Schwarz (1843-1921). German mathematician, mostly active in complex analysis. Some authors (mostly Russian ones) include Viktor Yakovlevich Bunyakovski (1804-1889). Russian mathematician.

3. Use 2. to prove $\|x\|^2 = \|x_1\|^2 + \|x_2\|^2 \geq \|x_1\|^2$.
4. Deduce Cauchy-Schwarz from $\|x_1\|^2 \leq \|x\|^2$.
5. Prove the claim about equality.

The above proof is the easiest to memorize (at least in outline) and reconstruct, but there are many others, e.g.:

5.6 Exercise Let V be a vector space with inner product $\langle\cdot,\cdot\rangle$ and define $\|x\| = \langle x,x\rangle^{1/2}$.

(i) For $x, y \in V$ and $t \in \mathbb{R}$, define $P(t) = \|x + ty\|^2$ and show this defines a quadratic polynomial in t with real coefficients.

(ii) Use the obvious fact that this polynomial takes values in $[0, \infty)$ for all $t \in \mathbb{R}$, thus also $\inf_{t\in\mathbb{R}} P(t) \geq 0$, to prove the Cauchy-Schwarz inequality.

5.7 Proposition *If $\langle\cdot,\cdot\rangle$ is an inner product on V then $\|x\| = +\sqrt{\langle x,x\rangle}$ is a norm on V. (An inner product $\langle\cdot,\cdot\rangle$ and a norm $\|\cdot\|$ related in this way are called compatible.)*

Proof. $\|x\| \geq 0$ holds by construction, and the third axiom in Definition 5.1 implies $\|x\| = 0 \Rightarrow x = 0$. We have

$$\|cx\| = \sqrt{\langle cx, cx\rangle} = \sqrt{c\overline{c}\langle x,x\rangle} = \sqrt{|c|^2\langle x,x\rangle} = |c|\|x\|,$$

thus $\|cx\| = |c|\|x\|$ for all $x \in V, c \in \mathbb{F}$. Finally,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle = \|x\|^2 + \|y\|^2 + \langle x,y\rangle + \langle y,x\rangle.$$

With $\operatorname{Re} z \leq |z|$ for all $z \in \mathbb{C}$ and the Cauchy-Schwarz inequality we have

$$\langle x,y\rangle + \langle y,x\rangle = \langle x,y\rangle + \overline{\langle x,y\rangle} = 2\operatorname{Re}\langle x,y\rangle \leq 2|\langle x,y\rangle| \leq 2\|x\|\|y\|,$$

thus

$$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$$

and therefore $\|x+y\| \leq \|x\| + \|y\|$, i.e. subadditivity. ■

In terms of the norm, the Cauchy-Schwarz inequality just becomes $|\langle x,y\rangle| \leq \|x\|\|y\|$.

5.8 Definition *A pre-Hilbert space (or inner product space) is a pair $(V, \langle\cdot,\cdot\rangle)$, where V is an $\mathbb{F}$-vector space and $\langle\cdot,\cdot\rangle$ an inner product on it. A Hilbert space is a pre-Hilbert space that is complete for the norm $\|\cdot\|$ obtained from the inner product.*

5.9 Remark 1. By the above, an inner product gives rise to a norm and therefore to a norm topology τ. Now the Cauchy-Schwarz inequality implies that the inner product $\langle\cdot,\cdot\rangle \to \mathbb{F}$ is jointly continuous:

$$\begin{aligned} |\langle x,y\rangle - \langle x',y'\rangle| &= |\langle x,y\rangle - \langle x,y'\rangle + \langle x,y'\rangle - \langle x',y'\rangle| \\ &= |\langle x, y-y'\rangle + \langle x-x', y'\rangle| \\ &\leq \|x\|\|y-y'\| + \|x-x'\|\|y'\|. \end{aligned}$$

2. If $\langle\cdot,\cdot\rangle$ be an inner product on H and $\|\cdot\|$ is the norm derived from it then

$$\|x\| = \sup_{y\in H, \|y\|=1} |\langle x,y\rangle| \quad \forall x \in H. \tag{5.2}$$

(For $x=0$ this is obvious, and for $x \neq 0$ it follows from $\langle x, \frac{x}{\|x\|}\rangle = \|x\|$.)

3. The restriction of an inner product on H to a linear subspace $K \subseteq H$ again is an inner product. Thus if H is a Hilbert space and K a closed subspace then K again is a Hilbert space (with the restricted inner product).

4. All spaces considered in Example 5.3 are complete, thus Hilbert spaces. For $\ell^2(S)$ this was proven in Section 4., and the claim for $\mathbb{C}^n$, thus also $M_{n\times n}(\mathbb{C})$, follows since $\mathbb{C}^n \cong \ell^2(S,\mathbb{C})$ when $\#S = n$. For $L^2(X,\mathcal{A},\mu)$ see books on measure theory like [29, 146]. □

5.10 Definition *Let $(H_1, \langle\cdot,\cdot\rangle_1), (H_2, \langle\cdot,\cdot\rangle_2)$ be pre-Hilbert spaces. A linear map $A : H_1 \to H_2$ is called*

- *isometric or an isometry if $\langle Ax, Ay\rangle_2 = \langle x, y\rangle_1 \quad \forall x, y \in H_1$.*
- *unitary if it is a surjective isometry.*

5.11 Remark Every unitary map is invertible with unitary inverse. Two Hilbert spaces H_1, H_2 are called unitarily equivalent or isomorphic if there exists a unitary $U : H_1 \to H_2$. □

If $(H_1, \langle\cdot,\cdot\rangle_1), (H_2, \langle\cdot,\cdot\rangle_2)$ are (pre-)Hilbert spaces then

$$\langle (x_1, x_2), (y_1, y_2)\rangle = \langle x_1, y_1\rangle_1 + \langle x_2, y_2\rangle_2$$

defines an inner product on $H_1 \oplus H_2$ turning it into a (pre-)Hilbert space. With this definition, $\|(x,y)\| = \langle (x,y),(x,y)\rangle^{1/2} = (\langle x,x\rangle_1 + \langle y,y\rangle_2)^{1/2} = (\|x\|_1^2 + \|y\|_2^2)^{1/2}$ (thus not $\|x\|_1 + \|y\|_2$!).

More generally, if $\{H_i, \langle\cdot,\cdot\rangle_i\}_{i\in I}$ is a family of (pre-)Hilbert spaces then

$$\bigoplus_{i\in I} H_i = \{\{x_i\}_{i\in I} \mid \sum_{i\in I} \langle x_i, x_i\rangle_i < \infty\}$$

with

$$\langle \{x_i\}, \{y_i\}\rangle = \sum_{i\in I} \langle x_i, y_i\rangle_i$$

is a (pre-)Hilbert space. (If $H_i = \mathbb{F}$ for all $i \in I$, this construction recovers $\ell^2(I,\mathbb{F})$, while the Banach space direct sum gives $\ell^1(I,\mathbb{F})$.)

5.12 Exercise Let $(V, \langle\cdot,\cdot\rangle)$ be a pre-Hilbert space and $\|\cdot\|$ the associated norm. Let $(V', \|\cdot\|')$ be the completion (as a normed space) of $(V, \|\cdot\|)$. Prove that V' is a Hilbert space.

5.2 The parallelogram and polarization identities

Given a normed space $(V, \|\cdot\|)$, it is natural to ask whether there exists an inner product $\langle\cdot,\cdot\rangle$ on V compatible with $\|\cdot\|$ (in the sense $\langle x, x\rangle = \|x\|^2 \ \forall x$).

5.13 Exercise Let $(V, \langle\cdot,\cdot\rangle)$ be a pre-Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Prove the parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall x, y \in V \tag{5.3}$$

and the polarization identities

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) \quad \text{if } \mathbb{F} = \mathbb{R}, \tag{5.4}$$

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2 \qquad \text{if } \mathbb{F} = \mathbb{C}. \tag{5.5}$$

For a map of (pre-)Hilbert spaces we have two a priori different notions of isometry, but they are equivalent:

5.14 Exercise Let $(H_1, \langle\cdot,\cdot\rangle_1), (H_2, \langle\cdot,\cdot\rangle_2)$ be (pre-)Hilbert spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let $\|\cdot\|_{1,2}$ be the norms induced by the inner products. Prove that a linear map $A : H_1 \to H_2$ is an isometry of normed spaces (i.e. $\|Ax\|_2 = \|x\|_1 \ \forall x \in H_1$) if and only if it is an isometry of pre-Hilbert spaces (i.e. $\langle Ax, Ay\rangle_2 = \langle x, y\rangle_1 \ \forall x, y \in H_1$).

5.15 Exercise Let S be a set with $\#S \geq 2$ and $p \in [1, \infty]$. Prove that the norm $\|\cdot\|_p$ on $\ell^p(S, \mathbb{F})$ satisfies the parallelogram identity if and only if $p = 2$.

5.16 Exercise (Jordan-von Neumann 1935) [28] Let $(V, \|\cdot\|)$ be a normed space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ satisfying (5.3). Define $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ by (5.4).

(i) Prove $\langle x, y\rangle = \langle y, x\rangle$ and $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$.

(ii) Prove $\langle tx, y\rangle = t\langle x, y\rangle$ for all $t \in \mathbb{N}$, then successively for $t \in \mathbb{Z}, \mathbb{Q}, \mathbb{R}$.

(iii) Prove that $\langle\cdot,\cdot\rangle$ is compatible with $\|\cdot\|$ and makes V a real inner product space.

(iv) If $\mathbb{F} = \mathbb{C}$, prove that (5.5) defines an inner product (Definition 5.1) compatible with $\|\cdot\|$. Hint: Prove and use a relationship between the right hand sides of (5.4) and (5.5).

Exercise 5.16 shows that (5.3) characterizes the Banach spaces that 'are' Hilbert spaces in the sense of admitting an inner product compatible with the norm. There are very many such criteria: The whole book [2] is dedicated to proving about 350 of them! To state just one more: A Banach space V is a Hilbert space (in the above sense) if and only if for every 2-dimensional subspace W there exists an idempotent $P = P^2 \in B(V)$ with $\|P\| = 1$ such that $PV = W$.

5.17 Exercise (i) If H is a Hilbert space and $x_1, \ldots, x_n \in H$, prove [this is easier without induction!] the generalized parallelogram identity

$$2^{-n} \sum_{s \in \{\pm 1\}^n} \left\| \sum_{i=1}^n s_i x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2. \tag{5.6}$$

(ii) Prove that a Banach space $(V, \|\cdot\|)$ is isomorphic to a Hilbert space (not necessarily isometrically) if and only if there is an inner product $\langle\cdot,\cdot\rangle$ on V such that the norm $\|x\|' = \langle x, x\rangle^{1/2}$ is equivalent to $\|\cdot\|$.

(iii) If V is a Banach space isomorphic to a Hilbert space, prove that there are $C' \geq C > 0$ such that

$$C \sum_{i=1}^n \|x_i\|^2 \leq 2^{-n} \sum_{s \in \{\pm 1\}^n} \left\| \sum_{i=1}^n s_i x_i \right\|^2 \leq C' \sum_{i=1}^n \|x_i\|^2 \tag{5.7}$$

for all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in V$.

5.18 Remark ⋆ A Banach space V in which the second inequality in (5.7) holds for some $C' > 0$, all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in V$ is said to have 'type 2'. It has 'cotype 2' if the analogous statement holds for the first inequality. (Type p and cotype p are defined analogously by

[28] Pascual Jordan (1902-1980). German theoretical physicist with contributions to quantum theory. (Not to be confused with the French mathematician Camille Jordan (1838-1922) to whom e.g. the Jordan normal form is due.)

replacing all $\|\cdot\|^2$ by $\|\cdot\|^p$.) By a remarkable theorem of Kwapień[29] (1972), every Banach space of type and cotype 2 (i.e. satisfying (5.7) with fixed $C, C' > 0$ for all n, x_i) is isomorphic to a Hilbert space! For a proof see e.g. [1, Theorem 7.4.1], [97, Vol. 1, Theorem 5.V.6].

Granting this, we have a criterion for a given Banach space $(V, \|\cdot\|)$ to be isomorphic (not necessarily isometrically) to a Hilbert space. We will encounter two more, cf. Remark 6.10.4 and Remark 9.13, but proving the converse directions is way too involved for these notes. □

5.3 Basic Hilbert space geometry

5.19 Definition *If H is a (pre-)Hilbert space then $x, y \in H$ are called orthogonal, denoted $x \perp y$, if $\langle x, y\rangle = 0$. If $S, T \subseteq H$ then $S \perp T$ means $x \perp y$ $\forall x \in S, y \in T$.*

It should be obvious that $x \perp y$ implies $cx \perp dy$ for all $c, d \in \mathbb{F}$.

5.20 Lemma (Pythagoras' theorem) *Let H be a (pre-)Hilbert space and $x_1, \ldots, x_n \in H$ mutually orthogonal, i.e. $i \neq j \Rightarrow \langle x_i, x_j\rangle = 0$. Let $x = x_1 + \cdots + x_n$. Then*

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

Proof. We have

$$\|x\|^2 = \langle x, x\rangle = \Big\langle \sum_i x_i, \sum_j x_j \Big\rangle = \sum_{i,j} \langle x_i, x_j\rangle = \sum_i \langle x_i, x_i\rangle = \sum_i \|x_i\|^2,$$

where we used $i \neq j \Rightarrow \langle x_i, x_j\rangle = 0$. ■

5.21 Remark If H is a Hilbert space, I is an infinite set and $\{x_i\}_{i\in I} \subseteq H$ is such that $\sum_i \|x_i\| < \infty$ then we can make sense of $x = \sum_{i\in I} x_i \in H$ by completeness and Proposition 3.15. If all x_i are mutually orthogonal then by taking the limit over finite subsets we again have $\|x\|^2 = \sum_{i\in I} \|x_i\|^2$. (This shows that $\sum_i \|x_i\| < \infty \Rightarrow \sum_i \|x_i\|^2 < \infty$, which also follows from the inclusion $\ell^1(S) \subseteq \ell^2(S)$ proven in Lemma 4.10.) □

5.22 Definition *Let V be an $\mathbb{F}$-vector space. Then $C \subseteq V$ is called convex if for all $x, y \in C$ and $t \in [0,1]$ we have $tx + (1-t)y \in C$. (Equivalently $tC + (1-t)C \subseteq C$ for all $t \in [0,1]$.)*

5.23 Exercise If V is an $\mathbb{F}$-vector space and $C \subseteq V$ is convex, prove that $\sum_{i=1}^N t_i x_i \in C$ whenever $x_1, \ldots, x_N \in C$ and $t_1, \ldots, t_N \geq 0$ satisfy $\sum_i t_i = 1$.

5.24 Proposition (Riesz lemma) [30] *Let H be a Hilbert space and $C \subseteq H$ a non-empty closed convex set. Then for each $x \in H$ there is a unique $y \in C$ minimizing $\|x - y\|$, i.e. $\|x - y\| = \inf_{z\in C} \|x - z\|$.*

Proof. We will prove this for $x = 0$, in which case the statement says that there is a unique element of C of minimal norm. For general $x \in H'$, let y' be the unique element of minimal norm in the convex set $C' = C - x$. Then $y = y' + x$ is the unique element in C minimizing $\|x - y\|$.

[29] Stanisław Kwapień (b. 1942), Polish mathematician, mostly in functional analysis and probability. The proof of his theorem indeed uses some probability theory.

[30] Frigyes Riesz (1880-1956). Hungarian mathematician and one of the pioneers of functional analysis. (The same applies to his younger brother Marcel Riesz (1886-1969).)

Let $d = \inf_{z \in C} \|z\|$ and pick a sequence $\{y_n\}$ in C such that $\|y_n\| \to d$. Since C is convex, we have $\frac{y_n+y_m}{2} \in C$, thus $\left\|\frac{y_n+y_m}{2}\right\| \geq d$ for all n, m. By the choice of $\{y_n\}$ we have $\|y_n\| \to d$. Thus for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $n \geq N$ implies $\|y_n\| < d + \varepsilon$. Thus if $n, m \geq N$, then with the parallelogram identity (5.3) we have

$$\begin{aligned} \|y_n - y_m\|^2 &= 2\|y_n\|^2 + 2\|y_m\|^2 - \|y_n + y_m\|^2 = 2\|y_n\|^2 + 2\|y_m\|^2 - 4\left\|\frac{y_n + y_m}{2}\right\|^2 \\ &< 4(d+\varepsilon)^2 - 4d^2 = 8d\varepsilon + 4\varepsilon^2. \end{aligned}$$

This implies that $\{y_n\}$ is a Cauchy sequence and therefore converges to some $y \in H$ by completeness of H, and closedness of C gives $y \in C$. By continuity of the norm, we have $\|y\| = \|\lim y_n\| = \lim \|y_n\| = d$.

If $y, y' \in C$ with $\|y\| = \|y'\| = d$ then $\frac{y+y'}{2} \in C$ by convexity, thus $\|(y+y')/2\|^2 \geq d^2$ by the definition of d. Now the parallelogram identity gives $0 \leq \|y - y'\|^2 = 4d^2 - 4\left\|\frac{y+y'}{2}\right\|^2 \leq 0$. Thus $\|y - y'\| = 0$, proving $y = y'$. ■

5.4 Closed subspaces, orthogonal complement, and orthogonal projections

5.25 Definition *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space and $S \subseteq H$. The orthogonal complement $S^\perp$ is defined as*

$$S^\perp = \{y \in H \mid \langle y, x\rangle = 0 \quad \forall x \in S\}.$$

5.26 Exercise Let H be a Hilbert space over $\mathbb{F}$ and $S, T \subseteq H$ arbitrary subsets. Prove:

(i) $T \subseteq S^\perp \Leftrightarrow S \perp T \Leftrightarrow S \subseteq T^\perp$.

(ii) $S^\perp \subseteq H$ is a closed linear subspace.

(iii) $S^\perp = \overline{\text{span}_{\mathbb{F}}(S)}^\perp$.

(iv) If $S \subseteq T$ then $T^\perp \subseteq S^\perp$.

(v) $S \subseteq S^{\perp\perp}$ and $S^\perp = S^{\perp\perp\perp}$.

A linear subspace of a vector space clearly is a convex subset. Now,

5.27 Theorem *Let H be a Hilbert space and $K \subseteq H$ a closed linear subspace. Define a map $P : H \to K$ by $Px = y$, where $y \in K$ minimizes $\|x - y\|$ as in Proposition 5.24. Also define $Qx = x - Px$. Then*

(i) $Qx \in K^\perp$ $\forall x$.

(ii) For each $x \in H$ there are unique $y \in K, z \in K^\perp$ with $x = y + z$, namely $y = Px, z = Qx$.

(iii) The maps P, Q are linear.

(iv) The map $P : H \to H$ satisfies $P^2 = P$ and $\langle Px, y\rangle = \langle x, Py\rangle$. The same holds for Q.

(v) The map $U : H \to K \oplus K^\perp$, $x \mapsto (Px, Qx)$ is an isomorphism of Hilbert spaces. In particular, $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$ $\forall x$.

Proof. (i) Let $x \in H, v \in K$. We want to prove $Qx \perp v$, i.e. $\langle x - Px, v\rangle = 0$. Since $y = Px$ is the element of K minimizing $\|x - y\|$, we have for all $t \in \mathbb{C}$

$$\|x - Px\| \leq \|x - Px - tv\|.$$

Taking squares and putting $z = x - y = x - Px$, this becomes $\langle z, z\rangle \le \langle z - tv, z - tv\rangle$, equivalent to

$$2\,\mathrm{Re}(t\langle v, z\rangle) \le |t|^2\|v\|^2.$$

With the polar decomposition $t = |t|e^{i\varphi}$, this inequality becomes $2\,\mathrm{Re}(e^{i\varphi}\langle v, z\rangle) \le |t|\|v\|^2$. Taking $|t| \to 0$, we find $\mathrm{Re}(e^{i\varphi}\langle v, z\rangle) = 0$, and since φ was arbitrary, we conclude $\langle v, z\rangle = 0$. In view of $z = x - y = x - Px$ this is what we wanted.

(ii) For each $x \in H$ we have $x = Px + Qx$ with $Px \in K$, $Qx \in K^\perp$, proving the existence. If $y, y' \in K, z, z' \in K^\perp$ such that $y + z = y' + z'$ then $y - y' = z' - z \in K \cap K^\perp = \{0\}$. Thus $y - y' = z' - z = 0$, proving the uniqueness.

(iii) If $x, x' \in H, c, c' \in \mathbb{F}$ then $cx + c'x' = P(cx + c'x') + Q(cx + c'x')$. But also

$$cx + c'x' = c(Px + Qx) + c'(Px' + Qx') = \underbrace{(cPx + c'Px')}_{\in K} + \underbrace{(cQx + c'Qy')}_{\in K^\perp}$$

is a decomposition of $cx + c'x'$ as a sum of vectors in K and $K^\perp$, respectively. Since such a decomposition is unique by (ii), we have $P(cx + c'x') = cPx + c'Px'$ and $Q(cx + c'x') = cQx + c'Qx'$, which just is the linearity of P and Q.

(iv) The definition of P clearly implies $Px = x$ if $x \in K$ (thus $P \restriction K = \mathrm{id}_K$). With $PH \subseteq K$ we have $P^2x = P(Px) = Px$ for all x, thus $P^2 = P$. And for all $x, y \in H$ we have

$$\langle Px, y\rangle = \langle Px, Py + Qy\rangle = \langle Px, Py\rangle = \langle Px + Qx, Py\rangle = \langle x, Py\rangle,$$

where we used the orthogonality of the images of P and Q. The proofs for Q are analogous.

(v) It is clear that U is a linear isomorphism. Furthermore, $Px \perp Qy$ implies

$$\langle x, y\rangle = \langle Px + Qx, Py + Qy\rangle = \langle Px, Py\rangle + \langle Qx, Qy\rangle = \langle Ux, Uy\rangle,$$

so that U is an isometry of Hilbert spaces. ■

5.28 Remark The above theorem remains valid if H is only a pre-Hilbert space, provided $K \subseteq H$ is finite-dimensional. We first note that the proof of Lemma 5.24 only uses completeness of $C \subseteq H$, not that of H. And we recall that finite-dimensional subspaces of normed spaces are automatically complete and closed, cf. Exercises 2.33 and 3.22. In the proof of Theorem 5.27 we use Lemma 5.24 with $C = K$, which is complete as just noted. □

5.29 Exercise Let H be a Hilbert space and $V \subseteq H$ a linear subspace. Prove:

(i) $V^\perp = \{0\}$ if and only if $\overline{V} = H$.

(ii) $V^{\perp\perp} = \overline{V}$.

5.30 Definition *Let V be a vector space and H a (pre-)Hilbert space.*

- *A linear map $P : V \to V$ is idempotent if $P^2 \equiv P \circ P = P$.*
- *A linear map $P : H \to H$ is self-adjoint[31] if $\langle Px, y\rangle = \langle x, Py\rangle$ for all $x, y \in H$.*
- *A bounded linear map $P : H \to H$ is an orthogonal projection if it is a self-adjoint idempotent.*

[31]Many authors write ‘hermitian’ instead of ‘self-adjoint’. We stick to the latter.

(In Theorem 7.32 we will prove that every self-adjoint $P : H \to H$ is automatically bounded, but this is not needed here. There are unbounded idempotents.)

We have seen that every closed subspace K of a Hilbert space gives rise to an orthogonal projection P with $PH = K$. Conversely, we have:

5.31 Exercise Let H be a Hilbert space and $P \in B(H)$ idempotent. Prove:

(i) $K = PH \subseteq H$ and $L = (\mathbf{1} - P)H$ are closed linear subspaces.

(ii) We have $K \perp L$ if and only if P is self-adjoint, i.e. an orthogonal projection.

(iii) If P is an orthogonal projection then it equals the P associated to K by Theorem 5.27.

In linear algebra, one has a notion of quotient spaces, cf. e.g. [55, Exercise 31 in Section 1.3]: If V is an $\mathbb{F}$-vector space and $W \subseteq V$ is a linear subspace, one defines an equivalence relation on V by $x \sim y \Leftrightarrow x - y \in W$ and then lets V/W denote the quotient space $V/\!\sim$, i.e. the set of $\sim$-equivalence classes. One shows that V/W again is a $\mathbb{F}$-vector space.

It is very natural to ask whether V/W again is a Hilbert (or Banach) space if that is the case for V. For Hilbert spaces this is quite easy:

5.32 Exercise Let H be a Hilbert space and $K \subseteq H$ a closed linear subspace. Prove that there is a linear isomorphism $H/K \to K^\perp$ of $\mathbb{F}$-vector spaces.

Conclude that the quotient space H/K of a Hilbert space H by a closed subspace K admits an inner product turning it into a Hilbert space.

5.33 Exercise Let H be a Hilbert space and K, L closed linear subspaces of H such that $\dim K < \infty$ and $\dim K < \dim L$. Prove that $L \cap K^\perp \neq \{0\}$.

5.5 The dual space H^* of a Hilbert space

If H is a Hilbert space, every $y \in H$ gives rise to a linear functional on H via $\varphi_y : x \mapsto \langle x, y\rangle$. The Cauchy-Schwarz inequality gives $|\varphi_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$, so that $\|\varphi_y\| \leq \|y\| < \infty$. For every $y \neq 0$ we have $\varphi_y(y) = \langle y, y\rangle = \|y\|^2 > 0$, implying $\|\varphi_y\| = \|y\|$. As a consequence, for every non-zero $x \in H$ there is a $\varphi \in H^*$ with $\varphi(x) \neq 0$. Thus H^* separates the points of H. But we have more:

5.34 Theorem ((F.) Riesz-Fréchet representation theorem) *If H is a Hilbert space and $\varphi \in H^*$ then there is a unique $y \in H$ such that $\varphi = \varphi_y = \langle \cdot, y\rangle$.*

Thus the map $H \to H^*$, $y \mapsto \varphi_y$ is an anti-linear isometric bijection.

Proof. If $\varphi = 0$, put $y = 0$ (every $y \neq 0$ gives $\varphi_y \neq 0$). Now assume $\varphi \neq 0$. Let $K = \ker \varphi = \varphi^{-1}(0)$. Then $K \subseteq H$ is a linear subspace that is closed (by continuity of φ) and proper (since $\varphi \neq 0$). By the preceding section, $K^\perp \neq \{0\}$. The dimension of $K^\perp$ is one. (Either by the algebraic fact that $\dim(H/K) = \operatorname{codim} K = 1$ or as follows: If $y_1, y_2 \in K^\perp \backslash \{0\}$ then $\varphi(y_1) \neq 0 \neq \varphi(y_2)$ implies $\varphi(y_1/\varphi(y_1) - y_2/\varphi(y_2)) = 0$ and therefore $y_1/\varphi(y_1) - y_2/\varphi(y_2) \in K \cap K^\perp = \{0\}$, thus y_1 and y_2 are linearly dependent.) Pick a non-zero $z \in K^\perp$ and put $y = \frac{\overline{\varphi(z)}}{\|z\|^2} z$. Then φ_y vanishes on K, and $\varphi_y(z) = \langle z, y\rangle = \left\langle z, \frac{\overline{\varphi(z)}}{\|z\|^2} z \right\rangle = \varphi(z)$. Thus $\varphi = \varphi_y$ on both K and $K^\perp = \mathbb{F}z$, thus on $H = K + K^\perp$. In view of the construction, y is unique. Since the inner product is anti-linear in the second variable, the final claim follows. ■

By the above, the map $H \to H^*$, $y \mapsto \varphi_y$ is an anti-linear isometric bijection of Banach spaces. If we want, we can use it to put an inner product on H^*.

5.35 Exercise Let H be a Hilbert space, $K \subseteq H$ a linear subspace (not necessarily closed!) and let $\varphi \in K^*$. Prove:

(i) There exists $\widehat{\varphi} \in H^*$ such that $\widehat{\varphi} \upharpoonright K = \varphi$.

(ii) Uniqueness of $\widehat{\varphi} \in H^*$ satisfying $\widehat{\varphi} \upharpoonright K = \varphi$ holds if and only $\overline{K} = H$.

(iii) There is a unique $\widehat{\varphi} \in H^*$ satisfying $\widehat{\varphi} \upharpoonright K = \varphi$ and $\|\widehat{\varphi}\| = \|\varphi\|$.

5.6 Orthonormal sets and bases

We begin by recalling the notion of bases from linear algebra: A finite subset $\{x_1, \ldots, x_n\}$ of a vector space V over the field k is called linearly independent if $\sum_{i=1}^n c_i x_i = 0$, where $c_1, \ldots, c_n \in k$, implies $c_1 = \cdots = c_n = 0$. (In particular, $x_i \neq 0\ \forall i$.) An arbitrary subset $B \subseteq V$ is called linearly independent if every finite subset $S \subseteq B$ is linearly independent. A linearly independent subset $B \subseteq V$ is called a (Hamel) basis if every $x \in V$ can be written as a linear combination of finitely many elements of B. This is equivalent to B being maximal, i.e. non-existence of linearly independent sets B' properly containing B. One now proves that any two bases of V have the same cardinality. All this is known from linear algebra, but the following possibly not:

5.36 Proposition *Every vector space V has a basis.*

Proof. If $V = \{0\}$, $\emptyset$ is a basis. Thus let V be non-zero and let $\mathcal{B}$ be the set of linearly independent subsets of V. The set $\mathcal{B}$ is partially ordered by inclusion $\subseteq$ and non-empty (since it contains $\{x\}$ for all $0 \neq x \in V$). We claim that every chain in (=totally ordered subset of) $(\mathcal{B}, \subseteq)$ has a maximal element: Just take the union $\widehat{B}$ of all sets in the chain. Since any finite subset of the union over a chain of sets is contained in some element of the chain, every finite subset of $\widehat{B}$ is linearly independent. Thus $\widehat{B}$ is in $\mathcal{B}$ and clearly is an upper bound of the chain. Thus the assumption of Zorn's Lemma is satisfied, so that $(\mathcal{B}, \subseteq)$ has a maximal element M. We claim that M is a basis for V: If this was false, we could find a $v \in V$ not contained in the span of M. But then $M \cup \{v\}$ would be a linearly independent set strictly larger than M, contradicting the maximality of M. ■

As soon as we study topological vector spaces, we can also talk about infinite linear combinations, which renders the linear algebra notion of basis quite irrelevant (except as a tool in some proofs). For Hilbert spaces we have the following natural notions:

5.37 Definition *Let H be a (pre-)Hilbert space. A subset $E \subseteq H$ is called*

- *orthogonal if $\langle x, y \rangle = 0$ whenever $x, y \in E,\ x \neq y$,*
- *orthonormal if it is orthogonal and every $x \in E$ is a unit vector, i.e. $\|x\| = 1$. (Equivalently, $\langle x, y \rangle = \delta_{x,y}\ \forall x, y \in E$.)*
- *orthonormal basis (ONB) if it is orthonormal and maximal. (I.e. there is no orthonormal set E' properly containing E.)*

5.38 Exercise Prove that every orthonormal set is linearly independent. What about orthogonal sets?

5.39 Lemma *Let H be a (pre-)Hilbert space and $E \subseteq H$ an orthonormal set. Then*

$$\sum_{e \in E} |\langle x, e \rangle|^2 \leq \|x\|^2 \quad \forall x \in H, \tag{5.8}$$

which is called the the inequality of Bessel.[32]

Proof. Let E be a finite orthonormal set and $x \in V$. Define $y = x - \sum_{e \in E} \langle x, e \rangle e$. It is straightforward to check that $\langle y, e \rangle = 0$ for all $e \in E$, so that $E \cup \{y\}$ is an orthogonal set. In view of $x = y + \sum_{e \in E} \langle x, e \rangle e$, Pythagoras' theorem (Lemma 5.20) gives

$$\|x\|^2 = \|y\|^2 + \sum_{e \in E} \|\langle x, e \rangle e\|^2 = \|y\|^2 + \sum_{e \in E} |\langle x, e \rangle|^2,$$

which in view of $\|y\|^2 \geq 0$ implies (5.8) for all finite E. If E is infinite, $\sum_{e \in E} |\langle x, e \rangle|^2$ equals the supremum of $\sum_{e \in F} |\langle x, e \rangle|^2 \leq \|x\|^2$ over the finite subsets $F \subseteq E$, which thus also satisfies (5.8). ■

5.40 Lemma *For every orthonormal set E in a (pre-)Hilbert space H there is an orthonormal basis $\widehat{E}$ containing E. In particular every Hilbert space admits an ONB.*

Proof. The proof is essentially the same as that of Proposition 5.36: Let $\mathcal{B}$ be the set of orthonormal sets that contain E, partially ordered by inclusion. Then $E \in \mathcal{B}$, thus $\mathcal{B} \neq \emptyset$. A Zorn's lemma argument gives the existence of maximal element $\widehat{E}$ of the partially ordered set $(\mathcal{B}, \subseteq)$. Thus $\widehat{E}$ is maximal among the orthonormal sets containing E. If there is a unit vector $f \in E^\perp$ then $\widehat{E} \cup \{f\}$ is an orthonomal set containing E strictly larger than $\widehat{E}$, contradicting the maximality of $\widehat{E}$. Thus no such f exists, and $\widehat{E}$ is an ONB for H. ■

The following exercise is motivated by a frequently made mistake. If H is a Hilbert space with ONB E and $A \in B(H)$ then $\sup_{e \in E} \|Ae\| \leq \|A\|$, but equality need not hold at all!!

5.41 Exercise If H is a Hilbert space, $E \subseteq H$ is an ONB and $A \in B(H)$, define $\|A\|_E = \sup_{e \in E} \|Ae\|$.

(i) Prove that $\| \cdot \|_E$ is a norm on $B(H)$ and that $\|A\|_E \leq \|A\| \; \forall A \in B(H)$.

(ii) Show that for every $N \in \mathbb{N}$ there are H, E, A as above with $\dim H = N$ such that $\|A\| \geq \sqrt{N} \|A\|_E$.
(Thus there is no $C < \infty$ such that $\|A\| \leq C \|A\|_E$ for all H, E, A, certainly not $C = 1$.)

5.42 Theorem [33] *Let H be a Hilbert space and E an orthonormal set in H. Then the following are equivalent:*

(i) E is an orthonormal basis, i.e. maximal.

(ii a) If $x \in H$ and $x \perp e$ for all $e \in E$ then $x = 0$.

(ii b) The map $H \to \ell^2(E, \mathbb{F}), x \mapsto \{\langle x, e \rangle\}_{e \in E}$ (well-defined thanks to (5.8)) is injective.

(iii) $\overline{\mathrm{span}_{\mathbb{F}} E} = H$.

(iv) For every $x \in H$, there are numbers $\{a_e\}_{e \in E}$ in $\mathbb{F}$ such that $x = \sum_{e \in E} a_e e$.

(v) For every $x \in H$, the equality $x = \sum_{e \in E} \langle x, e \rangle e$ holds.

(vi a) For every $x \in H$, we have $\|x\|^2 = \sum_{e \in E} |\langle x, e \rangle|^2$. (Abstract Parseval[34] identity)

[32] Friedrich Bessel (1784-1846), German mathematician, now best known for certain differential equations.

[33] I dislike the approach of some textbooks that restrict this statement to finite or countably infinite orthonormal sets, which amounts to assuming H to be separable. I also find it desirable to understand how much of the theorem survives without completeness since the latter does not hold in situations like Example 5.49. See Remark 5.43.

[34] Marc-Antoine Parseval (1755-1836). French mathematician.

(vi b) The map $H \to \ell^2(E, \mathbb{F}), x \mapsto \{\langle x, e\rangle\}_{e\in E}$ is an isometric map of normed spaces, where $\ell^2(S)$ has the $\|\cdot\|_2$-norm.

(vii a) For all $x, y \in H$ we have $\langle x, y\rangle = \sum_{e\in E}\langle x, e\rangle\overline{\langle y, e\rangle} = \sum_{e\in E}\langle x, e\rangle\langle e, y\rangle$.

(vii b) The map $H \to \ell^2(E, \mathbb{F}), x \mapsto \{\langle x, e\rangle\}_{e\in E}$ is an isometric map of pre-Hilbert spaces.

Here all summations over E are in the sense of the unordered summation of Appendix A.1 (with $V = H$ in (iv),(v) and $V = \mathbb{F}$ in (vi a),(vii a)).

Proof. If (ii a) holds then E is maximal, thus (i). If (ii a) is false then there is a non-zero $x \in H$ with $x \perp e$ for all $e \in E$. Then $E \cup \{x/\|x\|\}$ is an orthonormal set larger than E, thus E is not maximal. Thus (i)⇔(ii a). The equivalence (ii a)⇔(ii b) follows from the fact that a linear map is injective if and only if its kernel is $\{0\}$.

(iii)⇒(i) If $\overline{\mathrm{span}_{\mathbb{F}} E} = H$ and $x \in H$ satisfies $x \perp E$ then also $x \perp (\overline{\mathrm{span}_{\mathbb{F}} E} = H)$, thus $x = 0$. Thus E is maximal and therefore a basis.

(ii a)⇒(iii) $K = \overline{\mathrm{span}_{\mathbb{F}} E} \subseteq H$ is a closed linear subspace. If $K \neq H$ then by Theorem 5.27 we can find a non-zero $x \in K^\perp$. In particular $x \perp e\ \forall e \in E$, contradicting (ii a). Thus $K = H$.

It should be clear that the statements (vi b) and (vii b) are just high-brow versions of (vi a), (vii a), respectively, to which they are equivalent. That (vii a) implies (vi a) is seen by taking $x = y$. Since Exercise 5.14 gives (vi b)⇒(vii b), we have the mutual equivalence of (vi a), (vi b), (vii a), (vii b).

(v)⇒(iv) is trivial. If (iv) holds then continuity of the inner product, cf. Remark 5.9.1, implies $\langle x, y\rangle = \sum_{e\in E} a_e\langle e, y\rangle$ for all $y \in H$. For $y \in E$, the r.h.s. reduces to a_y, implying (v).

(iv) means that every $x \in H$ is a limit of finite linear combinations of the $e \in E$, thus (iii) holds.

(v)⇒(vi a) For finite $F \subseteq E$ we define $x_F = \sum_{e\in F}\langle x, e\rangle e$. Pythagoras' theorem gives $\|x_F\|^2 = \sum_{e\in F}|\langle x, e\rangle|^2$. As $F \nearrow E$, the l.h.s. converges to $\|x\|^2$ by (iii) and the r.h.s. to $\sum_{e\in E}|\langle x, e\rangle|^2$. Thus (vi a) holds.

If (vi a) holds then for each $\varepsilon > 0$ there is a finite $F \subseteq E$ such that $\sum_{e\in E\backslash F}|\langle x, e\rangle|^2 < \varepsilon$. Since $x - x_F$ is orthogonal to each $e \in F$, we have $x - x_F \perp x_F$, to that $\|x\|^2 = \|x - x_F\|^2 + \|x_F\|^2$. Combining this with (iv a) and $\|x_F\|^2 = \sum_{e\in F}|\langle x, e\rangle|^2$ we find $\|x - x_F\|^2 = \sum_{e\in E\backslash F}|\langle x, e\rangle|^2 < \varepsilon$. Since ε was arbitrary, this proves that $\lim_{F\nearrow E} x_F = x$, thus (v).

It remains to prove (iii)⇒(v). The fact $\overline{\mathrm{span}_{\mathbb{F}} E} = H$ means that for every $x \in H$ and $\varepsilon > 0$ there are a finite subset $F \subseteq E$ and coefficients $\{a_e\}_{e\in F}$ such that $\|x - \sum_{e\in F} a_e e\| < \varepsilon$. On the other hand, Theorem 5.27 tells us that for each finite $F \subseteq E$ there is a unique $P_F(x) \in K_F = \mathrm{span}_{\mathbb{F}} F$ minimizing $\|x - P_F(x)\|$. Clearly $P_F \restriction K_F$ is the identity map and the zero map on $K_F^\perp$. Thus defining $P'_F(x) = \sum_{e\in F}\langle x, e\rangle e$, we have $P'_F = P_F$. Thus $\|x - \sum_{e\in F}\langle x, e\rangle e\| \le \|x - \sum_{e\in F} a_e e\| < \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this proves $\lim_{F\nearrow E}\|x - \sum_{e\in F}\langle x, e\rangle e\| = 0$, which is nothing other than the statement $x = \sum_{e\in E}\langle x, e\rangle e$. Since the finite sums $\sum_{e\in F}\langle x, e\rangle e$ are in H, the identity $x = \sum_{e\in E}\langle x, e\rangle e$ also holds in H for all $x \in H$. ■

5.43 Remark 1. The identity $\langle x, y\rangle = \sum_{e\in E}\langle x, e\rangle\langle e, y\rangle\ \forall x, y \in H$ in statement (vii a) is sometimes called 'inserting a partition of unity'. Physicists write $\mathbf{1} = \sum_{e\in E}|e\rangle\langle e|$.

2. If H is only a pre-Hilbert space, we still have
(i)⇔(ii a/b)⇐(iii)⇔(iv)⇔(v)⇔(vi a/b)⇔(vii a/b). This follows from the fact that completeness is not used in proving these equivalences, except in (iii)⇒(v) where it can be avoided by appealing to Remark 5.28.

Furthermore, we have equivalence of (iii), i.e. $\overline{\mathrm{span}_{\mathbb{F}} E} = H$ (in H), with $\overline{\mathrm{span}_{\mathbb{F}} E} = \widehat{H}$ (in the completion $\widehat{H}$) and with E being an ONB for $\widehat{H}$. The equivalence of the second and third statement comes from (i)⇔(iii), applied to $\widehat{H}$. If $\mathrm{span}_{\mathbb{F}} E$ is dense in H then it is dense in $\widehat{H}$

since H is dense in $\widehat{H}$. And the converse follows from the general topology fact that the closure in H of some $S \subseteq H \subseteq \widehat{H}$ equals $\overline{S} \cap H$, where $\overline{S}$ is the closure in $\widehat{H}$.

In Example 5.49 below, all statements (i)-(vii) hold despite the incompleteness of H. But in the absence of completeness the implication (i)⇒(iii) can fail! For a counterexample see Exercise 5.44. (In view of this, maximal orthonormal sets in pre-Hilbert spaces should not be called bases.) In [63] it is even proven that a pre-Hilbert space in which every maximal orthonormal set E has dense span actually is a Hilbert space. Equivalently, in every incomplete pre-Hilbert space there is a maximal orthonormal set E whose span is non-dense! There even are pre-Hilbert spaces (called pathological) in which no orthonormal set has dense span!

Actually, most of the non-trivial results, like $H \cong K \oplus K^\perp$ for closed subspaces K and Theorem 5.34, hold for a pre-Hilbert space if and only if it is a Hilbert space, see [63]. □

5.44 Exercise (Counterexample) Let $H = \ell^2(\mathbb{N}, \mathbb{F})$, let $f = \sum_{n=1}^\infty \delta_n/n \in H$ (equivalently, $f(n) = 1/n$). Now $K = \mathrm{span}_{\mathbb{F}}\{f, \delta_2, \delta_3, \ldots\}$ (no closure!) is a pre-Hilbert space. Prove:

(i) $E = \{\delta_2, \delta_3, \ldots\}$ is a maximal orthonormal set in K.

(ii) $f \notin \overline{\mathrm{span}_{\mathbb{F}} E}$, thus $\overline{\mathrm{span}_{\mathbb{F}} E} \neq K$ (both closures in K).

5.45 Theorem ((F.) Riesz-Fischer) [35][36] *Let H be a pre-Hilbert space and E an orthonormal set such that $\overline{\mathrm{span}_{\mathbb{F}} E} = H$. Then the following are equivalent:*

(i) H is a Hilbert space (thus complete).

(ii) The isometric map $H \to \ell^2(E, \mathbb{F})$, $x \mapsto \{\langle x, e\rangle\}_{e \in E}$ is surjective. I.e. for every $f \in \ell^2(E, \mathbb{F})$ there is an $x \in H$ such that $\langle x, e\rangle = f(e)$ for all $e \in E$.

Proof. (ii)⇒(i) We know from (iii)⇒(vii b) in Theorem 5.42 that the map $H \to \ell^2(E, \mathbb{F})$ is an isometry. If it is surjective then it is an isomorphism of pre-Hilbert spaces. Since $\ell^2(E, \mathbb{F})$ is complete by Lemma 4.8, so is H.

(i)⇒(ii) With $f \in \ell^2(E, \mathbb{F})$ we have $\sum_{e \in E} |f(e)|^2 < \infty$. This implies that for each $\varepsilon > 0$ there is a finite $F \subseteq E$ such that $\sum_{e \in E \backslash F} |f(e)|^2 < \varepsilon$. For each finite subset $F \subseteq E$ we define $x_F = \sum_{e \in F} f(e)e$. Whenever $U, U' \subseteq E$ are finite subsets containing F, the identity $x_U - x'_U = \sum_{e \in E}(\chi_U(e) - \chi_{U'}(e))f(e)e$ implies

$$\|x_U - x_{U'}\|^2 = \sum_{e \in E} |\chi_U(e) - \chi_{U'}(e)|^2 |f(e)|^2 \leq \varepsilon$$

since $|\chi_U - \chi_{U'}|$ vanishes on F and is bounded by one on $(U \cup U')\backslash F$. Thus $\{x_F\}_{F \subseteq E \text{ finite}}$ is a Cauchy net in H and therefore convergent to a unique $x \in H$ by completeness, cf. Lemma A.14. By continuity of the inner product, $\langle x_F, e\rangle$ converges to $f(e)$, so that $\langle x, e\rangle = f(e)$ for all $e \in E$. ■

5.46 Remark If the $f : E \to \mathbb{F}$ in (ii) satisfies $\sum_{e \in E} |f(e)| < \infty$, thus $f \in \ell^1(E, \mathbb{F})$, then we can simply put $x = \sum_{e \in E} f(e)e$, the convergence following from Proposition 3.15(iii). But for infinite E, we have a strict inclusion $\ell^1(E, \mathbb{F}) \subsetneqq \ell^2(E, \mathbb{F})$ (compare Exercise 4.14), so that the above proof is not redundant. Here we again see that infinite-dimensional Hilbert spaces contain series that are unconditionally but not absolutely convergent. □

[35] Ernst Sigismund Fischer (1875-1954). Austrian mathematician. Early pioneer of Hilbert space theory.

[36] Also the completeness of $L^2(X, \mathcal{A}, \mu; \mathbb{F})$ (see Lemma 4.8 for $\ell^2(S)$) is sometimes called Riesz-Fischer theorem.

5.47 Proposition *For a Hilbert space H, the following are equivalent:*

(i) H is separable in the topological sense, i.e. there is a countable dense set $S \subseteq H$.

(ii) H admits a countable orthonormal basis.

(iii) Every orthonormal basis for H is countable.

Proof. If $E \subseteq H$ is any ONB for H, Theorem 5.45 gives a unitary equivalence $H \cong \ell^2(E, \mathbb{F})$. By Proposition 4.17, $\ell^2(E, \mathbb{F})$ is separable if and only if E is countable. Combining these facts proves the implications (ii)⇒(i)⇒(iii), while (iii)⇒(ii) is trivial. ■

5.48 Remark One can prove that any two ONBs E, E' for a Hilbert space H have the same cardinality, i.e. there is a bijection between E and E', cf. e.g. [30, Proposition I.4.14]. (This does not follow from the linear algebra proof, since the latter uses a different notion of basis, the Hamel bases.) The common cardinality of all bases of H is called the dimension of H. □

5.49 Example Here is an application of Theorem 5.42: Let[37]

$$H = \{f \in C([0, 2\pi], \mathbb{C}) \mid f(0) = f(2\pi)\}.$$

One easily checks that $\langle f, g\rangle = (2\pi)^{-1} \int_0^{2\pi} f(x)\overline{g(x)}dx$ (Riemann integral) is an inner product, so that $(H, \langle\cdot,\cdot\rangle)$ is a pre-Hilbert space. (If a continuous function satisfies $\int |f(x)|^2 dx = 0$ then it is identically zero.) For $n \in \mathbb{Z}$, let $e_n(x) = e^{inx}$. It is straightforward to show that $E = \{e_n \mid n \in \mathbb{Z}\}$ is an orthonormal set, thus Bessel's inequality holds. For $f \in H$ we have

$$\langle f, e_n\rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-inx}dx,$$

which is the n-th Fourier coefficient $\widehat{f}(n)$ of f, cf. e.g. [157, 83]. In fact, in Fourier analysis one proves, cf. e.g. [157, Corollary 5.4], that the finite linear combinations of the e_n ('trigonometric polynomials') are dense in H, which is (iii) of Theorem 5.42. Thus all other statements in the theorem also hold. The weaker statement (ii a) is also well-known in Fourier analysis, cf. [157, Corollary 5.3]. Furthermore,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \|f\|^2 = \sum_{n\in\mathbb{Z}} |\langle f, e_n\rangle|^2 = \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2.$$

This is the original Parseval formula, cf. e.g. [157, Chapter 3, Theorem 1.3]. Note that H is not complete. Measure theory tells us that this completion is $L^2([0, 2\pi], \lambda; \mathbb{C})$, the measure being Lebesgue measure λ (defined on the σ-algebra of Borel sets). Now the map $L^2([0, 2\pi]) \to \ell^2(\mathbb{Z}, \mathbb{C})$, $f \mapsto \widehat{f}$ is an isomorphism of Hilbert spaces. This nice situation shows that the Lebesgue integral is much more appropriate for the purposes of Fourier analysis than the Riemann integral (as for most other purposes).

Note: If we consider $L^2([0, 2\pi], \lambda; \mathbb{R})$ instead, we must replace the basis E by $E_\mathbb{R} = \{\cos nx \mid n \in \mathbb{N}_0\} \cup \{\sin nx \mid n \in \mathbb{N}\}$. One easily checks that $\text{span}_\mathbb{C} E = \text{span}_\mathbb{C} E_\mathbb{R}$.

5.50 Exercise Prove that the pre-Hilbert space $H = C([0, 1])$ with inner product $\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt$ is not complete.

[37]The set of continuous 2π-periodic functions can be identified with $C(S^1, \mathbb{C})$ via $z = e^{ix}$. We write $f(x)$ when we consider f as a function on $[0, 2\pi]$ (or $\mathbb{R}$) and $f(z)$ if f is understood as a function on $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$.

5.7 Tensor products of Hilbert spaces

In this optional section, referenced only in Section 12.4 but important well beyond that, you are assumed to know[38] the notion of (algebraic) tensor product $V \otimes_k W$ of two vector spaces V, W over a field k. (In two sentences: $V \otimes_k W$ is the free abelian group spanned the pairs $(v, w) \in V \times W$, divided by the subgroup generated by all elements of the form $(v+v', w)-(v, w)-(v', w)$ and $(v, w+w')-(v, w)-(v, w')$ and $(cv, w)-(v, cw)$, where $v, v' \in V, w, w' \in W, c \in k$, the quotient being a k-vector space in the obvious way. If $v \in V, w \in W$ then the equivalence class $[(v, w)]$ is denoted $v \otimes w$.)

The crucial property is that given a bilinear map $\alpha : V \times W \to Z$ (where $V \times W$ is the Cartesian product) there is a unique linear map $\beta : V \otimes_k W \to Z$ such that $\beta(v \otimes w) = \alpha((v, w))$.

5.51 Lemma *Let $(H, \langle \cdot, \cdot \rangle_H), (H', \langle \cdot, \cdot \rangle_{H'})$ be pre-Hilbert spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then there is a unique inner product $\langle \cdot, \cdot \rangle_Z$ on $Z = H \otimes_{\mathbb{F}} H'$ such that $\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle_H \langle w, w' \rangle_{H'}$.*

Proof. Every element $z \in Z = H \otimes_{\mathbb{F}} H'$ has a representation $z = \sum_{k=1}^K v_k \otimes w_k$ with $K < \infty$. Given another $z' = \sum_{l=1}^L v_l' \otimes w_l' \in H \otimes_{\mathbb{F}} H'$, we must define

$$\langle z, z' \rangle_Z = \sum_{k=1}^K \sum_{l=1}^L \langle v_k, v_l' \rangle_H \langle w_k, w_l' \rangle_{H'}.$$

Since an element $z \in Z$ can have many representations of the form $z = \sum_{k=1}^K v_k \otimes w_k$, we must show that this is well-defined. Let thus $\sum_{k=1}^K v_k \otimes w_k = \sum_{\tilde{k}=1}^{\tilde{K}} \tilde{v}_{\tilde{k}} \otimes \tilde{w}_{\tilde{k}}$. Now, for fixed l the map $H \times H \to \mathbb{F}$, $(x, y) \mapsto \langle x, v_l' \rangle_H \langle y, w_l' \rangle_{H'}$ clearly is bilinear, thus it gives rise to a unique linear map $H \otimes_{\mathbb{F}} H' \to \mathbb{F}$. This implies

$$\sum_{k=1}^K \sum_{l=1}^L \langle v_k, v_l' \rangle_H \langle w_k, w_l' \rangle_{H'} = \sum_{\tilde{k}=1}^{\tilde{K}} \sum_{l=1}^L \langle \tilde{v}_{\tilde{k}}, v_l' \rangle_H \langle \tilde{w}_{\tilde{k}}, w_l' \rangle_{H'}.$$

The independence of $\langle z, z' \rangle_Z$ of the representation of z' is shown in the same way.

It is quite clear that $\langle \cdot, \cdot \rangle_Z$ is sesquilinear and satisfies $\langle z', z \rangle_Z = \overline{\langle z, z' \rangle_Z}$.

In order to study $\langle z, z \rangle_Z$ we may assume that $z = \sum_k v_k \otimes w_k$, where the w_k are mutually orthogonal. This leads to

$$\langle z, z \rangle_Z = \sum_k \langle v_k, v_k \rangle_H \langle w_k, w_k \rangle_{H'} = \sum_k \|v_k\|^2 \|w_k\|^2 \geq 0$$

and $\langle z, z \rangle_Z = 0 \Rightarrow z = 0$. ■

5.52 Definition *If H, H' are Hilbert spaces then $H \otimes H'$ is the Hilbert space obtained by completing the above pre-Hilbert space $(Z, \langle \cdot, \cdot \rangle_Z)$.*

5.53 Remark 1. We usually write the completed tensor products $\otimes$ without subscript to distinguish them from the algebraic ones.

2. If E, E' are ONBs in the Hilbert spaces H, H', respectively, then it is immediate that $E \times E'$ is an orthonormal set in the algebraic tensor product $H \otimes_k H'$, thus also in $H \otimes H'$. In fact its span is dense in $E \otimes E'$, so that it is an ONB.

[38] Unfortunately, this is often omitted from undergraduate linear algebra teaching. E.g., it does not appear in [55] despite the book's > 500 pages. See however [84, 95] which, admittedly, are aiming higher.

This leads to a pedestrian way of defining the tensor product $H \otimes H'$ of Hilbert spaces over $\mathbb{F}$: Pick ONBs $E \subseteq H, E' \subseteq H'$ and define $H \otimes H' = \ell^2(E \times E', \mathbb{F})$. By Remark 5.48, the outcome is independent of the chosen bases up to isomorphism. If $x \in H, x' \in H'$ then the map $E \times E' \to \mathbb{F}$, $(e, e') \mapsto \langle x, e\rangle_H \langle x', e'\rangle_{H'}$ is in $\ell^2(E \times E, \mathbb{F})$, thus defines an element $x \otimes x' \in H \otimes H'$. This map $H \times H' \to H \otimes H'$ is bilinear. But this definition is very ugly and unconceptual due to its reliance on a choice of bases. □

6 Subspaces and quotient spaces of Banach spaces

6.1 Quotient spaces of Banach spaces

If H is a Hilbert space and $K \subseteq H$ a closed subspace, we have seen in Exercise 5.32 that the orthogonal complement $K^\perp$ is isomorphic to the quotient space H/K. For this reason, quotient spaces of Hilbert spaces are hardly ever used: We can work with the easily defined orthogonal complements instead. In a Banach space context, we have no notion of orthogonality and therefore no orthogonal complements. This is the reason why quotient spaces of Banach spaces are important, and we will now study them in some detail. (In fact, there is more trouble with general Banach spaces, to be discussed in Section 6.2.)

6.1 Proposition *If V is a normed space, $W \subseteq V$ a linear subspace and V/W denotes the quotient vector space, we define $\|\cdot\|' : V/W \to [0,\infty)$ by $\|v + W\|' = \inf_{w\in W} \|v - w\|$. Then*

(i) $\|\cdot\|'$ is a seminorm on V/W, and the quotient map $Q : V \to V/W$ satisfies $\|Q\| \le 1$.

(ii) $\|\cdot\|'$ is a norm if and only if $W \subseteq V$ is closed.

(iii) If $W \subseteq V$ is closed, the topology on V/W induced by $\|\cdot\|'$ coincides with the quotient topology, and the quotient map $p : V \to V/W$ is open.

(iv) If V is a Banach space and $W \subseteq V$ is closed then $(V/W, \|\cdot\|')$ is Banach space.

(v) If V is a normed space with closed subspace W and $T \in B(V, E)$, where E is a normed space with $W \subseteq \ker T$ then there is a unique $T' \in B(V/W, E)$ such that $T'Q = T$. Furthermore, $\|T'\| = \|T\|$. T' is surjective if and only if T is surjective. T' is injective if and only if $W = \ker T$.

(vi) If $\mathcal{A}$ is a normed (resp. Banach) algebra and $\mathcal{I} \subseteq \mathcal{A}$ is a closed two-sided ideal, then $\mathcal{A}/I$ is a normed (resp. Banach) algebra.

Proof. (i) It is clear that $\|0\|' = 0$ (where we denote the zero element of V/W by 0 rather than W). For $x \in V, c \in \mathbb{F}\backslash\{0\}$ we have

$$\|c(x + W)\|' = \|cx + W\|' = \inf_{w\in W} \|cx - w\| = |c| \inf_{w\in W} \|x - w/c\| = |c| \inf_{w\in W} \|x - w\| = |c|\|x\|',$$

where we used that $W \to W, w \mapsto cw$ is a bijection. Now let $x_1, x_2 \in V$ and $\varepsilon > 0$. Then there are $w_1, w_2 \in W$ such that $\|x_i - w_i\| < \|x_i + W\|' + \varepsilon/2$ for $i = 1, 2$. Then

$$\begin{aligned} \|x_1 + x_2 + W\|' &= \inf_{w\in W} \|x_1 + x_2 + w\| \le \|(x_1 - w_1) + (x_2 - w_2)\| \\ &\le \|x_1 - w_1\| + \|x_2 - w_2\| < \|x_1 + W\|' + \|x_2 + W\|' + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have $\|x_1+x_2+W\|' \le \|x_1+W\|'+\|x_2+W\|'$, proving subadditivity of $\|\cdot\|'$. In view of $0 \in W$ it is immediate that $\|v + W\|' = \inf_{w\in W} \|v - w\| \le \|v\|$.

(ii) If $v \in V$, the definition of $\|\cdot\|'$ readily implies that $\|v+W\|' = 0$ if and only if $v \in \overline{W}$. Thus if W is closed then $w = v+W \in V/W$ has $\|w\|' = 0$ only if w is the zero element of V/W. And if W is non closed then every $v \in \overline{W}\backslash W$ satisfies $\|v+W\|' = 0$ even though $v+W \in V/W$ is non-zero. Thus $\|\cdot\|'$ is not a norm.

(iii) Continuity of $Q : (V, \|\cdot\|) \to (V/W, \|\cdot\|')$ follows from $\|Q\| \le 1$, see (i). Since Q is norm-decreasing, we have $Q(B^V(0,r)) \subseteq B^{V/W}(0,r)$ for each $r > 0$. And if $y \in V/W$ with $\|y\| < r$ then there is an $x \in V$ with $Q(x) = y$ and $\|x\| < r$ (but typically larger than $\|y\|$). Thus Q maps $B^V(0,r)$ onto $B^{V/W}(0,r)$ for each r. Similarly, $Q(B^V(x,r)) = B^{V/W}(Q(x),r)$, and from this it is easily deduced that $Q(U) \subseteq V/W$ is open for each open $U \subseteq V$. Thus Q is open (w.r.t. the norm topologies on $V, V/W$), which implies (cf. [108, Lemma 6.4.5]) that Q is a quotient map, thus the topology on V/W coming from $\|\cdot\|'$ is the quotient topology.

(iv) Let $\{y_n\} \subseteq V/W$ be a Cauchy sequence. Then we can pass to a subsequence $w_n = y_{i_n}$ such that $\|w_n - w_{n+1}\| < 2^{-n}$. Pick $x_n \in V$ such that $Q(x_n) = w_n$ and $\|x_n - x_{n+1}\| < 2^{-n}$. (Why can this be done?) Then $\{x_n\}$ is a Cauchy sequence converging to some $x \in V$ by completeness of V. With $y = Q(x)$ we have $\|y_n - y\| \le \|x_n - x\| \to 0$. Thus $y_n \to y$, and V/W is complete.

(v) If $y \in V/W$, and $x, x' \in V$ satisfy $Q(x) = Q(x') = y$ then $Q(x-x') = 0$, thus $x - x' \in \ker Q = W \subseteq \ker T$, implying $Tx = Tx'$. Thus putting $T'y = Tx$ gives rise to a well-defined map $T' : V/W \to Z$ satisfying $T'Q = T$. One easily checks that T' is linear. And using $Q(B^V(0,1)) = B^{V/W}(0,1)$ from the proof of (iii) we have

$$\begin{aligned} \|T'\| &= \sup\{\|T'y\| \mid y \in B^{V/W}(0,1)\} = \sup\{\|T'Q(x)\| \mid x \in B^V(0,1)\} \\ &= \sup\{\|Tx\| \mid x \in B^V(0,1)\} = \|T\|. \end{aligned}$$

The statements concerning injectivity and surjectivity of T' are pure algebra, but for completeness we give proofs: The statement about surjectivity follows from $T = T'Q$ together with surjectivity of Q, which gives $T(V) = T'(V/W)$. If $W \subsetneqq \ker T$, pick $x \in (\ker T)\backslash W$ and put $y = Q(x)$. Then $y \ne 0$, but $T'y = T'Qx = Tx = 0$, so that T' is not injective. Now assume $W = \ker T$. If $y \in \ker T'$ then pick $x \in V$ with $y = Q(x)$. Then $Tx = T'Qx = T'y = 0$, thus $x \in \ker T = W$, so that $y = Q(x) = 0$, proving injectivity of T'.

(vi) It is known from algebra that $\mathcal{A}/I$ is again an algebra. By the above, it is a normed (resp. Banach) space. It remains to prove that the quotient norm on $\mathcal{A}/I$ is submultiplicative. Let $c, d \in \mathcal{A}/I$ and $\varepsilon > 0$. Then there are $a, b \in \mathcal{A}$ with $Q(a) = c$, $Q(b) = d$, $\|a\| < \|c\| + \varepsilon$, $\|b\| < \|d\| + \varepsilon$ (see the exercise below). Then $\|cd\| = \|Q(ab)\| \le \|ab\| \le \|a\|\|b\| < (\|c\| + \varepsilon)(\|d\| + \varepsilon)$, and since this holds for all $\varepsilon > 0$, we have $\|cd\| \le \|c\|\|d\|$. ■

6.2 Exercise (i) If V is a normed space and $W \subseteq V$ is a closed subspace, prove that for every $y \in V/W$ and every $\varepsilon > 0$ there is an $x \in V$ with $Q(x) = y$ and $\|x\| \le \|y\| + \varepsilon$.

(ii) Give an example of a normed space V, a closed subspace W and $y \in V/W$ for which no $x \in V$ with $y = Q(x)$, $\|x\| = \|y\|$ exists.

We have seen that if V is Banach and $W \subseteq V$ is closed then W and V/W (with their inherited and quotient norms, respectively) are complete. The converse is also true:

6.3 Proposition *Let V be a normed space and $W \subseteq V$ a closed subspace. Then*

(i) If W and V/W are complete then so is V.

(ii) If W is complete and has finite codimension then V is complete. If W is finite-dimensional and V/W is complete then V is complete.

Proof. (i) Let $\{v_n\} \subseteq V$ be a Cauchy sequence. Since $Q : V \to V/W$ is bounded, the sequence $\{z_n = Q(v_n)\} \subseteq V/W$ is Cauchy, so that by completeness of V/W it converges to some $z \in V/W$. Pick $\widehat{z} \in V$ with $Q(\widehat{z}) = z$. By Exercise 6.2(i), for every $n \in \mathbb{N}$ we can find a $y_n \in V$ such that $Q(y_n) = z - z_n = Q(\widehat{z} - v_n)$ and $\|y_n\| \leq \|z - z_n\| + 2^{-n}$. With $z_n \to z$ this implies $y_n \to 0$. With $Q(y_n + v_n - \widehat{z}) = 0$, we have $y_n + v_n - \widehat{z} \in \ker Q = W \ \forall n$. Since $\{v_n\}$ and $\{y_n\}$ are Cauchy, so is $\{y_n + v_n - \widehat{z}\}$ so that by completeness of W we have $y_n + v_n - \widehat{z} \to w$ for some $w \in W$. In view of $y_n \to 0$ this implies $v_n \to w + \widehat{z} \in V$. Thus V is complete.

(ii) Recalling that finite codimensionality of W means $\dim(V/W) < \infty$, both statements are immediate by (i) and the completeness of finite-dimensional normed spaces (Exercise 2.33). ■

6.4 Exercise Use the quotient space construction of Banach spaces to

(i) give a new proof for the difficult ($\Leftarrow$) part of Exercise 3.9.

(ii) prove that every non-zero $\varphi \in V^*$ is an open map. (To be vastly generalized later.)

6.5 Exercise If V a Banach space and $W \subseteq V$ a closed subspace, prove that V is separable if and only if W and V/W are separable.

If V is a Banach space and $W \subseteq V$ a closed linear subspace, it is natural to ask how the dual spaces W^* and $(V/W)^*$ are related to V^*. This leads to the following definitions, which are closely related to the Hilbert space $\perp$, but *not* the same:

6.6 Definition *Let V be a normed space.*

- *For any $W \subseteq V$, the annihilator of W is $W^\perp = \{\varphi \in V^* \mid \varphi(x) = 0 \ \forall x \in W\} \subseteq V^*$.*
- *For any $\Phi \subseteq V^*$, the annihilator of Φ is $\Phi^\top = \{x \in V \mid \varphi(x) = 0 \ \forall \varphi \in \Phi\} \subseteq V$.*

(Some authors write ${}^\perp\Phi$ instead of $\Phi^\top$.)

One easily checks for any $W \subseteq V$ that $W^\perp \subseteq V^*$ is a closed linear subspace and $W^\perp = \overline{\mathrm{span}_{\mathbb{F}} W}^\perp$. (Similar statements hold for $\Phi^\top$.) We trivially have $W \subseteq (W^\perp)^\top$ and $\Phi \subseteq (\Phi^\top)^\perp$ for all $W \subseteq V$, $\Phi \subseteq V^*$, and $\{0\}^\perp = V^*$ and $\overline{W} = V \Rightarrow W^\perp = \{0\}$. (For the less trivial converses see Exercise 9.14.) It is easy that $\Phi \subseteq \{0\} \Leftrightarrow \Phi^\top = V$. Answering the question when $\Phi^\top = \{0\}$ requires more preparations, cf. Theorem 10.32(ii).

6.7 Exercise Let V be a Banach space and $W \subseteq V$ a closed subspace. Let $Q : V \to V/W$ be the quotient map. Prove that the map $\alpha : (V/W)^* \to V^*, \ \psi \mapsto \psi Q$ is injective and isometric and its image is $W^\perp \subseteq V^*$. Thus $(V/W)^* \cong W^\perp$ as Banach spaces.

Under the same assumptions, $W^* \cong V^*/W^\perp$, but this has to await Exercise 9.15.

If V is a normed space and $W, Z \subseteq V$ are closed linear subspaces, we will later see that $W + Z = \{w + z \mid w \in W, z \in Z\} \subseteq V$ can fail to be closed. If W and Z are both finite-dimensional then $W + Z$ is finite-dimensional, thus closed. More generally:

6.8 Exercise Let V be a normed space, $W \subseteq V$ a closed subspace and $Z \subseteq V$ a finite-dimensional subspace. Prove that $W + Z \subseteq V$ is closed. Hint: Use V/W.

For more on closedness of sums of closed subspaces see Exercises 7.13 and 7.44.

6.2 Complemented subspaces

If V is a vector space over any field $\mathbb{K}$ and $W \subseteq V$ is a linear subspace, it is known from linear algebra that we can find another subspace $Z \subseteq V$ such that $V = W + Z$ and $W \cap Z = \{0\}$. The proof is easy: Pick a (Hamel) basis E for W, extend it to a basis $E' \supseteq E$ of V and put $Z = \mathrm{span}_{\mathbb{F}}(E' \backslash E)$. Such a Z is called a (algebraic) complement for W. Since every $x \in V$ can be written as $x = w + z$ with $w \in W, z \in Z$ in a unique way ($w + z = w' + z' \Rightarrow w - w' = z' - z \in W \cap Z = \{0\}$) one says V is the internal direct sum of W and Z, or $V \cong W \oplus Z$.

If V is a Banach space and $W \subseteq V$ a closed subspace, it is natural ask for a complementary subspace Z with the above properties to be closed, too. We have seen that for every closed subspace K of a Hilbert space H there is a *closed* complement, namely the orthogonal complement $K^\perp \subseteq H$. Since $K^\perp$ is defined in terms of the inner product, it is not surprising that the situation will turn out more complicated for general Banach spaces, where no inner product is around. (Simply passing from an algebraic complement Z to its closure is no solution since there is no reason for $W \cap \overline{Z} = \{0\}$ to hold.) This leads us to define:

6.9 Definition *Let V be a topological vector space. A closed subspace $W \subseteq V$ is called complemented if there is a closed subspace $Z \subseteq V$ such that $V = W + Z$ and $W \cap Z = \{0\}$.*

6.10 Remark 1. In Exercise 7.15 we will prove that if V is a Banach space and $W, Z \subseteq V$ are complementary closed subspaces, the linear isomorphism $V \simeq W \oplus Z$ also is a homeomorphism, thus an isomorphism of Banach spaces.

2. By the comments preceding the definition, every closed subspace of a Hilbert space is complemented, and the same holds for Banach spaces isomorphic to a Hilbert space.

3. But 'most' infinite-dimensional Banach spaces have uncomplemented closed subspaces! The simplest example of an uncomplemented closed subspace probably is $c_0(\mathbb{N}, \mathbb{R}) \subseteq \ell^\infty(\mathbb{N}, \mathbb{R})$. For a proof, not entirely trivial, see Appendix B.3.3.

Another easily stated example is given by $X = C(S^1, \mathbb{C})$ with norm $\|\cdot\|_\infty$ and the closed subspace $Y = \{f \in X \mid \widehat{f}(n) = 0 \ \forall n < 0\}$, where $\widehat{f}(n) = \int_0^1 f(e^{2\pi i t}) e^{-2\pi i n t} dt$, $n \in \mathbb{Z}$ are the Fourier coefficients. Proving that $Y \subseteq X$ is not complemented, see e.g. [76, p. 163-4], boils down to rather classical analysis, namely the fact that $\sum_{n=1}^N \frac{\sin nx}{n}$ is bounded uniformly in $N \in \mathbb{N}$ and $x \in \mathbb{R}$, while $\sum_{n=1}^N \frac{e^{inx}}{n}$ is not, combined with Exercise 7.15.

4. In fact, by a remarkable theorem of Lindenstrauss and Tzafriri[39] (1971), cf. e.g. [1, Theorem 13.4.5], [97, Vol. 2, Chap. 1, Theorem V.1], every Banach space all closed subspaces of which are complemented is isomorphic to a Hilbert space. Cf. also Remark 9.13. □

On the positive side, we have some results and exercises:

6.11 Proposition *Let V be a normed space and $W \subseteq V$ a linear subspace.*

(i) If W is closed and has finite codimension (i.e. $\dim(V/W) < \infty$) then it is complemented.

(ii) If W is finite-dimensional then it is complemented.

Proof. (i) Finite-dimensionality of V/W is equivalent to the existence of a finite-dimensional algebraic complement $Z \subseteq V$ for W. (This is just linear algebra, but here is a proof: Let $\{e_1, \ldots, e_n\}$ be a basis for V/W. By surjectivity of $p : V \to V/W$ we can find $x_1, \ldots, x_n \in V$ such that $p(x_i) = e_i$. It is easy to see that $\{x_1, \ldots, x_n\}$ is linearly independent. Let $Z =$

[39] Joram Lindenstrauss (1936-2012). Lior Tzafriri (1936-2008). Israeli mathematicians with many contributions to Banach space theory.

$\text{span}_{\mathbb{F}}\{x_1, \ldots, x_n\}$. Now it is straightforward that Z is a complement for W.) Being finite-dimensional, it is automatically closed by Exercise 3.22.

(ii) The proof will be given in Section 9.3 since it requires tools still to be developed. ■

6.12 Exercise It is not true that every subspace $W \subseteq V$ with $\dim(V/W) < \infty$ of a Banach space V is closed! Find a counterexample! (Hint: try codimension one.)

6.13 Exercise Let V be a normed space and $P \in B(V)$ satisfying $P^2 = P$ (i.e. idempotent). Prove that $W = PV$ is a complemented subspace. (For a converse see Exercise 7.15.)

6.14 Exercise Let $V = C([0,2], \mathbb{R})$ with the $\|\cdot\|_\infty$-norm. Let $W = \{f \in V \mid f_{|(1,2]} = 0\}$.

(i) Prove that W is complemented.

(ii) Can you 'classify' all possible complements, i.e. put them in bijection with a simpler set?

For more on the subject of complemented subspaces see [102].

In the process of returning from Hilbert to the more general Banach spaces, the above discussion of quotient spaces and complements was the easiest part. The question of bases is much more involved for Banach spaces, as the very extensive two-volume treatment [153] of the subject attests. (Then again, the basics are quite accessible[40], cf. e.g. [102, 26, 1, 74], but unfortunately we don't have the time.) The same is true for the formidable subject of tensor products of Banach spaces, see e.g. [144]. Going into that would be pointless given that we already slighted the much simpler tensor products of Hilbert spaces.

A more tractable problem is the fact that in the absence of an inner product, the existence of non-zero bounded linear functionals is rather non-trivial and can in general only be proven non-constructively. We will do this in Section 9. (Of course, for spaces that are given very explicitly like $\ell^p(S, \mathbb{F})$, we may well have more concrete descriptions of the dual spaces as in Section 4.5.) But first we will prove two major theorems that are non-trivial even when restricted to Hilbert spaces. Both of them use Baire's theorem on complete metric spaces, cf. Appendix A.5 which should perhaps be read first.

7 Open Mapping Theorem and relatives

7.1 Open Mapping Theorem. Bounded Inverse Theorem

We begin with a few very easy equivalences:

7.1 Exercise Let E, F be normed spaces and $T \in B(E, F)$. Consider the statements

(i) T is open (i.e. $TU \subseteq F$ is open for each open $U \subseteq E$).

(ii) For every $\alpha > 0$ there exists $\beta > 0$ such that $B^F(0, \beta) \subseteq TB^E(0, \alpha)$.

(iii) There exist $\alpha, \beta > 0$ such that $B^F(0, \beta) \subseteq TB^E(0, \alpha)$.

(iv) There is a $C > 0$ such that for every $y \in F$ there exists $x \in E$ with $Tx = y$ and $\|x\| \leq C\|y\|$. (This is a more quantitative or 'controlled' version of surjectivity.)

[40] A Schauder basis for a Banach space is a sequence $\{e_n\}_{n \in \mathbb{N}}$ such that for every $x \in V$ there are unique $c_n \in \mathbb{F}$ such that $x = \sum_{n=1}^\infty c_n e_n$ in the sense of (possibly conditional) convergence of series. One then proves that there are continuous linear functionals φ_n such that $x = \sum_n \varphi_n(x)x$. Existence of a Schauder basis clearly implies separability of V, but while most 'natural' separable Banach spaces have Schauder bases, counterexamples exist!

(v) T is surjective.

Obviously (iv)⇒(v). Prove the easy equivalences (i)⇔(ii)⇔(iii)⇔(iv).

Remarkably, for Banach spaces also (v) is equivalent to (i)-(iv):

7.2 Theorem (Open Mapping Theorem (OMT), Schauder 1930) [41] *If E, F are Banach spaces then every surjective $T \in B(E, F)$ is open (and (iv) holds, which is often useful).*

Most proofs of this theorem are not very transparent. We follow the slightly better approach of [61], which makes the proof of Proposition 7.4 more palatable by isolating a lemma that also has other applications[42]. It deduces statement (iv) in Exercise 7.1 from completeness of E and an approximate form of surjectivity of $T \in B(E, F)$:

7.3 Lemma *Let E be a Banach space, F a normed space and $T \in B(E, F)$. Assume also that there are $m > 0$ and $r \in (0, 1)$ such that for every $y \in F$ there is an $x_0 \in E$ with $\|x_0\| \le m\|y\|$ and $\|y - Tx_0\| \le r\|y\|$. Then for every $y \in F$ there is an $x \in E$ such that $\|x\| \le \frac{m}{1-r}\|y\|$ and $Tx = y$. In particular, T is surjective.*

Proof. By linearity we may assume $\|y\| = 1$. By assumption, there is $x_0 \in E$ such that $\|x_0\| \le m$ and $\|y - Tx_0\| \le r$. Putting $y_1 = y - Tx_0$ we have $\|y_1\| \le r$, and applying the hypothesis to y_1, we find an $x_1 \in E$ with $\|x_1\| \le m\|y_1\| \le rm$ and $\|y - T(x_0 + x_1)\| = \|y_1 - Tx_1\| \le r\|y_1\| \le r^2$. Continuing this inductively[43] we obtain a sequence $\{x_n\} \subseteq E$ such that for all $n \in \mathbb{N}$ we have

$$\|x_n\| \le r^n m, \tag{7.1}$$

$$\|y - T(x_0 + x_1 + \cdots + x_n)\| \le r^{n+1}. \tag{7.2}$$

Now, (7.1) together with completeness of E implies, cf. Proposition 3.15, that $\sum_{n=0}^{\infty} x_n$ converges to an $x \in E$ with

$$\|x\| \le \sum_{n=0}^{\infty} \|x_n\| \le \sum_{n=0}^{\infty} r^n m = \frac{m}{1-r},$$

and taking $n \to \infty$ in (7.2) gives $y = Tx$. Not assuming $\|y\| = 1$ gives an extra factor $\|y\|$. ■

7.4 Proposition *If E is a Banach space, F a normed space, $T \in B(E, F)$ and there are $\alpha, \beta > 0$ such that $B^F(0, \beta) \subseteq \overline{TB^E(0, \alpha)}$ then the statements (i)-(iv) from Exercise 7.1 hold.*

Proof. The hypothesis clearly implies $\overline{B^F(0, \beta)} \subseteq \overline{TB^E(0, \alpha)}$. By linearity of T and the fact that multiplication with a non-zero scalar is a homeomorphism, thus commutes with the closure, this is equivalent to $\overline{B}^F(0, 1) \subseteq \overline{TB^E(0, \gamma)}$, [44] where $\gamma = \alpha/\beta$. Thus for each $y \in F$, $\|y\| \le 1$, $\varepsilon > 0$ there is $x \in E, \|x\| < \gamma$ such that $\|Tx - y\| < \varepsilon$. This, in turn, is equivalent to $\forall y \in F, \varepsilon > 0\ \exists x \in E : \|x\| < \gamma\|y\|, \|Tx - y\| < \varepsilon\|y\|$. (Why?) Since this precisely is the hypothesis of Lemma 7.3, we can conclude that every $y \in F$ equals Tx for an $x \in E$ with $\|x\| \le \frac{\gamma}{1-\varepsilon}\|y\|$. This is statement (iv) in Exercise 7.1. ■

[41] Juliusz Schauder (1899-1943). Polish mathematician. Born in Lwow (then Austria-Hungary, now Ukraine). Killed by the Gestapo.

[42] It leads to a quick proof of the Tietze extension theorem in topology, see Appendix A.6.1.

[43] Here we are using the axiom DC_ω of countable dependent choice, cf. Appendix A.4.

[44] In a normed space E we have $\overline{B(x, r)} = \overline{B}(x, r) := \{y \in E \mid d(x, y) \le r\}$, but in a metric space $\supseteq$ may fail!

Proof of Theorem 7.2. Since T is surjective, we have

$$F = TE = \bigcup_{n=1}^{\infty} \overline{TB^E(0,n)}.$$

Since F is a complete metric space and has non-empty interior $F^0 = F \neq \emptyset$, Corollary A.24 of Baire's theorem implies that at least one of the closed sets $\overline{TB^E(0,n)}$ has non-empty interior. Thus there are $n \in \mathbb{N}, y \in F, \varepsilon > 0$ such that $B^F(y,\varepsilon) \subseteq \overline{TB^E(0,n)}$. If $x \in B^F(0,\varepsilon)$ then $2x = (y+x) - (y-x)$, thus $2B^F(0,\varepsilon) \subseteq B^F(y,\varepsilon) - B^F(y,\varepsilon)$ and thus

$$B^F(0,\varepsilon) \subseteq \frac{1}{2}(B^F(y,\varepsilon) - B^F(y,\varepsilon)) \subseteq \frac{1}{2}\left(\overline{TB^E(0,n)} - \overline{TB^E(0,n)}\right) \subseteq \overline{TB^E(0,n)}.$$

Thus the hypothesis of Proposition 7.4 is satisfied with $\alpha = n, \beta = \varepsilon$, and we are done. ■

7.5 Exercise Let E be a Banach space, F a normed space and $A \in B(E,F)$. Prove: If A is open then F is complete, thus Banach.

Hint: Combine Exercise 7.1 and the method of proof of Proposition 6.1(iv).

7.6 Remark 1. The preceding exercise shows that $A \in B(E,F)$ is never open if E is a Banach space and F an incomplete normed space! On the other hand, if E is incomplete, openness of $A \in B(E,F)$ can fail even if A is bijective. See Exercise 7.23 below.

2. The proof of the OMT involved two applications of DC_ω, in the proof of the approximation Lemma 7.3 and via Baire's theorem. It actually is possible to replace the invocation of Baire's theorem by the (weak) uniform boundedness theorem, see [46]. This is interesting since the latter can be proven using only the axiom of countable choice, cf. Appendix B.4. But replacing DC_ω by AC_ω in proving the approximation Lemma 7.3 seems impossible.

3. Recent claims by various authors (Kesavan, Liebaug and Spindler, Velasco, ...) that OMT can be deduced from the Uniform Boundedness Theorem (next section) do not make much sense. From a logical point of view, 'A implies B' is true for every true statement B: Just ignore A and prove B from scratch. In order for 'A implies B' to have meaning, we must restrict the tools allowed in proving the implication, as in Theorem B.43, where no use of choice axioms is allowed. The proposed deduction of OMT from UBT is not of this type since is uses DC_ω, while DC_ω is sufficient to prove OMT without invoking UBT! (There is another non-rigorous but very common use of 'A implies B' in the sense of 'deducing B from A is much simpler than proving B without A'. As in: The non-existence of a retraction from a ball to its boundary implies the Brouwer fixed point theorem. But also this does not apply here.) □

7.7 Corollary (Bounded Inverse Theorem (BIT), Banach 1929) *If E, F are Banach spaces and $T \in B(E,F)$ is a bijection then also T^{-1} is bounded.*

Proof. By Theorem 7.2, T is open. Thus the inverse T^{-1} that exists by bijectivity (and clearly is linear) is continuous, thus bounded by Lemma 3.5 or property (iv) from Exercise 7.1. ■

Proof of Theorem 2.23. The hypothesis implies that $\mathrm{id}_V : (V, \|\cdot\|_1) \to (V, \|\cdot\|_2)$ is a continuous bijection, thus a homeomorphism by the BIT. Now Lemma 3.5 gives $\|\cdot\|_1 \leq c'\|\cdot\|_2$. ■

7.8 Exercise Prove:

(i) The Two Norm Theorem 2.23 implies the BIT.

(ii) The BIT implies the OMT. (Hint: Use a quotient map.)

7.9 Remark 1. The BIT is equivalent to the statement that every continuous linear bijection of Banach spaces is a homeomorphism. This is reminiscent of the statement that every continuous bijection of compact Hausdorff spaces is a homeomorphism. (The analogy ends when we look at the generalization: Every continuous map from a compact space to a Hausdorff space is a *closed* map. But there are other open mapping theorems, e.g. for topological groups.)

2. The Open Mapping Theorem can be generalized to the case where E is an F-space, i.e. a TVS admitting a complete translation-invariant metric. See [141, Theorem 2.11]. □

7.2 Some applications

With the Open Mapping Theorem at our disposal, we return to the quotient space construction:

7.10 Corollary *Let E, F be Banach spaces and $T \in B(E, F)$ surjective. Then*

(i) The topology of F coincides with the quotient topology, thus T is a quotient map.

(ii) The linear bijection $T' : E/\ker T \to F$ from Proposition 6.1 is a homeomorphism.

(iii) The quotient norm on $E/\ker T$ is equivalent to $\|T'(\cdot)\|_F$.

Proof. (i) Since T is continuous, $T^{-1}(U) \subseteq E$ is open whenever $U \subseteq F$ is open. And if $U \subseteq F$ is such that $F^{-1}(U)$ is open then by surjectivity of T we have $U = F(F^{-1}(U))$, which is open by the OMT. Thus $U \subseteq F$ is open if and only if the preimage $T^{-1}(U) \subseteq E$ is open, which is the definition of the quotient topology. (ii) The map T' is bounded and injective by construction and surjective by surjectivity of T. Thus by the BIT it is a homeomorphism. (Alternatively this follows from continuity of T' and the fact that F is homeomorphic to $E/\ker T$ by (i).) (iii) The fact that T' is a homeomorphism means that the two norms on $E/\ker T$ are equivalent. ■

The next remarkable application will have several uses later:

7.11 Exercise Let V, W be Banach spaces and $A \in B(V, W)$.

(i) If A is injective and $\dim(W/AV) < \infty$, prove that $AV \subseteq W$ is closed.
Hint: Use an algebraic complement of $AV \subseteq W$ and the BIT.

(ii) Remove the injectivity assumption in (i).

(iii) Prove: If A has dense image but is not surjective then $\dim(W/AV) = \infty$.

7.12 Remark 1. The above contrasts interestingly with Exercise 6.12.

2. The quotient W/AV is called the (algebraic) cokernel of A. Some authors define the cokernel as $W/\overline{AV}$. But we don't do this, since finite-dimensionality of $W/\overline{AV}$ (the topological cokernel) is a much weaker condition on A and doesn't imply closedness of AV. □

The OMT and BIT have many applications to the questions concerning closed linear subspaces, their sums and complementedness:

7.13 Exercise Let V be a Banach space and $K, L \subseteq V$ closed linear subspaces. Equip $W = K \oplus L$ with the norm $\|(k, l)\| = \|k\| + \|l\|$ (or an equivalent one like $\max(\|k\|, \|l\|)$).

(i) Prove that the following are equivalent:

(α) $K + L \subseteq V$ is closed.

(β) The (surjective) map $+ : W \to K + L$, $(k, l) \mapsto k + l$ is open.

(γ) There is a C such that for every $y \in K+L$ there are $k \in K$, $l \in L$ such that $k+l=y$ and $\|k\|+\|l\| \le C\|y\|$.

(ii) If $K \cap L = \{0\}$ prove that (α)-(γ) are also equivalent to

(δ) $\inf\{\|k-l\| \mid k \in K, l \in L, \|k\| = \|l\| = 1\} > 0$. Thus the unit spheres of K and L have positive distance (which they cannot have if $K \cap L \ne \{0\}$).

7.14 Proposition *Let V be a Banach space and $W, Z \subseteq V$ closed complementary subspaces, i.e. $W+Z=V$ and $W \cap Z = \{0\}$. Then $W \oplus Z$ equipped with the norm $\|(w,z)\|_s = \|w\|+\|z\|$ (or $\|(w,z)\|_m = \max(\|w\|,\|z\|)$) is a Banach space, and the map $\alpha : W \oplus Z \to V, (w,z) \mapsto w+z$ is an isomorphism of Banach spaces (i.e. a linear bijection and a homeomorphism).*

Proof. As closed subspaces, W, Z are Banach spaces, thus $W \oplus Z$ with one of the given norms is a Banach space. The assumptions imply that α is a linear bijection. Continuity of α follows from $\|\alpha((w,z))\| = \|w+z\| \le \|w\|+\|z\| = \|(w,z)\|_s \le 2\|(w,z)\|_m$. Now the BIT gives that α is open, thus a homeomorphism. ■

7.15 Exercise Let V be a Banach space.

(i) For $W \subseteq V$ complemented, prove:

- There is a bounded linear map $P \in B(V)$ with $P^2 = P$ and $W = PV$. (The converse was proven in Exercise 6.13.)
- Every closed $Z \subseteq V$ complementary to W is isomorphic to V/W as a Banach space.

(ii) Prove that an idempotent $P : V \to V$ is bounded if and only if the linear subspaces PV and $(\mathbf{1}-P)V$ of V are both closed.

(iii) Give an example of a Banach space and an unbounded idempotent on it.

7.16 Remark If V is a Banach space and $W \subseteq V$ a finite-dimensional subspace, thus automatically closed and complemented, proving the existence of an idempotent $P \in B(V)$ with $PV = W$ does not require the open mapping theorem, cf. the proof of Proposition 6.11(ii) given in Section 9.3. One can also prove bounds on $\|P\|$ in terms of $\dim W$. □

7.17 Definition *If $A : V \to W$ and $B : W \to V$ are linear maps satisfying $AB = \mathrm{id}_W$ then A is called a left inverse of B and B is called a right inverse or section of A.*

If $AB = \mathrm{id}_W$ then clearly A is surjective and B is injective. In a pure algebra context, it is not hard to prove that every surjective (resp. injective) linear map admits a linear right (resp. left) inverse. (Do it!) But for normed spaces and bounded linear maps, matters are more complicated:

7.18 Exercise (i) Let V be a Banach space and $W \subseteq V$ a closed subspace. Prove that the quotient map $p : V \to V/W$ has a bounded section if and only if W is complemented.

(ii) If V, W are Banach spaces and $A \in B(V,W)$, prove that A has a bounded right inverse if and only if A is surjective and $\ker A \subseteq V$ is complemented.

7.19 Exercise Let V, W be Banach spaces and $A \in B(V,W)$. Prove that A has a bounded left inverse if and only if A is injective and $AV \subseteq W$ is complemented.

7.20 Remark Since all closed subspaces of Hilbert spaces are complemented, we find that a bounded map between Hilbert spaces has a bounded right (resp. left) inverse if and only if it is surjective (resp. injective with closed image). □

7.21 Exercise (i) Let V be an infinite-dimensional Banach space. Prove that all finite-dimensional subspaces have empty interior (in V!), then use Baire's theorem to prove that V cannot have a countable Hamel basis. (Thus $\dim V > \aleph_0 = \#\mathbb{N}$.)

(ii) Let V be an infinite-dimensional Banach space and $\{x_n\}, \{\varphi_n\}$ sequences as in Exercise 9.17. For every $N \subseteq \mathbb{N}$ define $x_N = \sum_{n\in N} 2^{-n} x_n$. Now use Lemma B.24 to find a linearly independent family in V of cardinality $\mathfrak{c} = \#\mathbb{R}$, so that $\dim V \geq \mathfrak{c}$ (Hamel dimension).

(iii) Prove that every separable normed space V has cardinality at most $\mathfrak{c}$ and deduce $\dim V \leq \mathfrak{c}$.

(iv) Conclude that every infinite-dimensional separable Banach space has Hamel dimension $\mathfrak{c}$.

7.22 Remark 1. The result of Exercise 7.21(i) can be proven using Riesz' Lemma 12.2 instead of Baire's theorem, see [10], but (as in most such cases) the proof uses countable dependent choice DC_ω like the proof of Baire's theorem.

2. If the continuum hypothesis (CH) is true, Exercise 7.21 (ii) readily follows from (i)+(iii). But the proof of (ii) indicated above is independent of CH. □

7.23 Exercise Give counterexamples showing that both spaces appearing in the Bounded Inverse Theorem must be complete.

Hint: For complete E, incomplete F use ℓ^p spaces, and for E incomplete, F complete use $F = \ell^1(\mathbb{N}, \mathbb{R})$ and the fact that it has Hamel dimension $\mathfrak{c} = \#\mathbb{R}$, cf. Exercise 7.21(iv).

7.3 The Closed Graph Theorem (CGT). Hellinger-Toeplitz

We quickly look at an interesting result equivalent to the Bounded Inverse Theorem, but we will not need it afterwards.

If $f : X \to Y$ is a function, the graph of f is the set $\mathfrak{G}(f) = \{(x, f(x)) \mid x \in X\} \subseteq X \times Y$.

7.24 Exercise Let X be a topological space, Y a Hausdorff space and $f : X \to Y$ continuous. Prove that $\mathfrak{G}(f) \subseteq X \times Y$ is closed.

If E, F are normed spaces, we know that $\|(x, y)\| = \|x\| + \|y\|$ is a norm on $E \oplus F$, complete if E and F are. The projections $p_1 : E \oplus F \to E$, $p_2 : E \oplus F \to F$ are bounded.

7.25 Lemma *Let E, F be normed spaces and $T : E \to F$ a linear map (not assumed bounded). Then the following are equivalent:*

(i) The graph $\mathfrak{G}(T) = \{(x, Tx) \mid x \in E\} \subseteq E \oplus F$ of T is closed.

(ii) Whenever $\{x_n\}_{n\in\mathbb{N}} \subseteq E$ is a sequence such that $x_n \to x \in E$ and $Tx_n \to y \in F$, we have $y = Tx$.

Proof. Since $E \oplus F$ is a metric space, $\mathfrak{G}(T)$ is closed if and only if it contains the limit (x, y) of every sequence $\{(x_n, y_n)\}$ in $\mathfrak{G}(T)$ that converges to some $(x, y) \in E \oplus F$. But a sequence in $\mathfrak{G}(T)$ is of the form $\{(x_n, Tx_n)\}$, and $(x, y) \in \mathfrak{G}(T) \Leftrightarrow y = Tx$. ■

7.26 Remark Operators with closed graph (in particular unbounded ones) are often called closed. But this must not be confused with their closedness as a map, i.e. the property of sending closed sets to closed sets! Bounded linear operators between Banach spaces have closed graphs, but need not be closed maps. □

7.27 Theorem (Banach 1929) *If E, F are Banach spaces, then a linear map $T : E \to F$ is bounded if and only if its graph is closed.*

Proof. Let E, F be Banach spaces, and let $T : E \to F$ be linear. If T is bounded then it is continuous, thus $\mathfrak{G}(T)$ is closed by Exercise 7.24. Now assume that $\mathfrak{G}(T)$ is closed. The Cartesian product $E \oplus F$ with norm $\|(e, f)\| = \|e\| + \|f\|$ is a Banach space. The linear subspace $\mathfrak{G}(T) \subseteq E \oplus F$ is closed by assumption, thus a Banach space. Since the projection $p_1 : \mathfrak{G}(T) \to E$ is a bounded bijection, by Corollary 7.7 it has a bounded inverse $p_1^{-1} : E \to \mathfrak{G}(T)$. Then also $T = p_2 \circ p_1^{-1}$ is bounded. ■

7.28 Exercise Show that the Bounded Inverse Theorem (Corollary 7.7) can be deduced from the Closed Graph Theorem. (Thus the three main results of this section are 'equivalent'.)

We discuss a few typical applications of the CGT. (For another one see Exercise 8.5.)

7.29 Exercise Let $A : E \to F$ be a linear map of Banach spaces. Prove that A is bounded if and only if the following holds: If $\{x_n\}_{n \in \mathbb{N}} \subseteq E$ is such that $x_n \to 0$ and $Ax_n \to y$ then $y = 0$.

7.30 Exercise Let V be a Banach space over $\mathbb{F}$, $p \in [1, \infty]$ and $\{x_n\}_{n \in \mathbb{N}}$ a sequence in V such that for every $\varphi \in V^*$ the sequence $\{\varphi(x_n)\}_{n \in \mathbb{N}}$ is in $\ell^p(\mathbb{N}, \mathbb{F})$. Thus there is a (clearly linear) map $A : V^* \to \ell^p(\mathbb{N}, \mathbb{F})$, $\varphi \mapsto \{\varphi(x_n)\}_{n \in \mathbb{N}}$ (called the analysis map). Prove that A is bounded, thus $A \in B(V^*, \ell^p(\mathbb{N}, \mathbb{F}))$.

7.31 Exercise Let H be a Hilbert space and $\{z_n\}_{n \in \mathbb{N}} \subseteq H$ a sequence such that for each $f \in \ell^2(\mathbb{N}, \mathbb{C})$ there exists an $x \in H$ satisfying $\langle x, z_n \rangle = f(n)\ \forall n$. Prove that there exists D such that for each $f \in \ell^2(\mathbb{N}, \mathbb{C})$ there exists an $x \in H$ satisfying $\langle x, z_n \rangle = f(n)\ \forall n$ and $\|x\| \leq D\|f\|_2$.

Hint: Put $N = \{z_1, z_2, \ldots\}^\perp$, construct a map $\ell^2(\mathbb{N}, \mathbb{C}) \to H/N$ and prove its boundedness.

7.3.1 The Hellinger-Toeplitz theorem

The following is a very typical application of the Closed Graph Theorem:

7.32 Theorem (Hellinger-Toeplitz theorem (1928)) [45]

(i) If H, K are Hilbert spaces and $A : H \to K$, $B : K \to H$ are linear maps satisfying $\langle Ax, y \rangle = \langle x, By \rangle$ for all $x \in H, y \in K$ then A and B are bounded.

(ii) If H is a Hilbert space and a linear map $A : H \to H$ is self-adjoint (i.e. $\langle Ax, y \rangle = \langle x, Ay \rangle$ for all $x, y \in H$) then A is bounded.

Proof. (i) Let $\{x_n\} \subseteq H$ be a sequence converging to $x \in H$ and assume that $Ax_n \to y$. Then

$$\langle Ax, z \rangle = \langle x, Bz \rangle = \lim_{n \to \infty} \langle x_n, Bz \rangle = \lim_{n \to \infty} \langle Ax_n, z \rangle = \langle y, z \rangle \quad \forall z \in K,$$

implying $Ax = y$. Thus A has closed graph and therefore is bounded by Theorem 7.27. The proof for B is analogous. (ii) is just the special case $H = K, A = B$ of (i). ■

[45] Ernst David Hellinger (1883-1950), Otto Toeplitz (1881-1940). German mathematicians. Both were forced into exile in 1939. See also the T.-Hausdorff Theorem B.143.

7.33 Remark The Hellinger-Toeplitz Theorem shows that on a Hilbert space H there are no unbounded linear operators $A : H \to H$ satisfying $\langle Ax, y\rangle = \langle x, Ay\rangle \ \forall x, y$. This is a typical example of a 'no-go-theorem'. Occasionally such results are a nuisance. After all, the operator of multiplication by n on $\ell^2(\mathbb{N})$ 'obviously' is self-adjoint. What Hellinger-Toeplitz really says is that such an operator cannot be defined everywhere, i.e. on all of H. This leads to the notion of symmetric operators, and also illustrates that no-go theorems often can be circumvented by generalizing the setting. This is the case here, since the Hellinger-Toeplitz theorem only applies to operators that are defined everywhere. □

7.34 Definition *A symmetric operator on a Hilbert space H is a linear map $A : D \to H$, where $D \subseteq H$ is a dense linear subspace, that satisfies $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in D$,*

7.35 Exercise Let $H = \ell^2(\mathbb{N}, \mathbb{C}), D = \{f \in \ell^2(\mathbb{N}, \mathbb{C}) \mid \sum_n |nf(n)|^2 < \infty\} \subseteq H$ and $A : D \to H, \ (Af)(n) = nf(n)$. Prove:

(i) $D \subseteq H$ is a dense proper linear subspace.

(ii) $A : D \to H$ is symmetric and unbounded.

There is an extensive theory of unbounded linear operators defined on dense subspaces of a Hilbert space. Most books on (linear) functional analysis have a chapter on them, e.g., [118, 128, 30, 141]. This subject is very important for applications to differential equations and quantum mechanics, but since it is quite technical one should probably not approach it before one has mastered the material of this course.

7.4 Boundedness below. Invertibility

7.36 Definition *Let V, W be normed spaces and let $A : V \to W$ be a linear map. Then A is called bounded below*[46] *if there is a $\delta > 0$ such that $\|Ax\| \geq \delta\|x\| \ \forall x \in V$.*

(Equivalently, $\inf_{\|x\|=1} \|Ax\| > 0$.)

Boundedness below of a map clearly implies injectivity, but the converse is not true. E.g., $A \in B(\ell^2(\mathbb{N}, \mathbb{F}))$ defined by $(Af)(n) = f(n)/n$ is injective, but not bounded below.

7.37 Exercise Let V, W be be normed spaces, where V is finite-dimensional, and let $A : V \to W$ be an injective linear map. Prove that A is bounded below.

7.38 Lemma *Let V, W be normed spaces and $A : V \to W$ a linear bijection. Then*

$$\|A^{-1}\| = \left(\inf_{\|x\|=1} \|Ax\|\right)^{-1}.$$

In particular, A is bounded below if and only if its (set-theoretic) inverse A^{-1} is bounded.

Proof. Using the bijectivity of $x \mapsto Ax$, we have

$$\|A^{-1}\| = \sup_{y \in W\backslash\{0\}} \frac{\|A^{-1}y\|}{\|y\|} = \sup_{x \in V\backslash\{0\}} \frac{\|x\|}{\|Ax\|} = \left(\inf_{x \in V\backslash\{0\}} \frac{\|Ax\|}{\|x\|}\right)^{-1} = \left(\inf_{\|x\|=1} \|Ax\|\right)^{-1}.$$

[46] This terminology clashes with another one according to which a self-adjoint operator A is bounded below if $\sigma(A) \subseteq [c, \infty)$ for some $c \in \mathbb{R}$. Since we consider only bounded operators, we'll have no use for this notion. The problem could be avoided by writing 'bounded away from zero', as some authors do, but this is a bit tedious.

The second statement follows immediately. ■

Recall that the image $AV \subseteq W$ of a linear map $A : V \to W$ need not be closed (if V is infinite-dimensional). The following generalizes Corollary 3.23:

7.39 Lemma *If V is a Banach space, W is a normed space and $A : V \to W$ is a linear map that is bounded and bounded below then its image $AV \subseteq W$ is closed.*

Proof. Let $y \in \overline{AV}$. Then there is a sequence $\{y_n\} \subseteq AV$ converging to y. Pick $\{x_n\} \subseteq V$ with $Ax_n = y_n \ \forall n$. (They are unique, but this is not needed.) The sequence $\{y_n\}$ is Cauchy, and the same holds for $\{x_n\}$ by $\|x\| \leq \delta^{-1}\|Ax\|$. Since V is complete, x_n converges to some $x \in V$. With boundedness, thus continuity, of A we have $y = \lim y_n = \lim Ax_n = Ax \in AV$, so that AV is closed. ■

7.40 Definition *If V, W are normed spaces then $A \in B(V, W)$ is called invertible if there is a $B \in B(W, V)$ such that $BA = \mathrm{id}_V$ and $AB = \mathrm{id}_W$.*

7.41 Proposition *Let V, W be a Banach spaces and $A \in B(V, W)$. Then the following are equivalent:*

(i) A is invertible.

(ii) A is injective and surjective.

(iii) A is bounded below and has dense image.

Proof. (i)⇒(ii)+(iii). It is clear that invertibility implies injectivity and surjectivity, thus in particular dense image. Since A^{-1} is bounded, Lemma 7.38 gives that A is bounded below.

(ii)⇒(i) The set-theoretic inverse which exists by bijectivity, clearly linear, is bounded by the bounded inverse theorem (Corollary 7.7). Thus A is invertible.

(iii)⇒(i). By boundedness below, A is injective. And $AV \subseteq W$ is dense by assumption and closed by Lemma 7.39, thus $AV = W$. Thus A is injective and surjective. Now boundedness of the inverse A^{-1} follows from boundedness below of A and Lemma 7.38. Note: BIT not used. ■

7.42 Remark 1. Note that dense image is weaker than surjectivity, while boundedness below is stronger than injectivity. The point of criterion (iii) is that it can be quite hard to verify surjectivity of A directly, while density of the image usually is easier to establish.

2. The material on bounded below maps discussed so far, including (i)⇔(iii) in Proposition 7.41, was entirely elementary and could be moved to Section 3. □

7.43 Exercise Let V, W be Banach spaces and $A \in B(V, W)$. Prove:

(i) The following are equivalent:

(α) A is bounded below.

(β) A is injective and $AV \subseteq W$ is closed.

(γ) $\widetilde{A} : V \to AV$ is an isomorphism.

(ii) If $\ker A$ has a complement Z then $AV \subseteq W$ is closed ⇔ $A_{|Z}$ is bounded below.

(iii) If V is a Hilbert space then $AV \subseteq W$ is closed if and only if $A \upharpoonright (\ker A)^{\perp}$ is bounded below.

7.44 Exercise Let V be an infinite dimensional Banach space and $A \in B(V)$ injective.

(i) Putting $W = V \oplus V$ with norm $\|(a,b)\| = \|a\| + \|b\|$, prove that the subspaces

$$K = \{(x,0) \mid x \in V\} \subseteq W \qquad \text{and} \qquad L = \{(x,Ax) \mid x \in V\} \subseteq W$$

are infinite-dimensional, closed and $K \cap L = \{0\}$.

(ii) Prove that $K + L = \{(x, Ay) \mid x, y \in V\}$.

(iii) Prove that $K + L \subseteq W$ is non-closed when A is not bounded below.

(iv) For A not bounded below, prove that the unit spheres of K and L have distance zero (as they must by Exercise 7.13(ii)).

7.45 Exercise Let H be a Hilbert space and $A \in B(H)$ such that $|\langle Ax, x\rangle| \geq C\|x\|^2$ for some $C > 0$. Prove that A is invertible and $\|A^{-1}\| \leq C^{-1}$. (Such an A is called elliptic or coercive.)

7.46 Exercise Let V be a Banach space. Prove that $\{A \in B(V) \mid A \text{ bounded below }\} \subseteq B(V)$ is open.

8 Uniform Boundedness Theorem and applications

8.1 The Uniform Boundedness Theorem (UBT)

The Open Mapping Theorem concerned a single bounded operator. Now we consider families:

8.1 Definition *Let E, F be normed spaces and $\mathcal{F} \subseteq B(E,F)$ a family of bounded linear maps.*

- *$\mathcal{F}$ is called pointwise bounded if $\sup_{A\in\mathcal{F}} \|Ax\| < \infty$ for each $x \in E$.*
- *$\mathcal{F}$ is called uniformly bounded if $\sup_{A\in\mathcal{F}} \|A\| < \infty$.*

That uniform boundedness of $\mathcal{F}$ implies pointwise boundedness is trivial, but this is not:

8.2 Theorem *Let E be a Banach space, F a normed space and $\mathcal{F} \subseteq B(E,F)$. Then:*

(i) Either $\mathcal{F}$ is uniformly bounded or the set $\{x \in E \mid \sup_{A\in\mathcal{F}} \|Ax\| = \infty\} \subseteq E$ is dense G_δ[47].

(ii) [Helly 1912, Hahn, Banach 1922] If $\mathcal{F}$ is pointwise bounded then it is uniformly bounded.

Proof. (i) The map $F \to \mathbb{R}_{\geq 0},\ x \mapsto \|x\|$ is continuous and each $A \in \mathcal{F}$ is bounded, thus continuous. Therefore the map $f_A : E \to \mathbb{R}_{\geq 0},\ x \mapsto \|Ax\|$ is continuous for every $A \in \mathcal{F}$. Defining for each $n \in \mathbb{N}$

$$V_n = \{x \in E \mid \sup_{A\in\mathcal{F}} \|Ax\| > n\},$$

the definition of sup implies

$$V_n = \{x \in E \mid \exists A \in \mathcal{F} : \|Ax\| > n\} = \bigcup_{A\in\mathcal{F}} \{x \in E \mid \|Ax\| > n\} = \bigcup_{A\in\mathcal{F}} f_A^{-1}((n,\infty)),$$

which is open by continuity of the functions f_A.

Thus $X = \bigcap_{n\in\mathbb{N}} V_n$ is G_δ. And $X = \{x \in E \mid \sup_{A\in\mathcal{F}} \|Ax\| = \infty\}$ in view of the definition of the V_n. Now Baire's Theorem A.23 implies that X is dense if all the V_n are.

[47] A G_δ set in a topological space is an intersection of countably many open sets.

On the other hand, if V_n is non-dense for some $n \in \mathbb{N}$, there exists $x_0 \in E$ and $r > 0$ such that $B(x_0, r) \cap V_n = \emptyset$. This means $\sup_{A\in\mathcal{F}} \|A(x_0 + x)\| \le n$ for all $x \in E$ with $\|x\| < r$. With $x = (x_0 + x) - x_0$ and the triangle inequality we have

$$\|Ax\| \le \|A(x_0 + x)\| + \|Ax_0\| \le 2n \qquad \forall A \in \mathcal{F},\ x \in B(0, r).$$

This implies $\|A\| \le 2n/r$ for all $A \in \mathcal{F}$, thus $\mathcal{F}$ is uniformly bounded.

(ii) If $\mathcal{F}$ is not uniformly bounded then by (i) it clearly is not pointwise bounded. ■

8.3 Remark 1. We call (i) the strong and (ii) the weak version of the Uniform Boundedness Theorem, respectively. Mystifyingly, most expositions of uniform boundedness use Baire's theorem to prove the weak version without pointing out that the proof actually gives a much stronger result. One of the few exceptions is [140], the source of the elegant proof given above.

2. There is an incessant stream of publications purporting to give 'elementary' proofs (i.e. avoiding Baire) of Theorem 8.2(ii), but *all* of them (except [53]) use, usually without acknowledging it, the axiom DC_ω of countable dependent choice which, however, is logically equivalent over ZF to Baire's theorem! (See [17].) This also holds for [154], but the proof given there admits a tiny modification, discovered quite recently [53], that really only uses countable choice AC_ω. See Appendix B.4 for the beautiful argument (employing a version of the 'method of the gliding hump') that really is more elementary, in the precise sense of reverse mathematics, which concerns itself with the weakest axioms needed to prove desired results. (Cf. [158] for an engaging introduction.) □

8.4 Exercise (i) Deduce Theorem 8.2(ii) from its special case where E, F are both assumed complete.

(ii) Show that the completeness assumption on E in Theorem 8.2(ii) cannot be omitted.

The weak version of the UBT can also be deduced from the closed graph theorem:[48]

8.5 Exercise Let E, F be Banach spaces and $\mathcal{F} \subseteq B(E, F)$ a pointwise bounded family. Use the Closed Graph Theorem to prove that $\mathcal{F}$ is uniformly bounded, as follows:

(i) Prove that $F_\mathcal{F} = \{\{y_A\}_{A\in\mathcal{F}} \in F^\mathcal{F} = \mathrm{Fun}(\mathcal{F}, F) \mid \sup_{A\in\mathcal{F}} \|y_A\| < \infty\}$ is a Banach space.

(ii) Show that pointwise boundedness of $\mathcal{F}$ is equivalent to $TE \subseteq F_\mathcal{F}$.

(iii) Prove that the graph $\mathfrak{G}(T) \subseteq E \oplus F_\mathcal{F}$ of T is closed. (Thus T is bounded by Theorem 7.27.)

(iv) Deduce uniform boundedness of $\mathcal{F}$ from the boundedness of T.

(v) Remove the requirement that F be complete.

8.2 Applications of the weak UBT. Banach-Steinhaus

8.6 Exercise Let X, Y be Banach spaces, Z a normed space and $T : X \times Y \to Z$ a map such that $Y \to Z,\ y \mapsto T(x, y)$ and $X \to Z,\ x \mapsto T(x, y)$ are linear and bounded for each $x \in X$ and $y \in Y$, respectively. Prove that there is a $0 \le C < \infty$ such that $\|T(x, y)\| \le C\|x\|\|y\|$ for all $x \in X, y \in Y$.

[48] While this approach can be found in various books, it is unsatisfactory. After all, the weak UBT can be proven quite directly using only AC_ω, cf. Appendix B.4, whereas the proof via the CGT, besides being quite indirect, relies on Baire's theorem, without however yielding the strong UBT.

8.7 EXERCISE Give a proof of the Hellinger-Toeplitz Theorem 7.32 using Theorem 8.2(ii) instead of the Closed Graph Theorem 7.27. (This is interesting since the it shows that also Hellinger-Toeplitz depends only on the axiom AC_ω of countable choice.)

8.8 DEFINITION *Let E, F be normed spaces. A sequence (or net) $\{A_n\} \subseteq B(E,F)$ is strongly convergent if $\lim_{n\to\infty} A_n x$ exists for every $x \in E$.*

Under the above assumption, the map $A : E \to F, x \mapsto \lim_{n\to\infty} A_n x$ is easily seen to be linear. Now we write $A_n \stackrel{s}{\to} A$ or $A = s\text{-}\lim A_n$.

8.9 COROLLARY (BANACH-STEINHAUS) [49][50] *If E is a Banach space, F a normed space and the sequence $\{A_n\} \subseteq B(E,F)$ is strongly convergent then the map $A = s\text{-}\lim A_n$ is bounded, thus in $B(E,F)$.*

Proof. The convergence of $\{A_n x\} \subseteq F$ for each $x \in E$ implies boundedness of $\{A_n x \mid n \in \mathbb{N}\}$ for each x, so that $\mathcal{F} = \{A_n \mid n \in \mathbb{N}\} \subseteq B(E,F)$ is pointwise bounded and therefore uniformly bounded by Theorem 8.2(ii). Thus there is T such that $\|A_n\| \le T \ \forall n$, so that $\|A_n x\| \le T\|x\| \ \forall x \in E, n \in \mathbb{N}$. With $A_n x \to Ax$ this implies $\|Ax\| = \lim \|A_n x\| \le T\|x\|$ for all x, thus $\|A\| \le T < \infty$. ■

8.10 REMARK Clearly $A_n \stackrel{s}{\to} A$ is equivalent to $\|A_n - A\|_x \to 0$ for all $x \in E$, where $\|A\|_x := \|Ax\|$ is a seminorm on $B(E,F)$ for each $x \in E$. If $\|A\|_x = 0$ for all $x \in E$ then $Ax = 0 \ \forall x \in E$, thus $A = 0$. Thus the family $\mathcal{F} = \{\|\cdot\|_x \mid x \in E\}$ is separating and induces a locally convex topology on $B(E,F)$, the strong operator topology τ_{sot}. Norm convergence $\|A_n - A\| \to 0$ clearly implies strong convergence $A_n \stackrel{s}{\to} A$, but usually the strong (operator) topology is strictly weaker (despite its name) than the norm topology. See the following exercise for an example. □

8.11 EXERCISE Let H be a Hilbert space and $\{e_1, e_2, \ldots\}$ an orthonormal sequence in H (not necessarily an ONB). Define $\varphi_n \in H^* = B(H,\mathbb{F})$ by $\varphi_n(x) = \langle x, e_n\rangle$. Prove that $\varphi_n \to 0$ strongly, but not in norm.

8.12 EXERCISE Let $1 \le p < \infty$ and $V = \ell^p(\mathbb{N},\mathbb{F})$. For each $m \in \mathbb{N}$ define $P_m \in B(V)$ by $(P_m f)(n) = f(n)$ for $n \ge m$ and $(P_m f)(n) = 0$ if $n < m$. Prove $P_m \stackrel{s}{\to} 0$, but $\|P_m\| = 1 \ \forall m$, thus $P_m \stackrel{\|\cdot\|}{\not\to} 0$.

8.13 EXERCISE Let V be a separable Banach space and $\mathcal{B} \subseteq B(V)$ a bounded subset.

(i) Prove: If $S \subseteq V$ is dense and a net $\{A_\iota\} \subseteq \mathcal{B}$ satisfies $\|A_\iota x\| \to 0$ for all $x \in S$ then $\|A_\iota x\| \to 0$ for all $x \in V$, thus $A_\iota \to 0$ in the strong operator topology.

(ii) Prove that the topological space $(\mathcal{B}, \tau_{\text{sot}})$ is metrizable.

(iii) BONUS: Prove that (V, τ_{sot}) is not metrizable if V is infinite-dimensional.

[49] Hugo Steinhaus (1887-1972). Polish mathematician

[50] In the literature, one can find either this result or Theorem 8.2(ii) denoted as 'Banach-Steinhaus theorem'.

8.3 Appl. of the strong UBT: Many continuous functions with divergent Fourier series

For the preceding applications of the uniform boundedness theorem, the weak version was sufficient. Other applications use the contraposition for a (non-constructive) existence proof: If $\mathcal{F} \subseteq B(E,F)$ is not uniformly bounded then it is not pointwise bounded, thus there exists $x \in E$ with $\sup_{A\in\mathcal{F}} \|Ax\| = \infty$. For such applications, Theorem 8.2(i) is a definite improvement.

Let $f : \mathbb{R} \to \mathbb{C}$ be 2π-periodic, i.e. $f(x+2\pi) = f(x)\ \forall x$, and integrable over finite intervals. Define

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx \tag{8.1}$$

and

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)e^{ikx}, \quad n \in \mathbb{N}. \tag{8.2}$$

The fundamental problem of the theory of Fourier series is to find conditions for the convergence $S_n(f)(x) \to f(x)$ as $n \to \infty$, where convergence can be understood as (possibly almost) everywhere pointwise or w.r.t. some norm, like $\|\cdot\|_2$ (as in Example 5.49) or $\|\cdot\|_\infty$. Here we will discuss only continuous functions and we identify continuous 2π-periodic functions with continuous functions on S^1. It is not hard to show that $S_n(f)(x) \to f(x)$ if f is differentiable at x (or just Hölder continuous: $|f(x') - f(x)| \le C|x'-x|^D$ with $C, D > 0$ for x' near x) and that convergence is uniform when f is continuously differentiable (or the Hölder condition holds uniformly in x, x'). (See any number of books on Fourier analysis, e.g. [157, 83].)

Assuming only continuity of f one can still prove that $\lim_{n\to\infty} S_n(f)(x) = f(x)$ if the limit exists, but there actually exist continuous functions f such that $S_n(f)(x)$ diverges at some x. Such functions were first constructed in the 1870s using 'condensation of singularities', a relative and precursor of the gliding hump method, cf. Appendix B.3.5. Nowadays, most textbook presentations of such functions are based on Lemma 8.15 below combined with either the uniform boundedness theorem or constructions 'by hand', see e.g. [83, Section II.2], that are quite close in spirit to the uniform boundedness method.

However, individual examples of continuous functions with Fourier series divergent in a point can be produced in a totally constructive fashion, avoiding all choice axioms! (See [109] for a very classical example.) But using non-constructive arguments seems unavoidable if one wants to prove that there are many such functions as in the following:

8.14 Theorem *There is a subset $X \subseteq C(S^1)$ that is dense G_δ (in the $\|\cdot\|_\infty$-topology) such that the Fourier series $\{S_n(f)(0)\}_{n\in\mathbb{N}}$ diverges for each $f \in X$.*

Proof. Inserting (8.1) into (8.2) we obtain

$$S_n(f)(x) = \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx}\int_0^{2\pi} f(t)e^{-ikt}\,dt = \frac{1}{2\pi}\int_0^{2\pi} f(t)\left(\sum_{k=-n}^{n} e^{ik(x-t)}\right)dt = (D_n \star f)(x),$$

where $\star$ denotes convolution, defined for 2π-periodic f, g by $(f\star g)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(t)g(x-t)dt$, and

$$D_n(x) := \sum_{k=-n}^{n} e^{ikx} = \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}$$

is the Dirichlet kernel. The quickest way to check the last identity is the 'telescoping' calculation

$$(e^{ix/2}-e^{-ix/2})D_n(x)=\sum_{k=-n}^{n}e^{ix(k+1/2)}-\sum_{k=-n}^{n}e^{ix(k-1/2)}=e^{ix(n+1/2)}-e^{-ix(n+1/2)},$$

together with $e^{ix}-e^{-ix}=2i\sin x$. Since $D_n(x)$ is an even function, we have

$$\varphi_n(f):=S_n(f)(0)=\frac{1}{2\pi}\int_0^{2\pi}f(x)D_n(x)dx.$$

It is clear that the norm of the map $\varphi_n:(C(S^1),\|\cdot\|_\infty)\to\mathbb{C}$ is bounded above by $\|D_n\|_1$. For $g_n(x)=\operatorname{sgn}(D_n(x))$ we have $\varphi_n(g_n)=(2\pi)^{-1}\int_0^{2\pi}|D_n(x)|dx=:\|D_n\|_1$. While g_n is not continuous, we can find a sequence of continuous $g_{n,m}$ bounded by 1 such that $g_{n,m}\stackrel{m\to\infty}{\longrightarrow}g_n$ pointwise. Now Lebesgue's dominated convergence theorem implies $\varphi_n(g_{n,m})\to\varphi_n(g_n)=\|D_n\|_1$, which implies $\|\varphi_n\|=\|D_n\|_1$. By Lemma 8.15 below, $\|D_n\|_1\to\infty$ as $n\to\infty$. Thus the family $\mathcal{F}=\{\varphi_n\}\subseteq B(C(S^1),\mathbb{C})$ is not uniformly bounded. Now Theorem 8.2(i) implies that the set $X=\{f\in C(S^1,\mathbb{C})\mid\{S_n(f)(0)\}\text{ is unbounded}\}$ is dense G_δ. ■

8.15 Lemma *We have* $\|D_n\|_1\geq\frac{4}{\pi^2}\log n$ *for all* $n\in\mathbb{N}$.

Proof. Using $|\sin x|\leq|x|$ for all $x\in\mathbb{R}$, we compute

$$\begin{aligned}\|D_n\|_1 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_n(x)|dx\geq\frac{2}{\pi}\int_0^{\pi}\left|\sin\left(n+\frac{1}{2}\right)x\right|\frac{dx}{x}\\ &= \frac{2}{\pi}\int_0^{(n+1/2)\pi}|\sin x|\frac{dx}{x}\geq\frac{2}{\pi}\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi}\frac{|\sin x|}{x}dx\\ &\geq \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int_0^{\pi}\sin x\,dx=\frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k}\geq\frac{4}{\pi^2}\log n,\end{aligned}$$

where we used $\sum_{k=1}^n 1/k\geq\int_1^{n+1}dx/x=\log(n+1)>\log n$. ■

8.16 Remark Also the Bounded Inverse Theorem has an interesting application to Fourier analysis: For $f\in L^1([0,2\pi])$, we define the Fourier coefficients $\widehat{f}(n)=(2\pi)^{-1}\int_0^{2\pi}f(t)e^{-int}dt$ for all $n\in\mathbb{N}$. It is immediate that $\|\widehat{f}\|_\infty\leq\|f\|_1$, and is not hard to prove the Riemann-Lebesgue theorem $\widehat{f}\in c_0(\mathbb{Z},\mathbb{C})$ and injectivity of the resulting map $L^1([0,2\pi])\to c_0(\mathbb{Z},\mathbb{C}),f\mapsto\widehat{f}$, see e.g. [140, Theorem 5.15] or [83]. If this map was surjective, the Bounded Inverse Theorem would give $\|f\|_1\leq C\|\widehat{f}\|_\infty\ \forall f\in L^1([0,2\pi])$. For the Dirichlet kernel it is immediate that $\widehat{D_n}(m)=\chi_{[-n,n]}(m)$, thus $\|\widehat{D_n}\|_\infty=1$ for all $n\in\mathbb{N}$. Since we know that $\|D_n\|_1\to\infty$, we would have a contradiction. Thus $L^1([0,2\pi])\to c_0(\mathbb{Z},\mathbb{C}),f\mapsto\widehat{f}$ is not surjective. □

9 Duality: Hahn-Banach Theorem and applications

We have seen that every bounded linear functional $\varphi\in H^*$, where H is a Hilbert space, is of the form $\varphi=\varphi_y$ for a certain (unique) $y\in H$. Thus dual spaces of Hilbert spaces are completely understood. (The map $H\to H^*$, $y\mapsto\varphi_y$ is an anti-linear bijection.) For a general Banach space V, matters are much more complicated. The point of the Hahn[51]-Banach theorem

[51]Hans Hahn (1879-1934). Austrian mathematician who mostly worked in analysis and topology.

(which comes in many versions)[52] is to show that all Banach spaces admit many bounded linear functionals.

9.1 First version of Hahn-Banach over $\mathbb{R}$

We begin with a slight generalization of the notion of seminorms:

9.1 DEFINITION *If V is a real vector space, a map $p : V \to \mathbb{R}$ is called sublinear if it satisfies*

- *Positive homogeneity: $p(cv) = cp(v)$ for all $v \in V$ and $c > 0$.*
- *Subadditivity: $p(x + y) \le p(x) + p(y)$ for all $x, y \in V$.*

9.2 THEOREM *Let V be a real vector space and $p : V \to \mathbb{R}$ a sublinear function. Let $W \subseteq V$ be a linear subspace and $\varphi : W \to \mathbb{R}$ a linear functional such that $\varphi(w) \le p(w)$ for all $w \in W$. Then there is a linear functional $\widehat{\varphi} : V \to \mathbb{R}$ such that $\widehat{\varphi} \upharpoonright W = \varphi$ and $\widehat{\varphi}(v) \le p(v)$ for all $v \in V$.*

The heart of the proof is the special case where W has codimension one:

9.3 LEMMA *Let V, p, W, φ be as in Theorem 9.2 and $v' \in V$. Then there is a linear functional $\widehat{\varphi} : Y = W + \mathbb{R}v' \to \mathbb{R}$ such that $\widehat{\varphi} \upharpoonright W = \varphi$ and $\widehat{\varphi}(v) \le p(v)$ for all $v \in Y$.*

Proof. If $v' \in W$, there is nothing to do so that we may assume $v' \in V \backslash W$. Then every $x \in W + \mathbb{R}v'$ can be written as $x = w + cv'$ with unique $w \in W, c \in \mathbb{R}$. Thus if $d \in \mathbb{R}$, we can define $\widehat{\varphi} : W + \mathbb{R}v' \to \mathbb{R}$ by $w + cv' \mapsto \varphi(w) + cd$ for all $w \in W$ and $c \in \mathbb{R}$. Since $\widehat{\varphi}$ is linear and trivially satisfies $\widehat{\varphi} \upharpoonright W = \varphi$, it remains to show that d can be chosen such that $\varphi \le p$ holds on $Y = W + \mathbb{R}v'$, to wit

$$\widehat{\varphi}(w + cv') = \varphi(w) + cd \le p(w + cv') \qquad \forall w \in W, c \in \mathbb{R}. \tag{9.1}$$

For $c = 0$, this holds by assumption. If (9.1) holds for all $w \in W$ and $c \in \mathbb{R}$ then in particular

$$\varphi(w) \pm d \le p(w \pm v') \quad \forall w \in W. \tag{9.2}$$

And if (9.2) holds then by linearity of φ and positive homogeneity of p, for all $e > 0$ we have

$$\widehat{\varphi}(w \pm ev') = e\widehat{\varphi}(e^{-1}w \pm v') \overset{(9.2)}{\le} ep(e^{-1}w \pm v') = p(w \pm ev'),$$

thus the desired inequality (9.1) holds for all $w \in W, c \in \mathbb{R}$. Now (9.2) is equivalent to

$$\varphi(w) - p(w - v') \le d \le p(w' + v') - \varphi(w') \qquad \forall w, w' \in W.$$

Clearly this is possible if and only if $\varphi(w) - p(w - v') \le p(w' + v') - \varphi(w')$ for all $w, w' \in W$, which in turn is equivalent to $\varphi(w) + \varphi(w') \le p(w - v') + p(w' + v')\ \forall w, w'$. The latter inequality is indeed satisfied for all $w, w' \in W$ since $w + w' \in W$, so that

$$\varphi(w) + \varphi(w') = \varphi(w + w') \le p(w + w') = p((w - v') + (w' + v')) \le p(w - v') + p(w' + v'),$$

which holds since φ is additive and bounded by p on W and since p is subadditive. ■

[52] Important early results are due to Eduard Helly (1884-1943), another Austrian mathematician. See [101, p. 54-55].

If $n = \dim(V/W) < \infty$, proving the theorem amounts to applying the Lemma n times. But otherwise an infinite inductive procedure is required. This can be formalized via 'transfinite induction', but is easier to invoke Zorn's lemma, as in the proof of Proposition 5.36:

Proof of Theorem 9.2. If $W = V$, there is nothing to do, so assume $W \subsetneqq V$. Let $\mathcal{E}$ be the set of pairs (Z, ψ), where $Z \subseteq V$ is a linear subspace space containing W and $\psi : Z \to \mathbb{R}$ is a linear map extending φ such that $\psi(z) \leq p(z)\ \forall z \in Z$. Since $W \neq V$, Lemma 9.3 implies $\mathcal{E} \neq \emptyset$.

We define a partial ordering on $\mathcal{E}$ by $(Z, \psi) \leq (Z', \psi') \ \Leftrightarrow\ Z \subseteq Z',\ \psi' \restriction Z = \psi$. If $\mathcal{C} \subseteq \mathcal{E}$ is a chain, i.e. totally ordered by $\leq$, let $Y = \bigcup_{(Z,\psi)\in\mathcal{C}} Z$ and define $\psi_Y : Y \to \mathbb{R}$ by $\psi_Y(v) = \psi(v)$ for any $(Z, \psi) \in \mathcal{C}$ with $v \in Z$. This clearly is consistent and gives a linear map. Now (Y, ψ_Y) is an element of $\mathcal{E}$ and an upper bound for $\mathcal{C}$. Thus by Zorn's lemma there is a maximal element (Y_M, ψ_M) of $\mathcal{E}$. Now $\psi_M : Y_M \to \mathbb{R}$ is an extension of φ satisfying $\psi_M(y) \leq p(y)$ for all $y \in Y_M$, so we are done if we prove $Y_M = V$. If this is not the case, we can pick $v' \in V \backslash Y_M$ and use Lemma 9.3 to extend ψ_Y to $Y_M + \mathbb{R}v'$, but this contradicts the maximality of (Y_M, ψ_M). ■

9.4 Remark The above proof is even more non-constructive than the preceding ones in that it uses Zorn's lemma, which is equivalent to the Axiom of Choice (AC)[53]. If V is a separable normed space, we can replace AC by the weaker axiom DC_ω (countable dependent choice), cf. Exercise 9.8. But even in the generality of all Banach spaces there is the seldom cited fact that the Hahn-Banach theorem can be deduced over ZF from the restriction of Tychonov's theorem to Hausdorff spaces, which is strictly weaker than AC. See Appendix B.6.1. □

9.2 Hahn-Banach theorem for (semi)normed spaces

With the exception of Section B.6.3 we will not use Theorem 9.2 directly, but only the following consequence:

9.5 Theorem (Hahn-Banach Theorem (1927/9)) *Let V be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, p a seminorm on it, $W \subseteq V$ a linear subspace and $\varphi : W \to \mathbb{C}$ a linear functional such that $|\varphi(w)| \leq p(w)$ for all $w \in W$. Then there is a linear functional $\widehat{\varphi} : V \to \mathbb{C}$ such that $\widehat{\varphi} \restriction W = \varphi$ and $|\widehat{\varphi}(v)| \leq p(v)$ for all $v \in V$.*

Proof. $\mathbb{F} = \mathbb{R}$: This is an immediate consequence of Theorem 9.2 since a seminorm p is sublinear with the additional property $p(-v) = p(v) \geq 0$ for all v. In particular, $-\widehat{\varphi}(v) = \widehat{\varphi}(-v) \leq p(-v) = p(v)$, so that $-p(v) \leq \widehat{\varphi}(v) \leq p(v)$ for all $v \in V$, which is equivalent to $|\widehat{\varphi}(v)| \leq p(v)\ \forall v$.

$\mathbb{F} = \mathbb{C}$[54]: Assume $V \supseteq W \stackrel{\varphi}{\to} \mathbb{C}$ satisfies $|\varphi(w)| \leq p(w)\ \forall w \in W$. Define $\psi : W \to \mathbb{R}, w \mapsto \mathrm{Re}(\varphi(w))$, which clearly is $\mathbb{R}$-linear and satisfies the same bound. Thus by the real case just considered, there is an $\mathbb{R}$-linear functional $\widehat{\psi} : V \to \mathbb{R}$ extending ψ such that $|\widehat{\psi}(v)| \leq p(v)$ for all $v \in V$. Define $\widehat{\varphi} : V \to \mathbb{C}$ by

$$\widehat{\varphi}(v) = \widehat{\psi}(v) - i\widehat{\psi}(iv).$$

Again it is clear that $\widehat{\varphi}$ is $\mathbb{R}$-linear. Furthermore

$$\widehat{\varphi}(iv) = \widehat{\psi}(iv) - i\widehat{\psi}(-v) = \widehat{\psi}(iv) + i\widehat{\psi}(v) = i(\widehat{\psi}(v) - i\widehat{\psi}(iv)) = i\widehat{\varphi}(v),$$

[53] "Such reliance on awful non-constructive results is unfortunately typical of traditional functional analysis." [92]

[54] This was discovered only in 1938 by Henri Frederic Bohnenblust (1906-2000) and Andrew Florian Sobczyk (1915-1981), Swiss resp. Polish born American mathematicians.

proving that $\widehat{\varphi} : V \to \mathbb{C}$ is $\mathbb{C}$-linear. If $w \in W$ then

$$\begin{aligned}\widehat{\varphi}(w) &= \widehat{\psi}(w) - i\widehat{\psi}(iw) = \psi(w) - i\psi(iw) = \mathrm{Re}(\varphi(w)) - i\mathrm{Re}(\varphi(iw)) \\ &= \mathrm{Re}(\varphi(w)) - i\mathrm{Re}(i\varphi(w)) = \mathrm{Re}(\varphi(w)) + i\mathrm{Im}(\varphi(w)) = \varphi(w),\end{aligned}$$

so that $\widehat{\varphi}$ extends φ.

Given $v \in V$, let $\alpha \in \mathbb{C}$, $|\alpha| = 1$ be such that $\alpha\widehat{\varphi}(v) \geq 0$. Then $\alpha\widehat{\varphi}(v) = \widehat{\varphi}(\alpha v) = \mathrm{Re}(\widehat{\varphi}(\alpha v)) = \widehat{\psi}(\alpha v)$, so that $|\widehat{\varphi}(v)| = |\alpha\widehat{\varphi}(v)| = \widehat{\psi}(\alpha v) \leq p(\alpha v) = p(v)$. ■

9.6 Remark In Exercise 5.35 we saw (with a fairly easy proof) that bounded linear functionals defined on linear subspaces of Hilbert spaces always have unique norm-preserving extensions to the whole space. For a general Banach space V this uniqueness is far from true! (It holds if and only if V^* is strictly convex, cf. Section B.6.7 for definition and proof.) □

9.7 Exercise Give an example for a Banach space V, a linear subspace $W \subseteq V$ and $\varphi \in W^*$ such that there are multiple norm-preserving extensions $\widehat{\varphi} \in V^*$ of φ.

9.8 Exercise Give a proof of Theorem 9.5 in the case of a separable normed space $(V, \|\cdot\|)$ using only the axiom DC_ω of countable dependent choice (thus neither AC nor Zorn).

9.3 First applications of Hahn-Banach

9.9 Proposition *Let V be a normed space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.*

(i) For every $0 \neq x \in V$ there is a $\varphi \in V^*$ with $\|\varphi\| = 1$ such that $\varphi(x) = \|x\|$.

(ii) For each $x \in V$ we have $\|x\| = \sup_{\varphi \in V^*, \|\varphi\|=1} |\varphi(x)|$.

(iii) 'V^* separates the points of V': If $x, x' \in V$ and $\varphi(x) = \varphi(x')$ $\forall \varphi \in V^*$ then $x = x'$.

(iv) If $x \in V$ then $\widehat{x} : V^* \to \mathbb{F}$, $\varphi \mapsto \varphi(x)$ is in V^{**} with $\|\widehat{x}\| = \|x\|$. The map $\iota_V : V \to V^{**}$, $x \mapsto \widehat{x}$ is an isometric embedding.

(v) The image[55] $\iota_V(V) \subseteq V^{**}$ is closed if and only if V is complete (i.e. Banach).

Proof. (i) Let $W = \mathbb{F}x \subseteq V$. The linear functional $\varphi : W \to \mathbb{F}$, $cx \mapsto c\|x\|$ is isometric since $|\varphi(x)| = \|x\|$, thus $\|\varphi\| = 1$. By the Hahn-Banach Theorem 9.5 there exists a $\widehat{\varphi} \in V^*$ with $\widehat{\varphi}(x) = \varphi(x) = \|x\|$ and $\|\widehat{\varphi}\| = \|\varphi\| = 1$.

(ii) It is clear that $\sup_{\varphi \in V^*, \|\varphi\|=1} |\varphi(x)| \leq \|x\|$, and the converse inequality follows from (i).

(iii) Apply (ii) to $x - x'$.

(iv) One easily checks that $\widehat{x} : V^* \to \mathbb{F}$, $\varphi \mapsto \varphi(x)$ is a linear functional. If $x \in V$, $\varphi \in V^*$ then $|\widehat{x}(\varphi)| = |\varphi(x)| \leq \|x\|\|\varphi\|$. Thus $\|\widehat{x}\| \leq \|x\|$. By (i) there is $\varphi \in V^*$ with $\|\varphi\| = 1$ such that $\varphi(x) = \|x\|$. This gives $\|\widehat{x}\| = \|x\|$. Thus the map $\iota_V : V \to V^{**}$, $x \mapsto \widehat{x}$, which clearly is linear, is an isometric embedding.

(v) If V is complete then $\iota_V(V) \subseteq V^{**}$ is closed by Corollary 3.23 since ι_V is an isometry by (iv). Conversely, if $\iota_V(V) \subseteq V^{**}$ is closed then completeness of V^{**} (Proposition 3.25(ii))

[55] If $f : X \to Y$ is any function, from a category theory point of view one would call X the source (or domain) and Y the target (or codomain) of f and call the subset $f(X) \subseteq Y$ the image of f. I prefer to avoid the term 'range' since some authors use it for 'target' (thus Y) and others for 'image' (thus $f(X)$). The term 'image' is unambiguous since no reasonable person would use it intending Y.

implies that $\iota_V(V)$ is complete, thus also V since $\iota_V : V \to V^{**}$ is an isometric bijection. ■

It is customary to simply write ι instead of ι_V or to drop ι_V from the notation entirely, identifying V with its image $\iota_V(V)$ in V^{**}, so that $V \subseteq V^{**}$.

9.10 Corollary *Every normed space V embeds isometrically as a dense subspace into a Banach space $\widehat{V}$. The latter is unique up to isometric isomorphism and is called the completion of V.*

Proof. This can be proven by completing the metric space (V, d), where $d(x, y) = \|x - y\|$ and showing that the completion is a linear space, but this is a bit tedious. Alternatively, using the above result that $\iota_V : V \to V^{**}$ is an isometry, we can take $\widehat{V} = \overline{\iota_V(V)} \subseteq V^{**}$ as the definition of $\widehat{V}$ since this is a closed subspace of the complete space V^{**}, thus complete, and contains $\iota_V(V) \cong V$ as a dense linear subspace.

Uniqueness of the completion follows with the same proof as for metric spaces, cf. [108]. ■

The Hahn-Banach theorem allows to prove two claims made earlier:

9.11 Corollary *If V is a normed space and $\{x_n\} \subset V$ such that $\sum_{n=1}^{\infty} x_{\sigma(n)}$ converges for all permutations σ of $\mathbb{N}$ (i.e. unconditionally) then the sums do not depend on σ.*

Proof. For each $\varphi \in V^*$, continuity of φ we have

$$\varphi\Big(\sum_{n=1}^{\infty} x_{\sigma(n)}\Big) = \varphi\Big(\lim_{N\to\infty} \sum_{n=1}^{N} x_{\sigma(n)}\Big) = \lim_{N\to\infty} \sum_{n=1}^{N} \varphi(x_{\sigma(n)}) = \sum_{n=1}^{\infty} \varphi(x_{\sigma(n)}).$$

Thus the series on the r.h.s. converges for all σ. Since it takes values in $\mathbb{R}$ or $\mathbb{C}$, this implies absolute convergence and independence of $\sum_{n=1}^{\infty} \varphi(x_{\sigma(n)})$ of σ. Thus $\varphi\big(\sum_{n=1}^{\infty} x_{\sigma(n)}\big) = \varphi\big(\sum_{n=1}^{\infty} x_n\big)$ for all $\varphi \in V^*$. Now the claim follows from Proposition 9.9(iii). ■

9.12 Proposition *Let V be a normed space and $W \subseteq V$ a finite-dimensional subspace. Then*

(i) There exists an idempotent $P \in B(V)$ such that $W = PV$, thus W is complemented.

(ii) We can achieve $\|P\| \leq \dim W$.

Proof. (i) To begin with, W is closed by Exercise 3.22. If $E = \{e_1, \ldots, e_n\}$ is a basis for W, there are unique linear functionals $\varphi_i : W \to \mathbb{C}$ such that $w = \sum_{i=1}^{n} \varphi_i(w) e_i$ for each $w \in W$, equivalently, $\varphi_j(e_i) = \delta_{ij}$. Since W is finite-dimensional, the φ_i are automatically bounded by Exercise 3.7. Now by the Hahn-Banach Theorem 9.5 there are linear functionals $\widehat{\varphi}_i : V \to \mathbb{C}$ extending the φ_i with $\|\widehat{\varphi}_i\| = \|\varphi_i\|$. Then $P \in B(V)$, $v \mapsto \sum_{i=1}^{n} \widehat{\varphi}_i(v) e_i$ is bounded. If $w \in W$ then $P(w) = \sum_{i=1}^{n} \widehat{\varphi}_i(w) e_i = \sum_{i=1}^{n} \varphi_i(w) e_i = w$, thus $P \restriction W = \mathrm{id}_W$. This implies $W = PV$ and $P^2 = P$, so that W is complemented by Exercise 6.13. (A complement, obviously closed, then is $Z = (\mathbf{1} - P)V = \ker P = \bigcap_{i=1}^{n} \ker \widehat{\varphi}_i$.)

(ii) If E is an Auerbach basis, cf. Proposition 3.10, then $\|e_i\| = \|\varphi_i\| = 1\ \forall i$. Then also the $\widehat{\varphi}_i$ have norm one, so that $\|Px\| = \|\sum_i \varphi_i(x) e_i\| \leq \sum_i |\varphi_i(x)| \|e_i\| \leq \|x\| \sum_i \|\varphi_i\| \|e_i\| = n\|x\|$, proving the claim. ■

9.13 Remark With more effort one proves the Kadets-Snobar theorem (1971) giving an idempotent P with image W satisfying $\|P\| \leq \sqrt{\dim W}$ (which is almost optimal, but not quite). Cf. e.g. [103, Theorem 12.14] or [1, Theorem 13.1.7]). If V is a Hilbert space, one clearly has the orthogonal projections satisfying $\|P\| = 1$. If V is isomorphic to a Hilbert space, this implies a uniform bound $\|P\| \leq \lambda < \infty$ for all finite-dimensional subspaces. The converse is also true! Cf. e.g. [1, Theorem 13.4.3]. Combining this with the Kadets-Snobar theorem, it is not too difficult to prove the characterization of Hilbert spaces mentioned in Remark 6.10.4. □

Now we can continue the discussion of $\perp$ and $\top$ begun with Definition 6.6 and Exercise 6.7:

9.14 Exercise Let V be a normed space and $W \subseteq V$, $\Phi \subseteq V^*$ linear subspaces. Prove:

(i) $W^\perp = \{0\} \Leftrightarrow \overline{W} = V$.

(ii) $W^\perp = V^* \Leftrightarrow W = \{0\}$.

(iii) $\overline{\Phi} = V^* \Rightarrow \Phi^\top = \{0\}$.

(iv) $(W^\perp)^\top = \overline{W}$.

9.15 Exercise Let V be a normed space and $W \subseteq V$ a closed linear subspace. Construct an isometric linear bijection $\beta : V^*/W^\perp \to W^*$.

9.16 Exercise If V is Banach and $Z \subseteq V^*$ a closed subspace, construct an isometric isomorphism $V^*/Z \cong (Z^\top)^*$.

9.17 Exercise Let V be an infinite-dimensional Banach space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

(i) Use Hahn-Banach to construct sequences $\{x_n\}_{n\in\mathbb{N}} \subseteq V$ and $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq V^*$ such that $\|x_n\| = 1$ and $\varphi_n(x_n) \neq 0$ for all $n \in \mathbb{N}$ and $\varphi_n(x_m) = 0$ whenever $n \neq m$.

(ii) Prove that $\{x_n\}_{n\in\mathbb{N}}$ is linearly independent and that $x_n \notin \overline{\mathrm{span}_{\mathbb{F}}\{x_m \mid m \neq n\}}$ for all n.

9.18 Exercise Let V be a normed space and $x \in V$, $\varphi \in V^*$. Prove that $\iota_V(x) \in V^{**}$ and $\iota_{V^*}(\varphi) \in V^{***}$ satisfy $\iota_{V^*}(\varphi)(\iota_V(x)) = \varphi(x)$.

9.4 Reflexivity of Banach spaces

The following definition is immediately suggested by the fact that for every Banach space V there is an isometric embedding $\iota_V : V \hookrightarrow V^{**}$:

9.19 Definition *A Banach space V is called reflexive if the map $\iota_V : V \to V^{**}$ is surjective.*

9.20 Remark 1. Reflexivity of V means that every bounded linear functional $\psi \in V^{**}$ on V^* is of the form $\psi = \iota_V(x)$, thus $\varphi \mapsto \varphi(x)$ for some $x \in V$.

2. If V is reflexive, $\iota_V : V \to V^{**}$ is an isometric isomorphism of normed spaces. Now completeness of V^{**} implies completeness of V. For this reason there is little point in defining reflexivity for normed spaces.

3. There are Banach spaces V that are not reflexive, yet satisfy $V \cong V^{**}$ non-canonically! An example is the James[56] space, see e.g. [102, Section 4.5], which is also interesting since $V^{**}/\iota_V(V)$ is one-dimensional! For 'most' non-reflexive spaces this quotient is infinite-dimensional. (E.g. $c_0^{**}/\iota_{c_0}(c_0) \cong \ell^\infty/c_0$ is infinite-dimensional since otherwise $c_0 \subseteq \ell^\infty$ would be complemented.) □

9.21 Theorem *Let V be a Banach space. Then V is reflexive if and only if V^* is reflexive.*

[56] Robert Clarke James (1918-2004). American functional analyst.

Proof. $\Rightarrow$ Reflexivity of V means surjectivity of $\iota_V : V \to V^{**}$. Let $\varphi \in V^{***} = (V^{**})^*$. Then $\varphi' = \varphi \circ \iota_V \in V^*$, and we claim that $\varphi = \iota_{V^*}(\varphi')$. This would clearly imply surjectivity of $\iota_{V^*} : V^* \to V^{***}$, thus reflexivity of V^*. The claim means $\varphi(x^{**}) = \iota_{V^*}(\varphi')(x^{**})$ for all $x^{**} \in V^{**}$. By surjectivity of $\iota_V : V \to V^{**}$, this is equivalent to $\varphi(\iota_V(x)) = \iota_{V^*}(\varphi')(\iota_V(x))$ for all $x \in V$. The latter identity indeed is true since both sides equal $\varphi'(x)$, the l.h.s. by the definition of φ' and the r.h.s. by Exercise 9.18.

$\Leftarrow$ Assume that V is not reflexive. Then $\iota_V(V) \subseteq V^{**}$ is a proper closed subspace, so that $\iota_V(V)^\perp \neq \{0\}$ by Exercise 9.14. Let thus $0 \neq \varphi \in \iota_V(V)^\perp \subseteq V^{***}$. Since V^* is reflexive, we have $\varphi = \iota_{V^*}(\varphi')$ for some $\varphi' \in V^*$. Using Exercise 9.18 again, for each $x \in V$ we have $\varphi'(x) = \iota_{V^*}(\varphi')(\iota_V(x)) = \varphi(\iota_V(x)) = 0$ by $\varphi \in \iota_V(V)^\perp$. Thus $\varphi' = 0$, implying $\varphi = 0$, but this is a contradiction. ■

9.22 Remark For non-reflexive V none of the spaces $V^*, V^{**}, V^{***}, \ldots$ is reflexive, so that $V \subsetneqq V^{**} \subsetneqq V^{****} \subsetneqq \cdots$ and $V^* \subsetneqq V^{***} \subsetneqq V^{*****} \subsetneqq \cdots$, and we have two somewhat mysterious successions of ever larger spaces! There do not seem to be many general results about this, but see Lemma B.25(iv). Even understanding $C(X, \mathbb{R})^{**}$ for compact X is complicated, cf. [81]. □

9.23 Exercise Prove:

(i) Every finite-dimensional Banach space is reflexive.

(ii) Every Hilbert space is reflexive.

(iii) If $1 < p < \infty$ then $\ell^p(S, \mathbb{F})$ is reflexive.

(iv) If S is infinite then $c_0(S, \mathbb{F}), \ell^1(S, \mathbb{F}), \ell^\infty(S, \mathbb{F})$ are not reflexive.

9.24 Exercise (i) Prove that if V is reflexive then for each $\varphi \in V^*$ there exists an $x \in V$ such that $\|x\| = 1$ and $|\varphi(x)| = \|\varphi\|$. (We say '$\varphi$ attains its norm'.)

(ii) Identify the $\varphi \in c_0(\mathbb{N}, \mathbb{F})^*$ for which there exists $x \in c_0(\mathbb{N}, \mathbb{F})$ with $\|x\| = 1$ such that $\varphi(x) = \|\varphi\|$. Conclude that such φ are dense in $c_0(\mathbb{N}, \mathbb{F})^*$.

(ii) Prove (again) that $c_0(\mathbb{N}, \mathbb{C})$ is not reflexive.

9.25 Remark 1. The converse of the statement in Exercise 9.24(i) is also true: If every $\varphi \in V^*$ attains its norm, V is reflexive. But the proof, also due to R. C. James, is much harder and more than 10 pages long! See [102, Section 1.13].

2. On the other hand, Bishop and Phelps[57] proved that the result of Exercise 9.24(ii) holds for every Banach space V, i.e. the set of $\varphi \in V^*$ that attain their norm is dense in V^*. Cf. [16] or [102, Section 2.11].

3. See Appendix B.6.8 for the notion of uniform convexity, which is stronger than the strict convexity encountered earlier, and a proof of the fact that uniformly convex spaces are reflexive.

We will also prove prove that $L^p(X, \mathcal{A}, \mu)$ is uniformly convex for each measure space $(X, \mathcal{A}, \mu)$ and $1 < p < \infty$. This provides a proof of reflexivity of these spaces that does not use the relation between L^p and L^q. This in turn leads to a simple proof of surjectivity of the isometric map $L^q \to (L^p)^*$ known from Section 4.7 (reversing the logic of Exercise 9.23(iii)). □

9.26 Exercise Let V be a Banach space. Prove:

[57] Errett Albert Bishop (1928-1983), Robert Ralph Phelps (1926-2013), Amerian functional analysts. Around 1965 Bishop became a strong advocate of and contributor to constructive mathematics.

(i) If V^* is separable then V is separable.

(ii) For V infinite-dimensional separable, V^* can be separable or non-separable. (Examples!)

(iii) If V is separable and reflexive then V^* is separable.

9.27 Theorem (Pettis) [58] *Let V be a Banach space and $W \subseteq V$ a closed subspace. Then the following are equivalent:*

(i) V is reflexive.

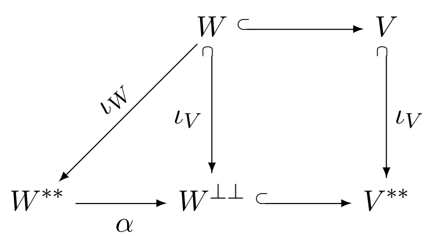
(ii) W and V/W are reflexive.

Proof. We begin with some preparations. Let $W \subseteq V$ be a closed subspace of W. $W^\perp \subseteq V^*$ is a closed subspace, thus $W^{\perp\perp}$ is a closed subspace of V^{**}. Explicitly,

$$W^{\perp\perp} = \{\psi \in V^{**} \mid \varphi \in V^*,\, \varphi\restriction W = 0 \Rightarrow \psi(\varphi) = 0\}. \tag{9.3}$$

If $w \in W$ and $\psi = \iota_V(w)$ then for each $\varphi \in V^*$ we have $\psi(\varphi) = \iota_V(w)(\varphi) = \varphi(w)$. Thus if $\varphi\restriction W = 0$ then $\psi(\varphi) = 0$. This proves $\iota_V(W) \subseteq W^{\perp\perp}$.

By Exercise 6.7 (dual space of quotient space) we have an isometric isomorphism $W^* \cong V^*/W^\perp$. Now Exercise 9.15 (dual space of subspace) gives an isometric isomorphism $\alpha : W^{**} \to W^{\perp\perp}$, where $W^{\perp\perp} \subseteq V^{**}$. We thus have the situation in this diagram:



(9.4)

Let $w \in W$. Now $\iota_W(w) \in W^{**}$ and $\iota_V(w) \in V^{**}$ are the linear functionals on W^* and V^*, respectively, given by evaluation at w. Thus for $\varphi \in V^*$ we have $\iota_V(w)(\varphi) = \varphi(w)$. On the other hand, $\varphi\restriction W \in W^*$, and $\iota_W(w)(\varphi\restriction W) = (\varphi\restriction W)(w) = \varphi(w)$, proving that the left triangle of the diagram commutes.

(i)⇒(ii) Now assume that V is reflexive, so that $\iota_V : V \to V^{**}$ is a bijection. Thus every $\psi \in V^{**}$ is of the form $\iota_V(v)$ for a unique $v \in V$. With this, (9.3) becomes

$$W^{\perp\perp} = \{\iota_V(v) \mid \varphi \in V^*,\, \varphi\restriction W = 0 \Rightarrow \varphi(v) = 0\} = \iota_V(W),$$

where we used that for every $v \in V\backslash W$ there exists a $\varphi \in V^*$ with $\varphi\restriction W = 0$, $\varphi(v) \neq 0$. Thus $\iota_V : W \to W^{\perp\perp}$ is a bijection. Since α is a bijection, also $\iota_W : W \to W^{**}$ is a bijection, thus W is reflexive.

By Theorem 9.21, V^* is reflexive. Thus the closed subspace $W^\perp \subseteq V^*$ is reflexive by what was just proven. Since $W^\perp \cong (V/W)^*$ by Exercise 6.7, $(V/W)^*$ is reflexive, thus V/W is reflexive using Theorem 9.21 again.

(ii)⇒(i) Let $\psi \in V^{**}$. Our aim is to find a $v \in V$ such that $\psi = \iota_V(v)$. We have a canonical isomorphism $\beta : (V/W)^* \to W^\perp \subseteq V^*$. Thus $\psi \circ \beta \in (V/W)^{**}$. Since V/W is reflexive, there exists $v + W \in V/W$ such that $\iota_{V/W}(v + W) = \psi \circ \beta$. Now for all $\varphi \in W^\perp \subseteq V^*$ we have

$$\psi(\varphi) = \psi \circ \beta \circ \beta^{-1}(\varphi) = \iota_{V/W}(v + W)(\beta^{-1}(\varphi)) = (\beta^{-1}(\varphi))(v + W) = \varphi(v) = \iota_V(v)(\varphi).$$

[58] Billy James Pettis (1913-1979), American mathematician who mostly worked in functional analysis.

Thus $\psi - \iota_V(v) \in V^{**}$ vanishes on $W^\perp$, so that $\psi - \iota_V(v) \in W^{\perp\perp}$. Since W is reflexive, ι_W is a bijection. Together with the fact that α is a bijection, this implies that $\iota_V : W \to W^{\perp\perp}$ is a bijection, thus $W^{\perp\perp} = \iota_V(W)$. Thus there exists $w \in W$ such that $\iota_V(w) = \psi - \iota_V(v)$, thus $\psi = \iota_V(v + w)$, proving surjectivity of ι_V. Thus V is reflexive. ■

In these notes, the Banach spaces $C_{(b)}(X, \mathbb{F}), C_0(X, \mathbb{F})$ of continuous functions do not play a very prominent role, since their study requires more general topology than the rest or our subjects, and also measure theory for the dual spaces. But the following is not too difficult:

9.28 Exercise Let X be a normal (T_4) topological space. Prove that the Banach space $C_b(X, \mathbb{F})$ is reflexive if and only if $\#X < \infty$. Hint: If $\#X = \infty$, pick distinct points $\{x_n\}_{n\in\mathbb{N}}$, use Urysohn's lemma to produce functions $f_n \in C(X, [0, 1])$ with disjoint supports and $f_n(x_m) = \delta_{n,m}$. Use these to produce an embedding $c_0 \hookrightarrow C_b(X, \mathbb{F})$.

9.5 The transpose of a bounded Banach space operator

9.29 Definition *If V, W are normed spaces over $\mathbb{F}$, $A : V \to W$ is linear and $\varphi \in W^*$ then $A^t\varphi := \varphi \circ A : V \to \mathbb{F}$ is linear (and bounded if A is bounded). This defines a linear map $A^t : W^* \to \mathrm{Lin}(V, \mathbb{F})$, called the transpose of A. If A is bounded then A^t maps $W^* \to V^*$.*

Some authors call A^t the 'adjoint' of A, but we stick to transpose in order to avoid confusion with the Hilbert space adjoint discussed below. Note that the transpose goes in the 'opposite direction'!

9.30 Exercise If V, W, Z are normed spaces and $A \in B(V, W), B \in B(W, Z)$, prove $(BA)^t = A^tB^t$ in $B(X^*, Z^*)$.

9.31 Lemma *If V, W are normed spaces over $\mathbb{F}$ and $A : V \to W$ is linear then $\|A^t\| = \|A\|$. Thus A^t is bounded if and only if A is bounded. The map $B(V, W) \to B(W^*, V^*)$, $A \mapsto A^t$ is isometric.*

Proof. The identity follows from the computation

$$\|A\| = \sup_{\substack{v\in V\\ \|v\|=1}} \|Av\| = \sup_{\substack{v\in V\\ \|v\|=1}} \sup_{\substack{\varphi\in W^*\\ \|\varphi\|=1}} |\varphi(Av)| = \sup_{\substack{\varphi\in W^*\\ \|\varphi\|=1}} \sup_{\substack{v\in V\\ \|x\|=1}} |\varphi(Av)| = \sup_{\substack{\varphi\in W^*\\ \|\varphi\|=1}} \|A^t\varphi\| = \|A^t\|,$$

where the the first and last identities are the definition of the norm, the second and fourth follow from Proposition 9.9(ii), and the third is the exchangeability of two suprema. (Note that we did not assume boundedness of A or A^t.) The rest is clear. ■

Now we can use W^* to test a linear map $A : V \to W$ for boundedness:

9.32 Exercise Let V, W be normed spaces and $A : V \to W$ a linear map. Prove: A is bounded if and only if $A^t\varphi = \varphi \circ A$ is bounded for all $\varphi \in W^*$. Hint: Use Lemma 9.31 and UBT or CGT.

The transposition operation can be iterated, giving $A^{tt} \in B(V^{**}, W^{**})$, etc.

9.33 Lemma *If V, W are normed spaces and $A \in B(V,W)$ then the diagram*

$$\begin{array}{ccc} V & \xrightarrow{\quad A \quad} & W \\ \Big\downarrow{\scriptstyle \iota_V} & & \Big\downarrow{\scriptstyle \iota_W} \\ V^{**} & \xrightarrow[\quad A^{tt} \quad]{} & W^{**} \end{array}$$

commutes, thus considering V and W as subspaces of V^{**}, W^{**}, we have $A^{tt} \upharpoonright V = A$.

Proof. Let $v \in V, \varphi \in W^*$. Then using the definition of ι_V, ι_W and of the transpose, we have

$$\iota_W(Av)(\varphi) = \varphi(Av) = (A^t\varphi)(v) = \iota_V(v)(A^t\varphi) = (A^{tt}\iota_V(v))(\varphi).$$

Now, $\iota_W(Av)$ and $(A^{tt}\iota_V(v))$ are in W^{**}, and the fact that they coincide on all $\varphi \in W^*$ means $\iota_W(Av) = A^{tt}\iota_V(v)$. And since this holds for all $v \in V$, we have $\iota_W A = A^{tt}\iota_V$, as claimed. ■

9.34 Remark 1. If V and W are reflexive, we can identify $V^{**} = V$ and $W^{**} = W$, obtaining $A^{tt} = A$, so that $B(V,W) \to B(W^*, V^*)$, $A \mapsto A^t$ is a bijection. But in general, the transposition map $B(V,W) \to B(W^*, V^*)$, $A \mapsto A^t$ is not surjective. For a characterization of the $A \in B(W^*, V^*)$ that are of the form $A = B^t$ with $B \in B(V,W)$ see Theorem 10.32(i).

2. If V, W are finite-dimensional, we know from linear algebra that $A \in B(V,W)$ is injective (surjective) if and only if $A^t \in B(W^*, V^*)$ is surjective (injective). In infinite dimensions this becomes more complicated since, as seen in Section 7.4, we must distinguish between injectivity and boundedness below and between dense image and surjectivity. □

9.35 Exercise Let V, W be Banach spaces and $A \in B(V,W)$. Prove:

(i) $\ker A^t = (AV)^\perp \subseteq W^*$, thus A^t is injective if and only if A has dense image $AV \subseteq W$.

(ii) If $A \in B(V,W)$ is invertible then $A^t \in B(W^*, V^*)$ is invertible.

(iii) If $A^t \in B(W^*, V^*)$ is invertible then $A \in B(V,W)$ is invertible. (Warning: We don't assume reflexivity of the spaces involved!)

Hint: (i) and (ii) are very easy. The proof of (iii) uses (i) and (ii).

9.36 Exercise Let V, W be Banach spaces and $A \in B(V,W)$.

(i) Prove $\ker A = (A^t W^*)^\top$.

(ii) Show that both (i) and Exercise 9.35(i) can be used to prove: If A^t has dense image then A is injective.

9.37 Exercise Let V, W be Banach spaces and $A \in B(V,W)$. Prove:

(i) If A is surjective then $A^t : W^* \to V^*$ is bounded below (thus injective).

(ii) If A is bounded below then $A^t : W^* \to V^*$ is surjective.

(iii) Deduce (not assuming reflexivity!) that A^t is surjective if and only if A is bounded below.

9.38 Proposition *Let V, W be Banach spaces and $A \in B(V,W)$. If A^t is bounded below then A is surjective.*

Proof. Let $C > 0$ be such that $\|A^t\varphi\| \geq C\|\varphi\|$ for all $\varphi \in W^*$. The set $Y = \overline{AB^V(0,1)} \subset W$ is closed, convex and balanced. Thus for each $z \in W \backslash Y$ by Proposition B.64 there exists $\varphi \in W^*$ such that $|\varphi(y)| \leq 1$ for all $y \in Y$ and $|\varphi(z)| > 1$. By the first of these properties, for all $x \in B^V(0,1)$ we have $|(A^t\varphi)(x)| = |\varphi(Ax)| \leq 1$, implying $\|A^t\varphi\| \leq 1$. Thus with the hypothesis,

$$C < C|\varphi(z)| \leq C\|z\|\|\varphi\| \leq \|z\|\|A^t\varphi\| \leq \|z\|.$$

By contraposition, if $w \in W$ satisfies $\|w\| \leq C$ then $w \in Y$. In particular, $B^W(0,C) \subseteq \overline{AB^V(0,1)}$. Now Proposition 7.4 gives the surjectivity of A. ■

9.39 Remark Summarizing our findings, for $A \in B(V,W)$ we have

A has dense image	$\Leftrightarrow$	A^t is injective
A is bounded below	$\Leftrightarrow$	A^t is surjective
A is surjective	$\Leftrightarrow$	A^t is bounded below
A is injective	$\Leftarrow$	A^t has dense image

From this, we easily deduce that A^{tt} is surjective (resp. bounded below) if and only if A is surjective (resp. bounded below). And if A^{tt} is injective then so is A, but this follows more directly from Lemma 9.33.

As to the last row, we will see that the condition on A^t equivalent to injectivity of A is somewhat weaker than (norm-)density of A^tW^*, cf. Theorem 10.32(i). □

9.40 Proposition *Let V, W be Banach spaces and $A \in B(V,W)$.*

(i) If A has closed image AV then $A^tW^* = (\ker A)^\perp \subseteq V^*$, thus A^t has closed image.

(ii) If A^t has closed image, so has A.

Proof. (i) is Exercise 9.41 below. (ii) Assume A^t has closed image. Following [141], define $Z = \overline{AV} \subseteq W$ and define $A' : V \to Z$ in the obvious way. Then A' has dense image, so that $(A')^t : Z^* \to V^*$ is injective by Exercise 9.35. By Hahn-Banach, every $\varphi \in Z^*$ has an extension $\widehat{\varphi} \in W^*$. Now for every $v \in V$ we have $(A^t\widehat{\varphi})(v) = \widehat{\varphi}(Av) = \varphi(Av) = ((A')^t\varphi)(v)$, implying $A^t\widehat{\varphi} = (A')^t\varphi$. Thus A^t and $(A')^t$ have the same images in V^*. Since $A^tW^* \subseteq V^*$ is closed by assumption, also $(A')^tZ^* \subseteq V^*$ is closed, thus complete. As an injective map with closed image, $(A')^t : Z^* \to (A')^tZ^*$ is invertible by the BIT, thus bounded below. Now Proposition 9.38 (applied to A') gives that $A' : V \to Z$ is surjective. Since Z is closed by definition, $AV = A'V = Z$ is closed. ■

9.41 Exercise Prove (i) of Proposition 9.40. Hint: Factorize A as $V \xrightarrow{A_1} V/\ker A \xrightarrow{A_2} AV \xrightarrow{A_3} W$ and use the BIT and HB theorems.

9.42 Exercise If V is a reflexive Banach space and $W \simeq V$ (not necessarily isometric isomorphism), prove that W is reflexive.

10 ⋆ Duality: Weak and weak ∗-topologies

Every Banach space has a canonical metric topology defined by the norm. But it also admits an equally canonical weaker topology, the weak topology. And on the dual space V^* of a Banach space V there is yet another topology, the weak-∗ topology. These topologies have many applications in functional analysis, cf. e.g. [30], but we will discuss only a few, most importantly:

- characterization of reflexive Banach spaces: Theorem 10.15.
- applications to ⊤ and transposes A^t of operators: Proposition 10.31, Theorem 10.32.
- characterizations of compact operators: Proposition 12.25 and Theorem 12.28.
- a (locally) compact topology relevant for the theory of commutative Banach algebras: Section 19.

10.1 The weak topology of a Banach space

10.1 Definition *If V is a normed space, the weak topology $\tau_{\rm w}$ is the topology on V induced by the family of seminorms $\mathcal{F} = \{\|\cdot\|_\varphi = |\varphi(\cdot)| \mid \varphi \in V^*\}$. Thus a net $\{x_\iota\} \subseteq V$ converges weakly to $x \in V$ if and only if $\varphi(x_\iota) \to \varphi(x)$ for all $\varphi \in V^*$.*

The weak topology is also called the $\sigma(V, V^*)$-topology (the topology on V induced by the linear functionals in V^*). The Hahn-Banach theorem immediately gives that $\mathcal{F}$ is separating, so that this topology is locally convex. It is clear that a norm-convergent net is weakly convergent since $|\varphi(x_\iota) - \varphi(x)| \le \|\varphi\|\|x_\iota - x\|$. This implies $\overline{S}^{\|\cdot\|} \subseteq \overline{S}^{\rm w}$ for every $S \subseteq V$ and $\tau_{\rm w} \subseteq \tau_{\|\cdot\|}$.

If $(H, \langle\cdot,\cdot\rangle)$ is a Hilbert space, Theorem 5.34 implies $\mathcal{F} = \{|\langle\cdot, y\rangle| \mid y \in H\}$, so that weak convergence of a net $\{x_\iota\}$ in a Hilbert space H is equivalent to convergence of the nets $\{\langle x_\iota, y\rangle\}$ for each $y \in H$.

Since the norm and weak topologies on a normed space are Hausdorff, Proposition 2.29 implies that they coincide if the space is finite-dimensional. On the other hand:

10.2 Proposition *Let V be an infinite-dimensional normed space. Then the weak topology $\tau_{\rm w}$ on V is strictly weaker than the norm-topology and not first countable. In particular $(V, \tau_{\rm w})$ is neither normable nor Fréchet nor an F-space.*

Proof. By the definition of $\tau_{\rm w}$, for every weakly open neighborhood U of 0 there are $\varphi_1, \ldots, \varphi_n \in V^*$ such that $\{x \in V \mid |\varphi_i(x)| < 1 \ \forall i = 1, \ldots, n\} \subseteq U$. Thus U contains the linear subspace $W = \bigcap_{i=1}^n \varphi_i^{-1}(0) \subseteq V$, whose codimension is $\le n$. Thus if V is infinite-dimensional then $\dim W$ is infinite, thus non-zero. On the other hand, it is clear that the (norm-)open ball $B(0,1)$ contains no linear subspace of dimension > 0. Thus $B(0,1) \notin \tau_{\rm w}$. Since $\tau_{\rm w} \subseteq \tau_{\|\cdot\|}$ was clear, we have $\tau_{\rm w} \subsetneqq \tau_{\|\cdot\|}$.

If we assume that $\tau_{\rm w}$ is first countable, $0 \in V$ has a countable open neighborhood base $\{U_n\}_{n\in\mathbb{N}}$. Replacing U_n by $\bigcap_{k=1}^n U_k$, we may assume $U_1 \supseteq U_2 \supseteq \cdots$. As seen above, being weakly open, each U_n contains a non-zero linear subspace V_n. For each n we can pick a $x_n \in V_n$. If now $\varphi \in V^*$ and $\varepsilon > 0$ are arbitrary, $U = \{x \in V \mid |\varphi(x)| < \varepsilon\}$ is a weakly open neighborhood of 0. Since $\{U_n\}$ is a shrinking weak neighborhood base of 0, there exists n_0 such that for all $n \ge n_0$ we have $x_n \in V_n \subseteq U_n \subseteq U$, thus $|\varphi(x_n)| < \varepsilon$, implying $\varphi(x_n) \to 0$. Since $\varphi \in V^*$ was arbitrary, we have proven $x_n \stackrel{\rm w}{\to} 0$. Since this holds for every choice of $\{x_n \in V_n\}$ and the V_n are vector spaces, we can choose $\{x_n\}$ such that $\|x_n\| \to \infty$. But this contradicts the fact that every weakly convergent sequence is norm-bounded, cf. Exercise 10.6 below. This contradiction shows that $\tau_{\rm w}$ is not first countable.

Now the last statement is trivial since for a TVS we have the implications normable ⇒ Fréchet ⇒ F-space ⇒ metrizable ⇒ first countable. ■

10.3 Exercise (i) Prove that the sequence $\{\delta_n\}_{n\in\mathbb{N}}$ has no weak limit in $\ell^1(\mathbb{N}, \mathbb{F})$.

(ii) Let $1 < p < \infty$. Prove that the sequence $\{\delta_n\}_{n\in\mathbb{N}} \subseteq \ell^p(\mathbb{N},\mathbb{F})$ converges to zero weakly, but not in norm.

(iii) If $f_n \stackrel{w}{\to} g$ in a Hilbert space and $\|f_n\| \to \|g\|$, prove that $\|f_n - g\| \to 0$.

(iv) (Bonus) Prove the result of (iii) for sequences in $(\ell^p(\mathbb{N},\mathbb{F}), \|\cdot\|_p)$, where $1 < p < \infty$.[59]

The deviant behavior of ℓ^1 in the preceding exercise can be understood as a consequence of the following surprising result:

10.4 Theorem (I. Schur 1920) [60] *If $g, \{f_n\}_{n\in\mathbb{N}} \subseteq \ell^1(\mathbb{N},\mathbb{F})$ and $f_n \stackrel{w}{\to} g$ then $\|f_n - g\|_1 \to 0$.*

10.5 Remark 1. Like the uniform boundedness theorem, this result can be proven using a beautiful gliding hump argument, cf. Section B.3.5, or using Baire's theorem.

2. Theorem 10.4 does not generalize to nets since the weak and norm topologies on $\ell^1(\mathbb{N},\mathbb{F})$ differ by Proposition 10.2 and nets can distinguish topologies, cf. e.g. [108, Section 5.1].

3. Banach spaces in which weak and norm convergence of sequences are equivalent are said to have the Schur property. All finite-dimensional spaces have it. See also Remark 12.31.1. □

10.6 Exercise Prove that every weakly convergent sequence in a normed space is norm-bounded. Hint: Uniform boundedness theorem. (This does not generalize to nets!)

10.7 Exercise Prove that every weakly compact subset of a Banach space is norm-bounded.

10.8 Exercise Let V be a Banach space. Prove that the (norm) closed unit ball $V_{\leq 1}$ is also weakly closed. Hint: Hahn-Banach.[61]

Given a linear map $A : E \to F$ between Banach spaces, one can consider its continuity w.r.t. different pairs of topologies on E and F: Norm-norm continuity (w.r.t. the norm topologies on both spaces), weak-norm continuity (the weak topology on E, the norm topology on F) and, similarly, norm-weak and weak-weak continuity. These notions are not all distinct:

10.9 Exercise Let V, W be Banach spaces and $A : V \to W$ a linear map. Prove that the following are equivalent:

(i) A is norm-norm continuous (equivalently, bounded).

(ii) A is norm-weak continuous.

(iii) A is weak-weak continuous.

10.10 Exercise With the same assumptions as above, prove that the following are equivalent:

(i) A is weak-norm continuous.

(ii) There are $\varphi_1, \ldots, \varphi_K \in V^*$ and $y_1, \ldots, y_K \in W$ such that $Ax = \sum_{k=1}^K \varphi_k(x) y_k \ \forall x \in V$.

(iii) A is bounded and $AV \subseteq W$ is finite-dimensional (i.e. A has 'finite rank').

[59] More generally, this implication holds for all uniformly convex Banach spaces, cf. e.g. [86, Proposition 9.11], [23, Sect. 3.7]. The L^p-spaces with $1 < p < \infty$ are uniformly convex, cf. Section B.6.8.

[60] Issai Schur (1874-1941). Russian mathematician. Studied and worked in Germany up to his emigration to Israel in 1939. Mostly known for his work in group and representation theory.

[61] More generally, every norm-closed convex set is weakly closed. But this is a bit harder.

10.11 Exercise If V is a normed space and $A \in B(V)$ has finite rank, prove that the trace (as known from linear algebra) of $A \upharpoonright AV \in \mathrm{End}(AV)$ coincides with $\sum_{i=1}^n \varphi_i(y_i)$ for any representation $A = \sum_{k=1}^K \varphi_k(\cdot) y_k$.

10.12 Definition *If V is a Banach space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, a sequence $\{x_n\} \subset V$ is called weakly Cauchy if the sequence $\{\varphi(x_n)\} \subset \mathbb{F}$ converges for every $\varphi \in V^*$. (Equivalently, $\{x_n\}$ is Cauchy in the sense of Remark 2.12.2 in the locally convex space (V, τ_{w}).)*

10.13 Exercise Prove that in a Banach space one has:

(i) Every weakly convergent sequence is weakly Cauchy.

(ii) Every weakly Cauchy sequence is norm-bounded.

(iii) Every weakly Cauchy sequence in a reflexive Banach space is weakly convergent.

(iv) Give an example of a sequence in a Banach space that is weakly Cauchy, but not weakly convergent. Hint: Try c_0.

10.14 Exercise Let V be a Banach space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Prove:

(i) If V^* is separable (thus also V by Exercise 9.26) then every norm-bounded sequence in V has a weakly Cauchy subsequence.

(ii) If V is reflexive and separable then every norm-bounded sequence in V has a weakly convergent subsequence.

(iii) Statement (ii) remains true without the separability hypothesis.

(iv) (Bonus) Show that the sequence $\{\delta_n\}_{n \in \mathbb{N}}$ in $\ell^1(\mathbb{N}, \mathbb{F})$ does not have a weakly Cauchy subsequence.[62]

The above (iii) means that every norm-bounded weakly closed subset of a reflexive Banach space is sequentially compact in the weak topology. (The converse also holds, see below.) Replacing sequential compactness by compactness we have the following, proven later:

10.15 Theorem *Let V be a Banach space. Then the following are equivalent:*

(i) $V_{\leq 1}$ is compact w.r.t. the weak topology.

(ii) V is reflexive. ($\Leftrightarrow$ V^* is reflexive by Theorem 9.21.)

10.16 Remark 1. If V is a finite-dimensional normed space then $V_{\leq 1}$ is compact w.r.t. the norm topology. For infinite-dimensional V this is false, as we will see in Section 12.1. But since the weak topology is weaker than the norm topology, a set can be weakly compact even though it is not norm compact.

2. For metric spaces, compactness and sequential compactness are equivalent, but the two properties are independent for general topological spaces. Despite the fact that weak topologies on infinite-dimensional Banach spaces are metrizable at best on bounded subsets (under separability assumptions), there is the Eberlein-Šmulyan[63] Theorem: For subsets of a Banach space

[62] In fact, given a Banach space V, every bounded sequence in V has a weakly Cauchy subsequence if and only if V has no subspace isomorphic to ℓ^1. This deep result is due to H. Rosenthal (1974). See e.g. [1, Chapter 11], [97, Vol. 1, Sect. 8.II].

[63] William Frederick Eberlein (1917-1986), American mathematician who worked in functional analysis, topology and related areas. Vitold Lvovich Šmulyan (1914-1944), Soviet mathematician, also known for the Krein-S. theorem. Killed while fighting in WW2.

weak compactness and weak sequential compactness are equivalent. (⇒ is not very difficult.) For proofs see more advanced texts or [173]. □

10.17 Exercise If H is a Hilbert space, prove that every orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$ in H converges weakly to zero.

10.2 The weak operator topology on $B(V)$

In Remark 8.10 we have encountered the strong (operator) topology on $B(V)$: A net $\{A_\iota\} \subseteq B(V)$ converges strongly to $A \in B(V)$ if $\|(A_\iota - A)x\| \to 0$ for all $x \in V$. Now we can have a brief look at the weak operator topology:

10.18 Definition *Let V be a Banach space. The weak operator topology τ_{wot} on $B(V)$ is generated by the family $\mathcal{F} = \{\|\cdot\|_{x,\varphi} : A \mapsto |\varphi(Ax)| \mid x \in V, \varphi \in V^*\}$ of seminorms. Thus $\{A_\iota\} \subseteq B(V)$ converges to $A \in B(V)$ w.r.t. τ_{wot} if and only if $\varphi((A_\iota - A)x) \to 0$ for all $x \in V, \varphi \in V^*$, i.e. $\{A_\iota x\} \subseteq V$ converges weakly to Ax for all $x \in V$. The family $\mathcal{F}$ is separating, so that τ_{wot} is Hausdorff. We write $A_\iota \stackrel{\text{wot}}{\longrightarrow} A$ or $A = \text{wot-lim}\, A_\iota$.*

There is little risk of confusing the weak topology on V with the weak operator topology on $B(V)$. But one might confuse the latter with the weak topology that $B(V)$ has as a Banach space, in particular since the above $\|\cdot\|_{x,\varphi}$ are in $B(V)^*$! However, when V is infinite-dimensional these seminorms do not exhaust (or span) the bounded linear functionals on $B(V)$, so that the weak operator topology on $B(V)$ is strictly weaker than the weak topology!

10.19 Exercise Let H be a Hilbert space.

(i) Given $A, \{A_\iota\} \subseteq B(H)$, prove that $A_\iota \stackrel{\text{wot}}{\longrightarrow} A$ if and only if $\langle A_\iota x, y\rangle \to \langle Ax, y\rangle$ for all $x, y \in H$.

(ii) Prove that the map $(B(H), \tau_{\text{wot}}) \to (B(H), \tau_{\text{wot}}),\ A \mapsto A^*$ is continuous.

(iii) Prove that the map $(B(H), \tau_{\text{sot}}) \to (B(H), \tau_{\text{sot}}),\ A \mapsto A^*$ is not continuous if $\dim H = \infty$.

10.3 The weak-$*$ topology on a dual space. Alaoglu's theorem

10.20 Definition *If V is a Banach space, the weak-$*$ topology τ_{w^*} (or $\sigma(V^*, V)$-topology) is the topology on the dual space V^* defined by the family $\mathcal{F} = \{\|\cdot\|_x \mid x \in V\}$ of seminorms, where $\|\varphi\|_x = |\widehat{x}(\varphi)| = |\varphi(x)|$. Thus a net $\{\varphi_\iota\}$ in V^* converges to $\varphi \in V^*$ if and only if $\varphi_\iota(x) \to \varphi(x)$ for every $x \in V$.*

10.21 Remark 1. Since $\varphi(x) = 0$ for all $x \in V$ means $\varphi = 0$, $\mathcal{F}$ is separating, thus the $\sigma(V^*, V)$-topology is Hausdorff and therefore locally convex.

2. If V is infinite-dimensional, the weak-$*$ topology τ_{w^*} is neither normable nor metrizable.

3. Since the weak-$*$ topology is induced by the linear functionals on V^* of the form $\widehat{x}$, which constitute a subset of V^{**}, it is weaker than the weak topology, thus also weaker than the norm topology: $\tau_{\text{w}^*} \subseteq \tau_{\text{w}} \subseteq \tau_{\|\cdot\|}$. As we know, the second inclusion is proper whenever V is infinite-dimensional. For the first, we have: □

10.22 Proposition *If V is a Banach space, the weak-$*$ topology $\sigma(V^*, V)$ on V^* coincides with the weak topology $\sigma(V^*, V^{**})$ if and only if V is reflexive.*

Proof. If V is reflexive then $V^{**} = \iota_V(V)$, so that the weak-$*$ topology $\sigma(V^*, V)$ on V^* coincides with the weak topology $\sigma(V^*, V^{**})$. If V is not reflexive, we have $V \subsetneqq V^{**}$. Now for $\psi \in V^{**} \backslash V$ it is clear that the linear functional ψ on V^* is $\sigma(V^*, V^{**})$-continuous, whereas Exercise 10.23 gives that it is not $\sigma(V^*, V)$-continuous. This proves $\sigma(V^*, V) \neq \sigma(V^*, V^{**})$. ■

10.23 Exercise Let V be an $\mathbb{F}$-vector space with algebraic dual space $V^\star$.

(i) For $\varphi, \psi_1, \ldots, \psi_n \in V^\star$ prove that $\varphi \in \text{span}_{\mathbb{F}}\{\psi_1, \ldots, \psi_n\} \Leftrightarrow \bigcap_{i=1}^n \ker \psi_i \subseteq \ker \varphi$. Hint: Use the map $V \to \mathbb{F}^n$, $x \mapsto (\psi_1(x), \ldots, \psi_n(x))$.

(ii) Let $W \subseteq V^\star$ be a linear subspace. Prove that a linear functional $\varphi : V \to \mathbb{F}$ is $\sigma(V, W)$-continuous if and only if $\varphi \in W$. Hint: Use (i).

10.24 Lemma *Let V be a Banach space over $\mathbb{F}$. Then*

(i) $(V, \tau_{\text{w}})^* = V^*$. (I.e., if $\varphi : V \to \mathbb{F}$ is linear, norm and weak continuity are equivalent.)

(ii) The continuous linear functionals on the locally convex space (V^*, τ_{w^*}) are precisely the functionals $\widehat{x} : \varphi \mapsto \varphi(x)$ for some $x \in V$.

Proof. Since the weak topology on V and the weak-$*$ topology on V^* are the $\sigma(V, V^*)$ and $\sigma(V^*, V)$ topologies, respectively, both claims are immediate consequences of Exercise 10.23(ii). ■

10.25 Remark Before we proceed, some comments are in order: While the norm and weak topologies are defined for each Banach space, the weak-$*$ topology is defined only on spaces that are the dual space V^* of a given space V. There are Banach spaces, like $c_0(\mathbb{N}, \mathbb{F})$, that are not isomorphic (isometrically or not) to the dual space of any Banach space, cf. Corollary B.26. And there are non-isomorphic Banach spaces with isomorphic dual spaces, cf. Corollary B.27. Thus to define the weak-$*$ topology on a Banach space V, it is *not* enough just to know that the latter is a dual space. We must choose a 'pre-dual' space W such that $W^* \cong V$. □

Recall that $V_{\leq 1}$ is norm compact if and only if V is finite-dimensional and weakly compact if and only if V is reflexive. The weak-$*$ topology on the dual of a non-reflexive Banach space V is strictly weaker than the weak topology, so that the closed unit ball of V^* has a chance of being weak-$*$ compact, and in fact this is the case unconditionally:

10.26 Theorem (Alaoglu's theorem (1940)) [64] *If V is a Banach space then the (norm)closed unit ball $(V^*)_{\leq 1} = \{\varphi \in V^* \mid \|\varphi\| \leq 1\}$ is compact in the $\sigma(V^*, V)$-topology.*

Proof. Define

$$Z = \prod_{x \in V} \{z \in \mathbb{C} \mid |z| \leq \|x\|\},$$

equipped with the product topology. Since the closed discs in $\mathbb{C}$ are compact, Z is compact by Tychonov's theorem. If $\varphi \in (V^*)_{\leq 1}$ then $|\varphi(x)| \leq \|x\| \; \forall x$, so that we have a map

$$f : (V^*)_{\leq 1} \to Z, \quad \varphi \mapsto \prod_{x \in V} \varphi(x).$$

Since the map $\varphi \mapsto \varphi(x)$ is continuous for each x, f is continuous (w.r.t. the weak-$*$ topology on $(V^*)_{\leq 1}$). It is trivial that V separates the points of V^*, thus f is injective. By definition,

[64] Leonidas Alaoglu (1914-1981). Greek mathematician. (Earlier versions due to Helly and Banach.)

a net $\{\varphi_\iota\}$ in $(V^*)_{\le 1}$ converges in the $\sigma(V^*, V)$-topology if and only if $\varphi_\iota(x)$ converges for all $x \in V$, and therefore if and only if $f(\varphi_\iota)$ converges. Thus $f : (V^*)_{\le 1} \to f((V^*)_{\le 1}) \subseteq Z$ is a homeomorphism.

Now let $z \in \overline{f((V^*)_{\le 1})} \subseteq Z$. Clearly, $|z_x| \le \|x\| \; \forall x \in X$. By Proposition A.12.2 there is a net in $f((V^*)_{\le 1})$ converging to z and therefore a net $\{\varphi_\iota\}$ in $(V^*)_{\le 1}$ such that $f(\varphi_\iota) \to z$. This means $\varphi_\iota(x) \to z_x \; \forall x \in V$. In particular $\varphi_\iota(\alpha x + \beta y) \to z_{\alpha x + \beta y}$, while also $\varphi_\iota(\alpha x + \beta y) = \alpha\varphi_\iota(x) + \beta\varphi_\iota(y) \to \alpha z_x + \beta z_y$. Thus the map $\psi : V \to \mathbb{C}, x \mapsto z_x$ is linear with $\|\psi\| \le 1$, to wit $\psi \in (V^*)_{\le 1}$ and $z = f(\psi)$. Thus $\overline{f((V^*)_{\le 1})} \subseteq f((V^*)_{\le 1})$, so that $f((V^*)_{\le 1}) \subseteq Z$ is closed.

Now we have proven that $(V^*)_{\le 1}$ is homeomorphic to the closed subset $f((V^*)_{\le 1})$ of the compact space Z, and therefore compact. ■

10.27 Remark We deduced Alaoglu's theorem from Tychonov's theorem, which is known to be equivalent to the axiom of choice (AC). But we only needed Tychonov as restricted to Hausdorff spaces, and the converse also holds. See Appendix B.5 where we also prove equivalence of these statements to the Ultrafilter Lemma (UL), a set theoretic axiom that is known to be strictly weaker than AC and its equivalents. UL also implies the Hahn-Banach theorem. □

10.28 Exercise Use Alaoglu's theorem to prove that every Banach space V over $\mathbb{F}$ admits a linear isometric bijection onto a closed subspace of $C(X, \mathbb{F})$ for some compact Hausdorff space X.

10.29 Exercise (i) Use Alaoglu's theorem to prove (ii)⇒(i) in Theorem 10.15.

(ii) Conclude that the closed unit ball of every Hilbert space is weakly compact.

(iii) Prove $\sigma(V, V^*) = \sigma(V^{**}, V^*) \restriction V$.

(iv) Use Theorem 10.30 and (iii) to prove (i)⇒(ii) in Theorem 10.15.

10.30 Theorem (Goldstine) [65] *If V is Banach then $V_{\le 1}$ is $\sigma(V^{**}, V^*)$-dense in $(V^{**})_{\le 1}$.*

The fairly non-trivial proof is relegated to the supplementary Section B.6.3.

We close by applying weak-$*$ topologies to $\Phi^\top$ and to transposes of operators:

10.31 Proposition *Let V be a Banach space and $W \subseteq V$, $\Phi \subseteq V^*$ linear subspaces. Then*

(i) $W^\perp \subseteq V^*$ is weak-$*$ closed.

(ii) $\left(\overline{\Phi}^{w^*}\right)^\top = \Phi^\top$.

(iii) $(\Phi^\top)^\perp = \overline{\Phi}^{w^*}$.

(iv) $\Phi^\top = \{0\}$ holds if and only if $\Phi \subseteq V^*$ is weak-$*$ dense.

Proof. (i) Let $\varphi, \{\varphi_\iota\}_{\iota \in I} \subseteq W^\perp$ such that $\varphi_\iota \stackrel{w^*}{\to} \varphi$. This means $\varphi_\iota(x) \to \varphi_\iota(x) \; \forall x \in V$. Thus $\varphi(w) = \lim \varphi_\iota(w) = 0$ for all $w \in W$, proving $\varphi \in W^\perp$.

(ii) Since $\overline{\Phi}^{w^*} \supseteq \Phi$, it is clear that $\left(\overline{\Phi}^{w^*}\right)^\top \subseteq \Phi^\top$. Let $x \in \Phi^\top$. If $\varphi \in \overline{\Phi}^{w^*}$ then there is a net $\{\varphi_\iota\}$ in Φ such that $\varphi_\iota \stackrel{w^*}{\to} \varphi$, thus $\varphi_\iota(x) \to \varphi(x)$. In view of $x \in \Phi^\top$ we have $\varphi_\iota(x) = 0$ for all ι, thus $\varphi(x) = 0$. Thus $x \in (\overline{\Phi}^{w^*})^\top$, proving the missing inclusion $\Phi^\top \subseteq \left(\overline{\Phi}^{w^*}\right)^\top$.

[65] Herman Heine Goldstine (1913-2004). American mathematician and computer scientist. Worked on very pure and very applied mathematics, like John von Neumann, with whom he collaborated on computers.

(iii) In view of (ii), $(\Phi^\top)^\perp = ((\overline{\Phi}^{w^*})^\top)^\perp$. Thus it suffices to prove $(\Phi^\top)^\perp = \Phi$ for weak-$*$ closed Φ. Since it is clear that $\Phi \subseteq (\Phi^\top)^\perp$, it remains to prove the converse inclusion. If $\varphi' \in (\Phi^\top)^\perp \backslash \Phi$, Corollary B.61, applied to the locally convex space (V^*, τ_{w^*}), gives a $\psi \in (V^*, \tau_{w^*})^*$ such that $\psi \upharpoonright \Phi = 0$ and $\psi(\varphi') \neq 0$. Now by Lemma 10.24 there is a unique $x \in V$ such that $\psi(\varphi) = \varphi(x)$ for all $\varphi \in V^*$. Clearly $x \neq 0$. In view of this we have $\varphi(x) = 0\ \forall \varphi \in \Phi$, thus $x \in \Phi^\top$. With $\varphi' \in (\Phi^\top)^\perp$ this implies $\psi(\varphi') = \varphi'(x) = 0$, which is a contradiction.

(iv) If $\overline{\Phi}^{w^*} = V^*$ then (ii) implies $\Phi^\top = (\overline{\Phi}^{w^*})^\top = (V^*)^\top = \{0\}$. And if $\Phi^\top = \{0\}$ then (iii) gives $\overline{\Phi}^{w^*} = (\{0\})^\perp = V^*$. ■

10.32 Theorem *Let V, W be Banach spaces.*

(i) If $A \in B(V, W)$ then A is injective if and only if $A^t \in B(W^*, V^*)$ has weak-$*$ dense image, thus $\overline{A^t W^*}^{w^*} = V^*$. (A priori this condition is weaker than norm-density!)

(ii) If $A \in B(W^*, V^*)$ then there exists $B \in B(V, W)$ such that $A = B^t$ if and only if A is weak-$*$-weak-$*$ continuous, i.e. continuous as a map $(W^*, \tau_{w^*}) \to (V^*, \tau_{w^*})$.

Proof. (i) This is immediate by combining Exercise 9.36(i) with Proposition 10.31(iv).

(ii) Let $B \in B(V, W)$, and let $\{\varphi_\iota\}$ be a net in W^* that converges to $\varphi \in W^*$ in the weak-$*$ topology, i.e. $\varphi_\iota(w) \to \varphi(w)$ for all $w \in W$. If $v \in V$ then $(B^t \varphi_\iota)(v) = \varphi_\iota(Bv) \to \varphi(Bv) = (B^t\varphi)(v)$, proving that $B^t \varphi_\iota \overset{w^*}{\to} B^t \varphi$, so that indeed B^t is weak-$*$-weak-$*$ continuous.

Now assume $A \in B(W^*, V^*)$ is weak-$*$-weak-$*$ continuous. Then for each $v \in V$ the linear functional $W^* \to \mathbb{F}, \varphi \mapsto (A\varphi)(v)$ is weak-$*$ continuous, thus by Lemma 10.24 there is a unique $w \in W$ such that $(A\varphi)(v) = \varphi(w)$. This defines a map $B : V \to W$ such that $(A\varphi)(v) = \varphi(Bv)$ for all $v \in V$, $\varphi \in W^*$. Thus $(A\varphi)(v) = (B^t\varphi)(v)$ for all v, φ, to wit $A = B^t$. Now Lemma 9.31 gives $\|B\| = \|B^t\| = \|A\| < \infty$. ■

10.33 Remark If V is reflexive, the weak and weak-$*$ topologies on V^* coincide by Proposition 10.22, so that with Corollary B.63 a linear subspace of V^* is norm-dense if and only if it is weak-$*$ dense. Thus for reflexive V the conditions of weak-$*$ density in Proposition 10.31(iv) and Theorem 10.32(i) reduce to norm-density. □

11 Hilbert space operators and their adjoints. Special classes of operators

In Section 9.5 we defined and studied the transpose $A^t \in B(F^*, E^*)$ of an operator $A \in B(E, F)$ between Banach spaces. For Hilbert spaces, we have natural identifications $H \cong H^*$, which leads to new aspects that give the theory of operators between Hilbert spaces a special flavor.

11.1 The adjoint of a bounded Hilbert space operator

If H is a Hilbert space then we have a canonical map $\gamma_H : H \to H^*$ given by $y \mapsto \varphi_y = \langle \cdot, y \rangle$. This map is antilinear and isometric, and by the representation Theorem 5.34 it is a bijection. This bijection in a sense makes the dual spaces of Hilbert spaces redundant to a large extent, so that it is desirable to eliminate them from considerations of the transpose:

11.1 Proposition *Let H_1, H_2 be Hilbert spaces. For $A \in B(H_1, H_2)$, define the Hilbert space adjoint $A^* : H_2 \to H_1$ as the composite map $H_2 \stackrel{\gamma_{H_2}}{\longrightarrow} H_2^* \stackrel{A^t}{\longrightarrow} H_1^* \stackrel{Y_{H_1}^{-1}}{\longrightarrow} H_1$. I.e. $A^* := \gamma_{H_1}^{-1} \circ A^t \circ \gamma_{H_2}$. Now*

(i) The map $A^* : H_2 \to H_1$ is linear and bounded, thus in $B(H_2, H_1)$.

(ii) The map $B(H_1, H_2) \to B(H_2, H_1), A \mapsto A^*$ is anti-linear.

(iii) For all $x \in H_1, y \in H_2$ we have $\langle Ax, y\rangle_2 = \langle x, A^*y\rangle_1$.

Proof. (i) Linearity of $A^* : H_2 \to H_1$ follows from its being the composite of the linear map A^t with the two anti-linear maps γ_{H_2} and $\gamma_{H_1}^{-1}$. Boundedness follows from $\|A^t\| = \|A\|$.

(ii) Additivity of $A \mapsto A^*$ is obvious. Let $A \in B(H_1, H_2), c \in \mathbb{C}, x \in H_2$. Then

$$(cA)^*(x) = \gamma_{H_1}^{-1} \circ (cA)^t \circ \gamma_{H_2}(x) = \gamma_{H_1}^{-1}(cA^t(\gamma_{H_2}(x))) = \bar{c}\gamma_{H_1}^{-1}(A^t(\gamma_{H_2}(x))) = \bar{c}A^*(x),$$

where we used the linearity of $A \mapsto A^t$ and anti-linearity of $\gamma_{H_1}^{-1}$, shows $(cA)^* = \bar{c}A^*$. (The anti-linearity of γ_{H_2} is irrelevant here.)

(iii) If $y \in H_2$ then $\gamma_{H_2}(y) \in H_2^*$ is the functional $\langle \cdot, y\rangle_2$. Then $(A^t \circ \gamma_{H_2})(y) \in H_1^*$ is the functional $x \mapsto \langle Ax, y\rangle_2$. Thus $z = A^*y = (\gamma_{H_1}^{-1} \circ A^t \circ \gamma_{H_2})(y) \in H_1$ is a vector such that $\langle x, z\rangle_1 = \langle Ax, y\rangle_2$ for all $x \in H_1$. This means $\langle x, A^*y\rangle_1 = \langle Ax, y\rangle_2 \; \forall x \in H_1, y \in H_2$, as claimed. ■

11.2 Remark Combining (iii) above with the Hellinger-Toeplitz theorem (Corollary 7.32), we see that a linear map $A : H_1 \to H_2$ has a Hilbert space adjoint if *and only if* it is bounded. □

There is a very useful bijection between bounded operators and bounded sesquilinear forms. It can be used to give an alternative (at least in appearance) construction of the adjoint A^* (and for many other purposes). It is based on the following observation: If $A \in B(H)$ satisfies $\langle Ax, y\rangle = 0$ for all $x, y \in H$ then $Ax = 0$ for all x, thus $A = 0$. Applying this to $A - B$ shows that $\langle Ax, y\rangle = \langle Bx, y\rangle \; \forall x, y$ implies $A = B$. Thus bounded operators are distinguished by their 'matrix elements' $\langle Ax, y\rangle$. This motivates the following developments.

11.3 Definition *Let V be an $\mathbb{F}$-vector space. A map $V \times V \to V, (x, y) \mapsto [x, y]$ is called sesquilinear if it is linear w.r.t. x and anti-linear w.r.t. y. A sesquilinear form $[\cdot, \cdot]$ is bounded if* $\sup_{\|x\|=\|y\|=1} |[x, y]| < \infty$.

11.4 Remark Recall that the inner product $\langle \cdot, \cdot\rangle$ on a (pre-)Hilbert space is sesquilinear and bounded by Cauchy-Schwarz. If $\mathbb{F} = \mathbb{R}$, the definition of course reduces to bilinearity. □

11.5 Proposition *Let H be a Hilbert space. Then there is a bijection between $B(H)$ and the set of bounded sesquilinear forms on H, given by $B(H) \ni A \mapsto [\cdot, \cdot]_A$, where $[x, y]_A = \langle Ax, y\rangle$.*

Proof. Let $A \in B(H)$. Sesquilinearity of $[\cdot, \cdot]_A = \langle Ax, y\rangle$ is an obvious consequence of sesquilinearity of $\langle \cdot, \cdot\rangle$ and linearity of A, and boundedness follows from Cauchy-Schwarz:

$$|[x, y]_A| = |\langle Ax, y\rangle| \leq \|Ax\|\|y\| \leq \|A\|\|x\|\|y\| \; \forall x, y.$$

Now let $[\cdot, \cdot]$ be a sesquilinear form bounded by M. Then for each $x \in H$, the map $\psi_x : H \to \mathbb{C}$, $y \mapsto \overline{[x, y]}$ is linear (thanks to the complex conjugation) and satisfies $|\psi_x(y)| \leq M\|y\|\|x\|$, thus $\psi_x \in H^*$. Thus by Theorem 5.34 there is a unique vector $z_x \in H$ such that $\psi_x = \varphi_{z_x}$, thus $\overline{[x, y]} = \psi_x(y) = \varphi_{z_x}(y) = \langle y, z_x\rangle \; \forall y$ and, taking complex conjugates, $\langle z_x, y\rangle = [x, y] \; \forall y$. Thus defining $A : H \to H$ by $Ax = z_x$ $\forall x$ we have $\langle Ax, y\rangle = [x, y]$ $\forall x, y$. Since the maps $x \mapsto \psi_x$ and $\psi_x \mapsto z_x$ are both anti-linear, their composite A is linear. And since $H \to H^*$, $z \mapsto \varphi_z$ is an isometry, we have $\|Ax\| = \|z_x\| = \|\varphi_{z_x}\| = \|\psi_x\| \le M\|x\|$, thus $A \in B(H)$. ■

11.6 Proposition *Let H be a Hilbert space and $A \in B(H)$. Then*

(i) There is a unique $B \in B(H)$ such that

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x, y \in H.$$

This B is denoted A^* and called the adjoint of A.

(ii) In particular $A = A^*$ if and only if $\langle Ax, y\rangle = \langle x, Ay\rangle$ $\forall x, y \in H$, in which case we call A self-adjoint.

Proof. (i) The map $(y, x) \mapsto \langle y, Ax\rangle$ is sesquilinear and bounded (by $\|A\|$). Thus by Proposition 11.5 there is a bounded $B \in B(H)$ such that $\langle By, x\rangle = \langle y, Ax\rangle$ $\forall x, y$. Taking the complex conjugate gives $\langle x, By\rangle = \overline{\langle By, x\rangle} = \overline{\langle y, Ax\rangle} = \langle Ax, y\rangle$, which is the wanted identity. Now (ii) is obvious. ■

11.7 Remark 1. In view of the identity $\langle Ax, y\rangle = \langle x, A^*y\rangle$ satisfied by the adjoint as defined above and Proposition 11.1(iii) (with $H_1 = H_2 = H$), it is clear that the two constructions of A^* give the same result (and in a sense are the same construction since both use Theorem 5.34).

2. Proposition 11.5 generalizes readily to a bijection between bounded linear maps $A : H_1 \to H_2$ and bounded sesquilinear forms $[\cdot, \cdot]$ on $H_1 \times H_2$. Then also the proof of Proposition 11.6 generalizes in this way and then produces the same adjoint $A^* \in B(H_2, H_1)$ as Proposition 11.1. (Of course $A \in B(H_1, H_2)$ can only be self-adjoint if $H_1 = H_2$.)

3. If $[\cdot, \cdot]$ is a sesquilinear form then also $[x, y]' := \overline{[y, x]}$ is a sesquilinear form, called the adjoint form. Looking at the above definition of A^*, one finds that A^* is the bounded operator associated with the form $[\cdot, \cdot]'$. Thus self-adjointness of A is equivalent to $[\cdot, \cdot]'_A = [\cdot, \cdot]_A$, i.e. $[\cdot, \cdot]_A$ being self-adjoint.

4. The following should be known from linear algebra, cf. e.g. [55]: If H is a Hilbert space, $A \in B(H)$ and E is an orthonormal basis for H then $\langle A^*e, f\rangle = \langle e, Af\rangle = \overline{\langle Af, e\rangle}$ for all $e, f \in E$. Thus the (possibly infinite) matrix describing A^* w.r.t. E is obtained from the matrix corresponding to A by transposition and complex conjugation. □

11.8 Lemma *The map $B(H) \to B(H), A \mapsto A^*$ satisfies*

(i) $(cA + dB)^* = \overline{c}A^* + \overline{d}B^*$ $\forall A, B \in B(H), c, d \in \mathbb{F}$ (antilinearity).

(ii) $(AB)^* = B^*A^*$ (anti-multiplicativity).

(iii) $A^{**} = A$ (involutivity).

(iv) $\mathbf{1}^* = \mathbf{1}$.

Proof. (i) Follows from

$$\begin{aligned}\langle x, (\overline{c}A^* + \overline{d}B^*)y\rangle &= c\langle x, A^*y\rangle + d\langle x, B^*y\rangle = c\langle Ax, y\rangle + d\langle Bx, y\rangle = \langle (cA + dB)x, y\rangle \\ &= \langle x, (cA + dB)^*y\rangle.\end{aligned}$$

(ii) For all $x, y \in H$ we have $\langle x, (AB)^*y\rangle = \langle (AB)x, y\rangle = \langle Bx, A^*y\rangle = \langle x, B^*A^*y\rangle$.

(iii) Complex conjugating $\langle Ax, y\rangle = \langle x, A^*y\rangle$ gives

$$\langle y, Ax\rangle = \overline{\langle Ax, y\rangle} = \overline{\langle x, A^*y\rangle} = \langle A^*y, x\rangle,$$

which shows that A is an adjoint of A^*. Uniqueness of the adjoint now implies $A^{**} = A$.

(iv) Obvious. ■

11.9 Proposition *Let H be a Hilbert space. Then for all $A \in B(H)$ we have*

(i) $\|A^*\| = \|A\|$. *(The $*$-operation is isometric.)*

(ii) $\|A^*A\| = \|A\|^2$. *("C^*-identity")*

Proof. (i) Similarly to Lemma 9.31, using (5.2) we have

$$\|A^*\| = \sup_{\|x\|=\|y\|=1} |\langle A^*x, y\rangle| = \sup_{\|x\|=\|y\|=1} |\langle x, Ay\rangle| = \sup_{\|x\|=\|y\|=1} |\langle Ay, x\rangle| = \|A\|.$$

(ii) On the one hand, $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$, where we used (i). On the other hand,

$$\|A\|^2 = \Big(\sup_{\|x\|=1} \|Ax\|\Big)^2 = \sup_{\|x\|=1} \|Ax\|^2 = \sup_{\|x\|=1} \langle Ax, Ax\rangle = \sup_{\|x\|=1} \langle A^*Ax, x\rangle \leq \|A^*A\|,$$

where the last inequality follows from Cauchy-Schwarz. ■

The following is a Hilbert space version of the results of Section 9.5, but much easier:

11.10 Lemma *Let H_1, H_2 be Hilbert spaces and $A \in B(H_1, H_2)$. Then*

(i) A^* *is invertible if and only if A is invertible, in which case* $(A^*)^{-1} = (A^{-1})^*$.

(ii) $\ker A^* = (AH_1)^\perp \subseteq H_2$. *Thus A^* is injective if and only if A has dense image:* $\overline{AH_1} = H_2$. *Analogously, A is injective if and only if A^* has dense image.*

(iii) A^* *has closed image if and only if A has closed image.*

(iv) A^* *is surjective if and only if A is bounded below. Similarly for* $A \leftrightarrow A^*$.

Proof. (i) If A is invertible then applying $*$ to $AA^{-1} = \mathbf{1} = A^{-1}A$ gives $(A^{-1})^*A^* = \mathbf{1} = A^*(A^{-1})^*$, as claimed. The converse implication follows from this together with $A^{**} = A$.

(ii) For $y \in H_2$ we have

$$A^*y = 0 \;\Leftrightarrow\; \langle A^*y, x\rangle = 0\ \forall x \in H_1 \;\Leftrightarrow\; \langle y, Ax\rangle = 0\ \forall x \in H_1 \;\Leftrightarrow\; y \in (AH_1)^\perp.$$

Thus A^* is injective if and only if $(AH_1)^\perp = \{0\}$, which is equivalent to $\overline{AH_1} = H_2$ by Exercise 5.29(i). Applying the fact just proven to A^* and using $A^{**} = A$ proves A injective $\Leftrightarrow$ A^* has dense image.

(iii) We will prove that closedness of AH_1 implies closedness of A^*H_2. Replacing A by A^* then gives the converse implication. Put $H_{10} = \ker A$ (which is closed) and $H_{11} = H_{10}^\perp$. Then we have a direct sum decomposition $H_1 = H_{10} \oplus H_{11}$ by Theorem 5.27. By assumption $H_{21} = AH_1$ is closed, so that with $H_{20} = H_{21}^\perp$ we also have $H_2 = H_{20} \oplus H_{21}$. Now A maps H_{11} injectively (since $H_{11} \cap \ker A = H_{11} \cap H_{11}^\perp = \{0\}$) onto $H_{21} = AH_1$. This defines an operator $A' \in B(H_{11}, H_{21})$ that is injective and surjective, thus invertible, so that A'^* is invertible by (i) and therefore has closed image $A'^*H_{21} = H_{11} \subseteq H_1$. Now closedness of A^*H_2 follows once we prove that $A^*(x_{20} + x_{21}) = A'^*x_{21}$ for all $x_{2i} \in H_{2i}$. Since A^* vanishes on $H_{20} = (AH_1)^\perp$ by (ii), it remains to prove that $A^* \restriction H_{21}$ coincides with A'^* followed by the inclusion $H_{11} \hookrightarrow H_1$. This follows from the computation

$$\langle (x_{10} + x_{11}), A^*x_{21}\rangle = \langle A(x_{10} + x_{11}), x_{21}\rangle = \langle Ax_{11}, x_{21}\rangle = \langle A'x_{11}, x_{21}\rangle = \langle x_{11}, A'^*x_{21}\rangle,$$

where $x_{ij} \in H_{ij}$.

(iv) By Exercise 7.43, A is bounded below if and only if it is injective and has closed image. By (ii), injectivity of A is equivalent to A^* having dense image, and by (iii) closedness of the image of A is equivalent to closedness of the image of A^*. The proof is concluded by appealing to the trivial fact that surjectivity of A^* is equivalent to the combination of closedness and density of its image. ■

If H is finite-dimensional, closedness of the images is automatic, while dense image is equivalent to surjectivity and boundedness below to injectivity. Thus (ii) and (iv) reduce to the well known facts from linear algebra that A^* is injective (surjective) if and only if A is surjective (injective). Finally, obvious analogues of (i)-(iv) also hold for $A \in B(H_1, H_2)$ with $H_1 \neq H_2$.

11.11 Exercise Let H be a Hilbert space and $A \in B(H)$.

(i) Show by example that injectivity of A and of A^* does not imply invertibility of A.

(ii) Prove: If A is bounded below and A^* is injective (or vice versa) then A is invertible.

11.12 Exercise Let H be a Hilbert space. Equip $\widehat{H} = H \oplus H$ with the obvious Hilbert space structure.

(i) Let $A \in B(H)$ and A^* its adjoint. Show (without using boundedness of A^*) that the graph $\mathfrak{G}(A^*) \subseteq \widehat{H}$ of A^* is the orthogonal complement in $\widehat{H}$ of a certain linear subspace.

(ii) Conclude that A^* is bounded. (Never mind that there are simpler ways of seeing this.)

11.2 Unitaries, isometries, coisometries, partial isometries

11.13 Lemma *Let H_1, H_2 be Hilbert spaces and $A \in B(H_1, H_2)$. Then*

(i) A is an isometry if and only if $A^*A = \mathrm{id}_{H_1}$, i.e. A^* is a left inverse of A.

(ii) If A is an isometry then

(α) $AH_1 \subseteq H_2$ is closed.

(β) $P = AA^* \in B(H_2)$ is an orthogonal projection, and $PH_2 = AH_1$.

(γ) $\ker A^* = (AH_1)^\perp = (1-P)H_2$.

(δ) The restriction $A^* \upharpoonright AH_1 : AH_1 \to H_1$ is unitary.

(iii) A is unitary if and only if $A^*A = \mathrm{id}_{H_1}$ and $AA^* = \mathrm{id}_{H_2}$.

Proof. (i) By definition, A is an isometry if $\langle Ax, Ay\rangle_2 = \langle x, y\rangle_1$ for all $x, y \in H_1$. Since the l.h.s. equals $\langle x, A^*Ay\rangle_1$, A is an isometry if and only if $\langle A^*Ax, y\rangle_1 = \langle x, y\rangle_1$ for all $x, y \in H_1$, which is equivalent to $A^*A = \mathrm{id}_{H_1}$.

(ii) (α) This is just Corollary 3.23. (β) We have $P^* = (AA^*)^* = A^{**}A^* = AA^* = P$ and $P^2 = AA^*AA^* = AA^* = P$, so that P is an orthogonal projection. We have $PAx = AA^*Ax = Ax$ for all $x \in H_1$, thus $AH_1 \subseteq PH_2$. And if $y \in PH_2$ then $y = Py = AA^*y$ (since $P^2 = P$), thus $y = A(A^*y)$ so that $PH_2 \subseteq AH_1$. (γ) The equality $\ker A^* = (AH_1)^\perp$ is Lemma 11.10 and the second comes from $(AH_1)^\perp = (PH_1)^\perp = (1-P)H_1$. ($\delta$) If $y \in AH_1$ then there is a unique $x \in H_1$ such that $y = Ax$. Now $A^*y = A^*Ax = x$. Thus $A^* \upharpoonright AH_1 : AH_1 \to H_1$ is isometric and surjective, thus unitary.

(iii) By definition, A is unitary if and only if it is isometric and surjective. Thus $A^*A = \mathrm{id}_{H_1}$ by (i) and $AA^* = \mathrm{id}_{H_2}$ by (ii)(β). Conversely, if $A^*A = \mathrm{id}_{H_1}$ and $AA^* = \mathrm{id}_{H_2}$ then A is isometric by (i) and surjective by (ii)(β), thus unitary. ■

11.14 Definition *If H_1, H_2 are Hilbert spaces then $A \in B(H_1, H_2)$ is called coisometry if $AA^* = \mathrm{id}_{H_2}$.*

It is clear that A is a coisometry if and only if A^* is an isometry. Thus by the above, we have that if A is a coisometry then $A \restriction (\ker A)^\perp$ is unitary, and also the converse is easily checked. This suggests the following generalization:

11.15 Definition *$V \in B(H_1, H_2)$ is a partial isometry if $V_{|(\ker V)^\perp} : (\ker V)^\perp \to H_2$ is an isometry.*

It is obvious that every isometry and every coisometry is a partial isometry. And the isometries, respectively coisometries, are just the partial isometries that are injective, respectively surjective. Partial isometries will be put to use shortly.

For simplicity, the following exercise is stated for $H = H_1 = H_2$, but it generalizes literally to $V \in B(H_1, H_2)$, where $H_1 \neq H_2$.

11.16 Exercise Prove that for $V \in B(H)$, the following are equivalent.

(i) V is a partial isometry.

(ii) V^* is a partial isometry.

(iii) V^*V is an orthogonal projection.

(iv) VV^* is an orthogonal projection.

(v) $VV^*V = V$ (trivially equivalent to $V^*VV^* = V^*$).

11.3 Polarization revisited

By Proposition 11.5 there is a bijection between bounded operators and bounded sesquilinear forms. For this reason, the following is useful:

11.17 Lemma *(i) Let $[\cdot, \cdot] : V \times V \to \mathbb{C}$ be a sesquilinear form. Then*

$$[x, y] = \frac{1}{4} \sum_{k=0}^{3} i^k [x + i^k y, x + i^k y] \quad \forall x, y \in V. \tag{11.1}$$

And $[\cdot, \cdot] : V \times V \to \mathbb{C}$ is self-adjoint, i.e. $\overline{[x, y]} = [y, x]\ \forall x, y$, if and only if $[x, x] \in \mathbb{R}\ \forall x$.

(ii) Every bilinear form $[\cdot, \cdot] : V \times V \to \mathbb{R}$ that is symmetric, i.e. $[x, y] = [y, x]\ \forall x, y$, satisfies

$$[x, y] = \frac{1}{4}\Big([x + y, x + y] - [x - y, x - y]\Big) \quad \forall x, y \in V.$$

Proof. The polarization identities are proven by the same computations as for (5.5), the proof of which only used sesquilinearity of $\langle \cdot, \cdot \rangle$, and (5.4), for which we also used the symmetry $\langle x, y \rangle = \langle y, x \rangle$ of inner products over $\mathbb{R}$.

If a sesquilinear form $[\cdot, \cdot]$ is selfadjoint and $x \in V$ then $\overline{[x, x]} = [x, x]$, thus $[x, x] \in \mathbb{R}$ for all $x \in V$. Conversely, if $[x, x] \in \mathbb{R}\ \forall x$ then with (11.1) we have

$$\overline{[x, y]} = \frac{1}{4} \sum_{k=0}^{3} i^{-k} [x + i^k y, x + i^k y] = \frac{1}{4} \sum_{k=0}^{3} i^k [x + i^{-k} y, x + i^{-k} y] = \frac{1}{4} \sum_{k=0}^{3} i^k [i^k x + y, i^k x + y] = [y, x],$$

thus $[\cdot, \cdot]$ is self-adjoint. ∎

11.18 Remark The above shows that in the case $\mathbb{F} = \mathbb{C}$ we can omit axiom (ii) in Definition 5.1 if we replace the assumption of linearity in x by sesquilinearity in x, y, since (ii) then follows from the positivity assumption $\langle x, x\rangle \geq 0\ \forall x$. □

Now we can show that Hilbert space operators tend to be determined by their diagonal elements[66] $\langle Ax, x\rangle$:

11.19 Lemma *Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A, B \in B(H)$.*

(i) If $\langle Ax, x\rangle = 0$ for all $x \in H$ and either $\mathbb{F} = \mathbb{C}$ or $A = A^*$ then $A = 0$.

(ii) If $\langle Ax, x\rangle = \langle Bx, x\rangle$ for all $x \in H$ and either $\mathbb{F} = \mathbb{C}$ or $A = A^*$ and $B = B^*$ then $A = B$.

Proof. (i) (i) Apply Lemma 11.17 to $[\cdot, \cdot]_A = \langle A\cdot, \cdot\rangle$, which is sesquilinear over $\mathbb{C}$ and symmetric if $\mathbb{F} = \mathbb{R}$ and $A = A^*$ by $\langle Ax, y\rangle = \langle x, Ay\rangle = \overline{\langle Ay, x\rangle} = \langle Ay, x\rangle$. For (ii) apply (i) to $A - B$. ■

11.20 Remark 1. Every antisymmetric real matrix $A \in M_{n\times n}(\mathbb{R})$, e.g. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, satisfies $\langle Ax, x\rangle = 0\ \forall x \in \mathbb{R}^n$. Thus the condition '$\mathbb{F} = \mathbb{C}$ or $A = A^*$' cannot be dropped.

2. If $A \in B(H)$ and we define $f : H \to \mathbb{F}$ by $f(x) = \langle Ax, x\rangle$, we have seen that A is completely determined by f (if $\mathbb{F} = \mathbb{C}$, while f determines only $A + A^*$ if $\mathbb{F} = \mathbb{R}$). This raises the natural problem of characterizing the functions $H \to \mathbb{F}$ of the form $x \mapsto \langle Ax, x\rangle$. □

11.21 Exercise Let H be a Hilbert space over $\mathbb{F}$ and $f : H \to \mathbb{F}$ a function. Prove that the following are equivalent:

(i) There exists $A \in B(H)$ such that $f(x) = \langle Ax, x\rangle\ \forall x \in H$.

(ii) The function f satisfies

(α) $f(x + y) + f(x - y) = 2f(x) + 2f(y)\ \forall x, y \in H$.

(β) $f(cx) = |c|^2 f(x)\ \forall x \in H,\ c \in \mathbb{F}$.

(γ) $|f(x)| \leq C\|x\|^2$ for some $C \geq 0$.

11.4 A little more on self-adjoint operators

The following shows that the self-adjoint operators are analogues of the real numbers, in some sense:

11.22 Proposition *Let H be a Hilbert space over $\mathbb{C}$ and $A \in B(H)$. Then $A = A^*$ is equivalent to $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in H$.*

Proof. $\Rightarrow$ If $A = A^*$ then $\langle Ax, x\rangle = \langle x, A^*x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}\ \forall x$, and the claim follows. $\Leftarrow$ If $\langle Ax, x\rangle \in \mathbb{R}\ \forall x$ then $\langle A^*x, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle} = \langle Ax, x\rangle\ \forall x$, and Lemma 11.19(ii) (applicable since $\mathbb{F} = \mathbb{C}$) gives $A = A^*$. ■

Note that if $\mathbb{F} = \mathbb{R}$, the statement $\langle Ax, x\rangle \in \mathbb{R}$ is trivially true and therfore implies nothing, in particular not $A = A^*$, again by Remark 11.20.

[66] This language is a bit misleading since we looking at $\langle Ax, x\rangle$ for all x, not just $x \in E$ for some fixed basis E.

11.23 Exercise (i) Let H be a Hilbert space, $A \in B(H)$ and $K \subseteq H$ a closed subspace such that $AK \subseteq K$. (We say K is A-invariant.) Prove that $AK^\perp \subseteq K^\perp$ is equivalent to $A^*K \subseteq K$.

In this situation, K is called *reducing* since then $A \cong A_{|K} \oplus A_{|K^\perp}$.

(ii) Deduce that every invariant subspace of a self-adjoint operator is reducing.

11.5 Normal operators

If $A \in B(H)$ then A is self-adjoint if $A = A^*$ and unitary if $A^*A = AA^* = \mathbf{1}$. It is obvious that either of these properties implies the following:

11.24 Definition *A Hilbert space operator $A \in B(H)$ is called normal if $AA^* = A^*A$.*

11.25 Remark On every complex Hilbert space $H \neq \{0\}$ there are normal operators that are neither self-adjoint nor unitary. Non-normal operators exist whenever $\dim H \geq 2$. □

11.26 Lemma *Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(H)$. Then*

$$A \text{ is normal} \quad \Leftrightarrow \quad \|Ax\| = \|A^*x\| \text{ for all } x \in H \quad \Leftrightarrow \quad A^* \text{ is normal.}$$

Proof. For all $x \in H$ we have

$$\|A^*x\|^2 = \langle A^*x, A^*x\rangle = \langle AA^*x, x\rangle, \qquad \|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle,$$

so that $\|A^*x\| = \|Ax\| \ \forall x$ is equivalent to $\langle A^*Ax, x\rangle = \langle AA^*x, x\rangle \ \forall x$ and therefore is implied by normality of A. If $\|Ax\| = \|A^*x\| \ \forall x$ then $\langle AA^*x, x\rangle = \langle A^*Ax, x\rangle \ \forall x$, so that Lemma 11.19(ii) gives $AA^* = A^*A$. (This also holds for $\mathbb{F} = \mathbb{R}$ since AA^* and A^*A are self-adjoint.) ■

11.27 Proposition *Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(H)$ normal. Then*

(i) $\ker A = \ker A^* = (AH)^\perp = (A^*H)^\perp$.

(ii) With $H_0 = \ker A$ and $H' = H_0^\perp$ we have $AH' \subseteq H'$ and $A^*H' \subseteq H'$ with dense images. (In particular, $\ker A$ is a reducing subspace.)

(iii) A is injective if and only if it has dense image.

(iv) A is invertible $\Leftrightarrow$ A is bounded below $\Leftrightarrow$ A is surjective.

Proof. (i) The first equality is immediate from Lemma 11.26. The rest follows by applying Lemma 11.10 to A and A^*.

(ii) Taking the orthogonal complement of $(AH)^\perp = \ker A = H_0$ gives $(AH)^{\perp\perp} = \overline{AH} = H'$ and similarly for A^*.

(iii) In view of $(AH)^\perp = \ker A$ established in (i), injectivity of A is equivalent to $(AH)^\perp = \{0\}$, which is equivalent to density of the image AH by Exercise 5.29(i).

(iv) Invertibility implies boundedness below and surjectivity, cf. Proposition 7.41. If a normal operator is bounded below, it is injective, so that it has dense image by (iii). Now boundedness below and dense image imply invertibility by Proposition 7.41. And surjectivity implies dense image, thus injectivity by (iii). Now injectivity and surjectivity give invertibility by the BIT. ■

11.28 Remark In Corollary 17.25 we will characterize invertibility of $A\restriction(\ker A)^\perp$ in terms of the spectrum of A. □

11.29 Exercise Describe (i) the unitary self-adjoint operators, (ii) the normal partial isometries.

11.30 Exercise Give an example of a normal operator $A \in B(H)$ such that there is a non-trivial A-invariant subspace $K \subseteq H$ that is not reducing. (See Exercise 11.23 for the terminology.)

11.31 Exercise Let H be a Hilbert space and $A \in B(H)$.

(i) Assuming $A = A^*$, prove $\|A^{2^n}\| = \|A\|^{2^n}$ $\forall n \in \mathbb{N}$.

(ii) Generalize (i) to all normal A.

11.6 Numerical range and radius

In view of Lemma 11.17 and Proposition 11.5, a Hilbert space operator A can be recovered completely from its diagonal elements $\langle Ax, x\rangle$ when $\mathbb{F} = \mathbb{C}$ or $A = A^*$. We should therefore be able to prove a more quantitative statement improving on Lemma 11.19 (and implying it).

11.32 Definition *Let H be a Hilbert space and $A \in B(H)$. Then we define*

- *the numerical range of A as $W(A) = \{\langle Ax, x\rangle \mid x \in H, \|x\| = 1\}$,*
- *the numerical radius of A as $|||A||| = \sup_{\lambda \in W(A)} |\lambda| = \sup_{\|x\|=1} |\langle Ax, x\rangle|$.*

(In quantum mechanics, cf. e.g. [92], $W(A)$ is the set of expectation values of A (if $A = A^*$).)

If $\mathbb{F} = \mathbb{C}$, Lemma 11.19(i) and Proposition 11.22 can now be restated concisely: For $A \in B(H)$ we have $W(A) = \{0\}$ if and only if $A = 0$ and $W(A) \subseteq \mathbb{R}$ if and only if $A = A^*$.

We now focus on $|||\cdot|||$ and refer to Appendix B.12.1 for more on $W(A)$.

11.33 Lemma *Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(H)$. Then*

(i) $|||A||| \le \|A\|$.

(ii) $|\langle Ax, x\rangle| \le \|x\|^2\, |||A|||$ *for all* $x \in H$.

(iii) If $|||A||| = 0$ *and either* $\mathbb{F} = \mathbb{C}$ *or* $A = A^*$ *then* $A = 0$.

Proof. (i) If $\|x\| = 1$ then $|\langle Ax, x\rangle| \le \|Ax\|\|x\| \le \|A\|$.

(ii) For $x = 0$ this is clear. For $x \neq 0$ we have $\|x\|^{-2}|\langle Ax, x\rangle| = |\langle A\frac{x}{\|x\|}, \frac{x}{\|x\|}\rangle| \le |||A|||$.

(iii) In view of (ii) this is just a restatement of Lemma 11.19(i). ■

11.34 Proposition *Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(H)$. Then*

(i) If $A = A^*$ *then* $\|A\| = |||A|||$.

(ii) If $\mathbb{F} = \mathbb{C}$ *then* $\frac{1}{2}\|A\| \le |||A|||$.

(iii) If $\mathbb{F} = \mathbb{C}$ *and* A *is normal then* $\|A\| = |||A|||$.

Proof. (i) If $x, y \in H$ are unit vectors then

$$\begin{aligned} |\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle| &\le |\langle A(x+y), x+y\rangle| + \langle A(x-y), x-y\rangle| \\ &\le (\|x+y\|^2 + \|x-y\|^2)\, |||A||| \\ &= 2\big(\|x\|^2 + \|y\|^2\big)\, |||A||| \;=\; 4\, |||A|||, \end{aligned}$$

where we used Lemma 11.33(ii), the parallelogram identity (5.3) and $\|x\| = \|y\| = 1$. Inserting

$$\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle = 2\langle Ax, y\rangle + 2\langle Ay, x\rangle \quad \forall x, y \in H,$$

proven by direct computation, we have

$$|\langle Ax, y\rangle + \langle Ay, x\rangle| \leq 2 \,|||\, A \,|||\, .$$

Assuming $Ax \neq 0$ and putting $y = \frac{Ax}{\|Ax\|}$, this becomes

$$\left| \|Ax\| + \|Ax\|^{-1}\langle A^2x, x\rangle \right| \leq 2 \,|||\, A \,||| \quad \text{whenever} \quad \|x\| = 1, \; Ax \neq 0. \tag{11.2}$$

If now $A = A^*$ then $\langle A^2x, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2$, so that (11.2) reduces to $\|Ax\| \leq |||A|||$ when $\|x\| = 1$, whence $\|A\| \leq |||A|||$. Combining this with Lemma 11.33(i), we are done.

(ii) If we replace A in (11.2) by αA where $|\alpha| = 1$ then $\|Ax\|$ and $|||A|||$ are unchanged but $\langle A^2x, x\rangle$ acquires a factor α^2. Since $\mathbb{F} = \mathbb{C}$, we can choose α such that $\alpha^2\langle A^2x, x\rangle = |\langle A^2x, x\rangle| \geq 0$. Then (11.2) becomes $\|Ax\| + \|Ax\|^{-1}|\langle A^2x, x\rangle| \leq 2 \,|||\, A|||$, which implies $\|Ax\| \leq 2 \,|||\, A|||$ for all unit vectors x with $Ax \neq 0$. This also holds if $Ax = 0$, thus $\|A\| = \sup_{\|x\|=1} \|Ax\| \leq 2 \,|||\, A|||$.

(iii) See the following exercise. ■

11.35 Exercise Let H be a complex Hilbert space and $A \in B(H)$. Prove:

(i) $|||A^2||| \leq |||A|||^2$. Hint: Adapt the proof of Proposition 11.34(ii).

(ii) If A is normal then $|||A||| = \|A\|$. Hint: Prove and use $\|A\|^{2^n} \leq 2 \,|||\, A \,|||^{2^n} \;\; \forall n \in \mathbb{N}$.

For $\mathbb{F} = \mathbb{R}$, Proposition 11.34(ii) and both parts of Exercise 11.35 are false, a counterexample being the A in Remark 11.20, satisfying $\|A\| = 1 = |||A^2|||$ and $AA^* = A(-A) = (-A)A = A^*A$, but $|||A||| = 0$. Over $\mathbb{C}$ the following shows that Proposition 11.34(ii) is optimal without further assumption on A:

11.36 Exercise Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in B(\mathbb{C}^2)$. Prove $W(A) = \{z \in \mathbb{C} \mid |z| \leq \frac{1}{2}\}$ and $|||A||| = \frac{1}{2}\|A\|$.

11.7 Positive operators and their square roots

Just as self-adjoint operators are the 'real' elements of $B(H)$, there also are positive ones:

11.37 Definition *If H is a Hilbert space then $A \in B(H)$ is called positive, abbreviated $A \geq 0$, if $\langle Ax, x\rangle \geq 0$ for all $x \in H$. If $\mathbb{F} = \mathbb{R}$ we also require $A = A^*$.*

11.38 Remark 1. Positivity of A is equivalent to $W(A) \subseteq [0, +\infty)$.

2. By Proposition 11.22, the assumption $\langle Ax, x\rangle \geq 0 \;\forall x$ automatically implies $A = A^*$ if $\mathbb{F} = \mathbb{C}$. But this is not true for $\mathbb{F} = \mathbb{R}$, and it seems undesirable to consider matrices as in Remark 11.20 as positive, for example since positivity should be preserved by extending scalars from $\mathbb{R}$ to $\mathbb{C}$. But also the definition without self-adjointness can be found in the literature. □

11.39 Exercise Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Prove:

(i) If $A_1, A_2 \in B(H)$ are positive and $\lambda_1, \lambda_2 \geq 0$ then $\lambda_1 A_1 + \lambda_2 A_2$ is positive.

(ii) For each $A \in B(H)$ we have $A^*A \geq 0$. In particular, $A = A^* \Rightarrow A^2 \geq 0$.

(iii) If $A \in B(H)$ is positive then BAB^* is positive for every $B \in B(H)$.

(iv) If $A, B \in B(H)$ are positive and $A + B = 0$ then $A = B = 0$.

(v) If $A \geq 0$ then $A^k \geq 0$ for all $k \in \mathbb{N}$.

11.40 THEOREM *If H is a Hilbert space and $A \in B(H)$ is positive then there is a unique positive $B \in B(H)$ such that $B^2 = A$. We call B the square root $A^{1/2}$ of A.*

Proof. Existence: It suffices to prove the claim under the additional assumption $\|A\| \leq 1$ since then for general $A \neq 0$ we can put $A^{1/2} = \|A\|^{1/2}\left(\frac{A}{\|A\|}\right)^{1/2}$.

Let $A \geq 0, \|A\| \leq 1$. Then for $x \in H, \|x\| = 1$ we have $0 \leq \langle Ax, x\rangle \leq 1$, thus $\langle(\mathbf{1} - A)x, x\rangle = 1 - \langle Ax, x\rangle \in [0,1]$, so that $\mathbf{1} - A \geq 0$. Furthermore, Proposition 11.34(i) implies $\|\mathbf{1} - A\| = \sup_{\|x\|=1}\langle(\mathbf{1} - A)x, x\rangle \leq 1$. Thus $A \mapsto \mathbf{1} - A$ is an involutive bijection from the set $\{A \geq 0 \mid \|A\| \leq 1\}$ to itself, so that it suffices to define $(\mathbf{1} - A)^{1/2}$ whenever $A \geq 0, \|A\| \leq 1$. (This is advantageous since $z \mapsto (1-z)^{1/2}$ is analytic at $z = 0$ while $z \mapsto z^{1/2}$ is not.)

The function $z \mapsto (1-z)^{1/2}$ is infinitely differentiable at $z = 0$, with (formal) Taylor series

$$(1-z)^{1/2} = \sum_{k=0}^{\infty} c_k z^k = 1 - \frac{1}{2}\left(\frac{z}{1!} + \frac{1}{2}\frac{z^2}{2!} + \frac{1}{2}\frac{3}{2}\frac{z^3}{3!} + \cdots\right), \tag{11.3}$$

where the c_k can be found by explicit computation or appeal to Newton's binomial theorem, cf. e.g. [57, Vol. 1, Theorem 7.6.4]. We see that $c_0 = 1$ and $c_k < 0$ for all $k \geq 1$. The power series has convergence radius one and does converge to $(1-z)^{1/2}$ whenever $|z| < 1$. This can be seen invoking either complex analysis or Newton's theorem. As $z \nearrow 1$, the l.h.s. converges to zero, thus $\sum_{k=1}^{\infty} c_k = -1$, implying $\sum_{k=0}^{\infty}|c_k| = 2$ and $|c_k| \leq 1\ \forall k$. Thus for $\|A\| \leq 1$ we have $\sum_{k=0}^{\infty}|c_k|\|A^k\| < \infty$, so that $\sum_{k=0}^{\infty} c_k A^k$ converges in norm by Proposition 3.15. We interpret the sum as $(1-A)^{1/2}$. To see that this is justified we note that squaring (11.3) gives $1 - z = \left(\sum_{k=0}^{\infty} c_k z^k\right)^2$, which by absolute convergence also holds with z replaced by A. Since A is self-adjoint with $\|A\| \leq 1$, also A^k has these properties for all $k \geq 1$. Thus for all k and $x \in H, \|x\| = 1$ we have $\langle A^k x, x\rangle \in [-1,1]$. (Actually $\cdots \in [0,1]$ since $A^k \geq 0\ \forall k \in \mathbb{N}$ by Exercise 11.39(v), but we don't need this.) With $\sum_{k=1}^{\infty}|c_k| = 1$ it follows that $\sum_{k=1}^{\infty} c_k\langle A^k x, x\rangle \in [-1,1]$, so that with $c_0 = 1$ we have

$$\left\langle(\mathbf{1} - A)^{1/2}x, x\right\rangle = \left\langle\left(\sum_{k=0}^{\infty} c_k A^k\right)x, x\right\rangle = \sum_{k=0}^{\infty} c_k\langle A^k x, x\rangle \geq 0.$$

Thus $(\mathbf{1} - A)^{1/2} \geq 0$. Finally defining $A^{1/2} = (\mathbf{1} - (\mathbf{1} - A))^{1/2}$ we are done.

Uniqueness: Let $A \in B(H)$ be positive and $B = A^{1/2}$ as constructed above. Let also C satisfy $C^2 = A$ and $C \geq 0$. Then $CA = CC^2 = C^2C = AC$, and since B is a limit of polynomials in A, we have $BC = CB$. Using this and $B^2 = A = C^2$ we have

$$(B-C)B(B-C) + (B-C)C(B-C) = (B-C)(B+C)(B-C) = (B^2 - C^2)(B-C) = 0.$$

Since $(B-C)B(B-C)$ and $(B-C)C(B-C)$ are positive by Exercise 11.39(iii), Exercise 11.39(iv) gives $(B-C)B(B-C) = (B-C)C(B-C) = 0$. Thus also their difference $(B-C)^3$ vanishes, and so does $(B-C)^4$. Since $B-C$ is self-adjoint, we have $\|B-C\|^4 = \|(B-C)^2\|^2 = \|(B-C)^4\| = 0$, thus $B = C$. ■

11.41 Remark 1. For a proof of Theorem 11.40 that uses the Weierstrass approximation theorem instead of a power series, cf. e.g. [118, Lemma 3.2.10 & Proposition 3.2.11].

2. In Section 17 we will develop a less ad hoc and much more general approach to applying (continuous) functions to (normal) operators, but this will require much work. □

11.8 The polar decomposition

Every complex number z can be written as $z = ru$, where $|u| = 1$ and $r \geq 0$. For $z \neq 0$ this decomposition is unique. It is natural to ask whether there is an analogue for $A \in B(H)$.

11.42 Definition *Let H be a Hilbert space and $A \in B(H)$. Then we define the absolute value $|A| = (A^*A)^{1/2}$. (Apply Theorem 11.40 to A^*A, which is positive by Exercise 11.39(ii).)*

11.43 Exercise Let $A \in B(H)$ be invertible. Prove (directly, without Proposition 11.44):

(i) $B \in B(H)$ is invertible if and only if B^2 is invertible.

(ii) $|A|$ is invertible.

(iii) $U = A|A|^{-1}$ is unitary.

Thus we have the polar decomposition $A = U|A|$ with U unitary and $|A|$ positive and invertible.

If A is not invertible, a form of polar decomposition still exists, but is more subtle:

11.44 Proposition (Polar decomposition) *Let H be a Hilbert space and $A \in B(H)$.*

(i) There exists a unique partial isometry V such that $A = V|A|$ and $\ker A = \ker V$.

(ii) If A is injective (invertible) then V is an isometry (unitary).

(iii) In addition, we have $|A| = V^*A$. [This follows trivially from (i) only if V is unitary.]

Proof. (i) For each $x \in H$ we have

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle = \langle (A^*A)^{1/2}x, (A^*A)^{1/2}x\rangle = \langle |A|x, |A|x\rangle = \||A|x\|^2, \quad (11.4)$$

implying $\ker |A| = \ker A$. If $|A|x = |A|x'$ then $|A|(x - x') = 0$, thus $x - x' \in \ker |A| = \ker A$, so that $Ax = Ax'$. Thus there is a well-defined linear map $V : |A|H \to H$ satisfying $V|A|x = Ax$. By (11.4) this map is isometric and therefore extends continuously to $\overline{|A|H}$ by Lemma 3.12. We extend V to all of H by having it send $\overline{|A|H}^\perp$ to zero, obtaining a partial isometry. We have

$$\ker V = \overline{|A|H}^\perp = \ker |A|^* = \ker |A| = \ker A,$$

where we used Lemma 11.10, the self-adjointness of $|A|$ and (11.4). It is clear from the definition that $V|A| = A$.

That V is uniquely determined by its properties is quite clear: It must send $|A|x$ to Ax, which determines it on $\overline{|A|H}$. And in view of $\ker A = \ker |A| = \ker |A|^* = \overline{|A|H}^\perp$, the requirement $\ker V = \ker A$ forces V to be zero on $\overline{|A|H}^\perp$.

(ii) It is trivial that an injective (bijective) partial isometry is an isometry (unitary).

(iii) Using (11.4) as in (i) there is a unique partial isometry W such that $WAx = |A|x \; \forall x \in H$ and $W \restriction (AH)^\perp = 0$. Now it is immediate that $WV \restriction \overline{|A|H} = \mathrm{id}$ and $VW \restriction \overline{AH} = \mathrm{id}$, while WV and VW vanish on $|A|H^\perp$ and $AH^\perp$ respectively. Thus $W = V^*$, and we are done. ■

11.45 Remark Exercise 11.16 generalizes without any difficulty to operators $V \in B(H, H')$ between different Hilbert spaces $H \neq H'$. Then, of course, we have $V^*V \in B(H), VV^* \in B(H')$. Also Proposition 11.44 generalizes to $A \in B(H, H')$. Then $|A| = (A^*A)^{1/2} \in B(H)$ and $V \in B(H, H')$. □

11.46 Exercise Let $\{\lambda_k\}_{k\in\mathbb{N}} \subseteq \mathbb{C}$ be a bounded sequence. Let $H = \ell^2(\mathbb{N}, \mathbb{C})$ and $A_\lambda \in B(H)$ defined by $A_\lambda e_k = \lambda_k e_{k+1}$ (where $e_k(n) = \delta_{k,n}$).

(i) Compute $(A_\lambda)^*, |A_\lambda|$ and the partial isometry V in the polar decomposition of A_λ.

(ii) Give a necessary and sufficient condition on $\{\lambda_k\}$ for A_λ to be normal.

11.47 Exercise Prove:

(i) If $A \in B(H)$ is bounded below, i.e. $\|Ax\| \geq C\|x\|$ with $C > 0$, then also $|A|$ is bounded below with the same constant.

(ii) If $A \in B(H)$ is bounded below then $|A|$ is invertible.

11.9 ⋆ The trace of positive operators

If A is an $n \times n$ matrix, its trace $\mathrm{Tr}(A)$ is defined as $\sum_{i=1}^n A_{ii}$. One easily shows $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, so that $\mathrm{Tr}(BAB^{-1}) = \mathrm{Tr}(A)$ for all invertible B. If V is a finite-dimensional vector space and $A \in \mathrm{End}\, V$, we pick a basis E of V and define $\mathrm{Tr}(A)$ in terms of the matrix representation of A w.r.t. E, which then is independent of E by the conjugation invariance.

If V is an infinite-dimensional Banach space and $A \in B(V)$, it is a rather non-trivial problem to define $\mathrm{Tr}(A)$ (unless A has finite rank, cf. Exercise 10.11). But for Hilbert spaces, one has the following:

11.48 Definition *If H is a Hilbert space, E an ONB for H and $A \in B(H)$ is positive, we define the trace*

$$\mathrm{Tr}_E(A) = \sum_{e\in E} \langle Ae, e\rangle \in [0, +\infty].$$

11.49 Remark 1. If H is finite-dimensional, $\mathrm{Tr}_E(A)$ coincides with our earlier definition and therefore does not depend on the choice of E.

2. The sum is uniquely defined since $\langle Ae, e\rangle \geq 0$ for all $e \in E$. (The trace of certain non-positive operators will be considered in Appendix B.11.), but establishing the E-independence now is less straightforward. We begin with two special cases. □

11.50 Exercise Let H be a Hilbert space over $\mathbb{F}$ and $P \in B(H)$ an orthogonal projection. Prove that $\mathrm{Tr}_E(P)$ equals the rank of P, i.e. the Hilbert space dimension of the closed subspace $PH \subseteq H$ as an $\mathbb{F}$-Hilbert space, for each ONB E. Hint: Begin with rank-one projections.

11.51 Lemma *Let H be a Hilbert space and $A \in B(H)$. Then*

(i) $\mathrm{Tr}_E(A^*A) = \mathrm{Tr}_E(AA^*)$ for each ONB E of H.

(ii) The quantity in (i) is independent of the ONB E.

Proof. (i) Using Parseval's identity ($\|x\|^2 = \sum_{e' \in E} |\langle x, e' \rangle|^2$), we have

$$\begin{aligned}
\mathrm{Tr}_E(A^*A) &= \sum_{e\in E} \langle A^*Ae, e\rangle = \sum_{e\in E} \langle Ae, Ae\rangle = \sum_{e\in E} \|Ae\|^2 = \sum_{e\in E}\sum_{e'\in E} |\langle Ae, e'\rangle|^2 \\
&= \sum_{e'\in E}\sum_{e\in E} |\langle Ae, e'\rangle|^2 = \sum_{e'\in E}\sum_{e\in E} |\langle e, A^*e'\rangle|^2 = \sum_{e'\in E} \|A^*e'\|^2 \\
&= \sum_{e'\in E} \langle A^*e', A^*e'\rangle = \sum_{e'\in E} \langle AA^*e', e'\rangle = \mathrm{Tr}_E(AA^*),
\end{aligned}$$

where the exchange of summations is justified since all summands are non-negative.

(ii) If E is an ONB and $U \in B(H)$ is unitary, we have

$$\begin{aligned}
\mathrm{Tr}_E(UAA^*U^*) &= \mathrm{Tr}_E(UA(UA)^*) = \mathrm{Tr}_E((UA)^*UA) \\
&= \mathrm{Tr}_E(A^*U^*UA) = \mathrm{Tr}_E(A^*A) = \mathrm{Tr}_E(AA^*),
\end{aligned} \tag{11.5}$$

where the second and fifth equalities come from (i) as applied to UA and A, respectively. If now E' is a second ONB then E and E' have the same cardinality, so that there is a unitary U such that U^* maps E onto E'. Together with (11.5) this gives $\mathrm{Tr}_E(AA^*) = \mathrm{Tr}_E(UAA^*U^*) = \mathrm{Tr}_{E'}(AA^*)$, proving the claim. ∎

11.52 Corollary *If $A \in B(H)$ is positive then $\mathrm{Tr}_E(A)$ is independent of the ONB E.*

Proof. Positivity of A implies that there is a positive $B \in B(H)$ such that $A = B^2 = B^*B$. Now the claim is immediate by Lemma 11.51(ii). ∎

(Attempts to prove the corollary without using square roots tend to have gaps.) The above considerations will be put to use in Section 12.4 (and B.11).

12 Compact operators

12.1 Compact Banach space operators

We have met compact topological spaces many times in this course. A subset Y of a topological space (X, τ) is compact if it is compact when equipped with the induced (=subspace) topology $\tau_{|Y}$. Recall that a metric space X, thus also a subset of a normed space, is compact if and only if it is sequentially compact, i.e. every sequence $\{x_n\}$ in X has a convergent subsequence. And $Y \subseteq X$ is called precompact (or relatively compact) if its closure $\overline{Y}$ is compact. If X is complete metric, precompactness of $Y \subseteq X$ is equivalent to total boundedness.

A subset Y of a normed space $(V, \|\cdot\|)$ is called bounded if there is an M such that $\|y\| \leq M \ \forall y \in Y$. A compact subset of a normed space is closed and bounded, but the converse, while true for finite-dimensional spaces by the Heine-Borel theorem, is false in infinite-dimensional spaces. This is particularly easy to see for an infinite-dimensional Hilbert space: Any ONB $B \subseteq H$ clearly is bounded. For any $e, e' \in B, e \neq e'$ we have $\|e - e'\| = \langle e - e', e - e'\rangle^{1/2} = \sqrt{2}$. Thus $B \subseteq H$ is closed and discrete. Since it is infinite, it is not compact.

Related to this: If H is a Hilbert space and $W \subsetneq H$ a proper closed subspace, pick $x \in W^\perp$ with $\|x\| = 1$. Then $\|x - w\| = \sqrt{1 + \|w\|^2}$ for each $w \in W$, thus $\mathrm{dist}(x, W) = \inf_{w\in W} \|w - x\| = 1$. This can be generalized:

12.1 Exercise If V is a reflexive Banach space and $W \subsetneq V$ a proper closed subspace, there exists $x \in V$ satisfying $\|x\| = 1$ and $\mathrm{dist}(x, W) = 1$. Hint: Use Exercise 9.24.

Without reflexivity, we have the following weaker, but elementary result:

12.2 LEMMA (F. RIESZ) *Let V be a normed space and $W \subsetneqq V$ a closed and proper linear subspace. Then for each $\theta \in (0,1)$ there is an $x_\theta \in V$ such that $\|x_\theta\| = 1$ and $\mathrm{dist}(x_\theta, W) \geq \theta$.*

Proof. Since W is proper and closed, there are $x_0 \in V \backslash W$ and $\varepsilon > 0$ such that $B(x_0, \varepsilon) \subseteq V \backslash W$. Thus $\lambda := \mathrm{dist}(x_0, W) = \inf_{w \in W} d(x_0, w) \geq \varepsilon > 0$. In view of $\theta \in (0,1)$, we have $\frac{\lambda}{\theta} > \lambda$, so that we can find $w_0 \in W$ with $\lambda \leq \|w_0 - x_0\| < \frac{\lambda}{\theta}$. Putting

$$x_\theta = \frac{w_0 - x_0}{\|w_0 - x_0\|},$$

we have $\|x_\theta\| = 1$. If $w \in W$ then

$$\|w - x_\theta\| = \left\| w - \frac{w_0 - x_0}{\|w_0 - x_0\|} \right\| = \frac{\| \|w_0 - x_0\| w - w_0 + x_0 \|}{\|w_0 - x_0\|} \geq \frac{\mathrm{dist}(x_0, W)}{\|w_0 - x_0\|} > \frac{\lambda}{\frac{\lambda}{\theta}} = \theta,$$

where the $\geq$ is due to $\|w_0 - x_0\| w - w_0 \in W$ and the $>$ comes from $\mathrm{dist}(x_0, W) = \lambda$ and $\|w_0 - x_0\| < \frac{\lambda}{\theta}$. Since this holds for all $w \in W$, we have $\mathrm{dist}(x_\theta, W) \geq \theta$. ■

12.3 PROPOSITION *If $(V, \|\cdot\|)$ is an infinite-dimensional normed space then:*

(i) Each closed ball $\overline{B}(x,r) = \{y \in V \mid \|x - y\| \leq r\}$ (with $r > 0$) is non-compact.

(ii) Every subset $Y \subseteq V$ with non-empty interior Y^0 is non-compact.

Proof. (i) Choose $x_1 \in V$ with $\|x_1\| = 1$. Then $\mathbb{C}x_1$ is a closed proper subspace, thus there exists $x_2 \in V$ with $\|x_2\| = 1$ and $\|x_1 - x_2\| \geq \frac{1}{2}$. Since V is infinite-dimensional, $V_2 = \mathrm{span}\{x_1, x_2\}$ is a closed proper subspace, thus there exists $x_3 \in V$ with $\mathrm{dist}(x_3, V_2) \geq \frac{1}{2}$, thus in particular $\|x_3 - x_i\| \geq \frac{1}{2}$ for $i = 1, 2$. Continuing in this way we can construct a sequence $\{x_i\} \subseteq V$ with $\|x_i\| = 1$ and $\|x_i - x_j\| \geq \frac{1}{2}\ \forall i \neq j$. The sequence $\{x_i\}$ clearly cannot have a convergent subsequence, thus the closed unit ball $\overline{B}(0,1)$ is non-compact. Since $x \mapsto \lambda x + x_0$ with $\lambda > 0$ is a homeomorphism, all closed balls are non-compact.

(ii) If $Y \subseteq V$ and $Y^0 \neq \emptyset$ then Y contains some open ball $B(x,r)$, thus also $\overline{B}(x, r/2)$, which is non-compact. Thus neither Y nor $\overline{Y}$ are compact. ■

In view of the above, it is interesting to look at linear operators that send sets $S \subseteq V$ to sets AS with 'better compactness properties'. There are several such notions:

12.4 EXERCISE Let V, W be normed spaces and $A : V \to W$ a linear map. Prove that the following conditions are equivalent and imply boundedness of A:

(i) The image $AV_{\leq 1} \subseteq W$ of the closed unit ball $V_{\leq 1}$ is precompact.

(ii) Whenever $S \subseteq V$ is bounded, AS is precompact ($\Leftrightarrow$ totally bounded if V is complete).

(iii) Given any bounded sequence $\{x_n\} \subseteq V$, the sequence $\{Ax_n\}$ has a convergent subsequence.

12.5 DEFINITION *Operators $A \in B(V,W)$ satisfying the above equivalent conditions are called compact. The set of compact operators $V \to W$ is denoted $K(V,W)$, and we put $K(V) = K(V,V)$.*

12.6 Remark 0. That compactness implies boundedness is good to know, but rarely important since it is a priori clear in most situations.

1. Some authors write $B_0(V)$ rather than $K(V)$, motivated by Exercise 12.13(iii) below.

2. In the older literature one can find 'completely continuous' as synonym for compact, but this should be avoided since complete continuity now is defined differently and in general is not equivalent to compactness.

3. If $A \in B(V,W)$ is compact and $V' \subseteq V$ is closed then the restriction $A_{|V'} \in B(V',W)$ is compact. If $A \in B(V)$ is compact, $W \subseteq V$ is closed and $AW \subseteq W$ then $A_{|W} \in B(W)$ is compact.

4. The Heine-Borel theorem implies that every linear operator on a finite-dimensional normed space (automatically bounded by Exercise 3.7) is compact. For infinite-dimensional spaces this is false since every closed ball is bounded but non-compact by Proposition 12.3. In particular the unit operator $\mathbf{1}_V$ is compact if and only if V is finite-dimensional.

5. Compactness can also be defined for non-linear maps between Banach spaces. But then the three versions above are no more equivalent and continuity is no more automatic. See Section B.15. □

Before we develop further theory, we should prove that (non-zero) compact operators on infinite-dimensional spaces exist. The following may be known from Exercise 10.10:

12.7 Definition *Let V,W be normed spaces and $A \in B(V,W)$. Then A has finite rank if its image $AV \subseteq W$ is finite-dimensional. The set of finite rank operators $V \to W$ is denoted $F(V,W)$. Again, $F(V) = F(V,V)$.*

12.8 Remark 1. For example, if $\varphi \in V^*$, $y \in W$ then $A \in B(V,W)$ defined by $A : x \mapsto \varphi(x)y$ has finite rank.

2. Note that boundedness of $A : V \to W$ does not follow from finite-dimensionality of AV, as we know from $W = \mathbb{F}$.

3. It is straightforward to check that $F(V,W) \subseteq B(V,W)$ is a linear subspace. And if $A \in B(V,W), B \in B(W,Z)$ and at least one of A,B has finite rank then $BA \in F(V,Z)$. In particular, $F(V) \subseteq B(V)$ is a two-sided ideal.

4. In Exercise 10.10 we proved $F(V,W) = \{A \in B(V,W) \mid A \text{ is weak-norm continuous}\}$. In Section 12.2 we will relate weak versions of weak-norm continuity to compactness.

5. If V is infinite-dimensional then for each $n \in \mathbb{N}$ we can find (many) finite-dimensional subspaces W of dimension n. Since W is complemented, we have $V \simeq W \oplus Z$. This gives an embedding $B(W) \hookrightarrow F(V)$. Thus $F(V)$ has closed subalgebras isomorphic to $B(W)$ for each finite-dimensional W.

6. Finite rank operators can be used to prove that an algebraic isomorphism $K(V) \cong K(W)$ or $B(V) \cong B(W)$ implies an isomorphism $V \simeq W$ of Banach spaces. For a more general statement and proof see Section B.7. □

12.9 Lemma *Finite rank operators are compact. Thus $F(V,W) \subseteq K(V,W)$ for all Banach space V,W.*

Proof. Let $A \in F(V,W)$. If $S \subseteq V$ is bounded then $AS \subseteq AV$ is bounded by boundedness of V. Since $AV \subseteq W$ is finite-dimensional, it is closed and has the Heine-Borel property so that $\overline{AS} \subseteq \overline{AV} = AV$ is compact. Thus A is compact. ■

12.10 Lemma *$K(V,W) \subseteq B(V,W)$ is a vector space. If $A \in B(V,W), B \in B(W,Z)$ and at least one of A, B is compact then $BA \in K(V,Z)$. In particular, $K(V) \subseteq B(V)$ is a two-sided ideal.*

Proof. Let $\{x_n\}$ be a bounded sequence in V. Since A, B are compact, we can find a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that Ax_{n_k} and Bx_{n_k} converge as $k \to \infty$. Then also $(cA + dB)x_{n_k}$ converges, thus $cA + dB$ is compact. Thus $K(V,W) \subseteq B(V,W)$ is a linear subspace.

Alternative argument: Let $A, B \in K(V)$ and $S \subseteq V$ bounded. Then $\overline{AS}$ and $\overline{BS}$ are compact, and so are $\overline{cAS}$, $\overline{dBS}$ if $c, d \in \mathbb{F}$. Thus also $\overline{cAS} + \overline{dBS}$ is compact (by joint continuity of the map $+ : V \times V \to V$), thus also $\overline{(cA + dB)S} \subseteq \overline{\overline{cAS} + \overline{dBS}} = \overline{cAS} + \overline{dBS}$.

Now let $A \in B(V,W)$, $B \in K(W,Z)$ and $S \subseteq V$ bounded. Then $\overline{BS}$ and $A\overline{BS} = \overline{ABS}$ are compact by compactness of B and continuity of A, respectively. And boundedness of A implies boundedness of AS, so that $\overline{BAS}$ has compact closure by compactness of B. Thus AB and BA are compact. ■

For the proof of the next result, we need the notion of total boundedness in metric spaces, see Appendix A.6.4. In particular we will use Exercise A.43(iii).

12.11 Proposition *$K(V,W) \subseteq B(V,W)$ is $\|\cdot\|$-closed for all Banach spaces V, W.*

Proof. Since $B(V,W)$ is a metric space, it suffices to prove that the limit A of every norm-convergent sequence $\{A_n\}$ in $K(V,W)$ is in $K(V,W)$. Let thus $\{A_n\} \subseteq K(V,W)$ and $A \in B(V,W)$ such that $\|A_n - A\| \to 0$. We want to prove that $AV_{\leq 1} \subseteq W$ is precompact. Since $(B(V,W), \|\cdot\|)$ is complete, by Exercise A.43(iii) this is equivalent to $AV_{\leq 1}$ being totally bounded. To show this, let $\varepsilon > 0$. Then there is an n such that $\|A_n - A\| < \varepsilon/3$. Since A_n is compact, $A_n V_{\leq 1}$ is precompact, thus totally bounded, so that there are $x_1, \ldots, x_n \in V_{\leq 1}$ such that $\bigcup_{i=1}^n B(A_n x_i, \varepsilon/3) \supseteq A_n V_{\leq 1}$. Equivalently, for each $x \in V_{\leq 1}$ (thus $\|x\| \leq 1$) there is $i \in \{1, \ldots, n\}$ such that $\|A_n x - A_n x_i\| < \varepsilon/3$. Thus

$$\|Ax - Ax_i\| \leq \|Ax - A_n x\| + \|A_n x - A_n x_i\| + \|A_n x_i - Ax_i\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

where we used $\|A - A_n\| < \varepsilon/3$ and $x, x_i \in V_{\leq 1}$. Thus $AV_{\leq 1} \subseteq \bigcup_{i=1}^n B(Ax_i, \varepsilon)$, proving that $AV_{\leq 1}$ is totally bounded, thus precompact. Thus A is compact. ■

12.12 Corollary *For all Banach spaces V, W, we have $\overline{F(V,W)} \subseteq K(V,W)$.*

12.13 Exercise Let $V = \ell^p(S,\mathbb{F})$, where S is an infinite set and $1 \leq p < \infty$. If $g \in \ell^\infty(S,\mathbb{F})$ and $f \in \ell^p(S,\mathbb{F})$ then $M_g(f) = gf$ (pointwise product) is in $\ell^p(S)$. This defines a linear map $\ell^\infty(S,\mathbb{F}) \to B(V), g \mapsto M_g$. Prove:

(i) $g \mapsto M_g$ is an algebra homomorphism.

(ii) $\|M_g\| = \|g\|_\infty$.

(iii) $M_g \in K(V)$ if and only if $g \in c_0(S,\mathbb{F})$.

12.14 Remark 1. We now have two classes of compact operators: The (rather commutative) one of multiplication by c_0-functions, and the operators that are norm-limits of finite rank operators. Actually, the first class is contained in the second. Why?

2. In view of Exercise 12.13, the closed subspace $K(V) \subseteq B(V)$ is a non-abelian analogue of $c_0 \subseteq \ell^\infty$. If H is a separable Hilbert space, one can use the fact that c_0 is not complemented in ℓ^∞ to prove that $K(H)$ is not complemented in $B(H)$. See [31].

3. If either V or W is finite-dimensional, we have $F(V,W) = K(V,W) = B(V,W)$. While $K(V) \neq B(V)$ whenever V is infinite-dimensional since $\mathbf{1} \in B(V)\backslash K(V)$, very deep recent (2011) work [5], see also [6, 179], has produced infinite-dimensional Banach spaces V for which the difference between $B(V)$ and $K(V)$ is minimal in the sense of $B(V) = \mathbb{C}\mathbf{1} + K(V)$[67]! It is a much older result that there are V, W *both* infinite-dimensional with $B(V,W) = K(V,W)$! (Of course then $V \not\cong W$.) For large classes of examples see Theorems 12.24 and B.30.

4. It is quite natural to ask whether $\overline{F(V,W)} = K(V,W)$. We will later prove $\overline{F(H)} = K(H)$ for each Hilbert space. A Banach space V is said to have the approximation property if $\overline{F(W,V)} = K(W,V)$ holds for every Banach space W. (It is known[68] that V has the approximation property if and only if for every compact set $X \subset V$ and $\varepsilon > 0$ there is a $T \in F(V)$ such that $\|x - Tx\| < \varepsilon \ \forall x \in X$, cf. e.g. [98, vol. 1]. Whether this is already implied by $\overline{F(V)} = K(V)$ seems to be still open.) Whether all Banach spaces have the approximation property was an open problem until Enflo[69] in 1973 [47] constructed a counterexample. His construction was quite complicated and his spaces were not very 'natural' (in the sense of having a simple definition and/or having been encountered previously). A simpler example, but still tricky and not natural, can be found in [34]. Somewhat later, very natural examples were found: The Banach space $B(H)$ does not have the approximation property whenever H is an infinite-dimensional Hilbert space, cf. [162]. (Note that this is about compact operators on $B(H)$, not compact operators in $B(H)$!) Most of this is well beyond the level of this course, but you should be able to understand [34]. □

None of the above examples of compact operators seems very relevant for applications, even within mathematics. Indeed the most useful compact operators perhaps are integral operators. We will briefly look at a class of them in Exercise 12.41. But there are very simple examples:

12.15 Definition *Let $V = C([0,C],\mathbb{F})$ for some $C > 0$, equipped with the norm $\|f\| = \sup_{x\in[0,C]} |f(x)|$. As we know, $(V, \|\cdot\|)$ is a Banach space. Define a linear operator, the Volterra[70] operator, by*

$$A : V \to V, \quad (Af)(x) = \int_0^x f(t)dt.$$

We have $\|Af\| = \sup_x |\int_0^x f(t)dt| \leq \int_0^C |f(t)|dt \leq C\|f\|$, thus $\|A\| \leq C < \infty$.

12.16 Proposition *The Volterra operator $A : V \to V$ is compact.*

The proof of this result makes essential use of the Arzelà-Ascoli Theorem[71] which characterizes the (pre)compact subsets of the metric space $(C(X,\mathbb{F}), \|\cdot\|_\infty)$ for compact X.

[67] Fittingly, this is not only the most recent, but also by far the most sophisticated result mentioned in these notes. See [179] for a brief introduction, an excellent introduction to what happened in Banach space theory since the 1970s.

[68] This was proven by Alexander Grothendieck (1928-2014), German-born mathematician (later French) who first made fundamental contributions to functional analysis and then revolutionized algebraic geometry.

[69] Per H. Enflo (1944-). Swedish mathematician, working mostly in functional analysis. He also made seminal contributions to the 'invariant subspace problem' by constructing an infinite-dimensional separable Banach space V and an $A \in B(V)$ such that the only closed subspaces $W \subseteq V$ with $AW \subseteq W$ are $W = 0$ and $W = V$.

[70] Vito Volterra (1860-1940). Italian mathematician and one of the early pioneers of functional analysis.

[71] You should have seen this theorem in Analysis 2 or Topology. See e.g. Appendix A.6.4 or [57, Vol. 2, Theorem 15.5.1]. It has many applications in classical analysis, for example Peano's existence theorem on differential equations.

Proof. We will prove that $\mathcal{F} = A\overline{B} \subseteq V$ is precompact by showing that it satisfies the hypotheses of Theorem A.45. If $x \in [0, C]$ and $f \in C([0, C], \mathbb{F})$ with $\|f\| \leq 1$ then

$$|(Af)(x)| = \left| \int_0^x f(t)dt \right| \leq C\|f\| \leq C < \infty,$$

showing that $\mathcal{F}$ is pointwise bounded. For each $f \in V$ with $\|f\| \leq 1$ we have

$$|(Af)(x) - (Af)(y)| = \left| \int_y^x f(t)dt \right| \leq |x - y|.$$

Since this is uniform in $f \in \overline{B}$ it shows that $\mathcal{F}$ is equicontinuous. ■

The above proof is very easily adapted to give compactness of $A : V \to V$ with $V = C([a, b], \mathbb{F})$ given by $(Af)(x) = \int_a^b K(x, y)f(y)dy$ for any $K \in C([a, b] \times [a, b], \mathbb{F})$. For another class of compact integral operators see Example 12.41.

12.17 Theorem (Schauder 1930) *Let V, W be Banach spaces and $A \in B(V, W)$. Then $A^t \in B(W^*, V^*)$ is compact if and only if A is compact.*

Proof. For simplicity we assume $V = W$. The general case requires only notational changes. As in the proof of Proposition 12.11, we use the equivalence of precompactness and total boundedness from Exercise A.43(iii). If $A \in B(V)$ is compact then $AV_{\leq 1} \subseteq V$ is precompact, thus totally bounded. Thus for every $\varepsilon > 0$ there are $x_1, \ldots, x_n \in V_{\leq 1}$ such that $\bigcup_{i=1}^n B(Ax_i, \varepsilon) \supseteq AV_{\leq 1}$. Equivalently, for every $x \in V_{\leq 1}$ there is an i such that $\|Ax - Ax_i\| < \varepsilon/3$. Now define a bounded linear map $B : V^* \to \mathbb{F}^n$ (where $\mathbb{F}^n$ has the norm $\|\cdot\|_\infty$) by $B : \varphi \mapsto (\varphi(Ax_1), \ldots, \varphi(Ax_n))$. Since $\mathbb{F}^n$ is finite-dimensional, B has finite rank, thus is compact, so that $BV^*_{\leq 1} \subseteq V^*$ is totally bounded. Thus there are $\psi_1, \ldots, \psi_m \in V^*_{\leq 1}$ such that for every $\psi \in V^*_{\leq 1}$ we have $\|B\psi - B\psi_j\| < \varepsilon/3$ for some j. This gives

$$\max_{1 \leq i \leq n} |(A^t\psi)(x_i) - (A^t\psi_j)(x_i)| = \max_{1 \leq i \leq n} |\psi(Ax_i) - \psi_j(Ax_i)| = \|B\psi - B\psi_j\| < \varepsilon/3. \tag{12.1}$$

If now $x \in V_{\leq 1}$ then $\|Ax - Ax_i\| < \varepsilon/3$ for some i, thus $|(A^t\psi)(x) - (A^t\psi)(x_i)| = |\psi(Ax - Ax_i)| < \varepsilon/3$ for every $\psi \in V^*_{\leq 1}$, and (12.1) gives $|(A^t\psi)(x_i) - (A^t\psi_j)(x_i)| < \varepsilon/3$. Thus

$$\begin{aligned}
&|(A^t\psi)(x) - (A^t\psi_j)(x)| \leq \\
&\quad |(A^t\psi)(x) - (A^t\psi)(x_i)| + |(A^t\psi)(x_i) - (A^t\psi_j)(x_i)| + |(A^t\psi_j)(x_i) - (A^t\psi_j)(x)| \\
&\quad < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}$$

Since this holds for all $x \in V_{\leq 1}$, we have $\|A^t\psi - A^t\psi_j\| \leq \varepsilon$, and since $\psi \in V^*_{\leq 1}$ was arbitrary we have total boundedness of $A^tV^*_{\leq 1}$, thus compactness of A^t.

Now assume that $A^t \in B(V^*)$ is compact. Then by the above, $A^{tt} \in B(V^{**})$ is compact. Since $V \subseteq V^{**}$ is a closed subspace that is A^{tt}-invariant with $A^{tt}{}_{|V} = A$ by Lemma 9.33, compactness of A follows from Remark 12.6.3. ■

12.18 Remark The above self-contained proof, taken from [110], has much in common with the proof of the Arzelà-Ascoli theorem, and indeed the latter can be used to prove Schauder's theorem, as has become the standard proof, cf. e.g. [94]. An alternative proof is sketched in Remark 12.20 below. Yet another proof uses the circle of ideas in Section 10 (weak topologies and Alaoglu's theorem) as in [30, Theorem VI.3.4]. □

The following is an instructive – and useful – characterization of compact operators which should go some way towards making the notion more intuitive:

12.19 Theorem (H. E. Lacey 1963) *Let V, W be Banach spaces and $A \in B(V, W)$. Then the following are equivalent:*

(i) A is compact.

(ii) For every $\varepsilon > 0$ there exists a closed subspace $Z \subseteq V$ of finite codimension such that $\|A \upharpoonright Z\| \leq \varepsilon$.

Proof. (We follow the nice exposition in [121].) (i)⇒(ii) Let $\varepsilon > 0$. Since $AV_{\leq 1}$ is precompact, there are $w_1, \ldots, w_n \in W$ such that $AV_{\leq 1} \subseteq \bigcup_{i=1}^n B(w_i, \varepsilon)$. By Proposition 9.9 we can find $\varphi_1, \ldots, \varphi_n \in W^*$ such that $\|\varphi_i\| = 1$ and $\varphi_i(w_i) = \|w_i\|$ for all i. Now

$$Z = \{v \in V \mid \varphi_1(Av) = \cdots = \varphi_n(Av) = 0\} = \bigcap_{i=1}^n \ker A^t \varphi_i$$

is a closed subspace of V (by continuity of A and the φ_i). As the intersection of finitely many spaces of codimension ≤ 1, Z has finite codimension $\leq n$. If now $z \in Z_{\leq 1} \subseteq V_{\leq 1}$, there exists an i such that $\|Az - w_i\| < \varepsilon$. And with $\varphi_i(Az) = 0$ for all $z \in Z$, $i \in \{1, \ldots, n\}$ we have

$$\|w_i\| = |\varphi_i(w_i)| \leq |\varphi_i(w_i - Az)| + |\varphi_i(Az)| < \varepsilon,$$

thus

$$\|Az\| \leq \|Az - w_i\| + \|w_i\| \leq \varepsilon + \varepsilon = 2\varepsilon.$$

Since this holds for all $z \in Z_{\leq 1}$, we have $\|A \upharpoonright Z\| \leq 2\varepsilon$.

(ii)⇒(i) If $\varepsilon > 0$, by assumption there is a closed subspace $Z \subseteq V$ of finite codimension such that $\|A \upharpoonright Z\| \leq \varepsilon$. By Proposition 6.11 (i), Z is complemented, thus by Exercise 7.15(i), there is an idempotent P_Z with $P_Z V = Z$. But this is easy to prove directly: We can find unit vectors $e_1, \ldots, e_n$, where $n = \mathrm{codim} Z$, such that $Y = \mathrm{span}_{\mathbb{F}}\{e_1, \ldots, e_n\}$ is an algebraic complement for Z. Then every $v \in V$ is of the form $v = z + y$ with unique $z \in Z, y \in Y$. Now define $P_Z : z + y \mapsto z$ and $P_Y : z + y \mapsto y$. Since Y is finite-dimensional, P_Y has finite rank, thus is compact. Thus with $\delta = \min(1, \varepsilon/\|A\|)$, there are $v_1, \ldots, v_m \in V_{\leq 1}$ such that

$$P_Y V_{\leq 1} \subseteq \bigcup_{i=1}^m B(P_Y v_i, \delta).$$

If now $v \in V_{\leq 1}$ there is an i such that $\|P_Y v - P_Y v_i\| < \delta$. With $P_M(v - v_i) = (v - v_i) - P_n(v - v_i)$ we have

$$\|P_Z v - P_Z v_i\| \leq \|v - v_i\| + \|P_Y v - P_Y v_i\| \leq 2 + \delta \leq 3,$$

so that, again using $P_Z + P_Y = \mathbf{1}$, and $\|AP_Z\| = \|A \upharpoonright Z\| \leq \varepsilon$ we have

$$\|Av - Av_i\| \leq \|AP_Z v - AP_Z v_i\| + \|AP_Y v - AP_Y v_i\| \leq 3\|AP_Z\| + \delta\|A\| \leq 4\varepsilon.$$

Thus $AV_{\leq 1} \subseteq \bigcup_{i=1}^m B(Av_i, 4\varepsilon)$. Since $\varepsilon > 0$ was arbitrary, $AV_{\leq 1}$ is precompact and therefore A compact. ■

12.20 Remark Very similarly to the proof of the preceding theorem one can prove [143] that compactness of $A \in B(V, W)$ is equivalent to the following statement, dual to the above (ii):

(iii) For every $\varepsilon > 0$ there exists a finite-dimensional subspace $Y \subseteq W$ such that $\|QA\| < \varepsilon$, where Q is the quotient map $W \to W/Y$.

Now one can give an alternative proof [143] of Schauder's theorem: Assume A is compact, thus (iii) holds. Thus for every $\varepsilon > 0$, there is a finite-dimensional subspace $Y \subseteq W$ such that $\|QA\| < \varepsilon$, where $Q : W \to W/Y$. Put $Z = Y^\perp \subseteq W^*$, which has finite codimension. Now $A^* \restriction Z = (QA)^t$, so that $\|A^* \restriction Z\| = \|(QA)^t\| = \|QA\| < \varepsilon$. Thus T^* satisfies (ii) in Theorem 12.19 and therefore is compact. The converse implication is proven as in Theorem 12.17. (But ultimately this is the same proof, the finite-(co)dimensional subspaces Y playing a role similar to that of the auxiliary $\mathbb{F}^n$ in the first proof.) □

12.21 Exercise Let V, W be Banach spaces and $A \in K(V, W)$. Prove:

(i) $\overline{AV} \subseteq W$ is separable.

(ii) If $AV \subseteq W$ is closed then AV is finite-dimensional.

The following result has applications to Sobolev spaces and PDEs:

12.22 Exercise Let X, Y, Z be Banach spaces, $T \in B(X, Y)$ compact and $S \in B(Y, Z)$ injective. Prove (possibly by contradiction) that for every $\varepsilon > 0$ there is a $C_\varepsilon \geq 0$ such that

$$\|Tx\|_Y \leq \varepsilon \|x\|_X + C_\varepsilon \|STx\|_Z \quad \forall x \in X.$$

(Note: If $\varepsilon \geq \|T\|$ then we can put $C_\varepsilon = 0$. The point is that $0 < \varepsilon < \|T\|$ is allowed!)

12.23 Exercise If V, W are Banach spaces, $A \in B(V, W)$ is called strictly singular if there is no infinite-dimensional subspace $Z \subseteq V$ such that $A \restriction Z$ is bounded below. Prove:

(i) $A \in B(V, W)$ is strictly singular if and only if there is no infinite-dimensional closed subspace $Z \subseteq V$ such that $A : Z \to AZ$ is an isomorphism.

(ii) Every compact operator is strictly singular.

In Remark B.33 we will see that there are non-compact strictly singular operators!

12.2 ⋆ Compactness vs. weak forms of weak-norm continuity

Recognizing compact operators can be difficult. So far we have seen that compactness of some integral operators can be proven using the Arzelà-Ascoli theorem (see also Section 12.4), and we know that operators in the norm-closure of the finite rank operators are compact. But our only general criterion so far is Theorem 12.19. The fact that the finite rank operators are precisely the weak-norm continuous operators, cf. Exercise 10.10, suggests that weak topologies can be brought to bear on proving compactness. We begin with an easy but remarkable instance:

12.24 Theorem *If V is a reflexive Banach space then all bounded linear operators $V \to \ell^1(\mathbb{N}, \mathbb{F})$ and $c_0(\mathbb{N}, \mathbb{F}) \to V$ are compact.*

Proof. Let $A \in B(V, \ell^1(\mathbb{N}, \mathbb{F}))$. By Exercise 10.14(iii), $V_{\leq 1}$ with the weak topology is sequentially compact. Thus every norm-bounded sequence $\{x_n\} \subset V$ has a weakly convergent subsequence $\{x_{n_i}\}$. Since A is weak-weak continuous by Exercise 10.9, $\{Ax_{n_i}\} \subset \ell^1(\mathbb{N}, \mathbb{F})$ converges weakly. Now Schur's Theorem 10.4 implies that $\{Ax_{n_i}\} \subset \ell^1(\mathbb{N}, \mathbb{F})$ converges in norm. Thus $\{Ax_n\}$ has a norm-convergent subsequence, so that A is compact.

Let $A \in B(c_0, V)$, thus $A^t \in B(V^*, c_0^*)$. Since V^* is reflexive and $c_0^* \cong \ell^1$, the above gives that A^t is compact, thus also A by Schauder's Theorem 12.17. ■

Remarkably, Theorem 12.24 requires no information about A other than its boundedness. But the proof relies on the rather exceptional Schur property of ℓ^1. In view of the proof it is clear that $B(V, W) = K(V, W)$ whenever V is reflexive and W has the Schur property. (See [90] for further generalizations.)

To obtain results of wider applicability, it will be useful to characterize compactness of an operator in terms of properties slightly weaker than weak-norm continuity. We will consider three such properties of an operator, beginning with:

12.25 PROPOSITION *Let V, W be Banach spaces and $A \in B(V, W)$. Then A is compact if and only if the restriction $A : V_{\leq 1} \to W$ is weak-norm continuous.*

Proof. Assume A is compact. Let $\{x_\iota\}$ be a net in the closed unit ball of V weakly converges to zero. Let $\varepsilon > 0$. By Theorem 12.19 there exists a closed subspace $Z \subseteq V$ of finite codimension such that $\|A \restriction Z\| \leq \varepsilon$. Since Z is complemented, there is a finite rank idempotent $P \in B(V)$ such that $(\mathbf{1} - P)V = Z$. Since AP has finite rank, it is weak-norm continuous by Exercise 10.10, thus $\|APx_\iota\| \to 0$. Now with $\|A(\mathbf{1} - P)\| \leq \varepsilon$ we have

$$\|Ax_\iota\| \leq \|APx_\iota\| + \|A(\mathbf{1} - P)x_\iota\| \leq \|APx_\iota\| + 2\varepsilon.$$

Since $\|APx_\iota\| \to 0$ and $\varepsilon > 0$ was arbitrary, we have $\|Ax_\iota\| \to 0$. Thus A is weak-norm continuous.

Now assume that $A \restriction V_{\leq 1}$ is weak-norm continuous, and let $\varepsilon > 0$. By the weak-norm continuity, $V_{\leq 1} \cap A^{-1}(B^W(0, \varepsilon)) \subseteq V_{\leq 1}$ is a weakly open neighborhood of $0 \in V$. Thus there exist $\varphi_1, \ldots, \varphi_n \in V^*$ such that

$$\{x \in V_{\leq 1} \mid |\varphi_i(x)| < 1 \ \forall i\} \subseteq V_{\leq 1} \cap A^{-1}(B^W(0, \varepsilon)). \tag{12.2}$$

Now $Z = \bigcap_{i=1}^n \ker \varphi_i \subseteq V$ is a closed subspace of codimension $\leq n$, and for every $z \in Z_{\leq 1}$ we have $\varphi_i(z) = 0 \ \forall i$, and therefore $\|Az\| < \varepsilon$ by (12.2). Thus for all $v \in Z$ we have $\|Av\| \leq \varepsilon \|A\|$, so that we have $\|A \restriction Z\| \leq \varepsilon$. Now A is compact by Theorem 12.19. ■

12.26 REMARK 1. An alternative proof of the implication ⇒ that uses the compactness of A directly goes like this: Let $\{x_\iota\}_{\iota \in I}$ be a net in $V_{\leq 1}$ that converges weakly to zero. Since the norm-continuous A is weak-weak continuous, $Ax_\iota \stackrel{w}{\to} 0$. If $\|Ax_\iota\| \not\to 0$ then there exists an $\varepsilon > 0$ such that for every $\iota \in I$ there is a $\iota' \geq \iota$ such that $\|Ax_{\iota'}\| \geq \varepsilon$. Using this we can construct a subnet $\{x_\sigma\}_{\sigma \in \Sigma}$ of $\{x_\iota\}$ such that $\|Ax_\sigma\| \geq \varepsilon$ for all $\sigma \in \Sigma$. Since $\overline{AV_{\leq 1}}$ is compact, the net $\{Ax_\sigma\}$ has an accumulation point $w \in W$, which clearly satisfies $\|w\| \geq \varepsilon$. Since $w \neq 0$ also is a weak accumulation point of $\{Ax_\sigma\}$, we have a contradiction with $Ax_\alpha \stackrel{w}{\to} 0$. Thus $\|Ax_\iota\| \to 0$.

This argument is just the obvious adaptation the proof of Theorem 12.28(ii) to nets. But it is less elementary than the one above in that it uses accumulation points and subnets of nets.

2. By general topology, cf. e.g. [142, 108], continuity of $A : (V_{\leq 1}, \tau_w) \to (W, \|\cdot\|)$ is equivalent to the statement that the net $\{Ax_\iota\} \subset W$ is norm-convergent for every weakly convergent bounded net $\{x_\iota\}_{\iota \in I}$ in V. But either formulation is hard to verify, and we would prefer a criterion involving only sequences. □

12.27 DEFINITION *A linear map $A : V \to W$ of Banach spaces is sequentially weak-norm continuous if $\{Ax_n\} \subset W$ is norm-convergent for every weakly convergent sequence $\{x_n\}$ in V.*

12.28 THEOREM *Let V, W be Banach spaces and $A : V \to W$ linear. Then*

(i) If A is sequentially weak-norm continuous, it is bounded.

(ii) If A is compact then it is sequentially weak-norm continuous.

(iii) Every bounded linear map from ℓ^1 to an arbitrary Banach space is sequentially weak-norm continuous, yet $\mathbf{1} : \ell^1 \to \ell^1$ is non-compact.

(iv) Let V, W be Banach spaces, where V is reflexive or V^* is separable. Then every sequentially weak norm continuous $A : V \to W$ is compact.

Proof. (i) If A is unbounded, we can find $\{x_n\} \subset V$ with $\|x_n\| = 1$ such that $\|Ax_n\| \geq n^2$. Now $y_n = \frac{x_n}{n} \to 0$ in norm, thus weakly, while Ay_n does not converge in norm since $\|Ay_n\| \geq n$. This is a contradiction.

(ii) It suffices to prove that $\|Ax_n\| \to 0$ whenever $\{x_n\}$ converges weakly to zero. By Exercise 10.6, $\{x_n\}$ is bounded. Since A is weak-weak continuous, we have $Ax_n \stackrel{w}{\to} 0$. If $\|Ax_n\| \not\to 0$ then there exist $\varepsilon > 0$ and a subsequence $y_i = \{x_{n_i}\}$ such that $\|Ay_i\| \geq \varepsilon$ for all i. By compactness of A there is a second subsequence $z_i = y_{m_i}$ such that Az_i converges in norm. The limit is non-zero since $\|Az_i\| \geq \varepsilon$ for all i. Thus Az_i converges to a non-zero limit also weakly, contradicting $Ax_n \stackrel{w}{\to} 0$. Thus $\|Ax_n\| \to 0$.

(iii) If $\{x_n\} \subset \ell^1$ is a weakly convergent sequence, it is norm-convergent by Schur's Theorem 10.4, thus $\{Ax_n\}$ is norm-convergent for any $A : \ell^1 \to V$. Thus A is sequentially weak-norm continuous. The second follows from infinite-dimensionality of ℓ^1.

(iv) If V is reflexive then every bounded sequence $\{x_n\} \subset V$ has a weakly convergent subsequence by Exercise 10.14(iii). Since A maps the latter to a norm convergent sequence in W, $\{Ax_n\}$ has a norm-convergent subsequence, thus A is compact.

If V^* is separable then every bounded sequence in V has a weakly Cauchy subsequence by Exercise 10.14(i). Now compactness of A follows as in the preceding case if we use the following lemma. ■

12.29 LEMMA *A sequentially weak-norm continuous operator between Banach spaces maps weakly Cauchy sequences to norm convergent ones.*

Proof. Let $\{i_n\}, \{j_n\}$ be increasing sequences in $\mathbb{N}$. Then for each $\varphi \in V^*$ the sequences $\{\varphi(x_{i_n})\}$ and $\{\varphi(x_{j_n})\}$ have the same limit $\lim_{n\to\infty} \varphi(x_n)$. Thus $\{x_{i_n} - x_{j_n}\}$ is a weakly null sequence, so that $\|A(x_{i_n} - x_{j_n})\| \to 0$ by the sequential weak-norm continuity of A. Now the following exercise gives that $\{Ax_n\}$ is Cauchy w.r.t. $\|\cdot\|$ and therefore norm-convergent. ■

12.30 EXERCISE Let (X, d) be a metric space and $\{x_n\}$ a sequence in X such that $d(x_{i_n}, x_{j_n}) \to 0$ for all strictly increasing sequences $\{i_n\}, \{j_n\}$ of natural numbers. Prove that $\{x_n\}$ is Cauchy.

12.31 REMARK 1. Our first application of Theorem 12.28(iv) is that no infinite-dimensional Banach space with the Schur property is reflexive or has separable dual. This follows since the Schur property of V means that id_V is sequentially weak-norm continuous, while id_V is non-compact for infinite-dimensional V.

2. In Section B.3.6, Theorem 12.28(iv) will be used to prove some interesting results about the automatic compactness of a large class of bounded operators.

3. Accepting Footnote 62, Theorem 12.28(iv) immediately generalizes to all V that have no subspace isomorphic to ℓ^1. □

We quickly look at a third property of weak-norm type:

12.32 Definition *If V, W are Banach spaces then a linear $A : V \to W$ is completely continuous if weak compactness of $X \subset V$ implies norm-compactness of $AX \subset W$.*

12.33 Proposition *Let V, W be Banach spaces and $A : V \to W$ linear. Then*

(i) If $A \upharpoonright V_{\leq 1}$ is weak-norm continuous then it is completely continuous and sequentially weak-norm continuous.

(ii) If A is completely continuous, it is bounded.

(iii) A is completely continuous if and only if it is sequentially weak-norm continuous.

Proof. (i) The hypothesis clearly implies weak-norm continuity on all bounded sets. If now $X \subset V$ is weakly compact, it is norm-bounded by Exercise 10.7, so that $AX \subset W$ is norm-compact by the hypothesis on A. Thus A is completely continuous. Since every weakly convergent sequence is bounded by Exercise 10.6, the sequential weak-norm continuity of A is equally obvious.

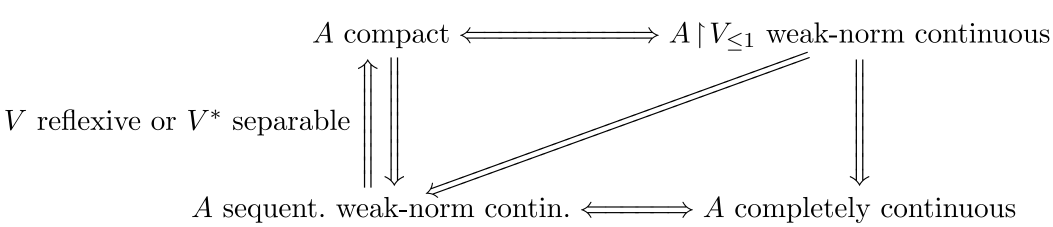
(ii) If A is unbounded, there is a sequence $\{x_n\}$ in V with $\|x_n\| = 1$ and $\|Ax_n\| \geq n^2$ for all n. Then $y_n = x_n/n$ converges to zero in norm, thus weakly. Thus $Y = \{y_n \mid n \in \mathbb{N}\} \cup \{0\}$ is weakly compact[72], thus AY is norm compact, thus bounded, contradicting $\|Ay_n\| \geq n$.

(iii) $\Rightarrow$: Assume $A : V \to W$ to be completely continuous and let $\{x_n\} \subset V$ converge weakly to $x \in V$. Since A is bounded (=norm-continuous) by (i), it is weak-weak continuous, thus $\{Ax_n\}$ converges weakly. If the set $\{Ax_n \mid n \in \mathbb{N}\}$ is finite, its linear span in W is finite-dimensional, thus weak and norm topology coincide on it, so that the weak convergence of $\{Ax_n\}$ implies norm convergence.

Thus from now on we may assume $\{Ax_n \mid n \in \mathbb{N}\}$ to be infinite. By weak convergence, $X = \{x_n \mid n \in \mathbb{N}\} \cup \{x\} \subseteq V$ is weakly compact. Thus $AX = \{Ax_n \mid n \in \mathbb{N}\} \cup \{Ax\} \subset W$ is norm-compact by complete continuity of A. Thus AX has at least one limit point $y \in AX$ (i.e. $\#(B(y,\varepsilon) \cap AX) = \infty$ for every $\varepsilon > 0$) since otherwise it would be discrete and infinite, contradicting compactness. For every limit point z of AX there is a subsequence of $\{Ax_n\}$ converging to z in norm, thus also weakly. But the weakly convergent sequence $\{Ax_n\}$ has precisely one weak limit point, so that it has exactly one limit point in the norm topology, thus converges in norm.

$\Leftarrow$: Assume $A : V \to W$ is sequentially weak-norm continuous and $X \subset V$ is weakly compact, thus bounded by Exercise 10.7. Let $\{y_n\}$ be a sequence in AX. Clearly there exists a sequence $\{x_n\}$ in X with $Ax_n = y_n \ \forall n$. Since X is weakly compact, it is weakly sequentially compact by the easier half of the Eberlein-Šmulyan theorem, cf. Remark 10.16, thus there is a weakly convergent subsequence $\{x_{n_i}\}$. By weak-norm continuity of A, $\{y_{n_i} = Ax_{n_i}\}$ is norm-convergent, proving norm-compactness of AX. Thus A is completely continuous. ■

Summarizing, for every linear $A : V \to W$ we have the following implications:



[72] In *any* topological space, convergence $x_n \to x$ implies compactness of $\{x_n \mid n \in \mathbb{N}\} \cup \{x\}$.

We close by mentioning another property involving weak topologies that an operator may have or not: $A \in B(V,W)$ is called weakly compact if it maps bounded sets of V to subsets of W that are precompact in the weak topology. See e.g. [102, Section 3.5].

12.3 Compact Hilbert space operators

12.34 Proposition *If H is a Hilbert space and $A \in B(H)$ is compact then A^* and $|A|$ are compact.*

Proof. By the polar decomposition, there is a partial isometry V such that $A = V|A|$ and $|A| = V^*A$. Since $K(H) \subseteq B(H)$ is an ideal (Lemma 12.10) the second identity and compactness of A give compactness of $|A|$. Since the adjoint of the first identity is $A^* = |A|V^*$, a second use of the ideal property of $K(H)$ gives compactness of A^*. ■

12.35 Remark 1. Compactness of A^* also follows from $A^* = \gamma^{-1} \circ A^t \circ \gamma$, compactness of A^t (Theorem 12.17) and the fact that $\gamma : H \stackrel{\cong}{\to} H^*,\ y \mapsto \langle \cdot, y \rangle$ is a homeomorphism.

2. In Corollary 14.15 we will prove, without any use of A^*, that for compact A and $\varepsilon > 0$ there is $F \in F(H)$ with $\|A - F\| < \varepsilon$, thus $\|A^* - F^*\| < \varepsilon$. By the following exercise, F^* is finite rank. Since $\varepsilon > 0$ was arbitrary, Corollary 12.12 again gives $A^* \in K(H)$. □

12.36 Exercise Let H be a Hilbert space.

(i) If $x, y \in H$, define $x \otimes y \in B(H)$ by $(x \otimes y)(z) = x\langle z, y\rangle$. Prove that $(x \otimes y)^* = (y \otimes x)$.

(ii) Prove that $A \in F(H) \Rightarrow A^* \in F(H)$ and $\dim A^*H = \dim AH$.
Hint: You can prove $A = \sum_{k=1}^K x_k \otimes y_k$ for $K = \dim AH$ and $x_1, y_1, \ldots, x_K, y_K \in H$ and use (i), but there also is an elegant direct proof.

12.37 Theorem *Let H be a Hilbert space and $A \in B(H)$. Then the following are equivalent:*

(α) For every $\varepsilon > 0$ there is an orthogonal projection P such that $\|PAP\| < \varepsilon$ and $PH \subseteq H$ has finite codimension.

(β) A is compact.

(γ) For every orthonormal sequence $\{e_n\}_{n\in\mathbb{N}} \subseteq H$ we have $Ae_n \to 0$.

(δ) For every orthonormal sequence $\{e_n\}_{n\in\mathbb{N}} \subseteq H$ we have $\langle Ae_n, e_n\rangle \to 0$.

The implication $(\gamma) \Rightarrow (\delta)$ follows trivially from Cauchy-Schwarz. As to the rest:

12.38 Exercise Let H be a Hilbert space and $A \in B(H)$.

(i) Prove $(\alpha) \Rightarrow (\beta)$ in Theorem 12.37.

(ii) Prove $(\beta) \Rightarrow (\gamma)$. Hint: Begin by proving weak convergence.

(iii) Prove $(\delta) \Rightarrow (\alpha)$ for self-adjoint A. Hint: Assuming that (α) is false, use Proposition 11.34(i) to construct an orthonormal sequence $\{e_n\}$ such that $\langle Ae_n, e_n\rangle \not\to 0$.

(iv) Deduce the general statement (δ) from the self-adjoint case.

12.4 ⋆ Hilbert-Schmidt operators: $L^2(H)$

If E is an ONB for a Hilbert space and $A \in B(H)$, we have seen in Lemma 11.51 that $\mathrm{Tr}_E(A^*A)$ does not depend on E, so that we can write $\mathrm{Tr}(A^*A)$, and that $\mathrm{Tr}(A^*A) = \mathrm{Tr}(AA^*)$.

12.39 Definition *For each $A \in B(H)$ we define*

$$\|A\|_2 = (\mathrm{Tr}(A^*A))^{1/2} \in [0,\infty], \qquad L^2(H) = \{A \in B(H) \mid \|A\|_2 < \infty\}.$$

The elements of $L^2(H)$ are called Hilbert-Schmidt[73] *operators.*

12.40 Proposition *Let H be a Hilbert space. Then*

(i) $\|A\| \le \|A\|_2 = \|A^*\|_2$ for all $A \in B(H)$. Thus $L^2(H)$ is self-adjoint.

(ii) For every $A, B \in L^2(H)$,

$$\langle A, B\rangle_{HS} = \mathrm{Tr}(B^*A) = \sum_e \langle B^*Ae, e\rangle$$

is absolutely convergent and independent of the ONB E. Now $\langle\cdot,\cdot\rangle_{HS}$ is an inner product on $L^2(H)$ such that $\langle A, A\rangle_{HS} = \|A\|_2^2$. And $(L^2(H), \langle\cdot,\cdot\rangle_{HS})$ is complete, thus a Hilbert space.

(iii) For all $A, B \in B(H)$ we have $\|AB\|_2 \le \|A\|\|B\|_2$ and $\|AB\|_2 \le \|A\|_2\|B\|$. Thus $L^2(H) \subseteq B(H)$ is a two-sided ideal.

(iv) We have $F(H) \subseteq L^2(H) \subseteq K(H)$ and $\overline{F(H)}^{\|\cdot\|_2} = L^2(H)$.

Proof. (i) If $x \in H$ is a unit vector, pick an ONB E containing x. Then $\|Ax\|^2 = \langle A^*Ax, x\rangle \le \mathrm{Tr}_E(A^*A) = \|A\|_2^2$. Thus $\|Ax\| \le \|A\|_2$ whenever $\|x\| = 1$, proving the inequality. And $\|A^*\|_2^2 = \mathrm{Tr}(AA^*) = \mathrm{Tr}(A^*A) = \|A\|_2^2$.

(ii) If E is any ONB (whose choice does not matter) for H, we have (as before)

$$\mathrm{Tr}(A^*A) = \sum_{e\in E}\langle A^*Ae, e\rangle = \sum_{e\in E}\langle Ae, Ae\rangle = \sum_{e\in E}\|Ae\|^2 = \sum_{e\in E}\sum_{e'\in E}|\langle Ae, e'\rangle|^2.$$

Thus $L^2(H)$ is the set of $A \in B(H)$ for which the matrix elements $\langle Ae, e'\rangle$ (w.r.t. the ONB E) are absolutely square summable. We therefore have a map

$$\alpha : L^2(H) \to \ell^2(E\times E),\ A \mapsto \{\langle Ae, e'\rangle\}_{(e,e')\in E^2}$$

that clearly is injective. (Recall that $\ell^2(S) = L^2(S,\mu)$, where μ is the counting measure.) To show surjectivity of α, let $f = \{f_{ee'}\} \in \ell^2(E\times E)$. Define a linear operator $A : H \to H$ by $A : e \mapsto \sum_{e'} f_{ee'}e'$. For each e, the r.h.s. is in H by square summability of f. If $x \in H$ then

$$\begin{aligned}\|Ax\|^2 &= \sup_{E'}\left\|\sum_e\langle x, e\rangle\sum_{e'\in E'} f_{ee'}e'\right\|^2 = \sup_{E'}\left\|\sum_{e'\in E'}\left(\sum_e\langle x, e\rangle f_{ee'}\right)e'\right\|^2\\ &= \sup_{E'}\sum_{e'}\left|\sum_e\langle x, e\rangle f_{ee'}\right|^2 \le \|x\|^2\sum_{e,e'}|f_{ee'}|^2,\end{aligned}$$

[73] Erhard Schmidt (1876-1959). Baltic German mathematician, contributions to functional analysis like Gram-Schmidt orthogonalization.

where the supremum is over the finite subsets $E' \subseteq E$, we used $|\langle x, e\rangle| \leq \|x\|$ and the change of summation order is allowed due to the finiteness of E'. This computation shows that $\|A\| \leq (\sum_{e,e'} |f_{ee'}|^2)^{1/2} < \infty$. Thus $A \in B(H)$ and $\alpha(A) = f$, so that α is surjective. Thus $\alpha : L^2(H) \to \ell^2(E \times E)$ is a linear bijection. Now $\ell^2(E \times E)$ is a Hilbert space (in particular complete) with inner product $(f, g) = \sum_{e,e'} f_{ee'} \overline{g_{ee'}}$, and pulling this inner product back to $L^2(H)$ along α we have

$$\begin{aligned} \langle A, B\rangle_{HS} &= \sum_{(e,e')\in E^2} \langle Ae, e'\rangle \overline{\langle Be, e'\rangle} = \sum_{(e,e')\in E^2} \langle Ae, e'\rangle\langle e', Be\rangle \\ &= \sum_e \langle Ae, Be\rangle = \sum_e \langle B^*Ae, e\rangle = \mathrm{Tr}(B^*A), \end{aligned}$$

where all sums converge absolutely. Lemma 11.51 gives that $\langle A, A\rangle_{HS} = \mathrm{Tr}(A^*A)$ is independent of the chosen ONB, and for general (A, B) this follows by the polarization identity.

From the above it is clear that $(L^2(H), \langle\cdot,\cdot\rangle_{HS})$ is isomorphic to the Hilbert space $(\ell^2(E \times E), \langle\cdot,\cdot\rangle)$, thus a Hilbert space. And the norm associated to $\langle\cdot,\cdot\rangle_{HS}$ is nothing other than $\|\cdot\|_2$.

(iii) For any ONB E we have

$$\|AB\|_2^2 = \mathrm{Tr}(B^*A^*AB) = \sum_{e\in E} \|ABe\|^2 \leq \|A\|^2 \sum_{e\in E} \|Be\|^2 = \|A\|^2 \mathrm{Tr}(B^*B) = \|A\|^2\|B\|_2^2,$$

proving $\|AB\|_2 \leq \|A\|\|B\|_2$. And $\|AB\|_2 = \|(AB)^*\|_2 = \|B^*A^*\|_2 \leq \|B^*\|\|A^*\|_2 = \|A\|_2\|B\|$, where we used the fact just proven and $\|A^*\|_2 = \|A\|_2$. The conclusion is obvious.

(iv) The inclusion $F(H) \subseteq L^2(H)$ is very easy and is left as an exercise. If $A \in L^2(H)$ and F is a finite subset of the ONB E, define $p_F = \sum_{e\in F} \langle\cdot, e\rangle e$ and $A_F = Ap_F$. Then $A_F \in F(H)$ and A

$$\|A - A_F\|_2^2 = \|A(1 - p_F)\|_2^2 = \sum_{\langle e,e'\rangle\in(E\backslash F)\times E} |\langle Ae, e'\rangle|^2.$$

This implies $\|A - A_F\|_2 \to 0$ as $F \nearrow E$, so that $L^2(H) = \overline{F(H)}^{\|\cdot\|_2}$.

Finally, by (i) we have $\|A - A_F\| \leq \|A - A_F\|_2 \to 0$, thus $A \in \overline{F(H)}^{\|\cdot\|} \subseteq K(H)$, where we used Corollary 12.12. This proves $L^2(H) \subseteq K(H)$. ■

12.41 Example (L^2-Integral operators) Let $(X, \mathcal{A}, \mu)$ be a measure space, and put $H = L^2(X, \mathcal{A}, \mu)$. Let $k : X \times X \to \mathbb{C}$ be measurable (w.r.t. the product σ-algebra $\mathcal{A} \times \mathcal{A}$) and assume $\int\int |k(x, y)|^2 \, d\mu(x)d\mu(y) < \infty$. (Thus $k \in L^2(X \times X, \mathcal{A} \times \mathcal{A}, \mu \times \mu)$.) Then

$$(Kf)(x) = \int_X k(x, y)f(y) \, d\mu(y)$$

defines a linear operator $K : H \to H$ whose Hilbert-Schmidt norm $\|K\|_2$ coincides with the norm $\|k\|_{L^2}$ of $k \in L^2(X \times X)$. Thus K is Hilbert-Schmidt, and in particular compact.

12.42 Exercise Prove the equality $\|K\|_2 = \|k\|_{L^2}$ of norms claimed in the above example.

If V, W are vector spaces over any field $\mathbb{K}$ then there is a canonical linear map $W \otimes_{\mathbb{K}} V^* \to \mathrm{Hom}_{\mathbb{K}}(V, W)$ sending $w \otimes_{\mathbb{K}} \varphi$ to the linear map $v \mapsto w\varphi(v)$. (Here V^* is the algebraic dual space and $\otimes_{\mathbb{K}}$ is the algebraic tensor product.) If V or W is finite-dimensional, this map is a bijection, but otherwise it is not. For Hilbert spaces, one has a statement that works irrespective of the dimensions:

12.43 Exercise Let H be a Hilbert space.

(i) Define $\overline{H}$ to be the vector space H with scalar action $(c, x) \mapsto \overline{c}x$ and inner product $\langle x, y\rangle_{\overline{H}} = \overline{\langle x, y\rangle}$. Prove that $\overline{H}$ is a Hilbert space.

(ii) Define a map $\alpha : H \otimes_{\text{alg}} \overline{H} \to F(H)$ by associating to $\sum_{i=1}^n x_i \otimes y_i$ the operator $z \mapsto \sum_{i=1}^n x_i \langle z, y_i\rangle$. Prove that α is linear and extends to an isometric bijection $\alpha : H \otimes \overline{H} \to L^2(H)$.

12.44 Remark See Appendix B.11 for the definition of $L^p(H)$ for all $p \in [1, \infty)$ and a more complete discussion of $L^1(H)$. □

12.45 Exercise Given a set S and $f \in \ell^\infty(S, \mathbb{F})$, define $H = \ell^2(S, \mathbb{C})$ and the multiplication operator $M_g : H \to H,\ f \mapsto gf$ (known from Exercise 12.13, where we saw $M_g \in K(H) \Leftrightarrow g \in c_0(S)$). Prove $|M_g| = M_{|g|}$ and $\|M_g\|_p = \|g\|_p$ for all $p \in [1, \infty)$. (Thus $M_g \in L^p(H) \Leftrightarrow g \in \ell^p(S, \mathbb{C})$.)

Part II: Spectral theory of operators and algebras

13 Spectrum of bounded operators and of Banach algebra elements

13.1 Spectra of bounded operators I: Definitions, first results

We now return to the discussion of invertible operators, specializing from $B(V, W)$ to $B(V)$, i.e. bounded linear maps from a normed space to itself. We recall some definitions from linear algebra without assuming finite-dimensionality of V:

13.1 Definition *Let V be a Banach space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(V)$.*

- *If $x \in V \backslash \{0\}$ and $\lambda \in \mathbb{F}$ such that $Ax = \lambda x$ then x is called an eigenvector of A with corresponding eigenvalue λ.*
- *Thus $\lambda \in \mathbb{F}$ is an eigenvalue of A if and only if $A - \lambda\mathbf{1}$ is not injective, i.e. $\ker(A - \lambda\mathbf{1}) \neq \{0\}$.*
- *The eigenspace of $\lambda \in \mathbb{F}$ is $\ker(A - \lambda\mathbf{1})$. (But 0 is not considered as eigenvector!)*
- *If $x \in V\backslash\{0\}$ and $(A - \lambda\mathbf{1})^n x = 0$ for some $n \in \mathbb{N}$ then x is called a generalized eigenvector for the eigenvalue λ. (If $(A - \lambda\mathbf{1})^n x = 0$ but $y = (A - \lambda\mathbf{1})^{n-1} x \neq 0$ then $(A - \lambda\mathbf{1})y = 0$, so that λ indeed is an eigenvalue even though x is not an eigenvector.)*
- *The geometric multiplicity of λ is $\dim \ker(A - \lambda\mathbf{1})$.*
- *The algebraic multiplicity of λ is $\dim \bigcup_{n \in \mathbb{N}} \ker(A - \lambda\mathbf{1})^n$.*
- *If V is finite-dimensional then the algebraic multiplicity of λ coincides with the multiplicity of λ as a zero of the characteristic polynomial $P(z) = \det(A - z\mathbf{1})$.*

If V is a finite-dimensional vector space and $A \in \operatorname{End} V$, it is well known that one has: A is injective $\Leftrightarrow$ A is surjective $\Leftrightarrow$ A is invertible. Thus in finite dimensions failure of $A - \lambda\mathbf{1}_V$ to be invertible for some $\lambda \in \mathbb{F}$ is equivalent to $\ker(A - \lambda\mathbf{1}_V) \neq \{0\}$, thus λ to being an eigenvalue.

It is extremely important that the equivalence of injectivity, surjectivity and invertibility fails in infinite dimensions, as the following standard examples illustrate:

13.2 DEFINITION *Let $V = \ell^p(\mathbb{N}, \mathbb{F})$, where $1 \le p \le \infty$. Define $L, R \in B(V)$ by*

$$(Lf)(n) = f(n+1), \qquad (Rf)(n) = \begin{cases} 0 & \text{if } n = 1 \\ f(n-1) & \text{if } n \ge 2 \end{cases}$$

Equivalently: $R\delta_n = \delta_{n+1}$, $L\delta_1 = 0$, $L\delta_n = \delta_{n-1}$ if $n \ge 2$, which is why we call L, R the left and right, respectively, shift operators on V.

It is immediate that R is injective, but not surjective (since $(Rf)(1) = 0\ \forall f \in V$) while L is surjective, but not injective (since Lf does not depend on $f(1)$). One easily checks $LR = \mathrm{id}_V$, while $RL \ne \mathrm{id}_V$ since $RL = P_2$ (notation from Exercise 8.12).

13.3 EXERCISE Consider the shift operators L, R on $\ell^p(\mathbb{N}, \mathbb{C})$, $p \in [1, \infty]$.

(i) Prove that R is an isometry.

(ii) In the Hilbert space case $p = 2$, prove $L^* = R$, $R^* = L$. Conclude that L a coisometry.

Passing to an infinite-dimensional Banach space V, there can be $A \in B(V)$ and $\lambda \in \mathbb{F}$ for which $A - \lambda\mathbf{1}$ is injective, but not surjective, thus not invertible. Such λ are not eigenvalues, but they turn out to be equally important as the former. This motivates the following:

13.4 DEFINITION *Let V be a Banach space over $\mathbb{F}$ and $A \in B(V)$. Then*

- *The spectrum[74] $\sigma(A)$ is the set of $\lambda \in \mathbb{F}$ for which $A - \lambda\mathbf{1}_V$ is not invertible.*
- *The point spectrum $\sigma_{\mathrm{p}}(A)$ consists of those $\lambda \in \mathbb{F}$ for which $A - \lambda\mathbf{1}_V$ is not injective. Equivalently, $\sigma_{\mathrm{p}}(A)$ consists of the eigenvalues of A.*
- *The continuous spectrum $\sigma_{\mathrm{c}}(A)$ consists of those $\lambda \in \mathbb{F}$ for which $A - \lambda\mathbf{1}_V$ is injective, but not surjective, while it has dense image, i.e. $\overline{(A - \lambda\mathbf{1}_V)V} = V$.*
- *The residual spectrum $\sigma_{\mathrm{r}}(A)$ consists of those $\lambda \in \mathbb{F}$ for which $A - \lambda\mathbf{1}_V$ is injective and $\overline{(A - \lambda\mathbf{1}_V)V} \ne V$.*

We have some immediate observations and comments:

- It is obvious by construction that the sets $\sigma_{\mathrm{p}}(A), \sigma_{\mathrm{c}}(A), \sigma_{\mathrm{r}}(A)$ are mutually disjoint and have $\sigma(A)$ as their union.
- Clearly $0 \in \sigma(A)$ is equivalent to non-invertibility of A and $0 \in \sigma_{\mathrm{p}}(A)$ to $\ker A \ne \{0\}$.
- If V is finite-dimensional then we know from linear algebra that injectivity and surjectivity of any $A \in B(V)$ are equivalent. Thus for all operators on a finite-dimensional space we have $\sigma_{\mathrm{c}}(A) = \sigma_{\mathrm{r}}(A) = \emptyset$, thus $\sigma(A) = \sigma_{\mathrm{p}}(A)$.
- If V is infinite-dimensional, the situation is much more complicated, thus more interesting. For example, the right shift R on $H = \ell^2(\mathbb{N})$ is injective, thus $0 \notin \sigma_{\mathrm{p}}(R)$, but with $\overline{RH} \ne H$ we find $0 \in \sigma_{\mathrm{r}}(R) \subseteq \sigma(R)$.
- If $\lambda \in \sigma_{\mathrm{p}}(A)$ then there is non-zero $x \in V$ with $Ax = \lambda x$. Then $A^n x = \lambda^n x\ \forall n \in \mathbb{N}$. With the definition of $\|A\|$ it follows that $|\lambda| \le \inf_{n\in\mathbb{N}} \|A^n\|^{1/n}$. (This can be smaller than $\|A\|$, e.g. if A is nilpotent, i.e. $A^n = 0$ for some $n \in \mathbb{N}$.) We'll prove a similar result for $\sigma(A)$.

[74]The choice of this term by Hilbert was nothing less than a stroke of genius since it turned out to fit exactly its later use in quantum theory.

- One reason for distinguishing continuous and residual spectrum is the (not quite perfect) duality between $\sigma_{\rm p}$ and $\sigma_{\rm r}$, cf. Exercise 13.9(ii)-(iii). Another is that $\sigma_{\rm r}$ often is empty, cf. e.g. Exercise 13.13.
- The continuous spectrum need not be 'continuous' in the sense of connected. But one can prove, cf. Exercise 13.68, that every isolated $\lambda \in \sigma(A)$ is an eigenvalue if either A is a normal operator on Hilbert space (Proposition 17.24(ii)) or some additional condition is satisfied, cf. Section B.10. This perhaps goes some way towards explaining the term 'continuous'.
- In Section B.10 we will define the discrete spectrum $\sigma_{\rm d}(A)$, a certain subset of the point spectrum $\sigma_{\rm p}(A)$, as well as several closely related essential spectra of A.
- Later we will prove that $\sigma(A)$ is always closed, while we will see in examples that $\sigma_{\rm p}, \sigma_{\rm c}, \sigma_{\rm r}$ need not be closed.

Proposition 7.41(iii) suggests to define another two interesting subsets of $\sigma(A)$:

13.5 Definition *Let V be a Banach space and $A \in B(V)$. Then define*

- *the approximate point spectrum* $\sigma_{\rm ap}(A) = \{\lambda \in \mathbb{F} \mid A - \lambda\mathbf{1} \text{ is not bounded below}\}$,
- *the compression spectrum* $\sigma_{\rm cp}(A) = \{\lambda \in \mathbb{F} \mid \overline{(A - \lambda\mathbf{1})V} \neq V\}$.

Again some immediate observations:

- $\sigma(A) = \sigma_{\rm ap}(A) \cup \sigma_{\rm cp}(A)$ (by Proposition 7.41(iii)), but $\sigma_{\rm ap}(A)$ and $\sigma_{\rm cp}(A)$ need not be disjoint, e.g. for $A = 0$.
- $\sigma_{\rm r}(A) = \sigma_{\rm cp}(A)\backslash\sigma_{\rm p}(A) \subseteq \sigma_{\rm cp}(A)$ is easily checked.
- $\sigma_{\rm p}(A) \subseteq \sigma_{\rm ap}(A)$ is obvious.
- $\sigma_{\rm c}(A) \subseteq \sigma_{\rm ap}(A)$, since $\lambda \in \sigma_{\rm c}(A)$ implies that $A - \lambda\mathbf{1}$ is not invertible, but has dense image, so that $\lambda \notin \sigma_{\rm cp}(A)$.
- $\sigma_{\rm ap}(A)$ is closed as consequence of Exercise 7.46. Thus $\overline{\sigma_{\rm p}(A) \cup \sigma_{\rm c}(A)} \subseteq \sigma_{\rm ap}(A)$.
- If $\sigma_{\rm r}(A) = \emptyset$ then $\sigma(A) = \sigma_{\rm app}(A)$.
- For more on $\sigma_{\rm ap}(A)$ see Exercises 13.13(iii), 13.65(ii).

13.6 Exercise Let V be a Banach space and $A, B \in B(V)$ with B invertible. Prove $\sigma(BAB^{-1}) = \sigma(A)$ and $\sigma_{\rm x}(BAB^{-1}) = \sigma_{\rm x}(A)$ for all ${\rm x} \in \{{\rm p}, {\rm c}, {\rm r}, {\rm app}, {\rm cp}\}$.

13.7 Exercise Compute $\sigma_{\rm p}(L)$ and $\sigma_{\rm p}(R)$ for the shift operators L, R on $\ell^p(\mathbb{N}, \mathbb{C})$ for all $p \in [1, \infty]$. (Of course, the p in $\sigma_{\rm p}$ has nothing to to with the p in ℓ^p.)

13.8 Exercise Let V, W be Banach spaces and $A \in B(V), B \in B(W)$. Define $C \in B(V \oplus W)$ by $C : (x, y) \mapsto (Ax, By)$. (Thus $C = A \oplus B$.)

(i) Prove $\sigma(C) = \sigma(A) \cup \sigma(B)$ and $\sigma_{\rm p}(C) = \sigma_{\rm p}(A) \cup \sigma_{\rm p}(B)$.

(ii) Compute $\sigma_{\rm c}(C)$ and $\sigma_{\rm r}(C)$. Warning: These are not simply unions as in (i)!

For operators on a Hilbert space, we can study how the spectra of A and A^* are related:

13.9 Exercise Let H be a Hilbert space and $A \in B(H)$. Use Lemma 11.10 to prove:

(i) $\sigma(A^*) = \sigma(A)^* := \{\overline{\lambda} \mid \lambda \in \sigma(A)\}$.[75]

(ii) If $\lambda \in \sigma_{\rm r}(A)$ then $\overline{\lambda} \in \sigma_{\rm p}(A^*)$.

(iii) If $\lambda \in \sigma_{\rm p}(A)$ then $\overline{\lambda} \in \sigma_{\rm p}(A^*) \cup \sigma_{\rm r}(A^*)$.

(iv) $\sigma_{\rm cp}(A) = \sigma_{\rm p}(A^*)^*$.

13.10 Remark Using Exercise 9.35 instead of Lemma 11.10, one has analogous results for the transpose $A^t \in B(V^*)$ of a Banach space operator $A \in B(V)$: $\sigma(A^t) = \sigma(A)$, $\sigma_{\rm r}(A) \subseteq \sigma_{\rm p}(A^t)$ and $\sigma_{\rm p}(A) \subseteq \sigma_{\rm p}(A^t) \cup \sigma_{\rm r}(A^t)$. (There is no complex conjugation since $A \mapsto A^t$ is linear.) □

Normal operators have very nice spectral properties, foreshadowing the spectral theorem:

13.11 Exercise Let H be a Hilbert space and $A \in B(H)$ normal. Prove:

(i) If $A^n x = 0$ for some $n \in \mathbb{N}$ then $Ax = 0$.

(ii) With $L_\lambda(A) = \bigcup_{n \in \mathbb{N}} \ker(A - \lambda \mathbf{1})^n$ we have $L_\lambda(A) = \ker(A - \lambda \mathbf{1})$, thus generalized eigenvectors are eigenvectors.

13.12 Exercise Let H be a Hilbert space and $A \in B(H)$ normal. Let x, x' be (non-zero) eigenvectors for the eigenvalues $\lambda, \lambda' \in \sigma_{\rm p}(A)$, respectively. Prove:

(i) $A^* x = \overline{\lambda} x$, thus x is eigenvector for A^* with eigenvalue $\overline{\lambda}$.

(ii) $\sigma_{\rm p}(A^*) = \sigma_{\rm p}(A)^*$.

(iii) If $\lambda \neq \lambda'$ then $x \perp x'$.[76]

(iv) Give an example of a non-normal $A \in B(H)$ for which (iii) fails.

13.13 Exercise (Spectra of normal operators) Let H be a Hilbert space and $A \in B(H)$ normal. Prove:

(i) $\sigma_{\rm r}(A) = \emptyset$. (No residual spectrum)

(ii) $\sigma_{\rm c}(A^*) = \sigma_{\rm c}(A)^*$.

(iii) $\sigma(A) = \sigma_{\rm ap}(A)$ (cf. Definition 13.5).

Keep in mind that self-adjoint operators are normal, so that the above results apply. But for non-normal operators we cannot expect such nice results to hold!

13.14 Exercise Let H be a Hilbert space and $A \in B(H)$. Use Exercise 13.13(iii) to prove:

(i) If A is self-adjoint then $\sigma(A) \subseteq \mathbb{R}$.

(ii) If A is unitary then $\sigma(A) \subseteq S^1$.

13.15 Exercise Let H be a Hilbert space and $A \in B(H)$. Recalling Definition 11.32, prove

(i) $\sigma_{\rm p}(A) \subseteq W(A)$.

(ii) $\sigma_{\rm r}(A) \subseteq W(A)$.

(iii) $\sigma_{\rm ap}(A) \subseteq \overline{W(A)}$.

(iv) Conclude that $\sigma(A) \subseteq \overline{W(A)}$.

---

[75] If $S \subseteq \mathbb{C}$ we write S^* for $\{\overline{s} \mid s \in S\}$ since $\overline{S}$ could be confused with the closure.

[76] You may have seen this before, but probably only for self-adjoint operators.

13.2 The spectrum in a unital Banach algebra

Since $B(E)$ is a unital Banach algebra for every Banach space E, all results proven in the rest of this section in particular apply to bounded operators on Banach spaces. Restricting these results to $B(E)$ would not simplify their proofs significantly. Whether $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ does not matter until Section 13.2.3.

13.2.1 The group of invertibles

13.16 Definition *If $\mathcal{A}$ is a unital algebra over $\mathbb{F}$ then* $\mathrm{Inv}\mathcal{A} = \{a \in \mathcal{A} \mid \exists b \in \mathcal{A} : ab = ba = \mathbf{1}\}$ *is the set of invertible elements of $\mathcal{A}$.*

It should be clear that $\mathrm{Inv}\mathcal{A} \subseteq \mathcal{A}$ is a group with the multiplication of $\mathcal{A}$ and unit $\mathbf{1}$.

13.17 Lemma *Let $\mathcal{A}$ be a unital normed algebra. Then* $\mathrm{Inv}\mathcal{A}$ *is a topological group (w.r.t. the norm topology).*

Proof. Since multiplication $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is jointly continuous (Remark 3.29), the same holds for its restriction to $\mathrm{Inv}\mathcal{A}$. It remains to show that the inverse map $\mathrm{Inv}\mathcal{A} \to \mathrm{Inv}\mathcal{A}, a \mapsto a^{-1}$ is continuous. To this purpose, let $a, a+h \in \mathrm{Inv}\mathcal{A}$ and define k by $(a+h)^{-1} = a^{-1} + k$. Then $\mathbf{1} = (a^{-1}+k)(a+h) = \mathbf{1} + a^{-1}h + ka + kh$, thus $a^{-1}h + ka + kh = 0$. Multiplying this on the right by a^{-1} we have $a^{-1}ha^{-1} + k + kha^{-1} = 0$, thus $k = -a^{-1}ha^{-1} - kha^{-1}$. Therefore $\|k\| \le \|a^{-1}\|^2\|h\| + \|k\|\|h\|\|a^{-1}\|$, which is equivalent to $\|k\|(1 - \|h\|\|a^{-1}\|) \le \|a^{-1}\|^2\|h\|$. Assuming $\|h\| < \|a^{-1}\|^{-1}$, the expression in brackets is positive, so that

$$\|k\| \le \frac{\|a^{-1}\|^2}{1 - \|h\|\|a^{-1}\|}\|h\|.$$

It follows that $\|h\| \to 0$ implies $\|k\| \to 0$, so that $a + h \to a$ in $\mathrm{Inv}\mathcal{A}$ entails $(a+h)^{-1} \to a^{-1}$. ■

So far, we know very little about $\mathrm{Inv}\mathcal{A}$. We might, in principle, have $\mathrm{Inv}\mathcal{A} = \mathbb{F}^*\mathbf{1}$, where $\mathbb{F}^* = \mathbb{F}\backslash\{0\}$. (This is indeed the case for the algebra $\mathcal{A} = \mathbb{F}[x]$ which is normed, but not Banach, with $\|P\| = \sup_{x\in[0,1]} |P(x)|$.) The next results provide invertible elements other than multiples of $\mathbf{1}$.

13.18 Definition *An element $a \in \mathcal{A}$ of an algebra is called nilpotent if $a^n = 0$ for some $n \in \mathbb{N}$.*

13.19 Lemma *Let $\mathcal{A}$ be a unital Banach algebra.*

(i) If $a \in \mathcal{A}$ is nilpotent then $\mathbf{1} - a \in \mathrm{Inv}\mathcal{A}$. (This is true in every unital algebra.)

(ii) If $a \in \mathcal{A}$, $\|a\| < 1$ then $\mathbf{1} - a \in \mathrm{Inv}\mathcal{A}$ and $(\mathbf{1} - a)^{-1} = \sum_{n=0}^\infty a^n$. [77]

(iii) $\mathrm{Inv}\mathcal{A} \subseteq \mathcal{A}$ is open. More precisely, if $a \in \mathrm{Inv}\mathcal{A}$ then $B(a, \|a^{-1}\|^{-1}) \subseteq \mathrm{Inv}\mathcal{A}$.

Proof. (i) If $a \in \mathcal{A}$ is nilpotent then the series $b = \sum_{n=0}^\infty a^n$ converges since it breaks off after finitely many terms. Now $(\mathbf{1} - a)b = b(\mathbf{1} - a) = \sum_{n=0}^\infty a^n - \sum_{n=1}^\infty a^n = \mathbf{1}$, thus $\mathbf{1} - a \in \mathrm{Inv}\mathcal{A}$.

(ii) If $\|a\| < 1$ then $\sum_{n=0}^\infty \|a^n\| \le \sum_{n=0}^\infty \|a\|^n < \infty$, so that the series $\sum_{n=0}^\infty a^n$ converges to some $b \in \mathcal{A}$ by completeness and Proposition 3.15. Now again $(\mathbf{1} - a)b = b(\mathbf{1} - a) = \sum_{n=0}^\infty a^n - \sum_{n=1}^\infty a^n = \mathbf{1}$, so that $\mathbf{1} - a$ is invertible with inverse b.

[77] In this context, the geometric series $\sum_{n=0}^\infty a^n$ is called the Neumann series, after the German mathematician Carl Gottfried Neumann (1832-1925).

(iii) If $a \in \mathrm{Inv}\mathcal{A}$ and $a' \in \mathcal{A}$ with $\|a - a'\| < \|a^{-1}\|^{-1}$ then $\|\mathbf{1} - a^{-1}a'\| = \|a^{-1}(a - a')\| \leq \|a^{-1}\|\|a - a'\| < 1$ so that $a^{-1}a' = \mathbf{1} - (\mathbf{1} - a^{-1}a') \in \mathrm{Inv}\mathcal{A}$ by (ii), thus $a' = a(a^{-1}a') \in \mathrm{Inv}\mathcal{A}$. This proves that $\mathrm{Inv}\mathcal{A}$ is open. ■

13.20 Exercise Let $\mathcal{A}$ be a unital algebra and $a, b \in \mathcal{A}$. Prove:

(i) If a and b are invertible then ab and ba are invertible.

(ii) If a and ab (or ba) are invertible then b is invertible.

(iii) Invertibility of ab need not imply invertibility of a or b.

(iv) If ab **and** ba are invertible then a and b are invertible. Express the inverses a^{-1}, b^{-1} in terms of $c = (ab)^{-1}, d = (ba)^{-1}$.

(v) If $a_1, \ldots, a_n \in \mathcal{A}$ commute mutually then $a_1 \cdots a_n \in \mathrm{Inv}\mathcal{A} \Leftrightarrow a_i \in \mathrm{Inv}\mathcal{A}\ \forall i$.

(vi) Invertibility of $\mathbf{1} - ab$ implies invertibility of $\mathbf{1} - ba$.
Hint: Assuming that $\mathcal{A}$ is Banach and $\|a\|\|b\| < 1$, find a formula for $(\mathbf{1} - ba)^{-1}$ in terms of $(\mathbf{1} - ab)^{-1}$. Now prove that the latter holds without the mentioned assumptions.

13.21 Exercise Let $\mathcal{A}$ be a unital Banach algebra. Prove:

(i) For all $b \in \mathrm{Inv}\mathcal{A}$, $a \in \mathcal{A}\backslash\mathrm{Inv}\mathcal{A}$ we have $\|b^{-1}\| \geq \|a - b\|^{-1}$.

(ii) For all $b \in \mathrm{Inv}\mathcal{A}$ we have $\|b^{-1}\| \geq (\mathrm{dist}(b, \mathcal{A}\backslash\mathrm{Inv}\mathcal{A}))^{-1}$.

13.2.2 The spectrum. Basic properties

13.22 Definition *Let $\mathcal{A}$ be a unital algebra over $\mathbb{F}$. The spectrum of $a \in \mathcal{A}$ is defined as*

$$\sigma(a) = \{\lambda \in \mathbb{F} \mid a - \lambda\mathbf{1} \notin \mathrm{Inv}\mathcal{A}\}.$$

The spectral radius of a is $r(a) = \sup\{|\lambda| \mid \lambda \in \sigma(a)\}$, where $r(a) = 0$ if $\sigma(a) = \emptyset$. (We will prove that $\sigma(a) \neq \emptyset$ for all $a \in \mathcal{A}$ if $\mathcal{A}$ is normed and $\mathbb{F} = \mathbb{C}$.) The map

$$R_a : \mathbb{F}\backslash\sigma(a) \to \mathcal{A},\ \lambda \mapsto (a - \lambda\mathbf{1})^{-1}$$

is called the resolvent (map). (Sometimes $\rho(a) = \mathbb{F}\backslash\sigma(a)$ is called the resolvent set.)

13.23 Remark 1. It is clear that for an element of the Banach algebra $B(E)$, where E is a Banach space, this definition is equivalent to Definition 13.4. But in the present abstract setting there is no analogue of the point, continuous, residual and compression spectra. (For a generalization of $\sigma_{\rm ap}$ to elements of abstract Banach algebras see Footnote 90.)

2. If $a \in \mathcal{A}$ and $b \in \mathrm{Inv}\mathcal{A}$ it is immediate that $\sigma(a) = \sigma(bab^{-1})$, thus $r(a) = r(bab^{-1})$.

3. Lemma 13.17 implies that the resolvent map $R_a : \mathbb{F}\backslash\sigma(a) \to \mathcal{A},\ \lambda \mapsto (a - \lambda\mathbf{1})^{-1}$ is continuous in every unital normed algebra.

4. Be warned that some authors define the resolvent $R_a(\lambda)$ as $(\lambda\mathbf{1} - a)^{-1}$. □

As to our standard examples of Banach algebras not of the form $B(V)$ with V Banach:

13.24 Exercise (i) Let X be a compact Hausdorff space. Recall that $(C(X, \mathbb{F}), \|\cdot\|_\infty)$ is a Banach algebra. For $f \in C(X, \mathbb{F})$, prove $\sigma(f) = f(X) \subseteq \mathbb{F}$.

(ii) As we saw in Section 4.6, $\ell^\infty(S, \mathbb{F})$ is a Banach algebra w.r.t. pointwise multiplication for every set S. If $f \in \ell^\infty(S, \mathbb{F})$, prove $\sigma(f) = \overline{f(S)}$.

We begin our study of the spectrum with two purely algebraic results:

13.25 LEMMA *Let $\mathcal{A}$ be a unital algebra. Then*

(i) If $a \in \mathcal{A}$ is nilpotent then $\sigma(a) = \{0\}$, thus $r(a) = 0$.

(ii) For all $a, b \in \mathcal{A}$ we have $\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$ and $r(ab) = r(ba)$.

Proof. (i) If $\lambda \neq 0$ then $\mathbf{1} - \frac{a}{\lambda}$ is invertible by Lemma 13.19(i). Thus also $\lambda\mathbf{1} - a$ is invertible, so that $\sigma(a) \subseteq \{0\}$. Since no nilpotent a is invertible (why?) we have $\sigma(a) = \{0\}$.

(ii) For all $\lambda \neq 0$ we have $\lambda \in \sigma(ab) \Leftrightarrow \lambda\mathbf{1} - ab \notin \mathrm{Inv}\mathcal{A} \Leftrightarrow \mathbf{1} - (\lambda^{-1}a)b \notin \mathrm{Inv}\mathcal{A} \Leftrightarrow \mathbf{1} - b(\lambda^{-1}a) \notin \mathrm{Inv}\mathcal{A} \Leftrightarrow \lambda\mathbf{1} - ba \notin \mathrm{Inv}\mathcal{A} \Leftrightarrow \lambda \in \sigma(ba)$, where the third equivalence comes from Exercise 13.20(vi). (For $\lambda = 0$ this argument does not work. As we know already, invertibility of ab and ba are independent.) The second statement is an obvious consequence. ■

13.26 DEFINITION *If $\mathcal{A}$ is a unital Banach algebra, $a \in \mathcal{A}$ is called quasi-nilpotent if $r(a) = 0$, equivalently $\sigma(a) \subseteq \{0\}$.*

As just proven, nilpotent ⇒ quasi-nilpotent. For examples of quasi-nilpotent elements that are not nilpotent, see Exercises 13.61, 13.63, 19.6.

13.27 PROPOSITION *If $\mathcal{A}$ is a unital Banach algebra and $a \in \mathcal{A}$ then*

(i) $\sigma(a)$ is closed.

(ii) $r(a) \leq \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \leq \|a\|$.

Proof. (i) If $a \in \mathcal{A}$ then $f_a : \mathbb{F} \to \mathcal{A}$, $\lambda \mapsto a - \lambda\mathbf{1}$ is continuous, thus $f_a^{-1}(\mathrm{Inv}\mathcal{A}) \subseteq \mathbb{F}$ is open by Lemma 13.19(iii). Now $\sigma(a) = \mathbb{F}\backslash f_a^{-1}(\mathrm{Inv}\mathcal{A})$ is closed.

(ii) If $\lambda \in \mathbb{F}, |\lambda| > \|a\|$ then $\|a/\lambda\| < 1$ so that $\mathbf{1} - a/\lambda \in \mathrm{Inv}\mathcal{A}$ by Lemma 13.19(ii). Thus $\lambda\mathbf{1} - a \in \mathrm{Inv}\mathcal{A}$, so that $\lambda \notin \sigma(a)$. This proves $r(a) \leq \|a\|$.

In each unital algebra we have the telescoping computation related to finite geometric sums

$$(z - \mathbf{1})(\mathbf{1} + z + z^2 + \cdots + z^{n-1}) = (z + \cdots + z^n) - (\mathbf{1} + \cdots + z^{n-1}) = z^n - \mathbf{1}. \qquad (13.1)$$

If $z - \mathbf{1}$ is not invertible, Exercise 13.20(v) implies that $z^n - \mathbf{1}$ is not invertible for any $n \in \mathbb{N}$. Applying this to $z = a/\lambda$, where $\lambda \in \sigma(a)\backslash\{0\}$, gives $\lambda^n \in \sigma(a^n)$[78], thus $r(a) \leq r(a^n)^{1/n}$, for all n. With $r(b) \leq \|b\|$ we have $r(a) \leq \inf_{n\in\mathbb{N}} r(a^n)^{1/n} \leq \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \leq \|a\|$. ■

13.28 REMARK 1. Already by the above we have: $\|a^n\|^{1/n} \to 0 \Rightarrow a$ is quasi-nilpotent.

2. There is another way of improving on $r(a) \leq \|a\|$: Let $\mathcal{A}$ be a unital Banach algebra, $a \in \mathcal{A}$ and $b \in \mathrm{Inv}\mathcal{A}$. Then with Remark 13.23.2 and Proposition 13.27 we have $r(a) = r(bab^{-1}) \leq \|bab^{-1}\|$, implying $r(a) \leq \inf_{b\in\mathrm{Inv}\mathcal{A}} \|bab^{-1}\|$. In favorable cases including $\mathcal{A} = B(H)$ one can prove this to be an equality, cf. Exercise 17.15.

3. The completeness assumption is essential: If $\mathcal{A} = \mathbb{F}[x]$ with the incomplete norm $\|P\| = \sup_{x\in[0,1]} |P(x)|$ then $\sigma(a) = \mathbb{F}$ for all $a \in \mathcal{A}\backslash\mathbb{C}\mathbf{1}$. □

13.29 EXERCISE Let $\mathcal{A}$ be a unital Banach algebra and $a, b \in \mathcal{A}$. Prove the first and second 'resolvent identities'

$$\begin{aligned} R_a(s) - R_a(t) &= (s - t)R_a(s)R_a(t) \quad \forall s, t \in \mathbb{F}\backslash\sigma(a), \\ R_a(s) - R_b(s) &= R_a(s)(b - a)R_b(s) \quad \forall s \in \mathbb{F}\backslash(\sigma(a) \cup \sigma(b)). \end{aligned}$$

[78] Thus $\{\lambda^n \mid \lambda \in \sigma(a)\} \subseteq \sigma(a^n)$. Later we will prove that equality holds if $\mathbb{F} = \mathbb{C}$.

13.30 Exercise Let $\mathcal{A}$ be unital Banach algebra and $a \in \mathrm{Inv}\mathcal{A}$. Prove:

(i) $\sigma(a^{-1}) = \{\lambda^{-1} \mid \lambda \in \sigma(a)\}$.

(ii) If $\|a\| \leq 1$ and $\|a^{-1}\| \leq 1$ then $\sigma(a) \subseteq S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$.

13.31 Exercise Let $\mathcal{A}$ be a unital Banach algebra, $a \in \mathcal{A}$. Prove that for all $\lambda \in \mathbb{F}\backslash\sigma(\mathcal{A})$:

(i) $r((a - \lambda\mathbf{1})^{-1}) = (\mathrm{dist}(\lambda, \sigma(a)))^{-1}$.

(ii) $\|R_a(\lambda)\| \geq (\mathrm{dist}(\lambda, \sigma(a)))^{-1}$

13.32 Remark 1. The result of (ii) could also be deduced from Exercise 13.21, but the approach via $r(R_a(\lambda))$ is more conceptual and will give the exact result for $\|R_a(\lambda)\|$ in certain situations, cf. Exercise 13.71.

2. Since $r(b) < \|b\|$ is perfectly possible, the above in general only gives a lower bound for $\|R_a(\lambda)\|$. Proving upper bounds tends to be harder. □

13.33 Exercise Let $\mathcal{A}$ be a unital Banach algebra and $a \in \mathcal{A}$ nilpotent. With $N = \min\{n \in \mathbb{N} \mid a^n = 0\}$ prove $\lim_{\lambda\to 0} |\lambda|^N \|(a - \lambda\mathbf{1})^{-1}\| = \|a^{N-1}\| \neq 0$. (Thus $\|(a - \lambda\mathbf{1})^{-1}\|$ behaves like $|\lambda|^{-N}\|a^{N-1}\|$ as $\lambda \to 0$.)

It is instructive to compare the above with Exercise 13.52.

13.34 Exercise Let $\mathcal{A}$ be a unital Banach algebra. For $a \in \mathcal{A}$ define $\zeta(a) = \inf\{\|ab\| \mid b \in \mathcal{A}, \|b\| = 1\} = 0$ and call a a topological left zero-divisor if $\zeta(a) = 0$. Prove:

(i) For $a \in \mathrm{Inv}\mathcal{A}$ we have $\zeta(a) = \|a^{-1}\|^{-1} > 0$. Thus $\zeta^{-1}(0) \subseteq \mathcal{A}\backslash\mathrm{Inv}\mathcal{A}$.

(ii) $|\zeta(a) - \zeta(b)| \leq \|a - b\| \ \forall a, b \in \mathcal{A}$.

(iii) If $a \in \partial\mathrm{Inv}\mathcal{A}$[79] then a is not invertible.

(iv) Every $a \in \partial\mathrm{Inv}\mathcal{A}$ is a topological left zero-divisor. Hint: Use Exercise 13.21.

(v) If $\lambda \in \partial\sigma(a)$ then $a - \lambda\mathbf{1}$ is a topological left zero-divisor.

(vi) For $\mathcal{A} = C(X, \mathbb{F})$, where X is compact, prove that $f \in \mathcal{A}$ is a topological (left) zero-divisor if and only if it non-invertible.

(vii) Give an example of a unital Banach algebra $\mathcal{A}$ and a non-invertible $a \in \mathcal{A}$ that is not a topological left zero-divisor.

13.35 Exercise Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{F}$.

(i) Use the ideas in the proof of Lemma 13.19 to give an alternative proof of the continuity of $\mathrm{Inv}\mathcal{A} \to \mathrm{Inv}\mathcal{A}$, $a \mapsto a^{-1}$.

(ii) BONUS: If E, F are normed spaces and $U \subseteq E$ is open, a map $f : U \to F$ is Fréchet differentiable at $x \in U$ if there is a bounded linear map $D \in B(E, F)$ such that

$$\frac{\|f(x+h) - f(x) - D(h)\|}{\|h\|} \to 0 \quad \text{as} \quad \|h\| \to 0.$$

Prove that $\mathrm{Inv}\mathcal{A} \to \mathrm{Inv}\mathcal{A}$, $a \mapsto a^{-1}$ is Fréchet differentiable. For $\mathbb{F} = \mathbb{C}$, conclude that the map $\mathbb{C}\backslash\sigma(a) \to \mathbb{C},\ \lambda \mapsto \varphi((a - \lambda\mathbf{1})^{-1})$ is holomorphic for each $a \in \mathcal{A}$, $\varphi \in \mathcal{A}^*$.

[79] If X is a topological space and $Y \subseteq X$ then $\partial Y = \overline{Y} \cap \overline{X\backslash Y}$ is the boundary of Y.

13.2.3 The spectral radius formula (Beurling-Gelfand theorem)

13.36 Lemma *Let $\mathcal{A}$ be a unital normed algebra and $a \in \mathcal{A}$. Put $\nu = \inf_{n\in\mathbb{N}} \|a^n\|^{1/n}$. Then*

(i) $\lim_{n\to\infty} \|a^n\|^{1/n} = \nu$.[80][81]

(ii) For all $\mu > \nu$ we have $\left(\frac{a}{\mu}\right)^n \to 0$ as $n \to \infty$, but $\left(\frac{a}{\nu}\right)^n \not\to 0$ provided $\nu > 0$.[82]

(iii) If $\nu = 0$ then $a \notin \mathrm{Inv}\mathcal{A}$, thus $0 \in \sigma(a)$.

Proof. (i) With $\|a^n\| \le \|a\|^n$ we trivially have

$$0 \le \nu = \inf_{n\in\mathbb{N}} \|a^n\|^{1/n} \le \liminf_{n\to\infty} \|a^n\|^{1/n} \le \limsup_{n\to\infty} \|a^n\|^{1/n} \le \|a\| < \infty. \tag{13.2}$$

By definition of ν, for every $\varepsilon > 0$ there is a k such that $\|a^k\|^{1/k} < \nu + \varepsilon$. Every $m \in \mathbb{N}$ is of the form $m = sk + r$ with unique $s \in \mathbb{N}_0$ and $0 \le r < k$ (division with remainder). Then

$$\|a^m\| = \|a^{sk+r}\| \le \|a^k\|^s \|a\|^r < (\nu+\varepsilon)^{sk}\|a\|^r,$$

$$\|a^m\|^{1/m} \le (\nu+\varepsilon)^{\frac{sk}{sk+r}} \|a\|^{\frac{r}{sk+r}}.$$

Now $m \to \infty$ means $\frac{sk}{sk+r} \to 1$ and $\frac{r}{sk+r} \to 0$, so that $\limsup_{m\to\infty} \|a^m\|^{1/m} \le \nu + \varepsilon$. Since this holds for every $\varepsilon > 0$, we have $\limsup_{m\to\infty} \|a^m\|^{1/m} \le \inf_{n\in\mathbb{N}} \|a^n\|^{1/n}$. Together with (13.2) this implies that $\lim_{m\to\infty} \|a^m\|^{1/m}$ exists and equals $\inf_{n\in\mathbb{N}} \|a^n\|^{1/n}$.

(ii) Let $\mu > \nu$, and choose μ' such that $\nu < \mu' < \mu$. Since $\|a^n\|^{1/n} \to \nu$ by (i), there exists n_0 such that $n \ge n_0 \Rightarrow \|a^n\|^{1/n} < \mu'$. For such n we have

$$\left\|\left(\frac{a}{\mu}\right)^n\right\| = \frac{\|a^n\|}{\mu^n} < \left(\frac{\mu'}{\mu}\right)^n \stackrel{n\to\infty}{\longrightarrow} 0.$$

This proves the first claim. On the other hand, by definition of ν we have $\|a^n\|^{1/n} \ge \nu$ for all $n \in \mathbb{N}$. With $\nu > 0$ this implies $\|(a/\nu)^n\| \ge 1\ \forall n$, and therefore $(a/\nu)^n \not\to 0$.

(iii) If $a \in \mathrm{Inv}\mathcal{A}$ then there is $b \in \mathcal{A}$ such that $ab = ba = \mathbf{1}$. Then $\mathbf{1} = a^n b^n$, thus with Remark 3.29 we have $1 \le \|\mathbf{1}\| = \|a^n b^n\| \le \|a^n\|\|b^n\| \le \|a^n\|\|b\|^n$. Taking n-th roots, we have $1 \le \|a^n\|^{1/n}\|b\|$, and taking the limit $n \to \infty$ gives the contradiction $1 \le \nu\|b\| = 0$. Thus if $\nu = 0$ then a is not invertible, so that $0 \in \sigma(a)$. ■

Essentially everything we did so far works for $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$. That the natural ONB for $L^2([0,2\pi],\lambda;\mathbb{F})$ depends on $\mathbb{F}$ is a triviality (but related to the significant fact that $x \mapsto e^{ix}$ is a homomorphism, while $x \mapsto \sin x, \cos x$ are not). The $\mathbb{R}/\mathbb{C}$-dependence in the polarization identities (Exercise 5.13 and Lemma 11.17) is more serious since it propagates to the fact that Lemma 11.19 and Propositions 11.22 and 11.34 are weaker over $\mathbb{R}$ than over $\mathbb{C}$.

By contrast, the rest of this section requires $\mathbb{F} = \mathbb{C}$, and the same applies whenever we use Theorem 13.39 and its consequences like Corollary 13.43 or Exercise 13.46. For this reason, from now on we assume $\mathbb{F} = \mathbb{C}$ throughout. If we still mention $\mathbb{C}$, it is only for emphasis. Identifying the few results that also hold over $\mathbb{R}$ (like like Lemma 14.1 through Corollary 14.4) usually is quite easy.

[80] This would be immediate if $n \mapsto \|a^n\|^{1/n}$ was decreasing, but this need not hold! See Exercise 13.64.

[81] More generally, if $\{c_n\}_{n\in\mathbb{N}} \subseteq [0,\infty)$ satisfies $c_{n+m} \le c_n c_m\ \forall n,m$ then $\lim_{n\to\infty} c_n^{1/n} = \inf_{n\in\mathbb{N}} c_n^{1/n}$.

[82] This is of course trivial if $\mu > \|a\|$, but $\mu > \nu$ is a weaker hypothesis when $\nu < \|a\|$.

13.37 LEMMA *Let $\mathcal{A}$ be a unital algebra over $\mathbb{C}$, and let $a \in \mathcal{A}$, $\lambda \in \mathbb{C}\backslash\{0\}, n \in \mathbb{N}$. Then $\left(\frac{a}{\lambda}\right)^n - \mathbf{1} \in \mathrm{Inv}\mathcal{A}$ if and only if $\lambda_k = e^{\frac{2\pi i}{n}k}\lambda \notin \sigma(a)$ for all $k = 1, \ldots, n$. In this case,*

$$\left(\left(\frac{a}{\lambda}\right)^n - \mathbf{1}\right)^{-1} = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{a}{\lambda_k} - \mathbf{1}\right)^{-1} \tag{13.3}$$

Proof. For $k \in \mathbb{Z}$ we write $e_k = e^{\frac{2\pi i k}{n}}$. It is obvious that $e_k^n = 1$ for all $k \in \mathbb{Z}$. Since $e_0, e_1, \ldots, e_{n-1}$ are mutually distinct, we have $z^n - 1 = \prod_{k=0}^{n-1}(z - e_k)$. This identity holds in every unital algebra, and replacing z by $a/\lambda \in \mathcal{A}$ and putting $\lambda_k = e_k\lambda$, it implies the first statement. On the other hand, it means that there is a partial fraction expansion[83]

$$\frac{1}{z^n - 1} = \sum_{k=0}^{n-1}\frac{c_k}{z - e_k}$$

for certain unique $c_k \in \mathbb{C}$. Multiplying this equation by $z - e_\ell$ and taking $z \to e_\ell$ gives

$$c_\ell = \lim_{z \to e_\ell}\frac{z - e_\ell}{z^n - 1}.$$

In view of

$$\lim_{z \to e_\ell}\frac{z^n - 1}{z - e_\ell} = \frac{1}{e_\ell}\lim_{z \to e_\ell}\frac{(z/e_\ell)^n - 1}{(z/e_\ell) - 1} = \frac{1}{e_\ell}\lim_{z \to e_\ell}\left(1 + (z/e_\ell) + \cdots + (z/e_\ell)^{n-1}\right) = \frac{n}{e_\ell},$$

where we used $e_\ell^n = 1$ and (13.1), we have $c_k = e_k/n$, so that

$$\frac{1}{z^n - 1} = \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{(z/e_k) - 1}.$$

Replacing z by a/λ herein and using $\lambda_k = \lambda e_k$ we obtain the identity (13.3) in $\mathcal{A}$. That this is justified follows from $\lambda_k \notin \sigma(a)$ for all k (thus also $\lambda^n \notin \sigma(a^n)$) and the following exercise. ■

13.38 EXERCISE Let $\mathcal{A}$ be a unital algebra over $\mathbb{C}$ and assume we have an identity $\sum_{i=1}^{I}\frac{f_i}{g_i} = 0$ in the field $\mathbb{C}(z)$ of rational functions. Prove $\sum_i \frac{f_i(a)}{g_i(a)} = 0$ for all $a \in \mathcal{A}$ for which all $g_i(a)$ are invertible, i.e. $\sigma(a) \cap g_i^{-1}(0) = \emptyset\ \forall i$.

13.39 THEOREM (BEURLING 1938-GELFAND 1939) [84][85] *Let $\mathcal{A}$ be a unital normed algebra over $\mathbb{C}$ (not necessarily complete) and $a \in \mathcal{A}$. Then $\sigma(a) \neq \emptyset$, and*

$$r(a) \geq \inf_{n\in\mathbb{N}}\|a^n\|^{1/n} = \lim_{n\to\infty}\|a^n\|^{1/n}. \tag{13.4}$$

If $\mathcal{A}$ is complete then equality holds in (13.4), which then is called the spectral radius formula.

[83] Most calculus books mention this, but a proof is rarely given. Here is a quick analytic proof in the case at hand, where all zeros of the denominator are simple: If $f(z) = (\prod_{j=1}^{n}(z - z_j))^{-1}$, we have $A_i = \lim_{z\to z_i}(z - z_i)f(z) = \prod_{j\neq i}(z_i - z_j)^{-1} \in \mathbb{C}\backslash\{0\}$ for each $i \in \{1, \ldots, n\}$. Now $f(z) - \frac{A_i}{z - z_i} = \frac{1}{z - z_i}[\prod_{j\neq i}(z - z_j)^{-1} - A_i]$, where the expression in square brackets is holomorphic near z_i, where it vanishes. Thus $f(z) - \frac{A_i}{z - z_i}$ extends holomorphically to a neighborhood of z_i. Continuing like this, $g(z) = f(z) - \sum_{i=1}^{n}\frac{A_i}{z - z_i}$ extends to an entire function. Since f and $\sum_{i=1}^{n}\frac{A_i}{z - z_i}$ tend to zero as $z \to \infty$, it follows that g is bounded, thus constant by Liouville's theorem. Since the constant must be zero, we have proven $f(z) = \sum_{i=1}^{n}\frac{A_i}{z - z_i}$. (For a purely algebraic treatment of the partial fraction expansion, including the case of multiple zeros of the denominator, see e.g. [93, Ch. IV, §4].)

[84] Arne Beurling (1905-1986). Swedish mathematician. Worked mostly on harmonic and complex analysis.

[85] Israel Moiseevich Gelfand (1913-2009). Outstanding Soviet mathematician. Many important contributions to many areas of mathematics, among which functional analysis and Banach algebras.

Proof. The equality of infimum and limit was Lemma 13.36(i). For $a \in \mathcal{A}$, define ν as before. If $\nu = 0$ then $0 \in \sigma(a)$ by Lemma 13.36(iii). Thus $\sigma(a) \neq \emptyset$ and (13.4) is trivially true.

From now on assume $\nu > 0$. Assume that there is no $\lambda \in \sigma(a)$ with $|\lambda| \geq \nu$. This implies that $(a - \lambda\mathbf{1})^{-1}$ exists for all $|\lambda| \geq \nu$ and depends continuously on λ by Lemma 13.17. The same holds (since $|\lambda| \geq \nu > 0$) for the slightly more convenient function

$$\phi : \{\lambda \in \mathbb{C} \mid |\lambda| \geq \nu\} \to \mathcal{A}, \quad \lambda \mapsto \left(\frac{a}{\lambda} - \mathbf{1}\right)^{-1}.$$

Now Lemma 13.37 gives for all λ with $|\lambda| \geq \nu$ and $n \in \mathbb{N}$ that $(\frac{a}{\lambda})^n - \mathbf{1} \in \text{Inv}\mathcal{A}$ with inverse given by (13.3). Pick any $\eta > \nu$. Since the annulus $\Lambda = \{\lambda \in \mathbb{C} \mid \nu \leq |\lambda| \leq \eta\}$ is compact, the continuous map $\phi : \Lambda \to \mathcal{A}$ is uniformly continuous. I.e., for every $\varepsilon > 0$ we can find $\delta > 0$ such that $\lambda, \lambda' \in \Lambda$, $|\lambda - \lambda'| < \delta \Rightarrow \|\phi(\lambda) - \phi(\lambda')\| < \varepsilon$. If $\nu < \mu < \nu + \delta$, we have $|\nu_k - \mu_k| = |\nu - \mu| < \delta$ and therefore $\|\phi(\nu_k) - \phi(\mu_k)\| < \varepsilon$ for all $n \in \mathbb{N}$ and $k = 1, \ldots, n$. Combining this with (13.3) we have $\|((\frac{a}{\nu})^n - \mathbf{1})^{-1} - ((\frac{a}{\mu})^n - \mathbf{1})^{-1}\| \leq \frac{1}{n}\sum_{k=1}^n \|\phi(\nu_k) - \phi(\mu_k)\| < \varepsilon \ \forall n \in \mathbb{N}$, so that:

$$\forall \varepsilon > 0 \ \ \exists \mu > \nu \ \ \forall n \in \mathbb{N} : \ \ \left\| \left((\frac{a}{\nu})^n - \mathbf{1} \right)^{-1} - \left((\frac{a}{\mu})^n - \mathbf{1} \right)^{-1} \right\| < \varepsilon. \tag{13.5}$$

By Lemma 13.36(ii), $\mu > \nu$ implies $(a/\mu)^n \to 0$ as $n \to \infty$. With continuity of the inverse map, $((a/\mu)^n - \mathbf{1})^{-1} \to -\mathbf{1}$. Thus for n large enough we have $\|((a/\mu)^n - \mathbf{1})^{-1} + \mathbf{1}\| < \varepsilon$, and combining this with (13.5) we have $\|((a/\nu)^n - \mathbf{1})^{-1} + \mathbf{1}\| < 2\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have $((a/\nu)^n - \mathbf{1})^{-1} \to -\mathbf{1}$ as $n \to \infty$ and therefore $(a/\nu)^n \to 0$. But this contradicts the other part of Lemma 13.36(ii), so that our assumption that there is no $\lambda \in \sigma(a)$ with $|\lambda| \geq \nu$ is false. Existence of such a λ obviously gives $\sigma(a) \neq \emptyset$ and $r(a) \geq \nu$, completing the proof of the main result. If $\mathcal{A}$ is complete then Proposition 13.27(ii) applies, and combining it with (13.4) the final claim follows. ■

13.40 Remark 1. We emphasize that only the last clause requires completeness of $\mathcal{A}$.

2. The standard proof of the above theorem requires completeness and uses the differentiability of the resolvent map R_a and a certain amount of complex analysis. The more elementary (which does not mean simple) proof given above, due to Rickart[86] ([130] 1958), shows that neither the completeness assumption nor the complex analysis are essential to the problem. (See also Exercise 13.53 and the subsequent remark.)

3. Even though we avoided complex analysis (holomorphicity etc.), it is clear that the proof only works over $\mathbb{C}$. In fact, the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$, the counterexample in Remark 11.20, has empty spectrum over $\mathbb{R}$ (as does every invertible antisymmetric real matrix). □

13.2.4 Applications, complements, exercises

13.41 Corollary ('Fundamental Theorem of Algebra') *Let $P \in \mathbb{C}[z]$ be a polynomial of degree $d \geq 1$. Then there is $\lambda \in \mathbb{C}$ with $P(\lambda) = 0$.*

Proof. We may assume that P is monic, i.e. the coefficient of the highest power z^d is 1. It is not hard to construct a matrix $a_P \in M_{d\times d}(\mathbb{C})$ such that $P(\lambda) = \det(\lambda\mathbf{1} - a_P)$ (do it!). Now $P^{-1}(0) = \sigma(a_P)$, and the claim follows since Theorem 13.39 gives $\sigma(a_P) \neq \emptyset$. ■

[86] Charles Earl Rickart (1913-2002). American mathematician, mostly operator algebraist.

13.42 Remark The above argument is not circular since the proof of Theorem 13.39 did not use the FTA but only some information about the exponential function in the complex domain (existence of n-th roots, in particular roots of unity). The same holds for the 'standard' proof of the FTA, cf. e.g. [108, Theorem 7.7.57], with which the above has much in common. Both proofs certainly are more elementary than those using complex analysis (Liouville's theorem) or topological arguments based on $\pi_1(S^1) \neq 0$. □

13.43 Corollary (Gelfand-Mazur)

(i) Every unital normed algebra over $\mathbb{C}$ other than $\mathbb{C}\mathbf{1}$ has non-zero non-invertible elements.

(ii) If $\mathcal{A}$ is a normed division algebra (i.e. unital with $\text{Inv}\mathcal{A} = \mathcal{A}\backslash\{0\}$) over $\mathbb{C}$ then $\mathcal{A} = \mathbb{C}\mathbf{1}$.

Proof. (i) Let $a \in \mathcal{A}\backslash\mathbb{C}\mathbf{1}$. By Theorem 13.39 we can pick $\lambda \in \sigma(a)$. Then $a - \lambda\mathbf{1}$ is non-zero and non-invertible. Now (ii) is immediate. ■

13.44 Remark 1. That every *finite-dimensional* division algebra over $\mathbb{C}$ is isomorphic to $\mathbb{C}$ is an easy consequence of algebraic closedness. (Why?) There are infinite-dimensional ones (like the field of $\mathbb{C}(z)$ of rational functions over $\mathbb{C}$), but they do not admit norms as a consequence of the above corollary, which does not assume finite-dimensionality of $\mathcal{A}$.

2. Over $\mathbb{R}$ a theorem of Hurwitz[87] says that there are precisely four division algebras admitting a norm, namely $\mathbb{R}, \mathbb{C}, \mathbb{H}$ (Hamilton's[88] quaternions, which everyone should know) and $\mathbb{O}$, the octonions of Graves[89]. But of these only $\mathbb{C}$ is an algebra over $\mathbb{C}$. For more on the fascinating subject of real division algebras see the 120 pages on the subject in [44]. □

The preceding corollaries only used $\sigma(a) \neq \emptyset$, but also the spectral radius formula will have many applications.

13.45 Exercise Let $\mathcal{A}$ be a unital normed algebra over $\mathbb{C}$ and $a \in \mathcal{A}$. Prove that a is quasi-nilpotent ($r(a) = 0$) if and only if $\lim_{n\to\infty} \|(za)^n\| = 0$ for all $z \in \mathbb{C}$.

If $\mathcal{A}$ is a Banach algebra with unit $\mathbf{1}$, $\mathcal{B} \subseteq \mathcal{A}$ a Banach subalgebra (=closed subalgebra) containing $\mathbf{1}$ and $b \in \mathcal{B}$ then we can consider the spectrum of b as an element of $\mathcal{A}$ or of $\mathcal{B}$, leading to $\sigma_{\mathcal{A}}(b), \sigma_{\mathcal{B}}(b)$ and the spectral radii $r_{\mathcal{A}}(b), r_{\mathcal{B}}(b)$.

13.46 Exercise Let $\mathcal{A}$ be a Banach algebra over $\mathbb{C}$ with unit $\mathbf{1}$. Prove:

(i) If $a, b \in \mathcal{A}$ with $ab = ba$ then $r(ab) \leq r(a)r(b)$.

(ii) If $\mathcal{B} \subseteq \mathcal{A}$ is a Banach subalgebra containing $\mathbf{1}$ then $r_{\mathcal{B}}(b) = r_{\mathcal{A}}(b)$ for all $b \in \mathcal{B}$.

Despite the above (ii), $\sigma_{\mathcal{B}}(b) = \sigma_{\mathcal{A}}(b)$ does not necessarily hold!

13.47 Lemma *Let $\mathcal{A}$ be a Banach algebra over $\mathbb{C}$ with unit $\mathbf{1}$ and $\mathcal{B} \subseteq \mathcal{A}$ a Banach subalgebra containing $\mathbf{1}$. Then*

(i) $\text{Inv}\mathcal{B} \subseteq \mathcal{B} \cap \text{Inv}\mathcal{A}$ and $\sigma_{\mathcal{A}}(b) \subseteq \sigma_{\mathcal{B}}(b)$ for all $b \in \mathcal{B}$.

(ii) $\sigma_{\mathcal{B}}(b) = \sigma_{\mathcal{A}}(b)$ holds for all $b \in \mathcal{B}$ if and only if $\text{Inv}\mathcal{B} = \mathcal{B} \cap \text{Inv}\mathcal{A}$.

[87] Adolf Hurwitz (1859-1919). German mathematician who worked on many subjects.

[88] Sir William Rowan Hamilton (1805-1865). Irish mathematician. Known particularly for quaternions and Hamiltonian mechanics. It was he who advocated the modern view of complex numbers as pairs of real numbers.

[89] John T. Graves (1806-1870). Irish jurist (!) and mathematician.

Proof. (i) The first statement is obvious. If $b - \lambda\mathbf{1}$ has an inverse in $\mathcal{B}$ then the latter also is an inverse in $\mathcal{A}$. Thus $\lambda \notin \sigma_{\mathcal{B}}(b) \Rightarrow \lambda \notin \sigma_{\mathcal{A}}(b)$.

(ii) Assume $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$. Then for all $b \in \mathcal{B}$, $\lambda \in \mathbb{F}$ we have that $b - \lambda\mathbf{1}$ is invertible in $\mathcal{B}$ if and only if it is invertible in $\mathcal{A}$, so that $\sigma_{\mathcal{B}}(b) = \sigma_{\mathcal{A}}(b)$. If $\mathrm{Inv}\mathcal{B} \neq \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$ then in view of $\mathrm{Inv}\mathcal{B} \subseteq \mathcal{B} \cap \mathcal{A}$ we have $\mathrm{Inv}\mathcal{B} \subsetneqq \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$. If now $b \in (\mathcal{B} \cap \mathrm{Inv}\mathcal{A}) \backslash \mathrm{Inv}\mathcal{B}$ then $0 \in \sigma_{\mathcal{B}}(b)$, while $0 \notin \sigma_{\mathcal{A}}(b)$. ■

Here is an example of a Banach subalgebra $\mathcal{B} \subseteq \mathcal{A}$ with $\mathrm{Inv}\mathcal{B} \subsetneqq \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$:

13.48 Example In Section 4.6 we saw that the Banach space $\mathcal{A} = \ell^1(\mathbb{Z}, \mathbb{C})$ with norm $\|\cdot\|_1$ becomes a Banach algebra when equipped with the convolution product $\star$. The functions $\delta_n(m) = \delta_{n,m}$ satisfy $\|\delta_n\| = 1$ and $\delta_n \star \delta_m = \delta_{n+m}$. In particular $\delta_n^{-1} = \delta_{-n}$ for each $n \in \mathbb{Z}$. Now Exercise 13.30 gives $\sigma(\delta_n) \subseteq S^1$ for all n.

Let $\mathcal{B} = \{f \in \mathcal{A} \mid f(n) = 0 \ \forall n < 0\} = \overline{\mathrm{span}_{\mathbb{C}}\{\delta_n \mid n \geq 0\}} \subseteq \mathcal{A}$. It is immediate that $\mathcal{B}$ is a closed subalgebra containing $\mathbf{1}$. If δ_1 had an inverse $c \in \mathcal{B}$, c would also be an inverse in $\mathcal{A}$, so that $c = \delta_{-1}$ by uniqueness of inverses. In view of $\delta_{-1} \notin \mathcal{B}$ we have $\delta_1 \in (\mathcal{B} \cap \mathrm{Inv}\mathcal{A}) \backslash \mathrm{Inv}\mathcal{B}$.

13.49 Exercise Let $\mathcal{A} = \ell^1(\mathbb{Z}, \mathbb{C})$ and $\mathcal{B} \subsetneqq \mathcal{A}$ as above. Prove, using no later results:

(i) $\sigma_{\mathcal{B}}(\delta_1) = \{z \in \mathbb{C} \mid |z| \leq 1\}$.

(ii) $\sigma_{\mathcal{A}}(\delta_1) = S^1$.

Since the spectra depend on the set of invertibles, one is interested in subalgebras $\mathcal{B} \subseteq \mathcal{A}$ satisfying $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$. For a very useful result in this direction see Theorem 16.19. Other examples are provided by the following exercise:

13.50 Exercise Let $\mathcal{A}$ be a unital Banach algebra with unit $\mathbf{1}$. For any subset $S \subseteq \mathcal{A}$, define the 'commutant' S' of S by

$$S' = \{t \in \mathcal{A} \mid st = ts \ \forall s \in S\}.$$

(i) For each $S \subseteq \mathcal{A}$, prove that $\mathcal{B} = S' \subseteq \mathcal{A}$ is a Banach subalgebra with unit $\mathbf{1}$ and that $\mathrm{Inv}(\mathcal{B}) = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$.

(ii) Let $S \subseteq T \subseteq \mathcal{A}$. Prove: (1) $T' \subseteq S'$, (2) $S \subseteq S''$, (3) $S' = S'''$.

(iii) Prove: $S \subseteq \mathcal{A}$ is commutative $\Leftrightarrow$ $S \subseteq S'$ $\Leftrightarrow$ S'' is commutative.

[Combining (i)-(iii) we have: If $S \subseteq \mathcal{A}$ is commutative then $\mathcal{B} = S'' \subseteq \mathcal{A}$ is a commutative Banach subalgebra containing $\mathbf{1}$ and S and satisfying $\mathrm{Inv}(\mathcal{B}) = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$.]

(iv) Prove that a subalgebra $\mathcal{B} \subseteq \mathcal{A}$ is maximal abelian (i.e. abelian and not properly contained in a larger abelian subalgebra of $\mathcal{A}$) if and only if $\mathcal{B} = \mathcal{B}'$. Conclude that every maximal abelian subalgebra $\mathcal{B}$ satisfies $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$.

(v) If H is a Hilbert space, $\mathcal{A} = B(H)$ and $S \subseteq \mathcal{A}$, prove that $S' \subseteq B(H)$ is also τ_{wot}-closed.

13.51 Exercise Let $V = C^1([0,1], \mathbb{C})$ (the differentiable functions $[0,1] \to \mathbb{C}$ with continuous derivative). For $f \in V$ define $\|f\| = \|f\|_\infty + \|f'\|_\infty$.

(i) Prove that $(V, \|\cdot\|)$ is a Banach space. (You may assume from analysis that if $f_n, g \in V$ and $f_n \to g$ and $f'_n \to h$ uniformly then $g' = h$.)

(ii) Show that V is a Banach algebra when multiplication of functions is defined point-wise, i.e. $(fg)(x) = f(x)g(x)$.

(iii) Let $g(x) = x$ for all $x \in [0,1]$. Compute the norm, the spectrum and the spectral radius of g.

(iv) If $E \subseteq [0,1]$ is a closed subset, show that

$$I_E = \{f \in V \mid f(x) = 0 \; \forall x \in E\}$$

is a closed two-sided ideal in V.

(v) For $a \in [0,1]$ let $I_a = \{f \in V \mid f(a) = f'(a) = 0\}$. Show that I_a is a closed ideal that is not of the form I_E as in (v) for any E.

13.52 Exercise Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$ and $a \in \mathcal{A}$ quasi-nilpotent.

(i) Prove that $\sum_{n=0}^{\infty} z^n a^n$ converges absolutely for all $z \in \mathbb{C}$ to $(1 - za)^{-1}$.

(ii) Prove that a resolvent bound $\|(a - \lambda\mathbf{1})^{-1}\| \leq C|\lambda|^{-D} \; \forall \lambda \neq 0$ implies $a^N = 0$ for $N = \lfloor D \rfloor$. Hint: Use Lemma B.170.

13.53 Exercise Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$, $a \in \mathcal{A}$ and $z \in \mathbb{C}\backslash\{0\}$ such that $zS^1 \cap \sigma(a) = \emptyset$ (i.e. there is no $\lambda \in \sigma(a)$ with $|\lambda| = |z|$). Let $p_n(a,z) = (\mathbf{1} - (a/z)^n)^{-1}$. Prove:

(i) If $|z| > r(a)$ then $\lim_{n\to\infty} p_n(a,z) = \mathbf{1}$.

(ii) If there is no $\lambda \in \sigma(a)$ with $|\lambda| < |z|$ then $\lim_{n\to\infty} p_n(a,z) = 0$.

(iii) The limit $p(a,z) = \lim_{n\to\infty} p_n(a,z) \in \mathcal{A}$ exists, commutes with a and depends only on $|z|$. Hint: Lemma 13.37.

(iv) $p(a,z) = \lim_{n\to\infty}(\mathbf{1} - (a/z)^{2n+1})^{-1} = \lim_{n\to\infty}(\mathbf{1} + (a/z)^{2n+1})^{-1}$.

(v) Use the preceding results to prove $p(a,z)^2 = p(a,z)$.

(vi) BONUS: $p(a,z) = -\frac{1}{2\pi i}\oint_C R_a(z)dz$, where C is the circle of radius $|z|$ around $0 \in \mathbb{C}$ with counterclockwise orientation. (This is used as the definition of $p(a,z)$ in the standard approach and for proving its properties using holomorphicity of $z \mapsto R_a(z)$.)

13.54 Remark The operators $((a/z)^n - \mathbf{1})^{-1} = -p_n$ already appeared in our (that is Rickart's) proof of the Beurling-Gelfand theorem, where we did not need their convergence as $n \to \infty$. The result of (vi) to the effect that the limit $p(a,z)$, known as (F.) Riesz idempotent, is given by a contour integral establishes a connection between Rickart's proof and the standard textbook proof via complex analysis. See also [132, §149]. □

Since we need a unit in order to define Inv$\mathcal{A}$ and $\sigma(a)$, the following construction is quite important when dealing with non-unital algebras:

13.55 Exercise (Unitization of Banach algebras) Let $\mathcal{A}$ be a Banach algebra over $\mathbb{F}$, possibly without unit. Define $\widetilde{\mathcal{A}} = \mathcal{A} \oplus \mathbb{F}$, which is an $\mathbb{F}$-vector space in the obvious way. For $(a,\alpha), (b,\beta) \in \widetilde{\mathcal{A}}$ define $(a,\alpha)(b,\beta) = (ab + \alpha b + \beta a, \alpha\beta)$ and $\|(a,\alpha)\|_1 = \|a\| + |\alpha|$. Prove:

(i) $\widetilde{\mathcal{A}}$ is an algebra with unit $(0,1)$, the map $\iota : \mathcal{A} \to \widetilde{\mathcal{A}}, a \mapsto (a,0)$ is an algebra homomorphism, and $\iota(A) \subseteq \widetilde{\mathcal{A}}$ is a two-sided ideal.

(ii) $(\widetilde{\mathcal{A}}, \|\cdot\|_1)$ is a normed algebra, and $\iota : \mathcal{A} \to \widetilde{\mathcal{A}}$ is an isometry.

(iii) $(\widetilde{\mathcal{A}}, \|\cdot\|_1)$ is complete.

13.3 Spectra of bounded operators II: Banach algebra methods

In this section, we apply our results on abstract Banach algebras to the study of some operators on Banach spaces.

13.56 Exercise Consider the left and right shift operators L, R of Definition 13.2 in the Hilbert space $\ell^2(\mathbb{N}, \mathbb{C})$.

(i) Prove $\sigma(L) = \sigma(R) = \overline{B}(0,1)$ (closed unit disc).

(ii) Determine σ_c and σ_r for L, R.

(iii) Find σ_{ap} and σ_{cp} for L, R.

Hint: Use Exercises 13.7, 13.9.

13.57 Exercise Let $H = \ell^2(\mathbb{N}, \mathbb{C})$ and define $A \in B(H)$ by $(Af)(n) = f(n)/n$. Determine $\sigma_p(A), \sigma_c(A), \sigma_r(A)$.

13.58 Exercise (Assuming some measure theory) Let $H = L^2([a,b], \mathbb{C})$, where $-\infty < a < b < \infty$, and define $A \in B(H)$ by $(Af)(x) = xf(x)$. Prove $\sigma_c(A) = [a,b]$ and $\sigma_p(A) = \sigma_r(A) = \emptyset$.

13.59 Exercise Prove that for every compact set $C \subseteq \mathbb{C}$ there is an operator $A \in B(H)$, where H is a separable Hilbert space, such that $\sigma(A) = C$.

Hint: Prove and use that C has a countable dense subset.

13.60 Exercise Prove that every quasi-nilpotent operator on a finite-dimensional complex Banach space is nilpotent and non-injective, thus $0 \in \sigma_p$.

13.61 Exercise Let $V = C([0,1], \mathbb{F})$ with (complete) norm $\|\cdot\|_\infty$. Define the Volterra operator $A : V \to V$ by $(Af)(x) = \int_0^x f(t)dt$. Prove

(i) A is injective.

(ii) A is bounded and satisfies $\|A^n\| = 1/n!$ for all $n \in \mathbb{N}$.

(iii) A is quasi-nilpotent, but not nilpotent, and $0 \in \sigma_r(A)$.

The next three exercises study an important class of bounded operators on $\ell^2(\mathbb{N})$, the 'weighted shift operators', generalizing the right shift R, and two typical applications:

13.62 Exercise (i) Let H be a Hilbert space and $A \in B(H)$. Define $p_n = \|A^n\|$ for $n \in \mathbb{N}_0$ and prove that $p_{i+j} \le p_i p_j \ \forall i, j$.

(ii) Let $\{\alpha_k\}_{k \in \mathbb{N}_0} \subseteq \mathbb{C}$. Put $H = \ell^2(\mathbb{N}_0, \mathbb{C})$ with natural ONB $\{e_k\}$ and define a linear 'weighted shift operator' $A : H \to H$ by $Ae_k = \alpha_k e_{k+1}$. Prove that $\|A\| = \sup_k |\alpha_k|$.

(iii) Let $\{p_k\}_{k \in \mathbb{N}_0}$ be given so that $p_0 = 1, \ p_k > 0 \ \forall k$ and $p_{i+j} \le p_i p_j \ \forall i, j$. Define

$$\alpha_0 = p_0 \qquad \text{and} \qquad \alpha_k = \frac{p_k}{p_{k-1}} \ \text{ if } \ k \ge 1.$$

Prove that the sequence $\{\alpha_k\}$ and the associated weighted shift operator A are bounded.

(iv) Let A be constructed from $\{p_k\}_{k \in \mathbb{N}_0}$ as in (iii). Prove $\|A^n\| = p_n \ \forall n \in \mathbb{N}_0$.

(v) (Bonus) Adapt the above construction to the case where $p_k = 0$ is allowed for some $k > 0$.

13.63 Exercise Let $H = \ell^2(\mathbb{N}, \mathbb{C})$ and define $A \in B(H)$ by $Ae_k = 2^{-k}e_{k+1}\ \forall k$. Prove that A is (i) injective, (ii) quasi-nilpotent, but (iii) not nilpotent.

13.64 Exercise Let $H = \ell^2(\mathbb{N}, \mathbb{C})$ and define $A \in B(H)$ by $Ae_k = \alpha_k e_{k+1}$, where $\alpha_k = 2$ for odd k and $\alpha_k = 1/2$ for even k. Compute $\|A^n\|$ for all n and show that $n \mapsto \|A^n\|^{1/n}$ is not monotonously decreasing.

13.65 Exercise Let V be a Banach space and $A \in B(V)$.

(i) Prove that A is not bounded below if and only if it is a topological left zero-divisor of $B(V)$ as defined in Exercise 13.34.[90]

(ii) Conclude that $\partial\sigma(A) \subseteq \sigma_{\mathrm{ap}}(A)$.

Note that the preceding exercises only used results up to Section 13.2.2! The next exercise is a (rather weak, given the strong hypothesis) converse of Exercise 13.8:

13.66 Exercise Let V be a complex Banach space and $A \in B(V)$ such that $\sigma(A)$ is disjoint from the circle $C = \{z \in \mathbb{C} \mid |z - z_0| = r\}$.

(i) Apply Exercise 13.53 to $A - z_0\mathbf{1}$ to obtain $P^2 = P \in B(V)$ satisfying $PA = AP$.

(ii) Prove that $AV_i \subseteq V_i$, where $V_1 = PV,\ V_2 = (1 - P)V$. Conclude that $A = A_1 \oplus A_2$ where $A_i = A{\restriction}V_i$.

(iii) Prove $V_1 = \{x \in V \mid \lim_{n\to\infty}(A - z_0\mathbf{1})^n x = 0\}$.

(iv) Deduce $\sigma(A_1) = \sigma(A) \cap B(z_0, r)$ and $\sigma(A_2) = \sigma(A)\backslash\overline{B}(z_0, r)$.

13.67 Remark The unnatural assumption that the two parts of the spectrum are separated by a circle can be removed using holomorphic functional calculus. Cf. e.g. [132, §149], [30, Chapter VII, §4]. For normal operators on Hilbert space, there is a more powerful approach, cf. Proposition 17.24. □

13.68 Exercise (Isolated points in the spectrum) Let V be a complex Banach space, $A \in B(V)$ and $\lambda \in \sigma(A)$ isolated. Pick $r > 0$ such that $\overline{B}(\lambda, r) \cap \sigma(A) = \{\lambda\}$, and put $C = \{z \mid |z - \lambda| = r\}$ and let P_λ, V_i, A_i be as constructed in Exercise 13.66. Prove that

(i) $A_1 - \lambda\mathbf{1}$ is quasi-nilpotent.

(ii) $A_2 - \lambda\mathbf{1}$ is invertible.

(iii) If $\dim V_1 < \infty$ then $\lambda \in \sigma_{\mathrm{p}}(A_1) \subseteq \sigma_{\mathrm{p}}(A)$.

Even though we developed the theory of the Riesz projector (Exercises 13.53, 13.66, 13.68) reasonably fully only for isolated points of the spectrum, it will suffice for applications to the spectral theory of compact operators in Section 14.4 and to the discussion of the discrete spectrum, cf. Section B.10.

[90] Thus if $\mathcal{A}$ is a unital Banach algebra and $a \in \mathcal{A}$, defining $\sigma_{\mathrm{ap}}(a) = \{\lambda \in \mathbb{F} \mid a - \lambda\mathbf{1}$ is topological left zero-divisor$\}$ is consistent with the usual definition when $\mathcal{A} = B(V)$. With this, (ii) also holds for Banach algebras.

13.4 Applications to normal Hilbert space operators

We will have more to say on abstract Banach algebras, in particular the subclass of C^*-algebras still to be defined. But before turning to these matters, we will consider some applications of the above to operator theory.

13.69 Proposition *If H is a complex Hilbert space and $A \in B(H)$ is normal then*

(i) $r(A) = \|A\|$.

(ii) There exists $\lambda \in \sigma(A)$ such that $|\lambda| = \|A\|$.

(iii) $\|A\| = |||A\,|||$ $\left(= \sup_{\|x\|=1} |\langle Ax, x\rangle| \right)$.

Proof. (i) By Exercise 11.31 we have $\|A^{2^n}\| = \|A\|^{2^n}$ for all normal A. Now Theorem 13.39 gives

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{n\to\infty} \|A^{2^n}\|^{1/2^n} = \lim_{n\to\infty} \left(\|A\|^{2^n}\right)^{1/2^n} = \|A\|. \tag{13.6}$$

(ii) Since $\sigma(A)$ is compact, the continuous real-valued function $\sigma(A) \to \mathbb{C}$, $\lambda \mapsto |\lambda|$ assumes its supremum $r(A)$, which equals $\|A\|$ by (i).

(iii) By Exercise 13.15 we have $r(A) \le |||A||| \le \|A\|$ for every $A \in B(H)$. Now for normal A the claim clearly follows from (i). ■

13.70 Remark Recall that Exercise 11.35 gave a direct proof of $|||A||| = \|A\|$ not using Theorem 13.39. Using this one can also give a direct proof [14] of (i),(ii): In Exercise B.141 it is shown that for every $A \in B(H)$ there exists a sequence $\{x_n\}$ with $\|x_n\| = 1$ such that $\langle Ax_n, x_n\rangle \to \lambda$, where $|\lambda| = |||A|||$, which equals $\|A\|$ by normality. Now

$$\begin{aligned} \|Ax_n - \lambda x_n\|^2 &= \langle Ax_n - \lambda x_n, Ax_n - \lambda x_n\rangle \\ &= \langle Ax_n, Ax_n\rangle - \lambda\langle x_n, Ax_n\rangle - \overline{\lambda}\langle Ax_n, x_n\rangle + |\lambda|^2\langle x_n, x_n\rangle. \end{aligned}$$

The sum of the three rightmost terms converges to $-|\lambda|^2 = -\|A\|^2$. Since $\langle Ax_n, Ax_n\rangle = \|Ax_n\|^2 \le \|A\|^2\ \forall n$, we have $\|Ax_n - \lambda x_n\| \to 0$, proving that $A - \lambda\mathbf{1}$ is not bounded below. Thus $\lambda \in \sigma(A)$, so that $r(A) \ge |\lambda| = \|A\|$. Combining this with $r(A) \le \|A\|$ from Proposition 13.27, we have $r(A) = \|A\|$. □

13.71 Exercise Let $A \in B(H)$ be normal. Prove:

(i) $\|A^n\| = \|A\|^n \quad \forall n \in \mathbb{N}$.

(ii) If $\sigma(A) = \{\lambda\}$ then $A = \lambda\mathbf{1}$.

(iii) For all $\lambda \in \mathbb{C}\backslash\sigma(A)$ we have $\|R_A(\lambda)\| = \|(A - \lambda\mathbf{1})^{-1}\| = (\mathrm{dist}(\lambda, \sigma(A)))^{-1}$.

14 Compact operators II: Spectral theorems

14.1 The spectrum of compact operators. Fredholm alternative

We now begin studying the spectrum of compact operators. Throughout, V is a Banach space.

14.1 Lemma *If $A \in B(V)$ is compact and $\lambda \in \mathbb{F}\backslash\{0\}$ then* $\ker(A - \lambda\mathbf{1})$ *is finite-dimensional.*

Proof. If $\lambda \notin \sigma(A)$ then this is trivial since $A - \lambda\mathbf{1}$ is invertible. In general, $V_\lambda = \ker(A - \lambda\mathbf{1})$ is the space of eigenvectors of A with eigenvalue λ. Clearly $A_{|V_\lambda} = \lambda\, \mathrm{id}_{V_\lambda}$, so that V_λ is an invariant subspace. Since V_λ is closed and $A_{|V_\lambda}$ is compact by Remark 12.6.3, V_λ must be finite-dimensional by Remark 12.6.4. ■

For $\lambda = 0$, the above does not hold since the zero operator on any V is compact.

14.2 Proposition (Fredholm alternative) [91] *Let V be a Banach space, $A \in B(V)$ compact and $\lambda \in \mathbb{F}\backslash\{0\}$. Then the following are equivalent:*

(i) $A - \lambda\mathbf{1}$ is invertible. (I.e. $\lambda \notin \sigma(A)$.)

(ii) $A - \lambda\mathbf{1}$ is injective.

(iii) $A - \lambda\mathbf{1}$ is surjective.

Proof. We know from Proposition 7.41 that (i) is equivalent to the combination of (ii) and (iii). It therefore suffices to prove (ii)⇔(iii).

(iii)⇒(ii): It suffices to do this for $\lambda = 1$. (Why?) Assume that $A - \mathbf{1}$ is not injective, but surjective. Then $(A - \mathbf{1})^n$ is surjective for all n. In view of $(A - \mathbf{1})^{n+1} = (A - \mathbf{1})(A - \mathbf{1})^n$ we have $\ker(A - \mathbf{1})^{n+1} \supseteq \ker(A - \mathbf{1})^n$. We claim that this inclusion is strict for each n, i.e. $\ker(A - \mathbf{1})^{n+1} \supsetneqq \ker(A - \mathbf{1})^n$: By non-injectivity of $A - \mathbf{1}$ we can find $y \in V\backslash\{0\}$ such that $(A - \mathbf{1})y = 0$. By surjectivity of $(A - \mathbf{1})^n$ there exists x such that $(A - \mathbf{1})^n x = y$. Now $(A - \mathbf{1})^{n+1}x = (A - \mathbf{1})y = 0$ while $(A - \mathbf{1})^n x = y \neq 0$, thus $x \in \ker(A - \mathbf{1})^{n+1} \backslash \ker(A - \mathbf{1})^n$.

Now by Riesz' Lemma 12.2, for each n we can find an $x_n \in \ker(A - \mathbf{1})^{n+1}$ such that $\|x_n\| = 1$ and $\mathrm{dist}(x_n, \ker(A - \mathbf{1})^n) \geq \frac{1}{2}$. If $n > m$ then $(A - \mathbf{1})^n A x_m = A(A - \mathbf{1})^n x_m = 0$ and $(A - \mathbf{1})^{n+1}x_n = 0$, implying $(A - \mathbf{1})x_n - Ax_m \in \ker(A - \mathbf{1})^n$. With the definition of $\{x_n\}$ it follows that for all $n > m$ (thus also $n < m$) we have

$$\|Ax_n - Ax_m\| = \|x_n + ((A - \mathbf{1})x_n - Ax_m)\| \geq \mathrm{dist}(x_n, \ker(A - \mathbf{1})^n) \geq \frac{1}{2}.$$

Thus $\{Ax_n\}$ has no convergent subsequence, contradicting the compactness of A.

(ii)⇒(iii): If $A - \lambda\mathbf{1}$ is injective but not surjective, one similarly proves $(A - \lambda\mathbf{1})^{n+1}V \subsetneqq (A - \lambda\mathbf{1})^n V$ for all n, which again leads to a contradiction with compactness of A. ■

14.3 Remark 1. The result fails for $\lambda = 0$ since there are compact injective operators that are not surjective.

2. In the above proof we have shown that if $A \in B(V)$ is compact and $\lambda \in \mathbb{F}\backslash\{0\}$ then we cannot have $\ker(A - \mathbf{1})^{n+1} \supsetneqq \ker(A - \mathbf{1})^n$ for all n or $(A - \lambda\mathbf{1})^{n+1}H \subsetneqq (A - \lambda\mathbf{1})^n H$ for all n. One says that $A - \lambda\mathbf{1}$ has finite ascent and descent.

3. Compactness is essential for this stabilization: For the shift operators on $V = \ell^p(\mathbb{N})$ we have $\ker L^{n+1} \supsetneqq \ker L^n$ and $R^{n+1}V \subsetneqq R^n V$ for all n. □

Fredholm's alternative has far-reaching consequences for the spectrum of A:

14.4 Corollary *If $A \in B(V)$ is compact then*

(i) $\sigma(A) \subseteq \sigma_{\mathrm{p}}(A) \cup \{0\}$. (Thus all non-zero elements of the spectrum are eigenvalues.)

(ii) If V is infinite-dimensional then $0 \in \sigma(A)$.

[91] Erik Ivar Fredholm (1866-1927). Swedish mathematician. Early pioneer of functional analysis through his work on integral equations.

Proof. (i) For $\lambda \neq 0$, with Proposition 14.2 we have: $\lambda \in \sigma(A) \Leftrightarrow A-\lambda\mathbf{1}$ not invertible $\Leftrightarrow A-\lambda\mathbf{1}$ not injective $\Leftrightarrow \lambda \in \sigma_{\rm p}(A)$.

(ii) If A was invertible then $\mathbf{1} = A^{-1}A$ would be invertible by Lemma 12.10, but this is false by Remark 12.6.3. Thus $0 \in \sigma(A)$. ∎

14.5 Exercise Show that a compact operator A can have 0 in any of $\sigma_{\rm p}(A), \sigma_{\rm c}(A), \sigma_{\rm r}(A)$.

14.6 Exercise Let H be a Hilbert space, $A \in K(H)$ and $\lambda \in \mathbb{C}\backslash\{0\}$. Show that each of the implications (ii)⇒(iii) and (iii)⇒(ii) in Proposition 14.2 can be deduced from the other.

14.2 ⋆ A glimpse of Fredholm operators

The content of this subsection is not needed for the proof of the spectral theorem, but puts the Fredholm alternative into perspective. (For more on Fredholm operators see Section B.9.)

Recall that if $A : V \to W$ is a linear map, the cokernel of A by definition is the linear quotient space W/AV. If V, W are Hilbert spaces and $AV \subseteq W$ is closed then, recalling Exercise 5.32 we may alternatively define the cokernel of A to be $(AV)^\perp \subseteq W$.

14.7 Proposition *Let $A \in B(V)$ be compact and $\lambda \in \mathbb{F}\backslash\{0\}$. Then*

(i) $\mathrm{coker}(A - \lambda\mathbf{1}) = V/((A - \lambda\mathbf{1})V)$ *is finite-dimensional.*

(ii) $(A - \lambda\mathbf{1})V \subseteq V$ *is closed.*

Proof. (i) By Theorem 12.17, A^t is compact, thus $\ker(A^t - \lambda\mathbf{1}_{V^*})$ is finite-dimensional by Lemma 14.1. Now

$$\ker(A^t - \lambda\mathbf{1}_{V^*}) = \ker[(A - \lambda\mathbf{1}_V)^t] = ((A - \lambda\mathbf{1}_V)V)^\perp \cong (V/((A - \lambda\mathbf{1}_V)V))^*,$$

where the second equality and the final isomorphism come from Exercises 9.35(i) and 6.7, respectively. Thus $(V/((A-\lambda\mathbf{1}_V)V))^*$ is finite-dimensional, implying (why?) finite-dimensionality of $V/((A - \lambda\mathbf{1}_V)V) = \mathrm{coker}(A - \lambda\mathbf{1}_V)$.

(ii) This follows from (i) and Exercise 7.11.

Alternatively, a direct proof goes as follows: By Lemma 14.1, $K = \ker(A - \lambda\mathbf{1})$ is finite-dimensional, thus closed by Exercise 3.22. Thus there is a closed subspace $S \subseteq V$ such that $V = K \oplus S$. (If V is a Hilbert space, we can just take $S = K^\perp$. For general Banach spaces this is the statement of Proposition 6.11.) The restriction $(A - \lambda\mathbf{1})_{|S} : S \to H$ is compact and injective. If $(A - \lambda\mathbf{1})_{|S}$ is not bounded below, we can find a sequence $\{x_n\}$ in S with $\|x_n\| = 1$ for all n and $\|(A - \lambda\mathbf{1})x_n\| \to 0$. Since A is compact, we can find a subsequence $\{x_{n_k}\}$ such that $\{Ax_{n_k}\}$ converges. We relabel, so that now $\{Ax_n\}$ converges. Now

$$x_n = \lambda^{-1}[Ax_n - (A - \lambda\mathbf{1})x_n].$$

Since $\{Ax_n\}$ converges and $\{(A - \lambda\mathbf{1})x_n\}$ converges to zero by choice of $\{x_n\}$, $\{x_n\}$ converges to some $y \in S$ (since $x_n \in S\ \forall n$ and S is closed). From $(A - \lambda\mathbf{1})x_n \to 0$ and $x_n \to y$ we obtain $(A - \lambda\mathbf{1})y = 0$, so that $y \in \ker(A - \lambda\mathbf{1}) = K$. Thus $y \in K \cap S = \{0\}$, which is impossible since $y = \lim_n x_n$ and $\|x_n\| = 1\ \forall n$. This contradiction shows that $(A-\lambda\mathbf{1})_{|S}$ is bounded below. Now Lemma 7.39 gives that $(A - \lambda\mathbf{1})H = (A - \lambda\mathbf{1})S$ is closed. ∎

14.8 Remark In fact one can prove more: If $A \in B(V)$ is compact and $\lambda \in \mathbb{C}\backslash\{0\}$ then

$$\dim\ker(A - \lambda\mathbf{1}) = \dim\mathrm{coker}(A - \lambda\mathbf{1}). \quad (14.1)$$

This clearly is much stronger than the equivalence (ii)⇔(iii) in Proposition 14.2, which amounts to the statement $\dim\ker(A-\lambda\mathbf{1})=0 \ \Leftrightarrow\ \dim\operatorname{coker}(A-\lambda\mathbf{1})=0$. See Remark 14.10.2. □

14.9 Definition *If V,W are Banach space then $A\in B(V,W)$ is called a Fredholm operator if both $\ker A$ and $\operatorname{coker} A$ are finite-dimensional.*

If A is Fredholm, $\operatorname{ind}(A)=\dim\ker A-\dim\operatorname{coker} A\in\mathbb{Z}$ is the (Fredholm) index of A.

Note that we do not require closedness of the image of A since Exercise 7.11 it is follows automatically by from the finite-dimensionality of W/AV.

14.10 Remark 1. If V,W are Banach spaces and $A\in B(V,W)$ then $A^t\in B(W^*,V^*)$ is Fredholm if and only if A is. In this case, $\operatorname{ind}(A^t)=-\operatorname{ind}(A)$. (Proposition B.105)

2. If A,B are Fredholm then so is AB and $\operatorname{ind}(AB)=\operatorname{ind}(A)+\operatorname{ind}(B)$. (Proposition B.104.)

3. If $A\in B(V)$ is compact and $\lambda\neq 0$ then Lemma 14.1 and Proposition 14.7 give that $A-\lambda\mathbf{1}$ is Fredholm, and (14.1) amounts to $\operatorname{ind}(A-\lambda\mathbf{1})=0$.

4. The latter identity follows immediately by combining the trivial fact that $\mathbf{1}$ is Fredholm with index zero with the following important stability result: If F is Fredholm and K is compact then $F+K$ is Fredholm and $\operatorname{ind}(F+K)=\operatorname{ind}(F)$. (Theorem B.108.)

5. Another important connection between compact and Fredholm operators is Atkinson's theorem: $A\in B(V)$ is Fredholm if and only there exists $B\in B(V)$ such that $AB-\mathbf{1}$ and $BA-\mathbf{1}$ are compact. (Equivalently, the image of A in the quotient algebra $B(V)/K(V)$ is invertible.) (Theorem B.107.) □

14.3 Spectral theorems for compact Hilbert space operators

14.11 Proposition *Let H be a non-zero complex Hilbert space and $A\in B(H)$ a compact normal operator. Then there is an eigenvalue $\lambda\in\sigma_{\mathrm{p}}(A)$ such that $|\lambda|=\|A\|$.*

Proof. If $A=0$ then it is clear that $\lambda=0$ does the job. Now assume $A\neq 0$. By Proposition 13.69(ii) there exists $\lambda\in\sigma(A)$ with $|\lambda|=\|A\|$. Since $\lambda\neq 0$, Corollary 14.4 gives $\lambda\in\sigma_{\mathrm{p}}(A)$. ■

In linear algebra one proves that for a matrix $A\in M_{n\times n}(\mathbb{C})$ the following are equivalent: A is normal, A can be diagonalized by a unitary matrix, $\mathbb{C}^n$ has an orthonormal basis consisting of eigenvectors of A, the (geometric) dimension of each eigenspace of A coincides with the (algebraic) multiplicity of the corresponding eigenvalue. Cf. e.g. [55, Theorem 6.16]. In basis-independent language, $A\in B(H)$ with H finite-dimensional is normal if and only if H admits an ONB consisting of eigenvectors of A. The following beautiful result generalizes this to compact normal operators:

14.12 Theorem (Spectral theorem for compact normal operators) *Let H be a complex Hilbert space and $A\in B(H)$ compact normal. Then*

(i) H is spanned by the eigenvectors of A.

(ii) There is an ONB E of H consisting of eigenvectors, thus $A=\sum_{e\in E}\lambda_e P_e$, where $P_e = e\otimes e: x\mapsto\langle x,e\rangle e$.

(iii) For each $\varepsilon>0$ there are at most finitely many $\lambda\in\sigma_{\mathrm{p}}(A)$ with $|\lambda|\geq\varepsilon$.

(iv) $\sigma_{\mathrm{p}}(A)$ is at most countable and has no accumulation points except perhaps 0, which is an accumulation point whenever $\sigma(A)$ is infinite.

(v) *We have* $\sigma(A) \subseteq \sigma_{\rm p}(A) \cup \{0\}$, *where* $0 \in \sigma_{\rm p}(A)$ *if and only if* A *has a kernel and* $0 \in \sigma_{\rm c}(A)$ *if and only if* A *is injective and* $\sigma(A)$ *is infinite.*

Proof. (i) Let $K \subseteq H$ be the smallest closed linear subspace containing $\bigcup_{\lambda \in \sigma_{\rm p}(A)} H_\lambda$, where $H_\lambda = \ker(A - \lambda \mathbf{1})$. Clearly K is an invariant subspace: $AK \subseteq K$. Exercise 13.12(i) implies that also $A^*K \subseteq K$. Now Exercise 11.23(i) gives that also $K^\perp$ is A-invariant: $AK^\perp \subseteq K^\perp$. If $K^\perp \neq \{0\}$ then $A_{|K^\perp}$ is compact and has eigenvectors by Proposition 14.11. Since this would contradict the definition of K, we have $K^\perp = 0$, proving that H is spanned by the eigenvectors of A.

(ii) By Exercise 13.12(ii), the eigenspaces for different eigenvalues of A are mutually orthogonal. Now the claim follows from (i) by choosing ONBs E_λ for each H_λ and putting $E = \bigcup_\lambda E_\lambda$.

(iii) Taking into account the unitary equivalence $H \cong \ell^2(E, \mathbb{C})$, cf. Theorem 5.45, this essentially is Exercise 12.13(iii).

(iv) This is an immediate consequence of (iii).

(v) Since A is normal, $0 \in \sigma_{\rm r}(A)$ is ruled out by Exercise 13.13(i). The statement about $0 \in \sigma_{\rm p}(A)$ is trivially true by definition. The one about $\sigma_{\rm c}(A)$ now follows from (iv) and the closedness of $\sigma(A)$. ■

14.13 Remark 1. The statements about $\sigma(A)$ actually hold for all compact operators on Banach spaces. (Instead of the orthogonality of eigenvectors for different eigenvalues, it suffices to use their linear independence.)

2. The common theme of 'spectral theorems' is that normal operators can be diagonalized, i.e. be interpreted as multiplication operators, compactness simplifying statement and proof considerably. Compare Theorem 18.4 for a result not requiring compactness.

3. If $A \in B(H)$ is compact normal and $f : \sigma(A) \to \mathbb{C}$ a function, we formally define $f(A)$ as $\sum_{e \in E} f(\lambda_e) P_e$, where E, λ_e, P_e are as in Theorem 14.12. The sum converges strongly to a bounded operator if and only if f is bounded, and $f(A)$ is compact if and only if $f(\lambda) \to 0$ as $\lambda \to 0$. Thus setting up a 'functional calculus' for compact normal operators is quite easy. Our discussion of not-necessarily-compact normal operators will proceed in the opposite order: We begin in Section 17 by constructing a functional calculus, which will then be used to prove spectral theorems. □

For non-normal operators one can only prove a weaker statement:

14.14 Proposition *Let* H *be a complex Hilbert space and* $A \in K(H)$. *Then are orthonormal sets (not necessarily bases!)* E *and* F *of* H, *a bijection* $E \to F, e \mapsto f_e$ *and positive numbers* $\{\beta_e\}$, *called the singular values of* A, *such that* $e \mapsto \beta_e$ *is in* $c_0(E, \mathbb{C})$ *and*

$$A = \sum_{e \in E} \beta_e f_e \langle \cdot, e \rangle.$$

The β_e *are the precisely the eigenvalues of* $|A|$.

Proof. $B = A^*A$ is compact and self-adjoint, so that by Theorem 14.12 there is an ONB E_B diagonalizing B, thus $B = \sum_{e \in E_B} \lambda_e P_e$. Clearly $E = \{e \in E_B \mid Ae \neq 0\}$ is orthonormal. For $e \in E$ put $f_e = \frac{Ae}{\|Ae\|}$. Now let $F = \{f_e \mid e \in E\}$. The $f_e = \frac{Ae}{\|Ae\|}$ are normalized by definition, and if $e, e' \in E, e \neq e'$ then $\langle Ae, Ae' \rangle = \langle e, A^*Ae' \rangle = 0$ since E diagonalizes A^*A, so that F is orthonormal. For all $x \in H$ we have

$$Ax = A \sum_{e \in E_B} \langle x, e \rangle e = \sum_{e \in E_B} \langle x, e \rangle Ae = \sum_{e \in E} \langle x, e \rangle Ae = \sum_{e \in E} \|Ae\| \langle x, e \rangle f_e,$$

and putting $\beta_e = \|Ae\| > 0$ we have the desired form.

Since E diagonalizes A^*A, we have $A^*Ae = \lambda_e e$ for all $e \in E$, where compactness of A^*A implies that $e \mapsto \lambda_e$ is in $c_0(E_B, \mathbb{C})$, cf. Theorem 14.12(iii). Now, $\|Ae\|^2 = \langle Ae, Ae\rangle = \langle e, A^*Ae\rangle = \lambda_e$, thus $\beta_e = \|Ae\| = \lambda_e^{1/2}$ implies that also $e \mapsto \beta_e$ is in $c_0(E)$. For the final claim, note that $|A|^2 e = A^*Ae = \lambda_e e$, thus $|A|e = \beta_e e$. ■

The following goes some way towards proving that Hilbert spaces have the approximation property:

14.15 Corollary *Let H be a complex Hilbert space, $A \in K(H)$ and $\varepsilon > 0$. Then there is a $B \in F(H)$ (finite rank) with $\|A - B\| \le \varepsilon$. Thus $K(H) = \overline{F(H)}^{\|\cdot\|}$.*

Proof. Pick a representation $A = \sum_{e\in E} \lambda_e f_e \langle \cdot, e\rangle$ as in the preceding proposition. Since $E \to \mathbb{C}$, $e \mapsto \lambda_e$ is in $c_0(E)$, there is a finite subset $F \subseteq E$ such that $|\lambda_e| < \varepsilon$ for all $e \in E\backslash F$. Define

$$B = \sum_{e\in F} \lambda_e \langle \cdot, f_e\rangle e,$$

which clearly has finite rank. If $x \in H$ then using the orthonormality of E and Bessel's inequality, we have

$$\|(A - B)x\|^2 = \|\sum_{e\in E\backslash F} \lambda_e \langle x, f_e\rangle e\|^2 = \sum_{e\in E\backslash F} |\lambda_e \langle x, f_e\rangle|^2 \le \varepsilon^2 \|x\|^2.$$

Thus $\|A - B\| \le \varepsilon$, so that $K(H) \subseteq \overline{F(H)}$. The converse inclusion was Corollary 12.12. ■

14.16 Remark 1. In the above, bases played a crucial role. Even though there is no notion of orthogonality in general Banach spaces, it turns out that Banach spaces having suitable bases do satisfy $K(H) = \overline{F(H)}^{\|\cdot\|}$, i.e. the approximation property. Cf. e.g. [102, Theorem 4.1.33].

2. If you like applications of complex analysis to functional analysis, see [128, Section VI.5] for an interesting alternative approach to compact operators. □

14.4 ⋆ Spectral theorem (Jordan normal form) for compact Banach space operators

We now return to compact operators on general Banach spaces, including of course non-normal compact operators on Hilbert spaces). We begin by reproving some assertions already shown for compact normal Hilbert spaces operators:

14.17 Proposition *Let V be a Banach space and $A \in B(V)$ compact. Then*

(i) For each $\varepsilon > 0$ there are at most finitely many $\lambda \in \sigma_p(A)$ with $|\lambda| \ge \varepsilon$.

(ii) $\sigma_p(A)$ is at most countable and has no accumulation points except perhaps 0, which is an accumulation point whenever $\sigma(A)$ is infinite.

Proof. We follow [102]. (i) For $\varepsilon > 0$ we define $\Sigma_\varepsilon = \{\lambda \in \sigma(A) \mid |\lambda| \ge \varepsilon\} = \sigma(A)\backslash B(0, \varepsilon)$, which is compact. We must show that all Σ_ε are finite. Assume this is not the case for some $\varepsilon > 0$. By compactness, Σ_ε then has an accumulation point. I.e. there is a $\lambda \in \Sigma_\varepsilon$ and a sequence $\{\lambda_n\}$ of mutually distinct elements of Σ_ε converging to λ. Thus we obtain a contradiction if we prove that every sequence of distinct non-zero eigenvalues must converge to zero.

Let thus $\{\lambda_n\}_{n\in\mathbb{N}} \subset \Sigma_\varepsilon$ be mutually distinct non-zero eigenvalues of A. Pick corresponding eigenvectors $\{x_n\}$ and put $W_n = \mathrm{span}_{\mathbb{F}}(x_1, \dots, x_n)$. Since we know from linear algebra that the sets $\{x_1, \dots, x_n\}$ all are linearly independent, we have $W_n \subsetneqq W_{n+1}$ for all n. Thus by Riesz' Lemma 12.2 there are unit vectors $y_{n+1} \in W_{n+1}$ such that $\mathrm{dist}(y_{n+1}, W_n) \geq 1/2$. Since $(A - \lambda_{n+1}\mathbf{1})$ kills x_{n+1}, we have $(A - \lambda_{n+1}\mathbf{1})(y_{n+1}) \in W_n$. Now for all $j > k > 1$ we have

$$\begin{aligned} A(\lambda_j^{-1}y_j) - A(\lambda_k^{-1}y_k) &= \lambda_j^{-1}(A - \lambda_j\mathbf{1})(y_j) - \lambda_k^{-1}(A - \lambda_k\mathbf{1})(y_k) + y_j - y_k \\ &= y_j - [-\lambda_j^{-1}(A - \lambda_j\mathbf{1})(y_j) + \lambda_k^{-1}(A - \lambda_k\mathbf{1})(y_k) + y_k]. \end{aligned}$$

Since the expression in square brackets lies in W_{j-1}, which has distance $\geq 1/2$ from y_j, we have $\|A(\lambda_j^{-1}y_j) - A(\lambda_k^{-1}y_k)\| \geq 1/2$, which clearly holds for all $j \neq k$. Thus the sequence $\{A(\lambda_j^{-1}y_j)\}$ has no convergent subsequence. Since A is compact, this proves that $\{\lambda_j^{-1}y_j\}$ has no bounded subsequence. With $\|y_j\| = 1 \ \forall j$, this proves $\lambda_j \to 0$, as desired.

(ii) This immediately follows from (i) in the same way as for compact normal Hilbert space operators. ■

By the above (which also holds over $\mathbb{R}$), every $\lambda \in \sigma(A)\backslash\{0\}$ is isolated. Restricting to $\mathbb{F} = \mathbb{C}$, Exercises 13.53, 13.66, 13.68 provide a Riesz idempotent $P_\lambda \in B(V)$ commuting with A and such that $\sigma(A \upharpoonright P_\lambda V) = \{\lambda\}$ and $\sigma(A \upharpoonright (\mathbf{1} - P_\lambda)V) = \sigma(A)\backslash\{\lambda\}$. Thus $V_\lambda = P_\lambda V$ is a closed A-invariant subspace. If $\lambda, \lambda' \in \sigma(A)\backslash\{0\}$ are distinct, P_λ and $P_{\lambda'}$ commute since both are limits of inverses of polynomials in A. It is easy to see that $P_\lambda P_{\lambda'} = P_{\lambda'} P_\lambda = 0$.

14.18 Proposition *Let V be a complex Banach space, $A \in B(V)$ compact and $\lambda \in \sigma(A)\backslash\{0\}$. Then*

(i) the generalized eigenspace $\bigcup_{n=1}^\infty \ker(A - \lambda\mathbf{1})^n$ of λ coincides with $V_\lambda = P_\lambda V$ and

(ii) is finite-dimensional,

(iii) $(A - \lambda\mathbf{1}) \upharpoonright V_\lambda$ is nilpotent.

Proof. Since V_λ is invariant under A, the restriction $A \upharpoonright V_\lambda$ is compact. By construction of P_λ, we have $\sigma(A \upharpoonright V_\lambda) = \{\lambda\}$. Since $\lambda \neq 0$, we have $0 \notin \sigma(A \upharpoonright V_\lambda)$ so that $A \upharpoonright V_\lambda$ is invertible. Thus $A \upharpoonright V_\lambda$ is compact and invertible, implying that V_λ is finite-dimensional. Since $\sigma(A \upharpoonright V_\lambda) = \{\lambda\}$, $(A - \lambda\mathbf{1}) \upharpoonright V_\lambda$ is quasi-nilpotent, thus nilpotent by Exercise 13.60. Thus for every $x \in V_\lambda$ we have $(A - \lambda\mathbf{1})^n x = 0$ for some n, proving $V_\lambda \subseteq \bigcup_{n=1}^\infty \ker(A - \lambda\mathbf{1})^n$. On the other hand, the fact that $A - \lambda\mathbf{1}$ is invertible on $(\mathbf{1} - P_\lambda)V$ implies the converse inclusion $\bigcup_{n=1}^\infty \ker(A - \lambda\mathbf{1})^n \subseteq V_\lambda$. This concludes the proof. ■

14.19 Remark An alternative proof for the finite-dimensionality of the generalized eigenspace (but not its coincidence with $V_\lambda = P_\lambda V$) proceeds as follows:

Let $B \in B(V)$ be arbitrary and $n \in \mathbb{N}$. If $x \in \ker B^{n+1}$ then $Bx \in \ker B^n$, so that B restricts to a linear map $\ker B^{n+1} \to \ker B^n$. And $Bx \in \ker B^{n-1} \subseteq \ker B^n$ if and only if $x \in \ker B^n$. Thus B induces a linear map $\ker B^{n+1} / \ker B^n \to \ker B^n / \ker B^{n-1}$, so that $\dim\ker B^{n+1} - \dim\ker B^n \leq \dim\ker B^n - \dim\ker B^{n-1}$. (For $n = 1$ this is $\dim\ker B^2 - \dim\ker B \leq \dim\ker B$.) Now by induction (or a telescoping sum) we have $\dim\ker B^n \leq n \dim\ker B$.

Putting $B = A - \lambda\mathbf{1}$, by the proof of Proposition 14.2 there is a d such that $\ker B^{n+1} = \ker B^n$ for all $n \geq d$. Thus $\dim \bigcup_{n=1}^\infty \ker B^n = \dim\ker B^d \leq d \dim\ker B < \infty$, where we used Lemma 14.1. □

Now we can proceed as known from finite-dimensional linear algebra and find for each $\lambda \in \sigma(A)\backslash\{0\}$ a basis for V_λ with respect to which A is given by a block diagonal matrix with a finite number of standard Jordan blocks, the number of these blocks equaling the geometric multiplicity $\dim \ker(A - \lambda\mathbf{1})$ of λ. We refer to the literature, cf. e.g. [55, 84, 95].

If $\sigma(A)$ is finite, also 0 is an isolated point of $\sigma(A)$ (if it is in it), so that we have a Riesz projector P_0. Now we have an isomorphism $V \simeq \bigoplus_{\lambda \in \sigma(A)} V_\lambda$. Note that V_0 need not be finite-dimensional, since there is an ample supply of compact quasi-nilpotent operators in infinite dimensions, e.g. the classical Volterra operator and the weighted shift operators with weight function decreasing fast enough. An attempt at classifying them would lead us too far.

If $\sigma(A)$ is infinite, matters are more complicated. At least for every $\varepsilon > 0$ we have an isomorphism $V \simeq \bigoplus_{\lambda \in \sigma(A), |\lambda| > \varepsilon} V_\lambda \bigoplus V_{\leq \varepsilon}$, where $V_{\leq \varepsilon}$ is a closed A-invariant subspace with $\sigma(A \upharpoonright V_{\leq \varepsilon}) \subseteq \overline{B}(0, \varepsilon)$. In this case $0 \in \sigma(A)$ is not isolated, so that we cannot define a Riesz projector, but we could put $V_0 = \bigcap_{\varepsilon > 0} V_{\leq \varepsilon}$. Now V_0 can be zero (as for $A \in K(\ell^2)$ defined by $(Af)(n) = f(n)/n$) or non-empty, in which case $A \upharpoonright V_0$ again is compact quasi-nilpotent.

For more on spectral theory and normal forms of compact operators see e.g. [24, 135].

Our last major target in this course is proving Theorem 18.4, an analogue of Theorem 14.12 for Hilbert space operators $A \in B(H)$ that are normal but not necessarily compact. This will require extensive preparations, but the mathematics needed is itself a central part of modern functional analysis.

15 Some functional calculus for Banach algebras

15.1 Characters. Spectrum of a Banach algebra

We now develop a new perspective on the spectrum that will prove very powerful, allowing to obtain results that would be hard to reach in other ways. For example: If $\mathcal{A}$ is a unital Banach algebra and $a, b \in \mathcal{A}$. What can we say about $\sigma(a+b)$ or $\sigma(ab)$? Using only the definition of the spectrum this seems quite difficult. In this section we require $\|\mathbf{1}\| = 1$.

15.1 Definition *If $\mathcal{A}, \mathcal{B}$ are $\mathbb{F}$-algebras, an (algebra) homomorphism $\alpha : \mathcal{A} \to \mathcal{B}$ is an $\mathbb{F}$-linear map such that also $\alpha(aa') = \alpha(a)\alpha(a')$ $\forall a, a' \in \mathcal{A}$. If $\mathcal{A}, \mathcal{B}$ are unital, α is called unital if $\alpha(\mathbf{1}_\mathcal{A}) = \mathbf{1}_\mathcal{B}$. Algebra homomorphisms from an $\mathbb{F}$-algebra to $\mathbb{F}$ are called characters. An algebra isomorphism is a bijective algebra homomorphism.*

15.2 Lemma *If $\mathcal{A}, \mathcal{B}$ are unital algebras and $\alpha : \mathcal{A} \to \mathcal{B}$ is a unital algebra homomorphism then $\sigma_\mathcal{B}(\alpha(a)) \subseteq \sigma_\mathcal{A}(a)$ $\forall a \in \mathcal{A}$.*

Proof. If $\lambda \notin \sigma_\mathcal{A}(a)$ then $a - \lambda\mathbf{1}_\mathcal{A} \in \mathcal{A}$ has an inverse b. Then $\alpha(b)$ is an inverse for $\alpha(a - \lambda\mathbf{1}_\mathcal{A}) = \alpha(a) - \lambda\mathbf{1}_\mathcal{B} \in \mathcal{B}$, thus $\lambda \notin \sigma_\mathcal{B}(\alpha(a))$. ■

15.3 Lemma *Let $\mathcal{A}$ be a unital Banach algebra. Then every non-zero character $\varphi : \mathcal{A} \to \mathbb{F}$ satisfies $\varphi(\mathbf{1}) = 1$, $\varphi(a) \in \sigma(a)$ $\forall a \in \mathcal{A}$ and $\|\varphi\| = 1$, thus φ is continuous.*

Proof. Since $\varphi \neq 0$ we can find $a \in A$ with $\varphi(a) \neq 0$. Now $\varphi(a) = \varphi(a\mathbf{1}) = \varphi(a)\varphi(\mathbf{1})$, and dividing by $\varphi(a)$ gives $\varphi(\mathbf{1}) = 1$. Thus every non-zero character is a unital homomorphism, so that Lemma 15.2 gives $\sigma_\mathbb{F}(\varphi(a)) \subseteq \sigma_\mathcal{A}(a)$. With $\sigma_\mathbb{F}(x) = \{x\}$ we have $\varphi(a) \in \sigma(a)$, thus $|\varphi(a)| \leq r(a) \leq \|a\|$ by Proposition 13.27, whence $\|\varphi\| \leq 1$. Since we require $\|\mathbf{1}\| = 1$, we also have $\|\varphi\| \geq |\varphi(\mathbf{1})|/\|\mathbf{1}\| = 1$. ■

15.4 Definition *If $\mathcal{A}$ is a unital Banach algebra, the spectrum $\Omega(\mathcal{A})$ of $\mathcal{A}$ is the set of non-zero characters $\varphi : \mathcal{A} \to \mathbb{F}$.*

15.5 Exercise Let X be a compact Hausdorff space and $\mathcal{A} = C(X, \mathbb{F})$. For every $x \in X$ define $\varphi_x : \mathcal{A} \to \mathbb{F},\ f \mapsto f(x)$. Prove:

(i) φ_x is a non-zero character of $\mathcal{A}$, thus $\varphi_x \in \Omega(\mathcal{A})$, for each $x \in X$.

(ii) The map $\iota : X \to \Omega(\mathcal{A}),\ x \mapsto \varphi_x$ is injective.

(iii) For each $f \in \mathcal{A}$ we have $\sigma(f) = \{\varphi(f) \mid \varphi \in \Omega(\mathcal{A})\}$. Do not use Proposition 15.7!

15.6 Remark Since characters are bounded, we have $\Omega(\mathcal{A}) \subseteq \mathcal{A}^*$. Thus every topology of $\mathcal{A}^*$ restricts to a topology on $\Omega(\mathcal{A})$. It will turn out that the 'right' one is the weak-$*$ topology from Section 10.3, for example since it makes the map $\iota : X \to \Omega(\mathcal{A})$ in the preceding exercise a homeomorphism. But we defer this discussion to Section 19.1. □

One could hope that $\sigma(a) = \{\varphi(a) \mid \varphi \in \Omega(\mathcal{A})\}$ holds for every unital Banach algebra $\mathcal{A}$ and $a \in \mathcal{A}$. While the inclusion $\supseteq$ always holds, for equality one needs more:

15.7 Proposition *Let $\mathcal{A}$ be a <u>commutative</u> unital Banach algebra <u>over $\mathbb{C}$</u>. Then*

(i) If $\varphi \in \Omega(\mathcal{A})$ then $\ker\varphi \subseteq \mathcal{A}$ is a maximal ideal (i.e. not contained in a larger proper ideal).

(ii) Every maximal ideal in $\mathcal{A}$ is the kernel of a unique $\varphi \in \Omega(\mathcal{A})$. In particular, $\Omega(\mathcal{A}) \neq \emptyset$.[92]

(iii) For each $a \in \mathcal{A}$ we have

$$\sigma(a) = \{\varphi(a) \mid \varphi \in \Omega(\mathcal{A})\}. \tag{15.1}$$

Proof. (i) It should be clear that $M = \ker\varphi$ is an ideal, and $M \neq \mathcal{A}$ since $\varphi \neq 0$. This ideal has codimension one since $\mathcal{A}/M \cong \varphi(\mathcal{A}) = \mathbb{C}$ and therefore is maximal.

(ii) Now let $M \subseteq \mathcal{A}$ be a maximal ideal. Since maximal ideals are proper, no element of M is invertible. If $b \in M$ satisfied $\|\mathbf{1} - b\| < 1$ then Lemma 13.19(i) would give invertibility of $b = \mathbf{1} - (\mathbf{1} - b)$, a contradiction. (This is the only place where completeness is used.) We thus have $\|\mathbf{1} - b\| \geq 1$ for all $b \in M$, implying $\mathbf{1} \notin \overline{M}$. Thus $\overline{M}$ is a proper ideal containing M. Since M is maximal, we have $\overline{M} = M$, thus M is closed. Now by Proposition 6.1(vi), $\mathcal{A}/M$ is a normed algebra, and by a well-known argument from commutative algebra, the maximality of M implies that $\mathcal{A}/M$ is a field, thus a division algebra. Thus $\mathcal{A}/M \cong \mathbb{C}$ by the Gelfand-Mazur theorem (Corollary 13.43), so that there is a unique isomorphism $\alpha : \mathcal{A}/M \to \mathbb{C}$ sending $\mathbf{1} \in \mathcal{A}/M$ to $1 \in \mathbb{C}$. If $p : \mathcal{A} \to \mathcal{A}/M$ is the quotient homomorphism then $\varphi = \alpha \circ p : \mathcal{A} \to \mathbb{C}$ is a non-zero character with $\ker\varphi = M$. This φ clearly is unique. Now $\Omega(\mathcal{A}) \neq \emptyset$ follows from the fact that every commutative unital algebra has maximal ideals (by a standard Zorn argument).

(iii) We already know that $\{\varphi(a) \mid \varphi \in \Omega(\mathcal{A})\} \subseteq \sigma(a)$, so that it remains to prove that for every $\lambda \in \sigma(a)$ there is a $\varphi \in \Omega(\mathcal{A})$ such that $\varphi(a) = \lambda$. If $\lambda \in \sigma(a)$ then $a - \lambda\mathbf{1} \notin \operatorname{Inv}\mathcal{A}$. Thus the ideal $I = (a - \lambda\mathbf{1})\mathcal{A} \subseteq \mathcal{A}$ does not contain $\mathbf{1}$ and therefore is proper. Using Zorn's lemma, we can find a maximal ideal $M \supseteq I$. By (ii) there is a $\varphi \in \Omega(\mathcal{A})$ such that $\ker\varphi = M$. Since $a - \lambda\mathbf{1} \in I \subseteq M = \ker\varphi$, we have $\varphi(a - \lambda\mathbf{1}) = 0$, and with $\varphi(\mathbf{1}) = 1$ we have $\varphi(a) = \lambda$. ■

[92] If $\mathcal{A}$ is an algebra over a field k, one must take care to distinguish between ring ideals $I \subseteq \mathcal{A}$ and algebra ideals. Both are closed under addition and under multiplication by elements of $\mathcal{A}$. Algebra ideals are linear subspaces, thus also closed under multiplication by the scalars in k. Every algebra ideal clearly is a ring ideal, and the converse holds if $\mathcal{A}$ has a unit (as is assumed here) since $cx = (c\mathbf{1})x \in I$ for each $c \in k$ and $x \in I$. But a non-unital algebra can have ring ideals that are not algebra ideals.

15.8 Exercise Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$ and $a, b \in \mathcal{A}$.

(i) If $\mathcal{A}$ is abelian, prove $\sigma(a+b) \subseteq \sigma(a)+\sigma(b)$ and $\sigma(ab) \subseteq \sigma(a)\sigma(b)$. Conclude that $r(a+b) \le r(a)+r(b)$ and $r(ab) \le r(a)r(b)$. (No use of Exercise 13.46!)

(ii) Prove that the results of (i) also hold for non-abelian $\mathcal{A}$ provided $ab = ba$.

(iii) Give examples of non-commuting $a, b \in \mathcal{A} = M_{2\times 2}(\mathbb{C})$ for which everything in (i) fails.

Hint: For (ii), use an abelian subalgebra and Exercise 13.50.

15.9 Exercise Prove by way of counterexamples that the statements of Proposition 15.7(ii)+(iii) all can fail if we drop the commutativity assumption or replace $\mathbb{C}$ by $\mathbb{R}$.

15.10 Exercise Let $\mathcal{A}$ be a Banach algebra over $\mathbb{F}$ and $\widetilde{\mathcal{A}}$ its unitization (Exercise 13.55). Define $\varphi_\infty : \widetilde{\mathcal{A}} \to \mathbb{F}$, $(a, \alpha) \mapsto \alpha$.

(i) Prove $\varphi_\infty \in \Omega(\widetilde{\mathcal{A}})$.

(ii) Prove that every $\varphi \in \Omega(\mathcal{A})$ has a unique extension to $\widehat{\varphi} \in \Omega(\widetilde{\mathcal{A}})$.

(iii) Prove $\Omega(\widetilde{\mathcal{A}}) = \{\widehat{\varphi} \mid \varphi \in \Omega(\mathcal{A})\} \cup \{\varphi_\infty\}$.

(iv) If $\mathcal{A}$ is non-unital and $a \in \mathcal{A}$, define $\sigma(a)$ as $\sigma_{\widetilde{\mathcal{A}}}(a)$. For $\mathcal{A}$ commutative non-unital over $\mathbb{C}$ and $a \in \mathcal{A}$, prove

$$\sigma(a) = \{\varphi(a) \mid \varphi \in \Omega(\mathcal{A})\} \cup \{0\}.$$

15.11 Exercise Consider the commutative unital Banach algebra $\mathcal{A} = \ell^1(\mathbb{Z}, \mathbb{C})$ with convolution product $\star$ and unit $\mathbf{1} = \delta_0$. Prove:

(i) For every $\varphi \in \Omega(\mathcal{A})$ we have $\varphi(\delta_1) \in S^1$.

(ii) If $\varphi_1, \varphi_2 \in \Omega(\mathcal{A})$ satisfy $\varphi_1(\delta_1) = \varphi_2(\delta_1)$ then $\varphi_1 = \varphi_2$.

(iii) For every $z \in S^1$ prove that $\varphi_z(f) = \sum_{n\in\mathbb{Z}} f(n) z^n$ defines an element of $\Omega(\mathcal{A})$.

(iv) The map $S^1 \to \Omega(\mathcal{A})$, $z \mapsto \varphi_z$ is a bijection.

(v) For every $f \in \mathcal{A}$ we have

$$\sigma(f) = \Big\{ \sum_{n\in\mathbb{Z}} f(n) z^n \;\Big|\; z \in S^1 \Big\}.$$

(vi) If $f \in \mathcal{A}$ is quasi-nilpotent then $f = 0$. Hint: f can be recovered from $\sum_n f(n) z^n$.

Note how difficult it would be to prove the result of (v) using only the definition of the spectrum! But Fourier analysis provides an instructive perspective on the result, cf. Section 19.2.

15.12 Exercise Let $f \in \ell^1 = \ell^1(\mathbb{Z}, \mathbb{C})$. With $f_m(n) = f(n-m)$, prove that $\mathrm{span}_{\mathbb{C}}\{f_m \mid m \in \mathbb{Z}\}$ (the finite linear combinations of the translates f_m of f) is dense in $\ell^1(\mathbb{Z}, \mathbb{C})$ if and only if $\widehat{f}(z) = \sum_{n\in\mathbb{Z}} f(n) z^m$ vanishes for no $z \in S^1$.

Hint: Use the closed ideal $I_f = \overline{f\ell^1} \subseteq \ell^1$ generated by f and Exercise 15.11.

15.13 Remark The result of Exercise 15.12 is the simplest of a whole family of 'span of translates' results. These were initiated by a theorem of N. Wiener[93]: Given $f \in L^1(\mathbb{R}, \lambda)$ (λ is Lebesgue measure), the linear span of the translates of f is dense in $L^1(\mathbb{R}, \lambda)$ if and only if the Fourier transform $\widehat{f}(\xi) = \int f(x) e^{-i\xi x} dx$ (which is a continuous function $\mathbb{R} \to \mathbb{C}$) vanishes for no $\xi \in \mathbb{R}$. Already this is harder to prove. Cf. e.g. [141, Theorem 9.5] or [27, Chapter 2]. □

[93] Norbert Wiener (1894-1964). American mathematician with important contributions to harmonic and functional analysis and many other fields. See Theorem 19.9 for a related result of his.

15.2 Baby version of holomorphic functional calculus

Functional calculus is concerned with defining $f(a)$ when f is a (suitable) function and a is an element of a Banach or C^*-algebra or $B(H)$. So far, we have only considered the rather trivial case $f_1 : x \mapsto 1/x$ (for invertible elements of any Banach algebra) and $f_2 : z \mapsto z^{1/2}$ (for positive Hilbert space operators) The next question is: Determine $\sigma(f(a))$. Does it equal $f(\sigma(a))$? (For f_1 it does by Exercise 13.30.) These are the basic questions addressed by the many different 'functional calculi' that there are: holomorphic, continuous, Borel, etc.

While functional calculus has many applications, our main one will be the proof of spectral theorems for normal Hilbert space operators (not necessarily compact) in Section 18.

Defining $f(a)$ poses no problem in the simplest case, which surely is $f = P$, a polynomial:

15.14 Definition *If $\mathcal{A}$ is a unital algebra, $a \in \mathcal{A}$ and $P(x) = c_n x^n + \cdots + c_1 x + c_0$ is a polynomial, we put $P(a) = c_n a^n + \cdots + c_1 a + c_0 \mathbf{1}$.*

15.15 Exercise (Polynomial functional calculus) Let $\mathcal{A}$ be a unital normed algebra over $\mathbb{C}$ and $P \in \mathbb{C}[z]$ with $n = \deg P$.

- (i) Prove that the map $\mathbb{C}[x] \to \mathcal{A}, P \mapsto P(a)$ is a homomorphism of unital $\mathbb{C}$-algebras.
- (ii) Prove $\sigma(P(a)) = P(\sigma(a)) := \{P(\lambda) \mid \lambda \in \sigma(a)\}$.
  Hint: Consider the case $n = 0$ separately. For $n \geq 1$ use a factorization $P(z) - \lambda = c_n \prod_{k=1}^n (z - z_k)$ (which exists by algebraic completeness of $\mathbb{C}$).
- (iii) Why did we assume $\mathcal{A}$ to be normed?

The result of Exercise 15.15 can be generalized in various directions, requiring different proofs. The following suffices for our purposes:

15.16 Proposition *Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$ and let $f(z) = \sum_{n=0}^\infty c_n z^n$ be a power series with convergence radius $R > 0$. Then for all $a \in \mathcal{A}$ satisfying $\|a\| < R$ we have:*

- *(i) The series $f(a) = \sum_{n=0}^\infty c_n a^n$ in $\mathcal{A}$ converges absolutely.*
- *(ii) $\sigma(f(a)) = f(\sigma(a))$. [Spectral mapping theorem]*

Proof. (i) We have $\sum_{n=0}^\infty \|c_n a^n\| \leq \sum_{n=0}^\infty |c_n| \, \|a\|^n$, which converges since $\|a\| < R$ and the power series $\sum c_n z^n$ and $\sum |c_n| z^n$ have the same convergence radius R (as follows from $R^{-1} = \limsup_n |c_n|^{1/n}$). Now use Proposition 3.15(ii).

(ii) Let $\mathcal{B} = \{a\}'' \subseteq \mathcal{A}$. This is a commutative unital Banach algebra with $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$, and $f(a) \in \mathcal{B}$. Since every $\varphi \in \Omega(\mathcal{B})$ is continuous and a unital homomorphism, we have

$$\begin{aligned}\varphi(f(a)) &= \varphi\left(\lim_{N\to\infty} \sum_{n=0}^N c_n a^n\right) = \lim_{N\to\infty} \varphi\left(\sum_{n=0}^N c_n a^n\right) \\ &= \lim_{N\to\infty} \sum_{n=0}^N c_n \varphi(a)^n = \sum_{n=0}^\infty c_n \varphi(a)^n = f(\varphi(a)).\end{aligned}$$

(Note that $\sum_{n=0}^\infty c_n \varphi(a)^n$ converges absolutely since $|\varphi(a)| \leq r(a) \leq \|a\| < R$.)

Applying (15.1) to $f(a) \in \mathcal{B}$, we have

$$\sigma_\mathcal{B}(f(a)) = \{\varphi(f(a)) \mid \varphi \in \Omega(\mathcal{B})\} = \{f(\varphi(a)) \mid \varphi \in \Omega(\mathcal{B})\} = \{f(\lambda) \mid \lambda \in \sigma_\mathcal{B}(a)\} = f(\sigma_\mathcal{B}(a)).$$

Since Exercise 13.50(i) gives $\sigma_\mathcal{B}(b) = \sigma_\mathcal{A}(b)$ for all $b \in \mathcal{B}$, we have $\sigma_\mathcal{A}(f(a)) = \sigma_\mathcal{B}(f(a))$. ■

15.17 Exercise Let $\mathcal{A}$ be a unital Banach algebra and $a \in \mathcal{A}$. Let $\sum_{n=0}^\infty c_n z^n$ be a power series with convergence radius $R > 0$. Prove that for the absolute convergence of $\sum_{n=0}^\infty c_n a^n$ it suffices that $r(a) < R$.

For $P \in \mathbb{C}[x]$, continuity of the map $a \mapsto P(a)$ is evident in every topological algebra. The analogous result for a power series requires more work:

15.18 Exercise Let $\mathcal{A}$ be a unital Banach algebra.

(i) Prove $\|a^n - b^n\| \le n\|a - b\|(\max(\|a\|, \|b\|))^{n-1}$ for all $a, b \in \mathcal{A}, n \in \mathbb{N}$. Hint: telescope.

(ii) Let $f(z) = \sum_{n=0}^\infty c_n z^n$ have convergence radius $R > 0$. Prove that the map $a \mapsto f(a)$ is uniformly continuous on $\overline{B}_{\mathcal{A}}(0, r)$ for each $r < R$ and continuous on $B_{\mathcal{A}}(0, R)$.

If one defines $\mathcal{H}_R$ to be the set of functions defined by power series with convergence radius $\ge R$ then $\mathcal{H}_R$ is easily checked to be a commutative algebra (which coincides with the algebra of functions holomorphic on $B(0, R)$). Now for every $a \in \mathcal{A}$ with $\|a\| < R$ (or just $r(a) < R$) one has a unital homomorphism $\mathcal{H}_R \to \mathcal{A},\ f \mapsto f(a)$. This can be generalized considerably, leading to the fully fledged holomorphic functional calculus, cf. e.g. [79, 94, 152], which we do not discuss since it is insufficient for our purposes: We want to make sense of $f(A)$ when $A \in B(H)$ is normal and $f : \sigma(A) \to \mathbb{C}$ is just continuous. This will be done in Section 17.

15.19 Example (Exponential function) Of course every power series $f(z) = \sum_{n=0}^\infty c_n z^n$ with infinite convergence radius can be 'applied' to every $a \in \mathcal{A}$. For example $\exp(a) = e^a = \sum_{n=0}^\infty \frac{a^n}{n!}$ converges for every $a \in \mathcal{A}$. By the spectral mapping theorem we have $\sigma(e^a) = \exp(\sigma(a))$.

15.20 Exercise Let $\mathcal{A}$ be a unital Banach algebra. Prove $\exp(a) = \lim_{n\to\infty} \left(\mathbf{1} + \frac{a}{n}\right)^n$ for all $a \in \mathcal{A}$.

15.21 Exercise Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $a \in \mathcal{A}$.

(i) Give an elementary proof of $\{e^\lambda \mid \lambda \in \sigma(a)\} \subseteq \sigma(e^a)$ (i.e. without using the spectral mapping theorem).

(ii) If $\mathbb{F} = \mathbb{C}$ and $\|e^{ita}\| = 1$ for all $t \in \mathbb{R}$, prove $\sigma(a) \subseteq \mathbb{R}$.

(iii) Show by example that $\sigma(a) \subseteq \mathbb{R}$ does not imply $\|e^{ita}\| = 1\ \forall t \in \mathbb{R}$.

15.22 Exercise (One parameter groups I) Let $\mathcal{A}$ be a Banach algebra with unit $\mathbf{1}$ over $\mathbb{R}$ or $\mathbb{C}$. Let $\exp : \mathcal{A} \to \mathcal{A}$ be defined as above. Prove:

(i) $\|e^a - \mathbf{1}\| \le e^{\|a\|} - 1\ \forall a \in \mathcal{A}$.

(ii) If $ab = ba$ then $e^{a+b} = e^a e^b = e^b e^a$.

(iii) The map $\mathbb{R} \to \mathcal{A},\ t \mapsto W(t) = e^{ta}$ is a one-parameter group (thus $W(0) = \mathbf{1}$ and $W(s+t) = W(s)W(t)\ \forall s, t \in \mathbb{R}$).

(iv) $t \mapsto e^{ta}$ is norm-continuous. Do this using (1a) and (1c), not Exercise 15.18.

(v) $\lim_{t\to 0} \frac{e^{ta} - \mathbf{1}}{t} = a$.

(vi) If $a, b \in \mathcal{A}$ and $e^{ta}e^{sb} = e^{sb}e^{ta}$ for all $s, t \in \mathbb{R}$ then $ab = ba$.

15.23 Remark One can find 2×2 matrices a, b such that $e^{a+b} = e^a e^b = a^b e^a$, but $ab \neq ba$, as well as non-commuting 2×2 matrices a, b such that $e^{a+b} = e^a e^b \neq e^b e^a$. There is an extensive literature on this phenomenon. We mention one interesting result [150]: If $\mathcal{A}$ is a unital Banach algebra, $a, b \in \mathcal{A}$ such that $e^a e^b = e^b e^a$ and $\sigma(a)$ and $\sigma(a)$ are $2\pi i$-congruence free then $ab = ba$. Here $\Omega \subseteq \mathbb{C}$ is called $2\pi i$-congruence free if $\lambda, \lambda' \in \Omega$, $\lambda - \lambda' \in 2\pi i\mathbb{Z}$ implies $\lambda = \lambda'$. (In particular $e^a e^b = e^b e^a$ does imply $ab = ba$ if $\sigma(a), \sigma(b) \subset \mathbb{R}$.) □

15.24 Exercise (One parameter groups II) Let $\mathcal{A}$ be a unital Banach algebra.

(i) Local inverse for exp.

(a) The logarithm function (more precisely the branch for which $z > 0 \Rightarrow \log z \in \mathbb{R}$) has a unique power series expansion $g(z) = \sum_{n=1}^{\infty} c_n (z-1)^n$ around $z = 1$. Prove that it has convergence radius one.

(b) For $a \in \mathcal{A}$ with $\|a - \mathbf{1}\| < 1$ we can define $\log(a) \in \mathcal{A}$ using the power series g. Prove: If $a, b \in \mathcal{A}$ commute and $\|a - \mathbf{1}\| < 1$, $\|b - \mathbf{1}\| < 1$ then $\log a$ and $\log b$ commute.

(c) Prove: If $\|a\| < \log 2$ then $\|e^a - \mathbf{1}\| < 1$ and $\log(e^a) = a$. And if $\|b - \mathbf{1}\| < 1$ then $\exp(\log b) = b$.

(d) Let $0 \in U = B(0, \frac{\log 2}{2}) \subset \mathcal{A}$ and $V = \exp(U) \ni \mathbf{1}$. Prove that $\exp : U \to V$ is a homeomorphism with inverse $\log : V \to U$.

(e) Prove that if $a, b \in V$ commute and $ab \in V$ then $\log(ab) = \log a + \log b$.

(ii) Let V be a Banach space, $\varepsilon > 0$ and $f : (-\varepsilon, \varepsilon) \to V$ continuous and satisfying $f(0) = 0$ and $f(s+t) = f(s) + f(t)$ whenever $s, t, s+t \in (-\varepsilon, \varepsilon)$. Prove that there exists a unique $x \in V$ such that $f(t) = tx \ \forall t \in (-\varepsilon, \varepsilon)$.

(iii) Now let $\mathbb{R} \to \mathcal{A}, t \mapsto W(t)$ be a norm-continuous one parameter group. Use the above results to prove that there is a unique $a \in \mathcal{A}$ such that $W(t) = e^{ta}$ for all $t \in \mathbb{R}$.

15.25 Remark For most applications of one-parameter groups, norm-continuity is too strong a requirement. To get further one considers one-parameter groups (or semigroups, defined only for $t \geq 0$) in $B(V)$ that are only strongly continuous, i.e. $\lim_{t \to 0} W(t)x = x \ \forall x \in V$. A typical result then is Stone's theorem according to which the unitary one-parameter groups on Hilbert spaces are of the form $W(t) = e^{itA}$, where A is a possibly unbounded self-adjoint operator, cf. e.g. [128]. The subject of operator semigroups is huge, cf. [50] for an introduction. □

15.26 Exercise Let $\mathcal{A}$ be a commutative unital Banach algebra and $a_1, \ldots, a_n \in \mathcal{A}$.

(i) Define the joint spectrum

$$\sigma(a_1, \ldots, a_n) = \{(\varphi(a_1), \ldots, \varphi(a_n)) \mid \varphi \in \Omega(\mathcal{A})\} \subseteq \mathbb{C}^n.$$

Prove $\sigma(a_1, \ldots, a_n) \subseteq \sigma(a_1) \times \cdots \times \sigma(a_n)$.

(ii) Prove $(\lambda_1, \ldots, \lambda_n) \notin \sigma(a_1, \ldots, a_n) \quad \Leftrightarrow \quad \exists b_1, \ldots, b_n \in \mathcal{A} : \sum_{i=1}^n b_i (a_i - \lambda_i \mathbf{1}) = \mathbf{1}$.
Hint: For $\lambda \in \mathbb{C}^n$, use the ideal $I_\lambda = \sum_{i=1}^n \mathcal{A}(a_i - \lambda_i \mathbf{1})$ generated by the $a_i - \lambda_i \mathbf{1}$.

(iii) Let $R_1, \ldots, R_n > 0$ and assume that $\sum_{j \in \mathbb{N}_0^n} c_{j_1, \ldots, j_n} z_1^{j_1} \cdots z_n^{j_n}$ converges whenever $|z_i| < R_i \ \forall i$, defining an analytic function f on this domain. Assuming $\|a_i\| < R_i \ \forall i$, define $f(a_1, \ldots, a_n) \in \mathcal{A}$ in analogy to the case $n = 1$ above. Prove

$$\sigma(f(a_1, \ldots, a_n)) = f(\sigma(a_1, \ldots, a_n)) = \{f(\lambda_1, \ldots, \lambda_n) \mid (\lambda_1, \ldots, \lambda_n) \in \sigma(a_1, \ldots, a_n)\}.$$

15.27 Remark 1. Example: $\sigma(a+b) = \{\alpha+\beta \mid (\alpha,\beta) \in \sigma(a,b)\}$, improving on Exercise 15.8.

2. In more general situations, like non-abelian algebras and non-commuting operators, there exist various definitions of the joint spectrum, not necessarily equivalent. □

16 Basics of C^*-algebras

16.1 Involutions. Definition of C^*-algebras

The properties of the adjoint map $A \mapsto A^*$ on $B(H)$ motivate some definitions:

16.1 Definition *Let $\mathcal{A}$ be a $\mathbb{C}$-algebra. A map $* : \mathcal{A} \to \mathcal{A}$ satisfying antilinearity, antimultiplicativity and involutivity, i.e. (i)-(iii) in Lemma 11.8, is called an involution or $*$-operation. An algebra with a chosen $*$-operation is called a $*$-algebra. A $*$-homomorphism $\alpha : \mathcal{A} \to \mathcal{B}$ of $*$-algebras is a homomorphism satisfying $\alpha(a^*) = \alpha(a)^*$ $\forall a \in \mathcal{A}$.*

16.2 Lemma *Let $\mathcal{A}$ be a unital $*$-algebra. Then*

(i) $\mathbf{1}^* = \mathbf{1}$.

(ii) If $a \in \mathcal{A}$ is invertible then a^* is invertible and $(a^*)^{-1} = (a^{-1})^*$.

(iii) $\sigma(a^*) = \sigma(a)^* := \{\overline{\lambda} \mid \lambda \in \sigma(a)\}$.

Proof. (i) $\mathbf{1}^* = \mathbf{1}\mathbf{1}^* = \mathbf{1}^{**}\mathbf{1}^* = (\mathbf{1}\mathbf{1}^*)^* = (\mathbf{1}^*)^* = \mathbf{1}$. The proofs of (ii) and (iii) are identical to those given earlier for Hilbert space operators. ■

16.3 Definition *If $\mathcal{A}$ is a Banach algebra and $* : \mathcal{A} \to \mathcal{A}$ an involution then $\mathcal{A}$ is called a*

- *Banach $*$-algebra if $\|a^*\| = \|a\|$ $\forall a \in \mathcal{A}$.*
- *C^*-algebra if $\|a^*a\| = \|a\|^2$ $\forall a \in \mathcal{A}$.*[94]

16.4 Lemma *Every C^*-algebra is a Banach $*$-algebra. If it has a unit $\mathbf{1}$ then $\|\mathbf{1}\| = 1$.*

Proof. With the C^*-identity and submultiplicativity we have $\|a\|^2 = \|a^*a\| \leq \|a^*\|\|a\|$, thus $\|a\| \leq \|a^*\|$ for all $a \in \mathcal{A}$. Replacing a by a^* herein gives the converse inequality, thus $\|a^*\| = \|a\|$.

If $\mathbf{1}$ is a unit then $\|\mathbf{1}\|^2 = \|\mathbf{1}^*\mathbf{1}\| = \|\mathbf{1}^*\| = \|\mathbf{1}\|$, and since $\|\mathbf{1}\| \neq 0$ this implies $\|\mathbf{1}\| = 1$. ■

16.5 Remark 1. Clearly $B(H)$ is a C^*-algebra for each Hilbert space H. Since this holds also for real Hilbert spaces, it shows that one can discuss Banach $*$-algebras and C^*-algebras over $\mathbb{R}$. But we will consider only complex ones.

2. There is no special name for the non-complete variants of the above definitions. But a submultiplicative norm on a $*$-algebra satisfying the C^*-identity is called a C^*-norm, whether $\mathcal{A}$ is complete w.r.t. it or not. Completion of a $*$-algebra w.r.t. a C^*-norm gives a C^*-algebra, and this is an important way of constructing new C^*-algebras. □

It is easy to find Banach $*$-algebras that are not C^*-algebras:

16.6 Exercise For $n \in \mathbb{N}$, consider $\mathcal{A} = M_{n\times n}(\mathbb{C})$ with the usual $*$-algebra structure. Prove that $\|a\| = \left(\sum_{i,j=1}^n |a_{i,j}|^2\right)^{1/2}$ defines a norm that satisfies submultiplicativity and $\|a^*\| = \|a\|$. Thus $(\mathcal{A}, \|\cdot\|)$ is a Banach $*$-algebra. Prove that it is not a C^*-algebra when $n \geq 2$.

[94] The original definition by Gelfand and Naimark (1942) had the additional axiom that $a^*a + \mathbf{1}$ be invertible for each a. This turned out to be redundant, cf. Proposition 17.6.

16.7 Exercise Recall the Banach algebra $\mathcal{A} = \ell^1(\mathbb{Z}, \mathbb{C})$ from Example 13.48. Show that both $f^*(n) = \overline{f(n)}$ and $f^*(n) = \overline{f(-n)}$ define involutions on $\mathcal{A}$ making it a Banach $*$-algebra. Show that neither of them satisfies the C^*-identity.

16.8 Exercise Equip the Banach algebra from Exercise 13.51 with the $*$-operation $f^*(x) = \overline{f(x)}$. Is it a C^*-algebra?

16.9 Lemma *Let X be a compact space. For $f \in C(X, \mathbb{C})$, define f^* by $f^*(x) = \overline{f(x)}$. Then $C(X, \mathbb{C})$ is a C^*-algebra. The same holds for $C_b(X, \mathbb{C})$, where X is arbitrary, thus also for $\ell^\infty(S, \mathbb{C})$.*

Proof. We know that $C(X, \mathbb{C})$ equipped with the norm $\|f\| = \sup_x |f(x)|$ is a Banach algebra. It is immediate that $*$ is an involution. The computation

$$\|f^*f\| = \sup_x |\overline{f(x)}f(x)| = \sup_x |f(x)|^2 = \big(\sup_x |f(x)|\big)^2 = \|f\|^2$$

proves the C^*-identity. It is clear that this generalizes to the bounded continuous functions on any space X. ■

In a sense, the examples $B(H)$ and $C(X, \mathbb{C})$ for compact X are all there is: One can prove, as we will do in Theorem 19.12, that every commutative unital C^*-algebra is isometrically $*$-isomorphic to $C(X, \mathbb{C})$ for some compact Hausdorff space X, determined uniquely up to homeomorphism. (For example one has $\ell^\infty(S, \mathbb{C}) \cong C(\beta S, \mathbb{C})$, where βS is the Stone-Čech compactification of (S, τ_{disc}).) And one can prove that every C^*-algebra is isometrically $*$-isomorphic to a norm-closed $*$-subalgebra of $B(H)$ for some Hilbert space H. See e.g. [110].

16.10 Exercise If $\mathcal{A}$ is a C^*-algebra and $\widetilde{\mathcal{A}}$ its unitization (Exercise 13.55), the norm $\|(a, \alpha)\|_1 = \|a\| + |\alpha|$ on $\widetilde{\mathcal{A}}$ usually fails to be a C^*-norm. Define $\|(a, \alpha)\| = \sup_{b \in \mathcal{A}, \|b\| \le 1} \|ab + \alpha b\|$. Prove:

(i) $\|\cdot\|$ is an algebra norm on $\widetilde{\mathcal{A}}$ if and only if $\mathcal{A}$ is non-unital, which is assumed from now on.

(ii) It satisfies $\|(a, \alpha)\| \le \|(a, \alpha)\|_1$ and $\|(a, 0)\| = \|a\|$ $\forall a \in \mathcal{A}$, thus $\iota : \mathcal{A} \hookrightarrow \widetilde{\mathcal{A}}$ is an isometry.

(iii) $\|\cdot\|$ is a C^*-norm.

(iv) $(\widetilde{\mathcal{A}}, \|\cdot\|)$ is complete and the norms $\|\cdot\|_1, \|\cdot\|$ on $\widetilde{\mathcal{A}}$ are equivalent.

16.2 Some classes of elements in a C^*-algebra and their spectra

Inspired by $B(H)$ we define:

16.11 Definition *Let $\mathcal{A}$ be a $\mathbb{C}$-algebra with an involution $*$. Then $a \in \mathcal{A}$ is called*

- *self-adjoint if $a = a^*$. We put $\mathcal{A}_{sa} = \{a \in \mathcal{A} \mid a = a^*\}$.*
- *normal if $aa^* = a^*a$. (I.e. a and a^* 'commute'.)*
- *unitary if $aa^* = a^*a = \mathbf{1}$. (Obviously $\mathcal{A}$ needs to be unital.)*
- *orthogonal projection if $a^2 = a = a^*$.*

A subset S of a $*$-algebra is called self-adjoint if $S = S^* := \{s^* \mid s \in S\}$.

16.12 Exercise Let $\mathcal{A}$ be a Banach $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed self-adjoint two-sided ideal. Prove:

(i) $\mathcal{A}/\mathcal{I}$ has a natural $*$-operation so that $p : \mathcal{A} \to \mathcal{A}/\mathcal{I}$ is a $*$-homomorphism.

(ii) With this $*$-operation and the quotient norm, $\mathcal{A}/\mathcal{I}$ is a Banach $*$-algebra.

(iii) If $\alpha : \mathcal{A} \to \mathcal{B}$ is a $*$-homomorphism such that $\mathcal{I} \subseteq \ker \alpha$ then the induced map $\alpha' : \mathcal{A}/\mathcal{I} \to \mathcal{B}$, cf. Proposition 6.1(v), is a $*$-homomorphism.

16.13 Remark If $\mathcal{A}$ is a C^*-algebra one can prove that every closed two sided ideal $\mathcal{I} \subseteq \mathcal{A}$ automatically is self-adjoint and that $\mathcal{A}/\mathcal{I}$ is a C^*-algebra. But the proofs would lead us too far, cf. e.g. [110, Theorems 3.1.3, 3.1.4]. □

The self-adjoint, respectively unitary, elements of the C^*-algebra $\mathcal{A} = \mathbb{C}$ are the real numbers and the phases ($|z| = 1$). Thus self-adjoint and unitary elements of a C^*-algebra should be thought of as generalized real numbers and phases, respectively. Also the real-imaginary decomposition generalizes:

16.14 Lemma *If $\mathcal{A}$ is a $*$-algebra and $a \in \mathcal{A}$, we define* $\mathrm{Re}(a) = \frac{a+a^*}{2}$, $\mathrm{Im}(a) = \frac{a-a^*}{2i}$. *Now*

(i) $\mathrm{Re}(a), \mathrm{Im}(a)$ *are self-adjoint and* $a = \mathrm{Re}(a) + i\,\mathrm{Im}(a)$.

(ii) The representation $a = b + ic$ with b, c self-adjoint is unique for each a.

(iii) a is self-adjoint if and only if $\mathrm{Im}(a) = 0$.

(iv) a is normal if and only if $\mathrm{Re}(a)$ *and* $\mathrm{Im}(a)$ *commute.*

Proof. Mostly trivial computations. We only prove (ii): If b, b', c, c' are self-adjoint and $b + ic = b' + ic'$ then $b - b' = i(c' - c)$. This implies $b - b' = (b - b')^* = -i(c' - c) = -(b - b')$. Thus $b = b'$ and in turn $c = c'$. ■

16.15 Exercise Prove that a $*$-algebra is commutative if and only if every element is normal.

There are several reasons why normal elements are important. An element $a \in \mathcal{A}$ is normal if and only if there is a $*$-closed commutative subalgebra $\mathcal{B} \subseteq \mathcal{A}$ containing a. We will see that normal elements behave like functions on a (locally) compact space.

While $ab = ba$ clearly implies $a^*b^* = b^*a^*$, it need not follow that a^* commutes with b (or equivalently a with b^*). To see this just pick any non-normal $a \in \mathcal{A}$ and take $b = a$. But:

16.16 Theorem (Fuglede 1950) *Let $\mathcal{A}$ be a unital C^*-algebra, and let a, b be commuting elements at least one of which is normal. Then $a^*b = ba^*$ (and $ab^* = b^*a$).*

The theorem is quite remarkable, and asked for a proof one probably wouldn't know where to begin. For matrices it actually is quite easy: Normality of a implies that a is diagonalizable, i.e. $a = \sum_i \lambda_i P_i$, where the λ_i are the (distinct) eigenvalues and the P_i orthogonal projections onto the corresponding eigenspaces, see e.g. [55, Theorem 6.16]. Now $ab = ba$ implies $P_i b = bP_i\ \forall i$. Taking adjoints gives $P_i b^* = b^* P_i\ \forall i$, whence $ab^* = b^*a$. With effort, this argument can be extended to operators on infinite-dimensional Hilbert spaces, cf. [160]. But there is a much more elegant argument that works in all C^*-algebras, for which we refer to Section B.13.1.

16.17 Proposition *Let $\mathcal{A}$ be a unital C^*-algebra. Then*

(i) If $a \in \mathcal{A}$ is normal then $r(a) = \|a\|$.

(ii) If $u \in \mathcal{A}$ is unitary then $\|u\| = 1$ and $\sigma(u) \subseteq S^1$.

(iii) If $a \in \mathcal{A}$ is self-adjoint then $\sigma(a) \subseteq \mathbb{R}$.

Proof. (i) The proof of Proposition 13.69(i) works identically in every abstract C^*-algebra.

(ii) By unitarity of u we have $\|u\|^2 = \|u^*u\| = \|\mathbf{1}\| = 1$, thus $\|u\| = 1 = \|u^*\|$. With $u^{-1} = u^*$ we also have $\|u^{-1}\| = 1$, and Exercise 13.30(ii) gives $\sigma(u) \subseteq S^1$.

(iii) Given $\lambda \in \sigma(a)$, write $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$. Applying $\sigma(a + z\mathbf{1}) = \sigma(a) + z \ \forall z \in \mathbb{C}$ (why?) to $z = -\alpha + in\beta$, where $n \in \mathbb{N}$, we have $i\beta(n+1) = \alpha + i\beta - \alpha + in\beta \in \sigma(a - \alpha\mathbf{1} + in\beta\mathbf{1})$. Thus with $r(c) \leq \|c\|$ (Proposition 13.27), the C^*-identity and $\|\mathbf{1}\| = 1$ we have

$$\begin{aligned}(n^2+2n+1)\beta^2 = |i\beta(n+1)|^2 \ &\leq \ r(a-\alpha\mathbf{1}+in\beta\mathbf{1})^2 \ \leq \ \|a-\alpha\mathbf{1}+in\beta\mathbf{1}\|^2 \\ = \|(a-\alpha\mathbf{1}-in\beta\mathbf{1})(a-\alpha\mathbf{1}+in\beta\mathbf{1})\| \ &= \ \|(a-\alpha\mathbf{1})^2+n^2\beta^2\mathbf{1}\| \ \leq \ \|a-\alpha\mathbf{1}\|^2+n^2\beta^2.\end{aligned}$$

This simplifies to $(2n+1)\beta^2 \leq \|a - \alpha\mathbf{1}\|^2 \ \forall n \in \mathbb{N}$, which implies $\beta = 0$, thus $\lambda \in \mathbb{R}$. ■

16.18 Remark 1. Since (i) implies $\|a\| = \|a^*a\|^{1/2} = r(a^*a)^{1/2}$ for all $a \in \mathcal{A}$ and the spectral radius $r(a)$ by definition depends only on the algebraic structure of $\mathcal{A}$, the latter also determines the norm, which therefore is unique in a C^*-algebra! But note that the conclusion $\|\cdot\|_1 = \|\cdot\|_2$ for C^*-norms $\|\cdot\|_{1,2}$ only follows if $\mathcal{A}$ is complete with respect to both norms!

2. The proof of (iii) is short and uses only $r(c) \leq \|c\|$ and the C^*-identity. A less direct, but perhaps more insightful argument uses the exponential function, cf. Example 15.19: If $a = a^* \in \mathcal{A}$ then $u_t = e^{ita}$ with $t \in \mathbb{R}$ satisfies $u_t^* = e^{-ita}$, thus $u_t u_t^* = u_t^* u_t = \mathbf{1}$, so that u_t is unitary. Now (ii) gives $\sigma(u_t) \subseteq S^1$, and $\sigma(a) \subseteq \mathbb{R}$ follows from the spectral mapping theorem or Exercise 15.21.

For another application of the unitarity of e^{ia} for self-adjoint a see Section B.13.1. □

The applications of Proposition 16.17(i) discussed in Section 13.4 (with the exception of the result on $|||A|||$) hold in all abstract C^*-algebras. Proposition 16.17(iii) can be used to improve on the results of Section 13.2, showing that C^*-algebras are better behaved than general Banach algebras:

16.19 Theorem *Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{B} \subseteq \mathcal{A}$ a C^*-subalgebra (=closed self-adjoint subalgebra) containing $\mathbf{1}$. Then $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$ and $\sigma_{\mathcal{B}}(b) = \sigma_{\mathcal{A}}(b)$ for all $b \in B$.*

Proof. Since the inclusion $\mathrm{Inv}\mathcal{B} \subseteq \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$ is clear we need to prove for every $b \in \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$ that $b^{-1} \in \mathcal{B}$. Suppose first that $b = b^* \in \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$. Proposition 16.17(iii) then gives that $b - it\mathbf{1}$ is invertible in $\mathcal{B}$ for all $t \in \mathbb{R}\backslash\{0\}$ and invertible in $\mathcal{A}$ for all $t \in \mathbb{R}$. Lemma 13.17 implies that the function $f : \mathbb{R} \to \mathcal{A}, \ t \mapsto (b - it\mathbf{1})^{-1}$ is continuous. For $t \in \mathbb{R}\backslash\{0\}$, $b - it\mathbf{1}$ is invertible in $\mathcal{B}$, so that uniqueness of inverses gives $f(t) \in \mathcal{B}$ for all $t \neq 0$. Now continuity of f and closedness of $\mathcal{B} \subseteq \mathcal{A}$ imply $f(0) \in \mathcal{B}$. Since $f(0) = b^{-1}$, we have $b^{-1} \in \mathcal{B}$, thus $b \in \mathrm{Inv}\mathcal{B}$.

Let now $b \in \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$, so that b has an inverse $a \in \mathcal{A}$. By Lemma 16.2, also b^* is invertible in $\mathcal{A}$, thus the same holds for bb^*. Since $bb^* \in \mathcal{B}$ is self-adjoint, it has an inverse $c \in \mathcal{B}$ by the first half of the proof, in particular $bb^*c = \mathbf{1}$. Combining with $ab = \mathbf{1}$ we have $a = abb^*c = b^*c$. In view of $b, c \in \mathcal{B}$ we have $b^{-1} = a \in \mathcal{B}$. This finishes the proof of $\mathrm{Inv}\mathcal{B} = \mathcal{B} \cap \mathrm{Inv}\mathcal{A}$. The rest follows by Lemma 13.47(ii). ■

16.3 Positive elements of a C^*-algebra I

16.20 Definition *If $\mathcal{A}$ is a unital C^*-algebra then $a \in \mathcal{A}$ is called positive, or $a \geq 0$, if $a = a^*$ and $\sigma(a) \subseteq [0, \infty)$. The set of positive elements of $\mathcal{A}$ is denoted $\mathcal{A}_+$.*

If X is a compact Hausdorff space and $f \in C(X, \mathbb{C}) = \mathcal{A}$ then $\sigma_{\mathcal{A}}(f) = f(X)$, thus $f \geq 0$ is equivalent to $f(x) \geq 0 \ \forall x \in X$.

16.21 Exercise Give an example of a unital C^*-algebra $\mathcal{A}$ and $a \in \mathcal{A}$ showing that $\sigma(a) \subseteq [0, \infty)$ does not imply $a = a^*$!

16.22 Exercise Let $\mathcal{A}$ be a unital C^*-algebra. Without using results proven later, prove:

(i) If $a \in \mathcal{A}_{sa}$ then $a^2 \geq 0$.

(ii) If $a, b \in \mathcal{A}$ are positive and $a + b = 0$ then $a = b = 0$.

(iii) If $a, b \in \mathcal{A}$ are positive and $ab = ba$ then $a + b$ is positive.

(iv) If $c \in \mathcal{A}$ is normal then c^*c is positive. Hint: Lemma 16.14.

16.23 Definition *If $\mathcal{A}$ is a C^*-algebra and $a, b \in \mathcal{A}_{sa}$ satisfy $b - a \geq 0$ we write $a \leq b$ and, equivalently, $b \geq a$.*

16.24 Exercise Let $\mathcal{A}$ be a unital C^*-algebra. Prove:

(i) The binary relation $\leq$ on $\mathcal{A}_{sa}$ is reflexive and anti-symmetric (i.e. $a \leq b \leq a \ \Rightarrow \ a = b$).

(ii) If $a \in \mathcal{A}_{sa}$ then $-\|a\|\mathbf{1} \leq a \leq \|a\|\mathbf{1}$.

(iii) If $a \in \mathcal{A}_{sa}$ satisfies $-c\mathbf{1} \leq a \leq c\mathbf{1}$ with $c \geq 0$ then $\|a\| \leq c$

(iv) If $a, b \in \mathcal{A}$ are positive with $\|a\| \leq 1$, $\|b\| \leq 1$ then $\|a - b\| \leq 1$.

(v) $a \in \mathcal{A}$ is positive if and only if $a = a^*$ and there is a $t \geq 0$ such that $\|a - t\mathbf{1}\| \leq t$.

(vi) If $a, b \in \mathcal{A}$ are positive then $a + b$ is positive. (No assumption that $ab = ba$!)

(vii) $\leq$ is transitive, and $(\mathcal{A}_{sa}, \leq)$ is a partially ordered set.

17 Continuous functional calculus for C^*-algebras

17.1 Continuous functional calculus for self-adjoint elements

Our goal is to make sense of $f(a)$, where a is a normal element of some arbitrary C^*-algebra $\mathcal{A}$, for all functions $f \in C(\sigma(a), \mathbb{C})$, in such a way that $f \mapsto f(a)$ is a $*$-homomorphism. (If you don't care for this generality, you may substitute $\mathcal{A} = B(H)$.) We will first do this for self-adjoint elements and then generalize to normal ones. We cannot hope to go beyond this: If $f(z) = z^2\overline{z}$ then it is not clear whether to define $f(a)$ as a^2a^* or aa^*a or a^*a^2 when $aa^* \neq a^*a$. [Yet 'quantization theory', motivated by quantum theory, tries to do it.] For normal a, this problem does not arise.

17.1 Proposition *Let $\mathcal{A}$ be a unital C^*-algebra, $a \in \mathcal{A}$ normal and P a polynomial. Then*

$$\|P(a)\| = \sup_{\lambda \in \sigma(a)} |P(\lambda)| = \|P_{|\sigma(a)}\|_\infty.$$

Proof. Normality of a implies that $P(a)$ is normal. Thus

$$\|P(a)\| = r(P(a)) = \sup_{\lambda \in \sigma(P(a))} |\lambda| = \sup_{\lambda \in \sigma(a)} |P(\lambda)|,$$

where the first equality is due to Proposition 16.17(i), the second is the definition of r and the third comes from the spectral mapping theorem (Proposition 15.16(ii) or Exercise 15.15). ■

Even though we are after a result for all normal operators, we first consider self-adjoint operators:

17.2 THEOREM *Let $\mathcal{A}$ be a unital C^*-algebra and $a = a^* \in \mathcal{A}$. Then there is a unique continuous $*$-homomorphism $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{A}$ such that $\alpha_a(P) = P(a)$ for all polynomials. (Usually we will write $f(a)$ instead of $\alpha_a(f)$.) It satisfies*

(i) α_a is an isometry: $\|\alpha_a(f)\| = \sup_{\lambda \in \sigma(a)} |f(\lambda)|$.

(ii) The image of α_a is the smallest C^*-subalgebra $\mathcal{B} \subseteq \mathcal{A}$ containing $\mathbf{1}$ and a. The map $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{B}$ is a $*$-isomorphism. $f(a)$ is self-adjoint if and only if f is real-valued.

(iii) $\sigma(\alpha_a(f)) = f(\sigma(a)) = \{f(\lambda) \mid \lambda \in \sigma(a)\}$. (Spectral mapping theorem)

(iv) If $f \in C(\sigma(a), \mathbb{R})$, $g \in C(f(\sigma(a)), \mathbb{C})$ then $\alpha_a(g \circ f) = \alpha_{\alpha_a(f)}(g)$, or just $g(f(a)) = (g \circ f)(a)$. (We require f to be real-valued in order for $f(a)$ to be self-adjoint.)

Proof. (i) By Propositions 13.27 and 16.17(iii), we have $\sigma(a) \subseteq [-\|a\|, \|a\|]$. By the classical Weierstrass approximation theorem, cf. Theorem A.32, for every continuous continuous function $f : [c, d] \to \mathbb{C}$ and $\varepsilon > 0$ there is a polynomial P such that $|f(x) - P(x)| \leq \varepsilon$ for all $x \in [c, d]$. We cannot apply this directly since $\sigma(a)$, while contained in an interval, need not be an entire interval. But using Tietze's Extension Theorem A.31, we can find (very non-uniquely) a continuous function $g : [-\|a\|, \|a\|] \to \mathbb{C}$ that coincides with f on $\sigma(a)$. Now this g can be approximated uniformly by polynomials thanks to Weierstrass' theorem. (Alternatively, apply the more abstract Stone-Weierstrass theorem directly to f.) In any case, the restriction of the polynomials to $\sigma(a)$ is dense in $C(\sigma(a), \mathbb{C})$ w.r.t. $\|\cdot\|_\infty$. By Proposition 17.1, the map $C(\sigma(a), \mathbb{C}) \supseteq \mathbb{C}[x]_{|\sigma(a)} \to \mathcal{A}$, $P \mapsto P(a)$ is an isometry. Thus applying Lemma 3.12 we obtaining a unique isometry $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{A}$ extending $P \mapsto P(a)$. Since $C(\sigma(a), \mathbb{C})$ is complete, its image under α_a is closed, thus equal to the closure $C^*(\mathbf{1}, a)$ of $\{P(a) \mid P \in \mathbb{C}(x)\}$. Thus (i) is proven up to the claim that α_a is a $*$-homomorphism. This is left as an exercise.

(ii) Since $\alpha_a : \mathbb{C}[x] \to \mathcal{A}$ is a $*$-homomorphism, $\mathcal{B} := \alpha_a(C(\sigma(a), \mathbb{C})) \subseteq \mathcal{A}$ is a $*$-subalgebra. And since α_a is an isometry by (i) and $(C(\sigma(a), \mathbb{C}), \|\cdot\|_\infty)$ is complete, $\mathcal{B}$ is closed, thus a C^*-algebra. Since α_a maps the constant-one function to $\mathbf{1} \in \mathcal{A}$ and the inclusion map $\sigma(a) \hookrightarrow \mathbb{C}$ to a, $\mathcal{B}$ contains $\mathbf{1}, a$. Conversely, the smallest C^*-subalgebra of $\mathcal{A}$ containing $\mathbf{1}$ and a clearly is obtained by taking the norm-closure of the set $\{P(a) \mid P \in \mathbb{C}[z]\}$, which is contained in the image of α_a. As the continuous extension of a $*$-homomorphism, $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{A}$ is a $*$-homomorphism. Finally $f(a)^* = \alpha_a(f)^* = \alpha_a(f^*)$. Since α_a is injective, this equals $f(a) = \alpha_a(f)$ if and only $f = f^*$, which is equivalent to real-valuedness of f.

(iii) Let $f \in C(\sigma(a), \mathbb{C})$. Then clearly $\alpha_a(f) \in \mathcal{B}$. Now

$$\sigma_{\mathcal{A}}(\alpha_a(f)) = \sigma_{\mathcal{B}}(\alpha_a(f)) = \sigma_{C(\sigma(a),\mathbb{C})}(f) = f(\sigma(a)),$$

where the equalities come from Theorem 16.19, from the fact that $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{B}$ is a $*$-isomorphism, and from Exercise 13.24, respectively.

(iv) If $\{P_n\}$ is a sequence of polynomials converging to f uniformly on $\sigma(a)$ and $\{Q_n\}$ is a sequence of polynomials converging to g uniformly on $\sigma(f(a))$, then $Q_n \circ P_n$ converges uniformly to $g \circ f$, thus $Q_n(P_n(a)) = (Q_n \circ P_n)(a)$ converges to $(g \circ f)(a)$. On the other hand, $\{Q_n(P_n(a))\}$ converges uniformly to $g(f(a))$. ■

17.3 EXERCISE Prove that α_a is a $*$-homomorphism.

17.4 REMARK The isometric $*$-isomorphism $\alpha_a : C(\sigma(a), \mathbb{C}) \to \mathcal{B} = C^*(\mathbf{1}, a)$ is a special case of the Gelfand isomorphism for commutative unital C^*-algebras proven in Section 19 (which will go in the opposite direction $\pi : \mathcal{A} \to C(\Omega(\mathcal{A}), \mathbb{C})$). A general commutative C^*-algebra $\mathcal{A}$ is not generated by a single element a, so that we'll need to find a substitute for $\sigma(a)$. It shouldn't be surprising that this will be $\Omega(\mathcal{A})$ (with a suitable topology). □

17.2 Positive elements of a C^*-algebra II. Absolute value

Using the functional calculus, we can continue the considerations begun in Section 16.3.

17.5 Exercise (i) Define $f_+, f_- : \mathbb{R} \to \mathbb{R}$ by $f_+(x) = \max(x, 0)$, $f_-(x) = -\min(x, 0)$. Prove: 1. $f_+f_- = 0$, 2. $f_\pm(x) = (|x| \pm x)/2$, 3. $f_\pm \in C(\mathbb{R}, \mathbb{R})$.

(ii) Let now $\mathcal{A}$ be a unital C^*-algebra and $a = a^* \in \mathcal{A}$. Define $a_\pm \in \mathcal{A}$ by functional calculus as $a_\pm = f_\pm(a)$. Prove: 1. $a_+ - a_- = a$ and $a_+ + a_- = |a|$, 2. $a_+a_- = a_-a_+ = 0$, 3. $a_+ \geq 0$, $a_- \geq 0$.

17.6 Proposition *If $\mathcal{A}$ is a unital C^*-algebra and $a \in \mathcal{A}$ then a^*a is positive.*

Proof. First a preparatory argument: Assume $c \in \mathcal{A}$ is such that $-c^*c$ is positive. Then by Lemma 13.25(ii) we have $\sigma(-cc^*)\backslash\{0\} = \sigma(-c^*c)\backslash\{0\}$, thus $-cc^*$ is positive. Writing $c = a + ib$ with a, b self-adjoint, we have $c^*c + cc^* = (a - ib)(a + ib) + (a + ib)(a - ib) = 2a^2 + 2b^2$, thus $c^*c = 2a^2 + 2b^2 - cc^*$. Using $-cc^* \geq 0$ just proven and Exercises 16.22(i) and 16.24(vi), this implies $c^*c \geq 0$. Combining $-c^*c \geq 0$ and $c^*c \geq 0$ gives $\sigma(c^*c) \subseteq [0, \infty) \cap (-\infty, 0] = \{0\}$. This implies $\|c\|^2 = \|c^*c\| = r(c^*c) = 0$, thus $c = 0$.

We turn to the proof of the claim. Let $a \in \mathcal{A}$ be arbitrary. Then $b = a^*a$ is self-adjoint, thus with Exercise 17.5(ii) we have $b = b_+ - b_-$ with $b_\pm \geq 0$ and $b_+b_- = 0$. Putting $c = ab_-$ we have $-c^*c = -b_-a^*ab_- = -b_-(b_+ - b_-)b_- = b_-^3$, which is positive (spectral mapping theorem). Now the preparatory step gives $ab_- = c = 0$. This implies $-b_-^2 = (b_+ - b_-)b_- = bb_- = a^*ab_- = 0$, thus $b_- = 0$. (Since $d = d^*$, $d^2 = 0$ implies $d = 0$.) Now we have $a^*a = b = b_+ \geq 0$. ■

17.7 Exercise Let $\mathcal{A}$ be a unital C^*-algebra and $a, b \in \mathcal{A}$ with $a \geq 0$. Use Proposition 17.6 to prove that $bab^* \geq 0$. Conclude that $a, c \in \mathcal{A}_{sa}$, $a \leq c \Rightarrow bab^* \leq bcb^*$.

17.8 Proposition *Let $\mathcal{A}$ be a unital C^*-algebra. If $a \in \mathcal{A}$ is positive then there is a positive $b \in \mathcal{A}$ such that $b^2 = a$, unique in $C^*(\mathbf{1}, a)$ (and in $\mathcal{A}$). We write $b = \sqrt{a}$.*

Proof. In view of $a \geq 0$ we have $\sigma(a) \subseteq [0, \infty)$. Now continuity of the function $[0, \infty) \to [0, \infty)$, $x \mapsto +\sqrt{x}$ allows us to define $b = \sqrt{a}$ by the continuous functional calculus. It is immediate by construction that $b = b^*$, and the spectral mapping theorem gives $\sigma(b) \subseteq [0, \infty)$, thus $b \geq 0$. Now $b^2 = (a^{1/2})^2 = a$ since $\sqrt{x}^2 = x$. If $c \in C^*(\mathbf{1}, a)$ is positive and $c^2 = a$ then $c = b$. This follows from the $*$-isomorphism $C^*(\mathbf{1}, a) \cong C(\sigma(a), \mathbb{C})$ and the fact that positive square roots are unique in the function algebra $C(\sigma(a), \mathbb{C})$ (why?). The stronger result that positive square roots are unique even in $\mathcal{A}$ can be proven as in Theorem 11.40 if we have replacements of Exercises 11.39(iii) and (iv) valid for abstract C^*-algebras. These are provided by Exercises 17.7 and 16.22(ii), respectively. ■

For elements of the C^*-algebra $B(H)$, where H is a Hilbert space, we have two competing definitions of positivity. Luckily there is no conflict:

17.9 Proposition *If H is a complex Hilbert space and $A \in B(H)$, the following are equivalent:*

(i) C^*-algebraic positivity: $A = A^*$ and $\sigma(A) \subseteq [0, +\infty)$.

(ii) Operator positivity: $\langle Ax, x\rangle \geq 0$ for all $x \in H$, equivalently $W(A) \subseteq [0, +\infty)$.

Proof. If A is C^*-positive then by Proposition 17.8 there is a $B = B^* \in B(H)$ such that $A = B^2 = B^*B$. Thus A is operator positive by Exercise 11.39(ii).

If A is operator positive then $A = A^*$ by Proposition 11.22, and using Exercise 13.15 we have $\sigma(A) \subseteq \overline{W(A)} \subseteq [0, \infty)$. Thus A is C^*-positive. ■

17.10 Exercise Let $\mathcal{A}$ be a unital C^*-algebra. Prove $|a| = \sqrt{a^2}$ for all $a \in \mathcal{A}_{sa}$.

With Proposition 17.6 we can define the 'absolute value' of all elements, not only the self-adjoint ones:

17.11 Definition *If $\mathcal{A}$ is a unital C^*-algebra and $a \in \mathcal{A}$, we define $|a| = (a^*a)^{1/2}$.*

By construction, $|a|$ is positive. And if a is positive then $|a| = a$. These properties are similar to those of $|z|$ for $z \in \mathbb{C}$, but some care is required (compare Exercise 19.17):

17.12 Exercise (i) Let A be a unital C^*-algebra and $a \in A$. Prove that $|a| = |a^*|$ holds if and only if a is normal.

(ii) Find counterexamples in $\mathcal{A} = M_{2\times 2}(\mathbb{C})$ disproving $|a + b| \le |a| + |b|$ and $|ab| \le |a|\,|b|$.

In Section 11.8 we have proven polar decomposition for the C^*-algebras $\mathcal{A} = B(H)$, but this does not generalize to arbitrary C^*-algebras:

17.13 Exercise (i) If $\mathcal{A}$ is a unital C^*-algebra and $a \in \mathrm{Inv}(\mathcal{A})$, prove that $|a| \in \mathrm{Inv}\mathcal{A}$. Use this to define $u \in \mathcal{A}$ by $a = u|a|$ and prove that u is unitary.

(ii) Give an example of a unital C^*-algebra $\mathcal{A}$ and $a \in \mathcal{A}$ such that there is no $b \in \mathcal{A}$ with $a = b|a|$.

17.14 Remark We have seen in Exercise 17.13 that in a unital C^*-algebra one does not always have polar decomposition. However, there is a class of particularly nice C^*-subalgebras $\mathcal{A} \subseteq B(H)$, the von Neumann algebras, such that for each $A \in \mathcal{A}$ one has not only $|A| \in \mathcal{A}$, but also $V \in \mathcal{A}$, where V is as in Proposition 11.44. Cf. e.g. [110, Theorem 4.1.10]. □

In Remark 13.28 we showed $r(a) \le \inf_{b\in\mathrm{Inv}\mathcal{A}} \|bab^{-1}\|$ for every element of a unital Banach algebra. For C^*-algebras, we can now prove this to be an equality:

17.15 Exercise Let $\mathcal{A}$ be a unital C^*-algebra. Let $a \in \mathcal{A}$ with $r(a) < 1$.

(i) Prove that $c = \sum_{n=0}^\infty (a^*)^n a^n$ converges and $c \ge \mathbf{1}$.

(ii) Define $b = c^{1/2}$ and prove $b \ge \mathbf{1}$ and $b \in \mathrm{Inv}\mathcal{A}$.

(iii) Prove $\|bab^{-1}\| < 1$. Hint: Use the C^*-identity and $b^2 = c$.

(iv) Use (i)-(iii) to prove $r(a) = \inf_{b\in\mathrm{Inv}\mathcal{A}} \|bab^{-1}\|$ for all $a \in \mathcal{A}$.

The results of (i)-(iii) can also be used to relate $\sigma(A)$, where $A \in B(H)$, to the sets $\overline{W(BAB^{-1})}$, where $B \in \mathrm{Inv}\, B(H)$, cf. Appendix B.12.1.

17.3 Continuous functional calculus for normal elements

17.16 Theorem *Theorem 17.2 literally extends to all normal elements of a unital C^*-algebra, with $\mathbb{R}$ replaced by $\mathbb{C}$.*

The proof of Theorem 17.2 does not generalize immediately.[95] The reason is that the spectrum of a normal operator need not be contained in $\mathbb{R}$. (In fact, for normal a we have $\sigma(a) \subseteq \mathbb{R} \Leftrightarrow a = a^*$, cf. Exercise 17.20.) If that happens, the polynomials, restricted to $\sigma(a)$,

[95] "It is a well-known technical nuisance that the proof of the spectral theorem for normal operators involves some difficulties that do not arise in the Hermitian case". [67]

fail to be uniformly dense in $C(\sigma(a),\mathbb{C})$. (All functions that are uniform limits of polynomials on sufficiently large subsets of $\mathbb{C}$ are holomorphic so that, e.g. $f(z) = \mathrm{Re}\, z$ cannot be approximated by polynomials in $z = x + iy$.) But with $\sigma(a) \subseteq \mathbb{C} \cong \mathbb{R}^2$ and considering functions on (a subset of) $\mathbb{C}$ as functions of two real variables, the polynomials in x, y are dense in $C(\sigma(a),\mathbb{C})$ by the higher dimensional version of the classical Weierstrass theorem, cf. Theorem A.38. Thus also the polynomials in $z = x + iy$ and $\overline{z} = x - iy$ are dense[96]. If $P = \sum_{i,j=0}^N c_{ij} z^i \overline{z}^j \in \mathbb{C}[z,\overline{z}]$ we define $P^*(z,\overline{z}) = \sum_{i,j=0}^N \overline{c_{ij}} z^j \overline{z}^i$. This turns $\mathbb{C}[z,\overline{z}]$ into a $*$-algebra.

There is a unique unital homomorphism α_a from $\mathbb{C}[z,\overline{z}]$ to $\mathcal{A}$ sending z to a and $\overline{z}$ to a^*, and we need to adapt Proposition 17.1 to this setting. For this we need another lemma:

17.17 LEMMA *Let $\mathcal{A}$ be a unital C^*-algebra. Then every character $\varphi \in \Omega(\mathcal{A})$ satisfies $\varphi(c^*) = \overline{\varphi(c)}$ for all $c \in \mathcal{A}$, i.e. is a $*$-homomorphism.*

Proof. We have $c = a + ib$, where $a = \mathrm{Re}(c)$, $b = \mathrm{Im}(c)$ are self-adjoint. Now $\sigma(a) \subseteq \mathbb{R}$ by Proposition 16.17(iii), thus $\varphi(a) \in \sigma(a) \subseteq \mathbb{R}$ by Lemma 15.3. Similarly $\varphi(b) \in \mathbb{R}$. Thus

$$\varphi(c^*) = \varphi(a - ib) = \varphi(a) - i\varphi(b) = \overline{\varphi(a) + i\varphi(b)} = \overline{\varphi(a + ib)} = \overline{\varphi(c)},$$

where the third equality used that $\varphi(a), \varphi(b) \in \mathbb{R}$ as shown before. ■

17.18 PROPOSITION *Let $\mathcal{A}$ be a unital C^*-algebra and $a \in \mathcal{A}$ normal. Then*

(i) There is a unique homomorphism $\alpha_a : \mathbb{C}[z,\overline{z}] \mapsto \mathcal{A}$ sending z to a and $\overline{z}$ to a^*. It is a $*$-homomorphism.

(ii) For every $P \in \mathbb{C}[z,\overline{z}]$ we have $\sigma(P(a,a^*)) = \{P(\lambda,\overline{\lambda}) \mid \lambda \in \sigma(a)\}$ and

$$\|P(a,a^*)\| = \sup_{\lambda\in\sigma(a)} |P(\lambda,\overline{\lambda})|. \tag{17.1}$$

Proof. (i) It is clear that we must define α_a by $\sum_{i,j=0}^N c_{ij} z^i \overline{z}^j \mapsto \sum_{i,j=0}^N c_{ij} a^i (a^*)^j$. Using the normality of a it is straightforward to see that $P \mapsto P(a,a^*)$ is a $*$-homomorphism.

(ii) Since a is normal, $\mathcal{B} = C^*(\mathbf{1}, a) \subseteq \mathcal{A}$ is commutative, so that Proposition 15.7 applies, and using that the $\varphi \in \Omega(\mathcal{B})$ are $*$-homomorphisms by Lemma 17.17, we have

$$\begin{aligned}\sigma_{\mathcal{B}}(P(a,a^*)) &= \{\varphi(P(a,a^*)) \mid \varphi \in \Omega(\mathcal{B})\} = \Big\{\varphi\Big(\sum_{i,j=0}^N c_{ij} a^i (a^*)^j\Big) \mid \varphi \in \Omega(\mathcal{B})\Big\} \\ &= \Big\{\sum_{i,j=0}^N c_{ij} \varphi(a)^i \overline{\varphi(a)}^j \mid \varphi \in \Omega(\mathcal{B})\Big\} = \{P(\lambda,\overline{\lambda}) \mid \lambda \in \sigma(a)\}.\end{aligned}$$

Appealing to Theorem 16.19, we get $\sigma_{\mathcal{A}}(P(a,a^*)) = \sigma_{\mathcal{B}}(P(a,a^*))$. Since $P(a,a^*)$ is normal, with Proposition 16.17(i) we have $\|P(a,a^*)\| = r(P(a,a^*)) = \sup_{\lambda\in\sigma(a)} |P(\lambda,\overline{\lambda})|$. ■

17.19 REMARK The main difficulty in the construction of the continuous functional calculus for normal operators is proving (17.1). (By contrast, Proposition 17.1 has an elementary proof since it only uses the spectral mapping theorem for polynomials.) The above proof is efficient and elegant, but it relies on Zorn's lemma via Proposition 15.7. Avoiding this at the present level of generality is possible, but quite cumbersome, cf. e.g. [155, Section 7.4], which makes massive

[96] We allow ourselves the harmless sloppiness of not distinguishing between elements of the ring $\mathbb{C}[z,\overline{z}]$ (where $z, \overline{z}$ are independent variables) and the functions $\mathbb{C} \to \mathbb{C}, z \mapsto f(z,\overline{z})$ induced by them.

use of holomorphic functional calculus including a version for several commuting operators. For normal Hilbert space operators there proofs [14, 174] (see also [113, Prop. 8.21]) that are reasonably elementary, but tricky and ad hoc. □

Proof of Theorem 17.16 Essentially as that of Theorem 17.2, now using the density of the polynomials in $z, \overline{z}$ in $C(\sigma(a), \mathbb{C})$ as explained before and using Proposition 17.18 instead of Proposition 17.1. ■

17.20 Exercise Let $\mathcal{A}$ be a unital C^*-algebra and $a \in \mathcal{A}$ normal. Prove

(i) If $\sigma(a) \subseteq \mathbb{R}$ then a is self-adjoint.

(ii) If $\sigma(a) \subseteq S^1$ then a is unitary.

17.21 Exercise Let $\mathcal{A}, \mathcal{B}$ be unital C^*-algebras, $\alpha : \mathcal{A} \to \mathcal{B}$ a unital $*$-homomorphism, $a \in \mathcal{A}$ normal and $f \in C(\sigma(a), \mathbb{C})$. Prove that $\alpha(f(a)) = f(\alpha(a))$. Why is the r.h.s. even defined?

17.22 Exercise (i) Let $\mathcal{A}$ be a unital C^*-algebra and $u \in \mathcal{A}$ unitary (thus $\sigma(u) \subseteq S^1$). Prove that if $\sigma(u) \neq S^1$, there exists $a \in \mathcal{A}_{sa}$ such that $e^{ia} = u$.

(ii) Give an example of a unital C^*-algebra $\mathcal{A}$ and a unitary $u \in \mathcal{A}$ such that there is no $a \in \mathcal{A}_{sa}$ with $u = e^{ia}$.

The following is a preview of later developments in infinitely many dimensions:

17.23 Exercise Let H be a finite-dimensional Hilbert space and $A \in B(H)$ normal. Prove:

(i) $A = \sum_{i=1}^n \lambda_i P_i$, where the P_i are the orthogonal projections onto the eigenspaces of A and the λ_i are the associated eigenvalues.

(ii) For any function $f : \{\lambda_1, \ldots, \lambda_n\} \to \mathbb{C}$, the $f(A)$ provided by continuous functional calculus coincides with $f(A) = \sum_{i=1}^n f(\lambda_i) P_i$.

With the help of continuous functional calculus for normal operators, we can improve on the results obtained in Exercises 13.66 and 13.68 (for Banach space operators):

17.24 Proposition *Let H be a Hilbert space and $A \in B(H)$ normal.*

(i) If $\Sigma \subseteq \sigma(A)$ is clopen with $\emptyset \neq \Sigma \neq \sigma(A)$ and $P = \chi_\Sigma(A) \in B(H)$ then with $H_1 = PH, H_2 = (1-P)H$ we have $AH_i \subseteq H_i$, $i = 1, 2$, thus $A = A_{|H_1} \oplus A_{|H_2}$. The restrictions $A_{|H_1}$, $A_{|H_2}$ are normal, and $\sigma(A_{|H_1}) = \Sigma$, $\sigma(A_{|H_2}) = \sigma(A) \backslash \Sigma$.

(ii) If $\lambda \in \sigma(A)$ is isolated then $\ker(A - \lambda\mathbf{1}) \neq \{0\}$, thus $\lambda \in \sigma_{\mathrm{p}}(A)$, and $(A - \lambda\mathbf{1}) \restriction \ker(A - \lambda\mathbf{1})^\perp$ is invertible.

Proof. (i) The assumption $\Sigma \neq 0$ implies $\sigma(\chi_\Sigma(A)) = \chi_\Sigma(\sigma(A)) \neq \{0\}$, thus $P \neq 0$. Similarly, $\sigma \neq \sigma(A)$ implies $P \neq \mathbf{1}$. Since P commutes with A, the subspaces $H_1 = PH$, $H_2 = (1-P)H$ are mapped into themselves by A and every $f(A)$, where $f \in C(\sigma(A), \mathbb{C})$. Normality of $A_i = A_{|H_i}$ is clear. Now there are unital $*$-homomorphisms $\pi_i : C(\sigma(A), \mathbb{C}) \to B(H_i)$ such that $f(A) = \pi_1(f) \oplus \pi_2(f)$ for all $f \in C(\sigma(A), \mathbb{C})$. Since Σ is clopen, the $*$-homomorphism $C(\sigma(A), \mathbb{C}) \to C(\Sigma, \mathbb{C}) \oplus C(\sigma(A) \backslash \Sigma, \mathbb{C})$, $f \mapsto (f_{|\Sigma}, f_{|\sigma(A)\backslash\Sigma})$ is an isomorphism. (The inverse sends (f_1, f_2) to $\widehat{f}_1 + \widehat{f}_2$ where $\widehat{f}_i$ is the extension of f_i to all of $\sigma(A)$ that vanishes on the complement of the domain of f_i.) Now the composite $C(\Sigma, \mathbb{C}) \oplus C(\sigma(A) \backslash \Sigma, \mathbb{C}) \to C(\sigma(A), \mathbb{C}) \to B(H_1) \oplus B(H_2)$ sends (f_1, f_2) to $(\pi_1(f_1), \pi_2(f_2))$. If now z_1, z_2 are the inclusion maps from Σ and $\sigma(A) \backslash \Sigma$, respectively, to $\mathbb{C}$, we have $A = A_1 \oplus A_2 = \pi_1(z_1) \oplus \pi_2(z_2)$. Thus $\sigma(A_1) = \sigma(\pi_1(z_1)) \subseteq \sigma(z_1) = \Sigma$. Analogously $\sigma(A_2) \subseteq \sigma(A) \backslash \Sigma$. Now in view of $\sigma(A) = \sigma(A_1) \cup \sigma(A_2)$, we have $\sigma(A_1) = \Sigma$, $\sigma(A_2) \subseteq \sigma(A) \backslash \Sigma$.

(ii) Since λ is isolated, $\Sigma = \{\lambda\} \subseteq \sigma(A)$ is clopen. Applying (i) to Σ gives H_1, H_2, A_1, A_2 with $\sigma(A_1) = \{\lambda\}$ and $\sigma(A_2) = \sigma(A) \backslash \{\lambda\}$. Now Exercise 13.71(ii) gives $A_1 = \lambda \mathbf{1}$, so that $H_1 = \ker(A - \lambda \mathbf{1})$. And $(A_2 - \lambda \mathbf{1}) \in B(H_2)$ is invertible. ■

Now we are are in a position to answer the question raised in Remark 11.28:

17.25 Corollary *Let H be a Hilbert space and $A \in B(H)$ normal. With $H' = (\ker A)^\perp$ we have $AH' \subseteq H'$, and the following are equivalent:*

(i) $A \restriction H' \in B(H')$ is invertible ($\Leftrightarrow$ surjective $\Leftrightarrow$ bounded below).

(ii) $0 \notin \sigma(A)$ or $0 \in \sigma(A)$ is isolated.

Proof. Recall from Proposition 11.27 that A maps H' to itself.

(ii)⇒(i) If $0 \notin \sigma(A)$ then A is invertible, thus $H' = H$ and $A \restriction H'$ is invertible. If $0 \in \sigma(A)$ is isolated then Proposition 17.24(ii) gives a decomposition $H = H_1 \oplus H_2$, where $H_1 = \ker A$ and $H_2 = (\ker A)^\perp = H'$, with $A \restriction H_2 = A \restriction H'$ invertible.

(i)⇒(ii) If A is injective, the assumption of invertibility of $A \restriction H'$ becomes invertibility of A, implying $0 \notin \sigma(A)$. If A is not injective then invertibility of $A \restriction H'$ means $0 \notin \sigma(A \restriction H')$. Since $\sigma(A \restriction H') \subseteq \mathbb{C}$ is closed, there is an open neighborhood $U \subset \mathbb{C}$ of 0 such that $U \cap \sigma(A \restriction H') = \emptyset$. On the other hand, the spectrum of $A \restriction \ker A$ obviously is $\{0\}$. Since $\sigma(A) = \{0\} \cup \sigma(A \restriction H')$, we have $\sigma(A) \cap U = \{0\}$, so that $0 \in \sigma(A)$ is isolated. ■

18 Spectral theorems for normal Hilbert space operators

18.1 Spectral theorem: Multiplication operator version

In this section we will prove several spectral theorems for normal Hilbert space operators, not requiring compactness. We assume some nodding acquaintance with measure and integration theory or willingness to learn the basics.

Let H be a Hilbert space, $A \in B(H)$ normal and $x \in H$. Then the map

$$\varphi_{A,x} : C(\sigma(A), \mathbb{C}) \to \mathbb{C}, \quad f \mapsto \langle f(A)x, x \rangle$$

is a bounded linear functional on the Banach space $(C(\sigma(A), \mathbb{C}), \|\cdot\|_\infty)$ since $\|f(A)\| = \|f\|_\infty$. If f is positive (i.e. takes values in $[0, \infty)$) then $\sigma(f(A)) \subseteq [0, \infty)$ by the spectral mapping theorem, so that $f(A) \geq 0$ and $\langle f(A)x, x \rangle \geq 0$ by Proposition 17.9. Thus $\varphi_{A,x}$ is a bounded positive linear functional on $C(\sigma(A), \mathbb{C})$. Now by the Riesz-Markov-Kakutani theorem, cf. Section A.7, there is a unique finite positive measure $\mu_{A,x}$ on the Borel σ-algebra of $\sigma(A)$ such that

$$\int f \, d\mu_{A,x} = \varphi_{A,x}(f) = \langle f(A)x, x \rangle \qquad \forall f \in C(\sigma(A), \mathbb{C}). \tag{18.1}$$

Taking $f = 1 = \text{const.}$, we have $f(A) = \mathbf{1}$, so that $\mu_{A,x}(\sigma(A)) = \int 1 \, d\mu_{A,x} = \|x\|^2 < \infty$. Now we have the Hilbert space $L^2(\sigma(A), \mu_{A,x})$, where we omit the Borel σ-algebra from the notation.

18.1 Definition *Let H be a Hilbert space, $A \in B(H)$ normal and $x \in H$. Then x is called $*$-cyclic for A if $\overline{\text{span}_\mathbb{C}\{A^n(A^*)^m x \mid n, m \in \mathbb{N}_0\}} = H$.*

18.2 Remark A vector x is called cyclic for A if $\overline{\text{span}_{\mathbb{C}}\{A^n x \mid n \in \mathbb{N}_0\}} = H$. Clearly the two notions are equivalent for self-adjoint A, but a normal operator can be $*$-cyclic but not cyclic (e.g. the shift on $\ell^2(\mathbb{Z}, \mathbb{F})$). For the present purpose, $*$-cyclicity is the right notion. □

18.3 Proposition *Let H be a complex Hilbert space, $A \in B(H)$ normal and $x \in H$ $*$-cyclic for A. Then there is a unitary $U : H \to L^2(\sigma(A), \mu_{A,x})$ such that $UAU^* = M_z$, where $(M_z f)(z) = zf(z)$ for all $f \in L^2(\sigma(A), \mu_{A,x})$ and $\mu_{A,x}$-almost all $z \in \sigma(A)$.*

Thus A is unitarily equivalent to a multiplication operator.

Proof. For all $f \in C(\sigma(A), \mathbb{C})$ we have

$$\|f(A)x\|^2 = \langle f(A)x, f(A)x\rangle = \langle f(A)^* f(A)x, x\rangle = \langle (\overline{f}f)(A)x, x\rangle = \int |f|^2 \, d\mu_{A,x} = \|f\|_2^2,$$

where the third equality comes from the $*$-homomorphism property of the functional calculus and the fourth from (18.1). Thus if we equip $C(\sigma(A), \mathbb{C})$ with the seminorm $\|f\|_2 = (\int |f|^2 \, d\mu_{A.x})^{1/2}$, the map $\alpha : C(\sigma(A), \mathbb{C}) \to H$, $f \mapsto f(A)x$ is isometric. With $\mathcal{L}^2(\sigma(A), \mu_{A,x}) = \{f : \sigma(A) \to \mathbb{C} \mid f \text{ measurable}, \|f\|_2 < \infty\}$ we have $C(\sigma(A), \mathbb{C}) \subseteq \mathcal{L}^2(\sigma(A), \mu_{A,x})$ (since $\mu_{A,x}$ is finite). Recall that $L^2(\sigma(A), \mu_{A,x})$ is defined as the quotient space of $\mathcal{L}^2(\sigma(A), \mu_{A,x})$ w.r.t. the equivalence relation $\sim$ defined by $f \sim g \Leftrightarrow \|f - g\|_2 = 0 \Leftrightarrow f = g$ μ-a.e. Since $f \sim g$ implies $\|f(A)x - g(A)x\| = \|f - g\|_2 = 0$, the map α descends to an isometric map $\alpha' : C(\sigma(A), \mathbb{C})/\!\sim \to H$ such that the triangle in

$$\begin{array}{ccccc}
\mathcal{L}^2(\sigma(A), \mu_{A,x}) & \supseteq & C(\sigma(A), \mathbb{C}) & & \\
\downarrow & & \downarrow & \searrow^{\alpha} & \\
L^2(\sigma(A), \mu_{A,x}) & \supseteq & C(\sigma(A), \mathbb{C})/\!\sim & \xrightarrow[\alpha']{} & H
\end{array}$$

commutes. It is a fact from measure theory that $C(\sigma(A), \mathbb{C})/\!\sim \,\subseteq L^2(\sigma(A), \mu_{A,x})$ is a dense linear subspace, cf. e.g. [140, Theorem 3.14]. Thus by Lemma 3.12 there is a unique isometric map $\widehat{\alpha'} : L^2(\sigma(A), \mu_{A,x}) \to H$ extending α'. With $\{f(A)x \mid f \in C(\sigma(A), \mathbb{C})\} \supseteq \{A^i (A^*)^j x \mid i, j \in \mathbb{N}_0\}$, the assumption that x be $*$-cyclic implies that α, thus α', has dense image in H. Since $L^2(\sigma(A), \mu_{A,x})$ is complete, its image under $\widehat{\alpha'}$ is closed and dense, thus all of H. Thus $\widehat{\alpha'}$ is unitary, and we define $U = \widehat{\alpha'}^* : H \to L^2(\sigma(A), \mu_{A,x})$. For $f \in C(\sigma(A), \mathbb{C})$ we have $U^*[f] = f(A)x$, thus

$$(UAU^*)([f])(z) = (UAf(A)x)(z) = (U(zf)(A)x)(z) = [zf](z),$$

and by density of $C(\sigma(A), \mathbb{C})/\!\sim$ in $L^2(\sigma(A), \mu_{A,x})$, this holds for all $f \in L^2(\sigma(A), \mu_{A,x})$ and $\mu_{A,x}$-almost all $z \in \sigma(A)$. ■

Not every normal operator $A \in B(H)$ admits a $*$-cyclic vector, see Exercise 18.9. But we always have:

18.4 Theorem (Spectral theorem for normal operators) *Let H be a complex Hilbert space and $A \in B(H)$ normal. Then there exists a family $\{\mu_\iota\}_{\iota \in I}$ of finite Borel measures on $\sigma(A)$ and a unitary $U : H \to \bigoplus_{\iota \in I} L^2(\sigma(A), \mu_\iota)$[97] such that $UAU^* = \bigoplus_{\iota \in I} M_z$, i.e.*

$$(UAU^* f)_\iota(z) = z f_\iota(z) \qquad \forall f = \{f_\iota\} \in \bigoplus_{\iota \in I} L^2(\sigma(A), \mu_\iota), \ z \in \sigma(A). \tag{18.2}$$

[97] Here $\bigoplus$ is the Hilbert space direct sum defined at the end of Section 5.1.

Proof. Let $\mathcal{F}$ be the family of subsets $F \subseteq H$ such that for $x, y \in F, x \neq y$ we have $f(A)x \perp f'(A)y$ for all $f, f' \in C(\sigma(A), \mathbb{C})$. We partially order $\mathcal{F}$ by inclusion. One easily checks that $\mathcal{F}$ satisfies the hypothesis of Zorn's lemma. (Given a totally ordered subset $\mathcal{C} \subseteq \mathcal{F}$, $\bigcup \mathcal{C}$ is in $\mathcal{F}$, thus an upper bound for $\mathcal{C}$.) Thus there is a maximal element $M \in \mathcal{F}$. For each $x \in M$ we put $H_x = \overline{\{f(A)x \mid f \in C(\sigma(A), \mathbb{C})\}}$. By construction these H_x are mutually orthogonal and $f(A)H_x \subseteq H_x \ \forall x$. Putting $K = \bigoplus_{x \in M} H_x$, we have $f(A)K \subseteq K$ for all $f \in C(\sigma(A), \sigma)$, thus also $f(A)^* K \subseteq K$ since $f(A)^* = \overline{f}(A)$. With Exercise 11.23 this means that K and $K^\perp$ are invariant under all $f(A)$. If $K^\perp \neq \{0\}$ then for every non-zero $y \in K^\perp$ we have $M \cup \{y\} \in \mathcal{F}$, which contradicts the maximality of M. Thus $K = H$.

For every $x \in M$ we have that $x \in H_x$ is $*$-cyclic for the restriction of A to H_x, so that we can apply Proposition 18.3 to obtain unitaries $U_x : H_x \to L^2(\sigma(A), \mu_{A,x})$ such that $U_x A = M_z U_x$. Defining $U : H \to \bigoplus_{x \in M} L^2(\sigma(A), \mu_{A,x})$ by sending $y \in H_x$ to $U_x y \in L^2(\sigma(A), \mu_{A,x})$ and extending linearly, U is unitary. It is clear that we have $UAU^* = \bigoplus_{x \in M} M_z$. Now we are done (with the obvious identifications $I = M$ and $\mu_x = \mu_{A,x}$). ■

18.5 Remark 1. Once the maximal family M of vectors has been picked, the construction is canonical. But there is no uniqueness in the choice of that family. (This is similar to the non-uniqueness of the choices of ONBs in the eigenspaces $\ker(A - \lambda \mathbf{1})$ that we make in proving Theorem 14.12.) For much more on this (in the self-adjoint case) see [128, Section VII.2].

2. Theorem 18.4 is perfectly compatible with Theorem 14.12: If A is compact normal and E is an ONB diagonalizing it then the H_ι in Theorem 18.4 are precisely the one-dimensional spaces $\mathbb{C}e$ for $e \in E$ and the measure μ_ι corresponding to $H_\iota = \mathbb{C}e$ is the δ-measure on $P(\sigma(A))$ defined by $\mu(S) = 1$ if $\lambda_e \in S$ and $\mu(S) = 0$ otherwise. (To be really precise, one should take the non-uniqueness in both theorems into account.)

3. If A is as in the theorem and $g \in C(\sigma(A), \mathbb{C})$ then the continuous functional calculus gives us a normal operator $g(A)$. We now have

$$Ug(A)U^* = \bigoplus_{\iota \in I} M_g.$$

(This is an obvious consequence of (18.2) when g is a polynomial and follows by a density argument in general.) If one took Theorem 18.4 as given, this could even be used to *define* the continuous functional calculus. This would be circular since we used the continuous functional calculus to prove the theorem, or rather Proposition 18.3 on which it relied, but it shows that the continuous functional calculus and the spectral theorem are 'equivalent' in the sense of being easily deducible from each other.

4. The statement of Theorem 18.4 may not quite be what we expected, given the slogan 'normal operators are multiplication operators', since there is a direct sum involved. But this can be fixed when H is separable: □

18.6 Corollary *Let H be a separable complex Hilbert space and $A \in B(H)$ normal. Then there exists a finite measure space $(X, \mathcal{A}, \mu)$, a function $g \in L^\infty(X, \mathcal{A}, \mu; \mathbb{C})$ and a unitary $W : H \to L^2(X, \mathcal{A}, \mu; \mathbb{C})$ such that $WAW^* = M_g$.*

Proof. We apply Theorem 18.4. Since H is separable, the index set I is at most countable, and we write $I = \{1, \ldots, N\}$ where $N \in \mathbb{N} \cup \{\infty\}$ with $\infty = \#\mathbb{N}$. Now we put $X = I \times \sigma(A) = \bigoplus_{i \in I} \sigma(A)$ and for $Y \subseteq X$ we put $Y_i = p_2(p_1^{-1}(i)) = \{x \in \sigma(A) \mid (i, x) \in Y\} \subseteq \sigma(A)$. We define

$\mathcal{A} \subseteq P(X)$ and $\mu : \mathcal{A} \to [0, \infty]$ by

$$\begin{aligned} \mathcal{A} &= \{Y \subseteq X \mid Y_i \in \mathcal{B}(\sigma(A)) \ \forall i \in I\}, \\ \mu(Y) &= \sum_{i \in I} \mu_i(Y_i). \end{aligned}$$

Using the countability of I it is straightforward to check that $\mathcal{A}$ is a σ-algebra on X and μ a (positive) measure on $(X, \mathcal{A})$. With (18.1) we have $\mu_i(\sigma(A)) = \|x_i\|^2$. Thus if we choose the cyclic vectors x_i such that $\|x_i\| = 2^{-i}$ then $\mu(X) = \sum_i \mu(\{i\} \times \sigma(A)) = \sum_i \mu_i(\sigma(A)) < \infty$, so that the measure space $(X, \mathcal{A}, \mu)$ is finite. Now we define a linear map

$$V : \bigoplus_{i \in I} L^2(\sigma(A), \mu_i) \to L^2(X, \mathcal{A}, \mu), \quad \{f_i\}_{i \in I} \mapsto f \ \text{ where } \ f((i, x)) = f_i(x).$$

From the way $(X, \mathcal{A}, \mu)$ was constructed, it is quite clear that V is unitary. (Check this!) Now $W = VU : H \to L^2(X, \mathcal{A}, \mu)$, where U comes from Theorem 18.4, is unitary. In view of $(UAU^* f)_i(\lambda) = \lambda f_i(\lambda)$, defining $g : X \to \mathbb{C}, \ (i, x) \mapsto x$ (which is bounded by $r(A) = \|A\|$), we have $WAW^* = M_g$. ■

18.7 Exercise Use the above results to prove that for every normal $A \in B(H)$, where H is a Hilbert space of dimension ≥ 2, there is a proper closed subspace $\{0\} \neq K \subsetneq H$ such that $AK \subseteq K$.

For a bit more on the existence of invariant subspaces see Section B.8.

18.8 Exercise (i) Let $\Sigma \subseteq \mathbb{C}$ be compact and non-empty and μ be a finite positive Borel measure on Σ. Put $H = L^2(\Sigma, \mu)$ and define $A \in B(H)$ by $A = M_z$, thus $(Af)(z) = zf(z)$ for $f \in H$. Prove:

$$\begin{aligned} \sigma(A) &= \{\lambda \in \Sigma \mid \forall \varepsilon > 0 : \mu(B(\lambda, \varepsilon)) > 0\}, \\ \sigma_{\mathrm{p}}(A) &= \{\lambda \in \Sigma \mid \mu(\{\lambda\}) > 0\}, \end{aligned}$$

where $B(\lambda, \varepsilon)$ denotes the open ε-disc around λ.

(ii) Prove that $L^2(B(\lambda, \varepsilon), \mu)$ is infinite-dimensional for each $\varepsilon > 0$ if $\lambda \in \sigma(A) \backslash \sigma_{\mathrm{p}}(A)$.

(iii) Let $A \in B(H)$ be normal. With the terminology of Theorem 18.4 prove

$$\begin{aligned} \sigma(A) &= \{\lambda \in \mathbb{C} \mid \exists i \in I \ \forall \varepsilon > 0 : \mu_\iota(B(\lambda, \varepsilon)) > 0\}, \\ \sigma_{\mathrm{p}}(A) &= \{\lambda \in \mathbb{C} \mid \exists \iota \in I : \mu_\iota(\{\lambda\}) > 0\}. \end{aligned}$$

18.9 Exercise Let H be a Hilbert space and $A \in B(H)$ non-zero and normal. Prove that A admits a $*$-cyclic vector if and only if H is separable and the algebra $\{A, A^*\}'$ is commutative.

18.2 Borel functional calculus for normal operators

In the preceding section we used the continuous functional calculus to prove the spectral theorem for normal operators. Now we will turn the logic around and use the spectral theorem to extend the functional calculus to a larger class of functions!

18.10 Definition *If (X, τ) is a topological space, $B^\infty(X, \mathbb{C})$ denotes the set of bounded functions $X \to \mathbb{C}$ that are measurable with respect to the Borel σ-algebra $B(X, \tau)$.*

18.11 Lemma *Let (X,τ) be a topological space. Then*

(i) If $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of Borel measurable functions $X \to \mathbb{C}$ converging pointwise to f then is Borel measurable.

(ii) $(B^\infty(X,\mathbb{C}), \|\cdot\|_\infty)$, equipped with pointwise multiplication and $*$-operation is a C^*-algebra.

Proof. (i) It is an elementary fact of measure theory, cf. e.g. [29, Proposition 2.1.5], that the pointwise limit of a sequence of measurable functions (whatever the σ-algebra) is measurable.

(ii) Every sequence in $B^\infty(S,\mathbb{C})$ that is Cauchy w.r.t. $\|\cdot\|_\infty$ converges pointwise everywhere, thus is measurable by (i), and clearly bounded. Thus $B^\infty(S,\mathbb{C})$ is complete. It is a C^*-algebra since product and $*$-operation satisfy submultiplicativity and the C^*-identity. ■

For a normal element $a \in \mathcal{A}$ of a C^*-algebra, we cannot make sense of $f(a)$ if f is not continuous. But the C^*-algebra $B(H)$ has much more structure, and it turns out there is a Borel functional calculus extending the continuous functional calculus:

18.12 Theorem *Let H be a complex Hilbert space and $A \in B(H)$ normal. Then:*

(i) There is a unique unital $*$-homomorphism $\alpha_A : B^\infty(\sigma(A),\mathbb{C}) \to B(H)$ extending the continuous functional calculus $C(\sigma(A),\mathbb{C}) \to B(H)$ and satisfying $\|\alpha_A(f)\| \le \|f\|_\infty$. Again we write more suggestively $f(A) = \alpha_A(f)$.

(ii) If $B \in B(H)$ commutes with A then B commutes with $g(A)$ for all $g \in B^\infty(\sigma(A),\mathbb{C})$.

(iii) If $\{f_n\}_{n\in\mathbb{N}} \subseteq B^\infty(\sigma(A),\mathbb{C})$ is a bounded sequence converging pointwise to f then $f \in B^\infty(\sigma(A),\mathbb{C})$ and $f_n(A) \stackrel{w}{\to} f(A)$, i.e. w.r.t. τ_{wot}, cf. Definition 10.18 and Exercise 10.19(i). (And $\|f_n - f\|_\infty \to 0 \Rightarrow \|f_n(A) - f(A)\| \to 0$.)

Proof. (i) For all $x, y \in H$, the map

$$\varphi_{x,y} : C(\sigma(A),\mathbb{C}) \to \mathbb{C}, \quad f \mapsto \langle f(A)x, y\rangle$$

is a linear functional on $C(\sigma(A),\mathbb{C})$ that is bounded since $\|f(A)\| = \|f\|_\infty$. Thus by the Riesz-Markov-Kakutani Theorem A.56 there exists a unique complex Borel measure $\mu_{x,y}$ on $\sigma(A)$ such that $\varphi_{x,y}(f) = \int f\, d\mu_{x,y}$ for all $f \in C(\sigma(A),\mathbb{C})$. Since $\varphi_{x,y}$ depends in a sesquilinear way on (x,y), the same holds for $\mu_{x,y}$, and $|\mu_{x,y}(\sigma(A))| = |\langle x,y\rangle| \le \|x\|\|y\|$. Thus if $f \in B^\infty(\sigma(A),\mathbb{C})$, the map $\psi_f : H^2 \to \mathbb{C}$ defined by $(x,y) \mapsto \int f\, d\mu_{x,y}$ is a sesquilinear form that is bounded since $|\psi_{x,y}(f)| \le \|f\|_\infty \|x\|\|y\|$. Thus by Proposition 11.5 there is a unique $A_f \in B(H)$ such that $\langle A_f x, y\rangle = \psi_{x,y}(f)$ for all $x, y \in H$. It satisfies $\|A_f\| \le \|f\|_\infty$. Define $\alpha : B^\infty(\sigma(A),\mathbb{C}) \to B(H)$ by $f \mapsto A_f$. If $f \in C(\sigma(A),\mathbb{C})$ then $\psi_{x,y}(f) = \langle f(A)x, y\rangle\ \forall x, y$, implying $A_f = f(A)$. Thus α_A extends the continuous functional calculus.

It remains to be shown that α_A is a $*$-homomorphism. Linearity is quite obvious. Since the continuous functional calculus is a $*$-homomorphism, for $f \in C(\sigma(A),\mathbb{C})$ we have $\overline{f}(A) = f(A)^*$, thus

$$\int f\, d\mu_{x,y} = \langle f(A)x, y\rangle = \langle x, f(A)^*y\rangle = \langle x, \overline{f}(A)y\rangle = \overline{\langle \overline{f}(A)y, x\rangle} = \overline{\int \overline{f}\, d\mu_{y,x}} = \int f\, d\overline{\mu_{y,x}},$$

implying $\mu_{y,x} = \overline{\mu_{x,y}}$. Now for all $f \in B^\infty(\sigma(A),\mathbb{C})$ the above computation can be read backwards, giving $\alpha_A(\overline{f}) = \alpha_A(f)^*$. Since the continuous functional calculus is a homomorphism, for all $f, g \in C(\sigma(A),\mathbb{C})$ we have

$$\int (fg)\, d\mu_{x,y} = \langle (fg)(A)x, y\rangle = \langle f(A)g(A)x, y\rangle = \langle g(A)x, f(A)^*y\rangle = \int g\, d\mu_{x,\overline{f}(A)y}.$$

The fact that this holds for all $f, g \in C(\sigma(A), \mathbb{C})$ implies $f\,\mu_{x,y} = \mu_{x,\overline{f}(A)y}$. Thus for all $f \in C(\sigma(A), \mathbb{C})$, $g \in B^\infty(\sigma(A), \mathbb{C})$ we have

$$\langle (fg)(A)x, y\rangle = \int fg\, d\mu_{x,y} = \int g\, d\mu_{x,\overline{f}(A)y} = \langle g(A)x, \overline{f}(A)y\rangle = \langle f(A)g(A)x, y\rangle,$$

so that $(fg)(A) = f(A)g(A)$. As above, we deduce from this that $f\,\mu_{x,y} = \mu_{x,\overline{f}(A)y}$ for all $f \in B^\infty(\sigma(A), \mathbb{C})$, and then $(fg)(A) = f(A)g(A)$ for all $f, g \in B^\infty(\sigma(A), \mathbb{C})$.

(ii) By normality of A and Fuglede's Theorem 16.16 we have $BA^* = A^*B$, thus B commutes with $C^*(\mathbf{1}, A)$, so that $Bf(A) = f(A)B$ for all $f \in C(\sigma(A), \mathbb{C})$. Thus

$$\varphi_{Bx,y}(f) = \langle f(A)Bx, y\rangle = \langle Bf(A)x, y\rangle = \langle f(A)x, B^*y\rangle = \varphi_{x,B^*y}(f) \quad \forall x, y, f.$$

This implies $\mu_{Bx,y} = \mu_{x,B^*y}$ for all x, y, whence

$$\langle f(A)Bx, y\rangle = \int f\, d\mu_{Bx,y} = \int f\, d\mu_{x,B^*y} = \langle f(A)x, B^*y\rangle \quad \forall x, y \in H,\ f \in B^\infty(\sigma(A), \mathbb{C}),$$

thus $f(A)B = Bf(A)$ for all $f \in B^\infty(\sigma(A), \mathbb{C})$.

(iii) Measurability of the limit function f follows from Lemma 18.11(i). If $\|f_n\| \le M\ \forall n$ then clearly $\|f\| \le M$. Thus $f \in B^\infty(\sigma(A))$. For all $x, y \in H$ we have

$$\langle \alpha_A(f_n)x, y\rangle = \int f_n\, d\mu_{x,y} \longrightarrow \int f\, d\mu_{x,y} = \langle \alpha_A(f)x, y\rangle,$$

where convergence in the center is a trivial application of the dominated convergence theorem, using boundedness of $\mu_{x,y}$ and $\|f_n\|_\infty \le M$ for all n. This proves $\alpha_A(f_n) \stackrel{\text{w}}{\to} \alpha_A(f)$. The final claim clearly follows from $\|f_n(A) - f(A)\| = \|(f_n - f)(A)\| \le \|f_n - f\|_\infty$. ■

The above construction of the Borel functional calculus was independent of the Spectral Theorem 18.4. We now wish to understand their relationship. This is the first step:

18.13 Exercise Let $\Sigma \subseteq \mathbb{C}$ be compact and λ a finite positive Borel measure on Σ. Let $H = L^2(\Sigma, \lambda; \mathbb{C})$ and $g \in B^\infty(\Sigma, \mathbb{C})$.

(i) Prove that the multiplication operator $M_g : H \to H,\ [f] \mapsto [gf]$ satisfies

$$\|M_g\| = \operatorname{ess\,sup}_\mu |g| = \inf\{t \ge 0 \mid \lambda(\{x \in X \mid |g(x)| > t\}) = 0\} \le \|g\|_\infty.$$

(ii) Let $A = M_z \in B(H)$, where $z : \Sigma \hookrightarrow \mathbb{C}$. Prove that $g(A)$ as defined by the Borel functional calculus coincides with M_g.

18.14 Corollary *Let $A \in B(H)$ be normal and $g \in B^\infty(\sigma(A), \mathbb{C})$. Then*

(i) $\sigma(g(A)) \subseteq \overline{g(\sigma(A))}$.

(ii) If $h \in B^\infty(\overline{g(\sigma(A))}, \mathbb{C}))$ *then* $h(g(A)) = (h \circ g)(A)$.

Proof. Let $U : H \to \bigoplus_{\iota \in I} L^2(\sigma(A), \mu_\iota)$ as in Theorem 18.4, so that $UAU^* = \bigoplus_\iota M_z$. The projectors P_ι onto the subspaces $L^2(\sigma(A), \mu_\iota)$ of the direct sum commute with A (and A^*), thus also with $g(A)$ for each $g \in B^\infty(\sigma(A), \mathbb{C})$ by Theorem 18.12(ii). Thus the Borel functional calculus respects the direct sum decomposition of A (no matter how the maximal set M was chosen). It is a pure formality to show that if $V : H \to H'$ is unitary then $Vg(A)V^* = g(VAV^*)$.

Thus with the direct sum decomposition $UAU^* = \bigoplus_\iota M_z$ we have $Ug(A)U^* = \bigoplus_\iota g(M_z) = \bigoplus_\iota M_g$, where the second equality comes from Exercise 18.13(ii).

(i) If $\lambda \notin \overline{g(\sigma(A))}$ then $M_g - \lambda\mathbf{1}$ has a bounded inverse with norm $\leq \mathrm{dist}(\lambda, \overline{g(\sigma(A))})^{-1}$. Thus all $M_g - \lambda\mathbf{1}$ in the direct sum decomposition of $A - \lambda\mathbf{1}$ have inverses with uniformly bounded norms. Thus $A - \lambda\mathbf{1}$ has a bounded inverse.

(ii) Under the assumption on h, we have

$$Uh(g(A))U^* = \bigoplus_\iota h(M_g) = \bigoplus_\iota M_{h\circ g} = U(h \circ g)(A)U^*.$$

(This is too sloppy, but the reader should be able to make it precise.) ■

18.15 Remark 1. Since it turns out that $g(A) = U^*(\bigoplus_\iota M_g)U$ for all $g \in B^\infty(\sigma(A), \mathbb{C})$, one might try to take this as the definition of $g(A)$. But apart from being very inelegant, it has the problem that one must prove the independence of $g(A)$ thus defined of the choice of the maximal set $M \subseteq H$ in the proof of the spectral theorem. This would not be difficult if every Borel measurable function was a pointwise limit of a sequence of continuous functions. But this is false, making such an approach quite painful. (Compare Lusin's theorem in, e.g., [140].)

2. We cannot hope to prove $\|g(A)\| = \|g\|_\infty$ for all $g \in B^\infty(\sigma(A), \mathbb{C})$ since it is true only if $\sigma(A) = \sigma_{\mathrm{p}}(A)$! Since singletons in $\mathbb{C}$ are closed, thus Borel measurable, we can change g arbitrarily for some $\lambda \in \sigma(A)$ without destroying the measurability of g, making $\|g\|_\infty$ as large as we want. But if $\lambda \in \sigma(A)\backslash\sigma_{\mathrm{p}}(A)$, Exercise 18.8 gives $\mu_\iota(\{\lambda\}) = 0\ \forall\iota \in I$, so that this change of g does not affect the norms, cf. Exercise 18.13, $\mathrm{ess\,sup}_{\mu_i}|g|$ of the multiplication operators making up $g(A)$ and therefore does not affect $\|g(A)\|$.

3. Let $A \in B(H)$ be normal and consider the C^*-algebra $\mathcal{A} = C^*(\mathbf{1}, A) \subseteq B(H)$. Then $g(A) \in \mathcal{A}$ for continuous g, but for most non-continuous g we have $g(A) \notin \mathcal{A}$. For this reason there is no Borel functional calculus in abstract C^*-algebras. (But $g(A)$ is always contained in the von Neumann algebra $\mathrm{vN}(A) = \overline{C^*(A, \mathbf{1})}^{\mathrm{wot}}$ generated by A. This follows from Theorem 18.12(ii) and von Neumann's 'double commutant theorem'.) □

As a remarkable application, we have the following complement to Exercise 17.22:

18.16 Lemma *If H is a complex Hilbert space and $U \in B(H)$ is unitary, there is a self-adjoint $A \in B(H)$ such that $e^{iA} = U$.*

Proof. Since U is unitary, we have $\sigma(U) \subseteq S^1$. Define $f : S^1 \to \mathbb{C}$ as the unique inverse of the map $(-\pi, \pi] \to S^1, x \mapsto e^{ix}$. The function f is continuous on S^1 except at -1, where it has a jump. Thus it is Borel measurable, so that we can define a normal operator $A = f(U) \in B(H)$ by Borel functional calculus (Theorem 18.12). By Corollary 18.14 we have $\sigma(A) \subseteq \overline{f(\sigma(U))} \subseteq [-\pi, \pi]$ (which together with normality of A implies $A = A^*$) and $e^{iA} = U$. ■

18.17 Proposition *Let H be a complex Hilbert space. Then the open set $\mathrm{Inv}B(H) \subset B(H)$ of invertible operators is path-connected.*

Proof. Let $A \in \mathrm{Inv}B(H)$, and let $A = V|A|$ be its polar decomposition. Then $|A|$ is positive and invertible, and V is unitary, cf. Exercise 11.43. Since $\sigma(|A|) \subseteq (0, +\infty)$, we can use continuous functional calculus to define $B = \log|A| \in B(H)$, satisfying $e^B = |A|$. By the preceding lemma, there is a self-adjoint D with $e^{iD} = V$. Now $g : [0, 1] \to \mathrm{Inv}B(H), t \mapsto e^{itD}e^{tB}$ is a continuous path in $\mathrm{Inv}\,B(H)$ (since e^X is invertible for all X) such that $g(0) = \mathbf{1}$ and $g(1) = e^{iD}e^B = V|A| = A$. Now a standard argument produces paths between any two invertible operators. ■

18.18 Remark The result also holds for infinite-dimensional real Hilbert spaces, but then it requires a different proof. This is surprising since $\mathrm{GL}(n,\mathbb{R})$ is not path-connected for $n \in \mathbb{N}$ but has two path-components, while $\mathrm{GL}(n,\mathbb{C})$ is path-connected. In fact, one has this much stronger theorem of Kuiper[98] [89]: $\mathrm{Inv}\, B(H)$ is contractible in the norm topology for all infinite-dimensional real or complex Hilbert spaces. □

18.3 Normal operators vs. projection-valued measures

There is yet another perspective on the spectral theorem/functional calculus, provided by projection-valued measures:

18.19 Definition *Let H be Hilbert space and $\Sigma \subseteq \mathbb{C}$ a compact subset. Let $\mathcal{B}(\Sigma)$ be the Borel σ-algebra on Σ. A projection-valued measure relative to (H,Σ) is a map $P : \mathcal{B}(\Sigma) \to B(H)$ such that*

(i) $P(S)$ is an orthogonal projection for all $S \in \mathcal{B}(\Sigma)$.

(ii) $P(\emptyset) = 0,\ P(\Sigma) = \mathbf{1}$.

(iii) $P(S \cap S') = P(S)P(S')$ for all $S, S' \in \mathcal{B}(\Sigma)$.

(iv) For all $x, y \in H$, the map $E_{x,y} : \mathcal{B}(\Sigma) \to \mathbb{C},\ S \mapsto \langle P(S)x, y\rangle$ is a complex measure. (Equivalently, if the $\{S_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\Sigma)$ are mutually disjoint then $\sum_n P(S_n)$ converges weakly to $P(\bigcup_n S_n)$.)

Note that (iii) implies $P(S)P(S') = P(S')P(S)$ for all $S, S' \in \mathcal{B}(\Sigma)$.

18.20 Proposition *Let H be a complex Hilbert space and $A \in B(H)$ normal. Put $\Sigma = \sigma(A)$. For each $S \in \mathcal{B}(\Sigma)$, define $P_A(S) = \chi_S(A)$ by Borel functional calculus. Then $S \mapsto P_A(S)$ is a projection-valued measure relative to (H,Σ), also called the spectral resolution of A.*

Proof. If $g = \chi_S$ for $S \in \mathcal{B}(\Sigma)$, $g(A)$ is a direct sum of operators of multiplication by χ_S, which clearly all are idempotent. And since $g = \chi_S$ is real-valued, $g(A)$ is self-adjoint. Thus each $P_A(S) = \chi_S(A)$ is an orthogonal projection. $P_A(\emptyset) = 0$ is clear, and $P_A(\Sigma) = 1(A) = \mathbf{1}_H$ (since the constant 1 function is continuous). Property (iii) is immediate from $\chi_{S\cap S'} = \chi_S \chi_{S'}$. Finally, if $x, y \in H$ let $Ux = \{f_\iota\}_{\iota\in I},\ Uy = \{g_\iota\}_{\iota\in I}$. Then

$$E_{x,y}(S) = \langle P_A(S)x, y\rangle = \sum_{\iota\in I} \int_{\sigma(A)} \chi_S(z) f_\iota(z)\overline{g_\iota(z)}\, d\mu_\iota(z).$$

From this it is clear that $S \mapsto E_{x,y}(S)$ is additive on countable families $\{S_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\Sigma)$ by absolute convergence of $\sum_\iota \int f_\iota \overline{g_\iota}\, d\mu_\iota$. ■

18.21 Exercise Let $A \in B(H)$ be normal and $\Sigma \subseteq \sigma(A)$ a Borel set. Prove $\sigma(A_{|P(\Sigma)H}) \subseteq \overline{\Sigma} \cup \{0\}$. Bonus: State and prove a better result.

18.22 Exercise Let $A \in B(H)$ be a normal operator and let P_A be the corresponding spectral measure. Prove:

(i) $\lambda \in \sigma(A)$ if and only if $P_A(\sigma(A) \cap B(\lambda, \varepsilon)) \neq 0$ for each $\varepsilon > 0$.

(ii) $\lambda \in \sigma(A)$ is an eigenvalue if and only if $P_A(\{\lambda\}) \neq 0$.

[98] Nicolaas Hendrik Kuiper (1920-1994), Dutch mathematician.

We have thus seen that every normal operator gives rise to a projection valued measure. The converse is also true, and we have a bijection between normal operators and projection-valued measures:

18.23 Proposition *Let H be a complex Hilbert space, $\Sigma \subseteq \mathbb{C}$ a compact subset and P a projection-valued measure relative to (H, Σ). Then*

(i) For every $f \in B^\infty(\Sigma, \mathbb{C})$ there is a unique $\alpha(f) \in B(H)$ such that

$$\langle \alpha(f)x, y\rangle = \int f dE_{x,y} \qquad \forall x, y \in H$$

and $\|\alpha(f)\| \leq \|f\|_\infty$ $\forall f$. We also write, somewhat symbolically, $\alpha(f) = \int f(z)\, dP(z)$.

(ii) The map $\alpha : B^\infty(\Sigma, \mathbb{C}) \to B(H)$ is a unital $*$-homomorphism. $\alpha(f)$ is normal for each f.

(iii) Put $A = \alpha(z) \in B(H)$, where $z : \Sigma \hookrightarrow \mathbb{C}$ is the inclusion map. Then $\sigma(A) \subseteq \Sigma$ and $\alpha(f) = f(A)$ for each $f \in C(\Sigma, \mathbb{C})$.

(iv) The maps from normal elements to projection-valued measures (Proposition 18.20) and and conversely ((i)-(iii) above) are mutually inverse.

Proof. (i) It is clear that the map $(x, y) \mapsto E_{x,y}(S) = \langle P(S)x, y\rangle$ is sesquilinear for each $S \in \mathcal{B}(\Sigma)$. For each $f \in B^\infty(\Sigma, \mathbb{C})$, we have $|\int f\, dE_{x,y}| \leq \|f\|_\infty \|x\|\|y\|$. Thus $[x, y]_f = \int f dE_{x,y}$ is a sesquilinear form with norm $\leq \|f\|_\infty$. Thus by Proposition 11.5 there is a unique $A_f \in B(H)$ such that $\langle A_f x, y\rangle = [x, y]_f$ $\forall x, y \in H$ and $\|A_f\| \leq \|f\|_\infty$. Now put $\alpha(f) := A_f$.

(ii) The inequality has already been proven. It is clear that $f \mapsto A_f = \alpha(f)$ is linear. If 1 is the constant one function, we have $\langle \alpha(1)x, y\rangle = \int 1\, dE_{x,y} = E_{x,y}(\Sigma) = \langle x, y\rangle$ since $P(\Sigma) = \mathbf{1}$. Thus $\alpha(1) = \mathbf{1}$. It remains to prove $\alpha(fg) = \alpha(f)\alpha(g)$ and $\alpha(f)^* = \alpha(\overline{f})$. We first do this for characteristic functions of measurable sets S, T: $f = \chi_S, g = \chi_T$. Now

$$\langle \alpha(\chi_S)x, y\rangle = \int_S dE_{x,y} = \langle P(S)x, y\rangle.$$

Thus $\alpha(\chi_S\chi_T) = \alpha(\chi_{S\cap T}) = E(S \cap T) = E(S)E(T) = \alpha(\chi_S)\alpha(\chi_T)$. By linearity of α we now have $\alpha(fg) = \alpha(f)\alpha(g)$ for all simple functions, i.e. finite linear combinations of characteristic functions. The latter are $\|\cdot\|_\infty$-dense in $B^\infty(\Sigma, \mathbb{C})$. (By a proof very similar to that of Lemma 4.13 in the case of ℓ^∞. Note that a measurable function is simple if and only if it assumes only finitely many values.) Now the identity follows for all f, g by continuity of α.

Furthermore, $\alpha(\chi_S)^* = P(S)^* = P(S) = \alpha(\chi_S)$, so that $\alpha(f)^* = \alpha(\overline{f})$ for simple functions. Now apply the same density and continuity argument as above.

In view of $f(A) = \alpha(f)$, the normality of $f(A)$ follows from

$$f(A)f(A)^* = \alpha(f)\alpha(f)^* = \alpha(ff^*) = \alpha(f^*f) = \alpha(f)^*\alpha(f) = f(A)^*f(A).$$

(iii) $\sigma(A) \subseteq \Sigma$ is clear. Since $\alpha(1) = \mathbf{1}$ and $\alpha(z) = A$ by definition, we have $\alpha(P) = P(A)$ for each polynomial. More generally, since α is a $*$-homomorphism, a polynomial in $z, \overline{z}$ is sent by α to the corresponding polynomial in A, A^*. These polynomials are $\|\cdot\|_\infty$-dense in $C(\Sigma, \mathbb{C})$ by Weierstrass Theorem A.38, so that the continuity proven in (ii) implies that $\alpha(f) = f(A)$ as produced by the continuous functional calculus.

(iv) Left as an exercise. ■

We close the discussion of spectral theorems with the advice of looking at the paper [67] and [152, Chapter 5] by two masters of functional analysis.

19 ⋆ The Gelfand homomorphism for commutative Banach and C^*-algebras

19.1 The topology of $\Omega(\mathcal{A})$. The Gelfand homomorphism

Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$ and $\Omega(\mathcal{A})$ its spectrum. For each $a \in \mathcal{A}$ define (as in Section 9.3)

$$\widehat{a} : \Omega(\mathcal{A}) \to \mathbb{C}, \ \varphi \mapsto \varphi(a).$$

We now want a topology τ on $\Omega(\mathcal{A})$ such that $\widehat{a}$ is continuous for each $a \in \mathcal{A}$, thus $\widehat{a} \in C((\Omega(\mathcal{A}), \tau), \mathbb{C})$. Since $\Omega(\mathcal{A}) \subseteq (\mathcal{A}^*)_{\leq 1}$ by Lemma 15.3, we could take τ to be the restriction of the norm topology of $\mathcal{A}^*$ to $\Omega(\mathcal{A})$ (i.e. the relative topology). But we can also take the weakest topology making all $\widehat{a} : \Omega(\mathcal{A}) \to \mathbb{C}$ continuous. This is nothing other than the restriction to $\Omega(\mathcal{A})$ of the weak-$*$ topology or $\sigma(\mathcal{A}^*, \mathcal{A})$-topology on $\mathcal{A}^*$.

19.1 Proposition *Let $\mathcal{A}$ be a unital Banach algebra. Let τ be the restriction of the weak-$*$ topology to $\Omega(\mathcal{A}) \subseteq (\mathcal{A}^*)_{\leq 1}$. Then $(\Omega(\mathcal{A}), \tau)$ is compact Hausdorff.*

Proof. By Alaoglu's theorem, $((\mathcal{A}^*)_{\leq 1}, \tau_{w^*})$ is compact. Thus it suffices to prove that $\Omega(\mathcal{A}) \subseteq (\mathcal{A}^*)_{\leq 1}$ is weak-$*$ closed. Let $\{\varphi_\iota\}$ be a net in $\Omega(\mathcal{A})$ that converges to $\psi \in \mathcal{A}^*$ w.r.t. the $\sigma(\mathcal{A}^*, \mathcal{A})$-topology. Then for all $a, b \in \mathcal{A}$ we have $\psi(ab) = \lim_\iota \varphi_\iota(ab) = \lim_\iota \varphi_\iota(a)\varphi_\iota(b) = \psi(a)\psi(b)$, so that $\psi \in \Omega(\mathcal{A})$. Thus $\Omega(\mathcal{A}) \subseteq (\mathcal{A}^*)_{\leq 1}$ is $\sigma(\mathcal{A}^*, \mathcal{A})$-closed. ■

The above works whether or not $\mathcal{A}$ is commutative, but we'll now restrict to commutative $\mathcal{A}$ since $\Omega(\mathcal{A})$ can be very small otherwise. We begin by completing Exercise 15.5:

19.2 Proposition *Let X be a compact Hausdorff space and $\mathcal{A} = C(X, \mathbb{C})$. Then the map $X \to \Omega(\mathcal{A})$, $x \mapsto \varphi_x$ is a homeomorphism (with the weak-$*$ topology on $\Omega(\mathcal{A})$).*

Proof. Injectivity was already proven in Exercise 15.5. In order to prove surjectivity, let $\varphi \in \Omega(\mathcal{A})$ and put $M = \ker \varphi$. Then $M \subseteq \mathcal{A}$ is a proper closed subalgebra (in fact an ideal), and it is self-adjoint by Lemma 17.17 since $\mathcal{A}$ is a C^*-algebra. If $x, y \in X$, $x \neq y$, pick $f \in \mathcal{A}$ with $f(x) \neq f(y)$. With $g = f - \varphi(f)\mathbf{1}$ we have $\varphi(g) = 0$, thus $g \in M$. This proves that M separates the points of X, yet it is not dense in $\mathcal{A}$. Now the incarnation Corollary A.41 of the Stone-Weierstrass theorem implies that there must be an $x \in X$ at which M vanishes identically, i.e. $\varphi_x(f) = 0$ for all $f \in M$. Now for every $f \in \mathcal{A}$ we have $f - \varphi(f) \in M$, thus $\varphi_x(f - \varphi(f)\mathbf{1}) = 0$, which is equivalent to $\varphi_x(f) = \varphi(f)$. Thus $\iota : X \to \Omega(\mathcal{A})$ is surjective.

If $\{x_\iota\} \subseteq X$ such that $x_\iota \to x$ then $\varphi_{x_\iota}(f) = f(x_\iota) \to f(x) = \varphi_x(f)$ for every $f \in \mathcal{A}$ by continuity of f. But this precisely means that $\varphi_{x_\iota} \to \varphi_x$ w.r.t. the weak-$*$ topology. Thus ι is continuous. As a continuous bijection of compact Hausdorff spaces it is a homeomorphism. ■

19.3 Definition *Let $\mathcal{A}$ be a unital Banach algebra. Then its radical is the set of quasi-nilpotent elements:* $\mathrm{rad}\mathcal{A} = \{a \in \mathcal{A} \mid r(a) = 0\}$. *We call $\mathcal{A}$ semisimple if* $\mathrm{rad}\mathcal{A} = \{0\}$.

19.4 Proposition *If $\mathcal{A}$ is a unital commutative Banach algebra, the map*

$$\pi : \mathcal{A} \to C(\Omega(\mathcal{A}), \mathbb{C}), \quad a \mapsto \widehat{a} \tag{19.1}$$

is a unital homomorphism, called the Gelfand homomorphism (or representation) of $\mathcal{A}$, and $\|\pi(a)\| = r(a) \leq \|a\|$ for all $a \in \mathcal{A}$. Thus $\ker \pi = \mathrm{rad}\mathcal{A}$, and π is injective if and only if $\mathcal{A}$ is semisimple.

Proof. It is clear that π is linear, and $\widehat{\mathbf{1}}(\varphi) = \varphi(\mathbf{1}) = 1$ for all φ. Let $a, b \in \mathcal{A}$, $\varphi \in \Omega(\mathcal{A})$. Then

$$\pi(ab)(\varphi) = \widehat{ab}(\varphi) = \varphi(ab) = \varphi(a)\varphi(b) = \widehat{a}(\varphi)\widehat{b}(\varphi) = \pi(a)(\varphi)\pi(b)(\varphi) = (\pi(a)\pi(b))(\varphi),$$

where we used multiplicativity of φ and the fact that the multiplication on $C(\Omega(\mathcal{A}), \mathbb{C})$ is pointwise, shows that $\pi(ab) = \pi(a)\pi(b)$, thus π is an algebra homomorphism. We have

$$\|\widehat{a}\| = \sup_{\varphi\in\Omega(\mathcal{A})} |\widehat{a}(\varphi)| = \sup_{\varphi\in\Omega(\mathcal{A})} |\varphi(a)| = \sup_{\lambda\in\sigma(a)} |\lambda| = r(a) \leq \|a\|,$$

where we used (15.1) and Proposition 13.27. In particular, $\ker \pi = r^{-1}(0) = \text{rad}\mathcal{A}$. ■

The Gelfand homomorphism can fail to be surjective or injective or both. See Section 19.2 for an important example for the failure of surjectivity and Exercise 19.6 for a non-trivial unital Banach algebra with very large radical.

19.5 Proposition *Let $\mathcal{A}$ be a commutative unital Banach algebra and $a \in \mathcal{A}$ such that $\mathcal{A}$ is generated by $\{\mathbf{1}, a\}$. Then the map $\widehat{a} : \Omega(\mathcal{A}) \to \sigma(a)$ is a homeomorphism.*

The same conclusion holds if $a \in \text{Inv}\mathcal{A}$ and $\mathcal{A}$ is generated by $\{\mathbf{1}, a, a^{-1}\}$.

Proof. We know from (15.1) that $\widehat{a}(\Omega(\mathcal{A})) = \sigma(a)$, thus $\widehat{a}$ is surjective. Assume $\widehat{a}(\varphi_1) = \widehat{a}(\varphi_2)$, thus $\varphi_1(a) = \varphi_2(a)$. Since the φ_i are unital homomorphisms, this implies $\varphi_1(a^n) = \varphi_2(a^n)$ for all $n \in \mathbb{N}_0$, so that φ_1, φ_2 agree on the polynomials in a. Since the latter are dense in $\mathcal{A}$ by assumption and the φ_i are continuous, this implies $\varphi_1 = \varphi_2$. Thus $\widehat{a} : \Omega(\mathcal{A}) \to \sigma(a)$ is injective, thus a continuous bijection. Since $\Omega(\mathcal{A})$ is compact and $\sigma(a) \subseteq \mathbb{C}$ Hausdorff, $\widehat{a}$ is a homeomorphism. This proves the first claim.

For the second claim, note that $\varphi(a)\varphi(a^{-1}) = \varphi(aa^{-1}) = \varphi(\mathbf{1}) = 1$, thus $\varphi(a^{-1}) = \varphi(a)^{-1}$, for each $\varphi \in \Omega(\mathcal{A})$. This implies that $\varphi_1(a^n) = \varphi_2(a^n)$ also holds for negative $n \in \mathbb{Z}$. Now φ_1, φ_2 agree on all Laurent polynomials in a, thus on $\mathcal{A}$ by density and continuity. The rest of the proof is the same. ■

19.6 Exercise Let $\alpha : \mathbb{N}_0 \to (0, \infty)$ be a map satisfying $\alpha(0) = 1$ and $\alpha_{n+m} \leq \alpha_n\alpha_m$ $\forall n, m$. For $f : \mathbb{N}_0 \to \mathbb{C}$, define $\|f\| = \sum_{n\in\mathbb{N}_0} \alpha_n |f(n)|$, and $\mathcal{A} = \{f : \mathbb{N}_0 \to \mathbb{C} \mid \|f\| < \infty\}$. For $f, g \in \mathcal{A}$, define $f \cdot g$ by $(f \cdot g)(n) = \sum_{\substack{u,v\in\mathbb{N}_0 \\ u+v=n}} f(u)g(v)$.

(i) Prove that $(\mathcal{A}, \cdot, \mathbf{1}, \|\cdot\|)$ is a commutative Banach algebra with unit $\mathbf{1} = \delta_0$.

(ii) Prove that δ_1 generates $\mathcal{A}$ and $r(\delta_1) = \lim_{n\to\infty} \alpha_n^{1/n}$.

(iii) Find a sequence $\{\alpha_n\}$ satisfying the above requirements such that δ_1 is quasi-nilpotent.

(iv) Conclude that with the $\{\alpha_n\}$ from (iii) we have $\text{rad}\mathcal{A} = \{f \in \mathcal{A} \mid f(0) = 0\}$.

19.7 Remark 1. Since every commutative unital Banach algebra has at least one non-zero character φ, the worst that can happen is $\text{rad}\mathcal{A} = \varphi^{-1}(0)$, which has codimension one, as in the preceding exercise.

2. If $\mathcal{A}$ is a non-unital Banach algebra and $a \in \mathcal{A}$ one defines $\sigma(a) = \sigma_{\widetilde{\mathcal{A}}}(a)$, where $\widetilde{\mathcal{A}}$ is the unitization of $\mathcal{A}$ considered in Exercise 13.55. Now one defines $r(a) = \sup_{\lambda\in\sigma(a)} |\lambda|$ and $\text{rad}\mathcal{A} = r^{-1}(0) \subseteq \mathcal{A}$ as before. Now for the non-unital subalgebra $\mathcal{A}' = \{f \in \mathcal{A} \mid f(0) = 0\}$ of the $\mathcal{A}$ from Exercise 19.6 one easily proves $\widetilde{\mathcal{A}'} \cong \mathcal{A}$, thus $r(a) = 0$ $\forall a \in \mathcal{A}'$ and $\text{rad}\mathcal{A}' = \mathcal{A}'$. □

19.8 Exercise Let $\mathcal{A}$ be a commutative unital Banach algebra generated by $\{a_1, \ldots, a_n\} \subseteq \mathcal{A}$. Prove that the map $s : \Omega(\mathcal{A}) \to \mathbb{C}^n$, $\varphi \mapsto (\varphi(a_1), \ldots, \varphi(a_n))$ is a homeomorphism of $\Omega(\mathcal{A})$ onto the joint spectrum $\sigma(a_1, \ldots, a_n)$, the latter being a closed subspace of $\sigma(a_1) \times \cdots \times \sigma(a_n)$.

19.2 Application: Absolutely convergent Fourier series

Let $(\mathcal{A} = \ell^1(\mathbb{Z}, \mathbb{C}), \|\cdot\|, \star, \mathbf{1})$ be the unital Banach algebra from Section 4.6. In Exercise 15.11 we used characters to compute the spectra of elements of $\mathcal{A}$ and proved that it is semisimple. We now give a new interpretation of these somewhat ad-hoc arguments in the light of Fourier analysis and the Gelfand homomorphism.

By the semisimplicity of $\mathcal{A}$, the Gelfand homomorphism $\pi : \mathcal{A} \to C(\Omega(\mathcal{A}), \mathbb{C})$ is injective. (In Exercise 15.11 we have proven a bijection $S^1 \to \Omega(\mathcal{A})$. Since the Banach algebra $\mathcal{A}$ is generated by δ_1 and $\delta_1^{-1} = \delta_{-1}$, Proposition 19.5 amplifies this to a homeomorphism $\Omega(\mathcal{A}) \to S^1$.) But π is neither isometric nor surjective: Its image consists precisely of

$$\mathcal{W} = \Big\{ g \in C(S^1, \mathbb{C}) \mid \sum_{n \in \mathbb{Z}} |\widehat{g}(n)| < \infty \Big\}.$$

This is an algebra since $\mathcal{A}$ is. (To see this without reference to π, note that $\widehat{f \cdot g}(n) = (\widehat{f} \star \widehat{g})(n)$ and use the fact that $\ell^1(\mathbb{Z})$ is closed under convolution.) While $\mathcal{A}$ inherits the norm $\|\cdot\|_\infty$ from $C(S^1, \mathbb{C})$, it is not closed in this norm and π is not an isometry. The norm on $\mathcal{W}$ for which $\pi : \mathcal{A} \to \mathcal{W}$ is an isometric isomorphism is $\|g\|_{\mathcal{W}} = \sum_{n \in \mathbb{Z}} |\widehat{g}(n)|$. Since $\mathcal{W}$ is generated by the function $z \mapsto z$, it follows that $\Omega(\mathcal{W})$ consists of the point evaluations $\{\varphi_z \mid z \in S^1\}$ as for $C(S^1, \mathbb{C})$. Now the result of Exercise 15.11 is obvious since it follows from the isometric isomorphism $\pi : (\mathcal{A}, \|\cdot\|_1) \to (\mathcal{W}, \|\cdot\|_{\mathcal{W}})$.

For the $g \in \mathcal{W}$ the Fourier series converges absolutely uniformly to g, but we have proven in Section 8.3 that $C(S^1, \mathbb{C})$ has a dense subset of functions whose the Fourier series does not even converge pointwise everywhere. (Our proof was non-constructive, but as we remarked, individual examples can be produced constructively.) Functions in $C(S^1, \mathbb{C}) \backslash \mathcal{W}$ can actually be written down even more concretely: With some effort (see [109] for an exposition) the series $\sum_{n=2}^{\infty} \frac{\sin nx}{n \log n}$ can be shown to converge uniformly to some $f \in C(S^1, \mathbb{C})$, and its Fourier coefficients are not absolutely summable since $\sum_{n=2}^{\infty} (n \log n)^{-1} = \infty$. (That the convergence is not unconditional follows from the fact that it isn't at $x = \pi/2$.)

We now turn the non-surjectivity of $\pi : \ell^1(\mathbb{Z}) \to C(S^1)$ into a virtue! For $g \in C(S^1, \mathbb{C})$ define $\|g\|_{\mathcal{W}} = \sum_{n \in \mathbb{Z}} |\widehat{g}(n)|$. Thus $\mathcal{W} = \{g \in C(S^1, \mathbb{C}) \mid \|g\|_{\mathcal{W}} < \infty\}$. We have seen that the Gelfand representation of $\ell^1(\mathbb{Z})$ is an isometric isomorphism $(\ell^1(\mathbb{Z}), \|\cdot\|_1) \to (\mathcal{W}, \|\cdot\|_{\mathcal{W}})$. Now we have:

19.9 Theorem *If $g \in \mathcal{W}$ satisfies $g(z) \neq 0$ $\forall z \in S^1$ and $h \in C(S^1, \mathbb{C})$ is its multiplicative inverse $h(z) = 1/g(z)$ then $h \in \mathcal{W}$ (thus h has absolutely convergent Fourier series).*

Proof. Let $f = \pi^{-1}(g) \in \ell^1(\mathbb{Z})$. We have seen that $\Omega(\mathcal{A}) = S^1$ and $\varphi_z(f) = g(z)$ for all $z \in S^1$. Now the assumption $g(z) \neq 0\ \forall z$ implies that $0 \notin \sigma(f) = \{\varphi_z(f) \mid z \in S^1\}$, so that f is invertible in $\ell^1(\mathbb{Z})$. Thus $\pi(f) = g \in \mathcal{W}$ is invertible. Since the product on $\mathcal{W}$ is pointwise multiplication, this proves that $h = g^{-1} \in \mathcal{W}$, thus h has absolutely convergent Fourier series. ■

19.10 Remark The first proof of this theorem due to Wiener was more involved. The above proof due to Gelfand was one of the first successes of his theory of commutative C^*-algebras. But now there is a much simpler and quite definitive proof using only convergence of the geometric/Neumann series in the Banach algebra $\mathcal{W}$. See [114] or [27, Section 2.5]. □

19.11 Exercise Prove that the Gelfand homomorphism $\pi : \ell^1(\mathbb{Z}, \mathbb{C}) \to C(S^1, \mathbb{C})$ (seen above to be injective) is not bounded below.

19.3 C^*-algebras. Continuous functional calculus revisited

In discussing when the Gelfand homomorphism $\pi : \mathcal{A} \to C(\Omega(\mathcal{A}), \mathbb{F})$ is an isomorphism, we limit ourselves to the case where $\mathcal{A}$ is a C^*-algebra over $\mathbb{C}$.

19.12 Theorem (Gelfand isomorphism) *If $\mathcal{A}$ is a commutative unital C^*-algebra then the Gelfand homomorphism $\pi : \mathcal{A} \to C(\Omega(A), \mathbb{C})$ is an isometric $*$-isomorphism.*

Proof. For all $a \in \mathcal{A}, \varphi \in \Omega(\mathcal{A})$, using Lemma 17.17 we have

$$\pi(a^*)(\varphi) = \widehat{a^*}(\varphi) = \varphi(a^*) = \overline{\varphi(a)} = \varphi^*(a) = \widehat{a}(\varphi^*) = \pi(a)(\varphi^*).$$

Thus $\pi(a^*) = \pi(a)^*$, so that π is a $*$-homomorphism, and $\pi(\mathcal{A}) \subseteq C(\Omega(\mathcal{A}), \mathbb{C})$ is self-adjoint.

Since $\mathcal{A}$ is commutative, all $a \in \mathcal{A}$ are normal, thus satisfy $r(a) = \|a\|$ by Proposition 16.17(i). Together with $\|\pi(a)\| = r(a)$ for all a this implies that π is an isometry, thus injective. Since $\mathcal{A}$ is complete, this implies that the image $\pi(\mathcal{A}) \subseteq C(\Omega(\mathcal{A}), \mathbb{C})$ is complete, thus closed.

If $\varphi_1 \neq \varphi_2$ then there is an $a \in \mathcal{A}$ such that $\varphi_1(a) \neq \varphi_2(a)$, thus $\pi(a)(\varphi_1) = \widehat{a}(\varphi_1) \neq \widehat{a}(\varphi_2) = \pi(a)(\varphi_2)$. This proves that $\pi(\mathcal{A}) \subseteq C(\Omega(\mathcal{A}), \mathbb{R})$ separates the points of $\Omega(\mathcal{A})$. Since π is also unital, the Stone-Weierstrass theorem (Corollary A.39) gives $\pi(\mathcal{A}) = \overline{\pi(\mathcal{A})} = C(\Omega(\mathcal{A}), \mathbb{C})$. ■

19.13 Proposition *Let $\mathcal{A}, \mathcal{B}$ be commutative unital C^*-algebras. If $\alpha : \mathcal{A} \to \mathcal{B}$ is a unital $*$-homomorphism, define $\alpha^* : \Omega(\mathcal{B}) \to \Omega(\mathcal{A}),\ \psi \mapsto \psi \circ \alpha$. Then*

(i) α^* is continuous w.r.t. the weak-$*$ topologies on $\Omega(\mathcal{A}), \Omega(\mathcal{B})$.

(ii) If α is surjective then α^* is injective.

(iii) If α is injective then α^* is surjective.

Proof. (i) Let $\{\psi_\iota\}$ be a net in $\Omega(\mathcal{B})$ weak-$*$ convergent to ψ. Thus $\psi_\iota(b) \to \psi(b)\ \forall b \in \mathcal{B}$. Then in particular $\psi_\iota(\alpha(a)) \to \psi(\alpha(a))\ \forall a \in \mathcal{A}$. Thus $\alpha^*(\psi_\iota) = \psi_\iota \circ \alpha \stackrel{w^*}{\longrightarrow} \psi \circ \alpha = \alpha^*(\psi)$, proving continuity of α^*.

(ii) This is trivial: If $\psi, \psi' \in \Omega(\mathcal{B})$ and $\alpha^*(\psi) = \alpha^*(\psi')$ then $\psi \circ \alpha = \psi' \circ \alpha$, and surjectivity of α implies $\psi = \psi'$.

(iii) Assume $\alpha^*(\Omega(\mathcal{B})) \subsetneqq \Omega(\mathcal{A})$. By (i) and weak-$*$ compactness of $\Omega(\mathcal{B})$, $\alpha^*(\Omega(\mathcal{B}))$ is a compact, thus closed, proper subset of $\Omega(\mathcal{A})$. Since $\Omega(\mathcal{A})$ is compact Hausdorff, it is normal, and by Urysohn's Lemma there is a non-zero $f \in C(\Omega(\mathcal{A}), \mathbb{C})$ such that $f \upharpoonright \alpha^*(\Omega(\mathcal{B})) = 0$. By the Gelfand isomorphism, there is a non-zero $a \in \mathcal{A}$ such that $f = \widehat{a}$. If now $\psi \in \Omega(\mathcal{B})$ then we have $\psi \circ \alpha(a) = \widehat{a}(\psi \circ \alpha) = f(\alpha^*(\psi)) = 0$. Since this holds for all $\psi \in \Omega(\mathcal{B})$ we have $\sigma_{\mathcal{B}}(\alpha(a)) = \{0\}$, thus $\alpha(a) = 0$. But in view of $a \neq 0$ this contradicts the injectivity of α. This contradiction proves the surjectivity of α^*. ■

The result of (iii) is equivalent to the following: Every character on a unital C^*-subalgebra of a commutative C^*-algebra has an extension to a character of the larger algebra.

19.14 Remark 1. Theorem 19.12 can be strengthened to a (contravariant) equivalence of categories between the categories of commutative unital C^*-algebras and unital $*$-homomorphisms and of compact Hausdorff spaces and continuous maps.

2. With some work, the assumption of $\mathcal{A}$ having a unit can be dropped, cf. e.g. [110]. One finds that every commutative C^*-algebra is isometrically $*$-isomorphic to $C_0(X, \mathbb{C})$ for a locally compact Hausdorff space X, unique up to homeomorphism. X is compact if and only if $\mathcal{A}$ is unital. And the equivalence of categories mentioned above extends to a contravariant

equivalence between the category of locally compact Hausdorff spaces and proper maps and the category of commutative $*$-algebras and non-degenerate homomorphisms.

3. The preceding comments in a sense end the theory of commutative C^*-algebras since the latter is reduced it to general topology. But the theory of non-commutative C^*-algebras is vast, see [79, 110] for accessible introductions, and it turns out that commutative C^*-algebras are a very useful tool for studying them, as results like Exercise 16.24 and Proposition 17.6 just begin to illustrate.

4. Comparing Theorem 19.12 with the non-surjectivity of the Gelfand-Homomorphism for $(\ell^1(\mathbb{Z},\mathbb{C}),\star)$ shows that $\ell^1(\mathbb{Z},\mathbb{C})$ does not admit a norm that would make it a C^*-algebra. But $\ell^1(\mathbb{Z},\mathbb{C})$ admits a non-complete C^*-norm $\|\cdot\|'$, and completing $\ell^1(\mathbb{Z},\mathbb{C})$ w.r.t. the latter yields a C^*-algebra $C^*(\mathbb{Z})$ that is isomorphic to $C^*(U)\subseteq B(\ell^2(\mathbb{Z},\mathbb{C}))$, where $U\in B(\ell^2(\mathbb{Z},\mathbb{C}))$ is the two-sided shift unitary. One also has $C^*(\mathbb{Z})\cong C(S^1,\mathbb{C})$, thus the C^*-completion 'adds' the continuous functions with non-absolutely convergent Fourier series. □

The following is a C^*-version of Proposition 19.5:

19.15 Proposition *Let $\mathcal{B}$ be a commutative unital C^*-algebra and $b\in\mathcal{B}$ such that $\mathcal{B}=C^*(\mathbf{1},b)$. Then the map $\widehat{b}:\Omega(\mathcal{B})\to\sigma(b)$ is a homeomorphism.*

Proof. The proof is similar to that of Proposition 19.5, enriched by the following argument: If $\varphi_1(b)=\varphi_2(b)$ then by Lemma 17.17 we have $\varphi_1(b^*)=\overline{\varphi_1(b)}=\overline{\varphi_2(b)}=\varphi_2(b^*)$. Thus φ_1 and φ_2 coincide on all polynomials in b and b^*, and therefore on $\mathcal{B}$. ■

Now we have another, perhaps more conceptual but certainly less elementary, proof of the continuous functional calculus for normal elements of a C^*-algebra (Theorem 17.16):

19.16 Theorem *Let $\mathcal{A}$ be a unital C^*-algebra and $a\in\mathcal{A}$ normal. Then*

(i) There is a unique unital $*$-homomorphism $\alpha_a: C(\sigma(a),\mathbb{C})\to\mathcal{A}$ such that $\alpha_a(z)=a$, where z is the inclusion map $\sigma(a)\hookrightarrow\mathbb{C}$. As in Section 17.1, we interpret $\alpha_a(f)$ as $f(a)$.

(ii) If $f\in C(\sigma(a),\mathbb{C}))$ then $\sigma(f(a))=f(\sigma(a))$.

(ii) If $f\in C(\sigma(a),\mathbb{C}))$ and $g\in C(f(\sigma(a)),\mathbb{C})$ then $(g\circ f)(a)=g(f(a))$.

Proof. (i) Let $\mathcal{B}=C^*(\mathbf{1},a)\subseteq\mathcal{A}$ be the closed $*$-subalgebra generated by $\{\mathbf{1},a\}$. Since a is normal, $\mathcal{B}$ is a commutative unital C^*-algebra, thus by Theorem 19.12, there is an isometric $*$-isomorphism $\pi:\mathcal{B}\to C(\Omega(\mathcal{B}),\mathbb{C})$. And by Proposition 19.15 we have a homeomorphism $\widehat{a}:\Omega(\mathcal{B})\to\sigma(a)$. Now we define α_a to be the composite of the maps

$$C(\sigma(a),\mathbb{C})\stackrel{\widehat{a}^t}{\to}C(\Omega(\mathcal{B}),\mathbb{C})\stackrel{\pi^{-1}}{\to}\mathcal{B}\hookrightarrow\mathcal{A},$$

where the first map is $\widehat{a}^t: f\mapsto f\circ\widehat{a}$. It is clear that α_a is a unital homomorphism. If $z:\sigma(a)\hookrightarrow\mathbb{C}$ is the inclusion, then $\alpha_a(z)=\pi^{-1}(z\circ\widehat{a})=\pi^{-1}(\widehat{a})=a$. Any continuous unital homomorphism $\alpha: C(\sigma(a))\to\mathcal{B}$ sending 1 to $\mathbf{1}_\mathcal{A}$ and z to a coincides with α_a on the polynomials $\mathbb{C}[x]$. Since the latter are dense in $C(\sigma(a),\mathbb{C})$ by Stone-Weierstrass, we have $\alpha=\alpha_a$.

(ii) As used above, C^*-subalgebra $\mathcal{B}=C^*(\mathbf{1},a)$ is abelian and there is an isometric $*$-isomorphism $\pi:\mathcal{B}\to C(\sigma(a),\mathbb{C})$. By construction of the functional calculus we have $\pi(f(a))=f\circ\iota$, where ι is the inclusion map $\sigma(a)\hookrightarrow\mathbb{C}$. Now, with Theorem 16.19 and Exercise 13.24 we have

$$\sigma_\mathcal{A}(f(a))=\sigma_\mathcal{B}(f(a))=\sigma_{C(\sigma(a))}(\pi(f(a)))=\sigma_{C(\sigma(a))}(f\circ\iota)=f(\sigma(a)).$$

(iii) This is essentially obvious, since applying f to a and g to $f(a)$ is just composition of maps on the right hand side of the Gelfand isomorphism. ■

It should be clear that Theorem 19.12 is of fundamental conceptual importance, but most of its applications just use Theorem 19.16, which we proved in Section 17.3 in a more elementary fashion (without weak-$*$ topology and Alaoglu's theorem, and using only the classical Weierstrass theorem). Genuine applications of Theorem 19.12 are harder to find. Here is one:

19.17 Exercise Let $\mathcal{A}$ be a unital C^*-algebra and let $a, b \in \mathcal{A}$ be commuting normal elements. Prove that the absolute values (Defin. 17.11) satisfy $|a+b| \le |a| + |b|$ and $|ab| = |a|\,|b|$.

Hint: Fuglede's theorem.

(For $\mathcal{A} = B(H)$ one has results under weaker hypotheses. See e.g. [107].)

19.18 Remark If $\mathcal{A}$ is a commutative unital C^*-algebra generated by $a_1, \ldots, a_n \in \mathcal{A}$, Exercise 19.8 gives a homeomorphism $\Omega(\mathcal{A}) \cong \sigma(a_1, \ldots, a_n)$. Combining this with the Gelfand isomorphism, we see that $\mathcal{A}$ is isometrically $*$-isomorphic to $C(\sigma(a_1, \ldots, a_n), \mathbb{C})$. □

A Some more advanced topics from topology and measure theory

A.1 Unordered infinite sums

If S is a finite set, A an abelian group and $f : S \to A$ a function, it is not hard to define $\sum_{s\in S} f(S)$ (even though few textbook authors bother to do so explicitly). One choses a bijection $\alpha : \{1, 2, \ldots, \#S\} \to S$ and defines $\sum_{s\in S} f(s) = \sum_{i=1}^{\#S} f(\alpha(i))$. The only slight difficulty is proving that the result does not depend on the choice of α.

In order to define infinite sums, we need a topology on A, and we restrict to the case of functions $f : S \to V$, where $(V, \|\cdot\|)$ is a normed space. Many authors of introductory texts (for a nice exception see [163, Vol. I, Section 8.2]) consider only those countable sums known as series, but for our purposes this is inadequate.

A.1 Definition *Let S be a set, $(V, \|\cdot\|)$ a normed space and $f : S \to V$ a function. We say that $\sum_{s\in S} f(s)$ exists or converges (or: f is summable over S) with sum $x \in V$ if for every $\varepsilon > 0$ there is a finite subset $T \subseteq S$ such that $\|x - \sum_{s\in U} f(s)\| < \varepsilon$ holds whenever $T \subseteq U \subseteq S$ with U finite.*

In many cases, the above will be applied to $V = \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $\|\cdot\| = |\cdot|$.

This notion of summation has some useful properties:

A.2 Proposition *Let V be a normed space over $\mathbb{F}$ and $f, g : V \to \mathbb{F}$.*

(i) If $\sum_{s\in S} f(s)$ exists then the sum $x \in V$ is uniquely determined.

(ii) If $\sum_{s\in S} f(s) = x$ and $\sum_{s\in S} g(s) = y$ then $\sum_{s\in S}(cf(s) + dg(s)) = cx + dy$ for all $c, d \in \mathbb{F}$.

(iii) If $f : S \to [0, \infty) \subseteq \mathbb{R}$ then $\sum_{s\in S} f(s)$ exists if and only if $\sup\{\sum_{t\in T} f(t) \mid T \subseteq S \text{ finite}\} < \infty$, in which case the two expressions coincide. These equivalent conditions imply that the set $\{s \in S \mid f(s) \neq 0\}$ is at most countable.

(iv) If $(V, \|\cdot\|)$ is complete and $\sum_{s\in S} \|f(s)\| < \infty$ then $\sum_{s\in S} f(s)$ exists, and $\|\sum_{s\in S} f(s)\| \le \sum_{s\in S} \|f(s)\|$.

(v) If $f : S \to \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is such that $\sum_{s\in S} f(s)$ exists then $\sum_{s\in S} |f(s)|$ exists, i.e. is finite.

The proofs of (i) and (ii) are straightforward and similar to those for the analogous statements about series. The equivalence in (iii) follows from monotonicity of the map $P_{\text{fin}}(S) \to [0, \infty)$, $T \mapsto \sum_{t\in T} f(t)$. If $\sum_{s\in S} |f(s)| < \infty$ then it follows that for every $\varepsilon > 0$ there are at most finitely many $s \in S$ such that $|f(s)| \ge \varepsilon$. In particular, for every $n \in \mathbb{N}$ the set $S_n = \{s \in S \mid |f(s)| \ge 1/n\}$ is finite. Since a countable union of finite sets is countable, we have countability of $\{s \in S \mid f(s) \ne 0\} = \bigcup_{n=1}^{\infty} S_n$. The proof of (iv) combines the argument in Proposition 3.15(iii) with Lemma A.14 below.

Statement (v) may be surprising at first sight since the analogous statement for series is false. Roughly, the reason is that our definition of $\sum_{s\in S} f(s)$ imposes no ordering on S, while, by a classical result of Riemann, the sum of a convergent series $\sum_{n=1}^{\infty} f(n)$ is invariant under reordering of the terms *only* if the series converges absolutely. The rigorous proof of (v), found e.g. in [12] or [108, Proposition 5.1.28], does not appeal to Riemann's result, but uses similar ideas. For $S = \mathbb{N}$ this is Proposition 3.16, the proof for general S being similar.

In discussing the spaces $\ell^p(S, \mathbb{F})$, the following (easy special case of Lebesgue's dominated convergence theorem) is useful:

A.3 Proposition (Discrete case of dominated convergence theorem) *Let S be a set and $\{f_n\}_{n\in\mathbb{N}}$ functions $S \to \mathbb{C}$. Assume that*

1. *For each $s \in S$, the limit $\lim_{n\to\infty} f_n(s)$ exists. Define $h : S \to \mathbb{C}, s \mapsto \lim_{n\to\infty} f_n(s)$.*
2. *There exists a function $g : S \to [0, \infty)$ such that $\sum_{s\in S} g(s) < \infty$ and $|f_n(s)| \le g(s)\ \forall s \in S$.*

Then

(i) $\sum_{s\in S} f_n(s)$ converges for each $n \in \mathbb{N}$. So does $\sum_{s\in S} h(s)$.

(ii) $\lim_{n\to\infty} \sum_{s\in S} f_n(s) = \sum_{s\in S} h(s)$. (Thus limit and summation can be interchanged.)

Proof. (i) Assumption 1. gives $|h(s)| \le g(s)\ \forall s$. Now assumption 2. implies convergence of $\sum_s h(s)$ and of $\sum_s f_n(s)$ for all n.

(ii) Let $\varepsilon > 0$. Since $\sum_s g(s) < \infty$, there is a finite subset $T \subseteq S$ such that $\sum_{s\in S\backslash T} g(s) < \frac{\varepsilon}{4}$. For each $t \in T$ there is an $n_t \in \mathbb{N}$ such that $n \ge n_t \Rightarrow |f_n(t) - h(t)| < \frac{\varepsilon}{2\#T}$. Put $n_0 = \max_{t\in T} n_t$. If $n \ge n_0$ then

$$\left|\sum_{s\in S} f_n(s) - \sum_{s\in S} h(s)\right| \le \left|\sum_{s\in T} f_n(s) - \sum_{s\in T} h(s)\right| + \left|\sum_{s\in S\backslash T} f_n(s) - \sum_{s\in S\backslash T} h(s)\right|.$$

The first term on the r.h.s. is bounded by

$$\sum_{s\in T} |f_n(s) - h(s)| \le \#T \cdot \frac{\varepsilon}{2\#T} = \frac{\varepsilon}{2}$$

due to the definition of n_0 and $n \ge n_0 \ge n_t$. And the second is bounded by

$$\sum_{S\backslash T} (|f_n(s)| + |h(s)|) \le 2 \sum_{S\backslash T} g(s) \le \frac{\varepsilon}{2},$$

where we used that g dominates $|f_n|$ and $|h|$, as well as the choice of T. Putting the two estimates together gives $n \ge n_0 \Rightarrow |\sum_{s\in S} f_n(s) - \sum_{s\in S} h(s)| \le \varepsilon$, completing the proof. ■

A.2 More on unconditional convergence of series

In the case $S = \mathbb{N}$ and $f(n) = x_n$, there is a connection between the summability of Definition A.1 and the notion of unconditionally convergent series:

A.4 THEOREM *Let $(V, \|\cdot\|)$ be a Banach space and $\{x_n\}_{n\in\mathbb{N}} \subset V$. Then the following are equivalent:*

(i) $\sum_{n\in\mathbb{N}} x_n$ *exists in the sense of Definition A.1.*

(ii) $\sum_{n=1}^{\infty} x_n$ *is unconditionally convergent, the sums of all rearrangements being equal.*

(iii) $\sum_{n=1}^{\infty} x_n$ *is unconditionally convergent.*

(iv) For every $\varepsilon > 0$ there exists a finite $S \subseteq \mathbb{N}$ such that for every finite subset $T \subset \mathbb{N}\backslash S$ we have $\|\sum_{t\in T} x_t\| < \varepsilon$.

(v) $\displaystyle\lim_{N\to\infty} \sup_{\varphi\in V^*_{\le 1}} \sum_{k=N+1}^{\infty} |\varphi(x_k)| = 0.$

(vi) $\sum_{n=1}^{\infty} c_n x_n$ *is convergent for all bounded sequences $\{c_n\}$ in $\mathbb{F}$. (The convergence then is unconditional.)*

(vii) $\sum_{k=1}^{\infty} x_{n_k}$ *converges for all $n_1 < n_2 < \cdots$. (Subseries convergence)*

Proof. (i)⇒(ii) Let $\sum_{n\in\mathbb{N}} x_n = x$. If now σ is any permutation of $\mathbb{N}$ and $\varepsilon > 0$ then by assumption there is a finite subset $T \subseteq \mathbb{N}$ such that $\|x - \sum_{n\in U} x_n\| < \varepsilon$ for every finite $U \subset \mathbb{N}$ containing T. Since $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection, there exists n_0 such that $n \ge n_0$ implies $T \subseteq \{\sigma(1), \ldots, \sigma(n)\}$. (We can take $n_0 = \max\{\sigma^{-1}(k) \mid k \in T\}$.) Thus $\|x - \sum_{k=1}^{n} x_{\sigma(k)}\| < \varepsilon$. This proves that all rearranged sums $\sum_{k=1}^{\infty} x_{\sigma(k)}$ converge to x.

(ii)⇒(iii) Trivial.

(iii)⇒(iv) Assume (iv) does not hold. By elementary logic this means that there is an $\varepsilon > 0$ such that for every finite $S \subseteq \mathbb{N}$ there exists a finite $T \subseteq \mathbb{N}\backslash S$ with $\|\sum_{t\in T} x_t\| \ge \varepsilon$. Using this we can construct a sequence $S_1, S_2, \ldots$ of mutually disjoint finite subsets $S_k \subseteq \mathbb{N}$ such that $\|\sum_{s\in S_k} x_s\| \ge \varepsilon$ for each k. Now we can find a permutation σ of $\mathbb{N}$ and a sequence $n_1 < n_2 < \ldots$ such that $S_k = \sigma(\{n_k, n_k + 1, \ldots, n_k + \#S_k - 1\})$ for all k. Thus for each k we have $\|\sum_{n=n_k}^{n_k+\#S_k-1} x_{\sigma(n)}\| = \|\sum_{s\in S_k} x_s\| \ge \varepsilon$, so that the series $\sum_{n=1}^{\infty} x_{\sigma(n)}$ is not Cauchy[99], thus divergent, contradicting (iii).

(iv)⇒(i) If $\varepsilon > 0$, let S be as in (i). If now $T, U \subseteq \mathbb{N}$ are finite sets containing S then $\|\sum_{t\in T} x_t - \sum_{u\in U} x_u\| \le \|\sum_{t\in T\backslash S} x_t\| + \|\sum_{u\in U\backslash S} x_u\| \le 2\varepsilon$. Thus $P_{\text{fin}}(\mathbb{N}) \to \mathbb{F}, T \mapsto \sum_{t\in T} x_t$ is a Cauchy net (Definition A.13) and therefore converges by Lemma A.14.

(iv)⇒(v) Let $\varepsilon > 0$ and $S \subseteq \mathbb{N}$ as provided correspondingly by assumption (iv). Put $N = \max(S)$. It then follows that $\|\sum_{t\in T} x_t\| < \varepsilon$ for each finite $T \subseteq \mathbb{N}$ with $\min(T) > N$. If $\varphi \in V^*$ and $L \ge K > N$, put

$$\begin{aligned} F^+ &= \{n \in \mathbb{N} \mid K \le n \le L,\ \operatorname{Re}\varphi(x_n) \ge 0\}, \\ F^- &= \{n \in \mathbb{N} \mid K \le n \le L,\ \operatorname{Re}\varphi(x_n) < 0\}. \end{aligned}$$

Clearly $\min(F^+) \ge K > N$, so that $\|\sum_{n\in F^+} x_n\| < \varepsilon$. Now for every $\varphi \in V^*_{\le 1}$ we have

$$\sum_{n\in F^+} |\operatorname{Re}\varphi(x_n)| = \sum_{n\in F^+} \operatorname{Re}\varphi(x_n) = \operatorname{Re}\varphi\Big(\sum_{n\in F^+} x_n\Big) \le \Big|\varphi\Big(\sum_{n\in F^+} x_n\Big)\Big| \le \|\varphi\| \Big\| \sum_{n\in F^+} x_n \Big\| < \varepsilon.$$

[99] A series $\sum_{n=1}^{\infty} x_n$ is Cauchy if the sequence $\{S_n\}$ of partial sums is Cauchy.

Essentially the same argument holds for F^-, so that $\sum_{n=K}^L |\mathrm{Re}\,\varphi(x_n)| < 2\varepsilon$. If $\mathbb{F}=\mathbb{C}$ a similar argument gives $\sum_{n=K}^L |\mathrm{Im}\,\varphi(x_n)| < 2\varepsilon$. With $|z| \le |\mathrm{Re}\,z| + |\mathrm{Im}\,z|$ we conclude $\sum_{n=K}^L |\varphi(x_n)| < 4\varepsilon$. Taking $L\to\infty$ gives $\sum_{n=K}^\infty |\varphi(x_n)| \le 4\varepsilon$. Since this holds for all $\varphi \in V^*_{\le 1}$, (v) follows.

(v)⇒(vi) We may clearly assume $|c_n| \le 1$ for all n. Proposition 9.9(i) implies for every $y\in V$ that $\|y\| = \sup_{\varphi\in V^*_{\le 1}} |\varphi(y)|$. Thus for $M > N$ we have

$$\begin{aligned}\Big\|\sum_{k=N+1}^M c_k x_k\Big\| &= \sup_{\varphi\in V^*_{\le 1}} \Big|\varphi\Big(\sum_{k=N+1}^M c_k x_k\Big)\Big| = \sup_{\varphi\in V^*_{\le 1}} \Big|\sum_{k=N+1}^M c_k \varphi(x_k)\Big| \\ &\le \sup_{\varphi\in V^*_{\le 1}} \sum_{k=N+1}^M |c_k|\,|\varphi(x_k)| \le \sup_{\varphi\in V^*_{\le 1}} \sum_{k=N+1}^M |\varphi(x_k)| \le \sup_{\varphi\in V^*_{\le 1}} \sum_{k=N+1}^\infty |\varphi(x_k)|.\end{aligned}$$

Since the rightmost expression tends to zero as $N\to\infty$ by (v), it follows that the series $\sum_{k=1}^\infty c_k x_k$ is Cauchy, thus convergent.

By the same proof, $\sum_{n=1}^\infty c_n d_n x_n$ converges for all choices of $\{d_n\}$ in $\{0,1\}^{\mathbb{N}}$. Thus the convergence of $\sum_{n=1}^\infty c_n x_n$ is unconditional by (vi)⇒(iii).

(vi)⇒(vii) This follows by rewriting $\sum_{k=1}^\infty x_{n_k}$ as $\sum_{n=1}^\infty c_n x_n$, where $c_n \in \{0,1\}$.

(vii)⇒(iv) Assume (iv) does not hold. Arguing as in the proof of (iii)⇒(iv) we have an infinite sequence $\{S_k\}$ of mutually disjoint finite subsets of $\mathbb{N}$ such that $\|\sum_{n\in S_k} x_n\| \ge \varepsilon$ for all k. It is clear that we can choose the S_k in such a way that $\max(S_k) < \min(S_{k+1})$ for all k. Putting $S = \bigcup_k S_k$ and $c_n = \chi_S(n)$, with $N_k = \#\bigcup_{i=1}^{k-1} S_i$ we have $\|\sum_{n=N_k+1}^{N_{k+1}} c_n x_n\| = \|\sum_{n\in S_k} x_n\| \ge \varepsilon$. Thus the series $\sum_{n=1}^\infty c_n x_n$ diverges since it is not Cauchy, contradicting (vii). ■

A.5 Remark 1. Statement (vi) expresses more clearly than all the others that unconditional convergence of a series does not rely on 'cancellations' between the summands, so that the convergence is not affected by reordering or omission of any number of summands.

2. Our proof of the hardest implication (iv)⇒(vi) was found in [74]. It is perhaps the most elementary one, using only Hahn-Banach. There are many other interesting proofs, see e.g. [102, 4.2.6-4.2.8], [77, Vol. 1, Proposition 4.1.5], [97, Vol. 1, Theorem II.7].

3. If the series $\sum_{n=1}^\infty x_n$ in V is unconditionally convergent and $\varphi\in V^*$ then $\sum_{n=1}^\infty \varphi(x_n)$ converges unconditionally, thus absolutely. Thus $\sum_{n=1}^\infty |\varphi(x_n)| < \infty\ \forall\varphi\in V^*$. A series with this property is called weakly unconditionally Cauchy (WUC). But the WUC property does not even imply conditional convergence, as is illustrated by series $\sum_{n=1}^\infty \delta_n$ in c_0, which is easily shown to be WUC using $c_0^* \cong \ell^1$. This is essentially the only counterexample: Bessaga and Pełczyński[100] proved that every WUC series in a Banach space V converges unconditionally if and only if V has no subspace isomorphic to c_0. Cf. e.g. [98, Theorem 2.e.4], [1, Theorem 2.4.11]. (This is similar to Rosenthal's ℓ^1-theorem mentioned in footnote 62, but much easier.)

4. While the WUC property of a series is weaker than unconditional convergence, weak-topology characterizations of unconditional convergence do exist. One of them is statement (v), being a uniform (in $\varphi\in V^*_{\le 1}$) version of the statement $\lim_{N\to\infty}\sum_{n=N+1}^\infty |\varphi(x_n)| \to 0$ that clearly follows from WUC. On the other hand, it is clear that unconditional convergence implies weak convergence of all subseries. The latter property of a series is somewhat stronger than being WUC, and indeed by the Orlicz[101]-Pettis theorem, cf. e.g. [1, Theorem 2.4.14], [97, Vol. 1, Theorem II.3], we can add the following to the list in Theorem 19.16:

(viii): Every subseries $\sum_{i=1}^\infty x_{n_i}$, where $n_1 < n_2 < \cdots$, converges weakly. □

[100] Czeslaw Bessaga (1932-2021), Aleksander Pełczyński (1932-2012). Polish functional analysts. Both were students of S. Mazur.

[101] Władysław Orlicz (1903-1990). Polish functional analyst and topologist. Also known for O. spaces.

A.6 Exercise Use Theorem A.4 to give a high-brow proof of Proposition 3.16.

A.3 Nets

The Definition A.1 of unordered sums is an instance of a much more general notion, the convergence of nets.

A.7 Definition *A directed set is a set I equipped with a binary relation $\leq$ on I satisfying*

1. *$a \leq a$ for each $a \in I$ (reflexivity).*
2. *If $a \leq b$ and $b \leq c$ for $a, b, c \in I$ then $a \leq c$ (transitivity).*
3. *For any $a, b \in I$ there exists a $c \in I$ such that $a \leq c$ and $b \leq c$ (directedness).*

A.8 Remark If only 1. and 2. hold, $(I, \leq)$ is called a pre-ordered set. Some authors, as e.g. [101], require in addition that $a \leq b$ and $b \leq a$ together imply $a = b$ (antisymmetry). Recall that a pre-ordered set with this property is called partially ordered. But the antisymmetry is an unnatural assumption in this context and is never used. □

A.9 Example 1. Every totally ordered set $(X, \leq)$ is a directed set. Only the directedness needs to be shown, and it follows by taking $c = \max(a, b)$. In particular $\mathbb{N}$ is a directed set with its natural total ordering.

2. If S is a set then the power set $I = P(S)$ with its natural partial ordering is directed: For the directedness, put $c = a \cup b$. The same works for the set $P_{\text{fin}}(S)$ of finite subsets of S, which appeared in the definition of unordered sums.

3. If (X, τ) is a topological space and $x \in X$, let $\mathcal{U}_x$ be the set of open neighborhoods of x. Now for $U, V \in \mathcal{U}_x$, define $U \leq V \Leftrightarrow U \supseteq V$, thus we take the reversed ordering. Then $(\mathcal{U}_x, \leq)$ is directed with $c = a \cap b$.

A.10 Definition *If X is a set, a net*[102] *in X is a map $I \to X$, $\iota \mapsto x_\iota$, where $(I, \leq)$ is a directed set.*

If (X, τ) is a topological space, a net $\{x_\iota\}_{\iota \in I}$ in X converges to $z \in X$ if for every open neighborhood U of z there is a $\iota_0 \in I$ such that $\iota \geq \iota_0 \;\Rightarrow\; x_\iota \in U$.

When this holds, we write $x_\iota \to z$ or $\lim_{\iota \in I} x_\iota = z$. (The second notation should only be used if X is Hausdorff, since this property is equivalent to uniqueness of limits of nets.)

A.11 Remark 1. With $I = \mathbb{N}$ and $\leq$ the natural total ordering, a net indexed by I just is a sequence, and this net converges if and only if the sequence does.

2. Unordered summation is a special case of a net limit: If S is any set, let I be the set of finite subsets of S and let $\leq$ be the ordinary (partial) ordering of subsets of S. If $T, U \in I$ let $V = T \cup U$. Clearly $T \leq V, U \leq V$, showing that $(I, \leq)$ is a directed set. (This is the same as Example A.9.2, except that now we only look at finite subsets of S.) Now given $f : S \to \mathbb{F}$, for every $T \in I$, thus every finite $T \subseteq S$, we can clearly define $\sum_{t \in T} f(t)$. Now

$$\sum_{s \in S} f(s) = \lim_{T \in I} \sum_{t \in T} f(t),$$

where the sum exists if and only if the limit exists. □

[102] Nets were invented by the American mathematicians Eliakim Hastings Moore (1862-1932) and his student Herman L. Smith (1892-1950). Moore made many contributions to many areas of mathematics.

Why nets? The reason is that sequences are totally inadequate for the study of topological spaces that do not satisfy the first countability axiom.[103] Given a metric space X and a subset $Y \subseteq X$, one proves that $x \in \overline{Y}$ if and only if there is a sequence $\{y_n\}$ in Y converging to x, but for general topological spaces this is false. Similarly, the statement that a function $f : X \to Y$ is continuous at $x \in X$ if and only if $f(x_n) \to f(x)$ for every sequence $\{x_n\}$ converging to x is true for metric spaces, but false in general! (It is instructive to work out counterexamples.)

On the other hand:

A.12 Proposition *1. Let X be a topological space and $Y \subseteq X$. If $\{y_\iota\}$ is a net in Y that converges to $x \in X$ then $x \in \overline{Y}$.*

2. Let X be a topological space and $Y \subseteq X$. Then for every $x \in \overline{Y}$ there exists a net $\{y_\iota\}$ in Y such that $y_\iota \to x$.

3. A topological space X is Hausdorff if and only if there exists no net $\{x_\iota\}$ in X that converges to two different points of X.

4. If X, Y are topological spaces, $f : X \to Y$ a function, and $x \in X$, then f is continuous at x if and only if $f(x_\iota) \to f(x)$ for every net $\{x_\iota\}$ in X converging to x.

For proofs see any decent book on topology or [108].

If (X, d) is a metric space, the problems with sequences mentioned above do not arise. Nevertheless, there are situations where the use of nets in X is useful, as in the proof of Theorem 5.42 and 5.45 where we considered nets indexed by the finite subsets of an ONB E. In this case one wants:

A.13 Definition *A net $\{x_\iota\}$, indexed by a directed set $(I, \le)$, in a metric space (X, d) is a Cauchy net if for every $\varepsilon > 0$ there is a $\iota_0 \in I$ such that $\iota, \iota' \ge \iota_0 \Rightarrow d(x_\iota, x_{\iota'}) < \varepsilon$.*

A.14 Lemma *In a complete metric space every Cauchy net converges.*

Proof. Let $\{x_\iota\}_{\iota \in I}$ be Cauchy. Then for every $n \in \mathbb{N}$ there is a $\iota_n \in I$ such that $\iota, \iota' \ge \iota_n \Rightarrow d(x_\iota, x_{\iota'}) < 1/n$. We can arrange that in addition $\iota_1 \le \iota_2 \le \cdots$ (using directedness to replace ι_2 by some ι_2' larger than ι_1 and ι_2 etc.). Put $y_n = x_{\iota_n}$ for $n \in \mathbb{N}$. If $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that $1/n_0 < \varepsilon$ then for $n, m \ge n_0$ we have $\iota_m, \iota_n \ge \iota_{n_0}$, so that $d(y_n, y_m) = d(x_{\iota_n}, x_{\iota_m}) < 1/n_0 < \varepsilon$. Thus $\{y_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence, which by completeness of (X, d) converges to some $y \in X$ with $d(y, y_n) \le 1/n \ \forall n$. If now $\varepsilon > 0$, pick $n \in \mathbb{N}$ with $1/n < \varepsilon/2$. Now for all $\iota \ge \iota_n$ we have $d(x_\iota, y) \le d(x_\iota, x_{\iota_n}) + d(x_{\iota_n}, y) < \frac{1}{n} + \frac{1}{n} < \varepsilon$, thus $x_\iota \to y$. ■

A.15 Remark In the above proof we cannot argue by saying that there is a sequence $\iota_1 \le \iota_2 \le \cdots$ such that for every $\iota \in I$ there is an n with $\iota \le \iota_n$. If this was true, we could replace the net by a sequence in the first place! □

A.4 Reminder of the choice axioms and Zorn's lemma

A.16 Definition *The Axiom of Choice (AC) is any of the following statements, which are easily shown to be equivalent:*

- *If $f : X \to Y$ is a surjective function then there exists a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.*

[103] Unfortunately many introductory books and courses sweep this problem under the rug and don't even mention nets (or their alternatives like filters).

- *If X is a set, there exists a function $s : P(X)\backslash\{\emptyset\} \to X$ such that $s(Y) \in Y$ for each $Y \in P(X)\backslash\{\emptyset\}$, i.e. $\emptyset \neq Y \subseteq X$.*
- *If $\{X_i\}_{i\in I}$ is a family of non-empty sets then $\prod_{i\in I} X_i \neq \emptyset$. Concretely, there exists a map $f : I \to \bigcup_{j\in I} X_j$ such that $f(i) \in X_i \ \forall i \in I$.*

A.17 Definition *Let $(X, \leq)$ be a partially ordered set. Then*

- *$m \in X$ is called a maximal element if $y \in X, y \geq m$ implies $y = m$.*
- *$u \in X$ is called an upper bound for $Y \subseteq X$ if $x \leq u$ holds for each $u \in y$. If $u \in Y$ then it is called largest element of Y (which is unique).*

A.18 Theorem *Given the Zermelo-Frenkel axioms of set theory, the Axiom of Choice is equivalent to Zorn's lemma, which says: If $(X, \leq)$ is a non-empty partially ordered set such that every totally ordered subset $Y \subseteq X$ has an upper bound then X has a maximal element.*

A.19 Definition *The Axiom of Countable Choice (AC_ω) is the first (or third) of the above versions of AC with the restriction that Y (respectively I) be at most countable.*

Many iterative constructions, as in the proof of Urysohn's lemma or of Lemma 7.3, require countably many choices where, however, the n-th choice must take into account the preceding ones. For this we need an axiom that is stronger than AC_ω:

A.20 Definition *The Axiom of Countable Dependent Choice (DC_ω) is the following: If X is a set and $R \subseteq X \times X$ is such that for every $x \in X$ there is a $y \in X$ such that $(x, y) \in R$ then there is a sequence $\{x_n\}_{n\in\mathbb{N}}$ in X such that $(x_n, x_{n+1}) \in R$ for all $n \in \mathbb{N}$.*

A.21 Remark 1. Like the other choice axioms, DC_ω is often used without any comment.

2. It is easy to prove $AC \Rightarrow DC_\omega \Rightarrow AC_\omega$. The converse implications have been proven false by constructing models of ZF set theory satisfying, say, AC_ω but not DC_ω. □

A.5 Baire's theorem

To provide context, we begin with some simple considerations. If Y_1, Y_2 are dense subsets of a topological space X, it does not follow that $Y_1 \cap Y_2$ is dense: Consider $X = \mathbb{R}$ and the dense sets $Y_1 = \mathbb{Q}, Y_2 = \mathbb{R}\backslash\mathbb{Q}$, for which $Y_1 \cap Y_2 = \emptyset$. But we have:

A.22 Lemma *Let (X, τ) be a topological space with $X \neq \emptyset$.*

(i) $Y \subseteq X$ is dense $\Leftrightarrow$ $X\backslash Y$ has empty interior $\Leftrightarrow$ $Y \cap W \neq \emptyset$ whenever $\emptyset \neq W \in \tau$.

(ii) If $Y_1, Y_2 \subseteq X$ are dense and Y_1 is open then $Y_1 \cap Y_2$ is dense.

(iii) If $Y_1, \ldots, Y_n \subseteq X$ are dense open subsets then $Y_1 \cap \cdots \cap Y_n$ is dense (and open).

Proof. (i) Easy exercise. (ii) Let $W \subseteq X$ be open and non-empty. Then by (i) and density of Y_1 the open set $W' = W \cap Y_1$ is non-empty. By (i) and density of Y_2 we have $W \cap Y_1 \cap Y_2 = W' \cap Y_2 \neq \emptyset$, so that using (i) once more we conclude that $Y_1 \cap Y_2$ is dense. (iii) Follows from (ii) by induction. ■

Saying something about infinite intersections requires more work and assumptions, as in:

A.23 Theorem (Baire) [104] *Let (X,d) be a complete metric space and $\{U_n\}_{n\in\mathbb{N}}$ a countable family of dense open subsets. Then $\bigcap_{n=1}^\infty U_n$ is dense in X.*

Proof. Let $W \subseteq X$ be open and non-empty. Since U_1 is dense, $W \cap U_1 \neq \emptyset$ by Lemma A.22, so we can pick $x_1 \in W \cap U_1$. Since $W \cap U_1$ is open, we can choose $\varepsilon_1 > 0$ such that $\overline{B(x_1, \varepsilon_1)} \subseteq W \cap U_1$. We may also assume $\varepsilon_1 < 1$. Since U_2 is dense, $U_2 \cap \overline{B(x_1, \varepsilon_1)} \neq \emptyset$ and we pick $x_2 \in U_2 \cap B(x_1, \varepsilon_1)$. By openness, we can pick $\varepsilon_2 \in (0, 1/2)$ such that $\overline{B(x_2, \varepsilon_2)} \subseteq U_2 \cap B(x_1, \varepsilon_1)$. Continuing this iteratively, we find points x_n and $\varepsilon_n \in (0, 1/n)$ such that $\overline{B(x_n, \varepsilon_n)} \subseteq U_n \cap B(x_{n-1}, \varepsilon_{n-1})\ \forall n$. If $i > n$ and $j > n$ we have by construction that $x_i, x_j \in \overline{B(x_n, \varepsilon_n)}$ and thus $d(x_i, x_j) \leq 2\varepsilon_n < 2/n$. Thus $\{x_n\}$ is a Cauchy sequence, and by completeness of (X, d) it converges to some $z \in X$. Since $n > k$ implies $x_n \in \overline{B(x_k, \varepsilon_k)}$, the limit z is contained in $\overline{B(x_k, \varepsilon_k)}$ for each k, thus

$$z \in \bigcap_n \overline{B(x_n, \varepsilon_n)} \subseteq W \cap \bigcap_n U_n,$$

so that $W \cap \bigcap_n U_n$ is non-empty. Since W was an arbitrary non-empty open set, Lemma A.22 gives that $\bigcap_n U_n$ is dense. ■

The following dual and equivalent reformulation is also useful:

A.24 Corollary *Let (X,d) be a complete metric space and $\{C_n\}_{n\in\mathbb{N}}$ a countable family of closed subsets with empty interior. Then $\bigcup_{n=1}^\infty C_n$ has empty interior.*

Proof. The sets $U_n = X \backslash C_n,\ n \in \mathbb{N}$, are open and $\overline{U_n} = \overline{X \backslash C_n} = X \backslash C_n^0 = X$ since the interiors C_n^0 are empty. Thus the U_n are dense so that $\bigcap_n U_n$ is dense by Baire's theorem, thus $\overline{\bigcap_n X \backslash C_n} = \overline{\bigcap_n U_n} = X$. Thus with $X \backslash \overline{Y} = (X \backslash Y)^0$ we have $(\bigcup_n C_n)^0 = (X \backslash \bigcap_n (X \backslash C_n))^0 = X \backslash \overline{\bigcap_n X \backslash C_n} = \emptyset$, i.e. $\bigcup_n C_n$ has empty interior. ■

A.25 Remark 1. There are many other ways of stating Baire's theorem, but most of the alternative versions introduce additional terminology (nowhere dense sets, meager sets, sets of first or second category, etc.) that obscures the matter unnecessarily.

2. An intersection $\bigcap_n U_n$ of a countable family $\{U_n\}_{n\in\mathbb{N}}$ of open sets is called a G_δ-set. (And a countable union of closed sets is called F_σ-set.)

3. The proof implicitly used the axiom DC_ω of countable dependent choice. (Making this explicit is an instructive but tedious exercise.) Remarkably, the (Zermelo-Frenkel) axioms of set theory (without any choice axiom) combined with Baire's theorem imply DC_ω, cf. [17].

4. Some results customarily proven using Baire's theorem can alternatively be proven without it. But in most cases, such alternative proofs will also use the axiom DC_ω and therefore not be better from a foundational (reverse mathematics) point of view. See also Remark 8.3.2. □

A typical application of Baire's theorem is the following (for a proof see, e.g., [108]):

A.26 Theorem *There is a $\|\cdot\|_\infty$-dense G_δ-set $F \subseteq C([0,1],\mathbb{R})$ such that every $f \in F$ is nowhere differentiable.*

Note that a single function $f \in C([0,1],\mathbb{R})$ that is nowhere differentiable can be written down quite explicitly and constructively, for example $f(x) = \sum_{n=1}^\infty 2^{-n}\cos(2^n x)$. But for proving that such functions are dense one needs Baire's theorem (or something related).

[104] René-Louis Baire (1874-1932). French mathematician, proved this for $\mathbb{R}^n$ in his 1899 doctoral thesis. The generalization is due to Hausdorff (1914).

A.6 On $C(X,\mathbb{F})$

We recall a few facts from general topology:

A.27 Definition *A topological space (X,τ) is called*

- T_1 *if $\{x\}\subseteq X$ is closed for each $x\in X$.*
- T_2 *or Hausdorff if for any $x,y\in X$, $x\neq y$ there are disjoint open U,V with $x\in U$, $y\in V$.*
- T_4 *or normal if it is T_1 and for any two disjoint closed sets $C,D\subseteq X$ there are disjoint open sets $U,V\subseteq X$ with $C\subseteq U$, $D\subseteq V$.*

It is immediate that T_4 implies T_1 and T_2, and for $T_2\Rightarrow T_1$ it suffices to fix x and pick for each $y\neq x$ an open U_y with $y\in U_y\not\ni x$. Then $X\backslash\{x\}=X\backslash\bigcup_{y\neq x}U_y$ is open.

One easily checks that for a T_1 space, the following is equivalent to normality: Whenever $C\subseteq U$ with C closed and U open, there is an open V such that $C\subseteq V\subseteq\overline{V}\subseteq U$.

A.28 Proposition *Every compact Hausdorff space is normal.*

A.29 Proposition (Urysohn's Lemma) *If X is a normal space and $C,D\subseteq X$ there are disjoint closed sets, there exists $f\in C(X,[0,1])$ such that $f\restriction C=0$ and $f\restriction D=1$.*

For proofs see e.g. [142].

A.30 Lemma *If X is a compact topological space (not necessarily Hausdorff), $C(X,\mathbb{F})$ with norm $\|f\|=\sup_{x\in X}|f(x)|$ is a Banach space.*

Proof. By compactness of X, every $f\in C(X,\mathbb{F})$ is bounded, thus has $\|f\|<\infty$. That the norm axioms are satisfied is easy enough. If $\{f_n\}\subset C(X,\mathbb{F})$ is a Cauchy sequence, so is $\{f_n(x)\}$ for each $x\in X$. Since $\mathbb{F}$ is complete, $g(x)=\lim_{n\to\infty}f_n(x)$ exists for each x. Let $\varepsilon>0$ and n_0 be such that $n,m\geq n_0$ implies $\sup_{x\in X}|f_n(x)-f_m(x)|=\|f_n-f_m\|<\varepsilon$. Taking $m\to\infty$ we obtain $\|f_n-g\|=\sup_{x\in X}|f_n(x)-g(x)|=\leq\varepsilon$ for all $n\geq n_0$. Thus the convergence $f_n(x)\to g(x)$ is uniform in x. This implies $g\in C(X,\mathbb{F})$, as shown in topology. (If $x\in X$ and $\varepsilon>0$, pick n such that $\|g-f_n\|<\varepsilon/3$. By continuity of f_n there is an open neighborhood $U\subseteq X$ of x such that $|f(x)-f(y)|<\varepsilon/3$ for all $y\in U$. Then for all $y\in U$ we have

$$|g(x)-g(y)|\leq|g(x)-f_n(x)|+|f_n(x)-f_n(y)|+|f_n(y)-g(y)|\leq\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon,$$

so that f is continuous at x. Since this holds works for all $x\in X$, g is continuous.) Thus every Cauchy sequence in $C(X,\mathbb{F})$ converges to an element of $C(X,\mathbb{F})$, proving completeness. ■

A.6.1 Tietze's extension theorem

While Urysohn's lemma belongs to point set topology, the following result can be given [61] a functional analytic twist:

A.31 Theorem (Tietze-Urysohn extension theorem) [105] *Let (X,τ) be a normal (T_4) topological space, $Y\subseteq X$ closed and $f\in C_b(Y,\mathbb{R})$. Then there exists $\widehat{f}\in C_b(X,\mathbb{R})$ such that $\widehat{f}_{|Y}=f$ and $\|\widehat{f}\|=\|f\|$.*

[105] H. F. F. Tietze (1880-1964), Austrian mathematician. He proved this for metric spaces (for which Urysohn's lemma is a triviality). The generalization to normal spaces is due to Urysohn.

Proof. Let $f \in C_b(Y, \mathbb{R})$, where we may assume $\|f\| = 1$, so that $f(Y) \subseteq [-1, 1]$. Let $A = f^{-1}([-1, -1/3])$ and $B = f^{-1}([1/3, 1])$. Then A, B are disjoint closed subsets of Y, which are also closed in X since Y is closed. Thus by Urysohn's Lemma, there is a $g \in C(X, [-1/3, 1/3])$ such that $g \upharpoonright A = -1/3$ and $g \upharpoonright B = 1/3$. Thus $\|g\|_X = 1/3$, and with $Tg = g_{|Y}$ one easily checks (do it!) that $\|Tg - f\|_Y \leq 2/3$. Now Lemma 7.3 is applicable with $m = 1/3$ and $r = 2/3$ and gives the existence of $\widehat{f} \in C(X, \mathbb{R})$ with $T\widehat{f} = f$ and $\|\widehat{f}\| = \|f\|$ (since $m/(1-r) = 1$). ■

The theorem is easily extended to $\mathbb{C}$-valued functions.

A.6.2 Weierstrass' theorem

The following fundamental theorem of Weierstrass[106] (1885) has been proven in many ways. A fairly standard proof due to E. Landau (1908) involves convolution of f with a sequence $\{g_n\}$ of functions that is a polynomial approximate unit, cf. e.g. [163, Vol. II, Section 3.8]. The following proof, given in 1913 by S. Bernstein[107], has the advantage of using no integration.

A.32 Theorem *Let $f \in C([a, b], \mathbb{F})$ and $\varepsilon > 0$. Then there exists a polynomial $P \in \mathbb{F}[x]$ such that $|f(x) - P(x)| \leq \varepsilon$ for all $x \in [a, b]$. (As always, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.)*

Proof. It clearly suffices to prove this for the interval $[0, 1]$. For $n \in \mathbb{N}$ and $x \in [0, 1]$, define

$$P_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Clearly P_n is a polynomial of degree at most n, called Bernstein polynomial. In view of

$$1 = 1^n = (x + (1-x))^n = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \tag{A.1}$$

we have

$$f(x) - P_n(x) = \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right)\right) \binom{n}{k} x^k (1-x)^{n-k},$$

thus

$$|f(x) - P_n(x)| \leq \sum_{k=0}^{n} \left|f(x) - f\left(\frac{k}{n}\right)\right| \binom{n}{k} x^k (1-x)^{n-k}. \tag{A.2}$$

Since $[0, 1]$ is compact and $f : [0, 1] \to \mathbb{F}$ is continuous, it is bounded and uniformly continuous. Thus there is M such that $|f(x)| \leq M$ for all x, and for each $\varepsilon > 0$ there is $\delta > 0$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Let $\varepsilon > 0$ be given, and chose a corresponding $\delta > 0$ as above. Let $x \in [0, 1]$. Define

$$A = \left\{k \in \{0, 1, \ldots, n\} \mid \left|\frac{k}{n} - x\right| < \delta\right\}.$$

[106] Karl Theodor Wilhelm Weierstrass (1815-1897). German mathematician and one of the fathers of rigorous analysis.

[107] Sergei Natanovich Bernstein (1880-1968). Russian/Soviet mathematician. Important contributions to approximation theory, probability, PDEs.

For all k we have $|f(x)-f(k/n)| \leq 2M$, and for $k \in A$ we have $|f(x)-f(k/n)| < \varepsilon$. Thus with (A.2) we have

$$\begin{aligned} |f(x)-P_n(x)| &\leq \varepsilon \sum_{k\in A} \binom{n}{k} x^k(1-x)^{n-k} + 2M \sum_{k\in A^c} \binom{n}{k} x^k(1-x)^{n-k} \\ &\leq \varepsilon + 2M \sum_{k\in A^c} \binom{n}{k} x^k(1-x)^{n-k}, \end{aligned} \tag{A.3}$$

where we used (A.1) again. In an exercise, we will prove the purely algebraic identity

$$\sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k}(k-nx)^2 = nx(1-x) \tag{A.4}$$

for all $n \in \mathbb{N}_0$ and $x \in [0,1]$ (in fact all $x \in \mathbb{R}$). Now, $k \in A^c$ is equivalent to $|\frac{k}{n} - x| \geq \delta$ and to $(k-nx)^2 \geq n^2\delta^2$. Multiplying both sides of the latter inequality by $\binom{n}{k} x^k(1-x)^{n-k}$ and summing over $k \in A^c$, we have

$$\begin{aligned} n^2\delta^2 \sum_{k\in A^c} \binom{n}{k} x^k(1-x)^{n-k} &\leq \sum_{k\in A^c} \binom{n}{k} x^k(1-x)^{n-k}(k-nx)^2 \\ &\leq \sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k}(k-nx)^2 \;=\; nx(1-x), \end{aligned} \tag{A.5}$$

where the last equality comes from (A.4). This implies

$$\sum_{k\in A^c} \binom{n}{k} x^k(1-x)^{n-k} \leq \frac{nx(1-x)}{n^2\delta^2} \leq \frac{1}{n\delta^2}, \tag{A.6}$$

where we used the obvious inequality $x(1-x) \leq 1$ for $x \in [0,1]$. Plugging (A.6) into (A.3) we have $|f(x)-P_n(x)| \leq \varepsilon + \frac{2M}{n\delta^2}$. This holds for all $x \in [0,1]$ since, by uniform continuity, δ depends only on ε, not on x. Thus for $n > \frac{2M}{\varepsilon\delta^2}$ we have $|f(x)-P_n(x)| \leq 2\varepsilon \;\forall x \in [0,1]$ and are done. ■

A.33 Exercise Prove (A.4). Hint: Use basic properties of the binomial coefficients or differentiate $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$ twice with respect to x and then put $y = 1-x$.

An immediate consequence of Theorem A.32 is the following:

A.34 Corollary *There exists a sequence $\{p_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}[x]$ of real polynomials that converges uniformly on $[0,1]$ to the function $x \mapsto \sqrt{x}$.*

The above corollary can also be proven directly:

A.35 Exercise Define a sequence $\{p_n\}_{n\in\mathbb{N}_0} \subseteq \mathbb{R}[x]$ of polynomials by $p_0 = 0$ and

$$p_{n+1}(x) = p_n(x) + \frac{x - p_n(x)^2}{2}. \tag{A.7}$$

Prove by induction that the following holds:

(i) $p_n(x) \leq \sqrt{x}$ for all $n \in \mathbb{N}_0, x \in [0,1]$.

(ii) The sequence $\{p_n(x)\}$ increases monotonously for each $x \in [0,1]$ and converges uniformly to $\sqrt{x}$.

A.6.3 The Stone-Weierstrass theorem

Theorem A.32 says that the polynomials, restricted to $[0,1]$ are uniformly dense in $C([0,1])$. Our aim is to generalize this, replacing replacing $[0,1]$ by (locally) compact Hausdorff spaces. In order to see what should take the place of polynomials, notice that a polynomial on $\mathbb{R}$ is a linear combination of powers x^n, and the latter can be seen as powers f^n (under pointwise multiplication) of the identity function $f = \mathrm{id}_{\mathbb{R}}$. Thus the polynomials are the unital subalgebra $P \subseteq C(\mathbb{R},\mathbb{R})$ generated by the single element $\mathrm{id}_{\mathbb{R}}$. Now, if X is a topological space and $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ then $C(X,\mathbb{F})$ is a unital algebra, and we will consider subalgebras (not necessarily singly generated) $A \subseteq C(X,\mathbb{F})$. Since the functions on a (locally) compact Hausdorff space separate points, we clearly need to impose the following if we want to prove $\overline{A} = C(X,\mathbb{F})$:

A.36 Definition *A subalgebra $A \subseteq C(X,\mathbb{F})$ separates points if for any $x,y \in X$, $x \neq y$ there is a $f \in A$ such that $f(x) \neq f(y)$.*

A.37 Theorem (M. H. Stone 1937) [108] *If X is compact Hausdorff and $A \subseteq C(X,\mathbb{R})$ is a unital subalgebra separating points then $\overline{A} = C(X,\mathbb{R})$.*

Proof. Replacing A by $\overline{A}$, the claim is equivalent to showing that $A = C(X,\mathbb{R})$. We proceed in several steps. We claim that $f \in A$ implies $|f| \in A$. Since f is bounded due to compactness, it clearly is enough to prove this under the assumption $|f| \leq 1$. With the p_n of Corollary A.34, we have $(x \mapsto p_n(f^2(x))) \in A$ since A is a unital algebra. Since $p_n \circ f^2$ converges uniformly to $\sqrt{f^2} = |f|$, closedness of A implies $|f| \in A$. In view of

$$\max(f,g) = \frac{f+g+|f-g|}{2}, \qquad \min(f,g) = \frac{f+g-|f-g|}{2},$$

and the preceding result, we see that $f,g \in A$ implies $\min(f,g), \max(f,g) \in A$. By induction, this extends to pointwise minima/maxima of finite families of elements of A.

Now let $f \in C(X,\mathbb{R})$. Our goal is to find $f_\varepsilon \in A$ satisfying $\|f - f_\varepsilon\| < \varepsilon$ for each $\varepsilon > 0$. Since A is closed, this will give $A = C(X,\mathbb{R})$.

If $a \neq b$, the fact that A separates points gives us an $h \in A$ such that $h(a) \neq h(b)$. Thus the function $h_{a,b}(x) = \frac{h(x)-h(a)}{h(b)-h(a)}$ is in A, continuous and satisfies $h(a) = 0$, $h(b) = 1$. Thus also $f_{a,b}(x) = f(a) + (f(b) - f(a))h_{a,b}(x)$ is in A, and it satisfies $f_{a,b}(a) = f(a)$ and $f_{a,b}(b) = f(b)$. This implies that the sets

$$U_{a,b,\varepsilon} = \{x \in X \mid f_{a,b}(x) < f(x) + \varepsilon\}, \qquad V_{a,b,\varepsilon} = \{x \in X \mid f_{a,b}(x) > f(x) - \varepsilon\}$$

are open neighborhoods of a and b, respectively, for every $\varepsilon > 0$. Thus keeping b,ε fixed, $\{U_{a,b,\varepsilon}\}_{a \in X}$ is an open cover of X, and by compactness we find a finite subcover $\{U_{a_i,b,\varepsilon}\}_{i=1}^n$. By the above preparation, the function $f_{b,\varepsilon} = \min(f_{a_1,b,\varepsilon}, \ldots, f_{a_n,b,\varepsilon})$ is in A. If $x \in U_{a_i,b,\varepsilon}$ then $f_{b,\varepsilon}(x) \leq f_{a_i,b,\varepsilon}(x) < f(x) + \varepsilon$ for all $x \in X$, and since $\{U_{a_i,b,\varepsilon}\}_{i=1}^n$ covers X, we have $f_{b,\varepsilon}(x) < f(x) + \varepsilon \ \forall x$. For all $x \in V_{b,\varepsilon} = \bigcap_{i=1}^n V_{a_i,b,\varepsilon}$, we have $f_{a_i,b,\varepsilon}(x) > f(x) - \varepsilon$, and therefore $f_b(x) = \min_i(f_{a_i,b,\varepsilon}) > f(x) - \varepsilon$. Now $\{V_{b,\varepsilon}\}_{b \in X}$ is an open cover of X, and we find a finite subcover $\{V_{b_j,\varepsilon}\}_{j=1}^n$. Then $f_\varepsilon = \max(f_{b_1,\varepsilon}, \ldots, f_{b_n})$ is in A. Now $f_\varepsilon(x) = \max_j(f_{b_j,\varepsilon}) \leq f(x) + \varepsilon$ holds everywhere, and for $x \in V_{b_j,\varepsilon}$ we have $f_\varepsilon(x) \geq f_{b_j,\varepsilon} > f(x) - \varepsilon$. Since $\{V_{b_j,\varepsilon}\}_j$ covers X, we conclude that $f_\varepsilon(a) \in (f(x) - \varepsilon, f(x) + \varepsilon)$ for all x, to wit $\|f - f_\varepsilon\| < \varepsilon$. ∎

[108] Marshall Harvey Stone (1903-1989). American mathematician, mostly active in topology and (functional) analysis.

Since the polynomial ring $\mathbb{R}[x]$ is an algebra, and the polynomials clearly separate the points of $\mathbb{R}$, Theorem A.37 recovers Theorem A.32. (This is not circular if one has used Exercise A.35 to prove Corollary A.34.) But we immediately have the higher dimensional generalization (which can also be proven by more classical methods, like approximate units):

A.38 Theorem *Let $X \subseteq \mathbb{R}^n$ be compact. Then the restrictions to X of the $P \in \mathbb{R}[x_1, \ldots, x_n]$ (considered as functions) are uniformly dense in $C(X, \mathbb{R})$.*

Having proven Theorem A.37, it is easy to generalize it to locally compact spaces or/and subalgebras of $C_{(0)}(X, \mathbb{C})$. Recall that a subset S of a $*$-algebra $\mathcal{A}$ is called self-adjoint if $S = S^* := \{s^* \mid s \in S\}$.

A.39 Corollary *If X is compact Hausdorff and $A \subseteq C(X, \mathbb{C})$ is a self-adjoint unital subalgebra separating points then $\overline{A} = C(X, \mathbb{C})$.*

Proof. Define $B = A \cap C(X, \mathbb{R})$. Let $f \in A$. Since $f^* \in A$, we also have $\mathrm{Re}(f) = \frac{f+f^*}{2} \in B$ and $\mathrm{Im}(f) = \frac{f-f^*}{2i} = -\mathrm{Re}(if) \in B$. Thus $A = B + iB$. It is obvious that $\mathrm{Re}(A) \subseteq C(X, \mathbb{R})$ is a unital subalgebra. If $x \neq y$ then there is $f \in C(X, \mathbb{C})$ such that $f(x) \neq f(y)$. Thus $\mathrm{Re}(f)(x) \neq \mathrm{Re}(f)(y)$ or $\mathrm{Re}(if)(x) \neq \mathrm{Re}(if)(y)$ (or both). Since $\mathrm{Re}(f), \mathrm{Re}(if) \in B$, we see that B separates points. Thus $\overline{B} = C(X, \mathbb{R})$ by Theorem A.37, implying $\overline{A} = \overline{B + iB} = \overline{B} + i\overline{B} = C(X, \mathbb{R}) + iC(X, \mathbb{R}) = C(X, \mathbb{C})$. ■

A.40 Definition *A subalgebra $A \subseteq C_0(X, \mathbb{F})$ vanishes at no point if for every $x \in X$ there is an $f \in A$ such that $f(x) \neq 0$.*

A.41 Corollary *If X is locally compact Hausdorff and $A \subseteq C_0(X, \mathbb{R})$ is a subalgebra separating points and vanishing at no point then $\overline{A} = C_0(X, \mathbb{R})$.*

Proof. Let $X_\infty = X \cup \{\infty\}$ be the one-point compactification of X. Recall that every $f \in C_0(X, \mathbb{R})$ extends to $\widehat{f} \in C(X_\infty, \mathbb{R})$ with $\widehat{f}(\infty) = 0$. Then $B = \{\widehat{f} \mid f \in A\} + \mathbb{R}\mathbf{1}$ clearly is a unital subalgebra of $C(X_\infty, \mathbb{R})$. We claim that B separates the points of X_∞. This is obvious for $x, y \in X$, $x \neq y$ since already A does that. Now let $x \in X$. Since A vanishes at no point, there is $f \in A$ such that $f(x) \neq 0$. Let $\widehat{f} \in C(X, \mathbb{R})$ be the extension to X_∞ with $\widehat{f}(\infty) = 0$. In view of $\widehat{f}(x) = f(x) \neq 0$, we see that B also separates ∞ from the points of X, so that Theorem A.37 gives $\overline{B} = C(X, \mathbb{R})$. In view of $\overline{B} = \overline{A} + \mathbb{R}\mathbf{1}$ and $C(X, \mathbb{R}) \upharpoonright X = C_0(X, \mathbb{R})$, we have $\overline{A} = \overline{B} \upharpoonright X = C_0(X, \mathbb{R})$. ■

A.42 Corollary *If X is locally compact Hausdorff and $A \subseteq C_0(X, \mathbb{C})$ is a self-adjoint subalgebra separating points and vanishing at no point then $\overline{A} = C_0(X, \mathbb{C})$.*

Proof. The proof just combines the ideas of the proofs of Corollaries A.39 and A.41. ■

A.6.4 The Arzelà-Ascoli theorem

Recall that a metric space (X, d) is totally bounded if for every $\varepsilon > 0$ there are $x_1, \ldots, x_n \in X$ such that $X = B(x_1, \varepsilon) \cup \cdots \cup B(x_n, \varepsilon)$. And: A metric space is compact if and only if it is complete and totally bounded, cf. e.g. [108]. The following will be needed on many occasions:

A.43 Exercise Let (X, d) be a metric space. Prove:

(i) If (X, d) is totally bounded and $Y \subseteq X$ then (Y, d) is totally bounded.

(ii) If (Y,d) is totally bounded and $Y \subseteq X$ is dense then (X,d) is totally bounded.

(iii) If (X,d) is complete and $Y \subseteq X$ then (Y,d) is totally bounded if and only if Y is precompact.

If (X,τ) is a topological space and (Y,d) metric, the set $C_b(X,Y)$ is topologized by the metric

$$D(f,g) = \sup_{x\in X} d(f(x),g(x)).$$

It is therefore natural to ask whether the (relative) compactness of a set $\mathcal{F} \subseteq C_b(X,Y)$ can be characterized in terms of the elements of $\mathcal{F}$, which after all are functions $f : X \to Y$. This will be the subject of this section, but we will restrict ourselves to compact X, for which $C(X,Y) = C_b(X,Y)$.

A.44 Definition *Let (X,τ) be a topological space and (Y,d) a metric space. A family $\mathcal{F}$ of functions $X \to Y$ is called equicontinuous if for every $x \in X$ and $\varepsilon > 0$ there is an open neighborhood $U \ni x$ such that $f \in \mathcal{F},\ x' \in U \Rightarrow d(f(x),f(x')) < \varepsilon$. Then $\mathcal{F} \subseteq C(X,Y)$.*

The point of course is that the choice of U depends only on x and ε, but not on $f \in \mathcal{F}$.

A.45 Theorem (Arzelà-Ascoli) [109] *Let (X,τ) be a compact topological space and (Y,d) a complete metric space. Then $\mathcal{F} \subseteq C(X,Y)$ is (pre)compact (w.r.t. the uniform topology τ_D) if and only if the following conditions are satisfied:*

(i) $\{f(x) \mid f \in \mathcal{F}\} \subseteq Y$ is (pre)compact for every $x \in X$.

(ii) $\mathcal{F}$ is equicontinuous.

Proof. $\Rightarrow$ If $f,g \in C(X,Y)$ then $d(f(x),g(x)) \leq D(f,g)$ for every $x \in X$. This implies that the evaluation map $e_x : C(X,Y) \to Y,\ f \mapsto f(x)$ is continuous for every x. Thus if $\mathcal{F}$ is compact, so is $e_X(\mathcal{F})$. And compactness of $\overline{\mathcal{F}}$ implies that $e_x(\overline{\mathcal{F}}) = \{f(x) \mid f \in \overline{\mathcal{F}}\}$ is compact, thus closed. Since $e_x(\overline{\mathcal{F}})$ contains $e_x(\mathcal{F})$, also $\overline{e_x(\mathcal{F})} \subseteq e_x(\overline{\mathcal{F}})$ is compact.

To prove equicontinuity, let $x \in X$ and $\varepsilon > 0$. Since $\overline{\mathcal{F}}$ is compact, $\mathcal{F}$ is totally bounded, thus there are $g_1,\ldots,g_n \in \mathcal{F}$ such that $\mathcal{F} \subseteq \bigcup_i B^D(g_i,\varepsilon)$. By continuity of the g_i, there are open $U_i \ni x,\ i = 1,\ldots,n$, such that $x' \in U_i \Rightarrow d(g_i(x),g_i(x')) < \varepsilon$. Put $U = \bigcap_i U_i$. If now $f \in \mathcal{F}$, there is an i such that $f \in B^D(g_i,\varepsilon)$, to wit $D(f,g_i) < \varepsilon$. Now for $x' \in U \subseteq U_i$ we have

$$d(f(x),f(x')) \leq d(f(x),g_i(x)) + d(g_i(x),g_i(x')) + d(g_i(x'),f(x')) < 3\varepsilon,$$

proving equicontinuity of $\mathcal{F}$ (at x, but x was arbitrary).

$\Leftarrow$ Following [72], we begin with a lemma:

A.46 Lemma *Let (X,d) be a metric space. Assume that for each $\varepsilon > 0$ there are a $\delta > 0$, a metric space (Y,d') and a continuous map $h : X \to Y$ such that $(h(X),d')$ is totally bounded and such that $d'(h(x),h(x')) < \delta$ implies $d(x,x') < \varepsilon$. Then (X,d) is totally bounded.*

Proof. For $\varepsilon > 0$, pick $\delta,(Y,d'),h$ as asserted. Since $h(X)$ is totally bounded, there are $y_1,\ldots,y_n \in h(X)$ such that $h(X) \subseteq \bigcup_i B(y_i,\delta) \subseteq Y$. Then $X = \bigcup_i h^{-1}(B(y_i,\delta))$. For each i choose $x_i \in X$ such that $h(x_i) = y_i$. Now $x \in h^{-1}(B(y_i,\delta)) \Rightarrow d'(h(x),y_i) < \delta \Rightarrow d(x,x_i) < \varepsilon$, so that $h^{-1}(B(y_i,\delta)) \subseteq B(x_i,\varepsilon)$. Thus $X = \bigcup_{i=1}^n B(x_i,\varepsilon)$, and (X,d) is totally bounded. ■

[109] Giulio Ascoli (1843-1896), Cesare Arzelà (1847-1912), Italian mathematicians. They proved special cases of this result, of which there also exist more general versions than the one above.

Let $\varepsilon > 0$. Since $\mathcal{F}$ is equicontinuous, for every $x \in X$ there is an open neighborhood U_x such that $f \in \mathcal{F},\ x' \in U_x \Rightarrow d(f(x), f(x')) < \varepsilon$. Since X is compact, there are $x_1, \ldots, x_n \in X$ such that $X = \bigcup_{i=1}^n U_{x_i}$. Now define $h : \mathcal{F} \to Y^{\times n} : f \mapsto (f(x_1), \ldots, f(x_n))$. Now $\widetilde{d}((y_1, \ldots, y_n), (y'_1, \ldots, y'_n)) = \sum_i d(y_i, y'_i)$ is a product metric on $Y^{\times n}$ making h continuous. By assumption $\overline{\{f(x) \mid f \in \mathcal{F}\}}$ is compact for each $x \in X$, thus $\overline{h(\mathcal{F})} \subseteq \prod_i \overline{\{f(x_i) \mid f \in \mathcal{F}\}} \subseteq Y^{\times n}$ is compact, thus $(h(\mathcal{F}), \widetilde{d})$ is totally bounded. If now $f, g \in \mathcal{F}$ satisfy $\widetilde{d}(h(f), h(g)) < \varepsilon$ then $d(f(x_i), g(x_i)) < \varepsilon\ \forall i$ by definition of $\widetilde{d}$. For every $x \in X$ there is i such that $x \in U_{x_i}$, thus

$$d(f(x), g(x)) \leq d(f(x), f(x_i)) + d(f(x_i), g(x_i)) + d(g(x_i), g(x)) < 3\varepsilon.$$

Since this holds for all $x \in X$, we have $D(f, g) \leq 3\varepsilon$. Thus the assumptions of Lemma A.46 are satisfied, and we obtain total boundedness, thus precompactness, of $\mathcal{F}$. ■

A.47 Remark 1. If $Y = \mathbb{R}^n$, as in most statements of the theorem, then in view of the Heine-Borel theorem the requirement of precompactness of $\{f(x) \mid f \in \mathcal{F}\}$ for each x reduces to that of boundedness, i.e. pointwise boundedness of $\mathcal{F}$. One can also formulate the theorem in terms of existence of uniformly convergent (or Cauchy) subsequences of bounded equicontinuous sequences in $C(X, \mathbb{R}^n)$.

2. We intentionally stated a more general version [72] of the theorem than needed in order to argue that the result belongs to general topology rather than functional analysis. For $Y = \mathbb{R}^n$ this is less clear, also since there are many alternative proofs of the theorem using various methods from topology and functional analysis, cf. e.g. [111]. □

A.6.5 Separability of $C(X, \mathbb{R})$

A.48 Proposition *Let X be a compact Hausdorff space. Then the following are equivalent:*

(i) X is second countable ($\Leftrightarrow$ metrizable).

(ii) The normed space $(C(X, \mathbb{R}), \|\cdot\|)$, where $\|\cdot\| = \|\cdot\|_\infty$, is second countable ($\Leftrightarrow$ separable).

Proof. (i)⇒(ii) Let $\mathcal{B} = \{U_1, U_2, \ldots\}$ be a countable base for the topology of X, and let

$$S = \{(n, m) \in \mathbb{N}^2 \mid \overline{U_n} \subseteq U_m\}.$$

For every $(n, m) \in S$ use Urysohn's Lemma to find a function $f_{(n,m)} \in C(X, [0, 1])$ such that $f_{(n,m)} \upharpoonright \overline{U_n} = 0,\ f_{(n,m)} \upharpoonright X \backslash U_m = 1$. Let $x, y \in F$, $x \neq y$. Since $X \backslash \{y\}$ is an open neighborhood of x and $\mathcal{B}$ is a base, there exists $m \in \mathbb{N}$ with $x \in U_m$. By normality of X there exists an open V such that $x \in V \subseteq \overline{V} \subseteq U_m$. Since $\mathcal{B}$ is a base there exists $n \in \mathbb{N}$ with $x \in U_n \subseteq V$. Now $\overline{U_n} \subseteq \overline{V} \subseteq U_m$, so that $(n, m) \in S$. Now $f_{(n,m)}(x) = 0$, $f_{(n,m)}(y) = 1$, so that the family $F_1 = \{f_{(n,m)} \mid (n, m) \in S\} \subset C(X, [0, 1])$ separates the points of X.

Let F_2 denote the set, clearly countable, of all finite products of elements of F_1. Interpreting the empty product as the function $\mathbf{1}$, we have $\mathbf{1} \in F_2$. Then also the set F_3 of finite linear combinations of elements of F_2 with $\mathbb{Q}$-coefficients is countable. Since $A = \overline{F_3}$ contains the finite linear combinations of elements of F_2 with coefficients in $\mathbb{R}$, it is a unital $\mathbb{R}$-algebra. Since already F_1 separates the points of X, the same holds for A. Thus the Stone-Weierstrass Theorem A.37 gives $C(X, \mathbb{R}) = A = \overline{F_3}$. Thus $C(X, \mathbb{R})$ has F_3 as countable dense subset.

(ii)⇒(i) Since $(C(X, \mathbb{R}), \|\cdot\|)$ is metric, second countability and separability are equivalent. Let $F \subseteq C(X, \mathbb{R})$ be a subset that is dense w.r.t. $\|\cdot\|$. Let $x, y \in X$, $x \neq y$. We claim that there is an $f \in F$ with $f(x) \neq f(y)$ (i.e. F separates the points of X). If this was false, the uniform density of F in $C(X,\mathbb{R})$ would imply $f(x) = f(y)$ for all $f \in C(X,\mathbb{R})$, which however is false by Urysohn's lemma. The map

$$\iota_F : X \to \prod_{f \in F} [\inf f, \sup f], \quad x \mapsto \prod_{f \in F} f(x)$$

is continuous by definition of the product topology, and it is injective since F separates points. Since X is compact and the product space Hausdorff, ι_F is an embedding, i.e. $\iota_X : X \to \iota_X(X)$ is a homeomorphism. If now F is countable, the countable product $\prod_{f \in F}[\inf f, \sup f]$ is second countable, thus also its subspace $\iota_F(X)$ which is homeomorphic to X. ■

A.49 Remark 1. If X is locally compact Hausdorff, it is most natural to consider $C_0(X,\mathbb{F})$. With the one-point (Alexandrov) compactification X_∞, one easily proves a Banach space isomorphism $C(X_\infty,\mathbb{F}) \cong C_0(X,\mathbb{F}) \oplus \mathbb{F}$, so that $C_0(X,\mathbb{F})$ is separable if and only if X_∞ is second countable. Second countablility of X_∞ implies that of X. The converse is also true, but is more work since it involves proving that $\infty \in X_\infty$ has a countable open neighborhood base. This is equivalent to X being hemicompact, i.e. there is a family $\{K_n \subseteq X\}_{n\in\mathbb{N}}$ of compact sets such that every compact $K \subseteq X$ is contained in some K_n. One can show that for a locally compact Hausdorff space, hemicompactness is equivalent to second countability, cf. e.g. [108]. Thus also for locally compact X one has that $C_0(X,\mathbb{F})$ is separable if and only if X is second countable.

2. For non-compact X, one can also study $C_b(X,\mathbb{F})$. At least if X is completely regular, it turns out that $C_b(X,\mathbb{F})$ is never separable for non-compact X. (For compact X we have $C_b(X,\mathbb{F}) = C(X,\mathbb{F})$ and are back in the situation of Proposition A.48.) □

A.6.6 ⋆ The Stone-Čech compactification

If X is a topological space, a compactification of X is a space $\widehat{X}$ together with a continuous map $\iota : X \to \widehat{X}$ such that $\iota(X) \subseteq X$ is a dense subset and $\iota : X \to \iota(X)$ is a homeomorphism.

You probably know the one-point or Alexandrov compactification of a topological space. (Usually it is considered only for locally compact spaces.) It is the smallest possible compactification in that it just adds one point.

But for many purposes, another compactification is more important, the Stone-Čech compactification. It is defined for spaces that have the following property:

A.50 Definition *A topological space X is completely regular or $T_{3.5}$ if it is T_1 and for every closed $C \subseteq X$ and $y \in X \backslash C$ there exists $f \in C(X,[0,1])$ such that $f \upharpoonright C = 0$ and $f(x) = 1$.*

All subspaces of a completely regular space are completely regular. By Urysohn's lemma, every normal space is completely regular, in particular every metrizable and every compact Hausdorff space. This implies that complete regularity is a necessary condition for a space X to have a compactification $\widehat{X}$ that is Hausdorff. In fact, it also is sufficient:

A.51 Theorem *Let X be a topological space. Then the following are equivalent:*

(i) X is completely regular.

(ii) There exists a compact Hausdorff space βX together with a dense embedding $X \hookrightarrow \beta X$ such that for every continuous function $f : X \to Y$, where Y is compact Hausdorff, there exists a continuous $\widehat{f} : \beta X \to Y$ such that $\widehat{f} \upharpoonright X = f$. (This $\widehat{f}$ is automatically unique by density of $X \subseteq \beta X$.)

A.52 Remark 1. The universal property (ii) implies that βX is unique up to homeomorphism. 'It' is called the Stone-Čech[110] compactification of X.

2. If X is completely regular and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ then the restriction map $C(\beta X, \mathbb{F}) \to C_b(X, \mathbb{F})$ given by $f \mapsto f \upharpoonright X$ is a bijection and an isometric isomorphism of Banach algebras.

3. There are many ways to prove the non-trivial implication (i)⇒(ii). We sketch two, referring to the literature for details.

(A) Define $Z = [0,1]^{C(X,[0,1])} = \prod_{f \in C(X,[0,1])} [0,1]$ with the product topology, which is compact Hausdorff. The map $\iota_X : X \to Z,\ x \mapsto \prod_{f \in C(X,[0,1])} f(x)$ is continuous and injective since the continuous functions on a completely regular space separate points. Using that they also separate points from closed sets, one proves that ι_X is an embedding, thus a homeomorphism $X \to \iota_X(X)$. Let $\beta X = \overline{\iota_X(X)} \subseteq Z$, which is compact Hausdorff. Now we can identify X with the dense subspace $\iota_X(X)$ of βX. By construction, for every $f \in C(X,[0,1])$ we have $p_f \circ \iota_X = f$, where p_f is the projection $Z \to [0,1]$ indexed by f. Thus $p_f \upharpoonright \beta X$ extends f to βX. This generalizes to any compact Hausdorff space Y instead of $[0,1]$ using the fact that every compact Hausdorff space is homeomorphic to a closed subset of a cube (which is proven as in the proof of (ii)⇒(i) Proposition A.48, but now taking $F = C(X,[0,1])$).

(B) Alternatively, one can use Gelfand duality for commutative C^*-algebras, cf. Section 19: If X is completely regular, $\mathcal{A} = C_b(X, \mathbb{C})$ with norm $\|f\| = \sup_x |f(x)|$ is a commutative unital C^*-algebra. As such it has a spectrum $\Omega(\mathcal{A})$, which is compact Hausdorff. We define $\beta X = \Omega(\mathcal{A})$. There is a map $\iota : X \to \Omega(\mathcal{A}), x \mapsto \varphi_x$, where $\varphi_x(f) = f(x)$. This map is continuous by definition of the topology on $\Omega(\mathcal{A})$. Using the complete regularity of X one proves that ι is an embedding, i.e. a homeomorphism of X onto $\iota(X) \subseteq \Omega(\mathcal{A})$. Now $\overline{\iota(X)} = \Omega(\mathcal{A})$ is seen as follows: $\overline{\iota(X)} \neq \Omega(\mathcal{A})$ would imply (using Urysohn or Tietze) that there are $f \in \mathcal{A} \backslash \{0\}$ such that $\iota(x)(f) = 0$ for all $x \in X$. This is a contradiction, since the elements of $\mathcal{A}$ are functions on X, so that $\iota(x)(f) = 0\ \forall x$ implies $f = 0$.

4. The first of the above constructions used the Tychonov theorem, but only for Hausdorff spaces. The second approach relies on Alaoglu's theorem to prove compactness of $\Omega(\mathcal{A})$. One can show that Alaoglu's theorem and the restriction of Tychonov's theorem to Hausdorff spaces are equivalent over the ZF axioms, see Section B.5. In fact, also Theorem A.51 is in this equivalence class.

5. If X is completely regular and non-compact, one can prove, cf. e.g. [51, 3.6.17], that no point $x \in \beta X \backslash X$ has a countable open neighborhood basis, so that βX is not second countable. Then $C_b(X, \mathbb{F}) \cong C(\beta X, \mathbb{F})$ is not separable. Not assuming complete regularity, not much can be said since one can find topological spaces – even regular (T_3) ones – on which every continuous $\mathbb{R}$-valued function is constant, in which case $C_b(X, \mathbb{F}) \cong \mathbb{F}$, see [51, 2.7.17, 2.7.18]. □

A.7 Some notions from measure and integration theory

A.53 Definition *If X is a set, a σ-algebra on X is a family $\mathcal{A} \subseteq P(X)$ of subsets such that*

1. *$\emptyset \in \mathcal{A}$.*
2. *If $A \in \mathcal{A}$ then $X \backslash A \in \mathcal{A}$.*
3. *If $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.*

A measurable space is a pair $(X, \mathcal{A})$ consisting of a set and a σ-algebra on it.

The closedness of $\mathcal{A}$ under complements implies that a σ-algebra also contains X and is closed under countable intersections. Obviously $P(X)$ is a σ-algebra.

[110]Eduard Čech (1893-1960). Czech mathematician, worked mostly in topology, e.g. Čech cohomology.

It is very easy to see that the intersection of any number of σ-algebras on X is a σ-algebra on X. Thus if $\mathcal{F} \subseteq P(X)$ is any family of subsets of X, we can define the σ-algebra generated by $\mathcal{F}$ as the intersection of all σ-algebras on X that contain $\mathcal{F}$.

If (X, τ) is a topological space, the σ-algebra on X generated by τ is called the Borel[111] σ-algebra $\mathcal{B}(X)$ of X. (We should of course write $\mathcal{B}(X, \tau)$...) Apart from the open sets, it contains the closed sets, the G_δ sets and many more. A function $f : X \to \mathbb{C}$ is called Borel measurable if $f^{-1}(U) \in \mathcal{B}(X)$ for every open $U \subseteq \mathbb{C}$. (This is equivalent to $f^{-1}(B) \in \mathcal{B}(X)$ for every $B \in \mathcal{B}(\mathbb{C})$.) If $(X, \mathcal{A})$ is a measurable space, $B^\infty(X, \mathbb{C})$ denotes the set of functions $f : X \to \mathbb{C}$ that are Borel-measurable and bounded, i.e. $\sup_{x \in X} |f(x)| < \infty$. It is not hard to check that this is an algebra (with the pointwise product).

A.54 DEFINITION *A positive measure on a measurable space $(X, \mathcal{A})$ is a map $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$ whenever $\{A_n\} \subseteq \mathcal{A}$ is a countable family of mutually disjoint sets. If $\mu(X) < \infty$ then μ is called finite (then $\mu(A) \leq \mu(X)$ $\forall A \in \mathcal{A}$).*

A Borel measure on a topological space (X, τ) is a positive measure on $(X, \mathcal{B}(X))$.

There is a notion of regularity of a measure. Since we will only consider measures on compact subsets of $\mathbb{C}$, which are second countable, regularity of all finite Borel measures is automatic. (This follows e.g. from [140, Theorem 2.18].)

For the definition of integration of real or complex valued functions w.r.t. a measure see any book on measure theory or the appendix of [101].

The counting measure on $(X, P(X))$ is defined by $\mu_c(A) = \#(A)$. It is easy to show that $f : X \to \mathbb{C}$ (obviously measurable) is μ_c-integrable if and only if $\sum_{x \in X} f(x)$ exists in the sense of Definition A.1, in which case $\int f(x)\, d\mu_c(x) = \sum_{x \in X} f(x)$.

A.55 DEFINITION *A complex measure on a measurable space $(X, \mathcal{A})$ is a map $\mu : \mathcal{A} \to \mathbb{C}$ such that $\mu(\emptyset) = 0$ and $\mu(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$ whenever $\{A_n\} \subseteq \mathcal{A}$ is a countable family of mutually disjoint sets.*

Note that complex measures are by definition bounded. Furthermore, if $\{A_n\}$ is a countable family of mutually disjoint sets then automatically $\sum_n |\mu(A_n)| < \infty$ since $\mu(\bigcup_n A_n)$ is invariant under permutations of the A_n. The supremum of this expression over the families $\{A_n\}_{n \in \mathbb{N}}$ of mutually disjoint measurable sets is called the total variation $\|\mu\|$ of μ.

For every bounded measurable function $f : X \to \mathbb{C}$ one can define the integral $\int_X f d\mu$ satisfying $|\int_X f d\mu| \leq \|f\|_\infty \|\mu\|$, thus defining a bounded linear functional on $L^\infty(X, \mathcal{B}(X), \mu; \mathbb{C})$.

A.56 THEOREM (F.RIESZ-MARKOV-KAKUTANI) [112] *Let X be a compact Hausdorff space and $\varphi : C(X, \mathbb{C}) \to \mathbb{C}$ a linear functional. Then*

- *(i) If φ is bounded, there is a finite complex measure μ on $(X, \mathcal{B}(X))$ such that $\|\mu\| = \|\varphi\|$ and $\varphi(f) = \int_X f d\mu$ for all $f \in C(X, \mathbb{C})$.*
- *(ii) If φ is positive, i.e. $\varphi(f) \geq 0$ whenever $f \geq 0$, then it is bounded, and there is a finite positive measure μ on $(X, \mathcal{B}(X))$ with $\mu(X) = \|\varphi\|$ and $\varphi(f) = \int_X f d\mu$ for all $f \in C(X, \mathbb{C})$. (This also works for $\mathbb{R}$ instead of $\mathbb{C}$.)*

For proofs see [29, Theorem 7.2.8] or [140, Theorem 2.14] in the real and [29, Theorem 7.3.6] or [140, Theorem 6.19] in the complex case. (For the implication positive $\Rightarrow$ bounded see Proposition B.158(i).)

[111] Emile Borel (1871-1956). French mathematician. One of the pioneers of measure theory.

[112] Andrey Andreyevich Markov (1903-1979), Soviet mathematician. Shizuo Kakutani (1911-2004), Japanese-American mathematician. There also are fixed point theorems due to Kakutani and to Markov-Kakutani.

B ⋆ Supplements for the curious

B.1 Functional analysis over fields other than $\mathbb{R}$ and $\mathbb{C}$?

The most general meaningful definition of (linear) functional analysis is as the theory of topological vector spaces over a topological field $\mathbb{F}$ and continuous linear maps between them. If (the topology of) $\mathbb{F}$ is discrete, we are effectively doing topological abelian group theory, and this would not be considered functional analysis. Thus we restrict ourselves to non-discrete topological fields. The general theory of topological fields is a thorny subject, almost unknown to non-specialists. (For reviews see [175, 167].) There would be no point in going into this here since in this course we considered general topological vector spaces only as a step towards spaces that are at least metrizable.

But there is a complete (in a sense) classification of the non-discrete locally compact fields. In characteristic zero, these are precisely $\mathbb{R}$, the p-adic fields $\mathbb{Q}_p$, where p runs through the prime numbers, and all their finite (thus algebraic) extensions. (And in characteristic $p \neq 0$ one has the finite extensions of $\mathbb{F}_p((x))$, the field of formal Laurent series over the finite field $\mathbb{F}_p$.) For a proof see e.g. [126]. While $\mathbb{R}$ has only one algebraic extension (namely $\mathbb{C}$), $\mathbb{Q}_p$ has infinitely many finite extensions, so that the algebraic closure of $\overline{\mathbb{Q}_p}$ (which is not complete!) is infinite-dimensional over $\mathbb{Q}_p$. Like $\mathbb{R}$ and $\mathbb{C}$, the p-adic fields and their finite extensions all have a norm, usually called 'valuation' or 'absolute value', i.e. a map $\mathbb{F} \to [0, \infty)$ satisfying $|x| = 0 \Leftrightarrow x = 0$, $|x + y| \leq |x| + |y|$ and $|xy| = |x||y|$. Note that the norm is strictly multiplicative, not just submultiplicative. The locally compact fields are complete w.r.t. their absolute value $|\cdot|$.

Books entitled 'Functional analysis' or 'Topological vector spaces' tend to work entirely over $\mathbb{R}$ and $\mathbb{C}$ unless the title contains 'p-adic', 'non-archimedian' or 'ultrametric' (but there are exceptions like [21, 112]). Nevertheless, functional analysis over p-adic fields is a well-studied subject, cf. e.g. [138, 125], but a somewhat exotic one since it only seems to have applications to number theory, algebraic geometry and related fields.[113]

In the remainder of this short section we briefly comment on the extent to which the theory covered in these notes remains valid over p-adic fields. As a rule of thumb, one must be very careful with theorems on normed/Banach spaces that involve $\mathbb{R}$ or $\mathbb{C}$ either in their statement or in the proof since then either the orderedness of $\mathbb{R}$ or the algebraic completeness of $\mathbb{C}$ tend to be used, while the p-adic fields are neither algebraically closed nor orderable! The Hahn-Banach theorem is a case in point since we first proved it for $\mathbb{R}$, making essential use of the orderedness of the base field $\mathbb{F} = \mathbb{R}$, thus not just of the set $[0, \infty)$ in which the norms take values, and then extended it to $\mathbb{C}$. (There nevertheless is a p-adic Hahn-Banach theorem, but with slightly different hypotheses and a different proof.)

Theorems not explicitly referring to $\mathbb{R}$ or $\mathbb{C}$ have a better chance of carrying over to p-adic functional analysis. For example, the open mapping theorem and both versions of the uniform boundedness theorem generalize without change. However, one has to be careful with the above rule since there are properties, like connectedness, shared by $\mathbb{R}$ and $\mathbb{C}$, but not enjoyed by the p-adic fields! There are other problems: There is no a priori relationship between the subsets $S_1 = \{|c| \mid c \in \mathbb{F}\}$ and $S_2 = \{\|x\| \mid x \in V\}$ of $[0, \infty)$. Thus given $x \in V \backslash \{0\}$ there may not be a $c \in \mathbb{F}$ such that $\|cx\| = 1$.

We also have to be very careful with results on Hilbert spaces, since scalars in $\mathbb{F}$ can be pulled out of inner products without picking up an absolute value: $\langle cx, y \rangle = c\langle x, y \rangle$. Indeed this leads to problems adapting the proof of Theorem 5.27. The same holds for the polarization

[113] The author is skeptical about claims of relevance of p-adic/ultrametric (functional) analysis to fundamental theoretical/mathematical physics (but statistical/condensed matter physics is another discussion).

identities.

We leave the discussion here and refer to the literature on p-adic (functional) analysis for more information. See e.g. [60, 136, 138, 125].

B.2 Even more on unconditional and conditional convergence

B.2.1 The Dvoretzky-Rogers theorem

B.1 Proposition *Let $n \in \mathbb{N}$ and V be a normed space with $\dim V \geq n^2$. Then there are unit vectors $x_1, \ldots, x_n \in V$ such that*

$$\Big\| \sum_{i=1}^n c_i x_i \Big\| \leq 8 \Big(\sum_{i=1}^n |c_i|^2 \Big)^{1/2} \quad \forall c_1, \ldots, c_n \in \mathbb{F}. \tag{B.1}$$

Before we prove the proposition, we consider its consequences:

B.2 Theorem (Dvoretzky-Rogers 1950) [114] *Let V be an infinite-dimensional Banach space and $\{c_n\}_{n\in\mathbb{N}}$ positive numbers such that $\sum_{n=1}^\infty c_n^2 < \infty$. Then there is an unconditionally convergent series $\sum_{n=1}^\infty x_n$ in V such that $\|x_n\| = c_n$ for all n.*

Proof. By $\sum_{j=1}^\infty c_j^2 < \infty$ we can choose integers $n_1 < n_2 < \cdots$ such that $\sum_{j=n_k}^\infty c_j^2 \leq 2^{-2k}$. For $i < n_1$ pick arbitrary vectors x_i with $\|x_i\| = c_i$. Since V is infinite-dimensional, Proposition B.1 provides, for every $k \in \mathbb{N}$, unit vectors $y_{n_k}, \ldots, y_{n_{k+1}-1} \in V$ such that with $x_i = c_i y_i$ we have

$$\Big\| \sum_{i=n_k}^{n_{k+1}-1} d_i x_i \Big\| \leq 8 \Big(\sum_{i=n_k}^{n_{k+1}-1} c_i^2 \Big)^{1/2} \leq 8 \cdot 2^{-k}$$

for all choices of $d_i \in \{0,1\}$. Thus all subseries of $\sum_{i=n_1}^\infty x_i = \sum_{k=1}^\infty \sum_{i=n_k}^{n_{k+1}-1} x_i$ converge, so that $\sum_{i=1}^\infty x_i$ converges unconditionally by Theorem A.4(vii)⇒(iii). ■

B.3 Corollary *In every infinite-dimensional Banach space there are series that converge unconditionally, but not absolutely.*

Proof. This is immediate, applying Theorem B.2 to any sequence $\{c_n\}$ of positive numbers such that $\sum_n c_n = \infty$ and $\sum_n c_n^2 < \infty$, like $c_n = 1/n$. ■

Proof of Proposition B.1. We can clearly assume $\dim V = n^2$. By Proposition 3.10, there exists an Auerbach basis $\{x_i\}$ be an Auerbach basis for V, i.e. $\|x_i\| = \|\varphi_i\| = 1$ for all i, where $\{\varphi_i\} \in V^*$ is the unique dual basis. Define an inner product on V by $\langle x, y\rangle = n^2 \sum_{i=1}^{n^2} \varphi_i(x)\overline{\varphi_i(y)}$, which induces the norm $\|x\|' = n\big(\sum_{i=1}^{n^2} |\varphi_i(x)|^2\big)^{1/2}$. Then

$$\|x\|'/n^2 \leq \max_i |\varphi_i(x)| \leq \|x\| \leq \sum_{i=1}^{n^2} |\varphi_i(x)| \leq \|x\|' \quad \forall x,$$

where the inequalities are due to, in turn, the definition of $\|\cdot\|'$ in terms of the $\varphi_i(x)$, $\|\varphi_i\| = 1$, the fact $x = \sum_i \varphi_i(x) x_i$ and $\|x_i\| = 1$, and the Cauchy-Schwarz inequality (applied to the vectors $\{\varphi_i(x)\}$ and $(1, \ldots, 1)$).

[114] Aryeh Dvoretzky (1916-2008), Russian-born Israeli mathematician, worked mostly in functional analysis and probability. Claude Ambrose Rogers (1920-2005), British mathematician, mostly in convex geometry.

****

Thus $\{y_i\}$ is an ONB for $(V, \langle\cdot,\cdot\rangle)$ such that $\|y_i\| \geq 1/8$ for each i.

Putting $x_i = y_i/\|y_i\|$, we have $\|x_i\| = 1$, $\|x_i\|' \leq 8$ and with the mutual orthogonality of the x_i

$$\Big\| \sum_{i=1}^n c_i x_i \Big\| \leq \Big\| \sum_{i=1}^n c_i x_i \Big\|' = \Big(\sum_{i=1}^n (|c_i| \|x_i\|')^2 \Big)^{1/2} \leq 8 \Big(\sum_{i=1}^n |c_i|^2 \Big)^{1/2}$$

for all $\{c_i\}$ in $\mathbb{F}$. ■

B.4 Remark Proposition B.1 (from [98]) suffices for Theorem B.2 and its proof only uses the very classical existence of Auerbach bases (1929). But it can be improved considerably if one uses slightly later theory like the John ellipsoid (1948) or Lewis' lemma (1979). Cf. e.g. [97, vol. 1, Proposition IV.1], where the 8 in (B.1) is replaced by 2 and the n^2 by $2n$.

And in 1960, Dvoretzky proved Dvoretzky's theorem, a much stronger result that started the 'local theory' of Banach spaces (which focuses on finite-dimensional subspaces): For every $k \in \mathbb{N}$ and $\varepsilon > 0$ there exists $N(k,\varepsilon) \in \mathbb{N}$ such that for every normed space $(V, \|\cdot\|)$ of dimension $\geq N(k,\varepsilon)$ there is a k-dimensional subspace $W \subseteq V$ such that $d(W,E) < 1 + \varepsilon$, where d is the Banach-Mazur distance and $E = \mathbb{R}^k$ with euclidean norm. In particular every infinite-dimensional Banach space has finite-dimensional subspaces of arbitrary large dimension and arbitrarily close to being Hilbert spaces. See e.g. [1] or [97, vol. 2]. □

B.2.2 Converses of Dvoretzky-Rogers

In general it is difficult to give necessary and sufficient conditions for unconditional convergence of a series. There is a relatively easy exception:

B.5 Exercise Let X be a compact Hausdorff space and $V = C(X, \mathbb{F})$ with norm $\|\cdot\|_\infty$. Prove that a series $\sum_{n=1}^\infty f_n$ in V is unconditionally convergent if and only if $\sum_{n=1}^\infty |f_n(x)| < \infty$ for all $x \in X$. Hint: Dini's theorem.

(Global absolute convergence means $\sum_{n=1}^\infty \sup_{x\in X} |f_n(x)| < \infty$. While it is quite plausible that the latter condition is strictly stronger than pointwise absolute convergence, the Dvoretzky-Rogers theorem proves that this is the case whenever X is infinite.)

B.6 Exercise Let H be a Hilbert space. Prove that a series $\sum_{n=1}^\infty x_n$ of mutually orthogonal terms converges unconditionally if and only if $\sum_{n=1}^\infty \|x_n\|^2 < \infty$.

It is plausible that unconditional convergence is harder to achieve if the summands are not mutually orthogonal. (For example, if $x \neq 0$ and $x_n = c_n x$ then unconditional convergence of $\sum_{n=1}^\infty x_n$ is equivalent to $\sum_{n=1}^\infty \|x_n\| < \infty$.) Therefore the following is not surprising:

B.7 Proposition (Orlicz) *If a series $\sum_{n=1}^\infty x_n$ in a Hilbert space converges unconditionally then $\sum_{n=1}^\infty \|x_n\|^2 < \infty$.*

Proof. Since $\sum_n x_n$ converges unconditionally, by Proposition (iv) there exists a finite $S \subset \mathbb{N}$ such that $\|\sum_{t\in T} x_t\| < 1$ for every finite $T \subset \mathbb{N}\backslash S$. Now for every finite $T \subset \mathbb{N}$ we have

$$\Big\| \sum_{t\in T} x_t \Big\| = \Big\| \sum_{t\in S\cap T} x_t + \sum_{t\in T\backslash S} x_t \Big\| \leq \Big\| \sum_{t\in S\cap T} x_t \Big\| + \Big\| \sum_{t\in T\backslash S} x_t \Big\| \leq \Big(\sum_{s\in S} \|x_s\| \Big) + 1 =: C.$$

Now given $n \in \mathbb{N}$ and $s_1, \ldots, s_n \in \{\pm 1\}$, putting $T_+ = \{i = 1, \ldots, n \mid \varepsilon_i = 1\}$, $T_- = \{i = 1, \ldots, n \mid \varepsilon_i = -1\}$ we have

$$\Big\| \sum_{i=1}^{n} s_i x_i \Big\| = \Big\| \sum_{t \in T_+} x_i - \sum_{t \in T_-} x_t \Big\| \le \Big\| \sum_{t \in T_+} x_i \Big\| + \Big\| \sum_{t \in T_-} x_i \Big\| \le 2C. \tag{B.2}$$

So far we haven't used any property of V, but now we do. With the generalized parallelogram identity (5.6), we have

$$\sum_{i=1}^{n} \|x_i\|^2 = 2^{-n} \sum_{s \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} s_i x_i \Big\|^2 \le (2C)^2$$

for all n, thus also $\sum_{i=1}^{\infty} \|x_i\|^2 \le (2C)^2 < \infty$. ■

Recall from Remark 5.18 that a Banach space V has cotype c if there exists $C < \infty$ such that for all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in V$

$$\sum_{i=1}^{n} \|x_i\|^c \le \frac{C}{2^n} \sum_{s \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} s_i x_i \Big\|^c. \tag{B.3}$$

Combining this with (B.2) immediately gives:

B.8 Proposition *If V is a Banach space of cotype $c < \infty$ then every unconditionally convergent series $\sum_{n=1}^{\infty} x_n$ in V satisfies $\sum_{n=1}^{\infty} \|x_n\|^c < \infty$.*

B.9 Exercise Let S be an infinite set and $p \in [1, \infty)$. Prove that $\ell^p(S, \mathbb{F})$ has cotype $\max(p, 2)$.

B.10 Corollary *Let S be a set, $p \in [1, \infty)$ and $\sum_{n=1}^{\infty} x_n$ an unconditionally convergent series in $\ell^p(S, \mathbb{F})$. Then $\sum_{n=1}^{\infty} \|x_n\|^c < \infty$, where $c = \max(p, 2)$.*

B.2.3 Conditional convergence

If a series $\sum_{n=1}^{\infty} x_n$ in a Banach space V is not unconditionally convergent, it is natural to ask about the set of sums of all convergent rearrangements. For $V = \mathbb{R}$, the latter is $\mathbb{R}$, as shown by Riemann. In finite dimensions we have:

B.11 Theorem (Lévy (1905), Steinitz (1913-4)) [115] *Let $\sum_{n=1}^{\infty} x_n$ be a conditionally convergent series in a finite-dimensional normed space V over $\mathbb{R}$. Then*

$$\Phi = \Big\{ \varphi \in V^* \ \Big| \ \sum_{n=1}^{\infty} |\varphi(x_n)| < \infty \Big\} \subseteq V^*$$

is a linear space, and the set

$$\Sigma = \Big\{ s \in V \mid \exists \sigma : s = \sum_{n=1}^{\infty} x_{\sigma(n)} \Big\}$$

of sums of convergent rearrangements equals $x_0 + \Phi^\top$, where $x_0 \in \Sigma$ and $\Phi^\top = \{x \in V \mid \varphi(x) = 0 \ \forall \varphi \in \Phi\}$.

[115] Paul Lévy (1886-1971), French mathematician. Best remembered for his work on probability theory. Ernst Steinitz (1871-1928), German mathematician who worked in many areas. Founder of the theory of fields.

The proof begins with an easy reduction to the case where $\Phi = \{0\}$, thus $\Sigma = V$. For a nice exposition see [69]. The straightforward generalization of the above to infinite-dimensional Banach spaces is false, since the set Σ of values of convergent rearrangements can fail to be an affine space, cf. e.g. [78]. (It can even consist of two points.) But there are a number of sufficient conditions under which $\Sigma = x_0 + \Phi^\top$ holds. For example the following result, which has a remarkable application to the 'universality' of the Riemann zeta function:

B.12 THEOREM (PECHERSKIĬ 1973) *[82, Appendix §6] If H is a real Hilbert space and $\sum_{n=1}^\infty x_n$ is such that $\Sigma \neq \emptyset$ and $\sum_{n=1}^\infty \|x_n\|^2 < \infty$ then $\Sigma = x_0 + \Phi^\top$.*

B.3 More on the spaces $\ell^p(S, \mathbb{F})$ and $c_0(S, \mathbb{F})$

B.3.1 Precompact subsets of $c_0(S, \mathbb{F})$ and $\ell^p(S, \mathbb{F})$, $1 \leq p < \infty$

One can characterize the precompact subsets of $\ell^p(S, \mathbb{F})$, $1 \leq p < \infty$ in a fashion quite similar to the Arzelà-Ascoli Theorem A.45. As in our discussion of the latter, we follow [72].

B.13 PROPOSITION *Let S be a set, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $1 \leq p \leq \infty$. Put $V = \ell^p(S, \mathbb{F})$ if $p < \infty$ and $V = c_0(S, \mathbb{F})$ if $p = \infty$. Then $X \subset V$ is (pre)compact if and only if*

(i) The set $\{f(s) \mid f \in X\} \subseteq \mathbb{F}$ is compact (bounded) for each $s \in S$.

(ii) For each $\varepsilon > 0$ there is a finite $F \subseteq S$ such that $\|(1 - \chi_F)f\|_p \leq \varepsilon$ for all $f \in X$. (I.e., the elements of X tend to zero at infinity uniformly.)

Proof. If X is compact, its image in $\mathbb{F}$ under the continuous evaluation map $p_s : f \mapsto f(s)$ is compact for each $s \in S$. If X is only precompact then $p_s(\overline{X})$ is compact, and continuity of p_s gives $\overline{p_s(X)} \subseteq p_s(\overline{X})$, so that $p_s(X)$ is precompact (=bounded).

By (pre)compactness of X, it is totally bounded. Thus if $\varepsilon > 0$ we can find $f_1, \ldots, f_n \in \ell^p(S, \mathbb{F})$ such that $X \subseteq \bigcup_{i=1}^n B(f_i, \varepsilon)$. Now there are finite sets $F_1, \ldots, F_n$ such that $\|(1 - \chi_F)f_i\|_p \leq \varepsilon/2$ for each $i = 1, \ldots, n$. Then $F = \bigcup_{i=1}^n F_i$ is finite. If $f \in V$ then pick i such that $\|f - f_i\|_p < \varepsilon/2$. Using the rather obvious facts $\|(1 - \chi_F)f_i\|_p \leq \|(1 - \chi_{F_i})f_i\|_p \leq \varepsilon/2$ and $\|(1 - \chi_F)(f - f_i)\|_p \leq \|f - f_i\|_p < \varepsilon/2$, we obtain

$$\|(1 - \chi_F)f\| \leq \|(1 - \chi_F)f_i\|_p + \|(1 - \chi_F)(f - f_i)\|_p \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

thus (ii). (Note that the existence of the F_i would fail for $(\ell^\infty(S, \mathbb{F}), \|\cdot\|_\infty)$.)

Now assume that (i) and (ii) hold and $\varepsilon > 0$. Then we can find a finite $F \subseteq S$ as in condition (ii). Write $F = \{s_1, \ldots, s_n\}$ and define a map $h : X \to \mathbb{F}^n$, $f \mapsto (f(s_1), \ldots, f(s_n))$. Equip $\mathbb{F}^n$ with the norm $\|\cdot\|_p$. By condition (i), $h(X) \subseteq \mathbb{F}^n$ is bounded/compact. Let $f, f' \in X$ such that $\|h(f) - h(f')\|_p \leq \varepsilon$. Since this is equivalent to $\|\chi_F(f - f')\|_p \leq \varepsilon$, we have

$$\|f - f'\|_p \leq \|\chi_F(f - f')\|_p + \|(1 - \chi)_F f\|_p + \|(1 - \chi)_F f'\|_p \leq 3\varepsilon.$$

This shows that the assumptions of Lemma A.46 are satisfied, so that X is totally bounded, thus precompact. Finally, assume the sets in (i) are compact. The total boundedness of X is equivalent to the statement that every sequence $\{f_n\}$ in X has a Cauchy subsequence $\{f_{n_i}\}$. The latter converges to some element $f \in V$. Now the closedness of $p_s(X) \subseteq \mathbb{F}$ for each s implies that the limit function f is in X. Thus X is compact. ■

B.14 Remark It is interesting to compare the conditions in Theorem A.45 and Proposition B.13. While the respective pointwise conditions (i) are identical, the conditions (ii) are totally different. In Theorem A.45 we are concerned with continuous function and therefore have a uniform (over the functions) version of continuity, while due to compactness there is no condition at infinity. On the other hand, in Proposition B.13 there are no continuity questions, but we need a uniform vanishing condition at infinity.

As one may expect, there is a generalization of Proposition B.13 to L^p-spaces, now involving three conditions. Note that pointwise conditions make no sense.

B.15 Theorem (Kolmogorov-M.Riesz theorem) [116] *Let $1 \le p < \infty$ and $X \subseteq L^p(\mathbb{R}^n, \lambda)$. Then X is precompact if and only if*

(i) X is $\|\cdot\|_p$-bounded.

(ii) $\lim_{R\to\infty} \sup_{f\in X} \left(\int_{|x|\ge R} |f(x)|^p d\lambda \right)^{1/p} = 0$ (condition at ∞)

(iii) $\lim_{y\to 0} \sup_{f\in X} \|f - f_y\|_p = 0$, where $f_y(x) = f(x-y)$ (variant of equicontinuity).

(Note that $\lim_{y\to\infty} \|f - f_y\|_p = 0$ holds for all $f \in L^p(\mathbb{R}^n)$.) For an elegant proof of the theorem, again using Lemma A.46, see [72]. In the literature one can find more restrictive versions (e.g. in [23, Theorem 4.26] condition (ii) is replaced by the stronger hypothesis that all $f \in X$ are supported in the same bounded set) as well as more general ones. See [169, §12] for a version on a locally compact group G, presented more readably in [88]. □

B.3.2 The dual space of $\ell^\infty(S, \mathbb{F})$

We have seen in in Theorem 4.19(v) that there are bounded linear functionals $\varphi \in \ell^\infty(S, \mathbb{F})^*$ that vanish on $c_0(S, \mathbb{F})$. Those clearly cannot be captured by the function $g(s) = \varphi(\delta_s)$ widely used in the proof of Theorem 4.19. This suggests to consider $\mu_\varphi(A) = \varphi(\chi_A)$ for arbitrary $A \subseteq S$ instead. If $A_1, \ldots, A_K$ are mutually disjoint, and $A = \bigcup_{k=1}^K A_k$ then $\chi_A = \sum_{k=1}^K \chi_{A_k}$, thus $\mu_\varphi(A) = \sum_{k=1}^K \mu_\varphi(A_k)$, so that μ_φ is finitely additive.[117]

B.16 Definition *If S is a set, a finitely additive finite $\mathbb{F}$-valued measure on S is a map $\mu : P(S) \to \mathbb{F}$ satisfying $\mu(\emptyset) = 0$ and $\mu(A_1 \cup \cdots \cup A_K) = \mu(A_1) + \cdots + \mu(A_K)$ whenever $A_1, \ldots, A_K$ are mutually disjoint subsets of S. The set of such μ, which we denote $fa(S, \mathbb{F})$, is a vector space via $(c_1\mu_1 + c_2\mu_2)(A) = c_1\mu_1(A) + c_2\mu_2(A)$. For $\mu \in fa(S, \mathbb{F})$ we define*

$$\begin{aligned} \|\mu\| &= \sup \left\{ \sum_{k=1}^K |\mu(A_k)| \mid K \in \mathbb{N},\ A_1, \ldots, A_K \subseteq S,\ i \ne j \Rightarrow A_i \cap A_j = \emptyset \right\}, \\ \|\mu\|' &= \sup_{A\subseteq S} |\mu(A)|. \end{aligned}$$

B.17 Theorem *(i) $\|\cdot\|$ and $\|\cdot\|'$ are equivalent norms on $fa(S, \mathbb{F})$. We write*

$$ba(S, \mathbb{F}) = \{\mu \in fa(S, \mathbb{F}) \mid \|\mu\|' < \infty \ (\Leftrightarrow \|\mu\| < \infty)\}.$$

(ii) $(ba(S, \mathbb{F}), \|\cdot\|)$ is a Banach space.

[116] Andrey Nikolaevich Kolmogorov (1903-1987). Soviet mathematician with countless contributions to classical, harmonic and functional analysis, dynamical systems, probability theory (which he founded on measure theory), etc.

[117] The discussion in this section strongly borrows from [43].

(iii) If $\varphi \in \ell^\infty(S,\mathbb{F})^*$ then $\|\mu_\varphi\| \le \|\mu\|$, thus we have a norm-decreasing linear map $\ell^\infty(S,\mathbb{F})^* \to ba(S,\mathbb{F}),\ \varphi \mapsto \mu_\varphi$.

Proof. (i) It is immediate from the definition $\|c\mu\| = |c|\|\mu\|$ and $\|c\mu\|' = |c|\|\mu\|'$ for all $c \in \mathbb{F}, \mu \in fa(S,\mathbb{F})$ and that $\|\mu\| = 0 \Leftrightarrow \mu = 0 \Leftrightarrow \|\mu\|' = 0$. Also $\|\mu_1+\mu_2\|' \le \|\mu_1\|' + \|\mu_2\|'$ is quite obvious. Now

$$\begin{aligned}\|\mu_1+\mu_2\| &= \sup\left\{\sum_{k=1}^K |\mu_1(A_k)+\mu_2(A_k)| \mid \cdots\right\} \le \sup\left\{\sum_{k=1}^K |\mu_1(A_k)| + |\mu_2(A_k)| \mid \cdots\right\} \\ &\le \sup\left\{\sum_{k=1}^K |\mu_1(A_k)| \mid \cdots\right\} + \sup\left\{\sum_{k=1}^K |\mu_2(A_k)| \mid \cdots\right\} = \|\mu_1\| + \|\mu_2\|.\end{aligned}$$

Thus $\|\cdot\|, \|\cdot\|'$ are norms on $fa(S,\mathbb{F})$. The definition of $\|\cdot\|$ clearly implies $|\mu(A)| \le \|\mu\|$ for each $A \subseteq S$, whence $\|\mu\|' \le \|\mu\|$.

Assume $\mu \in fa(S,\mathbb{R})$ and $\|\mu\|' < \infty$. If $A_1,\dots,A_K \subseteq S$ are mutually disjoint, put

$$A_+ = \bigcup\{A_k \mid \mu(A_k) \ge 0\}, \qquad A_- = \bigcup\{A_k \mid \mu(A_k) < 0\}.$$

Now by finite additivity, $\sum_k |\mu(A_k)| = \mu(A_+) + \mu(A_-) \le 2\|\mu\|'$ since $|\mu(A_\pm)| \le \|\mu\|'$. Taking the supremum over the families $\{A_k\}$ gives $\|\mu\| \le 2\|\mu\|'$.

If $\mu \in fa(S,\mathbb{C})$, writing $\mu = \operatorname{Re}\mu + i\operatorname{Im}\mu$ we find $\|\mu\| \le 4\|\mu\|'$. Thus $\|\mu\|' \le \|\mu\| \le 4\|\mu\|'$ for all μ, and the two norms are equivalent.

(ii) Here it is more convenient to work with the simpler norm $\|\cdot\|'$. Now let $\{\mu_n\}$ be a Cauchy sequence in $ba(S,\mathbb{F})$. Then $|\mu_n(A) - \mu_m(A)| \le \|\mu_n - \mu_m\|'$, so that $\{\mu_n(A)\}$ is Cauchy, thus convergent. Define $\mu(n) = \lim_n \mu_n(A)$. It is clear that $\mu(\emptyset) = 0$. If $A_1,\dots,A_K$ are mutually disjoint then

$$\mu(A_1\cup\cdots\cup A_K) = \lim_{n\to\infty} \mu_n(A_1\cup\cdots\cup A_K) = \lim_{n\to\infty}(\mu_n(A_1)+\cdots+\mu_n(A_K)) = \mu(A_1)+\cdots+\mu(A_K),$$

so that μ is finitely additive. Since $\{\mu_n\}$ is Cauchy, for every $\varepsilon > 0$ there is n_0 such that $n,m \ge n_0$ implies $\|\mu_m - \mu_n\|' < \varepsilon$. In particular there is n_0 such that $\|\mu_m\|' \le \|\mu_{n_0}\|' + 1$ for $m \ge n_0$. This implies boundedness of μ. And taking $m \to \infty$ in $|\mu_n(A) - \mu_m(A)| \le \|\mu_n - \mu_m\|' < \varepsilon$ gives $\|\mu_n - \mu\|' \le \varepsilon$, so that $\|\mu_n - \mu\|' \to 0$. Thus $ba(S,\mathbb{F})$ is complete (w.r.t. $\|\cdot\|'$, thus also w.r.t. $\|\cdot\|$).

(iii) It is clear that $\ell^\infty(S,\mathbb{F})^* \to fa(S,\mathbb{F}),\ \varphi \mapsto \mu_\varphi$ is linear. Now let $A_1,\dots,A_K \subseteq S$ be mutually disjoint. Then

$$\sum_{k=1}^K |\mu_\varphi(A_k)| = \sum_{k=1}^K \overline{\operatorname{sgn}(\mu_\varphi(A_k))}\mu_\varphi(A_k) = \sum_{k=1}^K \overline{\operatorname{sgn}(\mu_\varphi(A_k))}\varphi(\chi_{A_k}) = \varphi\left(\sum_{k=1}^K \overline{\operatorname{sgn}(\mu_\varphi(A_k))}\chi_{A_k}\right).$$

Since the A_k are mutually disjoint and $|\operatorname{sgn}(z)| \le 1$, we have $\|\sum_{k=1}^K \overline{\operatorname{sgn}(\mu_\varphi(A_k))}\chi_{A_k}\|_\infty \le 1$, so that $\sum_{k=1}^K |\mu_\varphi(A_k)| \le \|\varphi\|$. Taking the supremum over the finite families $\{A_k\}$ gives $\|\mu_\varphi\| \le \|\varphi\|$. ■

B.18 Theorem *(i) For each $\mu \in ba(S,\mathbb{F})$ there is a unique linear functional $\int_\mu \in \ell^\infty(S,\mathbb{F})^*$ such that $\int_\mu(\chi_A) = \mu(A)$ for all $A \subseteq S$. It satisfies $\|\int_\mu\| \le \|\mu\|$.*

(ii) The maps $\alpha : \ell^\infty(S,\mathbb{F})^* \to ba(S,\mathbb{F}),\ \varphi \mapsto \mu_\varphi$ and $\int : ba(S,\mathbb{F}) \to \ell^\infty(S,\mathbb{F})^*, \mu \mapsto \int_\mu$ are mutually inverse and isometric, thus $\ell^\infty(S,\mathbb{F})^* \cong ba(S,\mathbb{F})$.

Proof. (i) If $f \in \ell^1(S, \mathbb{F})$ has finite image, write $f = \sum_{k=1}^K c_k \chi_{A_k}$, where the A_k are mutually disjoint, and define

$$\int f \, d\mu = \sum_{k=1}^K c_k \mu(A_k).$$

(We write $\int_\mu(f)$ or $\int f \, d\mu$ according to convenience.) If $f = \sum_{l=1}^L c'_l \chi_{A'_l}$ is another representation of f, then using finite additivity of μ it ia straightforward to check, using the finite additivity of μ, that $\sum_{k=1}^K c_k \mu(A_k) = \sum_{l=1}^L c'_l \mu(A'_l)$, so that $\int f \, d\mu$ is well-defined. Now $\int cf \, d\mu = c \int f \, d\mu$ for $c \in \mathbb{F}$ is obvious, and $\int (f+g) d\mu = \int f \, d\mu + \int g \, d\mu$ for all finite-image functions follows from the fact that $f+g$ again is a finite-image function and the representation independence of $\int$. Thus $\int_\mu : f \mapsto \int f \, d\mu$ is a linear functional on the bounded finite image functions. It is clear that this is the unique linear functional sending χ_A to $\mu(A)$ for each $A \subseteq S$. Now

$$\left| \int f \, d\mu \right| \le \sum_{k=1}^K |c_k| \, |\mu(A_k)| \le \|f\|_\infty \sum_{k=1}^K |\mu(A_k)| \le \|f\|_\infty \|\mu\|.$$

Thus $\int_\mu$ is a bounded functional, and since the bounded finite-image functions are dense in $\ell^\infty(S, \mathbb{F})$ by Lemma 4.13, $\int_\mu$ has a unique extension to a linear functional $\int_\mu \in \ell^\infty(S, \mathbb{F})^*$ with $\| \int_\mu \| \le \|\mu\|$.

(ii) If $\mu \in ba(S, \mathbb{F})$ then by definition of $\int_\mu$, we have $\int \chi_A \; d\mu = \mu(A)$ for all $A \subseteq S$. Thus $\alpha \circ \int = \mathrm{id}_{ba(S,\mathbb{F})}$.

If $\varphi \in \ell^\infty(S, \mathbb{F})$ then in view of the definition of $\int$ we have $\int \chi_A \, d\mu_\varphi = \mu_\varphi(A) = \varphi(\chi_A)$ for all $A \subseteq S$. Thus φ and $\int_{\mu_\varphi}$ coincide on all characteristic functions, thus on all of $\ell^\infty(S, \mathbb{F})$ by linearity, density of the finite-image functions and the $\| \cdot \|_\infty$ continuity of φ and $\int_{\mu_\varphi}$. Thus $\int \circ \alpha = \mathrm{id}_{\ell^\infty(S,\mathbb{F})^*}$.

Since the maps α and $\int$ are mutually inverse and both norm-decreasing, they actually both are isometries. ■

This completes the determination of $\ell^\infty(S, \mathbb{F})^*$. (Note that we did not use the completeness of $ba(S, \mathbb{F})$ proven in Theorem B.17(ii). Thus it would also follow from the isometric bijection $ba(S, \mathbb{F}) \cong \ell^\infty(S, \mathbb{F})^*$ just established.)

B.19 Exercise Given $\mu \in ba(S, \mathbb{F})$, prove that μ is $\{0,1\}$-valued if and only if $\int_\mu \in \ell^\infty(S, \mathbb{F})^*$ is a character, i.e. $\int_\mu(fg) = \int_\mu(f) \int_\mu(g)$ for all $f, g \in \ell^\infty(S, \mathbb{F})$.

Since $\ell^\infty(S, \mathbb{F})^*$ has a closed subspace $\iota(\ell^1(S, \mathbb{F}))$, it is interesting to identify the corresponding subspace of $ba(S, \mathbb{F})$.

B.20 Definition *A finitely additive measure $\mu \in ba(S, \mathbb{F})$ is called countably additive if for every countable family $\mathcal{A} \subseteq P(S)$ of mutually disjoint sets we have*

$$\mu\left(\bigcup \mathcal{A} \right) = \sum_{A \in \mathcal{A}} \mu(A)$$

and totally additive if the same holds for any family of mutually disjoint sets. The set of countably and totally additive measures on S are denoted $ca(S, \mathbb{F})$ and $ta(S, \mathbb{F})$, respectively.

B.21 Proposition *For $\mu \in ba(S, \mathbb{F})$, consider the following statements:*

(i) There is $g \in \ell^1(S,\mathbb{F})$ such that $\mu(A) = \sum_{s\in A} g(s)$ for all $A \subseteq S$.

(ii) $\int_\mu \in \ell^\infty(S,\mathbb{F})^*$ is 'normal', thus $\int f\,d\mu = \lim_\iota \int f_\iota\,d\mu$ for every net $\{f_\iota\} \in \mathbb{F}^S$ that is pointwise convergent and uniformly bounded.

(iii) μ is totally additive.

(iv) μ is countably additive.

Then (i)⇔(ii)⇔(iii)⇒(iv). If S is countable then also (iv)⇒(iii).

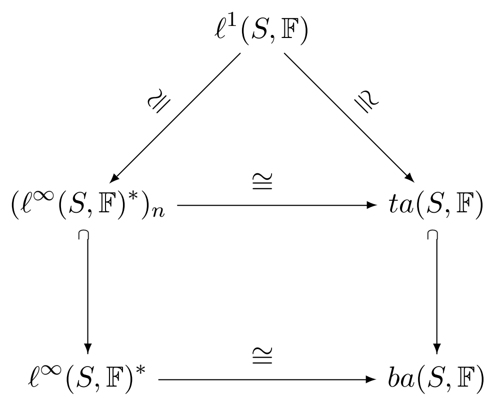
Proof. (i)⇒(ii) If μ is of the given form then clearly $\int_\mu \chi_A\,d\mu = \mu(A) = \sum_{s\in A} g(s)$ for each $A \subseteq S$. By the way $\int_\mu$ is constructed from μ, it is clear that $\int f\,d\mu = \sum_{s\in S} f(s)g(s)$ for all $f \in \ell^\infty(S,\mathbb{F})$. Thus $\int_\mu = \varphi_g$, and normality of $\int_\mu$ follows from Proposition A.3.

(ii)⇒(iii) We know that we can recover μ from $\int_\mu$ as $\mu(A) = \int \chi_A\,d\mu$. Let $\mathcal{A}$ be a family of mutually disjoint subsets of S. Then the net $\{f_F = \chi_{\bigcup F}\}$, indexed by the finite subsets $F \subseteq \mathcal{A}$, is uniformly bounded and converges pointwise to χ_B, where $B = \bigcup \mathcal{A}$. Now normality of $\int_\mu$ implies that $\mu(B) = \int_\mu \chi_B\,d\mu = \lim_F \int f_F = \lim_F \sum_{A\in F} \mu(A) = \sum_{A\in\mathcal{A}} \mu(A)$, which is additivity of μ.

(iii)⇒(i) If we put $g(s) = \mu(\{s\})$ then additivity of μ means that $\mu(A) = \sum_{s\in A} g(s)$ for all $A \subseteq S$, convergence being absolute. Now the finiteness of $\mu(S)$ gives $\|g\|_1 < \infty$.

(iii)⇒(iv) is trivial. If S is countable then a family of mutually disjoint non-empty subsets of S is at most countable, so that (iii) and (iv) are equivalent. ■

Thus we have the situation of the following diagram:



where $ta(S,\mathbb{F})$ can be replaced by $ca(S,\mathbb{F})$ if S is countable.

B.3.3 $c_0(\mathbb{N},\mathbb{F}) \subseteq \ell^\infty(\mathbb{N},\mathbb{F})$ is not complemented

B.22 Definition *We say that a Banach space V has property S if there is a countable subset $C \subseteq V^*$ separating the points of V. I.e., if $x \in V$ and $\varphi(x) = 0$ for all $\varphi \in C$ then $x = 0$.*

If V has property S then every closed subspace $W \subseteq V$ has property S. (And so would non-closed subspaces, but they are not Banach.)

It is easy to see that V has property S whenever V^* is separable. But this is not a necessary condition: $V = \ell^\infty(\mathbb{N},\mathbb{C})$ has property S, as we see by taking $C = \{\varphi_n\}_{n\in\mathbb{N}}$, where $\varphi_n(f) = f(n)$. But V^* ($\cong ba(\mathbb{N},\mathbb{C})$) is not separable, since by Exercise 9.26 this would imply separability of V, which is false by Exercise 4.18(i).

B.23 Theorem (Phillips 1939, Sobczyk 1940) *The closed subspace $c_0(\mathbb{N},\mathbb{R}) \subseteq \ell^\infty(\mathbb{N},\mathbb{R})$ is not complemented.*

Proof. (Whitley (1966).) From now on we abbreviate $\ell^\infty(\mathbb{N},\mathbb{F})$ and $c_0(\mathbb{N},\mathbb{F})$ as ℓ^∞, c_0. Our strategy for proving that $c_0 \subseteq \ell^\infty$ is not complemented is the following: If $c_0 \subseteq \ell^\infty$ had a complementary closed subspace W, Exercise 7.15 would give $\ell^\infty \cong c_0 \oplus W$, thus $\ell^\infty/c_0 \cong W$. Since W would have property S, it would follow that $Q = \ell^\infty/c_0$ has property S, but we will prove that it doesn't!

The idea for doing so is to produce an uncountable subset $\mathcal{F} \subseteq Q$ such that each functional $\varphi \in Q^*$ is non-zero only on countably many elements of $\mathcal{F}$. Then for any countable $C \subseteq Q^*$ the set $\mathcal{F}' = \bigcup_{\varphi\in C}\{q \in Q \mid \varphi(q) \neq 0\}$ is countable, so that the family $C \subseteq Q^*$ vanishes identically on the uncountable set $\mathcal{F}\backslash\mathcal{F}'$. It therefore cannot separate the elements of $\mathcal{F}$, let alone those of Q. Thus Q does not have property S and we are done. For the construction of such an $\mathcal{F}$ we use the following lemma:

B.24 Lemma *Every countably infinite set X admits a family $\{X_\lambda\}_{\lambda\in\Lambda}$ of subsets of X such that*

(i) Λ has cardinality $\mathfrak{c} = \#\mathbb{R}$, in particular it is uncountable.

(ii) X_λ is infinite for each $\lambda \in \Lambda$.

(iii) $X_\lambda \cap X_{\lambda'}$ is finite for all $\lambda, \lambda' \in \Lambda$, $\lambda \neq \lambda'$.

Proof. Take $Y = (0,1) \cap \mathbb{Q}$ and $\Lambda = (0,1)\backslash\mathbb{Q}$. Clearly Y is countable and Λ is uncountable (since removing a countable set from one of cardinality $\mathfrak{c}$ does not change the cardinality). For each $\lambda \in \Lambda$ pick a sequence $\{a_n\} \subseteq Y$ converging to λ (for example $a_n = \lfloor n\lambda\rfloor/n$) and put $Y_\lambda = \{a_n \mid n \in \mathbb{N}\}$. That each Y_λ is infinite follows from the irrationality of λ and the rationality of the a_n. If $\lambda \neq \lambda'$ and $a_n \to \lambda, a'_n \to \lambda'$ then there exists n_0 such that $n, n' \geq n_0 \Rightarrow \max(|a_n - \lambda|, |a'_{n'} - \lambda'|) < |\lambda - \lambda'|/2$, so that $a_n \neq a'_{n'}$. This implies $\#(Y_\lambda \cap Y_{\lambda'}) < \infty$. We thus have a family of subsets of Y with all desired properties. For an arbitrary countably infinite set X the claim now follows using a bijection $X \cong Y$. ■

Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of subsets of $\mathbb{N}$ as provided by the lemma. For $\lambda \in \Lambda$, the characteristic function $\chi_{X_\lambda} : \mathbb{N} \to \{0,1\} \subseteq \mathbb{C}$ clearly is in ℓ^∞. Let $p : \ell^\infty \to Q = \ell^\infty/c_0$ be the quotient map. Now let $q_\lambda = p(\chi_{X_\lambda})$ and $\mathcal{F} = \{q_\lambda \mid \lambda \in \Lambda\}$. If $\lambda, \lambda' \in \Lambda, \lambda \neq \lambda$ the symmetric difference $X_\lambda \Delta X_{\lambda'} = (X_\lambda \cup X_{\lambda'})\backslash(X_{\lambda'} \cap X_\lambda)$ is infinite by (ii) and (iii). Thus $\chi_{X_\lambda} - \chi_{X_{\lambda'}} \neq c_0 = \ker p$, so that $\lambda \mapsto q_\lambda$ is injective, thus with (i) we see that $\mathcal{F}$ is uncountable.

Let now $\varphi \in Q^*$, $m, n \in \mathbb{N}$ and let $\lambda_1, \ldots, \lambda_m \in \Lambda$ be mutually distinct and such that $|\varphi(q_{\lambda_i})| \geq 1/n \ \forall i = 1,\ldots,m$. For each i pick t_i with $|t_i| = 1$ such that $t_i\varphi(q_{\lambda_i}) = |\varphi(q_{\lambda_i})|$. Put $f = \sum_{i=1}^m t_i\chi_{X_{\lambda_i}} \in \ell^\infty$. Since the sets X_{λ_i} have pairwise finite intersections, the function f has absolute value larger than one only on a subset of the finite set $\bigcup_{j,k=1}^m X_{\lambda_j} \cap X_{\lambda_k}$ and absolute value one on the infinite set $(\bigcup_i X_{\lambda_i})\backslash(\bigcup_{j,k} X_{\lambda_j} \cap X_{\lambda_k})$. This implies that $\|p(f)\| = \inf_{g\in c_0} \|f - g\|_\infty = 1$. Thus

$$\|\varphi\| \geq |\varphi(p(f))| = \left|\sum_{i=1}^m t_i\varphi(p(\chi_{X_{\lambda_i}}))\right| = \left|\sum_{i=1}^m t_i\varphi(q_{\lambda_i})\right| = \sum_{i=1}^m |\varphi(q_{\lambda_i})| \geq \frac{m}{n}.$$

Thus $m \leq n\|\varphi\| < \infty$, so that for each $\varphi \in Q^*$ and $n \in \mathbb{N}$ there cannot be more than m distinct $\lambda \in \Lambda$ with $|\varphi(q_\lambda)| \geq 1/n$. If there was an uncountable $\mathcal{F}' \subseteq \mathcal{F}$ with $\varphi(q) \neq 0 \ \forall q \in \mathcal{F}'$, there would have to be an $n \in \mathbb{N}$ such that $|\varphi(q)| \geq 1/n$ for infinitely (in fact uncountably) many $q \in \mathcal{F}'$, contradicting what we just proved. This completes the proof. ■

B.3.4 $c_0(\mathbb{N},\mathbb{F})$ is not a dual space. Spaces with multiple pre-duals

Recall that we write $\cong$ for isometric isomorphism and $\simeq$ for isomorphism of Banach spaces.

B.25 Lemma *Let V be a Banach space. Then:*

(i) $P = \iota_{V^*} \circ (\iota_V)^* \in B(V^{***})$ *is an idempotent and* $PV^{***} = \iota_{V^*}(V^*)$.

(ii) $\iota_{V^*}(V^*) \subseteq V^{***}$ *is a complemented subspace.*

(iii) If a Banach space W is isomorphic to V^* with V Banach then $\iota_W(W) \subseteq W^{**}$ is complemented.

(iv) $V^{***}/V^* \simeq (V^{**}/V)^*$. *(We omitted the ι's for simplicity.)*

Proof. (i) Since ι_V, ι_{V^*} are bounded, with Lemma 9.31 we have boundedness of P. Let $\varphi \in V^*$ and $x \in X$. Then

$$(P\iota_{V^*}(\varphi))(\iota_V(x)) = \iota_{V^*}((\iota_V)^*(\iota_{V^*}(\varphi)))(\iota_V(x)) = [(\iota_V)^*(\iota_{V^*}(\varphi))](x) = \varphi(x) = \iota_{V^*}(\varphi)(\iota_V(x)),$$

where we used Exercise 9.18 several times, proves $P\iota_{V^*}(\varphi) = \iota_{V^*}(\varphi)$. Thus $P \restriction \iota_{V^*}(V^*) = \mathrm{id}$. On the other hand, it follows directly from the definition of P that $PV^{***} \subseteq \iota_{V^*}(V^*)$. Combining these two facts gives $P^2 = P$ and $PV^{***} = \iota_{V^*}(V^*)$.

(ii) This is an immediate consequence of (ii) and Exercise 6.13.

(iii) If $T : W \to V^*$ is an isomorphism then we have isomorphisms $T^* : V^{**} \to W^*$ and $T^{**} : W^{**} \to V^{***}$. Using this it is straightforward to deduce the claim from (ii).

(iv) By Exercise 6.7 we have $(V^{**}/V)^* \cong V^\perp \subseteq V^{***}$. And by (ii), $V^{***} \simeq V^* \oplus W$, where $W \simeq V^{***}/V^*$, the isomorphism being given by $x^{***} \mapsto (Px^{***}, (1-P)x^{***})$ with P as in (i). Thus $PV^{***} \simeq V^*$ and $V^{***}/V^* \simeq (1-P)V^{***}$. Thus the claimed isomorphism follows if we prove that the subspaces $V^\perp$ and $(1-P)V^{***}$ of V^{***} are equal.

Now, $x^{***} \in (1-P)V^{***}$ means $(1-P)x^{***} = x^{***}$, thus $Px^{***} = 0$. Since $P = \iota_{V^*} \circ (\iota_V)^*$, where ι_{V^*} is injective, this is equivalent to $(\iota_V)^*(x^{***}) = 0$. By the definition of the transpose, this means that $x^{***} \circ \iota_V = 0$. Since this is the same as $x^{***} \in \iota_V(V)^\perp$, we are done. ∎

B.26 Corollary *$c_0(\mathbb{N},\mathbb{F})$ is not isomorphic to the dual space of any Banach space.*

Proof. We again abbreviate $c_0(\mathbb{N},\mathbb{F})$ as c_0 etc. We know that $c_0^* \cong \ell^1$ and $c_0^{**} \cong \ell^\infty$, the canonical map $\iota_{c_0} : c_0 \to c_0^{**}$ just being the inclusion map $c_0 \hookrightarrow \ell^\infty$. By Theorem B.23, $c_0 \subseteq \ell^\infty$ is not complemented. Combining this with Lemma B.25(iii), the claim follows. ∎

B.27 Corollary *Let $X = c_0 \oplus (\ell^\infty/c_0)$. Then $X \not\simeq \ell^\infty$, but $X^* \simeq (\ell^\infty)^*$.*

Proof. $X \simeq \ell^\infty$ would imply that $c_0 \subseteq \ell^\infty$ is complemented, which it is not by Theorem B.23. Thus $X \not\simeq \ell^\infty$. With $c_0^* \cong \ell^1$ we have $X^* \simeq c_0^* \oplus (\ell^\infty/c_0)^* \simeq \ell^1 \oplus (\ell^\infty/c_0)^*$.

On the other hand, since $\ell^1 \cong c_0^*$ is a dual space, we see that $\ell^1 \subseteq (\ell^1)^{**} \cong (\ell^\infty)^*$ is complemented by Lemma B.25(iii). Thus $(\ell^\infty)^* \simeq \ell^1 \oplus (\ell^\infty)^*/\ell^1$ by Exercise 7.15(i). Now Lemma B.25(iv) with $V = c_0$ gives $(\ell^\infty)^*/\ell^1 \simeq (\ell^\infty/c_0)^*$, so that $X^* \simeq (\ell^\infty)^*$. ∎

One can also find Banach spaces V with $V^* \cong \ell^1$, while $V \not\simeq c_0$. But this is a bit more involved.

B.3.5 Schur's theorem for $\ell^1(\mathbb{N},\mathbb{F})$

As on earlier occasions, we abbreviate $\ell^1 = \ell^1(\mathbb{N},\mathbb{F})$.

B.28 Theorem (I. Schur 1921) *If $g, \{f_n\}_{n\in\mathbb{N}} \subseteq \ell^1(\mathbb{N},\mathbb{F})$ and $f_n \stackrel{w}{\to} g$ then $\|f_n - g\|_1 \to 0$.*

Proof. It clearly suffices to prove this for $g = 0$, thus $\ell^1 \ni f_n \stackrel{w}{\to} 0 \Rightarrow \|f_n\|_1 \to 0$.

Assume that $f_n \stackrel{w}{\to} 0$, but $\|f_n\| \not\to 0$. Since $\delta_m \in \ell^\infty \cong (\ell^1)^*$, the first fact clearly implies $f_n(m) = \varphi_{\delta_m}(f_n) \stackrel{n\to\infty}{\longrightarrow} 0$ for all m. And by the second assumption there exists $\varepsilon > 0$ such that $\|f_n\|_1 \geq \varepsilon$ for infinitely many n. Using this, we inductively define $\{n_k\}, \{r_k\} \subseteq \mathbb{N}$ as follows:

(a) Let n_1 be the smallest number for which $\|f_{n_1}\|_1 \geq \varepsilon$.

(b) Let r_1 be the smallest number for which $\sum_{i=1}^{r_1} |f_{n_1}(i)| \geq \frac{\varepsilon}{2}$ and $\sum_{i=r_1+1}^{\infty} |f_{n_1}(i)| \leq \frac{\varepsilon}{5}$.

For $k \geq 2$:

(c) Let n_k be the smallest number such that $n_k > n_{k-1}$ and $\|f_{n_k}\|_1 \geq \varepsilon$ and $\sum_{i=1}^{r_{k-1}} |f_{n_k}(i)| \leq \frac{\varepsilon}{5}$.

(d) Let r_k be the smallest number such that $r_k > r_{k-1}$ and $\sum_{i=r_{k-1}+1}^{r_k} |f_{n_k}(i)| \geq \frac{\varepsilon}{2}$ and $\sum_{i=r_k+1}^{\infty} |f_{n_k}(i)| \leq \frac{\varepsilon}{5}$.

The reader should convince herself that the existence of such n_k, r_k follows from our assumptions!

Now define $\{c_i\}_{i\in\mathbb{N}}$ by $c_i = \overline{\operatorname{sgn}(f_{n_k}(i))}$ where k is uniquely determined by $r_{k-1} < i \leq r_k$ with $r_0 = 0$. Now clearly $c = \{c_i\} \in \ell^\infty$, and for all k we have, using the lower bound in (b),(d),

$$\sum_{i=r_{k-1}+1}^{r_k} c_i f_{n_k}(i) = \sum_{i=r_{k-1}+1}^{r_k} |f_{n_k}(i)| \geq \frac{\varepsilon}{2},$$

while using $|c_i| \leq 1$ and the upper bounds in (b),(c),(d) we have

$$\left|\sum_{i=1}^{r_{k-1}} c_i f_{n_k}(i)\right| \leq \sum_{i=1}^{r_{k-1}} |f_{n_k}(i)| \leq \frac{\varepsilon}{5}, \qquad \left|\sum_{i=r_k+1}^{\infty} c_i f_{n_k}(i)\right| \leq \sum_{i=r_k+1}^{\infty} |f_{n_k}(i)| \leq \frac{\varepsilon}{5}.$$

Thus $|\varphi_c(f_{n_k})| \geq \frac{\varepsilon}{2} - \frac{\varepsilon}{5} - \frac{\varepsilon}{5} = \frac{\varepsilon}{10} > 0$ for all k, so that $\varphi_c(f_n) \not\to 0$. Since this contradicts the assumption $f_n \stackrel{w}{\to} 0$, we must have $\|f_n\|_1 \to 0$. ■

B.29 Remark 1. The above proof followed the argument in [9] quite closely. The method of proof is called the 'gliding (or sliding) hump method' (which is a variant of the earlier method with the equally colorful name 'condensation of singularities'). The gliding hump is precisely the dominant contribution to $\varphi_c(f_{n_k})$ coming from the i in the interval $\{r_k+1, \ldots, r_{k+1}\}$, which moves to infinity as $k \to \infty$. For a high-brow interpretation of Schur's theorem in terms of Banach space bases see [1, Section 2.3]. But also this discussion uses gliding humps!

2. With the appearance of the proof of the uniform boundedness theorem using Baire's theorem the gliding hump method fell a bit out of fashion. (But never completely, cf. e.g. [161].) Also Schur's theorem can be proven using Baire's theorem, cf. e.g. [30, Proposition V.5.2], but this is conceptually and technically more complicated.

3. Note also that the determination of the n_k, r_k in the above proof was deterministic, using no choice axiom at all. In this sense the proof is better than the alternative one using Baire's theorem, thus countable dependent choice, which nevertheless is instructive. (But of course also the above proof is non-constructive in the somewhat extremist sense of intuitionism since the necessary $\varepsilon > 0$ cannot be found algorithmically.) In the same vein, in next section we will give a gliding-hump proof for the weak uniform boundedness theorem that only uses the axiom AC_ω of countable choice. □

B.3.6 Compactness of all bounded linear maps $c_0 \to \ell^q \to \ell^p$ $(1 \le p < q < \infty)$

In Theorem 12.24 we proved that for a reflexive Banach V space all bounded linear operators $c_0 \to V$ and $V \to \ell^1$ are compact. Applying this to the reflexive spaces $\ell^p = \ell^p(\mathbb{N}, \mathbb{F})$ with $1 < p < \infty$, we see that for such p all bounded maps $c_0 \to \ell^p$ and $\ell^p \to \ell^1$ are compact. We now give an alternative proof that in addition gives compactness of all bounded maps $\ell^q \to \ell^p$ for $1 < p < q < \infty$ and $c_0 \to \ell^1$. Thus all bounded linear maps between the spaces c_0, ℓ^p $(1 \le p < \infty)$ that go in the opposite direction of the bounded inclusion maps $\ell^1 \hookrightarrow \ell^p \hookrightarrow \ell^q \hookrightarrow c_0$ are compact!

B.30 THEOREM (H. R. PITT 1936) *Let $1 \le p < q < \infty$ and let $V \subseteq \ell^q$ or $V \subseteq c_0$ be a closed subspace. Then every bounded linear map $V \to \ell^p$ is compact, thus $B(V, \ell^p) = K(V, \ell^p)$.*

We follow the elementary proof by Delpech [38], based upon the following

B.31 LEMMA *(i) Let $1 \le r < \infty$, $x \in \ell^r$ and $\{y_n\} \subset \ell^r$ a weak null-sequence. Then*

$$\limsup_{n\to\infty} \|x + y_n\|^r = \|x\|^r + \limsup_{n\to\infty} \|y_n\|^r. \tag{B.4}$$

(ii) If $x \in c_0$ and $\{y_n\} \subset c_0$ is a weak null-sequence then

$$\limsup_{n\to\infty} \|x + y_n\| = \max(\|x\|, \limsup_{n\to\infty} \|y_n\|). \tag{B.5}$$

Proof. (i) Since the evaluation map $\varphi_m : \ell^r \to \mathbb{F}$, $y \mapsto y(m)$ is in $(\ell^r)^*$, the weak convergence $y_n \xrightarrow{w} 0$ implies $y_n(m) \to 0$ for all m. From this it is quite clear that (B.4) holds if x has finite support $\mathrm{supp}(x) = \{m \in \mathbb{N} \mid x(m) \ne 0\}$ (i.e. $x \in c_{00}$) since in particular $y_n(m) \to 0$ for all $m \in \mathrm{supp}(x)$. For general $x \in \ell^r$ and $\varepsilon > 0$ we pick $x' \in c_{00}$ with $\|x - x'\| < \varepsilon$. Since $\|x' + y_n\| - \varepsilon \le \|x + y_n\| \le \|x' + y_n\| + \varepsilon$ for all n we find

$$\left(\|x'\|^r + \limsup_{n\to\infty} \|y_n\|^r\right)^{1/r} - \varepsilon \le \limsup_{n\to\infty} \|x + y_n\| \le \left(\|x'\|^r + \limsup_{n\to\infty} \|y_n\|^r\right)^{1/r} + \varepsilon.$$

Combined with the fact that $\|x'\| \to \|x\|$ as $\varepsilon \to 0$, (B.4) follows.

The proof of (ii) is essentially the same. ■

Proof of Theorem B.30. By Exercise 9.23, ℓ^q is reflexive for $1 < q < \infty$, thus same holds for every closed $V \subseteq \ell^q$ by Theorem 9.27. If $V \subseteq c_0$ is closed then $V^* \cong c_0^*/V^\perp$ by Exercise 9.15. Since $c_0^* \cong \ell^1$ is separable, the same holds for V^* by Exercise 6.5. In either case, V is reflexive or has separable dual. Thus compactness of $A \in B(V, \ell^p)$ follows from Theorem 12.28(iv) if we prove sequential weak-norm continuity. It suffices to prove that $x_n \xrightarrow{w} 0$ implies $\|Ax_n\| \to 0$. If $\varepsilon > 0$ we pick $y_\varepsilon \in V$ with $\|y_\varepsilon\| = 1$ and $\|Ay_\varepsilon\| > 1 - \varepsilon$. Let $t > 0$. In the case $V \subseteq \ell^q$ we apply Lemma B.31 to both sides of the inequality $\|Ay_\varepsilon + A(tx_n)\| \le \|y_\varepsilon + tx_n\|$, obtaining

$$\left(\|Ay_\varepsilon\|^p + t^p \limsup_{n\to\infty} \|Ax_n\|^p\right)^{1/p} \le \left(\|y_\varepsilon\|^q + t^q \limsup_{n\to\infty} \|x_n\|^q\right)^{1/q}. \tag{B.6}$$

Since $\{x_n\}$ is weakly convergent, Exercise 10.6 gives $\|x_n\| \le M \; \forall n$. With $\|y_\varepsilon\| = 1$ and $\|Ay_\varepsilon\| > 1 - \varepsilon$, (B.6) becomes

$$\limsup_{n\to\infty} \|Ax_n\|^p \le \frac{1}{t^p}\left((1 + t^q M^q)^{p/q} - (1 - \varepsilon)^p\right),$$

which holds for all $t > 0$ and $\varepsilon \in (0,1)$. Putting $t = \varepsilon^{1/q}$ and using $(1+x)^\alpha = 1 + \alpha x + o(x)$ as $x \searrow 0$, we have

$$\limsup_{n\to\infty} \|Ax_n\|^p \leq \frac{1}{\varepsilon^{p/q}}\big((1+\frac{p}{q}\varepsilon M^q) - (1-p\varepsilon) + o(\varepsilon)\big) = \frac{\varepsilon}{\varepsilon^{p/q}}\big(\frac{p}{q}M^q + p + o(1)\big),$$

which vanishes as $\varepsilon \to 0$ since $p < q$. This finishes the proof in the case $V = \ell^q$.

In the case $V \subseteq c_0$ we proceed similarly, but use part (ii) of the Lemma on the r.h.s., obtaining

$$\Big(\|Ay_\varepsilon\|^p + t^p \limsup_{n\to\infty} \|Ax_n\|^p\Big)^{1/p} \leq \max\big(\|y_\varepsilon\|, t \limsup_{n\to\infty} \|x_n\|\big),$$

leading similarly to the above to

$$\limsup_{n\to\infty} \|Ax_n\|^p \leq \frac{1}{t^p}\Big(\max(1, tM)^p - (1-\varepsilon)^p\Big).$$

Putting $t = \varepsilon^{1/2p}$, for ε and thus t small enough, this becomes

$$\limsup_{n\to\infty} \|Ax_n\|^p \leq \frac{1}{\varepsilon^{1/2}}\Big(1 - (1-\varepsilon)^p\Big) = \frac{p\varepsilon + o(\varepsilon)}{\varepsilon^{1/2}},$$

which again vanishes as $\varepsilon \to 0$. ■

B.32 Corollary *Let $V, W \in \{\ell^p \mid 1 \leq p < \infty\} \cup \{c_0\}$ with $V \neq W$. Then there is no isomorphism from an infinite-dimensional subspace of V to a subspace of W. Equivalently, every $A \in B(V,W)$ is strictly singular. In particular $V \not\simeq W$.*

Proof. Assume $K \subseteq V$, $L \subseteq W$ are infinite-dimensional closed subspaces and $A \in B(K,L)$ is an isomorphism, thus has a bounded inverse $B \in B(L,K)$. Then by Theorem B.30 either A or B is compact. But this implies that $BA = \mathrm{id}_K$ is compact, which is impossible by infinite-dimensionality of K. Thus no isomorphism $K \simeq L$ (let alone $V \simeq W$) can exist.

The equivalence of this statement to strict singularity was Exercise 12.23(i). ■

B.33 Remark 1. One says that the spaces c_0 and ℓ^p ($1 \leq p < \infty$) are uncomparable.

2. By the above, in particular the inclusion maps $\ell^p \hookrightarrow \ell^q \hookrightarrow c_0$, where $1 \leq p < q < \infty$, are strictly singular. But they are clearly non-compact since they send the bounded sequence $\{\delta_n\}$, which has no norm-convergent subsequence in any of these spaces, to itself. Taking $1 < p < q < \infty$ this shows that strict singularity of an operator does not imply compactness even if both spaces involved are reflexive. But one can prove that for all $A \in B(\ell^p)$, where $1 < p < \infty$, strict singularity does imply compactness. (Cf. e.g. [1, Problem 2.4].) By an easy reduction the same holds for all operators between not necessarily separable Hilbert spaces. □

B.4 The weak Uniform Boundedness Theorem using only AC_ω

The aim of this section is to give a proof of the following theorem that only uses the axiom AC_ω of countable choice. The proof uses gliding humps, but less evidently than in Section B.3.5.

B.34 Theorem *ZF+AC_ω $\Rightarrow$ If E is a Banach space, F a normed space and $\mathcal{F} \subseteq B(E,F)$ is pointwise bounded then $\mathcal{F}$ is uniformly bounded.*

Proof. Assume that $\mathcal{F}$ is not uniformly bounded. Then the sets $\mathcal{F}_n = \{A \in \mathcal{F} \mid \|A\| \geq 4^n\}$ are all non-empty, so that using AC_ω (axiom of countable choice), we can pick an $A_n \in \mathcal{F}_n$ for each $n \in \mathbb{N}$. By definition of $\|A_n\|$, the sets $X_n = \{x \in E \mid \|x\| \leq 1, \ \|A_n x\| \geq \frac{2}{3}\|A_n\|\}$ are all non-empty, to that using AC_ω again, we can choose an $x_n \in X_n$ for each $n \in \mathbb{N}$.

Applying the triangle inequality to $Az = \frac{1}{2}(A(y+z) - A(y-z))$ gives

$$\|Az\| = \frac{1}{2}\|(A(y+z) - A(y-z))\| \leq \frac{1}{2}(\|A(y+z)\| + \|A(y-z)\|) \leq \max(\|A(y+z)\|, \|A(y-z)\|).$$

Applying this inequality to $A = A_{n+1}$, $y = y_n$, $z = \pm 3^{-(n+1)}x_{n+1}$, recalling $\|A_n x_n\| \geq \frac{2}{3}\|A_n\|$, we see that for at least one of the signs $\pm$ we have

$$\|A_{n+1}(y_n \pm 3^{-(n+1)}x_{n+1})\| \geq 3^{-(n+1)}\|A_{n+1}x_{n+1}\| \geq 3^{-(n+1)}\frac{2}{3}\|A_{n+1}\|.$$

Thus defining a sequence $\{y_n\} \subseteq E$ inductively by $y_1 = x_1$ and

$$y_{n+1} = \begin{cases} y_n + 3^{-(n+1)}x_{n+1} & \text{if } \ \|A_{n+1}(y_n + 3^{-(n+1)}x_{n+1})\| \geq 3^{-(n+1)}\frac{2}{3}\|A_{n+1}\| \\ y_n - 3^{-(n+1)}x_{n+1} & \text{otherwise} \end{cases} \tag{B.7}$$

we have $\|A_n y_n\| \geq \frac{2}{3}3^{-n}\|A_n\|$ for all n. (For $n = 1$ this is true since $y_1 = x_1$.) Since (B.7) involves no further free choices, this inductive definition can be formalized in ZF (which we don't spell out here, see [53]).

With (B.7) and $\|x_n\| \leq 1$ for all n, we have $\|y_{n+1} - y_n\| \leq 3^{-(n+1)} \ \forall n$. Now for all $m > n$

$$\|y_m - y_n\| = \left\| \sum_{k=n}^{m-1} y_{k+1} - y_k \right\| \leq \sum_{k=n}^{\infty} 3^{-(k+1)} = 3^{-(n+1)}\frac{1}{1-\frac{1}{3}} = \frac{1}{2}3^{-n},$$

so that $\{y_n\}$ is a Cauchy sequence. By completeness of E we have $y_n \to y \in E$ with $\|y - y_n\| \leq \frac{1}{2}3^{-n}$. Another use of the triangle inequality gives

$$\|A_n y_n\| = \|A_n(y - y + y_n)\| \leq \|A_n y\| + \|A_n(y - y_n)\| \leq \|A_n y\| + \|A_n\|\|y - y_n\|,$$

so that with $\|y - y_n\| \leq \frac{1}{2}3^{-n}$, $\|A_n y_n\| \geq \frac{2}{3}3^{-n}\|A_n\|$ and $\|A_n\| \geq 4^n$ for all n we finally have

$$\|A_n y\| \geq \|A_n y_n\| - \|A_n\|\|y - y_n\| \geq \|A_n\|\left(\frac{2}{3}3^{-n} - \frac{1}{2}3^{-n}\right) = \frac{1}{6}3^{-n}\|A_n\| \geq \frac{1}{6}\left(\frac{4}{3}\right)^n \to \infty.$$

Thus $y \in E$ is a witness for the failure of pointwise boundedness of $\mathcal{F}$. ■

B.35 Remark The above argument was discovered only a few years ago and published [53] in 2017! □

B.5 Alaoglu's theorem ⇔ Tychonov $\restriction T_2$ ⇔ Ultrafilter lemma

Most proofs in functional analysis are quite non-constructive in that they rely on the Axiom of Choice (AC). But most results actually do not require the full strength of AC! We have seen that the axiom AC_ω of countable choice suffices for proving the weak uniform boundedness theorem, whereas the somewhat stronger axiom DC_ω of countable dependent choice (which is equivalent to Baire's theorem) yields the strong uniform boundedness and open mapping theorems, as well as Hahn-Banach for separable normed spaces.

In the proof of Alaoglu's Theorem 10.26 we have used Tychonov's theorem. The latter is well known to be equivalent (over ZF) to the Axiom of Choice and to Zorn's lemma. (For proofs see e.g. [108].) But an inspection of the proof shows that we used Tychonov's theorem only for a product of Hausdorff spaces. Thus:

B.36 Theorem *Over ZF, Tychonov's theorem for Hausdorff spaces $\Rightarrow$ Alaoglu's theorem.*

It turns out that Alaoglu's theorem actually is equivalent over ZF to Tychonov's theorem for Hausdorff spaces, as well as to several other statements, e.g. the Ultrafilter Lemma (UL), Alexander's subbase lemma, and the Boolean Prime Ideal Theorem. Furthermore, this class of equivalent statements is strictly weaker than the equivalent statements AC, Zorn and Tychonov, in the sense that there are models of the ZF axioms in which the former statements are true, but not the latter, see [70, 71]. We will prove Alaoglu $\Rightarrow$ UL below.

B.37 Definition *Let X be a set.*

- *A filter*[118] *on X is a family $\mathcal{F} \subseteq P(X)$ of subsets satisfying:*
  - *(i) If $F, G \in \mathcal{F}$ then $F \cap G \in \mathcal{F}$.*
  - *(ii) If $F \in \mathcal{F}$ and $G \supseteq F$ then $G \in \mathcal{F}$.*
  - *(iii) $\emptyset \notin \mathcal{F}$.*
  - *(iv) $\mathcal{F} \neq \emptyset$.*

  Notice that filters on X are partially ordered by inclusion (as subsets of $P(X)$).
- *An ultrafilter (or maximal filter) on a set is a filter that is not properly contained in another filter.*

B.38 Example 1. If X is any set and $x \in X$ then $\mathcal{F}_x = \{Y \subseteq X \mid x \in Y\}$ is an ultrafilter on X. The filters $\mathcal{F}_x$ are called prinicipal.

2. If (X, τ) is a topological space and $x \in X$ then $N \subseteq X$ is called a neighborhood of x if there exists $U \in \tau$ such that $x \in U \subseteq N$. Now the set $\mathcal{N}_x$ of neighborhoods of x is a filter on X. Only the closedness under finite intersections merits an argument: If $N_1, N_2 \in \mathcal{N}_x$ then there are open $U_i \subseteq N_i$ containing x. Then $U_1 \cap U_2$ is an open neighborhood of x, thus in $\mathcal{F}$, and so is $N_1 \cap N_2 \supseteq U_1 \cap U_2$. ($\mathcal{N}_x$ is principal if and only if x is isolated, i.e. $\{x\}$ is open.)

B.39 Lemma (Ultrafilter Lemma) *ZF+AC $\Rightarrow$ Every filter is contained in an ultrafilter.*

Proof. Let X be a set and $\mathcal{F}$ a filter on X. The family $\mathfrak{F}$ of all filters on X that contain $\mathcal{F}$ is a partially ordered set w.r.t. inclusion. If $\mathcal{C} \subseteq \mathfrak{F}$ is a totally ordered subset of $\mathfrak{F}$, we claim that the union $\bigcup \mathcal{C}$ of all elements of $\mathcal{C}$ is a filter that contains $\mathcal{F}$. That the union of any non-zero number of filters has the properties (ii), (iii) and (iv) in Definition B.37 is obvious, so that only (i) remains. Let $F_1, F_2 \in \bigcup \mathcal{C}$. By the total order of $\mathcal{C}$, there is a $\widetilde{\mathcal{F}} \in \mathcal{C}$ such that $F_1, F_2 \in \widetilde{\mathcal{F}}$ and thus $F_1 \cap F_2 \in \widetilde{\mathcal{F}} \subseteq \bigcup \mathcal{C}$. This proves requirement (i), thus $\bigcup \mathcal{C}$ is in $\mathfrak{F}$ and is an upper bound for the chain $\mathcal{C}$. Therefore Zorn's lemma applies and gives a maximal filter $\widehat{\mathcal{F}}$ containing $\mathcal{F}$. Since ultrafilters are just maximal filters, we are done. ■

Ultrafilters are characterized by a quite remarkable property:

B.40 Lemma *A filter $\mathcal{F}$ on X is an ultrafilter if and only if for every $Y \subseteq X$ exactly one of the alternatives $Y \in \mathcal{F}$, $X \backslash Y \in \mathcal{F}$ holds.*

[118] Filters were invented in 1937 by the French mathematician Henri Cartan (1904-2008), an important member of the Bourbaki group. Unsurprisingly, the best reference on filters is [21]. Preference for nets or filters is sometimes put as a question of American vs. European (in particular French) tastes, but this is simplistic. Most contemporary research in general topology is actually done in terms of filters, not nets.

Proof. We begin by noting that we cannot have both $Y \in \mathcal{F}$ and $X \backslash Y \in \mathcal{F}$ since (i) would imply $\emptyset = Y \cap (X \backslash Y) \in \mathcal{F}$, which is forbidden by (iii). Assume $\mathcal{F}$ contains Y or $X \backslash Y$ for every $Y \subseteq X$. This means that $\mathcal{F}$ cannot be enlarged by adding $Y \subseteq X$ since either already $Y \in \mathcal{F}$ or else $X \backslash Y \in \mathcal{F}$, which excludes $Y \in \mathcal{F}$. Thus $\mathcal{F}$ is an ultrafilter.

Now assume that $\mathcal{F}$ is an ultrafilter and $Y \subseteq X$. If there is an $F \in \mathcal{F}$ such that $F \cap Y = \emptyset$ then $F \subseteq X \backslash Y$, and property (ii) implies $X \backslash Y \in \mathcal{F}$. If, on the other hand, $Y \cap F \neq \emptyset \ \forall F \in \mathcal{F}$ then there is a filter $\widetilde{\mathcal{F}}$ containing $\mathcal{F}$ and Y. Since $\mathcal{F}$ is maximal, we must have $Y \in \mathcal{F}$. ■

B.41 Remark If $\mathcal{F}$ is a filter on a set $S \neq \emptyset$ then the map $\mu = \chi_{\mathcal{F}} : P(S) \to \{0, 1\}$ (that sends $A \subseteq S$ to 1 if $A \in \mathcal{F}$ and to 0 otherwise) clearly satisfies $\mu(\emptyset) = 0$, $\mu(S) = 1$, and for for disjoint $A_1, A_2 \subseteq S$ we have $\mu(A_1) + \mu(A_2) \leq \mu(A_1 \cup A_2) \leq 1$ since A_1 and A_2 cannot both be in $\mathcal{F}$. If equality holds for all disjoint non-empty A_1, A_2, then in particular for all $\emptyset \neq A_1 \neq S$ and $A_2 = S \backslash A_1$. Thus either A_1 or its complement A_2 belongs to $\mathcal{F}$. By Lemma B.40 this characterizes the ultrafilters on S. Conversely, one easily checks that if μ is a non-zero $\{0, 1\}$-valued finitely additive measure on S then $\mathcal{F} = \{A \subseteq S \mid \mu(A) = 1\}$ is an ultrafilter. □

B.42 Lemma *For a topological space (X, τ), the following are equivalent:*

(i) (X, τ) is compact (thus every open cover has a finite subcover).

(ii) Whenever $\mathcal{F} \subseteq P(X)$ is a family of closed subsets of X such that $\bigcap \mathcal{F} = \emptyset$ then there are $C_1, \ldots, C_n \in \mathcal{F}$ such that $C_1 \cap \cdots \cap C_n = \emptyset$.

(iii) If $\mathcal{F} \subseteq P(X)$ is a family of closed subsets of X with the finite intersection property (i.e. the intersection of any finite number of elements of $\mathcal{F}$ is non-empty) then $\bigcap \mathcal{F} \neq \emptyset$.

Proof. (i) and (ii) are dualizations of each other, using de Morgan's formulas, and (iii) is the contraposition of (ii). ■

B.43 Theorem *Over ZF, the following are equivalent:*

(i) The Ultrafilter Lemma.

(ii) Tychonov$\restriction T_2$: Any product of compact Hausdorff spaces is compact.

(iii) Alagolu's theorem.

Proof. (i)⇒(ii) This is 'just' general topology, and we do not reproduce it here. Cf. e.g. [108].

(ii)⇒(iii) Mentioned above.

(iii)⇒(i) Let $\mathcal{F}$ be a filter on the set X. Then $V = \ell^\infty(X, \mathbb{R})$ is a Banach space, and $\Sigma = (V^*)_{\leq 1}$ (which is a set of finitely additive measures on X, cf. Section B.3.2) is weak-$*$ compact, thus weak-$*$-closed, by Alaoglu's theorem. Every $x \in X$ gives rise to a bounded linear functional $\varphi_x \in \Sigma$, $f \mapsto f(x)$ with $\|\varphi_x\| = 1$. The map $\iota : X \to \Sigma, x \mapsto \varphi_x$ is injective. Now put $\overline{\mathcal{F}} = \{\overline{\iota(F)}^{w^*} \mid F \in \mathcal{F}\} \subseteq P(\Sigma)$. If $F_1, \ldots, F_n \in \mathcal{F}$ then by injectivity of ι and finite intersection property of $\mathcal{F}$ we have $\bigcap_k \overline{\iota(F_k)}^{w^*} \supseteq \bigcap_k \iota(F_k) = \iota(\bigcap_k F_k) \neq \emptyset$, so that $\overline{\mathcal{F}}$ has the finite intersection property. Since the sets $\overline{\iota(F)}^{w^*} \subseteq \Sigma$ are weak-$*$ closed, and Σ is weak-$*$ compact, Lemma B.42 gives $\bigcap \overline{\mathcal{F}} \neq \emptyset$. Pick $\psi \in \bigcap \overline{\mathcal{F}} \subseteq \Sigma \subseteq V^*$ and define a map $\mu : P(X) \to \mathbb{C}$, $S \mapsto \psi(\chi_S)$.

Now $\ell^\infty(X, \mathbb{R})$ is an algebra and each φ_x is a character. Since $\psi \in \overline{\iota(F)}^{w^*}$ for each $F \in \mathcal{F}$, it also is a character: We have $\psi = \lim_\lambda \varphi_\lambda$, where φ_λ is a net of characters converging in the weak-$*$ topology, thus

$$\psi(fg) = \lim_\lambda \varphi_\lambda(fg) = \lim_\lambda \varphi_\lambda(f)\varphi_\lambda(g) = \psi(f)\psi(g).$$

And χ_S is idempotent for each S, thus $\psi(\chi_S) = \psi(\chi_S^2) = \psi(\chi_S)^2$, implying $\mu(S) = \psi(\chi_S) \in \{0,1\}$ for all $S \subseteq X$. We have $\mu(X) = \psi(1) = 1$ (since $\varphi_x(1) = 1\ \forall x$), and $S \cap T = \emptyset$ implies $\chi_{S\cup T} = \chi_S + \chi_T$, so that $\mu(S \cup T) = \mu(S) + \mu(T)$. Thus μ is a finitely additive $\{0,1\}$-valued measure on X, and we know from Remark B.41 that $\widehat{\mathcal{F}} = \{Y \subseteq X \mid \mu(Y) = 1\}$ is an ultrafilter on X. If $Y \in \mathcal{F}$ then $\psi \in \bigcap \overline{\mathcal{F}} = \bigcap_{F\in\mathcal{F}} \overline{\iota(F)}^{\mathrm{w}^*}$ implies $\psi \in \overline{\iota(Y)}^{\mathrm{w}^*} = \overline{\{\varphi_x \mid x \in Y\}}^{\mathrm{w}^*}$. Since $\varphi_x(\chi_Y) = \chi_Y(x) = 1$ for all $x \in Y$, we have $\mu(Y) = \psi(\chi_Y) = 1$, thus $\mathcal{F} \subseteq \widehat{\mathcal{F}}$. We thus have embedded $\mathcal{F}$ into an ultrafilter. ■

B.44 Remark 1. Combining Theorems B.43 and B.45 we obtain the curious fact that over ZF Alaoglu's theorem (non-reversibly) implies Hahn-Banach.

2. Our proofs of OMT/BIT/CGT and UBT only use the ZF axioms and Baire's theorem, the latter being equivalent to DC_ω. Since we have seen that Alaolgu's theorem is equivalent to the Ultrafilter lemma (and to Tychonov for Hausdorff spaces and various other statements), all these theorems can be proven in the framework ZF+DC_ω+UL that is provably weaker than ZF+AC, see [124]. In Section B.6.1 we will prove that this also holds for Hahn-Banach. Yet, there are some results, like the Krein-Milman theorem that cannot be proven in ZF+DC_ω+UL. See Remark B.85. □

B.6 More on convexity and Hahn-Banach matters

B.6.1 Tychonov for Hausdorff spaces implies Hahn-Banach

Having seen that Alaoglu's theorem can be deduced over ZF from Tychonov's theorem for Hausdorff spaces, we now prove the same for the Hahn-Banach theorem:

B.45 Theorem *[Łoś & Ryll-Nardzewski (1951)]* [119] *Over ZF, Tychonov's theorem for Hausdorff spaces implies the Hahn-Banach Theorem 9.2.*

Proof. Let V, p, W, φ as in Theorem 9.2. Define $\mathcal{E} = \prod_{v\in V}[-p(-v), p(v)]$ with the product topology. Since we assume Tychonov's theorem for compact Hausdorff spaces, $\mathcal{E}$ is compact (and Hausdorff). Clearly every $e \in \mathcal{E}$ can be interpreted as a map $V \to \mathbb{R}$ satisfying the bound $-p(-v) \le e(v) \le p(v)\ \forall v$. For each $v \in V$ the coordinate map $e \mapsto e(v)$ is continuous, thus $\mathcal{E}' = \{e \in \mathcal{E} \mid e(w) = \varphi(w)\ \forall w \in W\} \subseteq \mathcal{E}$ is closed. For each finite-dimensional subspace $Z \subseteq V$ let $\mathcal{E}_Z = \{e \in \mathcal{E} \mid e \restriction Z \text{ is linear}\}$. Again using continuity of the coordinate maps $e \mapsto e(v)$, it follows that each

$$\mathcal{E}_Z = \bigcap_{x,x'\in Z} \{e \in \mathcal{E} \mid e(x + x') = e(x) + e(x')\} \cap \bigcap_{c\in\mathbb{R}, x\in Z} \{e \in \mathcal{E} \mid e(cx) = ce(x)\} \subseteq \mathcal{E}$$

is closed. If $Z \subseteq V$ is finite-dimensional, applying Lemma 9.3 a finite number of times, we find a linear extension ψ_Z of φ to $W + Z$ bounded by p. Defining $e \in \mathcal{E}$ by $e(x) = \psi_Z(x)$ for $x \in W + Z$ and $e(x) = p(x)$ otherwise, we have $e \in \mathcal{E}'_Z = \mathcal{E}' \cap \mathcal{E}_Z$, thus $\mathcal{E}'_Z \ne \emptyset$. If $Z, Z' \subseteq V$ are finite-dimensional then $\mathcal{E}'_Z \cap \mathcal{E}'_{Z'} \supseteq \mathcal{E}'_{Z+Z'} \ne \emptyset$. Thus the family $\{\mathcal{E}'_Z \mid Z \subseteq V \text{ fin. dim.}\}$ of closed subsets of $\mathcal{E}$ has the finite intersection property. Since $\mathcal{E}$ is compact, Lemma B.42 gives $\mathcal{E}_{HB} = \bigcap_{Z \text{ fin. dim.}} \mathcal{E}'_Z \ne \emptyset$. Pick any $e \in \mathcal{E}_{HB}$. Now $e : V \to \mathbb{R}$ coincides with φ on W, satisfies the p-bound and is linear on all finite-dimensional subspaces, thus globally. Thus it is a Hahn-Banach extension. ■

[119] Jerzy Łoś (1920-1988), Czesław Ryll-Nardzewski (1926-2015). Polish mathematicians. Ł. mostly worked in set theory and logic, R.-N. in functional analysis (R.-N. fixed point theorem), measure theory and probability.

B.46 Remark 1. The proof in [99] is phrased in terms of two rather more general theorems, but we chose to sacrifice the generality and simplify the argument considerably. (There also is a technical point: In the above proof we used distinguished coordinates $p(x)$ to make sure that the closed subsets $\mathcal{E}'_Z$ are non-empty. It is not clear to this author how to draw this conclusion (without invoking AC) in the generality of [99, Theorem 2] when the spaces $\{P_x\}_{x\in X_0}$ appearing there have nothing to do with each other. Perhaps they should be assumed pointed?)

2. In 1962, Luxemburg[120] [100] deduced Hahn-Banach from the UL by use of ultraproducts (non-standard analysis). The ultraproducts are not essential and can be removed, cf. [11] (or [108]). The resulting proof shares some features with the above, which however remains simpler.

3. In [123] it is shown that the Hahn-Banach theorem is strictly weaker than the equivalent statements of Tychonov for Hausdorff spaces, the Ultrafilter Lemma, etc. (But it still suffices for proving the Banach-Tarski paradox and the existence of sets that are not Lebesgue measurable, see [54, 117]. Thus the latter cannot be avoided without giving up Hahn-Banach.) □

B.6.2 Minkowski functionals. Criteria for normability and local convexity

We begin with an easy fact needed by all studies of convex sets in topological vector spaces:

B.47 Exercise Let V be a topological vector space and $A \subseteq V$ convex. Prove that the interior A^0 and the closure $\overline{A}$ are convex.

B.48 Proposition *Let V be a topological vector space and U a convex open neighborhood of 0. Define the 'Minkowski functional'[121] $\mu_U : V \to [0,\infty)$ of U by*

$$\mu_U(x) = \inf\{t \geq 0 \mid x \in tU\}.$$

Then μ_U is sublinear and continuous, and $U = \{x \in X \mid \mu_U(x) < 1\}$.

Proof. As $t \to \infty$ we have $t^{-1}x \to 0$. Since U is an open neighborhood of 0, we have $t^{-1}x \in U$ for t large enough. Thus $\mu_U(x) < \infty$ for each $x \in V$. It is quite obvious from the definition that $\mu_U(cx) = c\mu_U(x)$ for $c > 0$. Thus μ_U is positive-homogeneous. We have $\mu_U(x) < 1$ if and only if there exists $t \in (0,1)$ such that $x \in tU$. Thus $\mu_U(x) < 1 \Rightarrow x \in U$. And if $x \in U$ then openness of U implies that $(1-\varepsilon)x \in U$ for some $\varepsilon > 0$. Thus $\mu_U(x) < 1$, so that we have $U = \{x \in X \mid \mu_U(x) < 1\}$.

Let $x, y \in V$, and let $s, t > 0$ such that $x \in sU, y \in tU$. I.e. there are $a, b \in U$ such that $x = sa, y = tb$. Thus $x + y = sa + tb = (s+t)\frac{sa+tb}{s+t}$. Since $\frac{s}{s+t}a + \frac{t}{s+t}b \in U$ due to convexity of U, we have $x + y \in (s+t)U$. Thus $\mu_U(x+y) \leq s+t$, and since we have $x \in sU, y \in tU$ for all $s < \mu_U(x) + \varepsilon, t < \mu_U(y) + \varepsilon$ with $\varepsilon > 0$, the conclusion is $\mu_U(x+y) \leq \mu_U(x) + \mu_U(y)$, thus subadditivity. Being subadditive and positive homogeneous, μ_U is sublinear.

Let $\{x_\iota\}_{\iota\in I} \subseteq V$ be a net converging to zero. For each $n \in \mathbb{N}$, $n^{-1}U$ is an open neighborhood of zero. Thus there exists a $\iota_n \in I$ such that $\iota \geq \iota_n$ implies $x_\iota \in n^{-1}U$ and therefore, with the definition of μ_U, that $\mu_U(x_\iota) \leq n^{-1}$. Thus $\mu_U(x_\iota) \to 0$, which is continuity of μ_U at $0 \in V$.

If now $x_\iota \to x$ then the subadditivity of μ_U gives

$$\mu_U(x) - \mu_U(x - x_\iota) \leq \mu_U(x_\iota) \leq \mu_U(x) + \mu_U(x_\iota - x),$$

and since $\mu_U(x_\iota - x) \to 0$, we have $\mu_U(x_\iota) \to \mu_U(x)$, thus continuity of μ_U at all $x \in V$. ■

[120] W. A. J. Luxemburg (1929-2018). Dutch mathematician, mainly interested in non-standard analysis.

[121] Also called gauge functionals, gauge deriving from Minkowski's 'Aich…'.

B.49 Definition *Let V be a topological vector space and $0 \in U \subseteq V$. Then U is called*

- *balanced if $x \in U, |\lambda| \le 1 \Rightarrow \lambda x \in U$,*
- *bounded if for every open $W \ni 0$ there exists $\lambda > 0$ such that $\lambda U \subseteq W$.*

Note that if U is convex and contains zero, multiplication by $t \in [0,1]$ sends U into itself. Thus for checking balancedness it suffices to consider $|\lambda| = 1$.

B.50 Exercise Let (V, τ) be a TVS, where $\tau = \tau_d$ for a translation-invariant metric d. Prove:

(i) $B(0, r)$ is bounded in the TVS sense for each $r > 0$.

(ii) $X \subseteq V$ is bounded in the TVS sense if and only if $X \subseteq B(0, r)$ for some $r > 0$.

B.51 Proposition *Let (V, τ) be a topological vector space and U a convex open neighborhood of zero. Then*

(i) The Minkowski functional μ_U is a seminorm if and only if U is balanced.

(ii) If U is bounded then $\mu_U(x) = 0$ implies $x = 0$.

(iii) If U is balanced and bounded then $\|x\| = \mu_U(x)$ is a norm inducing the topology τ.

Proof. (i) Since μ_U is subadditive and positive-homogeneous, it is a seminorm if and only if $\mu_U(\lambda x) = \mu_U(x)$ for all $x \in V$ and $\lambda \in \mathbb{F}$ with $|\lambda| = 1$. If U is balanced then this is evidently satisfied. Now assume $\mu_U(\lambda x) = \mu_U(x)$. The openness of U implies that $\{t > 0 \mid x \in tU\} = (\mu_U(x), \infty)$. Thus if $|\lambda| = 1$ then the assumption $\mu_U(\lambda x) = \mu_U(x)$ implies that $x \in U$ if and only if $\lambda x \in U$. Thus U is balanced.

(ii) Assume that U is bounded and that $x \ne 0$. Since τ is T_1, there is an open $W \subseteq V$ such that $0 \in W \not\ni x$. Since U is bounded, there is $\lambda > 0$ such that $\lambda U \subseteq W$, which clearly implies $x \notin \lambda U$. Now the definition of μ_U implies $\mu_U(x) > \lambda > 0$.

(iii) Proposition B.48 and the above (i) and (ii) show that $\|\cdot\| = \mu_U$ is a continuous norm on V. Thus $x_n \to 0$ implies $\|x_n\| \to 0$. If we prove the converse implication then $\tau = \tau_{\|\cdot\|}$ follows since V is a topological vector space. Let $\{x_n\}$ be a sequence such that $\|x_n\| \to 0$, and let W be an open neighborhood of 0. Since U is bounded, there is $\lambda > 0$ such that $\lambda U \subseteq W$. Now, $\|x_n\| \to 0$ means that there is $n_0 \in \mathbb{N}$ such that $n \ge n_0 \Rightarrow \|x_n\| < \lambda/2$. With the definition of μ_U this implies $x_n \in \lambda U$, thus $x_n \in \lambda U \subseteq W$ for all $n \ge n_0$. This proves $x_n \to 0$. ■

We now know that a topological vector space is normable if the zero element has a balanced convex bounded open neighborhood. (The converse is easy.) But this can be improved:

B.52 Lemma *Let V be a topological vector space and U a convex open neighborhood of 0. Then there exists a balanced convex open neighborhood $U' \subseteq U$ of 0.*

Proof. Since multiplication by scalars is continuous, there exists an $\varepsilon > 0$ such that $\lambda U \subseteq U$ whenever $|\lambda| \le \varepsilon$. Thus with $W = |\varepsilon| U$ we have $tW \subseteq U$ whenever $|t| \le 1$. Put $Y = \bigcup_{|t| \le 1} tW \subseteq U$. By construction, Y is a balanced open neighborhood of 0.

For every $\lambda \in \mathbb{F}$ with $|\lambda| = 1$ it is clear that λU is a convex open neighborhood of 0. Putting $Z = \bigcap_{|\lambda|=1} \lambda U$, it is manifestly clear that Z is balanced and $0 \in Z$. Furthermore, U' is convex (as an intersection of convex sets). Since $tW \subseteq U$ for all $|t| = 1$, we have $Y \subseteq Z$, so that Z has non-empty interior Z^0. Now we put $U' = Z^0$ and claim that U' has the desired properties. Clearly U' is an open neighborhood of 0, as the interior of a convex set it is convex (Exercise B.47). If $|t| = 1$ then the map $Z \to Z, x \mapsto tx$ is a homeomorphism. Thus if $x \in Z^0 = U'$ then

$tx \in Z^0 = U'$, showing that $U' = Z^0$ is balanced. ■

Now we are in a position to prove geometric criteria for normability and local convexity of topological vector spaces:

B.53 THEOREM *Let V be a topological vector space. Then V is normable if and only if there exists a bounded convex open neighborhood of* 0.

Proof. If V is normable by the norm $\|\cdot\|$ then $B_{\|\cdot\|}(0,1) = \{x \in V \mid \|x\| < 1\}$ is clearly open, convex (and balanced). To show boundedness, let $W \ni 0$ be open. Then there is $\varepsilon > 0$ such that $B(0,\varepsilon) \subseteq W$. Now clearly $\varepsilon B(0,1) = B(0,\varepsilon) \subseteq W$, thus $B(0,1)$ is bounded.

If there exists a bounded convex open neighborhood U of 0 then by Lemma B.52 we can assume U in addition to be balanced. (The U' provided by the lemma is a subset of U, thus bounded if U is bounded.) Now by Proposition B.51(iii), μ_U is a norm inducing the given topology on V. ■

B.54 THEOREM *A topological vector space (V,τ) is locally convex in the sense of Definition 2.37 (i.e. the topology τ comes from a separating family $\mathcal{F}$ of seminorms) if and only if it is Hausdorff and the zero element has an open neighborhood base consisting of convex sets.*

Proof. Given a separating family $\mathcal{F}$ of seminorms and putting $\tau = \tau_{\mathcal{F}}$, a basis of open neighborhoods of 0 is given by the finite intersections of sets $U_{p,\varepsilon} = \{x \in V \mid p(x) < \varepsilon\}$, where $p \in \mathcal{F}$, $\varepsilon > 0$. Each of the $U_{p,\varepsilon}$ is convex and open, thus also their finite intersections.

And if τ has the stated property, Lemma B.52 gives that 0 has a neighborhood base consisting of balanced convex open sets. Defining $\mathcal{F} = \{\mu_U \mid U \text{ balanced convex open neighborhood of } 0\}$, each of the μ_U is a continuous seminorm by Propositions B.48 and B.51. Thus if $x_\iota \to 0$ then $\|x_\iota\|_U := \mu_U(x_\iota) \to 0$. And $\|x_\iota\|_U \to 0$ for all balanced convex open U implies that x_ι ultimately is in every open neighborhood of 0, thus $x_\iota \to 0$. Thus $\tau = \tau_{\mathcal{F}}$, and 2.36 gives that $\mathcal{F}$ is separating. ■

B.55 EXERCISE Let $0 < p < 1$.

(i) Prove that $(\ell^p(S,\mathbb{F}), \tau_{d_p})$ is normable if S is finite.

(ii) Prove that the open unit ball of $(\ell^p(S,\mathbb{F}), \tau_{d_p})$ does not contain any convex open neighborhood of 0 if S is infinite.

(iii) Prove that $(\ell^p(S,\mathbb{F}), \tau_{d_p})$ is neither normable nor locally convex if S is infinite.

B.6.3 Hahn-Banach separation theorems for locally convex spaces

The Hahn-Banach theorem in the sublinear functional version (Theorem 9.2) has an important geometric application, namely the fact that disjoint convex sets in a topological vector space can be separated by hyperplanes, i.e. sets $H = \{x \in X \mid \operatorname{Re}\varphi(x) = t\}$ for some $\varphi \in V^*$, $t \in \mathbb{R}$.

B.56 DEFINITION *Let V be a topological vector space over $\mathbb{F}$ and V^* the space of continuous linear functionals $V \to \mathbb{F}$. For $A, B \subseteq V$ we say that A and B are*

(i) separated if there exists $\varphi \in V^*$ with $\operatorname{Re}\varphi(a) < \inf_{b\in B} \operatorname{Re}\varphi(b)$ $\forall a \in A$.

(ii) strictly separated if there exist $\varphi \in V^*$, $\alpha \in \mathbb{R}$ with $\operatorname{Re}\varphi(a) < \alpha < \operatorname{Re}\varphi(b)$ $\forall a \in A, b \in B$.

(iii) very strictly separated if there exists $\varphi \in V^*$ with $\sup_{x\in A} \operatorname{Re}\varphi(x) < \inf_{x\in B} \operatorname{Re}\varphi(x)$.

B.57 Remark 1. Geometrically, (ii) and (iii) are equivalent to A, B being contained in two disjoint open half spaces (having positive distance in case (iii)), while (i) means that A, B are contained in an open half-space and its complement, respectively.

2. The sets $A = (0,1), B = (0,1)$ in $V = \mathbb{R}$ are strictly separated, but not very strictly, while $A = (0,1), B = [1,2)$ are only separated. An example for two closed sets that are non-strictly separated is $A = \{(x,y) \mid x, y > 0,\ xy \geq 1\}$, $B = \{(x,y) \mid x \leq 0\}$ in $V = \mathbb{R}^2$. □

B.58 Theorem *Let V be a topological vector space and $A, B \subseteq V$ disjoint non-empty convex subsets, A being open. Then*

(i) A and B are separated.

(ii) If also B is open, A and B are strictly separated.

Proof. (i) Pick $a_0 \in A, b_0 \in B$ and put $z = b_0 - a_0$ and $U = (A - a_0) - (B - b_0) = A - B + z$, which is a convex (as pointwise sum of two convex sets) open (since $U = \bigcup_{x \in -B - a_0 + b_0}(A + x)$) neighborhood of 0. Let $p = \mu_U$ be the associated Minkowski functional. As a consequence of $A \cap B = \emptyset$ we have $0 \notin A - B$, thus $z \notin U$, and therefore $p(z) \geq 1$.

Put $W = \mathbb{R}z$ and define $\psi : W \to \mathbb{R}$, $cz \mapsto c$. For $c \geq 0$ we have $\psi(cz) = c \leq cp(z) = p(cz)$. Thus by sublinearity of p and Theorem 9.2 there exists a linear functional $\varphi : V \to \mathbb{R}$ satisfying $\varphi \restriction W = \psi$, thus $\varphi(cz) = c$, and $\varphi(x) \leq p(x)\ \forall x \in V$. Thus also $-p(-x) \leq -\varphi(-x) = \varphi(x)$, and since $x \to 0$ implies $p(x) \to 0$, φ is continuous at zero, thus everywhere.

If now $a \in A$, $b \in B$ then $a - b + z \in U$, so that $p(a - b + z) < 1$. Thus

$$\varphi(a - b) + 1 = \varphi(a - b + z) \leq p(a - b + z) < 1,$$

thus $\varphi(a) < \varphi(b)$ for all $a \in B$ and $b \in B$. Thus the subsets $\varphi(A), \varphi(B)$ of $\mathbb{R}$ are disjoint. Since A, B are convex, they are connected. Consequently, $\varphi(A), \varphi(B)$ are connected, thus intervals. Since A is open, so is $\varphi(A)$ (open mapping theorem). If we put $s = \sup \varphi(A)$, we have $\varphi(a) < s \leq \varphi(b)$ for all $a \in A, b \in B$, and this is equivalent to $\varphi(a) \leq \inf_{b \in B} \varphi(b)$ for all $a \in A$.

$\mathbb{F} = \mathbb{C}$: Considering V as $\mathbb{R}$-vector space, apply the above to obtain a continuous $\mathbb{R}$-linear functional $\varphi_0 : V \to \mathbb{R}$ such that $\varphi(a) < \inf_{b \in B} \varphi(b)\ \forall a \in A$. Now define $\varphi : V \to \mathbb{C}$, $x \mapsto \varphi_0(x) - i\varphi_0(ix)$. This clearly is continuous and satisfies $\operatorname{Re} \varphi = \varphi_0$, so that the desired inequality holds. That φ is $\mathbb{C}$-linear follows from the same argument as in the proof of Theorem 9.5.

(ii) It suffices to consider $\mathbb{F} = \mathbb{R}$. With $\alpha = \inf_{b \in B} \operatorname{Re} \varphi(b)$, by (i) we have $\operatorname{Re} \varphi(a) < \alpha$ for all $a \in A$. Since B is open, $\varphi(B) \subseteq \mathbb{R}$ is open by Exercise 6.4 (or the open mapping theorem), so that φ does not assume its infimum α on B, whence $\alpha < \operatorname{Re} \varphi(b)\ \forall b \in B$. ■

B.59 Lemma *If V is a locally convex space and $K \subseteq U \subseteq V$ with K compact and U open, there is a convex open neighborhood $N \subseteq V$ of zero such that $K + N \subseteq U$.*

Proof. For brevity, we only prove the Lemma for normed spaces and leave the generalization to locally convex spaces as an Exercise.

For every $x \in K$ there exists $\varepsilon_x > 0$ such that $B(x, 2\varepsilon_x) \subseteq U$. Since $\{B(x, \varepsilon_x)\}_{x \in K}$ is an open cover of the compact set K, there are $x_1, \ldots, x_n$ such that $K \subseteq \bigcup_{i=1}^n B(x_i, \varepsilon_{x_i})$. Put $\varepsilon = \min(\varepsilon_{x_1}, \ldots, \varepsilon_{x_n}) > 0$ and $N = B(0, \varepsilon)$, which is open and convex. Now

$$K + N \subseteq \left(\bigcup_{i=1}^n B(x_i, \varepsilon_{x_i}) \right) + B(0, \varepsilon) \subseteq \bigcup_{i=1}^n B(x_i, 2\varepsilon_{x_i}) \subseteq U,$$

where we used $B(x_i, 2\varepsilon_{x_i}) \subseteq U\ \forall i$. ■

B.60 Corollary *Let V be a locally convex vector space and $A, B \subseteq V$ disjoint non-empty convex subsets with A compact and B closed. Then A and B are very strictly separated.*

Proof. We only discuss $\mathbb{F} = \mathbb{R}$, the changes for $\mathbb{F} = \mathbb{C}$ being the same as above. Applying Lemma B.59 to $K = A$, $U = V \backslash B$, we obtain a convex open $N \ni 0$ such that $A + N \subseteq V$. It is easy to show that $A + N$ is open and convex. (Do it!) Applying Theorem B.58(i) to $A + N$ and B, we obtain $\varphi \in V^*$ such that $\varphi(a) < \inf_{b \in B} \varphi(b)$ $\forall a \in A + N \supset A$. Since φ assumes its supremum on the compact set A, we have $\sup_{a \in A} \varphi(a) < \inf_{b \in B} \varphi(b)$. ■

B.61 Corollary *Let V be a locally convex vector space and $W \subsetneq V$ a proper closed subspace. Then for every $x \in V \backslash W$ there exists a continuous linear functional $\varphi \in V^*$ such that $\varphi \restriction W = 0$ and $\varphi(x) \neq 0$. In particular, for every $x \in V \backslash \{0\}$ there exists $\varphi \in V^*$ with $\varphi(x) \neq 0$.*

Proof. Let W, x as stated. Put $A = \{x\}$, $B = W$. Then A and B are disjoint closed convex subsets, and A is compact. Thus they are very strictly separated by Corollary B.60, thus there exists $\varphi \in V^*$ such that $\operatorname{Re} \varphi(x) > \sup_{w \in W} \operatorname{Re} \varphi(w)$. Since W is a linear subspace, finiteness of the supremum implies $\varphi \restriction W = 0$.

For the final claim, apply the above with $W = \{0\}$. ■

B.6.4 First applications of the separation theorems to Banach spaces

B.62 Proposition *Every norm-closed convex set in a Banach space is weakly closed.*

Proof. Let V be a Banach space and $X \subseteq V$ convex and norm-closed. If $y \in \overline{X}^{\mathrm{w}} \backslash X$ then $\{y\}$ is compact, so that applying Corollary B.60 to $(V, \tau_{\|\cdot\|})$ and $A = \{y\}, B = X$ there is a continuous linear functional $\varphi \in V^*$ such that $\inf_{x \in X} \operatorname{Re} \varphi(x) > \operatorname{Re} \varphi(y)$. If now $\{x_\iota\}$ is a net in X with $x_\iota \stackrel{\mathrm{w}}{\to} y$ then $\varphi(x_\iota) \to \varphi(y)$, but this contradicts $\inf_{x \in X} \operatorname{Re} \varphi(x) > \operatorname{Re} \varphi(y)$. Thus X is weakly closed. ■

B.63 Corollary *The weak and norm closures of a convex set in a Banach space coincide.*

Proof. Then $X \subseteq V$ be convex. Then also $\overline{X}^{\|\cdot\|}$ is convex, thus weakly closed, so that $\overline{X}^{\mathrm{w}} \subseteq \overline{\overline{X}^{\|\cdot\|}}^{\mathrm{w}} = \overline{X}^{\|\cdot\|}$. Combining this with the obvious inclusion $\overline{X}^{\mathrm{w}} \supseteq \overline{X}^{\|\cdot\|}$ we are done. ■

B.64 Proposition *If V is a Banach space, $X \subset V$ is closed, convex and balanced and $y \in V \backslash X$ then there exists $\varphi \in V^*$ such that $|\varphi(x)| \leq 1$ $\forall x \in X$ and $|\varphi(y)| > 1$.*

Proof. Let X, y be as stated. As in the proof of Proposition B.62 there exists $\varphi \in V^*$ such that $\inf_{x \in X} \operatorname{Re} \varphi(x) > \operatorname{Re} \varphi(y)$. With $\varphi' = -\varphi$ this becomes $\operatorname{Re} \varphi'(y) > \sup_{x \in X} \operatorname{Re} \varphi'(x)$. Now

$$|\varphi'(y)| \geq \operatorname{Re} \varphi'(y) > \sup_{x \in X} \operatorname{Re} \varphi'(x) = \sup_{x \in X} |\varphi'(x)|,$$

where the final equality is due to the fact that X is balanced, thus in particular closed under multiplication by all λ with $|\lambda| = 1$. Now $(\sup_{x \in X} |\varphi'(x)|)^{-1} \varphi'$ has the desired properties. ■

B.65 Theorem (Goldstine) *If V is a Banach space then $V_{\leq 1}$ is $\sigma(V^{**}, V^*)$-dense in $(V^{**})_{\leq 1}$.*

Proof. We abbreviate $\tau = \sigma(V^{**}, V^*)$. The unit ball $(V^{**})_{\leq 1}$ is τ-compact by Alaoglu's theorem, thus τ-closed, so that $B = \overline{V_{\leq 1}}^{\tau}$, which is convex by Exercise B.47, is contained in $(V^{**})_{\leq 1}$. If this inclusion is strict, pick $x^{**} \in (V^{**})_{\leq 1} \backslash \overline{V_{\leq 1}}^{\tau}$. Then x^{**} has a τ-open neighborhood U disjoint from B, and by Theorem B.54 there is a convex open $A \subseteq U$. Now Theorem B.58(i) applied to

(V^{**}, τ) and $A, B \subseteq V^{**}$ gives a $\tau = \sigma(V^{**}, V^*)$-continuous linear functional $\varphi \in (V^{**})^\star$ such that $\mathrm{Re}\,\varphi(a) < \inf_{b\in B} \mathrm{Re}\,\varphi(b)\ \forall a \in A$. Now Exercise 10.23 gives $\varphi \in V^* \subseteq V^{***}$.

Putting $\psi = -\varphi$ we have $\sup_{b\in B} \mathrm{Re}\,\psi(b) < \mathrm{Re}\,\psi(a)\ \forall a \in A$, which is more convenient. Since $\psi \in V^*$ and $B \supseteq V_{\leq 1}$, we have $\|\psi\| \leq \sup_{b\in B} \mathrm{Re}\,\psi(b)$. On the other hand, with $x^{**} \in A$ and $\|x^{**}\| \leq 1$, we have $\mathrm{Re}\,\psi(x^{**}) \leq |\psi(x^{**})| \leq \|x^{**}\|\|\psi\| \leq \|\psi\|$. Combining these findings, we have $\|\psi\| \leq \sup_{b\in B} \mathrm{Re}\,\psi(b) < \mathrm{Re}\,\psi(x^{**}) \leq \|\psi\|$, which is absurd. This contradiction proves $\overline{V_{\leq 1}}^{\tau} = (V^{**})_{\leq 1}$. ■

B.6.5 Convex hulls and closed convex hulls

B.66 Definition *If V is a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $X \subseteq V$ then the convex hull $\mathrm{conv}(X)$ of X is the intersection of all convex $Y \subseteq V$ containing X. (Clearly this is the smallest convex set containing X.)*

B.67 Exercise Prove: If V is a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $X \subseteq V$ then

$$\mathrm{conv}(X) = \Big\{ \sum_{i=1}^{I} t_i x_i \mid I \in \mathbb{N},\ x_i \in X,\ t_i \geq 0,\ \sum_{i=1}^{I} t_i = 1 \Big\}. \tag{B.8}$$

B.68 Theorem (Carathéodory 1911) [122] *If $X \subseteq \mathbb{R}^d$ then every point in $\mathrm{conv}(X)$ is a a convex combination of at most $d+1$ points of X. Thus (B.8) holds with fixed $I = d+1$.*

Proof. (Sketch) By (B.8), every element of $\mathrm{conv}(X)$ is a convex combination of finitely many points of X. Now one uses the fact that, just as more than n points in $\mathbb{R}^n$ must be linearly dependent, more than $n+1$ points in the n-dimensional affine space must be affinely dependent. Thus one can remove points from the convex combination until $I \leq d+1$. For details see e.g. [96, Theorem 2.23] (or Wikipedia). ■

B.69 Corollary *If $X \subseteq \mathbb{R}^d$ is compact then $\mathrm{conv}(X)$ is compact.*

Proof. The simplices $\Delta_d = \{(t_0, \ldots, t_d) \mid t_i \geq 0, \sum_i t_i = 1\}$ are compact, and by Carathéodory's theorem, $\mathrm{conv}(X)$ is the image of an obvious continuous map $X^{d+1} \times \Delta_d \to X$. If X is compact, so is $X^{d+1} \times \Delta_d$, so that $\mathrm{conv}(X)$ is compact, thus closed. ■

We will see below that closedness of $\mathrm{conv}(X)$ does not follow from closedness of X, and for infinite-dimensional V not even from compactness of X.

B.70 Exercise Prove: If V is a topological vector space and $X \subseteq V$ is open then $\mathrm{conv}(X)$ is open.

B.71 Exercise Let $X = \{(x, y) \mid x \in \mathbb{R},\ y \geq \frac{1}{1+x^2}\} \subset \mathbb{R}^2$ (which is closed). Prove that $\mathrm{conv}(X) = \{(x, y) \mid x \in \mathbb{R},\ y > 0\}$ (which is open, but not closed).

B.72 Exercise Let $V = \ell^2(\mathbb{N}, \mathbb{R})$ with standard basis $\{\delta_n\}$. Prove:

(i) $X = \{n^{-1}\delta_n \mid n \in \mathbb{N}\} \cup \{0\}$ is compact.

(ii) $\mathrm{conv}(X) \subseteq V$ is not closed (thus not compact).

[122] Constantin Carathéodory (1873-1950), Greek mathematician with many contributions to real and complex analysis and other areas and author of many textbooks.

B.73 Exercise Let V be a topological vector space and $X \subseteq V$. Prove that $\overline{\mathrm{conv}(X)}$ coincides with the intersection of all closed convex sets that contain X.

B.74 Definition *If V is a topological vector space and $X \subseteq V$ we call $\overline{\mathrm{conv}(X)}$ the closed convex hull of X, but mostly write $\overline{\mathrm{conv}}(X)$ for readability.*

B.75 Proposition (Mazur's compactness theorem (S. Mazur 1930)) *Let V be a normed space and $X \subseteq V$. Then*

(i) If X is totally bounded then so is $\mathrm{conv}(X)$.

(ii) If V is Banach and X is precompact then $\mathrm{conv}(X)$ is precompact and $\overline{\mathrm{conv}}(X)$ compact.

Proof. (i) Let $X \subseteq V$ be totally bounded, and let $\varepsilon > 0$. Then there are $z_1, \ldots, z_K \in X$ such that $X \subseteq \bigcup_{k=1}^K B(z_k, \varepsilon)$. Now $C = \mathrm{conv}(\{z_1, \ldots, z_K\}) \subseteq V$ is compact as the image in V of the compact set $\{(t_1, \ldots, t_K) \mid t_k \geq 0, \sum_{k=1}^K t_k = 1\} \subseteq \mathbb{R}^K$ under the continuous map $(t_1, \ldots, t_K) \mapsto \sum_{k=1}^K t_k z_k$. Thus there are $y_1, \ldots, y_L \in C \subseteq \mathrm{conv}(X)$ such that $C \subseteq \bigcup_{l=1}^L B(y_l, \varepsilon)$.

If now $y \in \mathrm{conv}(X)$ then by Exercise B.67 we have $y = \sum_{m=1}^M t_m x_m$ for certain $x_1, \ldots, x_M \in X$, $t_1, \ldots, t_M \geq 0$ with $\sum_{m=1}^M t_m = 1$. For each m pick k_m such that $\|x_m - z_{k_m}\| < \varepsilon$. Putting $y' = \sum_{m=1}^M t_m z_{k_m} \in C$, we have $\|y - y'\| = \|\sum_{m=1}^M t_m(x_m - z_{k_m})\| \leq \sum_{m=1}^M t_m \|x_m - z_{k_m}\| < \varepsilon$. Picking l such that $\|y' - y_l\| < \varepsilon$, we have $\|y - y_l\| \leq \|y - y'\| + \|y' - y_l\| < 2\varepsilon$. This proves $\mathrm{conv}(X) \subseteq \bigcup_{l=1}^L B(y_l, 2\varepsilon)$. Since $\varepsilon > 0$ was arbitrary, $\mathrm{conv}(X)$ is totally bounded.

(ii) Since V is complete, total boundedness and precompactness of its subsets are equivalent by Exercise A.43. Thus $\mathrm{conv}(X)$ is precompact by (i), and by definition this means that $\overline{\mathrm{conv}}(X)$ is compact. ■

B.76 Remark Giving a description of $\overline{\mathrm{conv}}(X)$ that is similarly explicit as (B.8) is difficult. The convex hull of $X = \{x_1, \ldots, x_n\}$ is compact, thus coincides with $\overline{\mathrm{conv}}(X)$. Now assume $X = \{x_1, x_2, \ldots\}$ is a bounded sequence, i.e. $\|x_i\| \leq C \ \forall i$. Then for all $\{t_i \geq 0\}$ with $\sum_i t_i = 1$ we have $\sum_i t_i \|x_i\| \leq C$, thus $\sum_i t_i x_i$ converges absolutely to some $x \in V$. It is easy to see that $x \in \overline{\mathrm{conv}}(X)$, but in general it is not clear that every $x \in \overline{\mathrm{conv}}(X)$ is of the form $\sum_i t_i x_i$. This can be proven if the sequence $\{x_n\}$ is convergent. See [102, Lemma 3.4.29] for the proof due to Grothendieck. □

B.6.6 Extreme points and faces of convex sets. The Krein-Milman theorem

B.77 Definition *Let V be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $X \subseteq V$ convex.*

(i) A subset $F \subseteq X$ is a face of X if it is convex and $tx + (1-t)y \in F$ with $x, y \in X, 0 < t < 1$ implies $x, y \in F$.

(ii) $x \in X$ is an extreme point if $\{x\}$ is a face. (Thus x is not a non-trivial convex combination.) The set of extreme points of X is denoted $E(X)$.

B.78 Exercise If $X \subseteq \mathbb{R}^2$ is a convex polygon (non-degenerate, i.e. with non-empty interior), prove that the faces of X are X, the edges of the boundary and the corner singletons.

B.79 Exercise Let X be a convex set. Prove that $x \in X$ is an extreme point if and only if $X \backslash \{x\}$ is convex.

B.80 Exercise Let $x \in F \subseteq X \subseteq V$, where V is a vector space, X is convex, F is a face of X and x an extreme point of F. Prove that x is an extreme point of X. (Thus $E(F) \subseteq E(X)$.)

B.81 Exercise Let $V = \mathbb{R}^3$ and let X be the closed convex hull of

$$S = \{(x, y, 0) \mid (x-1)^2 + y^2 = 1\} \cup \{(0,0,1), (0,0,-1)\}.$$

Determine the set E of extreme points of X and show that it its not closed.

B.82 Exercise Show that the closed unit ball in the Banach space c_0 has no extreme points.

B.83 Exercise Show that the closed unit ball in the Banach space $L^1([0,1], \lambda)$, where λ is Lebesgue measure on the Borel σ-algebra, has no extreme points.

B.84 Theorem (Krein-Milman 1940) [123] *Every non-empty compact convex subset of a locally convex space is the closed convex hull $\overline{\mathrm{conv}}(E)$ of its set E of extreme points. (In particular $E \neq \emptyset$.)*

Proof. Let V be locally convex and $X \subseteq V$ compact and convex. We may assume $X \neq \emptyset$ since otherwise there is nothing to prove. Let $\mathcal{F}$ be the set of compact faces of X. Then $X \in \mathcal{F}$, thus $\mathcal{F} \neq \emptyset$. Partially order $\mathcal{F}$ by reverse inclusion. If $\mathcal{C} \subseteq \mathcal{F}$ is a chain (totally ordered subset) then it has the finite intersection property, so that $F = \bigcap \mathcal{C}$ is non-empty by Lemma B.42. Assume $x \in F$ satisfies $x = ty + (1-t)z$, where $0 < t < 1$ and $y, z \in X$ with $y \neq z$. Since every $G \in \mathcal{C}$ contains x and is a face, it follows that $x \in G$. Thus $x \in \bigcap_{G \in \mathcal{C}} G = \bigcap \mathcal{C} = F$, proving that F is a face and compact. Thus F is an upper bound for $\mathcal{C}$, so that by Zorn's lemma $\mathcal{F}$ has a maximal element, thus a face that is minimal (in the sense of not having a proper subset that is a compact face).

We claim that F is a singleton. Assuming $x, y \in F$ with $x \neq y$, Corollary B.61 provides a continuous linear functional $\varphi \in V^*$ such that $\varphi(x - y) \neq 0$. Multiplying by a phase we can achieve $\mathrm{Re}\,\varphi(x) \neq \mathrm{Re}\,\varphi(y)$. Put $M = \sup_{x \in F} \mathrm{Re}\,\varphi(x)$ and $F' = \{x \in F \mid \mathrm{Re}\,\varphi(x) = M\}$. Then $F' \neq \emptyset$ since $\mathrm{Re}\,\varphi$ assumes its supremum on the compact set F. Now $F' \subseteq F$ is a closed subset and convex by linearity of φ. If $z, z' \in F$, $0 < t < 1$ such that $tz + (1-t)z' \in F'$ then in view of $t\mathrm{Re}\,\varphi(z) + (1-t)\mathrm{Re}\,\varphi(z') = M$ we must have $\mathrm{Re}\,\varphi(z) = \mathrm{Re}\,\varphi(z') = M$, thus $z, z' \in F'$, proving that $F' \subseteq F$ is a compact face. Since the face F is minimal by construction, we have $F' = F$. Thus $x, y \in F'$, implying the contradiction $\mathrm{Re}\,\varphi(x) = M = \mathrm{Re}\,\varphi(y)$. Thus F is a singleton. Since its element is an extreme point, we have $E \neq \emptyset$.

It remains to prove that $X_0 = \overline{\mathrm{conv}}(E)$ equals X. If this is not true, picking $z \in X \backslash X_0$, Corollary B.60 provides a $\varphi \in V^*$ such that $\sup_{x \in X_0} \mathrm{Re}\,\varphi(x) < \mathrm{Re}\,\varphi(z)$. Define $M = \sup_{x \in X} \mathrm{Re}\,\varphi(x)$ and $F = \{x \in X \mid \mathrm{Re}\,\varphi(x) = M\}$. Since X is compact, $\mathrm{Re}\,\varphi$ assumes its supremum, thus $F \neq \emptyset$. As before F is a compact face of X, and applying the above to F there exists an extreme point $y \in F$. Now $\mathrm{Re}\,\varphi(y) = M$. By Exercise B.80, y is an extreme point of X, thus $y \in E \subseteq X_0$. This implies the contradiction $M = \mathrm{Re}\,\varphi(y) \leq \sup_{x \in X_0} \mathrm{Re}\,\varphi(x) < \mathrm{Re}\,\varphi(z) \leq M$. Thus $\overline{\mathrm{conv}}(E) = X$. ■

B.85 Remark 1. With a little extra effort one can prove $E \subseteq S$ for every closed set $S \subseteq X$ such that $\overline{\mathrm{conv}}(S) = X$.

2. Exercise B.82 shows that 'compact' cannot be replaced by 'closed bounded' in Theorem B.84.

3. Also the local convexity of V is essential: In the metrizable, but not locally convex, TVS $L^p([0,1])$ with $0 < p < 1$ there exist [137] compact convex sets without extreme points. □

[123] Mark Grigorievich Krein (1907-1989), David Milman (1912-1982), Soviet mathematicians. (Milman emigrated to Israel.) Among other things, Krein is also known for the Tannaka-Krein duality theory for compact groups, Milman e.g. for the M.-Pettis theorem, cf. Section B.6.8).

B.86 COROLLARY *Let V be a Banach space. Then*

(i) $V^*_{\le 1} = \overline{\text{conv}}^{\text{w}^*}(E(V^*_{\le 1}))$ *(weak-$*$ closure).*

(ii) If V is reflexive then $V_{\le 1} = \overline{\text{conv}}^{\text{w}}(E(V_{\le 1}))$ *(weak closure).*

Proof. Since the weak and weak-$*$ topologies are locally convex, this follows by combining Krein-Milman with Alaoglu's Theorem for (i) and with the weak compactness of $V_{\le 1}$ for (ii). ∎

This proves that the closed unit ball of a Banach space has 'enough' extreme points whenever V is reflexive or a dual space. (In view of this Exercise B.82 again shows that c_0 is not a dual space.) While ℓ^1 and ℓ^∞ are non-reflexive, they are dual spaces, so that their unit balls have extreme points. They can be identified:

B.87 EXERCISE Let S be a set and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Prove:

(i) For $V = \ell^1(S, \mathbb{F})$, the set of extreme points of $V_{\le 1}$ is $\{c\delta_s \mid s \in S, c \in \mathbb{F}, |c| = 1\}$.

(ii) For $V = \ell^\infty(S, \mathbb{F})$, the set of extreme points of $V_{\le 1}$ is $\{f \in V \mid |f(s)| = 1 \ \forall s \in S\}$.

B.88 REMARK 1. The statement that the closed unit ball $V^*_{\le}$ has extreme points for every Banach space V implies AC. This is proven in less than a page in [13]. In view of Corollary B.86 we thus have the implication Alaoglu+KM $\Rightarrow$ AC. (But HB+KM $\not\Rightarrow$ AC.) Recall that over ZF, Alaoglu's theorem is equivalent to the Ultrafilter Lemma, Tychonov's theorem for Hausdorff spaces, etc. Since ZF+UL+DC$_\omega$ is provably weaker than AC, it follows that the Krein-Milman theorem cannot be proven in the framework ZF+UL+DC$_\omega$ (which does suffice for OMT, UBT, HBT and Alaoglu).

2. In Exercise B.67 we found a representation for all elements of the convex hull of a set. Doing a similar thing for the closed convex hull is much harder. One solution is the following Choquet-Bishop-de Leeuw[124] theorem: If V is a locally convex space, $X \subset V$ is compact and convex, and $x \in X$ then there is a probability measure μ on the set E of extreme points such that $x = \int_E y \, d\mu(y)$ as a weak integral: For every $\varphi \in V^*$ we have $\varphi(x) = \int_E \varphi(y) d\mu(y)$. For a thorough discussion of the rather technical subject see [119].

3. Using the Choquet-Bishop-de Leeuw theorem, one easily proves Rainwater's[125] theorem: If $\{x_n\}$ is a (norm)bounded sequence in a Banach space, then for the weak convergence $x_n \stackrel{w}{\to} x$ it is sufficient that $\varphi(x_n) \to \varphi(x)$ for all $\varphi \in E(V^*_{\le 1})$.

4. If V is locally convex and $X \subset V$ compact convex then $x \in X$ is called an exposed point if there is a $\varphi \in V^*$ such that $\text{Re}\, \varphi(x) > \text{Re}\, \varphi(y) \ \forall y \in X \backslash \{x\}$. It is easy to see that every exposed point is extremal, but that the converse need not hold. Still, one can prove that X is the closed convex hull of its set of exposed points provided X is metrizable, see [8]. Clearly this applies if V is Banach.

5. There is much more to be said about the relation between convex sets and linear (and non-linear!) functionals. We want at least to mention another important result of Bishop and Phelps: If V is a Banach space over $\mathbb{R}$ and $X \subset V$ is non-empty, convex, closed and bounded then the set of bounded linear functionals that assume their supremum on X is (norm-)dense in V^*:

$$\overline{\{\varphi \in V^* \mid \exists x_0 \in X : |\varphi(x_0)| = \sup_{x \in X} |\varphi(x)|\}} = V^*.$$

[124] Gustave Choquet (1915-2006), French analyst, proved this under the assumption of metrizability of X. This hypothesis was removed by to the American mathematicians Errett Albert Bishop (1928-1983), encountered earlier, and Karel de Leeuw (1930-1978, murdered by a PhD student).

[125] John Rainwater is a ficticious mathematician, see wikipedia or `http://at.yorku.ca/t/o/p/d/47.htm`, but "his" results are not.

(Note that this is trivial if X is compact.) This can fail in complex Banach spaces, but it does hold for the closed unit ball in a complex Banach space, giving the result mentioned in Remark 9.25.2.) See e.g. [102, Section 2.11]. □

B.6.7 Strictly convex Banach spaces and Hahn-Banach uniqueness

If V is a Banach space, let E be the set of extreme points of the closed unit ball $V_{\leq 1}$. If $0 < \|x\| < 1$ then with $t = \|x\|, y = x/\|x\|, z = 0$ we have $x = ty + (1-t)z$, thus $x \notin E$. Similarly $0 \neq E$, thus $E \subseteq V_1 = \{x \in V \mid \|x\| = 1\}$. The question is whether $V_1 \subseteq E$. Assume $x \in V_1$ is a non-trivial convex combination $x = ty + (1-t)z$ with $y, z \in V_{\leq 1}$. Then $1 = \|x\| \leq t\|y\| + (1-t)\|y\|$, which is possible only if $y, z \in V_1$. But this is all we can say in general since there are Banach spaces with $E = \emptyset$ or $\emptyset \neq E \subsetneqq V_1$, cf. Exercises B.82 and B.87.

B.89 PROPOSITION *For every Banach space V, the following are equivalent:*

(i) $x, y \in V$, $\|x\| = \|y\| = 1, x \neq y$ *implies* $\|x + y\| < 2$.

(ii) If $x, y \in V$ *satisfy* $\|x + y\| = \|x\| + \|y\|$ *then* $y = 0$ *or* $x = cy$ *with* $c \geq 0$.

(iii) The set of extreme points of the closed unit ball $V_{\leq 1}$ *equals* $V_1 = \{x \in V \mid \|x\| = 1\}$.

Proof. (i)⇒(ii) Let $x, y \in V$ satisfy $\|x + y\| = \|x\| + \|y\|$. If $x = 0$ or $y = 0$ then we are done. By rescaling and/or exchanging if necessary we may assume $1 = \|x\| \leq \|y\|$. Put $z = y/\|y\|$. Then

$$\begin{aligned} 2 \geq \|x + z\| &= \|x + y - (1 - \|y\|^{-1})y\| \geq \|x + y\| - (1 - \|y\|^{-1})\|y\| = \|x + y\| - \|y\| + 1 \\ &= \|x\| + \|y\| - \|y\| + 1 = 2, \end{aligned}$$

where we used $\|a - b\| \geq \|a\| - \|b\|$ and the assumptions $\|x + y\| = \|x\| + \|y\|$ and $\|x\| = 1$. This implies $\|x + z\| = 2$. Since $\|x\| = 1 = \|z\|$ (by assumption and by construction of z), the strict convexity implies $x = z = y/\|y\|$. Thus $y = \|y\|x$, and we have proven (ii).

(ii)⇒(i) Assume $x, y \in V$, $x \neq y$ and $\|x\| = \|y\| = 1$. If $\|x + y\| = 2$ was true then (ii) would give $x = cy$ with $c > 0$, but then $\|x\| = \|y\|$ gives $c = 1$, thus the contradiction $x = y$. Since $\|x + y\| \leq 2$ is obvious, we must have $\|x + y\| < \|x\| + \|y\| = 2$.

(ii)⇒(iii) Let E be the set of extreme points of $V_{\leq 1}$. As noted above, if $x \in V_1$ is a non-trivial convex combination $x = ty + (1-t)z$ with $y, z \in V_{\leq 1}$, we must have $y, z \in V_1$, implying $\|x\| = \|ty\| + \|(1-t)z\|$. Since (ii) holds, we have that y and z are related by a positive scalar, thus $y = z$. Thus x is an extreme point.

(iii)⇒(i) Let $x, y \in V_1$ with $x \neq y$. Then $(x + y)/2$ is a non-trivial convex combination and therefore not in $E = V_1$, so that $\|(x + y)/2\| < 1$. Thus $\|x + y\| < 2$, and we have (i). ■

B.90 DEFINITION *A Banach space V is called strictly convex if it satisfies the equivalent conditions in Proposition B.89.*

Combining the above with Exercises B.82, B.87, B.83 shows that the spaces c_0, ℓ^1, ℓ^∞ and $L^1([0,1], \lambda)$ are not strictly convex. On the other hand:

B.91 EXERCISE Prove that $\ell^p(S, \mathbb{F})$ is strictly convex for every S and $1 < p < \infty$.

Since every Hilbert space is isometrically isomorphic to some $\ell^2(S, \mathbb{F})$, it is strictly convex. This is very easy to show directly and also follows by combining Exercise 5.35 with the following:

B.92 Proposition (Taylor-Foguel) [126] *Let V be a Banach space. Then the following are equivalent:*

(i) V^* is strictly convex.

(ii) For every closed subspace $W \subseteq V$ and $\varphi \in W^*$ there is a unique $\widehat{\varphi} \in V^*$ with $\widehat{\varphi}_{|W} = \varphi$ and $\|\widehat{\varphi}\| = \|\varphi\|$.

Proof. (i)⇒(ii) Existence of $\widehat{\varphi}$ is guaranteed by Hahn-Banach. Assume $W \subseteq V$ is a closed subspace, $\varphi \in W^*$ and $\psi_1, \psi_2 \in V^*$ are such that $\psi_1 \neq \psi_2$, but $\psi_1 \upharpoonright W = \psi_2 \upharpoonright W = \varphi$, $\|\psi_1\| = \|\psi_2\| = \|\varphi\|$. Then $\psi' = (\psi_1 + \psi_2)/2$ satisfies $\psi' \upharpoonright W = \varphi$, thus

$$\|\varphi\| = \|\psi' \upharpoonright W\| \leq \|\psi'\| \leq \frac{\|\psi_1\| + \|\psi_2\|}{2} = \|\varphi\|$$

and therefore $\|\psi'\| = \|\varphi\|$, contradicting the consequence $\|(\psi_1 + \psi_2)/2\| < \|\psi_1\| = \|\psi_2\|$ of strict convexity of V^* .

(ii)⇒(i) Assume V^* is not strictly convex. Then there are $\varphi_1, \varphi_2 \in V^*$ with $\varphi_1 \neq \varphi_2$ and $\|\varphi_1\| = \|\varphi_2\| = 1$ and $\|\varphi_1 + \varphi_2\| = 2$. Then $W = \{x \in V \mid \varphi_1(x) = \varphi_2(x)\} \subseteq V$ is a closed linear subspace and proper (since $\varphi_1 \neq \varphi_2$). Put $\psi = \varphi_{1|W} = \varphi_{2|W} \in W^*$. We will prove $\|\psi\| = 1$. Then φ_1, φ_2 are distinct norm-preserving extensions of $\psi \in W^*$ to V, providing a counterexample for uniqueness of norm-preserving extensions.

Since $\varphi_1 - \varphi_2 \neq 0$, there exists $z \in V$ with $\varphi_1(z) - \varphi_2(z) = 1$. Now every $x \in V$ can be written uniquely as $x = y + cz$, where $y \in W$, $c \in \mathbb{C}$: Put $c = \varphi_1(x) - \varphi_2(x)$ and then $y = x - cz$. Now it is obvious that $y \in W$. Uniqueness of such a representation follows from $z \notin W$.

Since $\|\varphi_1 + \varphi_2\| = 2$, we can find a sequence $\{x_n\} \subseteq V$ with $\|x_n\| = 1$ $\forall n$ such that $\varphi_1(x_n) + \varphi_2(x_n) \to 2$. Since $|\varphi_i(x_n)| \leq 1$ for $i = 1, 2$ and all n, it follows that $\varphi_i(x_n) \to 1$ for $i = 1, 2$. Now write $x_n = y_n + c_n z$, where $\{y_n\} \subseteq W$ and $\{c_n\} \subseteq \mathbb{C}$. Then $c_n = \varphi_1(x_n) - \varphi_2(x_n) \to 0$. Thus $\|x_n - y_n\| = |c_n|\|z\| \to 0$, so that $\|y_n\| \to 1$. And with $c_n \to 0$ we have

$$\lim_{n\to\infty} \varphi_1(y_n) = \lim_{n\to\infty} \varphi_1(y_n + c_n z) = \lim_{n\to\infty} \varphi_1(x_n) = 1.$$

In view of $\{y_n\} \subseteq W$ and $\varphi_{1|W} = \psi$, we have $\psi(y_n) = \varphi_1(y_n) \to 1$. Together with $\|y_n\| \to 1$ this implies $\|\psi\| \geq 1$. Since the converse inequality is obvious, we have $\|\psi\| = 1$, as claimed. ■

B.6.8 Uniform convexity and reflexivity. Duality of L^p-spaces reconsidered

B.93 Definition *A Banach space V is called uniformly convex if for every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $x, y \in V$, $\|x\| = \|y\| = 1$ and $\|x - y\| \geq \varepsilon$ imply $\|(x + y)/2\| \leq 1 - \delta(\varepsilon)$.*

It is obvious that uniform convexity implies strict convexity. The converse is not true:

B.94 Exercise Let $V = \ell^1(\mathbb{N}, \mathbb{R})$, equipped with the norm $\|f\| = \|f\|_1 + \|f\|_2$. Prove that (i) $(V, \|\cdot\|)$ is a Banach space, (ii) strictly convex, but (iii) not uniformly convex.

B.95 Theorem (Milman-Pettis 1938/9) *Every uniformly convex Banach space is reflexive.*

Proof. (Following Ringrose [133]) Assume V is uniformly convex, but not reflexive. Let $S \subseteq V$ and $S^{**} \subseteq V^{**}$ be the unit spheres (sets of elements of norm one). Since $S = S^{**}$ easily implies $V = V^{**}$, we have $S \subsetneqq S^{**}$. If $x^{**} \in S^{**} \backslash S$ then by the obvious norm-closedness of $S \subseteq S^{**}$

[126] Angus Ellis Taylor (1911-1999), American mathematician, proved (i)⇒(ii) in 1939. Shaul Reuven Foguel (1931-2020), Israeli mathematician, proved (ii)⇒(i) in 1958.

there is $\varepsilon > 0$ such that $B(x^{**}, \varepsilon) \cap S = \emptyset$. Since $\|x^{**}\| = 1$, we can find $\varphi \in V^*$ with $\|\varphi\| = 1$ and $|x^{**}(\varphi) - 1| > 1 - \delta(\varepsilon)/2$. Now $U = \{y^{**} \in V^{**} \mid |y^{**}(\varphi) - 1| > 1 - \delta(\varepsilon)\} \subseteq V^{**}$ is a $\tau := \sigma(V^{**}, V^*)$-open neighborhood of x^{**}. By Goldstine's Theorem B.65, $V_{\leq 1} \subseteq (V^{**})_{\leq 1}$ is τ-dense. If $\{x_\alpha\} \subseteq V_{\leq 1}$ is a net τ-converging to $x \in S^{**}$ then $\|x_\alpha\| \to 1$ and $\frac{x_\alpha}{\|x_\alpha\|} \stackrel{\tau}{\to} x$. Thus $S \subseteq S^{**}$ is τ-dense, thus $S \cap U \neq \emptyset$. If now $y_1, y_2 \in S \cap U$ then $|\varphi(y_1) + \varphi(y_2)| > 2 - 2\delta(\varepsilon)$. With $\|\varphi\| = 1$ this implies $\|y_1 + y_2\| > 2 - 2\delta(\varepsilon)$. Thus by uniform convexity we have $\|y_1 - y_2\| < \varepsilon$. Since every net in S that τ-converges to x^{**} ultimately lives in U, picking any $y_1 \in S \cap U$ we have $\|x^{**} - y_1\| \leq \varepsilon$. But this contradicts the choice of ε. ■

The converse of the theorem is not true. In fact there are spaces that are reflexive and strictly convex, but not uniformly convex, but the construction [36] is laborious. Note also that the dual of a uniformly convex space need not be uniformly convex!

B.96 Theorem *For every measure space $(X, \mathcal{A}, \mu)$ and $1 < p < \infty$, the space $L^p(X, \mathcal{A}, \mu; \mathbb{F})$ is uniformly convex and reflexive.*

Proof. We follow [86]. Let $0 < \varepsilon \leq 2^{1-p}$. Then the set

$$Z = \left\{ (x, y) \in \mathbb{R}^2 \mid |x|^p + |y|^p = 2, \ \left| \frac{x - y}{2} \right|^p \geq \varepsilon \right\}$$

is closed and bounded, thus compact, and non-empty since $(2^{1/p}, 0) \in Z$. Since the function $\mathbb{R} \to \mathbb{R}$, $t \mapsto |t|^p$ is strictly convex, we have $\left|\frac{x+y}{2}\right|^p < \frac{|x|^p + |y|^p}{2}$ whenever $x \neq y$. Thus

$$\rho(\varepsilon) = \inf_{(x,y) \in Z} \left(\frac{|x|^p + |y|^p}{2} - \left| \frac{x + y}{2} \right|^p \right) > 0.$$

Now by homogeneity we have

$$\left| \frac{x - y}{2} \right|^p \geq \varepsilon \frac{|x|^p + |y|^p}{2} \quad \Rightarrow \quad \rho(\varepsilon) \frac{|x|^p + |y|^p}{2} \leq \frac{|x|^p + |y|^p}{2} - \left| \frac{x + y}{2} \right|^p. \tag{B.9}$$

Let now $0 < \varepsilon < 2^{1-p}$ and $f, g \in L^p(X, \mathcal{A}, \mu)$ with $\|f\|_p = \|g\|_p = 1$ and $\|(f + g)/2\|_p^p > 1 - \delta$. Writing f, g instead of $f(x), g(x)$, we put

$$M = \left\{ x \in X \mid \left| \frac{f - g}{2} \right|^p \geq \varepsilon \frac{|f|^p + |g|^p}{2} \right\}.$$

Now

$$\begin{aligned}
\left\| \frac{f - g}{2} \right\|_p^p &= \int_{X \setminus M} \left| \frac{f - g}{2} \right|^p + \int_M \left| \frac{f - g}{2} \right|^p \\
&\leq \varepsilon \int_{X \setminus M} \frac{|f|^p + |g|^p}{2} + \int_M \frac{|f|^p + |g|^p}{2} \\
&\leq \varepsilon \int_X \frac{|f|^p + |g|^p}{2} + \frac{1}{\rho(\varepsilon)} \int_M \left(\frac{|x|^p + |y|^p}{2} - \left| \frac{x + y}{2} \right|^p \right) \\
&\leq \varepsilon \int_X \frac{|f|^p + |g|^p}{2} + \frac{1}{\rho(\varepsilon)} \int_X \left(\frac{|x|^p + |y|^p}{2} - \left| \frac{x + y}{2} \right|^p \right) \\
&\leq \varepsilon + \frac{1}{\rho(\varepsilon)} - \frac{1 - \delta}{\rho(\varepsilon)} \ = \ \varepsilon + \frac{\delta}{\rho}.
\end{aligned}$$

(In the second row we used the definition of M and (4.1), in the third we used (B.9), which holds on M, in the forth the fact that the expression in brackets is non-negative on $X\backslash M$, and finally we used the assumptions $\|f\|_p^p \le 1, \|g\|_p^p \le 1$ and $\|(f+g)/2\|_p^p > 1-\delta$.) Now choosing $\delta < \varepsilon\rho(\varepsilon)$ we have $\|(f-g)/2\|_p^p \le 2\varepsilon$, thus uniform convexity (more precisely, an implication equivalent to it).

Reflexivity now follows from Theorem B.95. ■

B.97 Remark The uniform convexity of L^p for $1 < p < \infty$ was first proven by Clarkson in 1936 with a fairly complicated proof. (Reflexivity was known earlier thanks to F. Riesz' proof of $(L^p)^* \cong L^q$.) A simpler proof, still giving optimal bounds, can be found in [73]. □

Now we are in a position to complete the determination of $L^p(X, \mathcal{A}, \mu)^*$ for arbitrary measure space $(X, \mathcal{A}, \mu)$ and $1 < p < \infty$ without invocation of the Radon-Nikodym theorem:

B.98 Corollary *Let $1 < p < \infty$ and $(X, \mathcal{A}, \mu)$ any measure space. Then the canonical map $L^q(X, \mathcal{A}, \mu; \mathbb{F}) \to L^p(X, \mathcal{A}, \mu; \mathbb{F})^*$ is an isometric bijection.*

Proof. Let $(X, \mathcal{A}, \mu)$ be any measure space, $1 < p < \infty$ and q the conjugate exponent. We abbreviate $L^p(X, \mathcal{A}, \mu)$ to L^p. As discussed (without complete, but hopefully sufficient detail) in Section 4.7, the map $\varphi : L^q \to (L^p)^*,\ g \mapsto \varphi_g$ is an isometry, so that only surjectivity remains to be proven. Assume $\varphi(L^q) \subsetneqq (L^p)^*$. The subspace being closed (since L^q is complete and φ is an isometry), by Hahn-Banach there is a $0 \ne \psi \in (L^p)^{**}$ such that $\psi \restriction \varphi(L^q) = 0$. By reflexivity of L^p (Theorem B.96), there is an $f \in L^p$ such that $\psi = \iota_{L^p}(f)$. This implies $\varphi_g(f) = \psi(\varphi_g) = 0$ for all $g \in L^q$. With $\varphi_g(f) = \int fg\, d\mu = \varphi'_f(g)$, where $\varphi' : L^p \to (L^q)^*$ is the canonical map, this implies $\varphi'_f = 0$. Since φ' is an isometry, we have $f = 0$ and therefore $\psi = 0$, which is a contradiction. Thus $\varphi : L^q \to (L^p)^*$ is surjective. ■

B.7 The Eidelheit-Chernoff theorem

It is not unreasonable to ask whether the assignment $V \mapsto B(V)$ (or $V \mapsto K(V)$) is injective. This certainly is the case for finite-dimensional spaces, for the simple reason that $\dim B(V) = \dim K(V) = (\dim V)^2$ and the fact that all normed spaces of the same finite dimension are isomorphic.

Dropping the finiteness, we need the notion of algebra isomorphism, cf. Definition 15.1.

B.99 Theorem (Eidelheit 1940, Chernoff 1969/73) [127] *Let V, W be normed spaces and $\mathcal{A} \subseteq B(V)$, $\mathcal{B} \subseteq B(W)$ subalgebras containing the algebras $F(V)$ and $F(W)$, respectively, of finite rank operators. Then every algebra isomorphism $\alpha : \mathcal{A} \to \mathcal{B}$ is of the form $\alpha(A) = TAT^{-1}$ for some isomorphism $T : V \to W$ of Banach spaces.*

In particular, an algebraic isomorphism $K(V) \cong K(W)$ or $B(V) \cong B(W)$ implies $V \simeq W$.

Proof. We closely follow [7] with some extra details. Pick a rank-one idempotent $P \in \mathcal{A}$. Then $\alpha(P)^2 = \alpha(P^2) = \alpha(P)$, thus $\alpha(P) \in \mathcal{B}$ is an idempotent and non-zero (since α is injective). We claim it has rank one. [If not, pick a non-zero proper subspace of $W' = \alpha(P)W$ and an idempotent Q with image W'. It satisfies $Q\alpha(P) = \alpha(P)Q = Q$. Since $Q \in F(W) \subseteq \mathcal{B}$ and α is an isomorphism, there is an idempotent $P' \in \mathcal{A}$ with $\alpha(P') = Q$. It satisfies $P'P = PP' = P'$, thus is a non-zero proper subprojection of P, but this is impossible since P has rank one.]

[127] Meier Eidelheit (1910-1943), Polish mathematician, killed in the holocaust. Paul Robert Chernoff (1942-2017), American analyst.

Pick non-zero vectors $x_0 \in PV, y_0 \in \alpha(P)W$ and an $S \in B(V,W)$ with $Sx_0 = y_0$. [E.g., pick any $\varphi \in V^*$ with $\varphi(x_0) \neq 0$. Then $S : x \mapsto \varphi(x_0)^{-1}\varphi(x)y_0$ does the job.] Define linear maps

$$\psi_1 : \mathcal{A}P \to V,\ AP \mapsto APx_0, \quad \psi_2 : \mathcal{B}\alpha(P) \to W,\ B\alpha(P) \mapsto B\alpha(P)y_0.$$

Now ψ_1 injective quite trivially, and it is surjective since for every $x \in V$ there exists $A \in F(V)$ with $Ax_0 = x$ and since $F(V) \subseteq \mathcal{A}$. Similarly, ψ_2 is a linear bijection. Since the bijection $\alpha : \mathcal{A} \to \mathcal{B}$ restricts to a bijection $\alpha_0 : \mathcal{A}P \to \mathcal{B}\alpha(P)$, the composite $T = \psi_2 \circ \alpha_0 \circ \psi_1^{-1} : V \to W$ is a linear bijection. By definition of these maps, we have

$$TAPx_0 = \alpha(A)\alpha(P)y_0 = \alpha(A)\alpha(P)Sx_0 \quad \forall A \in \mathcal{A},$$

which is the same as saying $TAP = \alpha(A)\alpha(P)SP$. If now $A, A' \in \mathcal{A}$ then this implies

$$TAA'P = \alpha(A)\alpha(A')\alpha(P)SP = \alpha(A)TA'P.$$

Thus $TAA'x_0 = \alpha(A)TA'x_0$, and since every $x \in V$ is of the form $A'x_0$ for some $A \in F(V) \subseteq \mathcal{A}$, we have the identity $TA = \alpha(A)T$ (of linear maps $V \to W$). With invertibility of T this is the same as $\alpha(A) = TAT^{-1}$.

It remains to prove that T is bounded. Pick $y' \in W\backslash\{0\}$. For each $\varphi \in W^*$, we have the finite rank operator $y' \otimes \varphi \in F(W) : y \mapsto y'\varphi(y)$. Since $F(W) \subseteq \mathcal{B}$, there exists $A \in \mathcal{A}$ such that $y' \otimes \varphi = \alpha(A) = TAT^{-1}$. Thus $A = T^{-1}(y' \otimes \varphi)T = (T^{-1}y') \otimes (\varphi \circ T)$. Since A is bounded and $T^{-1}y' \neq 0$, this implies that $\varphi \circ T = T^t\varphi$ is bounded for each $\varphi \in W^*$. As proven in Exercise 9.32, this implies that T is bounded. ■

B.100 Corollary *Let V be a normed space. Then:*

(i) Every automorphism of $B(V)$ is inner, i.e. of the form $T \cdot T^{-1}$ for some $T \in \mathrm{Inv}B(V)$.

(ii) Every automorphism of $K(V)$ extends to an automorphism of $B(V)$.

B.101 Remark The conclusion of the theorem implies $\|\alpha\| \leq \|T\|\|T^{-1}\| < \infty$, so that α is bounded. Thus a purely algebraic hypothesis implies a continuity statement. This is an early instance of the subject of 'automatic continuity', cf. [32]. For another such statement see Theorem B.173. □

B.8 A bit more on invariant subspaces. Lomonosov's theorem

A classical question of functional analysis is the *invariant subspace problem*, asking whether every Banach space operator $A \in B(V)$ admits a proper A-invariant subspace. Several affirmative answers are known:

- If $A = 0$ then every subspace is A-invariant.
- If V is finite-dimensional with $\dim V \geq 2$, the Jordan normal form implies that every $A \in B(V)$ has an invariant subspace.
- If V is non-separable, pick any $x \in V\backslash \ker A$ and put $W = \overline{\mathrm{span}_{\mathbb{F}}\{A^n x \mid n \in \mathbb{N}_0\}}$. Then W is a non-zero A-invariant subspace and separable, thus proper.
- Every normal operator on a complex Hilbert space has proper invariant subspaces by Exercise 18.7.

Thus we are left with non-normal operators on infinite-dimensional separable spaces. In 1954, Aronszajn and Smith[128] proved that every compact operator on a complex Banach space has a non-trivial invariant subspace. Lomonosov[129] used Schauder's fixed point theorem, see Section B.15, to prove the following stronger statement:

B.102 Theorem (Lomonosov 1973) *On an infinite-dimensional complex Banach space V, every non-zero compact operator K has a non-trivial hyperinvariant subspace, i.e. a closed subspace $0 \neq W \subsetneqq V$ such that $AW \subseteq W$ for every $A \in B(V)$ that commutes with K.*

Proof. By Corollary 14.4 of Fredholm's alternative, all non-zero elements of $\sigma(K)$ are eigenvalues. The eigenspace V_λ corresponding to a non-zero eigenvalue λ is finite-dimensional by Lemma 14.1, thus proper, and is invariant not only for K but also every operator commuting with it: If $x \in V_\lambda$ and $AK = KA$ then $KAx = AKx = \lambda Ax$, thus $Ax \in V_\lambda$. We are therefore left with the case of compact K with $\sigma(K) = \{0\}$, i.e. the quasi-nilpotent ones. Rather than follow Lomonosov, we give the fantastic proof of Hilden (as presented by Michaels [104]) that avoids all fixed point theorems. It only uses that for quasi-nilpotent K the spectral radius formula gives $\lim \|K^n\|^{1/n} = r(K) = 0$, so that $\|(cK)^n\| \to 0$ for all $c \in \mathbb{C}$, cf. Exercise 13.45.

Since multiplying K by a non-zero number does not change $\{K\}'$, we may assume $\|K\| = 1$. Pick $x_0 \in V$ such that $\|Kx_0\| > 1$. Clearly $\|x_0\| > 1$. Let $\mathcal{B} = x_0 + V_{\leq 1} = \overline{B}(x_0, 1)$. Then $0 \notin \mathcal{B}$ and $0 \notin \overline{K\mathcal{B}} = Kx_0 + \overline{KV_{\leq 1}}$ (since $\|Kx_0\| > 1$ and $\overline{KV_{\leq 1}} \subseteq V_{\leq 1}$.) For each $y \in V$ put

$$W_y = \{Ay \mid A \in B(V),\ AK = KA\}.$$

Now W_y is a linear subspace of V containing y, so that it is non-zero whenever $y \neq 0$. If B commutes with K then $BW_y \subseteq W_y$. (Why?) Thus $\overline{W_y} \subseteq V$ is a closed hyperinvariant subspace for K, so that we are done if we can find a $y \neq 0$ with $\overline{W_y} \neq V$.

To prove this by contradiction, assume $\overline{W_y} = V\ \forall y \neq 0$. Then for each $y \neq 0$ there exists $A \in B(V)$ with $AK = KA$ and $\|Ay - x_0\| < 1$. With

$$U(A) = \{y \in V \mid \|Ay - x_0\| < 1\} = A^{-1}(B(x_0, 1))$$

this is equivalent to the statement $\bigcup_{A \in B(V), AK=KA} U(A) = V \backslash \{0\}$. Thus the $U(A)$ form an open (by continuity of A) cover of the compact (by compactness of K) subset $\overline{K\mathcal{B}}$ of $V\backslash\{0\}$. Thus there are $A_1, \ldots, A_n \in B(V)$ such that $A_iK = KA_i$ for all i and $K\mathcal{B} \subseteq \bigcup_{i=1}^n U(A_i)$.

Since $Kx_0 \in K\mathcal{B}$, we have $Kx_0 \in U(A_i)$ for some i_1, meaning $\|A_{i_1}Kx_0 - x_0\| < 1$, to wit $A_{i_1}Kx_0 \in \mathcal{B}$. Then $KA_{i_1}Kx_0 \in K\mathcal{B}$, thus $KA_{i_1}Kx_0 \in U(A_{i_2})$ for some i_2, meaning $A_{i_2}KA_{i_1}Kx_0 \in \mathcal{B}$. It is clear that we can iterate this construction, obtaining a sequence $\{i_m\}$ such that $A_{i_m}KA_{i_{m-1}}K \cdots A_{i_1}Kx_0 \in \mathcal{B}$ for all $m \in \mathbb{N}$. Since all A_i commute with K, this is equivalent to $A_{i_m}A_{i_{m-1}} \cdots A_{i_1}K^m x_0 \in \mathcal{B}$ for all m. With $c = \max(\|A_1\|, \ldots, \|A_n\|)$ this in turn is equivalent to

$$(c^{-1}A_{i_m})(c^{-1}A_{i_{m-1}}) \cdots (c^{-1}A_{i_1})(cK)^m x_0 \in \mathcal{B} \quad \forall m \in \mathbb{N}.$$

The l.h.s. of this tends to zero since $\|c^{-1}A_i\| \leq 1$ for all i and $\|(cK)^m\| \to 0$ by assumption, thus $0 \in \overline{\mathcal{B}}$. But this contradicts the fact that $\mathcal{B}$ is closed by definition and $0 \notin \mathcal{B}$. The only way out of this contradiction is thay there exists $y \neq 0$ with $\overline{W_y} \neq V$, which then is a non-trivial closed hyperinvariant subspace. ■

[128] Nachman Aronszajn (1907-1980), Polish-American mathematician who worked on functional analysis, mathematical logic. Kennan Tayler Smith (1926-2000), American mathematician.

[129] Victor Lomonosov (1946-2018), Russian-American mathematician who mostly worked on functional analysis.

B.103 Remark 1. While in finite dimensions every operator has non-trivial invariant subspaces, Lomonosov's theorem fails since $\mathbf{1}$ is compact, while there clearly is no non-trivial subspace that is invariant under all operators. This is also why we required $K \neq 0$ above.

2. With the Aronszajn-Smith result, the invariant subspace problem is reduced (over $\mathbb{C}$) to non-compact non-normal operators on separable spaces. In 1975/1987 Enflo (already encountered in connection with the approximation property) constructed an operator on a Banach space having no invariant subspace [48]. By now, more examples are known, even on ℓ^1 [127], but all live on non-reflexive spaces. It still is an open question whether all operators on reflexive spaces, or at least on separable Hilbert spaces must have an invariant subspace! (In a paper [49] from May 2023, Enflo claims to prove just this, but this has not yet been verified.) □

Why should anyone be interested in the existence of invariant subspaces? The arguments used to prove the existence of invariant subspaces in finite dimensions and for normal operators on Hilbert spaces are closely related to results (usually called 'spectral theorem' only for normal operators) giving representations of those operators by standard forms. Thus proving the existence of invariant subspaces will invariably lead to structural results on the family of operators in question.

As we saw in Section 14.4, for a compact operator $A \in B(V)$ is is not too difficult to construct a family $\{P_\lambda\}_{\lambda\in\sigma(A)}$ of mutually commuting idempotents satisfying $\sum_\lambda P_\lambda = \mathbf{1}$ (converging unconditionally in the strong operator topology) and such that each $V_\lambda = P_\lambda V$ is an A-invariant subspace with $\sigma(A \upharpoonright V_\lambda) = \{\lambda\}$. For $\lambda \neq 0$, each V_λ is finite-dimensional and $A - \lambda\mathbf{1} \upharpoonright V_\lambda$ nilpotent. But saying something non-trivial about $A \upharpoonright V_0$ requires the less trivial existence of invariant subspaces for quasi-nilpotent compact operators given by the above theorem. Using the latter, good results on normal forms for compact operators have indeed been proven by Ringrose[130] [134, 135] and Brodsky[131] [24].

B.9 More on Fredholm operators

We now study Fredholm operators, cf. Definition 14.9, in more depth.

We begin with some some considerations in pure linear algebra. If X, Y, Z are vector spaces over the same field k and $A : X \to Y, B : Y \to Z$ are linear maps with finite-dimensional kernels, it is easy to prove

$$\dim\ker A \leq \dim\ker(BA) \leq \dim\ker A + \dim\ker B.$$

An analogous inequality holds for the cokernels, but there is no exact additivity of the dimensions of the kernels or cokernels. Yet additivity does hold for the Fredholm indices! Generalizing the definition of Fredholm operators to arbitrary vector spaces and putting $\mathrm{ind}(A) = \dim\ker A - \dim\mathrm{coker} A$ we have:

B.104 Proposition *Let X, Y, Z be vector spaces and $A : X \to Y, B : Y \to Z$ Fredholm. Then BA is Fredholm and* $\mathrm{ind}(BA) = \mathrm{ind}(A) + \mathrm{ind}(B)$.

Proof. We follow the argument in [147]. If $A : X \to Y$ with X and Y finite-dimensional, we have $\dim X = \dim\ker A + \dim AX = \dim\ker A + \dim Y - \dim\mathrm{coker} A$, thus $\mathrm{ind}(A) = \dim X - \dim Y$. If also Z is finite-dimensional and $B : Y \to Z$ then

$$\mathrm{ind}(BA) = \dim X - \dim Z = (\dim X - \dim Y) + (\dim Y - \dim Z) = \mathrm{ind}(A) + \mathrm{ind}(B),$$

[130] John Robert Ringrose (1932-), British functional analyst, working also on operator algebras.
[131] Mikhail Samoilovich Brodskii (1913-1989), Soviet mathematician. Student of M. G. Krein.

proving the claim in the case of finite-dimensional spaces.

In the general case, we will find complementary subspaces $X_0, X_1 \subseteq X$ (thus $X = X_0 + X_1$, $X_0 \cap X_1 = \{0\}$) and similarly $Y_0, Y_1 \subseteq Y$ and $Z_0, Z_1 \subseteq Z$ such that X_0, Y_0, Z_0 are finite-dimensional, $AX_0 \subseteq Y_0, BY_0 \subseteq Z_0$ and $AX_1 = Y_1, BY_1 = Z_1$ with $X_1 \cap \ker A = \emptyset = Y_1 \cap \ker B$. Thus $BA : X_1 \to Z_1$ is a bijection, so that with the first half of the proof we have

$$\operatorname{ind}(BA) = \operatorname{ind}(BA : X_0 \to Z_0) = \operatorname{ind}(A : X_0 \to Y_0) + \operatorname{ind}(B : Y_0 \to Z_0) = \operatorname{ind}(A) + \operatorname{ind}(B).$$

It remains to construct the spaces in question. We need one fact from linear algebra: If $X_1, X_2 \subseteq X$ are linear subspaces of X with $X_1 \cap X_2 = \{0\}$, there exists a complementary subspace of X_1 containing X_2. (This follows by applying Zorn's lemma to the family of subspaces of X that contain X_2 and have trivial intersection with X_1.)

We put $X_0 = \ker(BA) = A^{-1}(\ker B) \supseteq \ker A$, which has dimension at most $\dim \ker A + \dim \ker B$. Letting X_1 be any complement of X_0 in X, we have $X_1 \cap \ker A = \{0\}$. Putting $Y_1 = AX_1$, the map $A : X_1 \to Y_1$ is a bijection, and since $A : X \to Y$ has finite cokernel, Y_1 has finite codimension in Y. We have $Y_1 \cap \ker B = \{0\}$ (since $BA \restriction X_1$ is injective, thus also $B \restriction AX_1 = Y_1$). Thus there is a complement Y_0, clearly finite-dimensional, of Y_1 in Y containing $\ker B$. Putting $Z_1 = BY_1$, by the same argument as earlier, $Z_1 \subseteq Z$ has finite codimension. Since $\ker B \subseteq Y_0$ by construction, we have $BY_0 \cap Z_1 = \{0\}$. Thus there exists a complement Z_0, clearly finite-dimensional, of Z_1 containing BY_0. Now all our claims are satisfied. ∎

Note that by the first part of the proof, every $A \in \operatorname{End} V$, where V is finite-dimensional, is Fredholm with index zero.

B.105 Proposition *Let V, W be Banach spaces and $A \in B(V, W)$. Then A^t is Fredholm if and only if A is Fredholm. Under these equivalent hypotheses we have* $\dim \ker(A^t) = \dim \operatorname{coker}(A)$ *and* $\dim \operatorname{coker}(A^t) = \dim \ker(A)$, *thus* $\operatorname{ind}(A^t) = -\operatorname{ind}(A)$.

Proof. Assume that A is Fredholm, in particular it satisfies $\dim(W/AV) < \infty$. By Exercise 7.11 this implies closedness of $AV \subseteq W$. By Exercise 9.35(i), $\ker A^t = (AV)^\perp \subseteq W^*$. And by Exercise 6.7 $(AV)^\perp$ is isometrically isomorphic to $(W/AV)^*$ as Banach spaces. Since W/AV is finite-dimensional, we have $\dim \ker(A^t) = \dim(W/AV)^* = \dim(W/AV) = \dim \operatorname{coker}(A)$.

Since A has closed image, the same is true for A^t by Exercise 9.41. By Exercise 9.36(i) we have $\ker A = (A^t W^*)^\top$, thus $(\ker A)^* \cong ((A^t W^*)^\top)^* \cong V^*/A^t W^*$, where the (isometric) isomorphism from Exercise 9.16. With finite-dimensionality of $\ker A$ we thus have $\dim \ker(A) = \dim(\ker A)^* = \dim(V^*/A^t W^*) = \dim \operatorname{coker}(A^t)$. Now it is clear that A^t is Fredholm with $\operatorname{ind}(A^t) = -\operatorname{ind}(A)$.

Now assume that A^t is Fredholm. As before, this implies that A^t has closed image. As a consequence of a result that we did not prove, A has closed image. Now we can argue as above, resulting in A being Fredholm. ∎

B.106 Corollary *Let H, H' be Hilbert spaces.*

(i) If $A \in B(H, H')$ is Fredholm then $A^* \in B(H', H)$ is Fredholm with $\operatorname{ind}(A^*) = -\operatorname{ind}(A)$.

(ii) If $A \in B(H)$ is Fredholm and normal then $\operatorname{ind}(A) = 0$.

Proof. (i) is immediate, combining Proposition B.105 with the definition of A^* in terms of A^t, cf. Proposition 11.1.

(ii) If $A \in B(H)$ is normal then Proposition 11.27(i) gives $\ker A = \ker A^*$, thus $\dim \ker A = \dim \ker(A^*)$. Now $\dim \operatorname{coker} A = \dim \ker(A^*) = \dim \ker A$, whence $\operatorname{ind}(A) = 0$. ∎

B.107 THEOREM (ATKINSON) [132] *Let V be an infinite-dimensional*[133] *Banach space and $A \in B(V)$. Then the following are equivalent:*

(i) A is Fredholm.

(ii) There exists a Fredholm B with $\mathrm{ind}(B) = -\mathrm{ind}(A)$ and $ABA = A$[134] *that $\mathbf{1} - AB$ and $\mathbf{1} - BA$ are finite rank idempotents.*

(iii) There exists $B \in B(V)$ such that $AB - \mathbf{1}$ and $BA - \mathbf{1}$ are compact.

(iv) The image of A in the Calkin[135] *algebra $\mathcal{C}(V) = B(V)/K(V)$ is invertible.*

Proof. (i)⇒(ii) Being Fredholm, A has finite-dimensional kernel and cokernel. Thus $\ker A$ is complemented, and we can find a closed complement $V_1 \subseteq V$. And since $AV \subseteq V$ is closed by Exercise 7.11 and has finite codimension, it has a closed complement V_2. Now $A \restriction V_1$ is injective, thus $A' : V_1 \to AV$ is a bounded linear bijection. By the BIT, its inverse $(A')^{-1} : AV \to V_1$ is bounded. Define $B : V \to V$ by $B = (A')^{-1}$ on AV and as zero on the complement V_2 of AV. By construction B is bounded. We have $\ker B = V_2$, thus $\dim \ker B = \dim \mathrm{coker} A < \infty$, and $BV = V_1$, thus $\dim \mathrm{coker} B = \dim \ker A < \infty$. Thus B is Fredholm with $\mathrm{ind}(B) = -\mathrm{ind}(A)$.

Now BA is the identity on V_1 and zero on $\ker A$, thus $BA = \mathbf{1} - P_1$, where P_1 is the unique idempotent with $P_1 V = \ker A$ and $(1 - P_1)V = V_1$. Since $\ker V$ is finite-dimensional, P_1 has finite rank. And $ABA = A(\mathbf{1} - P_1) = A - AP_1 = A$. Similarly, AB is the identity on AV and zero on the complement V_2 of AV. Thus $AB = \mathbf{1} - P_2$, where P_2 is the idempotent with image V_2. Since V_2 is finite-dimensional, P_2 has finite rank.

(ii)⇒(iii) is trivial. (iii)⇒(i) There exists $B \in B(V)$ such that $AB = \mathbf{1}+C, BA = \mathbf{1}+D$ with C, D compact. Since this implies $\ker A \subseteq \ker(\mathbf{1}+D)$, and Lemma 14.1 gives $\dim \ker(\mathbf{1}+D) < \infty$, we have $\dim \ker A < \infty$. On the other hand, $(1+C)V \subseteq AV$ and since $(\mathbf{1}+C)V \subseteq V$ has finite codimension by Proposition 14.7, so has $AV \subseteq V$, thus $\mathrm{coker} A = V/AV$ is finite-dimensional. Thus A is Fredholm.

(iii)⇔(iv) Invertibility of $q(A) \in \mathcal{C}(A)$ means that there is a $B \in B(V)$ such that $q(A)q(B) = \mathbf{1}_{\mathcal{C}(V)}$ and $q(B)q(A) = \mathbf{1}_{\mathcal{C}(V)}$. Since $K(V) \subseteq B(V)$ is a two-sided ideal, q is a homomorphism. Thus $q(A)q(B) = q(AB)$, which equals $\mathbf{1}_{\mathcal{C}(V)}$ if and only if $AB \in \mathbf{1} + K(H)$. Similarly for $q(B)q(A)$, so that the claim follows. ■

B.108 THEOREM *Let V be a Banach space. Then*

(i) The set $\mathrm{Fr}(V)$ of Fredholm operators on V is an open subset of $B(V)$.

(ii) The map $\mathrm{Fr}(V) \to \mathbb{Z}$, $A \mapsto \mathrm{ind}(A)$ is continuous.

(iii) If $A \in B(V)$ is Fredholm and $K \in K(V)$ is compact then $A + K$ is Fredholm and $\mathrm{ind}(A + K) = \mathrm{ind}(A)$.

Proof. (i) By Proposition 6.1 the quotient map $q : B(V) \to B(V)/K(V) = \mathcal{C}(V)$ is continuous and $\mathcal{C}(V)$ is a Banach algebra. Thus its set of invertibles is open by Lemma 13.19. Thus $q^{-1}(\mathrm{Inv}\, \mathcal{C}(V)) \subseteq B(V)$ is open, and since this coincides with $\mathrm{Fr}(V)$ by Theorem B.107, we are done.

[132] Frederick Valentine Atkinson (1916-2002), British mathematician.

[133] For finite-dimensional V, (i)-(iii) are trivially true. The same holds for (iv) if we are willing to consider 0 as unit element of $\{0\}$.

[134] An element a of an algebra $\mathcal{A}$ is called regular if there exists a $b \in \mathcal{A}$ such $aba = a$.

[135] John Williams Calkin (1909-1964). American mathematician. Worked at Los Alamos for 12 years on you guess what.

(ii) Let A be Fredholm and pick B as in Theorem B.107(ii). If now A' is Fredholm with $\|A - A'\| < \|B\|^{-1}$ then $\|AB - A'B\| < 1$, so that $D = \mathbf{1} + A'B - AB$ is invertible by Lemma 13.19(ii), thus $\mathrm{ind}(D) = 0$. Now $DA = A + A'BA - ABA = A + A'BA - A = A'BA$ using $ABA = A$. Since A, A', B, D are Fredholm, this implies $\mathrm{ind}(D) + \mathrm{ind}(A) = \mathrm{ind}(A') + \mathrm{ind}(B) + \mathrm{ind}(A)$. With $\mathrm{ind}(D) = 0$ and $\mathrm{ind}(B) = -\mathrm{ind}(A)$, we conclude $\mathrm{ind}(A') = \mathrm{ind}(A)$. Thus ind is locally constant (constant on an open neighborhood of each point) and therefore continuous.

(iii) If $A \in B(V)$ is Fredholm and K is compact then $q(A) \in \mathcal{C}(V)$ is invertible by Theorem B.107, thus also $q(A + K) = q(A)$ is invertible, so that $A + K$ is Fredholm. Now $t : [0,1] \to \mathrm{Fr}(V), t \mapsto A + tK$ is continuous, thus with (ii) also $t \mapsto \mathrm{ind}(A + tK)$ is continuous. Since a continuous map from a connected space to a discrete space is constant, we have $\mathrm{ind}(A) = \mathrm{ind}(A + K)$. ■

B.109 Corollary *If $A \in B(V)$ is compact and $\lambda \in \mathbb{F}\backslash\{0\}$ then $A - \lambda\mathbf{1}$ is Fredholm with index zero, thus* $\dim\ker(A - \lambda\mathbf{1}) = \dim\mathrm{coker}(A - \lambda\mathbf{1})$.

Proof. Apply Theorem B.108 to $\mathbf{1} - \frac{A}{\lambda}$, noting that $\mathbf{1}$ is Fredholm with index zero. ■

B.110 Proposition *If V is a Banach space and $A \in B(V)$ then the following are equivalent:*

(i) A is Fredholm with index zero.

(ii) There exists a compact $K \in K(V)$ such that $A + K$ is invertible.

Proof. (ii)⇒ If there exists $K \in K(V)$ such that $A+K$ is invertible then with Theorem B.108(iii) we have $\mathrm{ind}(K) = \mathrm{ind}(A + K) = 0$ since $A + K$ is invertible, thus has index zero.

(i)⇒(ii) If A is Fredholm with $\mathrm{ind}(A) = 0$ then $\ker A$ and $\mathrm{coker}A$ have the same finite dimension. Now $\ker A$ has a closed complement V_1, and AV (which we know to be automatically closed) has a finite-dimensional complement V_2. Now $\dim V_2 = \dim\mathrm{coker}A = \dim\ker A < \infty$. Thus we can find an invertible (and bounded, by finite-dimensionality) $B : \ker A \to V_2$. Since we have the isomorphisms $(\ker A) \oplus V_1 \simeq V \simeq V_2 \oplus AV$, we can define a bounded $K : V \to V$ as B on $\ker A$ and zero on V_1. Since K has finite rank, it is compact. And $A + K$ restricts to the invertible maps $B : \ker A \to V_2$ and $A' : V_1 \to AV$ and therefore is invertible. ■

B.111 Exercise Let V be a Banach space and $A \in B(V)$ Fredholm. Prove: There exists a compact K such that $A + K$ is surjective (resp. injective) if and only if $\mathrm{ind}(A) \geq 0$ (resp. ≤ 0).

If X is a topological space, recall that $\pi_0(X)$ is the set of path components of X, i.e. $X/{\sim}$, where $x \sim y$ if and only if there exists $p \in C([0,1], X)$ with $p(0) = x, p(1) = y$. The $\sim$-equivalence of x is denoted $[x]$. A continuous map $f : X \to Y$ induces a map $f_* : \pi_0(X) \to \pi_0(Y)$.

B.112 Proposition *Let V be a Banach space. Then $\pi_0(\mathrm{Fr}(V))$ admits a group structure such that $[A][B] = [AB]$ and a homomorphism* $\mathrm{ind} : \pi_0(\mathrm{Fr}(V)) \to \mathbb{Z}$ *such that* $\mathrm{ind}([A]) = \mathrm{ind}(A)$.

If $\iota : \mathrm{Inv}B(V) \to \mathrm{Fr}(V)$ is the inclusion map, we have $\ker\mathrm{ind} = \iota_*(\pi_0(\mathrm{Inv}\, V))$. *(Thus $\pi_0(\mathrm{Inv}B(V)) \to \pi_0(\mathrm{Fr}(V)) \to \mathbb{Z}$ is a short exact sequence.)*

Proof. The set $(\mathrm{Fr}(V), \cdot, \mathbf{1})$ is a monoid (defined like a group, but without inverses). If $A, A', B, B' \in \mathrm{Fr}(V)$ such that $A \sim A'$ and $B \sim B'$ then $AB \sim A'B'$, so that $[A][B] = [AB]$ defines a monoid structure on $\pi_0(\mathrm{Fr}(V))$. It actually is a group with unit $[\mathbf{1}]$: If $A \in \mathrm{Fr}(V)$ then by Theorem B.107(ii) there is a Fredholm B such that $P_1 = \mathbf{1} - BA$ is a finite rank idempotent. Define $C_t = BA + tP_1$. Since $\{C_t\}_{t\in[0,1]}$ is a continuous path in $\mathrm{Fr}(V)$ from BA to the identity, we have $[B][A] = [BA] = [\mathbf{1}]$ in $\pi_0(\mathrm{Fr}(V))$. Similarly $[A][B] = [\mathbf{1}]$.

The map $\text{ind} : \text{Fr}(V) \to \mathbb{Z}$ is a monoid homomorphism by Proposition B.104, and the local constancy of ind (Theorem B.108(ii)) implies that we have a well defined map $\pi_0(\text{Fr}(V)) \to \mathbb{Z}$, also denoted by ind, such that $\text{ind}([A]) = \text{ind}(A)$. This clearly is a group homomorphism.

The set $\text{Inv}(V)$ of invertible operators is a group, and the same holds for $\pi_0(\text{Inv}B(V))$ by the same argument as before. Since every invertible operator is Fredholm with index zero, the inclusion map $\iota : \text{Inv}B(V) \to \text{Fr}(V)$ induces a monoid homomorphism $\iota_* : \pi_0(\text{Inv}B(V)) \to \pi_0(\text{Fr}(V))$ such that the composite map $\pi_0(\text{Inv}B(V)) \stackrel{\iota_*}{\to} \pi_0(\text{Fr}(V)) \stackrel{\text{ind}}{\longrightarrow} \mathbb{Z}$ is zero. Thus $\iota_*(\pi_0(\text{Inv}B(V))) \subseteq \ker \text{ind}$.

It remains to prove that $\text{ind}([A]) = 0$ implies $[A] \in \iota_*(\pi_0(\text{Inv}B(V)))$. On the level of operators instead of path-components this amounts to proving that every Fredholm operator with index zero can be connected to an invertible operator by a continuous path in $\text{Fr}(V)$. Let thus $A \in B(V)$ with $\dim \ker A = \dim \text{coker} A < \infty$. Since the subspaces $\ker V, \text{coker} V \subseteq V$ are finite-dimensional, they are complemented, giving rise to direct sum decompositions $V = V_1 \oplus \ker V = V_2 \oplus \text{coker} V$ with respect to which A is described by the matrix $\begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix}$, where $A' : V_1 \to V_2$ is invertible. Since $\ker V$ and $\text{coker} V$ have the same dimensions, we can find an invertible $B : \ker A \to \text{coker} A$. For $t \in [0,1]$ let $A_t = \begin{pmatrix} A' & 0 \\ 0 & tB \end{pmatrix} \in B(V)$. Since $A_t \in \text{Fr}(V)$ for all t and $A_0 = A$, while A_1 is invertible, we have achieved our goal. ■

We now specialize to complex Hilbert spaces.

B.113 Exercise Let H be an infinite-dimensional Hilbert space and $n \in \mathbb{Z}$. Construct a Fredholm operator $A \in B(H)$ with $\text{ind}(A) = n$.

B.114 Theorem *If H is a complex Hilbert space, the set of Fredholm operators of index n on H is path-connected for every $n \in \mathbb{Z}$. Thus $\text{Fr}(H)$ is the disjoint union of open path-connected components, one for each $n \in \mathbb{Z}$. Equivalently, the map $\pi_0(\text{Fr}(H)) \to \mathbb{Z}$ is a bijection.*

Proof. Surjectivity of the index map $\pi_0(\text{Fr}(H)) \to \mathbb{Z}$ is immediate by Exercise B.113. Injectivity follows at once from Proposition B.112 if we prove $\pi_0(\text{Inv}B(H)) = 0$. But this is nothing other than path-connectedness of $\text{Inv}B(H)$, which we proved in Proposition 18.17 (for $\mathbb{F} = \mathbb{C}$). ■

We close with an interesting related result:

B.115 Theorem (Feldman & Kadison 1954) *Let H be a separable Hilbert space. Then*

$$\overline{\text{Inv}\, B(H)}^{\|\cdot\|} = \{A \in B(H) \mid AH \subseteq H \text{ is non-closed or } \dim \ker A = \dim \ker A^* \in \mathbb{N}_0 \cup \{\infty\}\}.$$

For an accessible proof (using the essential spectrum, see the next section) see [20].

B.10 Discrete and essential spectrum

B.10.1 Banach space operators

The spectra of an operator that we have defined so far are all very unstable under perturbation of the operator, e.g. by compact operators (or by operators of small norm). This motivates the search for interesting and relevant subsets of $\sigma(A)$, called essential sprectra, that have such invariance properties. We begin with two obvious candidates:

B.116 Definition *Let V be a complex Banach space and $A \in B(V)$. Define*

- $\sigma_{\text{ess},1}(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathbf{1} \text{ is not Fredholm}\}$.
- $\sigma_{\text{ess},2}(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathbf{1} \text{ is not Fredholm with index zero}\}$.

($\sigma_{\text{ess},1}(A)$ is called the Fredholm (essential) spectrum of A and $\sigma_{\text{ess},2}(A)$ the Weyl (essential) spectrum. These are many other definitions discussed in the literature, cf. e.g. [45].)

Some immediate observations:

- Since every invertible operator is Fredholm with index zero, we have
$$\sigma_{\text{ess},1}(A) \subseteq \sigma_{\text{ess},2}(A) \subseteq \sigma(A).$$
- In view of Theorem B.108(iii) it is evident that
$$\sigma_{\text{ess},1}(A+K) = \sigma_{\text{ess},1}(A) \quad \text{and} \quad \sigma_{\text{ess},2}(A+K) = \sigma_{\text{ess},2}(A) \qquad \forall A \in B(V), K \in K(V).$$
- This implies that $\sigma_{\text{ess},1}(K) = \sigma_{\text{ess},2}(K) = \{0\}$ for all compact K, in particular for all $A \in B(V)$ if V is finite-dimensional.
- If A is Fredholm of non-zero index then $0 \notin \sigma_{\text{ess},1}(A)$, while $0 \in \sigma_{\text{ess},2}(A)$. Thus the two essential spectra can differ.
- By Corollary B.106(ii), $\sigma_{\text{ess},1}(A) = \sigma_{\text{ess},2}(A)$ if A is a normal operator on Hilbert space.

Both essential spectra can be expressed in terms of the usual spectrum:

B.117 Lemma *If V is a complex Banach space and $A \in B(V)$ then with the quotient map $Q : B(V) \to B(V)/K(V)$*
$$\begin{aligned}\sigma_{\text{ess},1}(A) &= \sigma(Q(A)),\\ \sigma_{\text{ess},2}(A) &= \bigcap_{K \in K(V)} \sigma(A+K).\end{aligned}$$
Thus $\sigma_{\text{ess},1}(A)$ and $\sigma_{\text{ess},2}(A)$ are closed and non-empty.

Proof. The first statement is immediate by Atkinson's Theorem B.107 and implies that $\sigma_{\text{ess},1}(A)$ is closed and non-empty. And $\lambda \in \bigcap_{K \in K(V)} \sigma(A+K)$ is equivalent to the statement that $\lambda \in \sigma(A+K)$ for all compact K, thus $A + K - \lambda\mathbf{1}$ is non-invertible for all $K \in K(V)$. By Proposition B.110 this is equivalent to $A - \lambda\mathbf{1}$ not being Fredholm with index zero, thus to $\lambda \in \sigma_{\text{ess},2}(A)$. As an intersection of closed sets, $\sigma_{\text{ess},2}(A)$ is closed. And $\sigma_{\text{ess},2}(A) \supsetneq \sigma_{\text{ess},1}(A) \neq \emptyset$. ■

$\sigma_{\text{ess},2}(A)$ is the largest part of $\sigma(A)$ that is stable under compact perturbations of A:

B.118 Exercise If $\sigma' : B(V) \to P(\mathbb{C})$ is such that $\sigma'(A) \subseteq \sigma(A)$ and $\sigma'(A+K) = \sigma'(A)$ for all $A \in B(V)$ and $K \in K(V)$, prove that $\sigma'(A) \subseteq \sigma_{\text{ess},2}(A)$ for all $A \in B(V)$.

In this sense, $\sigma_{\text{ess},2}(A)$ is the 'best' definition of essential spectrum. Yet, we will have a look at a popular third definition, sitting between $\sigma_{\text{ess},2}(A)$ and $\sigma(A)$. The theory of the Riesz projector (Exercises 13.53, 13.66, 13.68) will play an important role. We begin with a preparatory result similar to Proposition B.110:

B.119 PROPOSITION *Let V be a complex Banach space and $A \in B(V)$. Then the following are equivalent:*

(i) *0 is an isolated point of $\sigma(A)$, and $P_0 V$ is finite-dimensional, where P_0 is the Riesz projector for $\lambda = 0$.*

(ii) *There are closed subspaces $V_1, V_2 \subseteq V$ with $V_1 + V_2 = V$ and $V_1 \cap V_2 = \{0\}$, where V_1 has non-zero finite dimension, such that $AV_i \subseteq V_i$. With $A_i = A \restriction V_i$, A_1 is nilpotent, and A_2 is invertible.*

(iii) *0 is an isolated point of $\sigma(A)$ and A is Fredholm of index zero.*

(iv) *0 is an isolated point of $\sigma(A)$ and A is Fredholm.*

(v) *A is not invertible, and there is a compact K such that $AK = KA$ and $A+K$ is invertible.*

Proof. (i)⇒(ii) With $V_1 = P_0 V$, $V_2 = (\mathbf{1} - P_0)V$, we have $AV_i \subseteq V_i$ with $\sigma(A \restriction V_1) = \{0\}$ and $\sigma(A \restriction V_2) = \sigma(A) \backslash \{0\}$. Thus $A_2 = A \restriction V_2$ is invertible and $A_1 = A \restriction V_1$ quasi-nilpotent, thus nilpotent by finite-dimensionality of V_1.

(ii)⇒(iii) By (ii) we have $A = A_1 \oplus A_2$, where A_2 is invertible and V_1 finite-dimensional. Thus A_1 is Fredholm of index zero, and the same holds for A. And $\sigma(A_1) = \{0\}$ while $0 \notin \sigma(A_2)$. Since $\sigma(A_2)$ is closed, there is an open neighborhood U of zero such that $U \cap \sigma(A_2) = \emptyset$, thus $U \cap \sigma(A) = \{0\}$. Thus 0 is an isolated point of $\sigma(A)$.

(iii)⇒(iv) This is trivial.

(iv)⇒(i) Since $0 \in \sigma(A)$ is isolated, it has a Riesz projector P_0. Let V_i, A_i as above. Then A_1 is quasi-nilpotent and A_2 is invertible. Together with the fact that A is Fredholm, this implies that $A_1 \in B(V_1)$ is Fredholm. Now Exercise B.120 below gives that V_1 is finite-dimensional.

(ii)⇒(v) It is clear that A is not invertible. Since A_1 is nilpotent, $A_1 + \mathbf{1}_{V_1} \in B(V_1)$ is invertible by Lemma 13.19. Now $K = \mathbf{1}_{V_1} \oplus 0$ is compact since $\dim V_1 < \infty$, commutes with $A = A_1 \oplus A_2$, and $A + K = (A_1 + \mathbf{1}_{V_1}) \oplus A_2$ is invertible.

(v)⇒(iii) By assumption, $B = A + K$ is invertible, implying that $A = B - K$ is Fredholm of index zero. Since B commutes with K (and A), with Exercise 15.8(ii) we have $\sigma(A) \subseteq \sigma(B) - \sigma(K)$. Since A is not invertible, we have $0 \in \sigma(A)$, thus there are $\lambda \in \sigma(B) \cap \sigma(K)$. Since $\sigma(B)$ is closed and does not contain zero, there is an $\varepsilon > 0$ such that $B(0, \varepsilon) \cap \sigma(B) = \emptyset$. Since $\sigma(K)$ has zero as only limit point, we see that $\sigma(B) \cap \sigma(K)$ is finite. If $\lambda \in \sigma(K) \backslash \{0\}$ and $r > 0$ small enough so that $B(\lambda, r) \cap \sigma(K) = \{\lambda\}$ then A commutes with $\mathbf{1} - ((K - \lambda\mathbf{1})/z)^n$ (which is invertible) for each n, thus also with the inverses and with their limit, the Riesz projector (for K!)

$$P_\lambda = \lim_{n \to \infty} (\mathbf{1} - ((K - \lambda\mathbf{1})/z)^n)^{-1}.$$

Thus A maps each of the spaces $V_\lambda = P_\lambda V$ into themselves, and the restrictions of A and K commute on each V_λ. Now $P = \sum_{\lambda \in \sigma(B) \cap \sigma(K)} P_\lambda$ is a finite sum, and we put $V_1 = PV$, $V_2 = (\mathbf{1} - P)V$. On V_2, both $B = A + K$ and A are invertible, while $A \restriction V_1$ is not. Since V_1 is finite-dimensional, A is Fredholm of index zero, and zero is an isolated point of $\sigma(A)$. ■

B.120 EXERCISE Let V be a complex Banach space and $A \in B(V)$ quasi-nilpotent and Fredholm. Prove that V is finite-dimensional.

B.121 DEFINITION *Let V be a complex Banach space and $A \in B(V)$. Then*

- *The discrete spectrum $\sigma_{\mathrm{d}}(A)$ of A is the set of $\lambda \in \mathbb{C}$ for which $A - \lambda\mathbf{1}$ satisfies the equivalent conditions in Proposition B.119.*

- *The Browder*[136] *essential spectrum is* $\sigma_{\text{ess},3}(A) = \sigma(A) \backslash \sigma_{\text{d}}(A)$.

B.122 Remark From property (v) in Proposition B.119 it is evident that $\sigma_{\text{d}}(A)$ is definitely not stable under compact perturbations! □

B.123 Proposition *Let V be a complex Banach space. Then the Browder essential spectrum $\sigma_{\text{ess},3}(A)$ satisfies*

$$\sigma_{\text{ess},2}(A) \subseteq \sigma_{\text{ess},3}(A) = \bigcap_{K \in K(V) \cap \{A\}'} \sigma(A+K) \ \subseteq \sigma(A) \tag{B.10}$$

and is is closed and non-empty.

Proof. $\sigma_{\text{ess},3} \subseteq \sigma(A)$ holds by definition. And $\sigma_{\text{ess},2}(A) \subseteq \sigma_{\text{ess},3}(A)$ clearly follows from the fact that one of the equivalent definitions of $\sigma_{\text{ess},3}(A)$ resulting from the above is

$$\sigma_{\text{ess},3}(A) = \{\lambda \in \sigma(A) \mid \lambda \text{ not isolated or } A - \lambda\mathbf{1} \text{ not Fredholm of index zero}\}. \tag{B.11}$$

If $\lambda \in \mathbb{C}$ is such that $A - \lambda\mathbf{1}$ is not Fredholm of index zero then on the one hand $\lambda \in \sigma_{\text{ess},3}(A)$. On the other, for all compact K Theorem B.108(iii) gives that, $A + K - \lambda\mathbf{1}$ is not Fredholm of index zero, thus not invertible, so that thus $\lambda \in \sigma(A+K)$. A fortiori, λ is in the intersection in (B.10). Since the latter is contained in $\sigma(A)$, in order to complete the proof of (B.10), it suffices to prove that "$\lambda \in \sigma_{\text{ess},3}(A) \Leftrightarrow \lambda \in \sigma(A+K)$ for all compact K commuting with A" holds under the assumption that $A - \lambda\mathbf{1}$ is Fredholm of index zero. It suffices to consider the case $\lambda = 0$, which amounts to the statement that 0 is a non-isolated point of $\sigma(A)$ if and only if there is no compact K commuting with A such that $A + K$ is invertible. This is exactly the equivalence (iii)⇔(v) in Proposition B.119. ■

If the Browder essential spectrum $\sigma_{\text{ess},3}(A)$ is bigger than the Weyl essential spectrum $\sigma_{\text{ess},2}(A)$, it cannot be stable under all compact perturbations of A. But the above shows that it still has a very nice stability property.

B.10.2 Normal Hilbert space operators. Weyl's theorem

We now restrict our attention to normal operators on Hilbert spaces.

B.124 Theorem *Let H be a complex Hilbert space and $A \in B(H)$ normal. Then*

$$\sigma_{\text{ess},3}(A) = \{\lambda \in \sigma(A) \mid \lambda \text{ not isolated or } \dim\ker(A - \lambda\mathbf{1}) = \infty\} \tag{B.12}$$

$$= \sigma_{\text{ess},1}(A) = \sigma_{\text{ess},2}(A). \tag{B.13}$$

(We simply call this the essential spectrum $\sigma_{\text{ess}}(A)$.) Whenever A, A' are normal and $A - A'$ is compact, $\sigma_{\text{ess}}(A) = \sigma_{\text{ess}}(A')$.

Proof. If $\lambda \in \sigma(A)$ is not isolated, by definition it is contained in (B.11) and in (B.12). If λ is isolated then by Proposition 17.24 there are mutually orthogonal A-invariant subspaces $H_1, H_2 \subseteq H$ with $H = H_1 + H_2$ such that $H_1 = \ker(A - \lambda\mathbf{1}) \neq 0$ and $(A - \lambda\mathbf{1}) \upharpoonright H_2$ is invertible. Thus $A - \lambda\mathbf{1}$ is Fredholm if and only if its restriction to H_1 is Fredholm. The latter being

136 Felix Earl Browder (1927-2016). American mathematician who mostly worked on functional analysis and differential equations.

identically zero, this is equivalent to $\dim H_1 = \dim\ker(A - \lambda\mathbf{1}) < \infty$. This proves the equality in (B.12).

We already know that for normal A we have $\sigma_{\text{ess},1}(A) = \sigma_{\text{ess},2}(A) \subseteq \sigma_{\text{ess},3}(A)$. Thus to prove equality of the three spectra we must show $\sigma_{\text{ess},3}(A) \subseteq \sigma_{\text{ess},1}(A)$. Let thus $\lambda \in \sigma_{\text{ess},3}(A)$. By (B.12), either $\ker(A - \lambda\mathbf{1})$ is infinite-dimensional or $\lambda \in \sigma(A)$ is not isolated. In the first case, $A - \lambda\mathbf{1}$ is not Fredholm, thus $\lambda \in \sigma_{\text{ess},1}(A)$, and we are done.

We are thus reduced to the situation where $\ker(A - \lambda\mathbf{1})$ is finite-dimensional and λ is not isolated. By Proposition 11.27(ii), we have a direct sum situation $H = (\ker(A - \lambda\mathbf{1})) \oplus H'$, where $H' = (\ker(A - \lambda\mathbf{1}))^\perp$. Since $\ker(A - \lambda\mathbf{1})$ is finite-dimensional, $A - \lambda\mathbf{1}$ is Fredholm if and only if $(A - \lambda\mathbf{1}) \restriction H'$ is. Since $(A - \lambda\mathbf{1}) \restriction H'$ is by construction injective, it has dense image by normality and Proposition 11.27(iii). Since $\lambda \in \sigma(A)$ is not isolated, Corollary 17.25 gives that $(A - \lambda\mathbf{1}) \restriction H'$ is not invertible. Since it is injective, it is not surjective. Now Exercise 7.11(iii) gives that $(A - \lambda\mathbf{1}) \restriction \ker(A - \lambda\mathbf{1})^\perp$ has infinite-dimensional cokernel. It thus is not Fredholm, thus also $A - \lambda\mathbf{1}$ is not Fredholm, implying $\lambda \in \sigma_{\text{ess},1}(A)$. This finishes the proof of the equality of the three spectra for normal operators. The invariance of $\sigma_{\text{ess},3}(A)$ under compact perturbations now follows from that of $\sigma_{\text{ess},2}(A)$. ■

B.125 Exercise If H is an infinite-dimensional Hilbert space and $A \in B(H)$ is normal, prove $\sigma_{\text{ess}}(A) \neq \emptyset$.

B.126 Remark The essential spectrum was introduced [170, 171] in 1909/10 by Weyl[137], who only considered self-adjoint operators, but allowed unbounded ones since he was studying differential equations. At that time functional analysis had just started developing, and the tools used prove Theorem B.124 were not yet available. (The bounded inverse theorem came in 1929 and Fredholm operators and Theorem B.108 around 1950!) Weyl's original approach to proving invariance of $\sigma_{\text{ess}}(A)$ under compact perturbations was quite different (but not totally) and is also interesting, which is why we briefly discuss it now. □

B.127 Definition *Let H be a Hilbert space and $A \in B(H)$, $\lambda \in \mathbb{C}$. A Weyl sequence for (A, λ) is an orthonormal sequence $\exists\{x_n\} \subset H$ such that $\|(A - \lambda\mathbf{1})x_n\| \to 0$.*

B.128 Proposition *Let H be a complex Hilbert space, $A \in B(H)$ normal and $\lambda \in \mathbb{C}$. Then the following are equivalent:*

(i) $\lambda \in \sigma_{\text{ess}}(A)$.

(ii) $P_A(B(\lambda, \varepsilon))H \subseteq H$ *is infinite-dimensional for each* $\varepsilon > 0$. *(Compare Proposition 18.20.)*

(iii) There exists a Weyl sequence for (A, λ).

Proof. (i)⇒(ii) If $\dim\ker(A - \lambda\mathbf{1}) = \infty$ then already $P_A(\{\lambda\})H = \ker(A - \lambda\mathbf{1})$ is infinite-dimensional. If λ is an accumulation point of eigenvalues, the linear span of the eigenspaces $\ker(A - \lambda'\mathbf{1})$ with $\lambda' \in B(\lambda, \varepsilon)$ is infinite-dimensional for each $\varepsilon > 0$. If neither of these holds, there is an $\varepsilon > 0$ such that the restriction of A to $H' = P(B(\lambda, \varepsilon'))H \cap (\ker(A - \lambda\mathbf{1}))^\perp)$ has purely continuous spectrum containing λ for all $\varepsilon' \in (0, \varepsilon)$. In this case, $H' \subseteq P(B(\lambda, \varepsilon))H$ is infinite-dimensional for all $\varepsilon > 0$ since otherwise $A \restriction H'$ would have λ as eigenvalue. (Compare also Exercise 18.8(ii).)

[137] Hermann Weyl (1885-1955). German mathematician, who worked in many areas of mathematics and mathematical physics, like real, complex and functional analysis, differential equations, Lie groups, quantum theory and relativity, as well as philosophy of mathematics.

(ii)⇒(iii) By (ii), for each $n \in \mathbb{N}$ we can pick $x_n \in P_A(B(\lambda, 1/n))H \cap \{x_1, \ldots, x_{n-1}\}'$ and $\|x_n\| = 1$. Then $\{x_n\}$ is an orthonormal sequence, and $\|(A - \lambda)x_n\| \leq 1/n \to 0$. Thus $\{x_n\}$ is a Weyl sequence.

(iii)⇒(i) Existence of a Weyl sequence implies that $A - \lambda\mathbf{1}$ is not bounded below, thus $\lambda \in \sigma(A)$. Assume $\lambda \notin \sigma_{\text{ess}}(A)$, thus $\lambda \in \sigma_{\text{d}}(A)$. Then $A = A_1 \oplus A_2$ with $A_1 = 0$ and $A_2 - \lambda\mathbf{1}_{H_2}$ invertible, thus bounded below. With $\|(A - \lambda\mathbf{1})x_n\| \to 0$, this implies $\|P_2 x_n\| \to 0$. Combined with $\|x_n\| = 1\ \forall n$, this gives $\|P_1 x_n\| \to 1$. Together with the fact that H_1 is finite-dimensional and the $\{x_n\} \subseteq H$ orthonormal, this produces a contradiction. Thus $\lambda \in \sigma_{\text{ess}}(A)$. ■

Now we have an alternative proof for the invariance of the essential spectrum under compact perturbations:

B.129 Theorem (Weyl) *Let V be a complex Hilbert space and $A, B \in B(H)$ normal such that $A - B$ is compact. Then $\sigma_{\text{ess}}(A) = \sigma_{\text{ess}}(B)$.*

Proof. If a Weyl sequence for A, λ exists, it is clear that $\lambda \in \sigma_{\text{app}} = \sigma(A)$.

Assume $\lambda \in \sigma_{\text{ess}}(A)$. Then there exists a Weyl sequence $\{x_n\}$. Now

$$\|(B - \lambda\mathbf{1})x_n\| \leq \|(A - \lambda\mathbf{1})x_n\| + \|(B - A)x_n\|.$$

Now $\|(A - \lambda\mathbf{1})x_n\| \to 0$ by $\lambda \in \sigma_{\text{ess}}(A)$, while $\|(B - A)x_n\| \to 0$ since $\{x_n\}$ is a weak null sequence and $B - A$ is compact, thus sequentially weak-norm continuous, cf. Section 12.2. Using Weyl's criterion again, $\lambda \in \sigma_{\text{ess}}(B)$. Thus $\sigma_{\text{ess}}(A) \subseteq \sigma_{\text{ess}}(B)$. The converse inclusion follows by $A \leftrightarrow B$. ■

B.130 Remark There are many generalizations of the theorem, e.g. to unbounded operators. Cf. e.g. [45] and [129, Section XIII.4]. These generalizations have many applications to differential equations and quantum mechanics. □

B.10.3 Applications

Apart from the mentioned applications of the essential spectrum to differential equations, there are applications to operator theory. We mention a few without proofs:

B.131 Theorem (Weyl-von Neumann-Berg) *If H is a separable complex Hilbert space, $A \in B(H)$ is normal and $\varepsilon > 0$, there is a compact $K \in B(H)$ with $\|K\| < \varepsilon$ such that $D = A - K$ is diagonal. (I.e. there exists an ONB E for H that diagonalizes D.)*

For the proof, which has nothing particular to do with the essential spectrum, see e.g. [33, Section II.4]. Combining this with the theory of the essential spectrum, we obtain:

B.132 Theorem *Let H be a separable complex Hilbert space and $A_1, A_2 \in B(H)$ be normal. Then $\sigma_{\text{ess}}(A_1) = \sigma_{\text{ess}}(A_2)$ if and only if there exists a unitary U such that $A_2 - UA_1U^*$ is compact. (One says A_1 and A_2 are 'essentially unitarily equivalent' or 'compalent').*

Proof. If A_2 is normal and U unitary, it is clear that UA_2U^* is normal and $\sigma_{\text{ess}}(UA_2U^*) = \sigma_{\text{ess}}(A_2)$. Thus if $A_1 - UA_2U^*$ is compact then $\sigma_{\text{ess}}(A_1) = \sigma_{\text{ess}}(UA_2U^*) = \sigma_{\text{ess}}(A_2)$.

Now assume $\sigma_{\text{ess}}(A_1) = \sigma_{\text{ess}}(A_2)$. By Weyl-von Neumann-Berg there are diagonal D_1, D_2 and compact K_1, K_2 such that $A_i = D_i + K_i$. Since essential spectra are stable under compact perturbations,

$$\sigma_{\text{ess}}(D_1) = \sigma_{\text{ess}}(A_1) = \sigma_{\text{ess}}(A_2) = \sigma_{\text{ess}}(D_2).$$

For a diagonal operator $D = \text{diag}(d_n)$ we have $\lambda \in \sigma_{\text{ess}}(D)$ if and only λ is an accumulation point of $\{d_n\}$, i.e. $\{n \in \mathbb{N} \mid |d_n - \lambda| < \varepsilon\}$ is infinite for every $\varepsilon > 0$. Thus the eigenvalue sequences $\{d_{1,n}\}, \{d_{2,n}\}$ of D_1, D_2 have the same limit points. Using this one can construct a permutation σ of $\mathbb{N}$ such that $|d_{1,n} - d_{2,\sigma(n)}| \to 0$. If $\{e_{1,n}\}, \{e_{2,n}\}$ are the ONBs diagonalizing D_1, D_2, respectively, there is a unique unitary U such that $Ue_{1,n} = e_{2,\sigma(n)}$ $\forall n$. Now $UD_1U^* - D_2$ is compact, thus $UA_1U^* - A_2 = UD_1U^* + UK_1U^* - D_2 - K_2$ is compact. ■

B.133 Theorem *Let H be a separable Hilbert space and $A, B \in B(H)$ normal. Then the following are equivalent:*

(i) There is a sequence $\{U_n\}$ of unitaries such that $B = \lim_{n\to\infty} U_n A U_n^*$. ('$A$ and B are approximately unitarily equivalent.')

(ii) $\sigma_{\text{ess}}(A) = \sigma_{\text{ess}}(B)$ and $\dim\ker(A - \lambda\mathbf{1}) = \dim\ker(B - \lambda\mathbf{1})$ for all $\lambda \in \mathbb{C}\backslash\sigma_{\text{ess}}(A)$.

Proof. See [33, Theorem II.4.4]. ■

B.134 Theorem (Brown-Douglas-Fillmore 1973) *Given $A \in B(H)$, we have $A = N + K$ with N normal and K compact if and only if $A^*A - AA^*$ is compact (A is 'essentially normal') and $\text{ind}(A - \lambda\mathbf{1}) = 0$ for all $\lambda \notin \sigma_{\text{ess},1}(A)$.*

Proof. $\Rightarrow$ is quite trivial: If $A = N + K$ with N normal, K compact, then

$$A^*A - AA^* = (N^* + K^*)(N + K) - (N + K)(N^* + K^*) = N^*N - NN^* - \text{compact} \in K(H),$$

and $\text{ind}(A - \lambda\mathbf{1}) = \text{ind}(N + K - \lambda\mathbf{1}) = \text{ind}(N - \lambda\mathbf{1}) = 0$ whenever $N - \lambda\mathbf{1}$ is Fredholm, i.e. $\lambda \notin \sigma_{\text{ess}}(N) = \sigma_{\text{ess},1}(A)$. For the much deeper converse see e.g. [33, 75]. ■

The above is just the tip of an iceberg. For more, see the references given above.

B.11 Trace-class operators: $L^1(H)$

B.135 Definition *If H is a Hilbert space, $A \in B(H)$ and $1 \le p < \infty$, we define*

$$\|A\|_p = (\text{Tr}(|A|^p))^{1/p} \in [0, \infty], \qquad L^p(H) = \{A \in B(H) \mid \|A\|_p < \infty\}.$$

For $p = 2$ this agrees with Definition 12.39 since $|A|^2 = A^*A$. For $p = 1$ it specializes to

B.136 Definition *Let H be a Hilbert space and $A \in B(H)$. Then*

$$\|A\|_1 = \text{Tr}|A|, \qquad L^1(H) = \{A \in B(H) \mid \text{Tr}|A| < \infty\}.$$

The elements of $L^1(H)$ are called trace class operators.

Trace class operators play an important role in von Neumann algebra theory. Our treatment is inspired by [128]. See also [92].

B.137 Theorem *Let H be any Hilbert space. Then*

(i) $\|A\| \le \|A\|_1$ for all $A \in B(H)$.

(ii) $\|A^*\|_1 = \|A\|_1$ for all $A \in B(H)$. Thus $L^1(H)$ is self-adjoint.

(iii) For all $\lambda \in \mathbb{C}$, $A, B \in B(H)$ we have $\|\lambda A\|_1 = |\lambda|\|A\|_1$ and $\|A + B\|_1 \le \|A\|_1 + \|B\|_1$. If $0 \ne A \in L^1(H)$ then $\|A\|_1 > 0$. Thus $(L^1(H), \|\cdot\|_1)$ is a normed vector space.

(iv) *For all* $A, B \in B(H)$ *we have* $\|AB\|_1 \le \|A\|\|B\|_1$ *and* $\|AB\|_1 \le \|A\|_1\|B\|$. *Thus* $L^1(H) \subseteq B(H)$ *is a two-sided ideal.*

(v) $F(H) \subseteq L^1(H) \subseteq K(H)$.

(vi) $\overline{F(H)}^{\|\cdot\|_1} = L^1(H)$.

(vii) *The normed space* $(L^1(H), \|\cdot\|_1)$ *is complete, thus a Banach space.*

(viii) $(L^1(H), \|\cdot\|_1)$ *is a Banach* $*$*-algebra.*

(ix) *For* $A \in B(H)$, *the following are equivalent:*

- *(α)* $A \in L^1(H)$, *i.e.* $\mathrm{Tr}(|A|) < \infty$.
- *(β)* A *is a finite linear combination of positive operators with finite trace.*
- *(γ)* $\sum_e |\langle VAe, e\rangle| < \infty$ *for each ONB* E *and each unitary* V.[138]
- *(δ)* $\sum_e |\langle VAe, e\rangle| < \infty$ *for some ONB* E *and each unitary* V. *Under this condition,* $\|A\|_1 \le \sup_{V\in U(H)} \sum_e |\langle VAe, e\rangle|$.
- *(ϵ)* $\sum_e |\langle Ae, e\rangle| < \infty$ *for each ONB* E.

(x) *For each ONB* E, *the unordered sum in*

$$\mathrm{Tr} : L^1(H) \to \mathbb{C},\ A \mapsto \sum_{e\in E} \langle Ae, e\rangle$$

is convergent in the sense of Section A.1 (thus absolutely convergent) and independent of the choice of E, *defining a linear functional on* $L^1(H)$.

(xi) *For all* $A \in L^1(H)$ *we have* $\mathrm{Tr}(A^*) = \overline{\mathrm{Tr}(A)}$.

(xii) *If* $A \in B(H), b \in L^1(H)$ *then* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ *and* $|\mathrm{Tr}(AB)| \le \|A\|\|B\|_1$. *In particular* $|\mathrm{Tr}(B)| \le \|B\|_1$, *thus* $\mathrm{Tr} \in (L^1(H), \|\cdot\|_1)^*$.

Proof. (i) Let $B \ge 0$ and $x \in H$ a unit vector. If E is an ONB containing x, we have $\langle Bx, x\rangle \le \mathrm{Tr}_E(B)$. Since B is positive, we have $\|B\| = \sup_{x,\|x\|=1}\langle Bx, x\rangle$, thus $\|B\| \le \mathrm{Tr}(B)$. If now $A \in B(H)$ with polar decomposition $A = U|A|$, then applying the above to $B = |A|$ and using $\|U\| \le 1$ gives

$$\|A\| = \|U|A|\,\| \le \|\,|A|\,\| \le \mathrm{Tr}|A| = \|A\|_1.$$

(ii) Let $A \in B(H)$ with polar decomposition $A = U|A|$ and $|A| = U^*A$. Since U is a partial isometry, it satisfies $U^*UU^* = U^*$, so that $U^*U|A| = U^*UU^*A = U^*A = |A|$. Thus

$$(U|A|U^*)^2 = U|A|U^*U|A|U^* = U|A|^2U^* = (U|A|)(U|A|)^* = AA^* = |A^*|^2.$$

Since $|A^*|$ and $U|A|U^*$ are both positive, taking roots gives $U|A|U^* = |A^*|$. Choosing an ONB E such that each $e \in E$ is either in $\ker U^*$ or in $(\ker U^*)^\perp$, thus $U^*e = 0$ or $U^*e = e$, we find

$$\|A^*\|_1 = \mathrm{Tr}_E|A^*| = \mathrm{Tr}_E(U|A|U^*) = \sum_e \langle |A|U^*e, U^*e\rangle \le \sum_e \langle |A|e, e\rangle = \mathrm{Tr}|A| = \|A\|_1.$$

Replacing A by A^* gives the opposite inequality so that $\|A^*\|_1 = \|A\|_1$.

[138]This is the same as $\sum_{i\in I} |\langle Ae_i, f_i\rangle| < \infty$ for any ONBs $\{e_i\}_{i\in I}, \{f_i\}_{i\in I}$.

(iii) The first statement follows from $|\lambda A| = |\lambda||A|$. For the second, let $A, B \in B(H)$ with polar decompositions $A = U|A|,\ B = V|B|,\ A+B = W|A+B|$. If E is an ONB and $F \subseteq E$ is finite,

$$\begin{aligned}\sum_{e\in F}\langle |A+B|e,e\rangle &= \sum_{e\in F}\langle W^*(A+B)e,e\rangle = \sum_{e\in F}(\langle W^*U|A|e,e\rangle + \langle W^*V|B|e,e\rangle) \\ &\leq \sum_{e\in F}|\langle W^*U|A|e,e\rangle| + \sum_{e\in F}|\langle W^*V|B|e,e\rangle|. \end{aligned} \tag{B.14}$$

Focusing on the first term of the r.h.s., we have

$$\begin{aligned}\sum_{e\in F}|\langle W^*U|A|e,e\rangle| &= \sum_{e\in F}|\langle |A|^{1/2}e, |A|^{1/2}U^*We\rangle| \leq \sum_{e\in F}\|\,|A|^{1/2}e\|\;\||A|^{1/2}U^*We\| \\ &\leq \left(\sum_{e\in F}\|\,|A|^{1/2}e\|^2\right)^{1/2}\left(\sum_{e\in F}\||A|^{1/2}U^*We\|^2\right)^{1/2}, \end{aligned} \tag{B.15}$$

where the first $\leq$ comes from applying Cauchy-Schwarz to the inner product $\langle |A|^{1/2}e, |A|^{1/2}U^*We\rangle$ in H, the second $\leq$ from Cauchy-Schwarz in $\mathbb{C}^{|F|}$.

The argument of the first square root in the r.h.s. of (B.15) is dominated by $\sum_{e\in E}\|\,|A|^{1/2}e\|^2 = \mathrm{Tr}|A|$, and for the argument of the second root we have

$$\begin{aligned}\sum_{e\in F}\||A|^{1/2}U^*We\|^2 &= \sum_{e\in F}\langle |A|^{1/2}U^*We, |A|^{1/2}U^*We\rangle = \sum_{e\in F}\langle W^*U|A|U^*We,e\rangle \\ &\leq \mathrm{Tr}(W^*U|A|U^*W\rangle. \end{aligned}$$

Now, picking an ONB E such that each $e \in E$ is either in $\ker W$ or in $(\ker W)^\perp$, we find $\mathrm{Tr}(W^*U|A|U^*W) \leq \mathrm{Tr}(U|A|U^*)$. Repeating the argument with U, we have $\mathrm{Tr}(U|A|U^*) \leq \mathrm{Tr}|A|$. Thus $\mathrm{Tr}(W^*U|A|U^*W) \leq \mathrm{Tr}|A|$, so that $\sum_{e\in F}\||A|^{1/2}U^*We\|^2 \leq \mathrm{Tr}|A|$. Inserting this in (B.15), we find

$$\sum_{e\in F}|\langle W^*U|A|e,e\rangle| \leq \mathrm{Tr}|A| = \|A\|_1.$$

Analogously one proves the bound $\sum_{e\in F}|\langle W^*V|B|e,e\rangle| \leq \mathrm{Tr}|B| = \|B\|_1$ for the other summand in (B.14). Now taking the limit $F \nearrow E$ we have $\|A+B\|_1 \leq \|A\|_1 + \|B\|_1$. In view of this, it is clear that $L^1(H)$ is a vector space.

(iv) Let $B = U|B|$ and $AB = V|AB|$ be the polar decompositions of B, AB, respectively. Using Proposition 11.44(iii), we have $|AB| = V^*AB = V^*AU|B| = W|B|$, where $W = V^*AU$. In view of $\|U\|, \|V\| \leq 1$ we have $\|W\| \leq \|A\|$. Using $W^*W \leq \|W^*W\|\mathbf{1} = \|W\|^2\mathbf{1}$ and Exercise 17.7 we have

$$|AB|^2 = |AB|^*|AB| = |B|W^*W|B| \leq \|W\|^2|B|^2.$$

In view of $0 \leq A \leq B \Rightarrow A^{1/2} \leq B^{1/2}$, cf. [110, Theorem 2.2.6], this implies $|AB| \leq \|W\|\,|B|$. Thus $\|AB\|_1 = \mathrm{Tr}|AB| \leq \|W\|\mathrm{Tr}|B| = \|W\|\|B\|_1 \leq \|A\|\|B\|_1$.

The other inequality follows by $\|AB\|_1 = \|(AB)^*\|_1 = \|B^*A^*\|_1 \leq \|B^*\|\|A^*\|_1 = \|A\|_1\|B\|$, where we used the bound just proven and (ii). That $L^1(H)$ is an ideal now is obvious.

(v) We have $(x \otimes y)^* = (y \otimes x)$, thus $(x\otimes y)^*(x\otimes y) = \|x\|^2(y\otimes y) = \|x\|^2\|y\|^2(e\otimes e)$ with $e = y/\|y\|$, so that taking roots gives $|x\otimes y| = \|x\|\|y\|(e\otimes e)$, which clearly is in $L^1(H)$. Since every element of $F(H)$ is a finite linear combination of such $x\otimes y$, we have $F(H) \subseteq L^1(H)$.

If $A \in L^1(H)$ the then $|A|^2 = A^*A \in L^1(H)$ since $L^1(H)$ is an ideal. Thus for any ONB E we have

$$\sum_{e\in E}\|\,Ae\|^2 = \sum_e\langle A^*Ae,e\rangle = \sum_e\langle |A|^2e,e\rangle = \mathrm{Tr}_E(|A|^2) < \infty.$$

Let $F \subseteq E$ be finite and $x \in F^\perp$ with $\|x\| = 1$. Then $F \cup \{x\}$ is an orthonormal set and can be completed to an ONB E. Thus $\sum_{e\in F} \| Ae\|^2 + \| Ax\|^2 \le \mathrm{Tr}(|A|^2)$, or

$$\|Ax\|^2 \le \mathrm{Tr}(|A|^2) - \sum_{e\in F} \|Ae\|^2.$$

Since the r.h.s. goes to zero as $F \nearrow E$, we have

$$\sup\left\{\|Ax\| \mid x \in F^\perp, \|x\| = 1\right\} \stackrel{F\nearrow E}{\longrightarrow} 0. \tag{B.16}$$

If P_F is the orthogonal projection onto $\mathrm{span}_{\mathbb{C}} F$ then $A_F := AP_F$ is a finite rank operator that converges in norm to A by (B.16). Thus $A \in \overline{F(H)}^{\|\cdot\|} = K(H)$, proving $L^1(H) \subseteq K(H)$.

(vi) Assume first $A \in L^1(H)^+$. By (v), A is compact. By Theorem 14.12, compact self-adjoint operators can be diagonalized, thus there is an ONB E such that $A = \sum_{e\in E} \lambda_e\, e \otimes e$. In our case, $\lambda_e \ge 0$ for all $e \in E$ and $\sum_e \lambda_e = \mathrm{Tr}(A) < \infty$. For a finite subset $F \subseteq E$ define $A_F := \sum_{e\in F} \lambda_e\, e \otimes e$, which is finite rank. Now $A - A_F = \sum_{e\in E\backslash F} \lambda_e\, e \otimes e \ge 0$. Thus $\|A - A_F\|_1 = \mathrm{Tr}(A - A_F) = \sum_{e\in E\backslash F} \lambda_e$. With $\sum_e \lambda_e < \infty$ this implies $\|A - A_F\|_1 \to 0$ as $F \nearrow E$, thus $A \in \overline{F(H)}^{\|\cdot\|_1}$.

Let now $A \in L^1(H)$ with polar decomposition $A = U|A|$. Then $|A| \in L^1(H)$, thus by the above for each $\varepsilon > 0$ there is $B \in F(H)$ with $\| \,|A| - B\|_1 < \varepsilon$. With (iv) and $\|U\| \le 1$ this implies $\|A - UB\|_1 = \|U(|A| - B)\|_1 \le \||A| - B\|_1 < \varepsilon$. Since $F(H)$ is an ideal, we have $UB \in F(H)$, finishing the proof of $L^1(H) \subseteq \overline{F(H)}^{\|\cdot\|_1}$. The converse is clear.

(vii) This can be proven directly, using $L^1(H) \subseteq K(H)$, cf. e.g. [118, Theorem 3.4.12]. Since we don't need the result soon, we will follow Murphy in deducing it later from the isometric isomorphism $L^1(H) \cong B(H)^*$ of normed spaces and completeness of the dual space $B(H)^*$.

(viii) It only remains to prove submultiplicativity: If $A, B \in L^1(H)$ then $\|AB\|_1 \le \|A\|\|B\|_1 \le \|A\|_1\|B\|_1$, where we used (i) and (iv).

(ix) $(\beta) \Rightarrow (\alpha)$: This is trivial since $L^1(H)$ is a vector space by (iii) and obviously contains the positive operators of finite trace.

$(\alpha) \Rightarrow (\beta)$ Assume $A \in L^1(H)$. By (ii), $A^* \in L^1(H)$ so that (iii) implies $\mathrm{Re}(A), \mathrm{Im}(A) \in L^1(H)$. If $A = A^* \in L^1(H)$, let $A = A_+ - A_-$ be the canonical decomposition with $A_\pm \ge 0$ and $A_+A_- = 0$. Then $|A| = A_+ + A_-$, so that $\mathrm{Tr}(A_\pm) \le \mathrm{Tr}|A| = \|A\|_1 < \infty$, implying $A_\pm \in L^1(H)$. Thus every trace class operator is a linear combination of four (or less) positive trace class operators.

$(\beta) \Rightarrow (\gamma)$ Let $A \in L^1(H)$ and E an ONB for H. By (β), we have $A = \sum_{k=1}^K \lambda_k A_k$ with $A_k \in L^1(H)^+\ \forall k$. Now $|\langle Ae, e\rangle| \le \sum_{k=1}^K |\lambda_k| \langle A_k e, e\rangle$ for all $e \in E$, so that $A_k \in L^1(H)^+\ \forall k$ implies $\sum_e |\langle Ae, e\rangle| < \infty$. Since $L^1(H)$ is an ideal, the same holds for VA instead of A.

$(\gamma) \Rightarrow (\delta) + (\epsilon)$ is trivial.

$(\delta) \Rightarrow (\alpha)$ Let $A = U|A|$ be the polar decomposition of $A \in B(H)$. Recall that U maps $\overline{|A|H} \subseteq H$ isometrically to $\overline{AH} \subseteq H$ and sends $\overline{|A|H}^\perp$ to zero. Let $V = U^*$ on $\overline{AH} = u\overline{|A|H} \subseteq H$. Since the closed subspaces $\overline{|A|H}$ and $\overline{AH}$ are unitarily equivalent, they have the same dimension, thus also $\overline{|A|H}^\perp$ and $\overline{AH}^\perp$ have the same dimension. Define $V : \overline{AH}^\perp \to \overline{|A|H}^\perp$ to be (any) isometry. Then V is unitary. Since $VU|A|e = U^*U|A|e = |A|e$ for all e, we have

$$\sum_{e\in E} |\langle VU|A|e, e\rangle| = \sum_{e\in E} |\langle |A|e, e\rangle| = \sum_{e\in E} \langle |A|e, e\rangle = \mathrm{Tr}|A|.$$

By assumption, the l.h.s. is finite, thus $\mathrm{Tr}|A| < \infty$. The bound on $\|\cdot\|_1$ is obvious in view of the preceding computation.

(ϵ) $\Rightarrow$ (α) Assumption (ε) clearly implies $\langle Ae_n, e_n\rangle \to 0$ for any orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$. Now Theorem 12.37 gives that A is compact. If $A = B + iC$ with B, C self-adjoint then B, C are compact, thus diagonalizable. Thus $B = \sum_{f\in F} \lambda_f f \otimes f$ for a certain ONB F and $\lambda_f \in \mathbb{R}$. Now (ϵ), which also holds for B, C, clearly implies $\|B\|_1 = \mathrm{Tr}|B| = \sum_f |\lambda_f| < \infty$, so that $B \in L^1(H)$. Similarly, $C \in L^1(H)$, thus also $A = B + iC$ is trace-class by (iii).

(x) Let E be an ONB. By (ix)(γ), the sum $\sum_{e\in E}\langle Ae, e\rangle$ is absolutely convergent for each $A \in L^1(H)$. This proves that $\mathrm{Tr}_E : L^1(H) \to \mathbb{C}$ is well-defined and linear. It remains to show that Tr_E is independent of E. By (ix)(β), we have $A = \sum_{k=1}^K \lambda_k A_k$ where K is finite and $A_k \in L^1(H)^+ \ \forall k$. If now F is another ONB, we have

$$\mathrm{Tr}_E(A) = \sum_k \lambda_k \mathrm{Tr}_E(A_k) = \sum_k \lambda_k \mathrm{Tr}_F(A_k) = \mathrm{Tr}_F(A),$$

where we used the linearity of Tr_E and Tr_F and the fact that $\mathrm{Tr}_E(A_k) = \mathrm{Tr}_F(A_k)$ by $A_k \geq 0$ and Lemma 11.51(ii). This proves $\mathrm{Tr}_E = \mathrm{Tr}_F$.

(xi) If $A \in L^1(H)$ and E is an ONB, we have

$$\mathrm{Tr}(A^*) = \sum_e \langle A^*e, e\rangle = \sum_e \langle e, Ae\rangle = \sum_e \overline{\langle Ae, e\rangle} = \overline{\mathrm{Tr}(A)},$$

where the last identity comes from absolute convergence and continuity of complex conjugation.

(xii) Let $B \in L^1(H)$ throughout. If U is unitary then $BU, UB \in L^1(H)$ and

$$\mathrm{Tr}_E(BU) = \sum_{e\in E}\langle BUe, e\rangle = \sum_{e\in E}\langle UBUe, Ue\rangle = \sum_{f\in F}\langle UBf, f\rangle = \mathrm{Tr}_F(UB),$$

where $F = UE$ is another ONB. Since Tr does not depend on the ONB, we have $\mathrm{Tr}(UB) = \mathrm{Tr}(BU)$. If now $A \in B(H)$ is arbitrary, we have $A = \sum_{l=1}^4 \lambda_l U_l$ where $\lambda_l \in \mathbb{C}$ and the U_l are unitaries. Then $\mathrm{Tr}(AB) = \sum_l \lambda_l \mathrm{Tr}(U_l B) = \sum_l \lambda_l \mathrm{Tr}(BU_l) = \mathrm{Tr}(BA)$, thus the first claim.

Let $A \in L^1(H)$ with polar decomposition $A = U|A|$. Then

$$\begin{aligned}
|\mathrm{Tr}(A)| &= |\sum_e \langle U|A|e, e\rangle| = |\sum_e \langle |A|^{1/2}e, |A|^{1/2}U^*e\rangle| \\
&\leq \sum_e \| \, |A|^{1/2}e\| \, \| \, |A|^{1/2}U^*e\| \\
&\leq \left(\sum_{e\in E} \| \, |A|^{1/2}e\|^2\right)^{1/2} \left(\sum_{e\in E} \| \, |A|^{1/2}U^*e\|^2\right)^{1/2}.
\end{aligned}$$

Now, as in the proof of (iii), $\sum_{e\in E} \| \, |A|^{1/2}e\|^2 = \mathrm{Tr}|A|$ and $\sum_{e\in E} \| \, |A|^{1/2}U^*e\|^2 \leq \mathrm{Tr}|A|$. This proves $|\mathrm{Tr}(A)| \leq \|A\|_1$ for all $A \in L^1(H)$.

If now $A \in L^1(H), B \in B(H)$ then $|\mathrm{Tr}(AB)| \leq \|AB\|_1 \leq \|A\|_1\|B\|$ by (iv). ■

Let H be a Hilbert space. By Theorem B.137(xii), we have linear maps $\alpha : L^1(H) \to B(H)^*, A \mapsto \mathrm{Tr}(A\bullet)$ and $\beta : B(H) \to L^1(H)^*, A \mapsto \mathrm{Tr}(A\bullet)$ such that $\|\alpha(A)\| \leq \|A\|_1$ and $\|\beta(A)\| \leq \|A\|$.

B.138 THEOREM *Let H be a Hilbert space. Then*

(i) $\beta : (B(H), \|\cdot\|) \to (L^1(H), \|\cdot\|_1)^*$ *is an isometric bijection.*

(ii) $\alpha_K : (L^1(H), \|\cdot\|_1) \to (K(H), \|\cdot\|)^*$ *is an isometric bijection.*

(iii) The map $\alpha : (L^1(H), \|\cdot\|_1) \to (B(H), \|\cdot\|)^*$ is isometric, but it is not surjective if H is infinite-dimensional.

Proof. (i) [110, Theorem 4.2.3].

(ii) [110, Theorem 4.2.1].

(iii) Since $\|\alpha(A)\| \le \|A\|_1$ for all $A \in L^1(H)$ and $\|\alpha(A) \restriction K(H)\| = \|A\|_1$ by (ii), it is clear that also $\|\alpha(A)\| = \|A\|_1$, thus α is isometric.

For each $x, y \in H$ we have $x \otimes y \in K(H)$, and $\mathrm{Tr}(A(x \otimes y)) = \mathrm{Tr}((Ax) \otimes y) = \langle Ax, y\rangle$. Thus if $\alpha(A) = \mathrm{Tr}(A\cdot)$ vanishes on the compact operators then $A = 0$, thus $\alpha(A) = 0$.

However, if H is infinite-dimensional, $K(H) \subsetneq B(H)$ is a proper closed ideal, and the quotient space $C(H) = B(H)/K(H)$ (the Calkin algebra) is non-trivial. Thus it admits a bounded non-zero functional ψ. If $p : B(H) \to C(H)$ is the quotient map then $\psi \circ p$ is a non-zero norm-continuous functional on $B(H)$ that vanishes on $K(H)$. Such a functional cannot be of the form $\mathrm{Tr}(A\cdot)$ with $A \in L^1(H)$, proving that α is not surjective. ■

B.139 Remark 1. The results (i), (ii), (iii) are non-commutative analogues of $\ell^\infty(S, \mathbb{F}) \cong \ell^1(S, \mathbb{F})^*$, $\ell^1(S, \mathbb{F}) \cong c_0(S, \mathbb{F})^*$, and Theorem 4.19(v), respectively.

2. One can show that $\alpha(L^1(H)) \subsetneq B(H)^*$ consists precisely of the linear functionals that are not only norm-continuous but also ultra-weakly continuous (or, equivalently, normal). Cf. e.g. [110]. This is analogous to Proposition B.21.

3. The analogy between $L^p(H)$ and $\ell^p(S, \mathbb{F})$ extends to $p \notin \{1, 2, \infty\}$: Each space $L^p(H) = \{A \in B(H) \mid \|A\|_p < \infty\}$, the '$p$-th Schatten class', is a two-sided ideal in $B(H)$ and in fact $L^p(H) \subseteq K(H)$ for all p. If $1 \le p \le q < \infty$, it is not hard to show that $\|A\| := \|A\|_\infty \le \|A\|_q \le \|A\|_p$, thus $L^q(H) \subseteq L^p(H) \subseteq K(H)$. For all $1 < p < \infty$ one has $L^p(H) = \overline{F(H)}^{\|\cdot\|_p}$, the Hölder type inequality $\|AB\|_1 \le \|A\|_p \|B\|_q$ for $A \in L^p(H), B \in L^q(H)$ and in fact the duality $L^p(H)^* \cong L^q(H)$. For all this see [151]. □

We close with a harder result:

B.140 Theorem (Grothendieck 1956, Lidskii 1959) *If H is a Hilbert space and $A \in L^1(H)$ then $\mathrm{Tr} A = \sum_{\lambda \in \sigma_{\mathrm{p}}(A)} n_\lambda \lambda$, where $n_\lambda = \dim L_\lambda(A)$ is the algebraic multiplicity of λ, proven finite in Proposition 14.18.*

The theorem is obvious if A is normal, thus diagonalizable by Theorem 14.12. For finite matrices, there are two standard approaches, using the Jordan normal form and the characteristic polynomial, respectively. Both can be adapted to infinite dimensions, in the first case using the results of Section 14.4: One writes H as a direct sum $H = H_0 \oplus_{\lambda \in \sigma(A)\setminus\{0\}} H_\lambda$ of invariant subspaces, where $A \restriction H_0$ is quasi-nilpotent and the H_λ are finite-dimensional and $A - \lambda \mathbf{1} \restriction H_\lambda$ nilpotent. This already implies $\mathrm{Tr} A = \mathrm{Tr}(A \restriction H_0) + \sum_{\lambda \in \sigma_{\mathrm{p}}(A)} n_\lambda \lambda$, and it remains to show that the trace of a quasi-nilpotent trace class operator vanishes. For this see e.g. [59, p. 101-103], [103, Lemma 16.32] or [94, Sect. 30.3]. For the adaptation of the characteristic polynomial approach using determinants in infinite dimensions, see [62, 151, 152].

B.12 More on numerical ranges

B.12.1 The numerical range $W(A)$ of a Hilbert space operator

We have already proven some results concerning the numerical range $W(A)$ and radius $|||A|||$ of a Hilbert space operator A, cf. Definition 11.32. See Propositions 11.34 and 11.22, Exercises

11.36 and 13.15, and Proposition 13.69(iii). These results show that these quantities are of some interest, and here we prove some further results about $W(A)$.

To begin with, $W(A)$ need not be closed:

B.141 Exercise (i) Give an example of a bounded Hilbert space operator A such that there is no $x \in H$, $\|x\| = 1$ such that $|\langle Ax, x\rangle| = \|\|A\|\|$.

(ii) Prove that despite (i) there always exists a sequence $\{x_n\}$ with $\|x_n\| = 1$ such that $\langle Ax_n, x_n\rangle \to \lambda$ where $\lambda \in \mathbb{C}$, $|\lambda| = \|\|A\|\|$.

B.142 Exercise Let H be a Hilbert space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $A \in B(H)$. Prove:

(i) $W(\alpha A + \beta \mathbf{1}) = \alpha W(A) + \beta$ for all $\alpha, \beta \in \mathbb{F}$.

(ii) $W(A^*) = W(A)^*$.

(iii) $W(UAU^*) = W(A)$ for every unitary $U : H \to H'$. (NB: In general $W(BAB^{-1}) \neq W(A)$ for invertible B!)

(iv) If $\mathbb{F} = \mathbb{C}$ then $W(A) = \{\lambda\}$ if and only if $A = \lambda\mathbf{1}$.

(v) If $\mathbb{F} = \mathbb{C}$ then $W(A)$ is contained in a line segment $[\gamma\delta] = \{t\gamma + (1-t)\delta \mid t \in [0, 1]\}$ if and only there are $\alpha, \beta \in \mathbb{C}$ such that $\alpha A + \beta\mathbf{1}$ is self-adjoint.

B.143 Theorem (Toeplitz-Hausdorff (1918/9)) [139] *Let H be a complex Hilbert space and $A \in B(H)$. Then the numerical range $W(A) \subseteq \mathbb{C}$ is convex.*

Proof. If $x, y \in H$ with $\|x\| = \|y\| = 1$ and $t \in [0, 1]$ we must show that $t\langle Ax, x\rangle + (1-t)\langle Ay, y\rangle \in W(A)$. Let P be the orthogonal projection onto $K = \mathbb{C}x + \mathbb{C}y \subseteq H$ and $A_K = PAP$ considered as element of $B(K)$. Then $t\langle Ax, x\rangle + (1-t)\langle Ay, y\rangle = t\langle A_K x, x\rangle + (1-t)\langle A_K y, y\rangle$. Thus if we prove that the r.h.s. is in $W(A_K)$, it is of the form $\langle A_K z, z\rangle = \langle Az, z\rangle$ for some $z \in K$, thus also the l.h.s. is in $W(A)$. We have therefore reduced the claim to proving the special case $H = \mathbb{C}^2$ (since there is nothing to prove if K is one-dimensional). There are many ways of doing this, most of which are quite computational. We will give the nice argument from [35].

If $x \in H$ is a unit vector then $P = x \otimes x$ (notation of Exercise 12.36) is an orthogonal projection of rank one, and every such projection arises in this way. Now for all $A \in B(H)$ we have $\langle Ax, x\rangle = \mathrm{Tr}(PA)$. (This is easily checked by computing the trace using an ONB containing the vector x.) Since $\mathrm{Tr}(P)$ equals the rank of P, denoting the set of rank one orthogonal projections by $\mathcal{P}_1$ we have $W(A) = \{\mathrm{Tr}(PA) \mid P \in \mathcal{P}_1\}$.

Now specialize to $H = \mathbb{C}^2$. For $(x, y, z) \in \mathbb{R}^3$ it is clear that $M(x, y, z) = \frac{1}{2}\begin{pmatrix} 1+z & x+iy \\ x-iy & 1-z \end{pmatrix}$ is self-adjoint with trace one, and every such matrix is of this form. A trivial computation gives

$$M(x,y,z)^2 = \frac{1}{2}\begin{pmatrix} \frac{1+x^2+y^2+z^2}{2} + z & x+iy \\ x-iy & \frac{1+x^2+y^2+z^2}{2} - z \end{pmatrix},$$

implying that $M(x, y, z)$ is an idempotent, thus a rank one orthogonal projection, if and only if $x^2 + y^2 + z^2 = 1$. Thus the map $(x, y, z) \mapsto M(x, y, z)$ restricts to a bijection $S^2 \to \mathcal{P}_1$. Now

$$\mathrm{Tr}(M(x,y,z)A) = \frac{\mathrm{Tr}(A)}{2} + \frac{1}{2}\mathrm{Tr}\left(\begin{pmatrix} z & x+iy \\ x-iy & z \end{pmatrix} A\right).$$

[139] Felix Hausdorff (1868-1942), German mathematician. Towering figure in the history of general topology and related areas like measure theory and functional analysis. Driven to suicide by the Nazis.

Since the second summand depends $\mathbb{R}$-linearly on (x, y, z) (for any fixed A), we are done once we show that the image of $S^2 \subseteq \mathbb{R}^3$ under every linear map $\alpha : \mathbb{R}^3 \to \mathbb{R}^2$ is convex. For dimensional reasons, α has a non-trivial kernel K. Now the image of S^2 under the orthogonal projection $p : \mathbb{R}^3 \to K^\perp \subset \mathbb{R}^3$ is a ball, thus convex, so that also $\alpha(S^2) = \alpha(p(S^2))$ is convex. ■

There are interesting other proofs that do not proceed by reduction to two dimensions, cf. e.g. [113, Exercise 8.9].

In view of Exercise 13.15 and Theorem B.143 it is clear that for all bounded Hilbert space operators we have $\mathrm{conv}(\sigma(A)) \subseteq \overline{W(A)}$. (Recall Definition B.66.) Already Exercise 11.36 shows that this need not be an equality. Again, normal operators are behaved nicely:

B.144 Exercise Let H be a Hilbert space and $A \in B(H)$ normal. Use the Spectral Theorem 18.4 to prove $\mathrm{conv}(\sigma(A)) = \overline{W(A)}$. (Note that $\mathrm{conv}(\sigma(A))$ is closed by Corollary B.69.)

For an alternative proof avoiding the spectral theorem (but still using continuous functional calculus) see [118, Exercise E4.4.5].

While $\mathrm{conv}(\sigma(A)) = \overline{W(A)}$ does not hold for all operators, there is a slightly more involved general fact, somewhat related to the numerical identity from Exercise 17.15(iv):

B.145 Theorem (S. Hildebrandt 1966) [140] *Let H be a Hilbert space and $A \in B(H)$. Then*

$$\mathrm{conv}(\sigma(A)) = \bigcap_{B \in \mathrm{Inv}\, B(H)} \overline{W(BAB^{-1})}.$$

Proof. For each invertible B we have $\sigma(A) = \sigma(BAB^{-1}) \subseteq \overline{W(BAB^{-1})}$. With the convexity of the numerical ranges we have

$$\mathrm{conv}(\sigma(A)) \subseteq \bigcap_{B \in \mathrm{Inv}\, B(H)} \overline{W(BAB^{-1})}. \tag{B.17}$$

Assume equality does not hold, thus there exists λ such that $\lambda \in \overline{W(BAB^{-1})}$ for all $B \in \mathrm{Inv}\, B(H)$ but $\lambda \notin \mathrm{conv}(\sigma(A))$. Since $\mathrm{conv}(\sigma(A))$ is a compact convex set, it is not hard to find an open disc D such that $\mathrm{conv}(\sigma(A)) \subseteq D$ while $\lambda \notin D$. Using $\sigma(\alpha A + \beta) = \alpha\sigma(A) + \beta$ and $W(\alpha A + \beta) = \alpha W(A) + \beta$, we can reduce to the situation where D is the open unit disc $U = B(0, 1)$ around zero. Then $\sigma(A) \subseteq \mathrm{conv}(\sigma(A)) \subseteq U$, so that $r(A) < 1$. Now Exercise 17.15(i)-(iii) provides $B \in \mathrm{Inv}\, B(H)$ such that $\|BAB^{-1}\| < 1$. Then also $\overline{W(BAB^{-1})} \subseteq U$, which contradicts the facts $\lambda \in \overline{W(BAB^{-1})}$ and $|\lambda| \geq 1$. This contradiction proves that the inclusion in (B.17) is an equality. ■

For much more on $W(A)$ see the book [64].

B.12.2 Numerical range in Banach algebras

In this section, which strongly leans on [18], we exlusively consider complex unital Banach algebras satisfying $\|\mathbf{1}\| = 1$.

In the discussion of commutative Banach and C^*-algebras, a large role was played by characters, i.e. non-zero algebra homomorphisms $\mathcal{A} \to \mathbb{C}$. A non-commutative algebra may fail to have any characters. Of course $\mathcal{A}^*$ separates the points of $\mathcal{A}$, but this set is too big. The following natural subset of $\mathcal{A}^*$ will turn out useful:

[140] Stefan Oscar Walter Hildebrandt (1936-2015). German mathematician who mostly worked on variational calculus.

B.146 Definition *If $\mathcal{A}$ is a unital complex Banach algebra, $\varphi \in \mathcal{A}^*$ is called a state if $\varphi(\mathbf{1}) = \|\varphi\| = 1$. The set of states of $\mathcal{A}$ is denoted $S(\mathcal{A})$.*

B.147 Proposition *Let $\mathcal{A}$ be a unital normed algebra over $\mathbb{C}$. For each $a \in \mathcal{A}$ the subsets*

$$\begin{aligned} V_1(a) &= \bigcap_{z\in\mathbb{C}} \overline{B}(z, \|a - z\mathbf{1}\|) = \{\lambda \in \mathbb{C} \mid |\lambda - z| \le \|a - z\mathbf{1}\| \ \forall z \in \mathbb{C}\}, \\ V_2(a) &= \{\varphi(a) \mid \varphi \in \mathcal{A}^*, \ \|\varphi\| = \varphi(\mathbf{1}) = 1\} \end{aligned}$$

of $\mathbb{C}$ coincide. The resulting set $V(a)$ is called the (algebraic) numerical range and $R(a) = \sup_{\lambda\in V(\lambda)} |a|$ the (algebraic) numerical radius of a.

Proof. Let $a \in \mathcal{A}$ and $\varphi \in \mathcal{A}^*$ with $\varphi(\mathbf{1}) = \|\varphi\| = 1$. Then $|\varphi(a) - z| = |\varphi(a - z\mathbf{1})| \le \|a - z\mathbf{1}\|$ holds for each $z \in \mathbb{C}$. Thus $\varphi(a) \in V_1(a)$, proving $V_2(a) \subseteq V_1(a)$.

If $a = c\mathbf{1}$ then $V_2(a) = \{c\}$, and with $z = c$ the inequality $|\lambda - z| \le \|a - z\mathbf{1}\|$ becomes $|\lambda - z| \le 0$, so that also $V_1(a) = \{c\}$. Thus $V_1(a) = V_2(a)$ for $a \in \mathbb{C}\mathbf{1}$. Now assume $a \notin \mathbb{C}\mathbf{1}$ and let $\lambda \in V_1(a)$. Put $W = \mathbb{C}\mathbf{1} + \mathbb{C}a$ and define $\varphi_0 \in W^*$ by $c\mathbf{1} + da \mapsto c + d\lambda$. Then $\varphi_0(\mathbf{1}) = 1$ and $\varphi_0(a) = \lambda$, and

$$\|\varphi_0\| = \sup_{(c,d)\neq(0,0)} \frac{|\varphi_0(c\mathbf{1} + da)|}{\|c\mathbf{1} + da\|} = \sup_{(c,d)\neq(0,0)} \frac{|c + d\lambda|}{\|c\mathbf{1} + da\|}.$$

For $d = 0$ the fraction on the r.h.s. is 1, so that with $\lambda \in V_1(a)$ we have

$$\|\varphi_0\| = \max\left(1, \sup_{z\in\mathbb{C}} \frac{|z + \lambda|}{\|z\mathbf{1} + a\|}\right) = 1.$$

Now by Hahn-Banach there exists an extension φ of φ_0 to V satisfying $\|\varphi\| = \|\varphi_0\| = 1$. Now $\varphi(a) = \lambda$, proving $V_1(a) \subseteq V_2(a)$. ■

B.148 Proposition *Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$. Then*

(i) $V(a) \subseteq \mathbb{C}$ is closed and convex,

(ii) $\sigma(a) \subseteq V(a)$ and $r(a) \le R(a) \le \|a\|$. In particular, $V(a) \neq \emptyset$.

(iii) $V(\alpha a + \beta\mathbf{1}) = \alpha V(a) + \beta$ for all $\alpha, \beta \in \mathbb{C}$.

Proof. (i) Since $\overline{B}(z, \|a - z\mathbf{1}\|)$ is closed and convex for each $z \in \mathbb{C}$ and these properties pass to arbitrary intersections, closedness and convexity of $V(a)$ are obvious from $V(a) = V_1(a)$. And $R(a) \le \|a\|$ follows from $V_1(a) \subseteq \overline{B}(0, \|a - 0\mathbf{1}\|)$, the $z = 0$ term in the intersection.

(ii) Assume $\lambda \notin V(a)$. This means that there exists $z \in \mathbb{C}$ such that $|\lambda - z| > \|a - z\mathbf{1}\|$. Then $\left\|\frac{a-z\mathbf{1}}{\lambda-z}\right\| < 1$, so that $\mathbf{1} - \frac{a-z\mathbf{1}}{\lambda-z} \in \mathrm{Inv}\,\mathcal{A}$ by Lemma 13.19. Thus $-(a - \lambda\mathbf{1}) = (\lambda - z)\mathbf{1} - (a - z\mathbf{1}) \in \mathrm{Inv}\,\mathcal{A}$, so that $\lambda \notin \sigma(a)$. Thus $\sigma(a) \subseteq V(a)$, implying also $r(a) \le R(a)$. Now $\sigma(a) \neq \emptyset$ implies $V(a) \neq \emptyset$. (iii) is rather obvious for $V_2(a)$ since the φ involved satisfy $\varphi(\mathbf{1}) = 1$. ■

B.149 Remark Given the similarity of the properties of $V(a)$ to those of the (spatial) numerical range $W(A)$ of a Hilbert space operator (except for the closedness of $V(a)$), it is natural to ask how the two definitions are related if $\mathcal{A} = B(H)$. It is easy to see that $\overline{W(A)} \subseteq V(A)$. In fact, this always is an equality. We postpone the proof to Theorem B.168 since it requires some preparations and we prefer to stick to the general Banach algebra situation for now. □

B.150 LEMMA *Let $\mathcal{A}$ be a unital complex Banach algebra and $a \in \mathcal{A}$. Then*

$$\inf\{\operatorname{Re}\lambda \mid \lambda \in V(a)\} \le \inf\{\|ab\| \mid b \in \mathcal{A}, \|b\| = 1\}.$$

(Note that the r.h.s. also appeared in Exercise 13.34.)

Proof. For $b \in \mathcal{A}$ with $\|b\| = 1$ put $V(a,b) = \bigcap_{z\in\mathbb{C}} \overline{B}(z, \|(a-z\mathbf{1})b\|)$. With $\|(a-z\mathbf{1})b\| \le \|a-z\mathbf{1}\|$ and the definition of $V_1(a)$ we have $V(a,b) \subseteq V_1(a)$. Since clearly $V(a,\mathbf{1}) = V_1(a)$, we have $\bigcup_{b\in\mathcal{A}_{=1}} V(a,b) = V_1(a) = V(a)$. For all $b \in \mathcal{A}, \|b\| = 1$ we have $V(a,b) \subseteq \overline{B}(0, \|ab\|)$, implying

$$\inf\{\operatorname{Re}\lambda \mid \lambda \in V(a)\} \le \inf\{\operatorname{Re}\lambda \mid \lambda \in V(a,b)\} \le \|ab\|.$$

Now take the infimum over $b \in \mathcal{A}$ with $\|b\| = 1$. ■

B.151 THEOREM *Let $\mathcal{A}$ be a unital complex Banach algebra and $a \in \mathcal{A}$. Then*

$$\begin{aligned} \sup\{\operatorname{Re}\lambda \mid \lambda \in V(a)\} &= \inf_{t>0} \frac{\|\mathbf{1}+ta\|-1}{t} \\ &= \lim_{t\searrow 0} \frac{\|\mathbf{1}+ta\|-1}{t} \\ &= \sup_{t>0} \frac{\log\|e^{ta}\|}{t} \\ &= \lim_{t\searrow 0} \frac{\log\|e^{ta}\|}{t}. \end{aligned} \tag{B.18}$$

Proof. Fix $a \in \mathcal{A}$, and pick $\varphi \in S(\mathcal{A})$, $t > 0$. Then

$$1 + t\operatorname{Re}\varphi(a) \le |1 + t\operatorname{Re}\varphi(a)| = |\operatorname{Re}(\varphi(\mathbf{1}+ta))| \le |\varphi(\mathbf{1}+ta)| \le \|\mathbf{1}+ta\|,$$

implying $\operatorname{Re}\varphi(a) \le \frac{\|\mathbf{1}+ta\|-1}{t}$. Since this holds for all $t > 0$ and $\varphi \in S(\mathcal{A})$, we conclude

$$\sup\{\operatorname{Re}\lambda \mid \lambda \in V(a)\} = \sup_{\varphi\in S(\mathcal{A})} \operatorname{Re}\varphi(a) \le \inf_{t>0} \frac{\|\mathbf{1}+ta\|-1}{t}. \tag{B.19}$$

Now put $\mu = \sup\{\operatorname{Re}\lambda \mid \lambda \in V(a)\}$ and let $t > 0$. Then

$$1 - t\mu = 1 - t\sup\{\operatorname{Re}\lambda \mid \lambda \in V(a)\} = \inf_{\varphi\in S(\mathcal{A})} \operatorname{Re}\varphi(1-ta) \le \inf\{\|(1-ta)b\| \mid b \in \mathcal{A}, \|b\| = 1\},$$

where the final inequality comes from Lemma B.150. From this one readily deduces

$$(1-t\mu)\|b\| \le \|(1-ta)b\| \quad \forall b \in \mathcal{A},$$

which for $b = 1 + ta$ gives

$$(1-t\mu)\|1+ta\| \le \|(1-ta)(1+ta)\| = \|1+t^2a^2\| \le 1 + t^2\|a^2\|.$$

Assuming $t\mu < 1$ this gives $\|1+ta\| \le (1-t\mu)^{-1}(1+t^2\|a^2\|)$, implying

$$\frac{\|1+ta\|-1}{t} \le \frac{\mu + t\|a^2\|}{1-t\mu}$$

and therefore

$$\limsup_{t\searrow 0} \frac{\|1+ta\|-1}{t} \le \mu.$$

Combining this with (B.19) we have

$$\limsup_{t\searrow 0}\frac{\|1+ta\|-1}{t}\le\sup\{\mathrm{Re}\,\lambda\mid\lambda\in V(a)\}\le\inf_{t>0}\frac{\|\mathbf{1}+ta\|-1}{t}\le\liminf_{t\searrow 0}\frac{\|\mathbf{1}+ta\|-1}{t},$$

which implies the equality of the first three expressions in (B.18).
****

■

B.152 Definition *Let $\mathcal{A}$ be a unital complex Banach algebra. Then $a\in\mathcal{A}$ then a is called*

- *dissipative if* $\mathrm{Re}\,\lambda\le 0\ \forall\lambda\in V(a)$.
- *hermitian if $V(a)\subseteq\mathbb{R}$. We put $H(\mathcal{A})=\{a\in\mathcal{A}\mid a \text{ hermitian}\}$.*
- *normal if there are commuting $b,c\in H(\mathcal{A})$ with $a=b+ic$.*

B.153 Corollary *If $\mathcal{A}$ is a unital Banach algebra and $a\in\mathcal{A}$ then*

(i) $\|e^{ita}\|=1\ \forall t\in\mathbb{R}$ *if and only if a is hermitian.*

(ii) $\|e^{ta}\|\le 1\ \forall t\in\mathbb{R}$ *if and only if a is dissipative.*

Proof. Both statements are immediate consequences of (B.18). ■

B.154 Theorem (Bohnenblust & Karlin 1955) [141] *Let $\mathcal{A}$ be a unital complex Banach algebra. Then*

$$\frac{\|a\|}{e}\le R(a)\le\|a\|\quad\forall a\in\mathcal{A}.$$

(Here $e=\exp(1)=2.718\ldots$.) In particular $a=0\Leftrightarrow V(a)=\{0\}$.

Proof. The inequality $R(a)\le\|a\|$ has already been proven. Let $a\ne 0$. Rescaling if necessary we may assume $R(a)=1$, thus $|\lambda|\le 1$ for all $\lambda\in V(a)$. Then (B.18) implies

$$\log\|e^a\|\le\sup\{\mathrm{Re}\,\lambda\,|\,\lambda\in V(a)\}\le 1,$$

which remains valid if we replace a by ca with $|c|=1$. Thus $\|e^{za}\|\le e\ \forall z\in S^1$.

Whenever a power series $f(z)=\sum_{n=0}^\infty c_nz^n$ has convergence radius $R>r(a)$ we can define $f(e^{2\pi it}a)$ and have[142]

$$\int_0^1 f(e^{2\pi it}a)e^{-2\pi imt}\,dt=\int_0^1\Big(\sum_{n=0}^\infty c_na^ne^{2\pi int}\Big)e^{-2\pi imt}dt=\sum_{n=0}^\infty c_na^n\int_0^1 e^{2\pi i(n-m)t}dt=c_ma^m. \tag{B.20}$$

for each $m\in\mathbb{N}_0$, where the interchange of integration and summation was justified by the uniform convergence of the series on bounded sets, and we used $\int_0^1 e^{2\pi i(n-m)t}dt=\delta_{n,m}$. Thus $|c_m|\|a^m\|\le\int_0^1\|f(e^{2\pi it}a)\|dt$. Applying this to $f=\exp$ and $m=1$ (thus $c_m=1$) we obtain $\|a\|\le\int_0^1\|\exp(e^{2\pi it}a)\|\,dt\le e=eR(a)$.

Now $V(a)=\{0\}\Rightarrow R(a)=0\Rightarrow\|a\|=0\Rightarrow a=0\Rightarrow V(a)=\emptyset$. ■

[141] Samuel Karlin (1924-2007). Polish-born American mathematician. After his work in pure analysis he made many contributions to mathematical economy and biology.

[142] If V is a Banach space and $f:[a,b]\to V$ a continuous function, the Riemann integral $\int_a^b f(t)dt$ is defined (and existence proven using the uniform continuity of f) as for $\mathbb{R}$-valued functions. One then has $\|\int_a^b f(t)dt\|\le\int_a^b\|f(t)\|dt$.

The surprising factor e cannot be improved without further assumptions. (Recall that for the numerical radius of an operator on a complex Hilbert space we proved the slightly stronger $|||A||| \geq \frac{\|A\|}{2}$, which we showed to be optimal.)

We mention without proofs (see e.g. [19]) two more results:

B.155 Theorem *If $\mathcal{A}$ is a unital complex Banach algebra and $a \in \mathcal{A}$ is normal then $V(a) = \mathrm{conv}(\sigma(a))$, thus $R(a) = r(a)$. (Compare Exercise B.144.) For hermitian a even $R(a) = \|a\|$.*

If $a \in H(\mathcal{A}) \cap iH(\mathcal{A})$ then $V(a) \subseteq \mathbb{R} \cap i\mathbb{R} = \{0\}$, which implies $a = 0$. Thus $H(\mathcal{A}) \cap iH(\mathcal{A}) = \{0\}$. It is natural to ask whether $H(\mathcal{A}) + iH(\mathcal{A}) = \mathcal{A}$. For C^*-algebras this is true since one can prove $H(\mathcal{A}) = \mathcal{A}_{sa}$, cf. Proposition B.161. This actually characterizes C^*-algebras:

B.156 Theorem (Vidav (1956), Palmer (1968)) *Let $\mathcal{A}$ be a unital complex Banach algebra. If $\mathcal{A} = H(\mathcal{A}) + iH(\mathcal{A})$ then $(b + ic)^* = b - ic$ for $a, b \in H(\mathcal{A})$ defines a $*$-operation and $\|a^*a\| = \|a\|^2 \; \forall a \in \mathcal{A}$, thus $\mathcal{A}$ is a C^*-algebra.*

For proofs of these results (and many more), see [19].

B.12.3 Positive functionals on and numerical range for C^*-algebras

B.157 Definition *Let $\mathcal{A}$ be a complex C^*-algebra. A linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ is called positive if $\varphi(a) \geq 0$ for all positive a, i.e. $a \in \mathcal{A}_+$. The set of positive functionals on $\mathcal{A}$ is denoted $\mathcal{A}^+$.*

Positive functionals have remarkable properties:

B.158 Proposition *Let $\mathcal{A}$ be a C^*-algebra and φ a positive functional. Then*

(i) φ is bounded. Thus $\mathcal{A}^+ \subseteq \mathcal{A}^*$.

(ii) $[a, b] = \varphi(ab^*)$ defines a sequilinear form on $\mathcal{A}$ that is self-adjoint and positive.

(iii) For all $a, b \in \mathcal{A}$ we have $|\varphi(ab^*)| \leq \big(\varphi(aa^*)\varphi(bb^*)\big)^{1/2}$.

(iv) If $\mathcal{A}$ is unital, $\varphi(a^*) = \overline{\varphi(a)}$ for all $a \in \mathcal{A}$.

Proof. (i) Let φ be positive. Assume that φ is unbounded on the positive elements of $\mathcal{A}$. Then there is a sequence $\{a_n\} \subset \mathcal{A}_+$ with $\|a_n\| = 1$ and $\varphi(a_n) \geq 2^n$ for each $n \in \mathbb{N}$. Define $a = \sum_{n=1}^{\infty} 2^{-n} a_n$, which clearly converges with a positive and $\|a\| \leq \sum_{n=1}^{\infty} 2^{-n} = 1$. Since $\sum_{n=1}^{N} 2^{-n} \leq a$ for all N, we have $a - \sum_{n=1}^{N} 2^{-n} a_n \geq 0$, thus

$$\varphi(a) \geq \varphi\Big(\sum_{n=1}^{N} 2^{-n} a_n\Big) = \sum_{n=1}^{N} 2^{-n} \varphi(a_n) \geq \sum_{n=1}^{N} 2^{-n} \cdot 2^n = N$$

holds for all N. But this contradicts the fact that $\varphi(a) \in [0, +\infty)$. Thus there exists C such that $\varphi(a) \leq C\|a\|$ for all $a \in \mathcal{A}_+$.

For every $a \in \mathcal{A}$ we have $a = b + ic$ with $b, c \in \mathcal{A}_{sa}$, where $\|b\| \leq \|a\|$ and $\|c\| \leq \|a\|$. And by Exercise 17.5, $b = b_+ - b_-$ with $b_\pm \in \mathcal{A}_+$ and $\|b_\pm\| \leq \|b\| \leq \|a\|$ and similarly for c. Now

$$|\varphi(a)| = |\varphi(b_+) - \varphi(b_-) + i\varphi(c_+) - i\varphi(c_-)| \leq \varphi(b_+) + \varphi(b_-) + \varphi(c_+) + \varphi(c_-) \leq 4C\|a\|.$$

Thus φ is bounded with $\|\varphi\| \leq 4C < \infty$.

(ii) It is obvious that $[\cdot,\cdot]$ is a sesquilinear form on $\mathcal{A}$. For all $a \in \mathcal{A}$ we have $aa^* \in \mathcal{A}_+$, thus $[a,a] = \varphi(aa^*) \geq 0$, so that $[\cdot,\cdot]$ is positive in the sense $[a,a] \geq 0\ \forall a$. A fortiori, $[a,a] \in \mathbb{R}$ for all a, so that $[\cdot,\cdot]$ is self-adjoint by Lemma 11.17(i), i.e. $\overline{[x,y]} = [y,x]$.

(iii) Every positive sesquilinear form satisfies a Cauchy-Schwarz inequality $|[a,b]| \leq \big([a,a][b,b]\big)^{1/2}$ with the same proof as for (5.1) since non-degeneracy $[x,x] = 0 \Rightarrow x = 0$ was not needed there. Expressing the inequality in terms of φ, the claim follows.

(iv) Since $[\cdot,\cdot]$ is self-adjoint by (i), we have $\varphi(ab^*) = [a,b] = \overline{[b,a]} = \overline{\varphi(ba^*)}$. Putting $b = \mathbf{1}$, the claim follows. ■

(Statement (iii) also holds for non-unital algebras, but the proof uses approximate units.) Be sure not to confuse $\mathcal{A}_+ \subseteq \mathcal{A}$ and $\mathcal{A}^+ \subseteq \mathcal{A}^*$!

B.159 Proposition (Bohnenblust & Karlin 1955) *Let $\mathcal{A}$ be a unital complex C^*-algebra. Then a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ is positive if and only if it is bounded and satisfies $\varphi(\mathbf{1}) = \|\varphi\|$.*

Proof. $\Rightarrow$ If φ is positive, it is bounded by Proposition B.158(i), and $\varphi(\mathbf{1}) > 0$ since $\mathbf{1} \in \mathcal{A}_+$. If $a \in \mathcal{A}_{\leq 1}$, with Proposition B.158(iii) we have

$$|\varphi(a)|^2 \leq \varphi(aa^*)\varphi(\mathbf{1}) \leq \|aa^*\|\|\varphi\|\varphi(\mathbf{1}) \leq \|\varphi\|\varphi(\mathbf{1}).$$

Since for each $\varepsilon > 0$ we can find $a \in \mathcal{A}_{\leq 1}$ with $|\varphi(a)| > \|\varphi\| - \varepsilon$, we have $(\|\varphi\| - \varepsilon)^2 \leq \|\varphi\|\varphi(\mathbf{1})$. Taking $\varepsilon \to 0$ gives $\|\varphi\| \leq \varphi(\mathbf{1})$ (if $\varphi \neq 0$, but the result also holds for $\varphi = 0$). On the other hand, with $\|\mathbf{1}\| = 1$ we have $\|\varphi\| \geq \varphi(\mathbf{1})$. Combining the inequalities we have $\|\varphi\| = \varphi(\mathbf{1})$.

$\Leftarrow$ Replacing φ by $\|\varphi\|^{-1}\varphi$, we may assume $\|\varphi\| = 1$. Let $a \in \mathcal{A}_{sa}$, $\|a\| \leq 1$. Write $\varphi(a) = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$. Then for each $s \in \mathbb{R}$,

$$\|a - is\mathbf{1}\|^2 = \|(a - is\mathbf{1})^*(a - is\mathbf{1})\| = \|(a + is\mathbf{1})(a - is\mathbf{1})\| = \|a^2 + s^2\mathbf{1}\| \leq 1 + s^2,$$

where we used the C^*-identity. Thus

$$\alpha^2 + (\beta - s)^2 = |\alpha + i\beta - is|^2 = |\varphi(a) - is|^2 = |\varphi(a - is\mathbf{1})|^2 \leq 1 + s^2,$$

implying $\alpha^2 + \beta^2 - 2s\beta \leq 1$. Since this must hold for all $s \in \mathbb{R}$, we have $\beta = 0$, thus $\varphi(a) \in \mathbb{R}$. Thus for $a = a^*$ we have $V(a) \subseteq \mathbb{R}$.

Now let $a \in \mathcal{A}_+$, $\|a\| \leq 1$. Then $0 \leq a \leq \mathbf{1}$, thus $\mathbf{1} - a$ is positive with $\|\mathbf{1} - a\| \leq 1$ (compare Exercise 16.24), so that $|\varphi(\mathbf{1} - a)| \leq 1$. Combined with $\varphi(\mathbf{1} - a) \in \mathbb{R}$ from the above, this gives $1 - \varphi(a) = \varphi(\mathbf{1} - a) \leq 1$, implying $\varphi(a) \geq 0$. Thus φ is positive. ■

B.160 Remark If $\mathcal{A}$ is a unital C^*-algebra, by states on $\mathcal{A}$ one usually means the $\varphi \in \mathcal{A}^*$ that are positive and normalized, i.e. $\varphi(\mathbf{1}) = 1$. In view of the above result, this is entirely consistent with the Banach algebraic Definition B.146 of states. □

Now we are in a position to use $V(a)$ to characterize the elements of a C^*-algebra that satisfy $a = 0, a = a^*$ and $a \geq 0$, respectively, in analogy to the results for $A \in B(H)$ in terms of $W(A)$.

B.161 Proposition *Let $\mathcal{A}$ be a unital complex C^*-algebra and $a \in \mathcal{A}$. Then*

(i) $a = 0 \ \Leftrightarrow\ \varphi(a) = 0$ *for all* $\varphi \in \mathcal{A}^+ \ \Leftrightarrow\ V(a) = \{0\}$.

(ii) $a = a^* \ \Leftrightarrow\ \varphi(a) \in \mathbb{R}$ *for all* $\varphi \in \mathcal{A}^+ \ \Leftrightarrow\ V(a) \subseteq \mathbb{R}$.

(iii) $a \geq 0 \;\Leftrightarrow\; \varphi(a) \geq 0$ *for all* $\varphi \in \mathcal{A}^+ \;\Leftrightarrow\; V(a) \subseteq [0, \infty)$.

Proof. By Proposition B.159 the positive functionals are precisely the positive multiples of the states φ appearing in the definition of $V_2(a)$. In view of this connection between $V(a)$ and $\mathcal{A}^+$, the three rightmost equivalences are obvious.

The leftmost implications $\Rightarrow$ are obvious in (i) and (iii). As to (ii), if $a = a^*$ there are $a_\pm \in \mathcal{A}_+$ such that $a = a_+ - a_-$. Now with $\varphi(a_\pm) \geq 0$ we have $\varphi(a) = \varphi(a_+) - \varphi(a_-) \in \mathbb{R}$.

To prepare the proof of the leftmost implications $\Leftarrow$, assume $a = a^* \neq 0$. Then a is normal, thus $r(a) = \|a\| > 0$, so that $\sigma(a)$ contains a $\lambda \neq 0$. Then $\lambda \in V(a)$ by Proposition B.148, so that by Proposition B.147 there exists $\varphi \in \mathcal{A}^*$ with $\varphi(\mathbf{1}) = \|\varphi\| = 1$ and $\varphi(a) = \lambda \neq 0$. Since φ is positive by Proposition B.159, we have proven that for every $a = a^* \neq 0$ there exists $\varphi \in \mathcal{A}^+$ with $\varphi(a) \neq 0$.

For arbitrary a we have $a = b + ic$ with $b, c \in \mathcal{A}_{sa}$. If $\varphi \in \mathcal{A}^+$ then $\varphi(a) = \varphi(b) + i\varphi(c)$, where $\varphi(b), \varphi(c) \in \mathbb{R}$. In view of this, $\varphi(a) \in \mathbb{R}$ implies $\varphi(c) = 0$, while $\varphi(a) = 0$ implies $\varphi(b) = \varphi(c) = 0$. Together with the result just proven (if $a = a^*$ and $\varphi(a) = 0\ \forall \varphi \in \mathcal{A}^+$ then $a = 0$) this finishes the proof of both (i) and (ii). And if $\varphi(a) \geq 0$ for all $\varphi \in \mathcal{A}^+$ then $a = a^*$ by (ii), and $\sigma(a) \subseteq V(a) \subseteq [0, \infty)$ gives positivity of a, finishing (iii). ■

B.162 Remark 1. We had already proven the equivalence $a = 0 \Leftrightarrow V(a) = \{0\}$ for all unital Banach algebras, but the above proof for C^*-algebras is considerably simpler.

2. We emphasise that $V(a) \subseteq \mathbb{R}$ implies $a = a^*$, which it not implied by $\sigma(a) \subseteq \mathbb{R}$.

3. Proposition B.161 shows that the states of a C^*-algebra can fulfill many of the tasks of the characters in the commutative case. One easily checks that every character is a state. But a state is a character only if it is multiplicative. By the Riesz-Markov-Kakutani theorem, the states on $C(X, \mathbb{C})$ are in bijection with the normalized positive measures on X, while the characters correspond to the Dirac measures $\delta_x, x \in X$, for which $\int_X f\, d\delta_x = f(x)$. □

B.163 Corollary *For a unital* C^**-algebra* $\mathcal{A}$ *and* $a \in \mathcal{A}$*, the following are equivalent:*

(i) a *is hermitian, i.e.* $V(a) \subseteq \mathbb{R}$.

(ii) $a = a^*$.

(iii) e^{ita} *is unitary for all* $t \in \mathbb{R}$.

(iv) $\|e^{ita}\| = 1\ \forall t \in \mathbb{R}$.

Proof. (i)⇔(ii) was part of Proposition B.161. (ii)⇒(iii) is elementary, see Remark 16.18.2. (iii)⇒(iv) is immediate by the C^*-identity. And Corollary B.153 gives (iv)⇔(i). ■

B.164 Remark In every Banach $*$-algebra one can still define self-adjointness by $a = a^*$ and unitarity by $u^*u = uu^* = \mathbf{1}$. (Confusingly, even in the Banach-$*$ literature 'hermitian' can be used for either of (i) or (ii).) One then still has (ii)⇔(iii), where $\Leftarrow$ follows by the computation

$$ia - ia^* = \frac{d}{dt}_{|t=0} e^{ita} \left(e^{ita}\right)^* = \frac{d}{dt}_{|t=0} e^{ita} e^{-ita} = 0.$$

And of course (i)⇔(iv) as in every Banach algebra. But in the absence of the C^*-identity neither (i)⇔(ii) nor (iii)⇔(iv) needs to hold. (It is not difficult to construct Banach $*$-algebras with $\|\mathbf{1}\| = 1$ having unitary elements with norm $\neq 1$.) In fact, one can prove that every Banach $*$-algebra in which (ii)⇒(iv) holds is a C^*-algebra, cf. [58]. □

We now study the relation between $V(a)$ for C^*-algebra elements to $\sigma(a)$ and to the spatial numerical range $W(A)$ in the case $\mathcal{A} = B(H)$.

B.165 Exercise Let $\mathcal{A}$ be a complex unital C^*-algebra and $a \in \mathcal{A}$.

(i) Prove $V(a) = \mathrm{conv}(\sigma(a))$ for commutative $\mathcal{A}$.

(ii) Prove $V(a) = \mathrm{conv}(\sigma(a))$ for normal a. (C^*-version of Exercise B.144.)

Before we can give the promised proof of $V(A) = \overline{W(A)}$ for $A \in B(H)$, we need some understanding of state spaces of concrete C^*-algebras, i.e. C^*-subalgebras of $B(H)$:

B.166 Lemma *Let $\mathcal{A}$ be a unital C^*-subalgebra. Then $S(\mathcal{A}) \subseteq \mathcal{A}^*$ is weak-$*$-closed and convex.*

Proof. Let $\varphi_1, \varphi_2 \in S(\mathcal{A})$ and $t \in [0,1]$. Then $\varphi = t\varphi_1 + (1-t)\varphi_2$ is positive and normalized, since $\varphi(\mathbf{1}) = 1$. Thus $\varphi \in S(\mathcal{A})$, proving convexity.

Let $\{\varphi_\iota\} \subset \mathcal{S}(\mathcal{A})$ be a net such that $\varphi_\iota \stackrel{w^*}{\to} \varphi \in \mathcal{A}^*$. With $\varphi_\iota(\mathbf{1}) = 1$ this implies $\varphi(\mathbf{1}) = 1$. By Alaoglu's theorem, $\mathcal{A}^*_{\leq 1}$ is weak-$*$ compact, thus weak-$*$ closed. With $\varphi_\iota \in \mathcal{A}^*_{\leq 1}$ this implies $\varphi \in \mathcal{A}^*_{\leq 1}$, thus $\|\varphi\| \leq 1$. Together with $\varphi(\mathbf{1}) = 1$ this gives $\|\varphi\| = 1$, thus $\varphi \in S(\mathcal{A})$. ■

B.167 Proposition *Let H be a complex Hilbert space and $\mathcal{A} \subseteq B(H)$ a C^*-subalgebra with $\mathbf{1}_H \in \mathcal{A}$. Then with $VS(\mathcal{A})) = \{\varphi_x = \langle \cdot\, x, x\rangle \mid x \in H, \|x\| = 1\} \subseteq S(\mathcal{A})$ we have $\overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*} = S(\mathcal{A})$.*

Proof. If $x \in H$, $\|x\| = 1$ then $\varphi_x : A \mapsto \langle Ax, x\rangle$ is in $\mathcal{A}^*$ and satisfies $\varphi_x(\mathbf{1}) = \|\varphi_x\| = 1$. Thus $VS(\mathcal{A}) \subseteq S(\mathcal{A})$, so that with Lemma B.166 we have $\overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*} \subseteq S(\mathcal{A})$. If $\overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*} \neq S(\mathcal{A})$, pick $\varphi_0 \in S(\mathcal{A})\backslash\overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*}$. Then by Corollary B.60, applied to $(\mathcal{A}^*, \tau_{w^*})$ there exist a $\psi \in \mathcal{A}^{**}$ and $t \in \mathbb{R}$ such that

$$\sup\{\mathrm{Re}\,\psi(\varphi) \mid \varphi \in \overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*}\} < t < \mathrm{Re}\,\psi(\varphi_0). \tag{B.21}$$

By Lemma 10.24(ii) there is a (unique) $a \in \mathcal{A}$ such that $\psi(\varphi) = \varphi(a)$ for all $\varphi \in \mathcal{A}^*$. Writing $a = b + ic$ with $b, c \in \mathcal{A}_{sa}$ and using that every $\varphi \in S(\mathcal{A})$ assumes real values on $\mathcal{A}_{sa}$, so that $\mathrm{Re}\,\varphi(a) = \varphi(b)$, (B.21) becomes

$$\sup\{\varphi(b) \mid \varphi \in \overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*}\} < t < \varphi_0(b). \tag{B.22}$$

In particular for $\varphi_x = \langle \cdot\, x, x\rangle$, where $x \in H$ is a unit vector, this implies $\langle bx, x\rangle \leq t$. This is equivalent to $\langle bx, x\rangle \leq t\|x\|^2$ for all $x \in H$, thus to $\langle (t\mathbf{1} - b)x, x\rangle \geq 0\ \forall x$. With Proposition 17.9 this is equivalent to $t\mathbf{1} - b \geq 0$. Since φ_0 is a state, thus positive, this implies $\varphi_0(t\mathbf{1} - b) \geq 0$, thus $\varphi_0(b) \leq t$. Since this contradicts (B.22), we indeed have $\overline{\mathrm{conv}(VS(\mathcal{A}))}^{w^*} = S(\mathcal{A})$. ■

B.168 Theorem *Let H be a complex Hilbert space. Then for all $A \in B(H)$ we have $V(A) = \overline{W(A)}$ and $R(A) = |||A|||$, where the numerical range $V(A)$ and radius $R(A)$ are taken in the Banach algebra $B(H)$.*

Proof. With $W(A) = \{\varphi(A) \mid \varphi \in VS(B(H)\}$ it is clear that $W(A) \subseteq V(A)$, thus $\overline{W(A)} \subseteq V(A)$ by closedness of $V(A)$. Given $\lambda \in V(A)$, pick $\varphi \in S(\mathcal{A})$ with $\varphi(A) = \lambda$. By Proposition B.167 there is a net $\{\varphi_\iota\} \subset \mathrm{conv}(VS(B(H)))$ such that $\varphi_\iota \stackrel{w^*}{\to} \varphi$. In particular, $\varphi_\iota(A) \to \varphi(A) = \lambda$. Since $\varphi_\iota \in \mathrm{conv}(VS(B(H)))$ implies $\varphi_\iota(A) \in \mathrm{conv}(W(A)) = W(A)$, we have proven $\lambda \in \overline{W(A)}$, thus $V(A) \subseteq \overline{W(A)}$. The equality $R(A) = |||A|||$ is now obvious. ■

B.13 Some more basic theory of C^*-algebras

B.13.1 The Fuglede-Putnam theorem

B.169 Theorem *Let $\mathcal{A}$ be a unital C^*-algebra over $\mathbb{C}$.*

(i) Let $a, c \in \mathcal{A}$. If a is normal and $ac = ca$ then $a^*c = ca^*$ (and $ac^* = c^*a$).

(ii) Let $a, b, c \in \mathcal{A}$. If a, b are normal and $ac = cb$ then $a^*c = cb^*$.

Proof. Obviously (i) is just the special case $a = b$ of (ii).

(ii) We define $f : \mathbb{C} \to \mathcal{A}, z \mapsto e^{za^*}ce^{-zb^*}$, where $e^a = \exp(a)$ is defined in terms of the power series as in Example 15.19. Expanding the two power series in the definition of f we have

$$f(z) = e^{za^*}ce^{-zb^*} = \left(\sum_{k=0}^{\infty} \frac{z^k(a^*)^k}{k!}\right) c \left(\sum_{l=0}^{\infty} \frac{(-z)^l(b^*)^l}{l!}\right) = \sum_{n=0}^{\infty} z^n d_n \quad \forall z \in \mathbb{C}$$

for certain $d_n \in \mathcal{A}$. (The reshuffling is justified by the absolute convergence of the series.) We only need $d_1 = a^*c - cb^*$, which is quite obvious. Thus the theorem follows if we prove $d_1 = 0$.

By induction, the assumption $ac = cb$ is seen to imply $a^n c = cb^n$. Multiplying by $z^n/n!$ and summing over $n \in \mathbb{N}_0$ gives $e^{za}c = ce^{zb}$ for all $z \in \mathbb{C}$, thus also $e^{za}ce^{-zb} = c$. Thus

$$f(z) = e^{za^*}ce^{-zb^*} = e^{za^*}(e^{-\overline{z}a}ce^{\overline{z}b})e^{-zb^*} = e^{za^*-\overline{z}a}ce^{\overline{z}b-zb^*} = e^{2i\mathrm{Im}(za^*)}ce^{-2i\mathrm{Im}(zb^*)},$$

where $e^{za^*}e^{-\overline{z}a} = e^{za^*-\overline{z}a}$ and $e^{\overline{z}b}e^{-zb^*} = e^{\overline{z}b-zb^*}$ hold due to normality of a and b, respectively. Now $\mathrm{Im}(za^*)$, $\mathrm{Im}(zb^*)$ are self-adjoint so that $e^{2i\mathrm{Im}(za^*)}$ and $e^{2i\mathrm{Im}(zb^*)}$ are unitary for all $z \in \mathbb{C}$, cf. Remark 16.18(ii), thus bounded. This proves that $f : \mathbb{C} \to \mathcal{A}$ is bounded. Now the following Lemma implies $d_1 = 0$, and we are done. ∎

B.170 Lemma *Let $\mathcal{A}$ be a unital Banach algebra and $\{d_n\}_{n\in\mathbb{N}_0} \subseteq \mathcal{A}$ such that $f : \mathbb{C} \to \mathcal{A}, z \mapsto \sum_{n=0}^{\infty} z^n d_n$ converges absolutely for all $z \in \mathbb{C}$. ($\Leftrightarrow \|d_n\|^{1/n} \to 0$.) If $\|f(z)\| \leq C|z|^M$ $\forall z \in \mathbb{C}$ then $d_n = 0$ for all $n > M$. In particular, if f is bounded then $d_n = 0$ $\forall n \geq 1$.*

Proof. Let $r > 0, m \in \mathbb{N}$. Then similarly to (B.20) we have

$$\int_0^{2\pi} e^{-imt} f(re^{it})dt = \int_0^{2\pi} e^{-imt} \left(\sum_{n=0}^{\infty} r^n e^{int} d_n\right) dt = \sum_{n=0}^{\infty} r^n d_n \int_0^{2\pi} e^{i(n-m)t}dt = 2\pi r^m d_m,$$

where used the uniform convergence of the series on bounded sets and $\int_0^{2\pi} e^{i(n-m)t}dt = 2\pi\delta_{n,m}$. With $\| \int_0^{2\pi} e^{-imt} f(re^{it})dt\| \leq \int_0^{2\pi} \|f(re^{it})\|dt \leq 2\pi C r^M$ we have

$$\|d_m\| = \frac{1}{2\pi r^m} \left\| \int_0^{2\pi} e^{-imt} f(re^{it})dt \right\| \leq C\frac{r^M}{r^m} \qquad \forall m \in \mathbb{N},\ r > 0.$$

Since the r.h.s. tends to zero as $r \to +\infty$ if $m > M$, the claim follows. ∎

B.171 Remark 1. Part (i) of the theorem was proven by Fuglede in 1950, (ii) by Putnam in 1951. The above elegant proof is due to Rosenblum (1958).[143]

2. For $\mathcal{A} = \mathbb{C}, M = 0$ the function f is entire and the lemma reduces to Liouville's theorem from complex analysis. But the general lemma can be deduced from Liouville's theorem: For

[143] Bent Fuglede (1925-2023), Danish mathematician. Calvin Richard Putnam (1924-2008), Marvin Rosenblum (1926-2003), American mathematicians.

every $\varphi \in \mathcal{A}^*$, the function $z \mapsto \varphi(f(z)) = \sum_{n=0}^{\infty} z^n \varphi(c^n)$ is entire and bounded, thus constant. Thus for all $z, z' \in \mathbb{C}$, $\varphi \in \mathcal{A}^*$ we have $\varphi(f(z) - f(z')) = 0$. The Hahn-Banach theorem now implies $f(z) - f(z') = 0\ \forall z, z'$, thus f is constant, so that $d_n = 0\ \forall n \geq 1$.

3. But as our proof of the lemma shows, no use of complex analysis (holomorphicity etc.) is needed if the function is a priori given by an everywhere convergent power series. Thus as in Rickart's proof of Theorem 13.39 the invocation of complex analysis can be replaced by much simpler harmonic analysis. As there, also here the integration can be removed: □

B.172 Exercise Give a proof of Lemma B.170 that does not use integration.

B.13.2 Homomorphisms of C^*-algebras

B.173 Theorem *Let $\mathcal{A}$ be a Banach $*$-algebra, $\mathcal{B}$ a C^*-algebra and $\alpha : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then $\|\alpha\| \leq 1$, thus α is continuous.*

Proof. Assume first that $\mathcal{A}, \mathcal{B}, \alpha$ are unital. Let $a \in \mathcal{A}$. By Lemma 15.2 we have $\sigma_{\mathcal{B}}(\alpha(a)) \subseteq \sigma_{\mathcal{A}}(a)$, thus $r_{\mathcal{B}}(\alpha(a)) \leq r_{\mathcal{A}}(a)$. Now the claim follows from

$$\|\alpha(a)\|^2 = \|\alpha(a)^*\alpha(a)\| = \|\alpha(a^*a)\| = r_{\mathcal{B}}(\alpha(a^*a)) \leq r_{\mathcal{A}}(a^*a) \leq \|a^*a\| \leq \|a\|^2,$$

where we used the C^*-identity for $\mathcal{B}$, the fact that α is a $*$-homomorphism, the fact that $r(b) = \|b\|$ for normal elements of the C^*-algebra $\mathcal{B}$, Lemma 15.2, $r(a) \leq \|a\|$ and the inequality $\|a^*a\| \leq \|a\|^2$ holding in every Banach $*$-algebra.

If α is non-unital, in particular if $\mathcal{A}$ or $\mathcal{B}$ has no unit, then $\widetilde{\alpha} : \widetilde{\mathcal{A}} \to \widetilde{\mathcal{B}}, (a, \alpha) \mapsto (\alpha(a), \alpha)$ is easily seen to be a unital $*$-homomorphism. Since $\widetilde{\mathcal{B}}$ is a C^*-algebra by Exercise 16.10, the above applies to $\widetilde{\alpha}$, and we have $\|\alpha\| \leq \|\widetilde{\alpha}\| \leq 1$. ■

B.174 Remark 1. The above result is one of many cases in the theory of C^*-algebras where the 'algebra dictates the analysis'. Further instances are Theorem B.176 and Corollary B.177.

2. For ($*$-)homomorphisms between general Banach ($*$-)algebras the question of continuity of homomorphisms is much more complicated, with connections to foundational matters like the continuum hypothesis, cf. [32]. □

B.175 Exercise Let $\mathcal{A}$ be a Banach $*$-algebra. For $a \in \mathcal{A}$ define $\|a\|_* = \sup_{(H,\pi)} \|\pi(a)\|_{B(H)}$, where H is a Hilbert space and $\pi : \mathcal{A} \to B(H)$ is a $*$-homomorphism. Prove that $\|\cdot\|_*$ is a C^*-seminorm on $\mathcal{A}$.

B.176 Theorem *Every injective $*$-homomorphism $\alpha : \mathcal{A} \to \mathcal{B}$ of C^*-algebras is an isometry, and $\alpha(\mathcal{A}) \subseteq \mathcal{B}$ is closed.*

Proof. If α is non-unital, define a unital homomorphism $\widetilde{\alpha} : \widetilde{\mathcal{A}} \to \widetilde{\mathcal{B}}$ as in the proof of Theorem B.173. Now $\widetilde{\alpha}$ is injective, thus if we prove the theorem in the unital case, we have that $\alpha = \widetilde{\alpha} \restriction \mathcal{A}$ is isometric. Since $\mathcal{A}$ is complete, this implies closedness of $\alpha(\mathcal{A}) \subseteq \mathcal{B}$. Thus α is an isometric $*$-isomorphism onto a C^*-subalgebra of $\mathcal{B}$.

Assume thus α to be unital. It suffices to prove $\|\alpha(a)\| = \|a\|$ for self-adjoint $a \in \mathcal{A}$, since then using the C^*-identity we have $\|\alpha(a)\|^2 = \|\alpha(a)^*\alpha(a)\| = \|\alpha(a^*a)\| = \|a^*a\| = \|a\|^2\ \forall a \in \mathcal{A}$.

Let thus $a = a^* \in \mathcal{A}$. We claim that $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(\alpha(a))$. [We cannot invoke Theorem 16.19 since we do not yet know $\alpha(\mathcal{A}) \subseteq \mathcal{B}$ to be closed!] This will imply $\|a\| = r_{\mathcal{A}}(a) = r_{\mathcal{B}}(\alpha(a)) = \|\alpha(a)\|$, as claimed. By (part of) Theorem 17.2 there are isometric $*$-homomorphisms $\gamma : C(\sigma(a), \mathbb{C}) \to \mathcal{A}$ and $\gamma' : C(\sigma(\alpha(a)), \mathbb{C}) \to \mathcal{B}$ continuously extending the maps $P \mapsto P(a)$

and $P \mapsto P(\alpha(a))$. Assuming $\sigma(\alpha(a)) \subsetneqq \sigma(a)$, we can find a non-zero $f \in C(\sigma(a), \mathbb{C})$ that vanishes on $\sigma(\alpha(a))$. (Since we are dealing with subsets of $\mathbb{R}$, this can be done by hand, without Urysohn's lemma.) If $\{P_n\}$ is a sequence of polynomials converging uniformly to f on $\sigma(a) \subseteq \mathbb{R}$ then the sequences $\{P_n(a)\}$ and $\{P_n(\alpha(a))\}$ converge uniformly to $\gamma(f) = f(a) \in \mathcal{A}$ and $\gamma'(f) = f(\alpha(a)) \in \mathcal{B}$, respectively. Since γ, γ' are isometric we have $\|f(a)\| = \|f\|_{\sigma(\mathcal{A})} \neq 0$ and $\|f(\alpha(a))\| = \|f\|_{\sigma(\alpha(a))} = 0$. Since α is a unital homomorphism, we have $\alpha(P_n(a)) = P_n(\alpha(a))\ \forall n$. The r.h.s. converges to $f(\alpha(a))$, the l.h.s. to $\alpha(f(a))$ by continuity of α (Theorem B.173), so that $\alpha(f(a)) = f(\alpha(a)) = 0$. Since $f(a) \neq 0$, this contradicts the injectivity of α. Thus we have $\sigma(\alpha(a)) = \sigma(a)$ as claimed. ■

B.177 Corollary *If $\alpha : \mathcal{A} \to \mathcal{B}$ is a $*$-homomorphism of C^*-algebras then $\alpha(\mathcal{A}) \subseteq \mathcal{B}$ is closed.*

Proof. $\mathcal{I} = \ker \alpha$ clearly is a two-sided self-adjoint ideal, and it is closed since α is continuous by Theorem B.173. Thus by Exercise 16.12 and Remark 16.13, $\mathcal{C} = \mathcal{A}/\mathcal{I}$ is a C^*-algebra, and the induced homomorphism $\gamma : \mathcal{C} \to \mathcal{B}$ is a $*$-homomorphism. By construction, γ is injective, thus an isometry with closed image $\gamma(\mathcal{C}) \subseteq \mathcal{B}$ by Theorem B.176. With $\alpha(\mathcal{A}) = \gamma(\mathcal{C})$ we are done. ■

B.14 Unbounded operators (mostly on Hilbert space)

B.14.1 Basic definitions. Closed and closable operators

B.178 Definition *A (possibly unbounded) operator on a Banach space V is a pair (D, A), where $D \subseteq V$ is a dense linear subspace and $A : D \to V$ is a linear map.*

In practice, we just denote an unbounded operator by a letter, say A, and denote its domain by $D(A), D_A$ or just D.

B.179 Remark 1. We ignore operators with non-dense range, since they are of very limited use. But assuming dense domain forces us to check for every construction of a new operator whether its domain is dense.

2. Writing 'unbounded operator' is undesirable since it would exclude the bounded ones. Since 'possibly unbounded operator' is unbearable, we will just write 'operator', 'densely defined linear' being implied.

3. If A is an operator and $t \in \mathbb{F}$ it is immediate how to define tA on the domain $D(A)$. □

B.180 Definition *Let A, B be operators on V. Then $A + B$ has domain $D(A) \cap D(B)$, on which it is defined as $x \mapsto Ax + Bx$.*[144] *And AB is defined as $x \mapsto A(Bx)$ on the domain $D(AB) = D(B) \cap B^{-1}(D(A)) = \{x \in D(B) \mid Bx \in D(A)\}$.*

Note that it is quite possible for $D(A + B)$ or $D(AB)$ to be non-dense or even trivial, i.e. $\{0\}$.

B.181 Definition *Let V be a Banach space.*

- *The graph of an operator A is $\mathcal{G}(A) = \{(x, Ax) \mid x \in D(A)\} \subseteq V \oplus V$.*
- *If A, B are operators with $D(A) \subseteq D(B)$ and $B \restriction D(A) = A$ (equivalently, $\mathcal{G}(A) \subseteq \mathcal{G}(B)$) then B is called an extension of A, denoted $A \subseteq B$.*

[144] In particular, $D(A + c\mathbf{1}) = D(A)$ with $A + c\mathbf{1} : x \mapsto Ax + cx$, which we will use often.

- *An operator A is closed if $\mathcal{G}(A) \subseteq V \oplus V$ is closed, and closable if $\overline{\mathcal{G}(A)}$ is the graph of an operator. That operator, the closure of A, is then denoted $\overline{A}$.*
- *If A is closed with domain D then a linear subspace $D_0 \subseteq D$ is a core for A if $\overline{A \upharpoonright D_0} = A$.*

B.182 Remark 1. If A is defined on $D = V$ then closedness of A (thus of $\mathcal{G}(A)$) is equivalent to boundedness by the closed graph theorem. But an operator whose domain is only dense can be closed without being bounded! In this case we cannot appeal to Lemma 3.12 to extend A to all of V.

2. Since $\overline{\mathcal{G}(A)}$ is a vector space, for closability of A it is necessary and suffiencient that $\overline{\mathcal{G}(A)}$ not contain $(0, a)$ with $a \neq 0$.

3. Trivially, closed $\Rightarrow$ closable. If A is closable then $\overline{A}$ is an extension of A. If A admits some closed extension B then A is closable: We have $\mathcal{G}(A) \subseteq \mathcal{G}(B)$, thus $\overline{\mathcal{G}(A)} \subseteq \mathcal{G}(B)$ by closedness of B. This implies that $\overline{\mathcal{G}(A)}$ is the graph of an operator $\overline{A}$. □

If V is a Banach space and $K, L \subseteq V$ are closed subspaces, it can happen that $K + L \subseteq V$ is non-closed. See Exercises 7.13 and 7.44. We focus on the case where $K \cap L = \{0\}$. Replacing V by $\overline{K + L}$, we may assume that $K + L$ is dense.

B.183 Exercise Let V be a Banach space and K, L closed linear subspaces such that $K \cap L = \{0\}$ and $\overline{K + L} = V$. Put $D = K + L$ and define $S : D \to D$ by $S(k + l) = k - l$ for all $k \in K, l \in L$. Prove:

(i) S is bounded if and only if $K + L = V$.

(ii) S is always closed.

B.14.2 Adjoints of unbounded Hilbert space operators

We now focus on Hilbert spaces.

B.184 Proposition *Let A be an operator on H with dense domain $D(A)$. Let*

$$D(A^*) = \{y \in H \mid \text{the functional } D(A) \ni x \mapsto \langle Ax, y\rangle \text{ is bounded}\}.$$

Then

(i) For each $y \in D(A^*)$ there is a unique $z_y \in H$ such that $\langle Ax, y\rangle = \langle x, z\rangle$ for all $x \in D(A)$. We put $A^*y = z_y$.

(ii) $D(A^*) \subseteq H$ is a linear subspace and the map $A^* : D(A^*) \to H$ is linear.

Proof. (i) Let $y \in D(A^*)$ and assume $z, z' \in H$ satisfy $\langle Ax, y\rangle = \langle x, z\rangle = \langle x, z'\rangle$ for all $x \in D(A)$. This implies $\langle x, z - z'\rangle = 0$ for all $x \in D(A)$. Since $D(A)$ is dense, $z - z' = 0$ follows.

(ii) Let $y \in D(A^*)$ and $c \in \mathbb{F}$. Then $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x \in D(A)$, implying $\langle Ax, cy\rangle = \langle x, cA^*y\rangle$ $\forall x \in D(A)$. It follows that $cy \in D(A^*)$ and $A^*(cy) = cA^*y$. Now let $y, y' \in D(A^*)$. Adding the equations $\langle Ax, y\rangle = \langle x, A^*y\rangle$ and $\langle Ax, y'\rangle = \langle x, A^*y'\rangle$ gives $\langle Ax, y + y'\rangle = \langle x, A^*(y + y')\rangle$, showing $y + y' \in D(A^*)$ and $A^*(y + y') = A^*y + A^*y'$. ■

Thus $\langle Ax, y\rangle = \langle x, A^*y\rangle$ $\forall x \in D(A), y \in D(A^*)$.

B.185 Exercise Prove that $A \subseteq B$ implies $B^* \subseteq A^*$.

B.186 Theorem *Let A be a densely defined operator on H. Then* *(i) The graph of A^* is closed.*

(ii) The domain $D(A^*)$ of A^* is dense if and only if A is closable.

(iii) If A is closable then $\overline{A}^* = A^*$ and $\overline{A} = A^{**}$. (In particular A closed $\Rightarrow A = A^{**}$.)

Proof. (i) For frequent later use we notice that for every unitary U on H and every linear subspace E (not necessarily closed) the identity $UE^\perp = (UE)^\perp$ holds. Now equip $H \oplus H$ with the obvious inner product: $\langle (a,b),(c,d)\rangle = \langle a,c\rangle + \langle b,d\rangle$. With this it is trivial to check that the linear operator $V : H \oplus H \to H \oplus H,\ (x,y) \mapsto (-y,x)$ is unitary.

We claim that the following holds for every densely defined A:

$$\mathcal{G}(A^*) = V\mathcal{G}(A)^\perp = (V\mathcal{G}(A))^\perp. \tag{B.23}$$

This follows from the chain of equivalences

$$\begin{aligned}
&(x,y) \in V\mathcal{G}(A)^\perp = (V\mathcal{G}(A))^\perp\\
&\quad\Leftrightarrow\quad \langle (x,y), V(z,Az)\rangle = 0 \quad \forall z \in D(A)\\
&\quad\Leftrightarrow\quad \langle (x,y), (-Az,z)\rangle = 0 \quad \forall z \in D(A)\\
&\quad\Leftrightarrow\quad \langle x, Az\rangle = \langle y,z\rangle \quad \forall z \in D(A)\\
&\quad\Leftrightarrow\quad \langle Az, x\rangle = \langle z,y\rangle \quad \forall z \in D(A)\\
&\quad\Leftrightarrow\quad x \in D(A^*), y = A^*x \;\Leftrightarrow\; (x,y) \in \mathcal{G}(A^*).
\end{aligned}$$

As an orthogonal complement, $\mathcal{G}(A)^\perp$ is closed, and the closedness of $\mathcal{G}(A^*)$ follows from (B.23).

(ii$\Leftarrow$) Since $\mathcal{G}(A) \subseteq H \oplus H$ is a linear subspace, we have

$$\overline{\mathcal{G}(A)} = \mathcal{G}(A)^{\perp\perp} = VV\mathcal{G}(A)^\perp = V(V\mathcal{G}(A)^\perp)^\perp = V(\mathcal{G}(A^*))^\perp \tag{B.24}$$

where we used $V^2 = -\mathbf{1}$, the commutativity of V and $\perp$ and (B.23).

If now $D(A^*)$ is not dense, we can find $x \in D(A^*)^\perp \backslash \{0\}$. Then for each $y \in D(A^*)$ we have $\langle (x,0),(y,A^*y)\rangle = \langle x,y\rangle = 0$, implying $(x,0) \in \mathcal{G}(A^*)^\perp$. Thus using (B.24) we find $(0,x) = V(x,0) \in V\mathcal{G}(A^*)^\perp = \overline{\mathcal{G}(A)}$. In view of $x \neq 0$ and $(0,0) \in \overline{\mathcal{G}(A)}$, this shows that $\overline{\mathcal{G}(A)}$ is not the graph of an operator and thus A is not closable.

(ii$\Rightarrow$)+(iii) Assuming that $D(A^*)$ is dense, we can define A^{**}. Replacing A by A^* in (B.23) gives $\mathcal{G}(A^{**}) = V\mathcal{G}(A^*)^\perp = VV\mathcal{G}(A)^{\perp\perp} = \mathcal{G}(A)^{\perp\perp} = \overline{\mathcal{G}(A)}$. Thus $\overline{\mathcal{G}(A)}$ is the graph of the operator A^{**}, showing that A is closable with $\overline{A} = A^{**}$.

The remaining claim in (iii) follows from the computation $A^* = \overline{A^*} = A^{***} = \overline{A}^*$, where we used in turn the closedness of A^*, the fact that A^* is closed, thus closable with $\overline{A^*} = A^{***}$, and the closability of A with $\overline{A} = A^{**}$. ■

B.187 Definition *An operator A with dense domain $D \subseteq H$ is called*

- *symmetric if $\langle Ax,y\rangle = \langle x,Ay\rangle \quad \forall x,y \in D$.*
- *self-adjoint if $A = A^*$, including $D(A^*) = D(A)$.*

We notice:

- A is symmetric if and only if $A \subseteq A^*$.
- Thus self-adjoint $\Rightarrow$ symmetric, and a symmetric A is self-adjoint if and only if $D(A^*) = D(A)$.
- Let $A \subseteq B$. Then B symmetric $\Rightarrow$ A symmetric. But the converse need not hold.

- If A is symmetric then $A \subseteq A^{**} \subseteq A^*$ (and conversely). This follows from the fact that A^* is closed and therefore contains the closure $\overline{A} = A^{**}$ of A.
- The closed symmetric operators are those satisfying $A = A^{**} \subseteq A^*$.

B.188 Exercise Prove that every symmetric operator is closable, and its closure is symmetric.

B.189 Exercise Give an example of $A \subsetneqq B \subseteq C$ with A, B symmetric but C not symmetric.

B.190 Exercise Let $H = \ell^2(\mathbb{N}, \mathbb{C})$ and $h : \mathbb{N} \to \mathbb{R}$. Define $D = \{f \in H \mid hf \in H\}$ (pointwise multiplication) and $Af = hf$. Prove that A is self-adjoint.

B.191 Definition *An operator A is essentially self-adjoint if it is closable with $\overline{A}$ self-adjoint.*

B.192 Exercise Prove that the following are equivalent:

(i) A is essentially self-adjoint.

(ii) $A \subseteq A^{**} = A^*$.

(iii) A and A^* are symmetric.

B.193 Lemma *If A is essentially self-adjoint then $\overline{A}$ is the only self-adjoint extension of A.*

Proof. Let B be a self-adjoint extension of A. Thus $A \subseteq B = B^*$. Since B^* is closed, this implies $A \subseteq \overline{A} \subseteq B = B^* \subseteq A^* = \overline{A}^* = \overline{A}$. Thus $\overline{A} = B$. ■

One can show that the converse is also true: If A has a unique self-adjoint extension B then $\overline{A}$ is self-adjoint and therefore coincides with B.

B.194 Definition *A symmetric operator A is maximal symmetric if every symmetric extension $B \supseteq A$ coincides with A.*

B.195 Exercise Prove that self-adjoint ⇒ maximal symmetric ⇒ closed symmetric.

The following diagram summarizes the implications:

$$\begin{array}{ccccccc}
A \text{ self}-\text{adjoint} & \Longrightarrow & A \text{ max. symm.} & \Longrightarrow & A \text{ closed symmetric} & \Longrightarrow & A \text{ closed} \\
\Big\Downarrow & & & & \Big\Downarrow & & \Big\Downarrow \\
A \text{ ess. s.}-\text{a.} & & \Longrightarrow & & A \text{ symmetric} & \Longrightarrow & A \text{ closable}
\end{array}$$

B.14.3 Basic criterion for (essential) self-adjointness

Verifying (essential) self-adjointness of an operator directly from the definitions can be tedious, making it desirable to have managable criteria. We need some preparations.

The kernel and image of an unbounded operator $H \supseteq D \stackrel{A}{\to} H$ are defined in the obvious way, namely as $\{x \in D \mid Ax = 0\}$ and AD, respectively.

B.196 Lemma *For every (densely defined, of course) operator A*

$$\ker A^* = (AD)^\perp.$$

Proof. The proof is essentially that of Lemma 11.10(ii) with some attention to the domains: We have $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x \in D(A), y \in D(A^*)$. Thus

$$y \in \ker A^* \subseteq D(A) \;\Leftrightarrow\; \langle Ax, y\rangle = 0 \;\forall x \in D(A) \;\Leftrightarrow\; y \in (AD)^\perp.$$

■

B.197 Theorem *Let A be a symmetric operator with domain $D \subseteq H$. Then the following are equivalent:*

(i) A is self-adjoint.

(ii) A is closed and $\ker(A^* \pm i) = \{0\}$. ($A^* \pm i$ both injective.)

(iii) $(A \pm i)D = H$. ($A \pm i$ both surjective.)

(For an unbounded A, one defines $A + c\mathbf{1}$ in the obvious way on the domain $D(A)$.)

Proof. (i)⇒(ii) Let $x \in D(A) = D(A^*)$. If $A^*x = ix$ then $Ax = ix$ by $A = A^*$, thus

$$i\langle x, x\rangle = \langle ix, x\rangle = \langle Ax, x\rangle = \langle x, A^*x\rangle = \langle x, Ax\rangle = \langle x, ix\rangle = -i\langle x, x\rangle,$$

implying $x = 0$ and therefore $\ker(A - i) = \{0\}$. The identity $\ker(A + i) = \{0\}$ is proven in the same way.

(ii)⇒(iii) Since $A^* + i$ is injective, Lemma B.196 gives $((A - i)D)^* = \{0\}$, so that $(A - i)D$ is dense. It remains to prove that $(A - i)D$ is closed. Using that A is symmetric we find

$$\begin{aligned} \|(A - i)x\|^2 &= \langle (A - i)x, (A - i)x\rangle = \|Ax\|^2 + \|x\|^2 - i\langle x, Ax\rangle + i\langle Ax, x\rangle \\ &= \|Ax\|^2 + \|x\|^2 \end{aligned}$$

and therefore $\|(A - i)x\| \geq \|x\|$ for all $x \in D(A)$. The rest of the proof is an adaptation of Lemma 7.39 to unbounded operators: By the inequality, the map $A - i : D(A) \to H$ is injective, thus $A - i : D(A) \to (A - i)D(A)$ is a bijection. If $\{y_n\}$ is a Cauchy sequence in $(A - i)D(A)$ then the inequality gives that $\{x_n = (A - i)^{-1}(y_n)\}$ is a Cauchy sequence in $D(A - i)$. Thus $\{(x_n, y_n)\} \subseteq \mathcal{G}(A - i)$ is Cauchy. The closedness of A implies closedness of $A - i$, thus of $\mathcal{G}(A - i)$, so that $(x_n, y_n) \to (x, y) \in \mathcal{G}(A - i)$. This proves that $\lim_n y_n = y \in \mathcal{G}(A - i)$, so that $(A - i)D$ is closed, completing the proof of surjectivity of $A - i$. For $A + i$ one argues analogously.

(iii)⇒(i) Let $y \in D(A^*)$. Since $(A - i)D = H$, there exists $x \in D(A)$ such that

$$(A - i)x = (A^* - i)y. \tag{B.25}$$

Since A is symmetric, $x \in D(A) \subseteq D(A^*)$ and $A^*x = Ax$. Thus (B.25) rewrites as

$$(A^* - i)(x - y) = 0. \tag{B.26}$$

Appealing to $(A + i)D = H$, Lemma B.196 gives $\ker(A^* - i) = \ker((A + i)^*) = H^\perp = \{0\}$, thus injectivity of $A^* - i$. Combining this with (B.26) we obtain $x - y = 0$, so that $y = x \in D(A)$. This proves $D(A^*) \subseteq D(A)$, thus $A = A^*$. ■

Similarly:

B.198 Corollary *Let A be symmetric. Then the following are equivalent:*

(i) A is essentially self-adjoint.

(ii) $\ker(A^* \pm i) = \{0\}$. *($A^* \pm i$ both injective.)*

(iii) $\overline{(A \pm i)D} = H$.

Proof. (i)⇒(ii) By assumption A is closable with self-adjoint closure $\overline{A}$. Now the theorem gives $\ker(\overline{A}^* \pm i) = \{0\}$, and with $\overline{A}^* = A^*$ we have (ii) of the corollary.

(ii)⇒(i) From Exercise B.188 we know that A is closable with $\overline{A}$ symmetric. With $A^* = \overline{A}^*$, hypothesis (ii) gives $\ker(\overline{A}^* \pm i) = \{0\}$, so that $\overline{A}$ is self-adjoint by implication (ii)⇒(i) in the theorem. Equivalently, A is essentially self-adjoint.

(ii)⇔(iii) This is immediate from Lemma B.196, applied to $A \pm i$. ■

B.15 Glimpse of non-linear FA: Schauder's fixed point theorem

In this final section we give a glimpse of non-linear functional analysis by proving Schauder's fixed point theorem, which is a generalization of Brouwer's fixed point theorem to Banach spaces.

B.199 Definition *A topological space X has the fixed-point property if for every continuous map $f : X \to X$ there is $x \in X$ such that $f(x) = x$, i.e. a fixed-point.*

B.200 Theorem (Brouwer, Hadamard, 1910) [145] *$[0,1]^n$ has the fixed point property. The same holds for every non-empty compact convex subset of $\mathbb{R}^n$.*

The second result follows from the first since such an X is homeomorphic to some $[0,1]^m$. There are many proofs of the first result. For what probably is the simplest proof (due to Kulpa) of the first statement, using only some easy combinatorics, see [108]. (Proofs using algebraic topology or analysis involve inessential elements and don't reduce the combinatorics.)

B.201 Theorem (Schauder 1930) *Every non-empty compact convex subset K of a normed vector space has the fixed point property.*

Proof. Let $(V, \|\cdot\|)$ be a normed vector space, $K \subseteq V$ a non-empty compact convex subset and $f : K \to K$ continuous. Let $\varepsilon > 0$. Since K is compact, thus totally bounded, there are $x_1, \ldots, x_n \in K$ such that $K \subseteq \bigcup_{i=1}^n B(x_i, \varepsilon)$. Thus if we define continuous functions $\alpha_i : X \to \mathbb{R}_+,\ i = 1, \ldots, n$ by

$$\alpha_i(x) = \max(\varepsilon - \|x - x_i\|, 0)$$

we see that for each $x \in K$ there is at least one i such that $\alpha_i(x) > 0$. Since the α_i are continuous, so is the map

$$P_\varepsilon : K \to K, \quad x \mapsto \frac{\sum_{i=1}^n \alpha_i(x) x_i}{\sum_{i=1}^n \alpha_i(x)}.$$

Since $P_\varepsilon(x)$ is a convex combination of those x_i for which $\|x - x_i\| < \varepsilon$, we have $\|P_\varepsilon(x) - x\| < \varepsilon$ for all $x \in K$. The finite-dimensional subspace $V_n = \mathrm{span}(x_1, \ldots, x_n) \subseteq V$ is isomorphic to some $\mathbb{R}^m$, and by Corollary 2.32 the restriction of the norm $\|\cdot\|$ to V_n is equivalent to the Euclidean norm on $\mathbb{R}^m$. Thus the convex hull $\mathrm{conv}(x_1, \ldots, x_n) \subseteq V_n$ into which P_ε maps is homeomorphic to a compact convex subset of $\mathbb{R}^m$ and thus has the fixed point property by Theorem B.200.

[145] Luitzen Egbertus Jan Brouwer (1881-1966). Dutch mathematician. Important contributions to topology, founding of intuitionism. Jacques Hadamard (1865-1963). French mathematician.

Thus if we define $f_\varepsilon = P_\varepsilon \circ f$ then f_ε maps $\text{conv}(x_1, \ldots, x_n)$ into itself and thus has a fixed point $x' = f_\varepsilon(x')$. Now,

$$\|x' - f(x')\| \leq \|x' - f_\varepsilon(x')\| + \|f_\varepsilon(x') - f(x')\| = \|f_\varepsilon(x') - f(x')\| = \|P_\varepsilon(f(x')) - f(x')\| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we find $\inf\{\|x - f(x)\| \mid x \in K\} = 0$. Since K is compact and $x \mapsto \|x - f(x)\|$ continuous, the infimum is assumed, thus f has a fixed point in K. ■

The use of methods/results from algebraic topology is quite typical for non-linear functional analysis. (But also linear functional analysis connects to algebraic topology, for example via K-theory, cf. e.g. [110, Chapter 7].)

B.202 Definition *Let V be a Banach space and $W \subseteq V$. A map $f : W \to V$ is called compact if it is continuous and $f(S) \subseteq V$ is precompact for every bounded $S \subseteq V$.*

B.203 Corollary *Let V be a Banach space and $C \subseteq V$ closed, bounded and convex. If $f : C \to V$ is compact and $f(C) \subseteq C$ then f has a fixed point in C.*

Proof. Since C is bounded and f is compact, $f(C) \subseteq V$ is precompact, thus $K = \overline{\text{conv}(f(C))}$ is compact by Mazur's Theorem B.75 and convex. Thus K has the fixed point property by Schauder's theorem. Since C is closed and convex, we have $K \subseteq C$, thus f is defined on K and maps it into $f(K) \subseteq f(C) \subseteq C$. Thus f has a fixed point $x \in K \subseteq C$. ■

C The mathematicians encountered in these notes

- Marc-Antoine Parseval (1755-1836)
- Friedrich Bessel (1784-1846)
- Augustin-Louis Cauchy (1789-1857)
- Viktor Yakovlevich Bunyakovski (1804-1889)
- Sir William Rowan Hamilton (1805-1865)
- Karl Theodor Wilhelm Weierstrass (1815-1897)
- Eduard Heine (1821-1881)
- Bernhard Riemann (1826-1866)
- Carl Gottfried Neumann (1832-1925)
- Karl Hermann Armandus Schwarz (1843-1921)
- Giulio Ascoli (1843-1896)
- Cesare Arzelà (1847-1912)
- Adolf Hurwitz (1859-1919)
- Otto Hölder (1859-1937)
- Vito Volterra (1860-1940)
- Eliakim Hastings Moore (1862-1932)
- David Hilbert (1862-1943)
- Hermann Minkowski (1864-1909)
- Jacques Hadamard (1865-1963)
- Erik Ivar Fredholm (1866-1927)
- Felix Hausdorff (1868-1942)
- Ernst Steinitz (1871-1928)
- Emile Borel (1871-1956)
- Constantin Carathéodory (1873-1950)
- René-Louis Baire (1874-1932)
- Issai Schur (1874-1941)
- Ernst Sigismund Fischer (1875-1954)
- Erhard Schmidt (1876-1959)
- Maurice Fréchet (1878-1973)
- Hans Hahn (1879-1934)
- Frigyes Riesz (1880-1956)
- Sergei Natanovich Bernstein (1880-1968)
- Otto Toeplitz (1881-1940)
- Luitzen Egbertus Jan Brouwer (1881-1966)
- Ernst David Hellinger (1883-1950)

- Eduard Helly (1884-1943)
- Hermann Weyl (1885-1955)
- Marcel Riesz (1886-1969)
- Paul Lévy (1886-1971)
- Hugo Steinhaus (1887-1972)
- Stefan Banach (1892-1945)
- Eduard Čech (1893-1960)
- Norbert Wiener (1894-1964)
- Juliusz Schauder (1899-1943)
- Herman Auerbach (1901-1942)
- John von Neumann (1903-1957)
- Andrey Andreyevich Markov (1903-1979)
- Andrey Nikolaevich Kolmogorov (1903-1987)
- Marshall Harvey Stone (1903-1989)
- Władysław Orlicz (1903-1990)
- Henri Cartan (1904-2008)
- Stanisław Mazur (1905-1981)
- Arne Beurling (1905-1986)
- Henry Frederic Bohnenblust (1906-2000)
- Nachman Aronszajn (1907-1980)
- Mark Grigorievich Krein (1907-1989)
- Meier Eidelheit (1910-1943)
- Angus Ellis Taylor (1911-1999)
- Shizuo Kakutani (1911-2004)
- David Milman (1912-1982)
- Billy James Pettis (1913-1979)
- Charles Earl Rickart (1913-2002)
- Herman Heine Goldstine (1913-2004)
- Israel Moiseevich Gelfand (1913-2009)
- Vitold Lvovich Šmulyan (1914-1944)
- Leonidas Alaoglu (1914-1981)
- Laurent Schwartz (1915-2002)
- Gustave Choquet (1915-2006)
- Frederick Valentine Atkinson (1916-2002)
- Aryeh Dvoretzky (1916-2008)
- William Frederick Eberlein (1917-1986)

- Robert Clarke James (1918-2004)
- Jerzy Łoś (1920-1988)
- Nicolaas Hendrik Kuiper (1920-1994)
- Claude Ambrose Rogers (1920-2005)
- Mychajlo Jossypowytsch Kadets (1923-2011)
- Calvin Richard Putnam (1924-2008)
- Bent Fuglede (1925-2023)
- Robert Ralph Phelps (1926-2013)
- Czesław Ryll-Nardzewski (1926-2015)
- Kennan Tayler Smith (1926-2000)
- Felix Earl Browder (1927-2016)
- Errett Albert Bishop (1928-1983)
- Alexander Grothendieck (1928-2014)
- Wilhelmus Anthonius Josephus Luxemburg (1929-2018)
- Karel de Leeuw (1930-1978)
- Shaul Reuven Foguel (1931-2020)
- Aleksander Pełczyński (1932-2012)
- Czeslaw Bessaga (1932-2021)
- John Robert Ringrose (b. 1932)
- Lior Tzafriri (1936-2008)
- Joram Lindenstrauss (1936-2012)
- Stefan Oscar Walter Hildebrandt (1936-2015)
- Haskell Paul Rosenthal (1940-2021)
- Paul Robert Chernoff (1942-2017)
- Stanisław Kwapień (b. 1942)
- Per Henrik Enflo (b. 1944)
- Victor Lomonosov (1946-2018).

Embarrassingly the above list contains no women. In the related areas of PDEs and variational calculus (which is functional analysis, but non-linear) there have been quite a few, in particular Sofia Kowalevskaya (1850-1891), Emmy Noether (1882-1935), Olga Ladyzhenskaya (1922-2004), Cathleen Synge Morawetz (1923-2017), Yvonne Choquet-Bruhat (b. 1923), Karen Uhlenbeck (b. 1942, Abel prize 2019), ... (In classical and harmonic analysis, Grace Chisholm Young (1868-1944), Nina Bari (1901-1961) and Dorothy Maharam Stone (1917-2014) come to mind.) But in linear functional analysis the first notable women probably are

- Mary Beth Ruskai (1944-2023), who worked on functional analytic questions of quantum theory.
- Nicole Tomczak-Jaegermann (1945-2022), who worked on Banach space theory, e.g. [165].
- Dusa McDuff (b. 1945), who after a brilliant PhD thesis (1970) on operator algebras switched to more geometric matters (symplectic topology and geometry).
- Other female operator algebraists: Marie Choda and Claire Anantharaman-Delaroche with first publications in 1962 and 1967, respectively.

D Results stated, but not proven

- 1905, 1913: Levy-Steinitz theorem on reordering series in finite dimension.
- 1911: Carathéodory's convexity theorem.
- 1929/1940: Orlicz-Pettis theorem.
- 1936: Kakutani/Birkhoff: TVS is metrizable $\Leftrightarrow$ 0 has countable neighborhood base.
- 1940s: Gelfand, A. E. Taylor, Dunford, Lorch: holomorphic functional calculus.
- 1940/1947: Eberlein-Šmulyan theorem.
- 1951: R.C. James' space with $V \cong V^{**}$ but $\iota_V(V) \subsetneqq V^{**}$.
- 1955: Grothendieck: approx prop. $\Leftrightarrow$ finite rank ops. approx id on compact subsets
- 1957/1964: R.C. James: Banach sp. is reflexive $\Leftrightarrow \forall \varphi \in V^* \; \exists 0 \neq x \in V : |\varphi(x)| = \|x\| \|\varphi\|$.
- 1958: Bessaga/Pelczinski: Banach space V does not contain $c_0 \;\Leftrightarrow$ every WUC series in V converges unconditionally.
- 1959: Lidskii's theorem.
- 1960/1: Dvoretzky's theorem.
- 1961/3: Bishop-Phelps theorems.
- 1965: Kuiper's theorem.
- 1971: Kadets-Snobar theorem.
- 1971: Lindenstrauss-Tzafriri: Banach space with all closed subspaces complemented is isomorphic to Hilbert space.
- 1972: Kwapień's theorem
- 1973: Enflo: Banach space without approximation property.
- 1973: Pecherskii's theorem.
- 1974: Rosenthal's ℓ^1 theorem.
- 1975/1987: Enflo: Banach space operator without invariant subspace.
- 1977: Blair: Baire's theorem $\Leftrightarrow$ DC_ω.
- 1981: Szankowski: $B(H)$ doesn't have approximation property.
- 2011: Argyros/Haydon: solution of the scalar-plus-compact problem (Banach space V with $B(V) = \mathbb{C}\mathbf{1} + K(V)$).

Note that apart from the last one and some new takes on old results like [7, 38, 53, 61] we have hardly even mentioned any results from the last 40 years!

E What next?

For general orientation, the article [166] and Dieudonné's book [40] are strongly recommended.

- General topological vector spaces, F-spaces, beginning with [141].
- Locally convex spaces and distributions, beginning with [141, 30, 94]
- Sobolev spaces. Applications of the latter and of distributions to PDEs, e.g. [52].
- Index theory of elliptic PDEs (Atiyah-Singer etc.)
- Much more on Banach spaces, beginning with [102], then [26, 98, 1, 97] etc.
- Connections between Banach spaces and classical/harmonic analysis, e.g. wavelets, Hardy spaces, and with probability theory, e.g. martingales. E.g. [176, 77].
- More operator theory on Banach and Hilbert spaces, e.g. [59, 135, 24, 68, 129, 152].
- Semigroup theory, e.g. [4, 50]
- Banach algebras [131, 18, 80].
- Connections between Banach algebras and complex analysis.
- C^*- and von Neumann algebras: [110, 79] and many other books.
- Interactions of operator algebras and operator theory, beginning with [110, 33, 42].
- Algebraic topology of operator algebras, non-commutative geometry.
- Non-linear functional analysis, e.g. [28, 37, 39, 115, 164, 177].
- Variational calculus (with applications to differential equations), non-linear optimization.
- Applications of operator theory in quantum mechanics, e.g. [92].
- Applications of operator algebras in statistical physics and quantum field theory. E.g. [22, 65].
- Non-archimedean/p-adic functional analysis, e.g. [138, 125].

F Approximate schedule for 14 lectures à 90 minutes

1. Sections 1-2.2: Introduction. Topological vector spaces, normed spaces.
2. Sections 2.3-3: Glimpse beyond normed spaces. More on normed spaces and bounded maps
3. Section 4: The spaces $\ell^p(S, \mathbb{F})$ and $c_0(S, \mathbb{F})$. Proofs of Hölder and Minkowski inequalities, dual spaces. Most other proofs omitted or just sketched.
4. Sections 5.1-5.5: Hilbert spaces up to and incl. H^*.
5. Section 5.6-6.1: Bases and tensor products of Hilbert spaces. Quotients of Banach spaces.
6. Section 6.2: complemented subspaces. Section 7: Open mapping thm. incl. Baire, closed graph theorem, boundedness below.
7. Section 8: Uniform boundedness theorem and applications.
8. Section 9: Hahn-Banach theorem and applications incl. reflexivity and transpose of operators.
9. Section 11: Hilbert space operators, beginning with adjoint. Self-adjoint, normal ops, etc.
10. Finish Hilbert space operators. Then Section 12 on compact operators.
11. Sections 13.1-13.2.2: spectra of operators, spectrum in a Banach algebra.
12. Sections 13.2.3 and 13.2.4: Beurling-Gelfand theorem and its applications.
13. Section 14: spectral theorems for compact operators (normal or not). Quick mention of Fredholm operators. Section 15: Characters vs. maximal ideals. Power series functional calculus.
14. Sections 16, 17: C^*-algebras, continuous functional calculus for normal operators. Spectral theorems for normal operators (only Section 18.1).

    −

15. Section 10 on weak and weak-$*$ topologies, first half of Section 12.2.
16. Second half of Section 12.2, Section 19.

Lectures 15 and 16 are **not** part of the course since the weeks 15-16 of the semester were scrapped a few years ago. If there is interest, I can give these two lectures (incl. homework) in January for 1 EC.

All papers appearing in the bibliography are cited somewhere, but not all books. Still, all are worth looking at.

References

[1] F. Albiac, N. J. Kalton: *Topics in Banach space theory.* 2nd. ed. Springer, 2016.

[2] D. Amir: *Characterizations of inner product spaces.* Birkhäuser, 1986.

[3] N. R. Andre, S. M. Engdahl, A. E. Parker: An Analysis of the First Proofs of the Heine-Borel Theorem. `https://www.maa.org/press/periodicals/convergence/an-analysis-of-the-first-proofs-of-the-heine-borel-theorem`

[4] D. Applebaum: *Semigroups of linear operators.* Cambridge University Press, 2019.

[5] S. A. Argyros, R. G. Haydon: A hereditarily indecomposable $\mathcal{L}_\infty$-space that solves the scalar-plus-compact problem. Acta Math. **206**, 1-54 (2011).

[6] S. A. Argyros, R. G. Haydon: Bourgain-Delbaen $\mathcal{L}_\infty$-spaces, the scalar-plus-compact property and related problems. Proc. Intern. Congr. Math., vol. 3 (2018).

[7] M. B. Asadi, A. Khosravi: An elementary proof of the characterization of isomorphisms of standard operator algebras. Proc. Amer. Math. Soc. **134**, 3255-3256 (2006).

[8] M. Bachir: On the Krein-Milman-Ky Fan theorem for convex compact metrizable sets. Illinois Journ. Math. **62**, 1-24 (2018).

[9] S. Banach: *Théorie des opérations linéaires*, 1932. Engl. transl.: *Theory of linear operations.* North-Holland, 1987.

[10] W. R. Bauer, R. H. Benner: The non-existence of a Banach space of countably infinite Hamel dimension. Amer. Math. Monthly **78**, 895-596 (1971).

[11] V. Baumann: W. A. J. Luxemburgs Beweis des Satzes von Hahn-Banach. Archiv Math. **18**, 271-272 (1967).

[12] A. F. Beardon: *Limits.* Springer, 1997.

[13] J. L. Bell, D. H. Fremlin: A geometric form of the axiom of choice. Fund. Math. **77**, 168-170 (1972/3).

[14] S. J. Bernau: The spectral theorem for normal operators. Journ. London Math. Soc. **40**, 478-486 (1965).

[15] G. Birkhoff, E. Kreyzig: The establishment of functional analysis. Histor. Math. **11**, 258-321 (1984).

[16] E. Bishop, R. R. Phelps: A proof that every Banach space is subreflexive. Bull. Amer. Math. Soc. **67**, 97-98 (1961).

[17] C. E. Blair: The Baire category theorem implies the principle of dependent choices. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **25**, 933-934 (1977).

[18] F. F. Bonsall, J. Duncan: *Complete normed algebras.* Springer, 1973

[19] F. F. Bonsall, J. Duncan: *Numerical ranges of operators on normed spaces and of elements of normed algebras* and *Numberical ranges II.* Cambridge University Press, 1971, 1973.

[20] R. Bouldin: The essential minimum modulus. Indiana Univ. Math. Journ. **30**, 513-517 (1981).

[21] N. Bourbaki: *Topological vector spaces. Chapters 1-5.* Springer, 1987.

[22] O. Bratteli, D. W. Robinson: *Operator algebras and quantum statistical mechanics.* Two volumes. 2nd. eds.. Springer, 1987, 1997.

[23] H. Brézis: *Functional analysis, Sobolev spaces and partial differential equations.* Springer, 2011.

[24] M. S. Brodskiĭ: *Triangular and Jordan representations of linear operators.* American Mathematical Society Translations, 1971.

[25] T. Bühler, D. A. Salamon: *Functional analysis.* American Mathematical Society, 2018.

[26] N. L. Carothers: *A short course on Banach space theory.* Cambridge University Press, 2005.

[27] D. Choimet, H. Queffélec: *Twelve landmarks of twentieth-century analysis.* Cambridge University Press, 2009.

[28] P. G. Ciarlet: *Linear and nonlinear functional analysis with applications.* Society for Industrial and Applied Mathematics, 2013.

[29] D. L. Cohn: *Measure theory.* 2nd. ed. Springer, 2013.

[30] J. B. Conway: *A course in functional analysis.* 2nd. ed. Springer, 2007.

[31] J. B. Conway: The compact operators are not complemented in $B(H)$. Proc. Amer. Math. Soc. **32**, 549-550.

[32] H. G. Dales: *Banach algebras and automatic continuity.* Oxford University Press, 2000.

[33] K. R. Davidson: *C^*-Algebras by example.* American Mathematical Society, 1996.

[34] A. M. Davie: The Banach approximation problem. J. Approx. Th. **13**, 392-394 (1975).

[35] C. Davis: The Toeplitz-Hausdorff theorem explained. Canad. Math. Bull. **14**, 245-246 (1971).

[36] M. M. Day: Reflexive Banach spaces not isomorphic to uniformly convex spaces. Bull. Amer. Math. Soc. **47**, 313-317 (1941).

[37] K. Deimling: *Nonlinear functional analysis.* Springer, 1985.

[38] S. Delpech: A short proof of Pitt's compactness theorem. Proc. Amer. Math. Soc. **137**, 1371-1372 (2008).

[39] Z. Denkowski, S. Migórski, N. S. Papageorgiou: *An introduction to nonlinear analysis: Theory.* Springer, 2003.

[40] J. Dieudonné: *History of functional analysis.* North-Holland, 1981.

[41] J.-L. Dorier: A general outline of the genesis of vector space theory. Hist. Math. **22**, 227-261 (1995).

[42] R. G. Douglas: *Banach algebra techniques in operator theory.* 2nd. ed. Springer, 1998.

[43] N. Dunford, J. T. Schwartz: *Linear operators. I. General theory.* Interscience Publishers, 1958, John Wiley & Sons, 1988.

[44] H.-D. Ebbinghaus et al.: *Numbers.* Springer, 1991.

[45] D. E. Edmunds, W. D. Evans: *Spectral theory and differential operators.* Oxford University Press, 1987, 2018.

[46] N. Eldredge: Answer to the question 'Is there a simple direct proof of the Open Mapping Theorem from the Uniform Boundedness Theorem?' on `mathoverflow.net`. https://mathoverflow.net/questions/190587/is-there-a-simple-direct-proof-of-the-open-mapping-theorem-from-the-uniform-boun

[47] P. Enflo: A counterexample to the approximation problem in Banach spaces. Acta Math. **130**, 309-317 (1973).

[48] P. Enflo: On the invariant subspace problem in Banach spaces. Sémin. d'anal. fonctionnelle (Maurey-Schwartz). Exp. no 14 et 15, p. 1-6 (1975-1976) On the invariant subspace problem for Banach spaces. Acta Math. **158**, 213-313 (1987).

[49] P. Enflo: On the invariant subspace problem in Hilbert spaces. Preprint 2023. `https://arxiv.org/abs/2305.15442`

[50] K.-J. Engel, R. Nagel: *One-parameter semigroups for linear evolution equations.* Springer, 2000.

[51] R. Engelking: *General topology.* Heldermann Verlag, 1989.

[52] L. C. Evans: *Partial differential equations.* 2nd. ed. American Mathematical Society, 2010.

[53] A. Fellhauer: On the relation of three theorems of analysis to the axiom of choice. J. Logic Analysis **9**, 1-23 (2017).

[54] M. Foreman, F. Wehrung: The Hahn-Banach theorem implies the existence of a non-Lebesgue measurable set. Fund. Math. **77**, 13-19 (1991).

[55] S. Friedberg, A. Insel, L. Spence: *Linear algebra.* 4th. ed. Pearson, 2014.

[56] T. W. Gamelin, R. E. Greene: *Introduction to topology.* 2nd. ed. Dover Publications, 1999.

[57] D. J. H. Garling: *A course in mathematical analysis.* Vol. 1 & 2. Cambridge University Press, 2013.

[58] B. W. Glickfeld: A metric characterization of $C(X)$ and its generalization to C^*-algebras. Illinois Journ. Math **10**, 547-556 (1966).

[59] I. C. Gohberg, M. G. Krein: *Introduction to the theory of linear nonselfadjoint operators in Hilbert Space.* American Mathematical Society Translations, 1969.

[60] F. Q. Gouvêa: *p-adic numbers. An introduction.* 3rd. ed. Springer, 2020.

[61] S. Grabiner: The Tietze extension theorem and the open mapping theorem. Amer. Math. Monthly **93**, 190-191 (1986).

[62] A. Grothendieck: La théorie de Fredholm. Bull. Soc. Math. France **84**, 319-384 (1956).

[63] S. Gudder: Inner product spaces. Amer. Math. Monthly **81**, 29-36 (1974), **82**, 251-252 (1975), **82**, 818 (1975).

[64] K. E. Gustafson, D. K. M. Rao: *Numerical range. The field of values of linear operators and matrices.* Springer, 1997.

[65] R. Haag: *Local quantum physics.* 2nd. ed. Springer, 1996.

[66] P. R. Halmos: *Introduction to Hilbert space and the theory of spectral multiplicity.* 2nd. ed. Chelsea, 1957.

[67] P. R. Halmos: What does the spectral theorem say? Amer. Math. Monthly **70**, 241-247 (1963).

[68] P. R. Halmos: *A Hilbert space problem book.* 2nd. ed. Springer, 1982.

[69] I. Halperin: Sums of a series, permitting rearrangements. C. R. Math. Rep. Acad. Sci. Can. **VIII**, 87-102 (1986).

[70] J. D. Halpern: The independence of the axiom of choice from the Boolean prime ideal theorem. Fundam. Math. **55**, 57-66 (1964).

[71] J. D. Halpern, A. Lévy: The Boolean prime ideal theorem does not imply the axiom of choice. *In:* D. Scott (ed.) *Axiomatic set theory.* Amer. Math. Soc., 1971.

[72] H. Hanche-Olsen, H. Holden: The Kolmogorov-Riesz compactness theorem. Expos. Math. **28**, 385-394 (2010).

[73] O. Hanner: On the uniform convexity of L^p and l^p. Ark. Mat. **3**, 239-244 (1955).

[74] C. Heil: *A basis theory primer.* Birkhäuser, 2011.

[75] N. Higson, J. Roe: *Analytic K-homology.* Oxford University Press, 2000.

[76] K. Hoffman: *Banach spaces of analytic functions.* Prentice-Hall, 1962.

[77] T. Hytönen, J. van Neerven, M. Veraar, L. Weis: *Analysis in Banach spaces.* Two Volumes. Springer, 2016, 2017.

[78] V. M. Kadets, M. I. Kadets: *Rearrangements of series in Banach spaces.* American Mathematical Society, 1991.

[79] R. V. Kadison, J. R. Ringrose: *Fundamentals of the theory of operator algebras.* Two volumes. Academic Press, 1983, 1986.

[80] E. Kaniuth: *A course in commutative Banach algebras.* Springer, 2009.

[81] S. Kaplan: *The bidual of* $C(X)$ *I.* North-Holland, 1985.

[82] A. A. Karatsuba, S. M. Voronin: *The Riemann zeta function.* Walter de Gruyter, 1992.

[83] Y. Katznelson: *An introduction to harmonic analysis.* 3rd. ed. Cambridge University Press, 2004.

[84] Y. & Y. R. Katznelson: *A (terse) introduction to linear algebra.* American Mathematical Society, 2008.

[85] I. Kleiner: *A history of abstract algebra.* Birkhäuser, 2007.

[86] V. Komornik: *Lectures on functional analysis and the Lebesgue integral.* Springer, 2016.

[87] E. Kreyszig: *Introductory functional analysis with applications.* Wiley, 1978.

[88] M. Krukowski: Natural proof of the characterization of relatively compact families in L^p-spaces on locally compact groups. Publ. Math. **67**, 687-713 (2023).

[89] N. H. Kuiper: The homotopy type of the unitary group of Hilbert space. Topology **3**, 19-30 (1965).

[90] H. E. Lacey, R. Whitley: Conditions under which all the bounded linear maps are compact. Math. Ann. **158**, 1-5 (1965).

[91] E. Landau: *Differential and integral calculus.* Chelsea, 1951.

[92] N. P. Landsman: *Foundations of quantum theory.* Springer, 2017. Freely available at `https://link.springer.com/book/10.1007%2F978-3-319-51777-3`

[93] S. Lang: *Undergraduate algebra.* 2nd. ed. Springer, 1990.

[94] P. D. Lax: *Functional analysis.* Wiley, 2002.

[95] P. D. Lax: *Linear algebra and its applications.* 2nd. ed. Wiley, 2007.

[96] S. R. Lay: *Convex sets and their applications.* Wiley, 1982.

[97] D. Li, H. Queffélec: *Introduction to Banach spaces: Analysis and probability.* Two Volumes. Cambridge University Press, 2018.

[98] J. Lindenstrauss, L. Tzafriri: *Classical Banach spaces. Vol. 1: Sequence spaces, Vol. 2: Function spaces.* Springer, 1977, 1979.

[99] J. Łoś, C. Ryll-Nardzewski: On the application of Tychonoff's theorem in mathematical proofs. Fund. Math. **88**, 233-237 (1951).

[100] W. A. J. Luxemburg: Two applications of the method of construction by ultrapowers to analysis. Bull. Amer. Math. Soc. **68**, 416-419 (1962).

[101] B. MacCluer: *Elementary functional analysis.* Springer, 2009.

[102] R. E. Megginson: *An introduction to Banach space theory.* Springer, 1998.

[103] R. Meise, D. Vogt: *Introduction to functional analysis.* Oxford University Press, 1997.

[104] A. J. Michaels: Hilden's simple proof of Lomonosov's invariant subspace theorem. Adv. Math. **25**, 56-58 (1977).

[105] D. F. Monna: *Functional analysis in historical perspective.* Oosthoek Publishing Company, 1973.

[106] G. H. Moore: The axiomatization of linear algebra: 1875-1940. Hist. Math. **22**, 262-303 (1995).

[107] M. H. Mortad: On the absolute value of the product and the sum of operators. Rend. Circ. Mat. Palermo, II. Ser **68**, 247-257 (2019).

[108] M. Müger: *Topology for the working mathematician.* (work in progress).
`http://www.math.ru.nl/~mueger/topology.pdf`

[109] M. Müger: Some examples of Fourier series.
`http://www.math.ru.nl/~mueger/PDF/Fourier.pdf`

[110] G. J. Murphy: *C^*-Algebras and operator theory.* Academic Press, 1990.

[111] G. Nagy: A functional analysis point of view on the Arzela-Ascoli theorem. Real. Anal. Exch. **32**, 583-586 (2006/7).
D. C. Ullrich: The Ascoli-Arzelà Theorem via Tychonoff's Theorem. Amer. Math. Monthly **110**, 939-940 (2003).
M. Wójtowicz: For eagles only: probably the most difficult proof of the Arzelà-Ascoli theorem – via the Stone-Čech compactification. Quaest. Math. **40**, 981-984 (2017).

[112] L. Narici, E. Beckenstein, G. Bachman: *Functional analysis and valuation theory.* Marcel Dekker, Inc. 1971.

[113] J. van Neerven: *Functional analysis.* Cambridge University Press, 2022.

[114] D. J. Newman: A simple proof of Wiener's $1/f$ theorem. Proc. Amer. Math. Soc. **48**, 264-265 (1975).

[115] L. Nirenberg: *Topics in nonlinear functional analysis.* American Mathematical Society, 1974.

[116] B. de Pagter, A. C. M. van Rooij: *An invitation to functional analysis.* Epsilon Uitgaven, 2013.

[117] J. Pawlikowski: The Hahn-Banach theorem implies the Banach-Tarski paradox. Fund. Math. **77**, 21-22 (1991).

[118] G. Pedersen: *Analysis now.* Springer, 1989.

[119] R. R. Phelps: *Lectures on Choquet's theorem.* 2nd. ed. Springer, 2001.

[120] J.-P. Pier: *Mathematical analysis during the 20th century.* Oxford University Press, 2001.

[121] A. Pietsch: *Eigenvalues and s-numbers.* Cambridge University Press, 1987.

[122] A. Pietsch: *History of Banach spaces and linear operators.* Birkhäuser, 2007.

[123] D. Pincus: The strength of the Hahn Banach theorem. *In:* A. Hurd, P. Loeb (eds.): *Victoria Symposium on Nonstandard Analysis.* LNM 369, Springer, 1974.

[124] D. Pincus: Adding dependent choice to the prime ideal theorem. *In:* R. O. Gandy, J. M. E. Hyland (eds.): *Logic Colloquium 76.* North-Holland, 1977.

[125] J. B. Prolla: *Topics in functional analysis over valued division rings.* North-Holland, 1982.

[126] D. Ramakrishnan, R. J. Valenza: *Fourier analysis on number fields.* Springer, 1999.

[127] C. J. Read: A solution to the Invariant Subspace Problem on the space ℓ_1. Bull. London Math. Soc. **17**, 305-317 (1985).

[128] M. Reed, B. Simon: *Methods of modern mathematical physics I: Functional analysis.* Academic Press, 1980.

[129] M. Reed, B. Simon: *Methods of modern mathematical physics IV: Analysis of operators.* Academic Press, 1980.

[130] C. E. Rickart: An elementary proof of a fundamental theorem in the theory of Banach algebras. Michigan Math. J. **5**, 75-78 (1958).

[131] C. E. Rickart: *General theory of Banach algebras.* Robert E. Krieger Publishing Co. Inc., 1960.

[132] F. Riesz, B. Sz.-Nagy: *Functional analysis.* Frederick Ungar Publ., 1955, Dover, 1990.

[133] J. R. Ringrose: A note on uniformly convex spaces. J. London Math. Soc. **34**, 92 (1959).

[134] J. R. Ringrose: Super-diagonal forms for compact linear operators. Proc. London Math. Soc. **12**, 367-384 (1962).

[135] J. R. Ringrose: *Compact non-self-adjoint operators.* Van Nostrand Reinhold, 1971.

[136] A. M. Robert: *A course in p-adic analysis.* Springer, 2000.

[137] J. W. Roberts: A compact convex set with no extreme points. Studia Mathematica **60**, 255-266 (1977).

[138] A. C. M. van Rooij: *Non-Archimedean functional analysis.* Marcel Dekker, Inc., 1978.

[139] W. Rudin: *Principles of mathematical analysis.* McGraw Hill, 1953, 1964, 1976.

[140] W. Rudin: *Real and complex analysis.* McGraw-Hill, 1966, 1974, 1986.

[141] W. Rudin: *Functional analysis.* 2nd. ed. McGraw-Hill, 1991.

[142] V. Runde: *A taste of topology.* Springer, 2005.

[143] V. Runde: A new and simple proof of Schauder's theorem.
`https://arxiv.org/abs/1010.1298`

[144] R. A. Ryan: *Introduction to tensor products of Banach spaces.* Springer, 2002.

[145] B. P. Rynne, M. A. Youngson: *Linear functional analysis.* 2nd. ed. Springer, 2008.

[146] D. A. Salamon: *Measure and integration.* European Mathematical Society, 2016.

[147] D. Sarason: The multiplication theorem for Fredholm operators. Amer. Math. Monthly **94**, 68-70 (1987).

[148] K. Saxe: *Beginning functional analysis*. Springer, 2002.

[149] E. Schechter: *Handbook of analysis and its foundations*. Academic Press, 1997.

[150] C. Schmoeger: Remarks on commuting exponentials in Banach algebras. Proc. Amer. Math. Soc. **127**, 1337-1338 (1999).

[151] B. Simon: *Trace ideals and their applications*. 2nd. ed. American Mathematical Society, 2005.

[152] B. Simon: *Operator theory*. American Mathematical Society, 2015.

[153] I. Singer: *Bases in Banach spaces I & II*. Springer, 1970, 1981.

[154] A. Sokal: A really simple elementary proof of the Uniform Boundedness Theorem. Amer. Math. Monthly **118**, 450-452 (2011).

[155] P. Soltan: *A primer on Hilbert space*. Springer, 2018.

[156] L. A. Steen: Highlights in the history of spectral theory. Amer. Math. Monthly **80**, 359-381 (1973).

[157] E. M. Stein, R. Shakarchi: *Fourier analysis*. Princeton University Press, 2005.

[158] J. Stillwell: *Reverse mathematics. Proofs from the inside out*. Princeton University Press, 2018.

[159] M. H. Stone: *Linear transformations in Hilbert space and their applications to analysis*. American Mathematical Society, 1932.

[160] V. S. Sunder: Fuglede's theorem. Indian J. Pure Appl. M. **46**, 415-417 (2015).

[161] C. Swartz: *Infinite matrices and the gliding hump*. World Scientific, 1996.

[162] A. Szankowski: $B(H)$ does not have the approximation property. Acta Math. **147**, 89-108 (1981).

[163] T. Tao: *Analysis I & II*. 3rd. ed. Springer, 2016.

[164] G. Teschl: *Topics in linear and nonlinear functional analysis*. American Mathematical Society, 2020.

[165] N. Tomczak-Jaegermann: *Banach-Mazur distances and finite-dimensional operator ideals*. Longman Scientific & Technical, 1989.

[166] A. M. Vershik: The life and fate of functional analysis in the twentieth century. *In:* A. A. Bolibruch, Yu. S. Osipov, Ya. G. Sinai (eds.): *Mathematical events of the twentieth century*. Springer, 2006.

[167] S. Warner: *Topological fields*. North-Holland, 1989.

[168] J. Weidmann: *Lineare Operatoren in Hilberträumen. 1: Grundlagen, 2: Anwendungen*. Teubner, 2000, 2003.

[169] A. Weil: *L'intégration dans les groupes topologiques et ses applications*. 2ème éd. Hermann, 1965.

[170] H. Weyl: Über beschränkte quadratische Formen, deren Differenz vollstetig ist. Rend. Circ. Mat. Palermo **27**, 373-392 (1909).

[171] H. Weyl: Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen. Math. Ann. **68**, 220-269 (1910).

[172] R. Whitley: Projecting m onto c_0. Amer. Math. Monthly **73**, 285-286 (1966).

[173] R. Whitley: An elementary proof of the Eberlein-Šmulyan theorem. Math. Ann. **172**, 116-118 (1967).

[174] R. Whitley: The spectral theorem for a normal operator. Amer. Math. Monthly **75**, 856-861 (1968).

[175] W. Więsław: On topological fields. Colloq. Math. **29**, 119-146 (1974).

[176] P. Wojtaszczyk: *Banach spaces for analysts*. Cambridge University Press, 1991.

[177] E. Zeidler: *Nonlinear functional analysis*. Volumes 1, 2A, 2B, 3, 4, 5. Springer, 1984-1990.

[178] R. J. Zimmer: *Essential results of functional analysis*. University of Chicago Press, 1990.

[179] A. Zsák: On the Solution of the scalar-plus-compact problem by Argyros and Haydon. EMS Newsletter, December 2018, pp. 8-15.

[180] Question "Why do we care about L^p spaces besides $p = 1, p = 2, p = \infty$?" on MathOverflow with answers. https://mathoverflow.net/questions/28147/why-do-we-care-about-lp-spaces-besides-p-1-p-2-and-p-infty